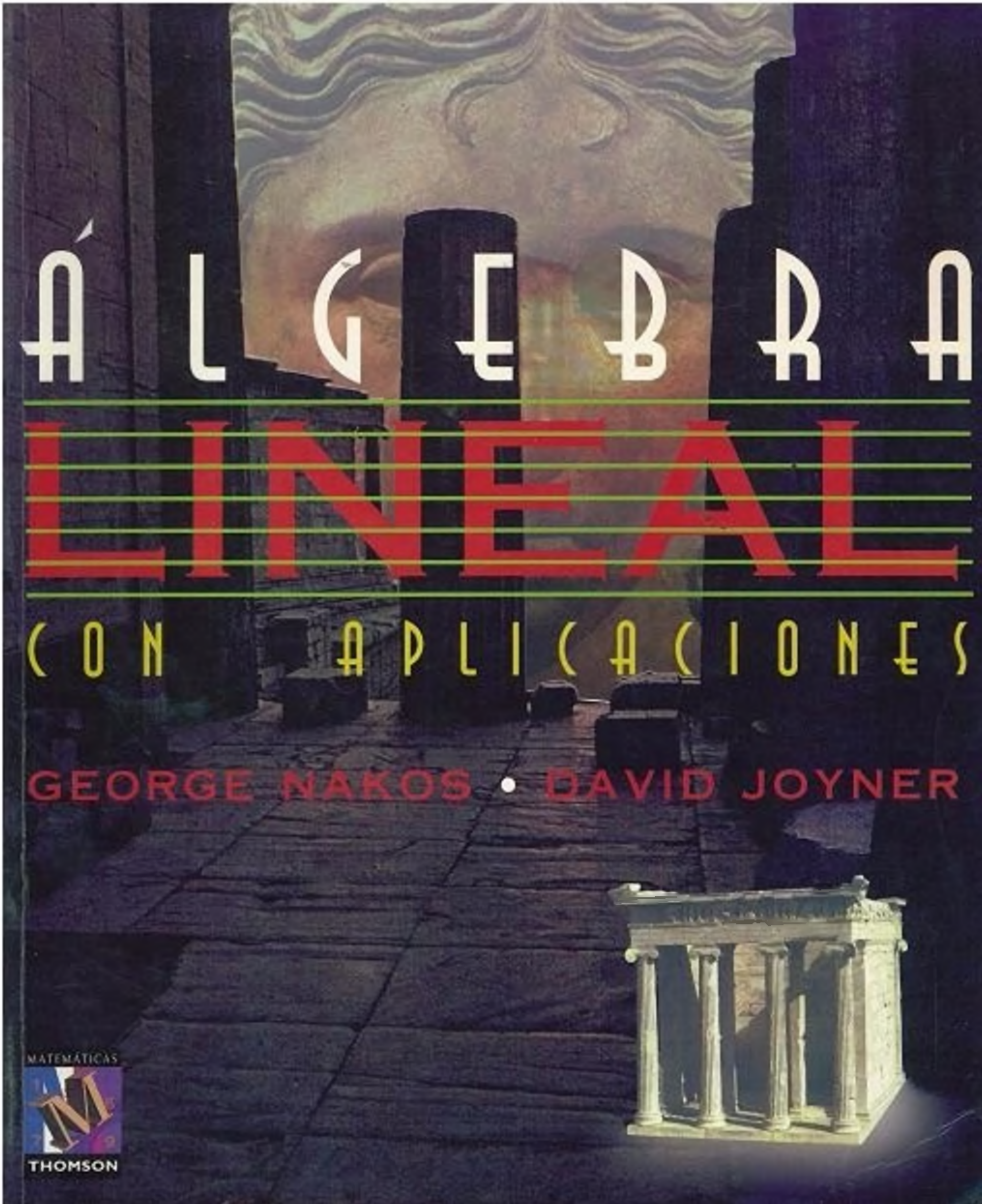


ÁLGEBRA LINEAL

CON APLICACIONES

GEORGE NAKOS • DAVID JOYNER

MATEMÁTICAS
THOMSON

Álgebra Lineal con Aplicaciones

George Nakos
U.S. Naval Academy

David Joyner
U.S. Naval Academy

ISBN 0-534-95526-6

Álgebra lineal con aplicaciones
ISBN 968-7529-86-5

ITP International Thomson Editores, S. A. de C. V. es una empresa de *International Thomson Publishing*. La marca registrada ITP se usa bajo licencia.

México y América Central
Séneca 53, Colonia Polanco
México, D. F. 11560
Tel. (525) 281-2906
Fax (525) 281-2656
clientes@mail.internet.com.mx
MÉXICO

El Caribe
Tel. (787) 758-7580
Fax (787) 758-7573
102154.1127@compuserve.com
Hato Rey, PUERTO RICO

América del Sur
Tel. (54-11)4325-2236
Fax (54-11)4328-1829
sdeluque@ba.net
Buenos Aires, ARGENTINA

España
Tel. (3491) 446-3350
Fax (3491) 445-6218
itesparaninfo.pedidos@mad.servicom.es
Madrid, ESPAÑA

Traducción
Ing. Virgilo González Pozo
Traductor profesional

Revisión técnica
Ana Elizabeth García Hernández
Escuela Superior de Ingeniería Química e Industrias Extractivas (ESIQIE)
Instituto Politécnico Nacional (IPN)

Director editorial y de producción: Miguel Ángel Toledo Castellanos
Editora de desarrollo: Leticia Medina Vigil
Editor de producción: René Garay Argueta
Corrector de estilo: Martha Alvarado
Tipografía: Inés Mendoza
Diseño de portada: Publicidad Williams-Lazarov
Lecturas: Martha Alvarado y Carlos Zúñiga

987654321 9VII9

Impreso en México
Printed in Mexico

1

Sistemas lineales

Tú has ordenado todo en medida, en número y en peso.
Sabiduría de Salomón, Capítulo 11, Versículo 20

Introducción

Muchas preguntas en ingeniería, física, matemáticas, economía y otras ciencias se reducen al problema de resolver un sistema lineal. El interés en la solución de esos sistemas es muy antiguo, como lo demuestra el *Problema del ganado* de Arquímedes (que se estudiará en la sección 1.5). Veamos un problema donde interviene un sistema lineal, del que se ocuparon los matemáticos de hace ochocientos años. Su solución se describe en la sección 1.4.

Fibonacci

Nuestra historia es acerca de Leonardo Pisano, matemático italiano (cerca 1175-1250 d.C.), mejor conocido como Fibonacci. Durante sus viajes, aprendió la "nueva aritmética" árabe, que después presentó al Occidente en su famoso libro *Liber abaci*. Dice la leyenda que el emperador Federico II de Sicilia invitó a Fibonacci y a otros sabios a participar en una especie de torneo de matemáticas, en el que se plantearon varios problemas. Uno de ellos era el siguiente:

> *Tres hombres poseen una sola pila de monedas, y sus partes son $\frac{1}{2}$, $\frac{1}{3}$ y $\frac{1}{6}$. Cada uno toma algo de dinero de la pila hasta que no queda nada. El primero regresa* $\frac{1}{2}$ de lo que tomó, el segundo $\frac{1}{3}$ y el tercero $\frac{1}{6}$. Cuando el total reintegrado se *divide por igual entre los tres, se descubre que cada uno posee lo que le corresponde. ¿Cuánto dinero había en la pila original, y cuánto tomó cada uno de esa pila?*

Fibonacci llegó a la solución: la cantidad total era 47, y las cantidades que tomaron los hombres de la pila fueron 33, 13 y 1. ¿Es correcto?

1.1 Introducción a los sistemas lineales

Objetivo del alumno para esta sección

Reconocer y emplear la eliminación para resolver un sistema lineal sencillo.

En esta sección presentamos las nociones de una ecuación lineal, y de un sistema de ecuaciones lineales. Describiremos cómo resolver sistemas "pequeños" por eliminación. En la práctica, los sistemas característicos se resuelven con computadora. Por lo general implican cientos o hasta miles de ecuaciones e incógnitas.

El conjunto de todos los números reales se representa por **R**. A menos que se indique otra cosa, un *escalar* es un número real. Por lo común la frase "x es un elemento de un conjunto A" se abrevia como sigue:

$$x \in A$$

Así, $x \in \mathbf{R}$ quiere decir que x es un escalar, o que x es un número real.

Ecuaciones lineales

DEFINICIÓN

(Ecuación lineal)

Una ecuación con n **variables** $x_1, \ldots, x_n$ es **lineal** si puede escribirse en la forma

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b \tag{1.1}$$

Las a_i son los **coeficientes**, y b es el **término constante** de la ecuación. Las variables también se llaman **incógnitas** o **indeterminadas**. Si $b = 0$, la ecuación se denomina **homogénea**. La que se obtiene a partir de (1.1), reemplazando b por 0, es la ecuación homogénea **asociada** con la ecuación (1.1). Si se ordenan las variables, la primera variable cuyo coeficiente es distinto de cero se llama **variable delantera**. Las demás son variables **libres**.

■ EJEMPLO 1 La ecuación

$$x_1 + x_2 + 4x_3 - 6x_4 - 1 = x_1 - x_2 + 2$$

es lineal, porque puede escribirse en la *forma canónica*, o normal (1.1):

$$0x_1 + 2x_2 + 4x_3 - 6x_4 = 3$$

Si las variables se ordenan de x_1 a x_4, x_2 es la variable delantera y x_1, x_3 y x_4 son libres. Los coeficientes son 0, 2, 4 y −6, y el término constante es 3. □

■ EJEMPLO 2 (de Celsius a Fahrenheit) La conversión normal de grados Celsius, C, a grados Fahrenheit, F, es una ecuación lineal en C y en F.

$$F = \frac{9}{5}C + 32 \tag{1.2}$$

□

■ EJEMPLO 3 Las ecuaciones lineales siguientes son homogéneas.

$$x_1 + 2x_2 - \sqrt{5}x_3 - x_4 = 0 \qquad x - y + z = (\operatorname{sen} 4)w$$

■ EJEMPLO 4 Las ecuaciones siguientes no son lineales, o **no lineales**:

$$xy - 3 = 2x \qquad x^2 - y = 1 \qquad \operatorname{sen} x + y = 0$$

Una **solución (particular)** de una ecuación lineal es una sucesión de números que, cuando se sustituyen en las variables, produce una ecuación que es una identidad.

Por ejemplo, $C = 5°$ y $F = 41°$ es una solución de (1.2), porque $\frac{9}{5} \cdot 5 + 32 = 41$. Por otro lado, $C = 5°$ y $F = 40°$ no es una solución, porque $\frac{9}{5} \cdot 5 + 32 \neq 40$.

El conjunto de todas las soluciones particulares se llama **conjunto solución**, el cual se obtiene despejando la variable delantera en función de las variables libres, y haciendo que cada variable libre tome cualquier valor escalar. Con esto se llega a un elemento genérico del conjunto solución, al cual se le llama **solución general**.

■ EJEMPLO 5 Determine la solución general de la ecuación

$$2x_1 + 0x_2 - 4x_3 = -2$$

SOLUCIÓN Despejaremos la variable delantera, x_1, para obtener $x_1 = 0x_2 + 2x_3 - 1$. Las variables libres x_2 y x_3 pueden tomar cualquier valor, por ejemplo $x_2 = s$ y $x_3 = r$. Por consiguiente, la solución general se expresa como

$$x_1 = 2r - 1, \qquad x_2 = s, \qquad x_3 = r \qquad \text{para toda } r, s \in \mathbf{R}$$

Las letras r y s empleadas para representar a las variables libres se llaman **parámetros**. El conjunto solución que acabamos de determinar es un *conjunto de dos parámetros*, o biparamétrico. Todas las soluciones particulares pueden encontrarse a partir de la solución general, asignando valores a los parámetros. Por ejemplo, $r = -1$ y $s = 2$ produce la solución particular $x_1 = -3$, $x_2 = 2$ y $x_3 = -1$.

Sistemas lineales

Un sistema lineal es un conjunto de ecuaciones lineales, como por ejemplo

$$\begin{aligned} 3x + 2y + z &= 39 \\ 2x + 3y + z &= 34 \\ x + 2y + 3z &= 26 \end{aligned} \tag{1.3}$$

Este sistema, con su solución, se encuentra en el libro chino de matemáticas *Nueve capítulos del arte matemático*[1] del siglo III a. C.

[1] Véase *A History of Mathematics*, de Carl Boyer, p. 219 (Nueva York: Wiley).

DEFINICIÓN

(Sistema lineal)

Un **sistema lineal** de m ecuaciones con n variables (o incógnitas) $x_1, \ldots, x_n$ es un conjunto de m ecuaciones lineales de la forma:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \tag{1.4}$$

Los números $a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn}$ son los **coeficientes** del sistema, y $b_1, b_2, \ldots, b_n$ son los **términos constantes.** Si todos los **términos constantes** son cero, el sistema se llama **homogéneo**. Cuando este último tiene los mismos coeficientes que el sistema (1.4) se dice que está **asociado** con (1.4).

Considérese el sistema

$$\begin{aligned} x_1 + 2x_2 \phantom{{}-2x_3} &= -3 \\ 2x_1 + 3x_2 - 2x_3 &= -10 \\ -x_1 \phantom{{}+3x_2} + 6x_3 &= 9 \end{aligned} \tag{1.5}$$

Sus coeficientes son, en orden, 1, 2, 0, 2, 3, −2, −1, 0, 6. Los términos constantes son −3, −10, 9. El sistema homogéneo asociado es

$$\begin{aligned} x_1 + 2x_2 \phantom{{}-2x_3} &= 0 \\ 2x_1 + 3x_2 - 2x_3 &= 0 \\ -x_1 \phantom{{}+3x_2} + 6x_3 &= 0 \end{aligned}$$

Se puede abreviar la escritura de un sistema lineal, anotando sólo sus coeficientes y términos constantes, siempre que estén especificados los nombres y el orden de las variables. El arreglo rectangular de los coeficientes y términos constantes de un sistema es su **matriz aumentada**. Por ejemplo, la matriz aumentada de las ecuaciones (1.5) es

$$\begin{bmatrix} 1 & 2 & 0 & -3 \\ 2 & 3 & -2 & -10 \\ -1 & 0 & 6 & 9 \end{bmatrix} \quad \text{o} \quad \left[\begin{array}{ccc:c} 1 & 2 & 0 & -3 \\ 2 & 3 & -2 & -10 \\ -1 & 0 & 6 & 9 \end{array}\right]$$

La segunda forma implica el uso de un separador para indicar dónde está la columna de los términos constantes. En general, una **matriz** es un arreglo rectangular de números. La **matriz de coeficientes** está formada por los coeficientes de un sistema. La matriz de una columna que muestra los términos constantes es el **vector de constantes**. La matriz de coeficientes y el vector de constantes del sistema (1.5) son, respectivamente:

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & -2 \\ -1 & 0 & 6 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -3 \\ -10 \\ 9 \end{bmatrix}$$

■ EJEMPLO 6 Escriba un sistema partiendo de una matriz aumentada.

$$\begin{bmatrix} 1 & 2 & 0 & -4 \\ 0 & 3 & -2 & -1 \end{bmatrix}$$

SOLUCIÓN Como la matriz aumentada tiene 4 columnas, el sistema tiene 3 variables. Si asignamos a las variables los nombres x_1, x_2 y x_3, entonces

$$\begin{aligned} x_1 + 2x_2 &= -4 \\ 3x_2 - 2x_3 &= -1 \end{aligned}$$

es un sistema que corresponde a la matriz aumentada dada. □

DEFINICIÓN

(Solución de un sistema lineal)

Una sucesión $r_1, r_2, \ldots, r_n$ de escalares es una **solución (particular)** del sistema (1.4) si todas las ecuaciones se satisfacen al sustituir $x_1 = r_1, \ldots, x_n = r_n$. El conjunto de todas las soluciones posibles es el **conjunto solución**. Cualquier elemento genérico del conjunto solución se llama **solución general**.

■ EJEMPLO 7 Demuestre que $x_1 = -15$, $x_2 = 6$ y $x_3 = -1$ es una solución particular del sistema (1.5).

SOLUCIÓN La sustitución $x_1 = -15$, $x_2 = 6$ y $x_3 = -1$ produce las siguientes afirmaciones correctas:

$$\begin{aligned} -15 + 2 \cdot 6 &= -3 \\ 2 \cdot (-15) + 3 \cdot 6 - 2 \cdot (-1) &= -10 \\ -(-15) + 6 \cdot (-1) &= 9 \end{aligned}$$ □

Si un sistema tiene soluciones se llama **consistente**; en cualquier otro caso se llama **inconsistente**. El sistema (1.5) es consistente. El sistema $x + y = 1, x + y = -1$, es inconsistente.

Un sistema lineal puede tener una, infinitas o ninguna solución. Esto se ilustra geométricamente para

$$\begin{array}{lll} y + x = 2 & y + x = 2 & y + x = 2 \\ y - x = 0 & 2y + 2x = 4 & y + x = 1 \end{array}$$

para el cual las líneas de ecuación se intersecan, coinciden o son paralelas (figura 1.1).

Gráficas de conjuntos solución

Sabemos que la gráfica de la ecuación $ax + by = c$ es una recta (excepto en los casos extremos $0x + 0y = 0$ y $0x + 0y = c \neq 0$). Entonces, en general, la gráfica del conjunto solución de un sistema de dos variables es la intersección de varias líneas rectas.

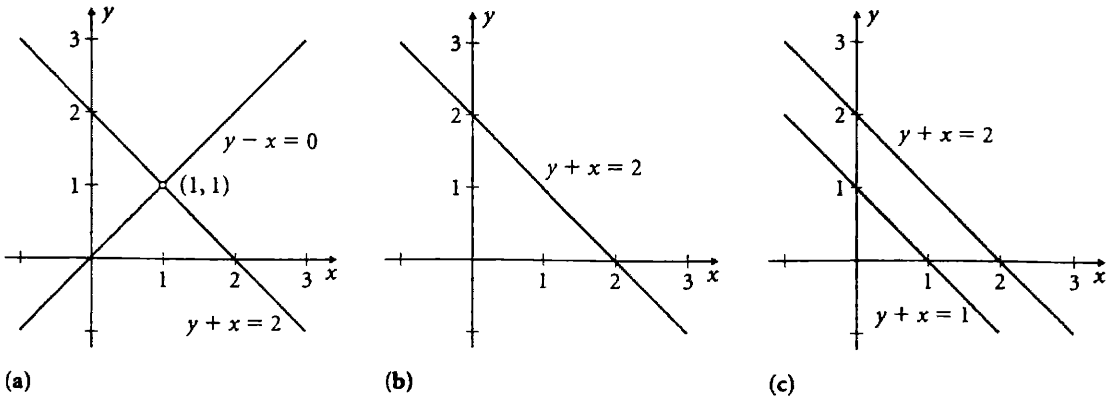


Figura 1.1 Soluciones: (a) solamente una, (b) un número infinito, (c) ninguna.

Aunque los planos se explicarán en el capítulo 2, debemos mencionar que la gráfica de la ecuación $ax + by + cz = d$ es un plano (excepto en los casos extremos $0x + 0y + 0z = 0$ y $0x + 0y + 0z = d \neq 0$). Por consiguiente, la gráfica del conjunto solución de un sistema de tres variables es, por lo común, la intersección de varios planos.

Observe que $x_1 = 0, \ldots, x_n = 0$ siempre es una solución de un sistema homogéneo. Se le llama **solución trivial**, o **solución cero**. A cualquier otra solución se le denomina **no trivial**. Por ejemplo, $x = 1$, $y = 1$ es una solución no trivial del sistema homogéneo

$$\begin{aligned} x - y &= 0 \\ -2x + 2y &= 0 \end{aligned}$$

Dos sistemas lineales que tienen los mismos conjuntos solución se llaman **equivalentes**. Los sistemas

$$\begin{aligned} -2x \phantom{{}-y} &= -2 \\ x - y &= 1 \end{aligned} \qquad \begin{aligned} 2x - 2y &= -2 \\ y &= 2 \end{aligned}$$

son equivalentes. Su solución común (y única) es $x = 1$, $y = 2$.

Solución de un sistema lineal

Los sistemas más fáciles de resolver tienen la forma triangular, o *de escalón*. En ellos, la variable delantera en cada ecuación se presenta a la derecha de la variable delantera de la ecuación escrita arriba. Esos sistemas se resuelven comenzando en la parte inferior y avanzando hacia arriba. Primero se soluciona la última ecuación, a continuación se sustituyen los valores en la próxima superior, con lo que también resolveremos esta ecuación. A este método se le llama **sustitución hacia atrás**.

■ EJEMPLO 8 Aplique la sustitución hacia atrás para resolver el sistema

$$\begin{aligned} x_1 - x_2 + x_3 - x_4 + 2x_5 - x_6 &= 1 \\ -2x_3 \phantom{{}-x_4} + 2x_5 \phantom{{}-x_6} &= 2 \\ -x_5 + x_6 &= 3 \end{aligned} \tag{1.6}$$

SOLUCIÓN La última ecuación implica que si x_6 es cualquier número, como por ejemplo r, entonces $x_5 = r - 3$. De acuerdo con la segunda ecuación, $x_3 = x_5 - 1 = r - 3 - 1 = r - 4$. Al despejar x_1, el resultado de la primera ecuación es $x_1 = 1 + x_2 - x_3 + x_4 - 2x_5 + x_6$. Pero x_2 y x_4 pueden tener cualquier valor, por ejemplo, $x_4 = s$, $x_2 = t$. Por consiguiente, $x_1 = 1 + t - (r - 4) + s - 2(r - 3) + r = -2r + s + t + 11$. Entonces, la solución general es

$$\begin{aligned} x_1 &= -2r + s + t + 11 \\ x_2 &= t \\ x_3 &= r - 4 \\ x_4 &= s \\ x_5 &= r - 3 \\ x_6 &= r \end{aligned} \qquad r, s, t \in \mathbf{R}$$

El conjunto solución es un conjunto infinito de *tres parámetros*. ☐

Por lo común, cuando se resuelve un sistema se eliminan las incógnitas para obtener un sistema equivalente en forma de escalón. A continuación se aplica la sustitución hacia atrás para determinar este último. Esta idea se refinará y describirá con más detalle en la sección 1.2.

Eliminaremos las incógnitas aplicando una secuencia de las siguientes operaciones en las ecuaciones, de tal manera que el sistema resultante quede en forma de escalón. *La aplicación de cualquiera de esas operaciones da como resultado sistemas equivalentes.*

DEFINICIÓN

(Operaciones elementales en ecuaciones)

Las **operaciones elementales en ecuaciones** de un sistema lineal consisten en lo siguiente.

(Eliminación) *Sume un múltiplo constante de una ecuación a otra.*

$\boxed{E_i + cE_j \to E_i}$

(Escalamiento) *Multiplique una ecuación por una constante distinta de cero.* $\boxed{cE_i \to E_i}$

(Intercambio) *Intercambie dos ecuaciones.* $\boxed{E_i \leftrightarrow E_j}$

Como la matriz aumentada de un sistema es una representación abreviada del mismo, podemos ahorrar tiempo y evitar faltas en la notación si trabajamos con la matriz aumentada. Las operaciones matriciales que corresponden a las operaciones elementales en ecuaciones se llaman operaciones elementales de renglón. En realidad, éstas pueden aplicarse a *cualquier* matriz.

DEFINICIÓN

(Operaciones elementales de renglón)

Las **operaciones elementales de renglón** de una matriz implican lo siguiente.

(Eliminación) *Sume un múltiplo constante de un renglón a otro renglón.* $\boxed{R_i + cR_j \to R_i}$

(Escalamiento) *Multiplique un renglón por una constante distinta de cero.* $\boxed{cR_i \to R_i}$

(Intercambio) *Intercambie dos renglones.* $\boxed{R_i \leftrightarrow R_j}$

■ EJEMPLO 9 Resuelva por eliminación el sistema

$$\begin{aligned} x_1 + 2x_2 \phantom{{}-2x_3} &= -3 \\ 2x_1 + 3x_2 - 2x_3 &= -10 \\ -x_1 \phantom{{}+3x_2} + 6x_3 &= 9 \end{aligned}$$

SOLUCIÓN Se tiene

$$\begin{aligned} x_1 + 2x_2 \phantom{{}-2x_3} &= -3 \\ 2x_1 + 3x_2 - 2x_3 &= -10 \\ -x_1 \phantom{{}+3x_2} + 6x_3 &= 9 \end{aligned} \quad \text{o} \quad \left[\begin{array}{rrrr} 1 & 2 & 0 & -3 \\ 2 & 3 & -2 & -10 \\ -1 & 0 & 6 & 9 \end{array}\right]$$

Al multiplicar la primera ecuación por -2 y al sumarla a la segunda, se eliminará x_1 de esta última. Cuando se suma la primera ecuación a la tercera también se elimina x_1 de la tercera. Todo esto puede abreviarse con $E_2 - 2E_1 \to E_2$, y $E_3 + E_1 \to E_3$, o también por $R_2 - 2R_1 \to R_2$, y con $R_3 + R_1 \to R_3$ en la matriz aumentada.

$$\begin{aligned} x_1 + 2x_2 \phantom{{}-2x_3} &= -3 \\ -x_2 - 2x_3 &= -4 \\ 2x_2 + 6x_3 &= 6 \end{aligned} \quad \text{o} \quad \left[\begin{array}{rrrcr} 1 & 2 & 0 & : & -3 \\ 0 & -1 & -2 & & -4 \\ 0 & 2 & 6 & & 6 \end{array}\right]$$

Para despejar x_2 de la tercera ecuación se lleva a cabo $E_3 + 2E_2 \to E_3$ (o $R_3 + 2R_2 \to R_3$ en la matriz aumentada).

$$\begin{aligned} x_1 + 2x_2 \phantom{{}-2x_3} &= -3 \\ -x_2 - 2x_3 &= -4 \\ 2x_3 &= -2 \end{aligned} \quad \text{o} \quad \left[\begin{array}{rrrcr} 1 & 2 & 0 & & -3 \\ 0 & -1 & -2 & : & -4 \\ 0 & 0 & 2 & : & -2 \end{array}\right]$$

El sistema tomó la forma de escalón. Empezamos por abajo y conforme ascendamos, se despejarán las incógnitas *arriba* de las variables delanteras de cada ecuación (sustitución hacia atrás). Para eliminar x_3 de la segunda ecuación, se suma $E_2 + E_3 \to E_2$ (o $R_2 + R_3 \to R_2$).

$$\begin{aligned} x_1 + 2x_2 \phantom{{}-2x_3} &= -3 \\ -x_2 \phantom{{}-2x_3} &= -6 \\ 2x_3 &= -2 \end{aligned} \quad \text{o} \quad \left[\begin{array}{rrrcr} 1 & 2 & 0 & : & -3 \\ 0 & -1 & 0 & : & -6 \\ 0 & 0 & 2 & : & -2 \end{array}\right]$$

Para eliminar x_2 de la primera ecuación se suma $E_1 + 2E_2 \to E_1$ (o $R_1 + 2R_2 \to R_1$).

$$\begin{aligned} x_1 \phantom{{}+2x_2-2x_3} &= -15 \\ -x_2 \phantom{{}-2x_3} &= -6 \\ 2x_3 &= -2 \end{aligned} \quad \text{o} \quad \left[\begin{array}{rrrcr} 1 & 0 & 0 & : & -15 \\ 0 & -1 & 0 & : & -6 \\ 0 & 0 & 2 & : & -2 \end{array}\right]$$

Por último, con $\frac{1}{2}E_3 \to E_3$ (o $\frac{1}{2}R_3 \to R_3$), y $(-1)E_2 \to E_2$ (o $(-1)R_2 \to R_2$) se obtiene

$$\begin{aligned} x_1 \phantom{{}+x_2+x_3} &= -15 \\ x_2 \phantom{{}+x_3} &= 6 \\ x_3 &= -1 \end{aligned} \quad \text{o} \quad \left[\begin{array}{rrrcr} 1 & 0 & 0 & : & -15 \\ 0 & 1 & 0 & & 6 \\ 0 & 0 & 1 & & -1 \end{array}\right]$$

Por consiguiente, $x_1 = -15$, $x_2 = 6$ y $x_3 = -1$ es la única solución del sistema.

■ EJEMPLO 10 Resuelva el sistema

$$\begin{aligned} x + 3y - z &= 4 \\ -2x + y + 3z &= 9 \\ 4x + 2y + z &= 11 \end{aligned}$$

SOLUCIÓN Trabajamos con la matriz aumentada para obtener

$$\begin{bmatrix} 1 & 3 & -1 & : & 4 \\ -2 & 1 & 3 & : & 9 \\ 4 & 2 & 1 & : & 11 \end{bmatrix} \quad \boxed{\begin{array}{l} R_2 + 2R_1 \to R_2 \\ R_3 - 4R_1 \to R_3 \end{array}} \quad \begin{bmatrix} 1 & 3 & -1 & : & 4 \\ 0 & 7 & 1 & : & 17 \\ 0 & -10 & 5 & : & -5 \end{bmatrix}$$

$$\boxed{R_3 + \tfrac{10}{7}R_2 \to R_3} \quad \begin{bmatrix} 1 & 3 & -1 & : & 4 \\ 0 & 7 & 1 & : & 17 \\ 0 & 0 & \frac{45}{7} & : & \frac{135}{7} \end{bmatrix} \quad \boxed{\tfrac{7}{45}R_3 \to R_3}$$

$$\begin{bmatrix} 1 & 3 & -1 & : & 4 \\ 0 & 7 & 1 & : & 17 \\ 0 & 0 & 1 & : & 3 \end{bmatrix} \quad \boxed{\begin{array}{l} R_2 - R_3 \to R_2 \\ R_1 + R_3 \to R_1 \end{array}} \quad \begin{bmatrix} 1 & 3 & 0 & : & 7 \\ 0 & 7 & 0 & : & 14 \\ 0 & 0 & 1 & : & 3 \end{bmatrix}$$

$$\boxed{\tfrac{1}{7}R_2 \to R_2} \quad \begin{bmatrix} 1 & 3 & 0 & : & 7 \\ 0 & 1 & 0 & : & 2 \\ 0 & 0 & 1 & : & 3 \end{bmatrix} \quad \boxed{R_1 - 3R_2 \to R_1} \quad \begin{bmatrix} 1 & 0 & 0 & : & 1 \\ 0 & 1 & 0 & : & 2 \\ 0 & 0 & 1 & : & 3 \end{bmatrix}$$

Por tanto, $x = 1$, $y = 2$ y $z = 3$. La solución *geométrica* es el punto $P(1, 2, 3)$, que es la figura de los tres planos definidos por las ecuaciones del sistema (figura 1.2). $\square$

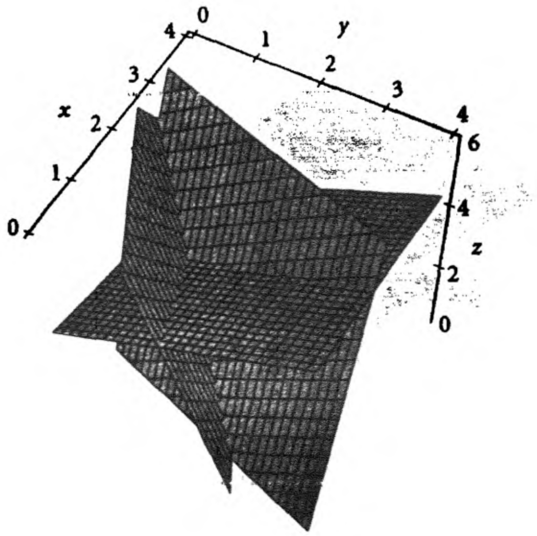


Figura 1.2 Solamente una solución: tres planos se intersecan en un punto.

■ EJEMPLO 11 (Soluciones infinitas) Resuelva el sistema

$$\begin{aligned} x + 2y - z &= 4 \\ 2x + 5y + 2z &= 9 \\ x + 4y + 7z &= 6 \end{aligned}$$

SOLUCIÓN Aplicando las operaciones elementales de renglón en la matriz aumentada del sistema se llega a $\begin{bmatrix} 1 & 0 & -9 & 2 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Ésta es la matriz aumentada del sistema

$$\begin{aligned} x_1 \qquad - 9x_3 &= 2 \\ x_2 + 4x_3 &= 1 \end{aligned}$$

Entonces, si $x_3 = r$ para cualquier escalar r, entonces $x_2 = -4r + 1$, y $x_1 = 9r + 2$, según las dos primeras ecuaciones. De esta manera, la solución general es

$$\begin{aligned} x_1 &= 9r + 2 \\ x_2 &= -4r + 1 \qquad r \in \mathbf{R} \\ x_3 &= r \end{aligned}$$

En este caso, la intersección de los planos definidos por las respectivas ecuaciones es una línea recta (figura 1.3). (Las líneas rectas se estudiarán en el capítulo 2.) $\square$

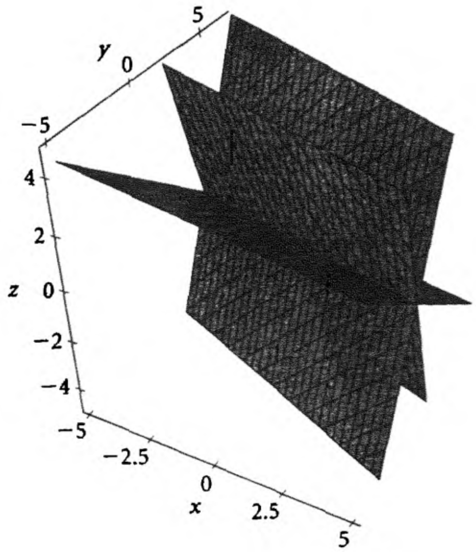


Figura 1.3 Los tres planos tienen una línea común.

■ EJEMPLO 12 (Sin soluciones) Resuelva el sistema

$$\begin{aligned} y - 2z &= -5 \\ 2x - y + z &= -2 \\ 4x - y \quad &= -4 \end{aligned}$$

SOLUCIÓN La matriz aumentada del sistema se reduce a

$$\begin{bmatrix} 2 & -1 & 1 & : & -2 \\ 0 & 1 & -2 & : & -5 \\ 0 & 0 & 0 & : & 5 \end{bmatrix}$$

El último renglón corresponde a la ecuación falsa $0x_3 = 5$. En consecuencia, el sistema es inconsistente. $\square$

Los tres casos gráficos de un sistema inconsistente de tres variables son (figura 1.4):

1. Tres planos paralelos entre sí
2. Dos planos paralelos, intersecados por el tercero
3. Tres planos que se intersecan, pero sin intersección común a los tres.

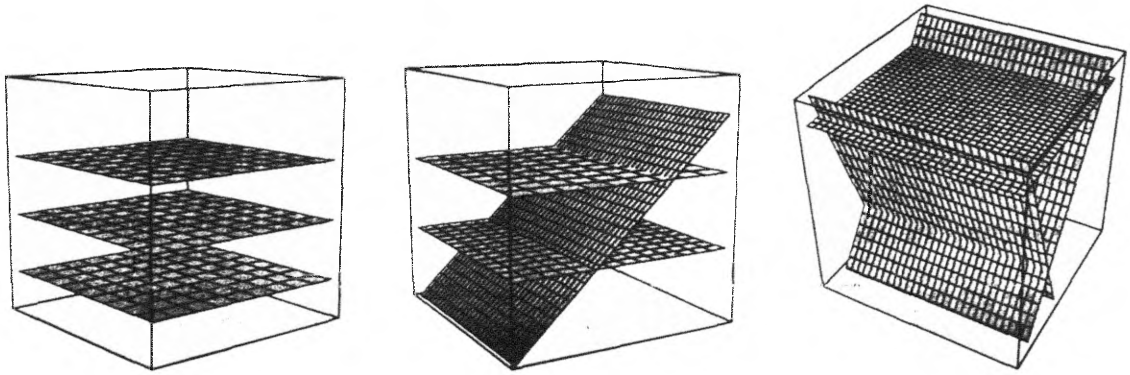


Figura 1.4 No hay soluciones.

■ EJEMPLO 13 (Empaquetamiento de libros) A una alumna se le asigna una nueva recámara. Al empacar sus libros, observa que si coloca 7 libros en cada caja, dejará uno fuera. Por otro lado, si pone 8 libros en cada caja, entonces la última caja sólo contiene un libro. ¿Cuántos libros y cuántas cajas hay?

SOLUCIÓN Sean x la cantidad de libros y y la cantidad de cajas. Entonces en el primer caso $7y = x - 1$, y en el segundo $8y = x + 7$. Al resolver este sistema de dos ecuaciones se obtiene $x = 57$ libros y $y = 8$ cajas. $\square$

Sistemas lineales con sistemas algebraicos computacionales

Maple

```
>solve({3*x+2*y=1,2*x-y=-3},{x,y});
```

$$\{x = \frac{-5}{7}, y = \frac{11}{y}\}$$

Mathematica

```
In[1]:=  Solve[{3x+2y==1,2x-y==-3},{x,y}]

Out[1]=            5        11
         {{x -> -(-), y -> --}}
                   7        7
```

MATLAB (Symbolic Toolbox)

```
>> solve('3*x+2*y=1','2*x-y=-3')
ans =
x = -5/7, y = 11/7
```

Ejercicios 1.1

Ecuaciones lineales

Para los ejercicios 1 a 4 considere las siguientes ecuaciones:

(a) $3x - 5 - x = 2x + 2y + 5$
(b) $2x + 3y - x = x + 3y - 1$
(c) $1 + x + y + z = 1$
(d) $x + y + z = 1 + y$
(e) $x + y + z = 1 + y - w + t$
(f) $xy + z = x - y$

1. Diga si cada ecuación es lineal o no lineal. Si es lineal, determine si es homogénea o no homogénea.

2. Para cada sistema lineal, escriba sus coeficientes, su término constante y su ecuación homogénea asociada.

3. Para cada ecuación lineal, ordene sus variables y diga cuál es la variable delantera y cuáles son las variables libres.

4. Para cada ecuación lineal determine, si es posible, la solución general y dos soluciones particulares.

5. ¿Cuáles de los puntos $P(2, -3, 0)$, $Q(2, -3, -1)$, $R(\frac{1}{2}, -\frac{1}{2}, -2)$ y $S(\frac{1}{2}, \frac{1}{2}, -2)$ están en el plano definido por la ecuación $x - y + z = -2$?

6. ¿Cuáles de los puntos $P(1, -1, 1)$, $Q(-2, 5, 3)$, $R(\frac{1}{2}, \frac{1}{7}, 10)$ y $S(0, 0, 0)$ están en el plano definido por la ecuación $2x - 7y + z = 10$

7. Encuentre todos los valores de a, tales que cada una de las siguientes ecuaciones tenga

(i) Exactamente una solución.
(ii) Soluciones infinitas.
(iii) Ninguna solución.

a. $a^2x - 2 = 4x + a$

b. $(a^2 - 4)x = 3$

c. $(a^2 - 4)x = 0$

d. $ax - a^2y = 3a$

8. Calcule todos los valores *reales* de a, tales que cada una de las ecuaciones de abajo tenga

(i) Infinitas soluciones.
(ii) Ninguna solución.

a. $a^2(x+y)-x-y-a+1=0$

b. $a(x+y)-x-y-a+1=0$

Sistemas lineales

9. Replantee el sistema lineal en la forma canónica.

$$\begin{aligned} 2x+4z+1&=0\\ 2z+2w-2&=x\\ -2x-z+3w&=-3\\ y+z+t&=w+4 \end{aligned}$$

Determine:

a. La matriz de coeficientes.
b. El vector de constantes.
c. La matriz aumentada.
d. El sistema homogéneo asociado.

10. Aplique la sustitución hacia atrás para resolver el sistema

$$\begin{aligned} x_1+2x_2+x_3\quad +x_5\quad &=-1\\ -2x_3\quad +4x_6&=\ \ 2\\ 4x_4-2x_5\quad &=\ \ 0 \end{aligned}$$

11. Emplee la sustitución hacia atrás para solucionar el sistema homogéneo asociado con el sistema del ejercicio 10.

12. Sea

$$M=\begin{bmatrix} 1 & -1 & 1 & -5 & 6 & -1 & 1\\ 0 & 0 & 0 & 0 & -1 & 1 & 0\\ 0 & 0 & -2 & 0 & 2 & 0 & 0 \end{bmatrix}$$

a. Escriba un sistema cuya matriz aumentada sea M.
b. Encuentre el sistema homogéneo asociado con el sistema lineal de la parte (a).
c. Aplique a M una operación elemental de renglón para que la matriz resultante corresponda a un sistema en forma de escalón.

13. Determine la solución general del sistema cuya matriz aumentada sea la matriz M, definida en el ejercicio 12.

14. Calcule la solución general del sistema homogéneo asociado al del ejercicio 13.

15. Encuentre la intersección de las rectas $y+x=1$ y $y-2x=\frac{1}{2}$.

16. Determine la intersección de las rectas $2y-3x=5$ y $y+2x=20$.

17. Sin resolver realmente los sistemas, demuestre que son equivalentes.

$$\begin{aligned} x-y+z&=1\\ 2x+2y-3z&=-2\\ -3x+4y+4z&=-1 \end{aligned}\qquad \begin{aligned} 4x-4y+4z&=4\\ 2x+2y-3z&=-2\\ 5y+2z&=-2 \end{aligned}$$

En los ejercicios 18 a 28 determine los sistemas consistentes y calcule sus soluciones generales.

18. $$\begin{aligned} -x+y-z&=1\\ -2x+y+3z&=10\\ 3x+y+2z&=3 \end{aligned}$$

19. $$\begin{aligned} y+2z&=6\\ 3x-3y-3z&=-15\\ x+3y+3z&=11 \end{aligned}$$

20. $$\begin{aligned} 3y+z&=-9\\ 3x+y&=-8\\ 3x+7y+2z&=-26 \end{aligned}$$

21. $$\begin{aligned} 3x+y+3z&=15\\ -x+3y-z&=-5\\ 2x+4y+2z&=9 \end{aligned}$$

22. $$\begin{aligned} -x+3y-2z&=-17\\ -2x-3y&=14\\ -3x-y-2z&=1 \end{aligned}$$

23. $$\begin{aligned} x+y-3z&=2\\ -3x+y+z&=6 \end{aligned}$$

24. $$\begin{aligned} \tfrac{1}{3}x+\tfrac{1}{3}y+z&=5\\ -\tfrac{1}{3}x-\tfrac{1}{2}y-\tfrac{1}{3}z&=-\tfrac{2}{3} \end{aligned}$$

25. $$\begin{aligned} 3y+z-w&=3\\ x+y-2z&=6\\ -2x+y+2z-w&=9 \end{aligned}$$

26. $$\begin{aligned} x+3y+z-w&=0\\ 3x+y+3z&=-2\\ 2x+6y+2z-2w&=2 \end{aligned}$$

27. $$\begin{aligned} x+y&=1\\ y+z&=1\\ z+w&=1\\ x+w&=1 \end{aligned}$$

28. $$\begin{aligned} x+y&=1\\ y+z&=1\\ z+w&=1\\ y+w&=1 \end{aligned}$$

29. Resuelva el viejo problema chino descrito en la sección 1.1.

En los ejercicios 30 a 34, utilice las matrices aumentadas que aparecen.

30. $\begin{bmatrix} -1 & 2 & 0 & -6 \\ 3 & -2 & -1 & 10 \\ 3 & 2 & 2 & -14 \end{bmatrix}$

31. $\begin{bmatrix} -1 & -1 & 0 & -1 \\ 0 & 2 & 1 & 3 \\ -2 & 1 & 3 & 7 \end{bmatrix}$

32. $\begin{bmatrix} -1 & -2 & -1 & 1 \\ -1 & -1 & 1 & 4 \\ 1 & 1 & -1 & 4 \end{bmatrix}$

33. $\begin{bmatrix} 2 & -1 & 3 & -4 \\ -1 & -3 & -1 & 7 \\ 2 & -1 & 0 & 8 \end{bmatrix}$

34. $\begin{bmatrix} -3 & -1 & 1 & 2 \\ 6 & 2 & -2 & -4 \\ 3 & 1 & -1 & -2 \end{bmatrix}$

35. Despeje θ del sistema:

$$\begin{aligned} \operatorname{sen}\theta - 4\cos\theta &= 4 \\ 4\operatorname{sen}\theta - 4\cos\theta &= 4 \end{aligned}$$

36. Despeje θ del sistema:

$$\begin{aligned} 2\operatorname{sen}\theta + \sqrt{2}\tan\theta &= 2\sqrt{2} \\ 4\operatorname{sen}\theta - 3\sqrt{2}\tan\theta &= -\sqrt{2} \end{aligned}$$

37. Considere al sistema homogéneo

$$\begin{aligned} a_1x + b_2y &= 0 \\ a_2x + b_2y &= 0 \end{aligned}$$

a. Demuestre que si $x = x_0$ y $y = y_0$ es una solución para ese sistema, también lo es $x = kx_0$, $y = ky_0$ para cualquier k constante.
b. Compruebe que si $x = x_1$, $y = y_1$ y $x = x_2$, $y = y_2$ son dos soluciones, entonces también lo es $x = x_1 + x_2$, $y = y_1 + y_2$.

Para los ejercicios 38 y 39 considere el sistema

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

38. Sea $a_1b_2 - b_1a_2 \neq 0$. Demuestre que
a. El sistema tiene exactamente una solución. Calcule esta solución.
b. El sistema homogéneo asociado tiene sólo la solución trivial.

39. Sea $a_1b_2 - b_1a_2 = 0$. Demuestre que
a. El sistema tiene un numérico infinito de soluciones o no las tiene.
b. El sistema homogéneo asociado tiene soluciones no triviales.

Aplicaciones

40. Pedro y Paula comienzan un juego de béisbol con baraja, y cada uno tiene la misma cantidad de cartas. Durante la primera ronda Pedro ganó 20 cartas, pero durante la segunda perdió $\frac{2}{3}$ de las que tenía al comenzar. Al final de la segunda ronda, Paula tiene tres veces las cartas que tiene Pedro. ¿Cuál fue la cantidad total de cartas?

41. Una profesora de matemáticas asigna 3 ejercicios, y pide a $\frac{1}{4}$ de sus alumnos que resuelvan el primero, a $\frac{3}{8}$ que resuelvan el segundo y a $\frac{5}{16}$ el tercero. Como 2 alumnos están ausentes, ¿cuál es la cantidad total de alumnos en el grupo?

42. Un estante contiene $\frac{3}{5}$ de la cantidad total de libros que están en el estante vecino. Si pasamos 10 libros del primero al segundo estante, éste tendrá el doble de libros que el primero. ¿Cuántos libros había en cada librero?

43. 45 marineros de la Academia Naval a bordo de un bote con 15 miembros de la tripulación tienen alimentos para 30 días. 12 días después, rescatan a 10 personas en una lancha inflable, con alimentos para 4 días. ¿Cuánto tiempo durará el suministro total de alimentos, si al personal en el bote se le da la misma ración diaria?

44. La tercera parte de un texto de matemáticas tiene un total de 90 secciones de ejercicios. Como en cada sección hay 25 ejercicios en promedio, y 15 ejercicios caben en una página, ¿cuántas páginas tiene el libro?

45. Determine los ángulos de un paralelogramo, que tienen la propiedad de que dos ángulos consecutivos difieren 20°.

46. Calcule la longitud de cada lado de un paralelogramo con 10 pies de perímetro y en el que dos lados consecutivos difieren 1 pie en su longitud.

47. Calcule la longitud de cada lado de un triángulo isósceles cuyo perímetro es 16 pies, y la diferencia de longitud de dos de sus lados desiguales es 2 pies.

48. Cuando se agrega un disco duro a una computadora personal, el sistema nuevo cuesta $1 400. Se sabe que $\frac{1}{3}$ del va-

lor de la computadora más $\frac{1}{5}$ del valor del disco duro dan un total de $400. ¿Cuál es el costo del disco duro?

49. En un crucigrama de 20 columnas, $\frac{1}{8}$ de la cantidad total es de cuadros negros, y son 5 menos que $\frac{1}{7}$ de ellos. ¿Cuál es la cantidad de renglones del crucigrama?

En los ejercicios 50 y 51 los sistemas resultantes son no lineales, pero ¿se podrían reducir a sistemas lineales?[2]

50. Media hora después de que uno de los dos autores salió de su casa conduciendo su coche al trabajo, encontró que un carril de la autopista había sido cerrado por un accidente. Esto ocasionó que el autor manejara, en promedio, a $\frac{2}{5}$ de su velocidad anterior durante el resto del camino, y por eso llegó 1 hora y 3 minutos tarde al trabajo. Si el accidente hubiera ocurrido 15 millas más adelante, el autor hubiera llegado 27 minutos más temprano. ¿A qué distancia de su trabajo vive el autor? ¿Cuál es su velocidad normal?

51. Dos botes de vela salen de Annapolis en horas distintas, lo que les da una separación de 4 millas, y tienen el mismo destino y la misma velocidad. Después de media hora de haber salido el segundo, el viento se calma bastante. Esto desacelera a los botes, que ahora viajan a $\frac{2}{3}$ de su velocidad anterior. El segundo bote llega a su destino 45 minutos después de la hora predicha. El primer bote sólo llega 35 minutos tarde. ¿Cuál es la distancia recorrida por los botes?

1.2 Eliminación de Gauss

Objetivos del alumno para esta sección

1. Reconocer una matriz en la forma renglón escalón y renglón escalón reducida
2. Aprender el proceso de eliminación de Gauss.

Ahora examinaremos con más detalle el método que usamos para resolver los sistemas lineales. Se llama *eliminación de Gauss*,[3] aunque data de una época anterior a Gauss. De hecho, se usó para resolver el sistema en el viejo libro chino citado en la sección 1.1. También describiremos una variante, la *eliminación de Gauss-Jordan*.

Forma de renglón escalón; eliminación de Gauss

Las matrices se describen con detalle en el capítulo 3. En este párrafo sólo presentaremos la notación necesaria para resolver con eficiencia los sistemas lineales.

Una matriz es un arreglo u ordenación rectangular de números, llamados **elementos**.[4] Los renglones de una matriz se numeran de arriba abajo, y las columnas de izquierda a derecha. Así, cuando decimos última columna nos referimos a la columna de la derecha. Los elementos se numeran de acuerdo con su posición (renglón, columna). Si una matriz tiene m renglones y n columnas, sé llama **tamaño** de $m \times n$.

[2] Ambos ejercicios están inspirados en un problema planteado por Sam Loyd, uno de los mejores redactores de rompecabezas y problemas de ajedrez de todos los tiempos.

[3] **Karl Friedrich Gauss** (1777–1855) se considera como uno de los más grandes matemáticos de todos los tiempos. Nació en Alemania y fue niño prodigio. A los 18 años formó un polígono regular de 17 lados con regla y compás, resolviendo un problema de 2 000 años de antigüedad. Escribió *Disquisitiones Arithmeticae*, otra maestra en la teoría de los números, y demostró el *teorema fundamental del álgebra*. En su época, se le llamaba príncipe de los matemáticos.

[4] N. del T.: Con frecuencia se les llama también "entradas".

A continuación se presentan M, N y P, que son matrices; M y N son de tamaño 3×4, pero el tamaño de P es de 2×3. El elemento $(2, 3)$ de P es 1. El elemento $(1, 2)$ de N es -6.

$$M = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & 0 & -3 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad N = \begin{bmatrix} 1 & -6 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad P = \begin{bmatrix} 0 & 2 & -6 \\ 0 & -1 & 1 \end{bmatrix}$$

El **renglón cero** de una matriz sólo incluye ceros. Un **renglón no cero** tiene cuando menos un elemento distinto de cero. De la misma forma se puede hablar de *columnas* cero y no cero. El primer elemento no cero de un renglón no cero se llama **elemento delantero**, o **elemento capital**. Si sucede que el elemento delantero es 1, se le llama **1 delantero**.

Para la matriz M es válido lo siguiente: sus primeros dos renglones son no cero. El tercero es cero. Sus elementos delanteros son -1 y -3. Su segunda columna es cero. Las demás son no cero. Observe que ninguno de los elementos delanteros de M son 1. En contraste, todos los elementos delanteros de N son 1.

DEFINICIÓN

(Forma de renglón escalón)

Una matriz puede tener las siguientes propiedades:

1. *Todos los renglones cero están en la parte inferior de la matriz.*
2. *El elemento delantero de cada renglón no cero después del primero se presenta a la derecha del elemento delantero del renglón anterior.*
3. *El elemento delantero de cualquier renglón no cero es 1.*
4. *Todos los elementos en la columna arriba y debajo de un 1 delantero son cero.*

Si una matriz satisface las dos primeras condiciones se dice que está en **forma de renglón escalón,** o simplemente en **forma de escalón**. Si una matriz satisface las cuatro condiciones se dice que está en **forma reducida de renglón escalón** (o tan sólo que está en **forma de escalón reducida**). Una matriz en forma de escalón reducida siempre está en forma de escalón.

Las matrices M y N están en forma de escalón. N además está en forma de escalón reducida. M no lo está, porque la condición 3 no se satisface. La matriz P tampoco lo está, porque no tiene la propiedad 2. Para estudiar otros ejemplos, sean

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix} \qquad E = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad F = \begin{bmatrix} 1 & 7 & 0 & 9 & 0 \\ 0 & 0 & 1 & -8 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$G = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

A, B, D, F, G y H están en forma de escalón, porque se satisfacen las dos primeras condiciones. De ellas, A, B, D y F están en forma de escalón reducida por tener las cuatro propiedades. Pero G y H no lo están. Para G, la condición 4 no se cumple. H no tiene la propiedad 3. C y E no están en la forma de escalón. Para C falla la condición 2. Para E falla la condición 1.

DEFINICIÓN

(Matrices equivalentes)

Dos matrices son **equivalentes (de renglón)** si una puede obtenerse de la otra mediante una sucesión finita de operaciones elementales de renglón. A veces se usa la notación

$$A \sim B$$

para indicar "la matriz A es equivalente a B".

■ EJEMPLO 14

$$A = \begin{bmatrix} 0 & 3 \\ 1 & 2 \\ -1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$$

son equivalentes, porque

$$A \boxed{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 \\ 0 & 3 \\ -1 & 1 \end{bmatrix} \boxed{R_3 + R_1 \to R_3} \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 0 & 3 \end{bmatrix} \boxed{R_3 - R_2 \to R_3} B$$

Se dice que una matriz *se convierte (o reduce) a la forma escalón (reducida)* si es equivalente a una matriz en la forma de escalón (reducida).

■ EJEMPLO 15 Demuestre que la matriz siguiente se convierte a la forma de escalón reducida.

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 3 & 3 & 1 & 0 & -1 \\ -2 & -2 & 1 & 1 & 0 \end{bmatrix}$$

SOLUCIÓN Esto se comprueba usando la sucesión de operaciones elementales de renglón siguientes.

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 3 & 3 & 1 & 0 & -1 \\ -2 & -2 & 1 & 1 & 0 \end{bmatrix} \boxed{\begin{array}{c} R_2 - 3R_1 \to R_2 \\ R_3 + 2R_1 \to R_3 \end{array}} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\boxed{R_3 - R_2 \to R_3} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

La última matriz está en forma de escalón reducida.

El método de solución en el ejemplo 15 es un caso especial del siguiente e importante algoritmo (proceso) que nos permite reducir cualquier matriz a la forma de escalón, o a la forma de escalón reducida.

Algoritmo 1

(Eliminación de Gauss)

Para convertir cualquier matriz a la forma de escalón reducida, proceda con los pasos siguientes:

Paso 1. *Vaya a la columna no cero extrema izquierda.*
Paso 2. *Si el primer renglón tiene un cero en la columna del paso 1, intercámbielo con uno que tenga un elemento no cero en la misma columna.*
Paso 3. *Obtenga ceros abajo del elemento delantero, sumando múltiplos adecuados del renglón superior a los renglones debajo de él.*
Paso 4. *Cubra el renglón superior y repita el mismo proceso comenzando con el paso 1, aplicado a la submatriz restante. Repita este proceso con el resto de los renglones. (En este punto la matriz ya está en forma de escalón.)*
Paso 5. *Comenzando con el último renglón no cero, avance hacia arriba: para cada renglón obtenga un 1 delantero e introduzca ceros arriba de él, sumando múltiplos adecuados a los renglones correspondientes.*

El ejemplo que sigue ilustra este proceso.

■ EJEMPLO 16 Emplee la eliminación de Gauss para determinar una forma de escalón reducida de la matriz

$$\begin{bmatrix} 0 & 3 & -6 & -4 & -3 & -5 \\ -1 & 3 & -10 & -4 & -4 & -2 \\ 4 & -9 & 34 & 0 & 1 & -21 \\ 2 & -6 & 20 & 2 & 8 & -8 \end{bmatrix}$$

SOLUCIÓN

Paso 1. *Vaya a la columna extrema izquierda no cero*: en este caso es la primera.
Paso 2. *Si el primer renglón tiene un cero en la columna del paso 1, intercámbielo con uno que tenga un elemento no cero en la misma columna.*

$$\boxed{R_1 \leftrightarrow R_2} \quad \begin{bmatrix} -1 & 3 & -10 & -4 & -4 & -2 \\ 0 & 3 & -6 & -4 & -3 & -5 \\ 4 & -9 & 34 & 0 & 1 & -21 \\ 2 & -6 & 20 & 2 & 8 & -8 \end{bmatrix}$$

Paso 3. *Obtenga ceros abajo del elemento delantero, sumando múltiplos adecuados del primer renglón a los que están por debajo de éste:*

$$\boxed{\begin{array}{l} R_3 + 4R_1 \to R_3 \\ R_4 + 2R_1 \to R_4 \end{array}} \quad \begin{bmatrix} -1 & 3 & -10 & -4 & -4 & -2 \\ 0 & 3 & -6 & -4 & -3 & -5 \\ 0 & 3 & -6 & -16 & -15 & -29 \\ 0 & 0 & 0 & -6 & 0 & -12 \end{bmatrix}$$

Paso 4. *Cubra el renglón superior y repita el mismo proceso, comenzando con el paso 1 aplicado a la submatriz restante. Haga lo mismo con el resto de los renglones.*

$$\begin{bmatrix} -1 & 3 & -10 & -4 & -4 & -2 \\ 0 & 3 & -6 & -4 & -3 & -5 \\ 0 & 3 & -6 & -16 & -15 & -29 \\ 0 & 0 & 0 & -6 & 0 & -12 \end{bmatrix}$$

$$\boxed{R_3 - R_2 \to R_3} \begin{bmatrix} -1 & 3 & -10 & -4 & -4 & -2 \\ 0 & 3 & -6 & -4 & -3 & -5 \\ 0 & 0 & 0 & -12 & -12 & -24 \\ 0 & 0 & 0 & -6 & 0 & -12 \end{bmatrix}$$

$$\boxed{R_4 - \tfrac{1}{2}R_3 \to R_4} \begin{bmatrix} -1 & 3 & -10 & -4 & -4 & -2 \\ 0 & 3 & -6 & -4 & -3 & -5 \\ 0 & 0 & 0 & -12 & -12 & -24 \\ 0 & 0 & 0 & 0 & 6 & 0 \end{bmatrix}$$

Paso 5. *Comenzando con el último renglón distinto de cero, avance hacia arriba: para cada renglón, obtenga un 1 delantero e introduzca ceros arriba de él, sumando los múltiplos adecuados a los renglones correspondientes.*

$$\boxed{\tfrac{1}{6}R_4 \to R_4} \begin{bmatrix} -1 & 3 & -10 & -4 & -4 & -2 \\ 0 & 3 & -6 & -4 & -3 & -5 \\ 0 & 0 & 0 & -12 & -12 & -24 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \boxed{\begin{matrix} R_3 + 12R_4 \to R_3 \\ R_2 + 3R_4 \to R_2 \\ R_1 + 4R_4 \to R_1 \end{matrix}}$$

$$\begin{bmatrix} -1 & 3 & -10 & -4 & 0 & -2 \\ 0 & 3 & -6 & -4 & 0 & -5 \\ 0 & 0 & 0 & -12 & 0 & -24 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \boxed{-\tfrac{1}{12}R_3 \to R_3} \begin{bmatrix} -1 & 3 & -10 & -4 & 0 & -2 \\ 0 & 3 & -6 & -4 & 0 & -5 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\boxed{\begin{matrix} R_2 + 4R_3 \to R_2 \\ R_1 + 4R_3 \to R_1 \end{matrix}} \begin{bmatrix} -1 & 3 & -10 & 0 & 0 & 6 \\ 0 & 3 & -6 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \boxed{\tfrac{1}{3}R_2 \to R_2}$$

$$\begin{bmatrix} -1 & 3 & -10 & 0 & 0 & 6 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \boxed{R_1 - 3R_2 \to R_1} \begin{bmatrix} -1 & 0 & -4 & 0 & 0 & 3 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\boxed{(-1)R_1 \to R_1} \begin{bmatrix} 1 & 0 & 4 & 0 & 0 & -3 \\ 0 & 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Observaciones

1. Los cuatro primeros pasos del algoritmo se llaman **paso directo** de la eliminación de Gauss. Convierten la matriz en forma de escalón. El paso 5 es el **paso inverso** (o sustitución hacia atrás), y es donde la matriz se lleva a su forma de escalón reducida.
2. Con frecuencia, el algoritmo se describe introduciendo 1 delanteros en el paso 3, con lo que hay una tendencia a introducir fracciones en una etapa temprana del cálculo, si los elementos son enteros. Esta variante es la preferida por los programadores en aritmética de punto flotante. Además, en el paso 2 el renglón con el elemento no cero de mayor valor

absoluto se cambia hasta arriba. A esto se le llama *pivoteo parcial*, y se describe en la sección 1.3. El pivoteo parcial ayuda a controlar los errores de redondeo.[5]

3. A veces conviene obtener 1 delanteros al final de todo el proceso.

Dos cosas que deben evitarse

1. Evitar el cambio del orden de los pasos del algoritmo (porque se puede terminar con la matriz inicial). Observe que el siguiente paso en la reducción de $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$ es $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$ y *no es* $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$. En otras palabras, primero es preciso reducir a la forma de escalón y después aplicar el paso 5. La única libertad permitida es al obtener los 1 delanteros.
2. Combinar varias operaciones elementales en una es una mala práctica, y puede originar errores. Por ejemplo, operaciones de la forma $cR_i + dR_j \to R_i$ y $cR_i + R_j \to R_i$ con $c \neq 1$ son no elementales, y deben evitarse. La operación correcta es $R_i + cR_j \to R_i$. En otras palabras, *el renglón que se sustituye no debe multiplicarse por un escalar* $\neq 1$.

Unicidad de la forma de escalón reducida: Pivotes

La eliminación de Gauss implica que cada matriz se convierte a la forma de escalón simple y a la reducida. Una matriz puede ser equivalente a varias matrices en forma de escalón, pero sólo a una en forma de escalón reducida. Ésta es la afirmación del siguiente teorema, cuya demostración se describe en la sección 4.6.

TEOREMA 1 **(Unicidad de la forma de escalón reducida)**

Toda matriz es equivalente a una, y sólo a una matriz en forma de escalón reducida.

Sean A, E y R tres matrices equivalentes tales que E esté en la forma de escalón y R en la de forma escalón reducida. Se dice que E es **una forma de escalón** de A. De acuerdo con el teorema 1, se dice que R es **la forma de escalón reducida** de A.

Observe que en cualquier forma de escalón de una matriz A, los elementos delanteros se encuentran en las mismas columnas. Esto es consecuencia de la unicidad de la forma de escalón reducida, y del hecho que después del paso 4 no se modifican las posiciones de los elementos delanteros. Se llaman **posiciones pivote**, o de pivoteo, de A. Las columnas fijas que contienen las posiciones pivote se llaman **columnas pivote**, o de pivoteo, de A. Un **pivote** es cualquier elemento no cero de una posición pivote.

---

[5] También, a diferencia de las personas, la computadora no necesita intercambio físico de renglones para mantener registro de las operaciones efectuadas. El intercambio de renglones es costoso y lento. Los programadores recurren a una especie de "renombrar renglones" que se llama *método vectorial de permutación*.

Por ejemplo, veamos la matriz A. Las matrices E y R son dos etapas de la reducción de A, siendo E una forma de escalón y R la forma de escalón reducida de A.

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 & 1 \\ 3 & -6 & -1 & 1 & 1 \\ 4 & -8 & 5 & -1 & 14 \end{bmatrix} \to E = \begin{bmatrix} 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & 4 & 0 \end{bmatrix}.$$

$$\to R = \begin{bmatrix} 1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Las columnas 1, 3 y 4 son las columnas pivote de A, E y R. Las posiciones pivote de A, E y R son las posiciones (1, 1), (2, 3) y (3, 4). Los pivotes de E son 1, −1, 4 y los de R son 1, 1, 1.

En general, *las matrices equivalentes tienen las mismas posiciones de columna pivote y las mismas posiciones pivote.*

■ EJEMPLO 17 Las columnas 1, 2, 4 y 5 son las columnas pivote de la matriz en el ejemplo 16. □

Solución de sistemas lineales

Veamos ahora cómo emplear la eliminación de Gauss para resolver cualquier sistema lineal. El proceso se aplica a la matriz aumentada del sistema. Produce una matriz en forma de escalón reducida, cuyo sistema correspondiente es equivalente al sistema dado y además fácil de resolver: primero se separan las variables en **delanteras** y **libres**. Las variables delanteras son las que corresponden a las posiciones pivote. Las variables restantes, si las hay, son libres. A continuación se escriben las variables delanteras en función de las variables libres, de las constantes o de ambas. Se acostumbra asignar nuevos nombres a las variables libres, y llamarlas **parámetros**. Los parámetros pueden asumir cualquier valor escalar.

■ EJEMPLO 18 (Soluciones infinitas) Resuelva el sistema

$$\begin{aligned} 3x_2 - 6x_3 - 4x_4 - 3x_5 &= -5 \\ -x_1 + 3x_2 - 10x_3 - 4x_4 - 4x_5 &= -2 \\ 2x_1 - 6x_2 + 20x_3 + 2x_4 + 8x_5 &= -8 \end{aligned}$$

SOLUCIÓN Por eliminación de Gauss en la matriz aumentada del sistema (la reducción real se deja como ejercicio), se obtiene

$$\begin{bmatrix} 1 & 0 & 4 & 0 & 1 & : & -3 \\ 0 & 1 & -2 & 0 & -1 & : & 1 \\ 0 & 0 & 0 & 1 & 0 & : & 2 \end{bmatrix}$$

Por tanto, el sistema original se reduce al sistema equivalente

$$\begin{aligned} x_1 \qquad + 4x_3 \qquad + x_5 &= -3 \\ x_2 - 2x_3 \qquad - x_5 &= 1 \\ x_4 \qquad &= 2 \end{aligned}$$

Como las columnas pivote son la 1, 2 y 4, x_1, x_2 y x_4 son las variables delanteras, y x_3 y x_5 son las variables libres; estas últimas pueden tener cualquier valor. Sean $x_5 = r, x_3 = s$. A continuación se escriben las variables delanteras x_1 y x_2 en función de r y s y obtenemos

$$\begin{aligned} x_1 &= -4s - r - 3 \\ x_2 &= 2s + r + 1 \\ x_3 &= s \qquad\qquad \text{para toda } r, s \in \mathbf{R} \\ x_4 &= 2 \\ x_5 &= r \end{aligned}$$

Ésta es la solución general del sistema. El conjunto solución es un conjunto biparamétrico infinito. ☐

■ EJEMPLO 19 (Una solución) Determine si los cinco planos definidos por las ecuaciones siguientes pasan por el mismo punto.

$$\begin{aligned} -x \qquad + z &= -2 \\ 2x - y + z &= 1 \\ -3x + 2y - 2z &= -1 \\ x - 2y + 3z &= -2 \\ 5x + 2y + 6z &= -1 \end{aligned}$$

SOLUCIÓN La reducción de la matriz aumentada por operaciones de renglón produce

$$\begin{bmatrix} 1 & 0 & 0 & : & 1 \\ 0 & 1 & 0 & : & 0 \\ 0 & 0 & 1 & : & -1 \\ 0 & 0 & 0 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

En consecuencia, $x = 1, y = 0$ y $z = -1$, por lo que los cinco planos pasan por el punto cuyas coordenadas son $(1, 0, -1)$. ☐

Observe que tan pronto se conocen las columnas pivote (final del paso 4 en el algoritmo 1), es posible predecir si el sistema es consistente o no: si la última es columna pivote, el sistema es inconsistente. Esto se debe a que uno de los renglones de la forma de escalón de la matriz aumentada tendrá la forma

$$[0 \quad 0 \quad 0 \qquad 0 \quad c]$$

con pivote c (por tanto $c \neq 0$). Pero esto corresponde a la ecuación

$$0x_1 + \cdots + 0x_n = c$$

que es falso sean cuales fueren los valores de las variables, porque $c \neq 0$. Si el sistema es inconsistente no es necesario continuar con el paso 5.

■ EJEMPLO 20 (Sin solución) Determine si el sistema es consistente o no.

$$\begin{aligned} x_1 \quad\quad + 2x_3 - 2x_4 &= 1 \\ -x_1 + x_2 \quad\quad + x_4 &= -2 \\ x_2 + 2x_3 - x_4 &= 1 \end{aligned}$$

SOLUCIÓN Mediante la reducción se llega a

$$\begin{bmatrix} 1 & 0 & 2 & -2 & : & 1 \\ 0 & 1 & 2 & -1 & : & -1 \\ 0 & 0 & 0 & 0 & : & 2 \end{bmatrix}$$

Como la última columna de la matriz aumentada es una columna de pivote, el sistema es inconsistente. $\square$

Naturalmente que durante *cualquier* etapa de la reducción, la presencia de un renglón como este implica que el sistema es inconsistente. Con frecuencia, no se necesita completar la eliminación de Gauss, como se muestra en la reducción $\begin{bmatrix} 1 & 1 & : & 0 \\ 1 & 1 & : & 2 \\ 3 & 4 & : & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & : & 0 \\ 0 & 0 & : & 2 \\ 3 & 4 & : & 5 \end{bmatrix}$

Algoritmo 2

(Solución de un sistema lineal)

Para resolver cualquier sistema lineal:

Paso 1. *Aplique la eliminación de Gauss a la matriz aumentada del sistema (paso directo). Si durante cualquier etapa de este proceso nota que la última columna es de pivote, deténgase. En este caso, el sistema es inconsistente. En caso contrario, continúe con el paso 2.*

Paso 2. *Termine la eliminación de Gauss. Escriba el sistema que corresponde a la forma de escalón reducida de la matriz aumentada, sin tener en cuenta las ecuaciones con ceros.*

Paso 3. *Separe las variables del sistema reducido en delanteras y libres (si las hay). Escriba las delanteras en función de las variables libres o de constantes.*

Ahora sacaremos algunas conclusiones importantes del estudio del algoritmo 2.

TEOREMA 2

(Existencia de soluciones)

Un sistema lineal es consistente si y sólo si la última columna de su matriz aumentada no es columna pivote, o bien si cualquier forma de escalón de la matriz aumentada no tiene un renglón de la forma

$$[0 \quad 0 \quad 0 \quad \cdots \quad 0 \quad : \quad c] \quad \text{en la que } c \neq 0.$$

El paso 1 no sólo nos indica si el sistema es consistente o no, sino también la cantidad de soluciones. Si un sistema consistente tiene variables libres, entonces el paso 3 señala que hay soluciones infinitas (dejando que los parámetros asuman cualquier valor). Si no hay variables libres, entonces las variables delanteras serán constantes, y únicamente obtendremos una solución. Como dichas variables corresponden a columnas pivote, vemos que un sistema consistente tiene sólo una solución si todas las columnas, con excepción de la última, son pivote. Todo lo anterior se resume como sigue:

TEOREMA 3

(Unicidad de soluciones)

1. Un sistema lineal consistente tiene solamente una solución si y sólo si no tiene variables libres.
2. Un sistema lineal consistente tiene solamente una solución siempre y cuando cada columna de la matriz aumentada, excepto la última, sea de pivote, y la última no sea columna pivote.

Debemos tener en cuenta que la presencia de variables libres no garantiza que haya una cantidad infinita de soluciones, porque el sistema puede ser inconsistente, como en el ejemplo 20.

■ **EJEMPLO 21** ¿Qué se puede decir acerca de los sistemas cuyas matrices aumentadas se reducen a la forma de escalón? (No es necesario especificar las literales constantes ni trazar las líneas de separación.)

$$\begin{bmatrix} 1 & a & b & d & g \\ 0 & 2 & c & e & h \\ 0 & 0 & 3 & f & i \\ 0 & 0 & 0 & 4 & j \end{bmatrix} \begin{bmatrix} 1 & a & b & c & e \\ 0 & 0 & 2 & d & f \\ 0 & 0 & 0 & 3 & g \end{bmatrix} \begin{bmatrix} 1 & a & b & d & f \\ 0 & 2 & c & e & g \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & a & b & c \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

SOLUCIÓN Los dos primeros sistemas son consistentes, porque sus últimas columnas no son de pivote (teorema 2). Los dos últimos sistemas son inconsistentes, porque sus últimas columnas sí son de pivote (teorema 2). El primer sistema nada más tiene una solución, porque cada columna es de pivote, con excepción de la última (teorema 3). El segundo sistema tiene soluciones infinitas, porque hay una columna no pivote (la segunda) que no es la última (teorema 3). □

UN ERROR FRECUENTE El segundo sistema tiene una solución, aun cuando $g = 0$. No debe confundirse $[0 \quad 0 \quad 0 \quad 3 \quad 0]$ con $[0 \quad 0 \quad 0 \quad 0 \quad 3]$.

Los teoremas 2 y 3 implican lo siguiente.

TEOREMA 4

(Cantidad de soluciones)

Para cualquier sistema lineal, sólo es válida una de las propiedades siguientes:

1. El sistema tiene solamente una solución.
2. El sistema posee soluciones infinitas.
3. El sistema no tiene soluciones.

Por último, nos concentraremos en el caso de un sistema lineal homogéneo.

TEOREMA 5

(Soluciones de sistemas lineales homogéneos)

1. Un sistema lineal homogéneo tiene sólo la solución trivial, o bien un número infinito de soluciones.
2. Un sistema lineal homogéneo tiene una gran cantidad de soluciones, siempre y cuando posea variables libres.
3. Si un sistema lineal homogéneo tiene más incógnitas que ecuaciones, entonces mostrará una infinidad de soluciones.

DEMOSTRACIÓN Cualquier sistema lineal homogéneo es consistente porque tiene la solución trivial. De esta forma, las partes 1 y 2 son consecuencia del teorema 3. Para demostrar la parte 3, observamos que como el sistema de la forma de escalón reducida de la matriz aumentada tiene más incógnitas que ecuaciones, deben existir variables libres. Por consiguiente, el sistema tiene un número infinito de soluciones, de acuerdo con la parte 2. □

Obsérvese que para los sistemas homogéneos, la presencia de variables libres *sí* garantiza esa infinidad de soluciones.

■ EJEMPLO 22 Demuestre que el sistema tiene soluciones no triviales.

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 - x_2 - x_3 &= 0 \end{aligned}$$

SOLUCIÓN Como es un sistema homogéneo con más incógnitas que ecuaciones, entonces muestra soluciones infinitas; por consiguiente, el sistema tiene un número infinito de soluciones *no triviales*. □

Eliminación de Gauss-Jordan

Una variante interesante de la eliminación de Gauss se presenta si, durante el paso directo, se producen 1 delanteros y después ceros abajo y **arriba** de los anteriores. Así, cuando se termina el paso directo, la matriz se encontrará en su forma de escalón reducida. A este método se le llama **eliminación de Gauss-Jordan**.[6]

■ EJEMPLO 23 Encuentre la forma de escalón reducida de A aplicando la eliminación de Gauss-Jordan.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -2 \\ 0 & 2 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

[6] **Wilhelm Jordan** (1842–1899), ingeniero alemán, autor del conocido *Libro de bolsillo de geometría práctica*. Según Gewirtz, Sitomer y Tucker, "inventó el algoritmo de reducción a pivotes llamado eliminación de Gauss-Jordan, para aplicaciones geodésicas".

SOLUCIÓN Primero se escala el segundo renglón para obtener el 1 delantero. A continuación se obtienen ceros abajo y *arriba* de este 1. Después se repite el proceso.

$$A \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & \boxed{1} & -1 \\ 0 & 2 & 1 \\ 0 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 3 \\ 0 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

□

Reducción por operaciones de renglón con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> rref(matrix([[1,2,3],[2,2,3],[3,3,3]]));
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Mathematica

```
In[1]:=
     RowReduce[{{1,2,3},{2,2,3},{3,3,3}}]
Out[1]=
     {{1, 0, 0}, {0, 1, 0}, {0, 0, 1}}
```

MATLAB

```
>> rref([1 2 3; 2 2 3; 3 3 3])
ans =
     1     0     0
     0     1     0
     0     0     1
```

Ejercicios 1.2

Forma de escalón

En los ejercicios 1 a 5 ponga cada matriz en una de las siguientes categorías:

1. Forma de escalón, pero *no* forma de escalón reducida.
2. Forma de escalón reducida.
3. Diferente de la forma de escalón.

1. a. $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

c. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **d.** $\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$

2. a. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

c. $\begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ d. $\begin{bmatrix} 1 & 0 & 0 & -6 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

3. a. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 4 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ d. $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4. a. $\begin{bmatrix} 1 & 9 & 0 & -7 & 0 \\ 0 & 0 & 1 & 8 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

b. $\begin{bmatrix} 0 & 1 & 8 & 6 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 8 & 0 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 6 & 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$

5. a. $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 & 0 & 7 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 5 & 0 & -1 & 0 \\ 0 & 0 & 1 & 9 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 4 & 0 & 5 & 0 & 6 \\ 0 & 0 & 1 & 4 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$

Matrices equivalentes: reducción con operaciones de renglón

6. Demuestre que cada una de las operaciones elementales de renglón es *reversible*. En otras palabras, si se usa una operación para obtener la matriz B a partir de A, hay una operación elemental de renglón que invertirá el efecto de la primera, transformando B de nuevo en A.

7. Aproveche el ejercicio 6 para demostrar que

Si $A \sim B$, entonces $B \sim A$.

8. Demuestre que

Si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

9. Emplee los resultados de los ejercicios 7 y 8 para demostrar la equivalencia entre $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix}$ primero compruebe que ambas son equivalentes a $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

10. Demuestre que

$$\begin{bmatrix} 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

11. ¿Son equivalentes las dos matrices siguientes?

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & -1 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & \frac{7}{9} \\ 0 & 1 & \frac{10}{9} \end{bmatrix}$$

12. Demuestre que las matrices siguientes *no son* equivalentes.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & -1 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 10 \end{bmatrix}$$

13. Compruebe que las matrices siguientes *no son* equivalentes.

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

Sea $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

14. Demuestre que si $ad - bc \neq 0$, la forma de escalón reducida de $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es I.

15. Aplique el resultado del ejercicio 14 para comprobar que la matriz siguiente se reduce a I para cualquier θ.

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

16. Reduzca las matrices siguientes a la forma de escalón.

a. $\begin{bmatrix} 0 & -1 & 0 & -1 \\ 2 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$ b. $\begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$

17. Calcule dos formas de escalón distintas para cada una de las matrices siguientes.

a. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 0 & -1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

En los ejercicios 18 a 22 determine la forma de escalón reducida de cada matriz.

18. $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{bmatrix}$

19. $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \end{bmatrix}$

20. $\begin{bmatrix} -1 & -6 & 0 & 1 & 0 & -2 \\ 5 & 30 & 1 & -6 & 0 & 13 \\ 4 & 24 & 1 & -5 & 1 & 15 \end{bmatrix}$

21. $\begin{bmatrix} -1 & -4 & 0 & -5 & 0 & -6 \\ 5 & 20 & 1 & 29 & 0 & 34 \\ 4 & 16 & 1 & 24 & 1 & 30 \end{bmatrix}$

22. $\begin{bmatrix} 0 & -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$

Sistemas lineales

En los ejercicios 23 a 33 resuelva los sistemas de ecuaciones.

23. $$\begin{aligned} x \quad + z + w &= -5 \\ x \quad - z + w &= -1 \\ x + y + z + w &= -3 \\ 2x \quad + 2z \quad &= -2 \end{aligned}$$

24. $$\begin{aligned} x_1 - 8x_2 \quad + 7x_4 \quad &= 9 \\ -2x_1 + 16x_2 - x_3 - 20x_4 \quad &= -24 \\ 2x_1 - 16x_2 + 6x_3 + 50x_4 + x_5 &= 51 \end{aligned}$$

25. $$\begin{aligned} x_1 \quad + 4x_3 + 5x_4 \quad &= 0 \\ -2x_1 - x_2 - 10x_3 - 16x_4 \quad &= -6 \\ 2x_1 + 6x_2 + 20x_3 + 46x_4 + x_5 &= 33 \end{aligned}$$

26. $$\begin{aligned} 6x_1 \quad - x_3 + 4x_4 &= 0 \\ 2x_1 - x_2 - x_3 - 6x_4 &= -6 \\ 16x_1 - 2x_2 - 4x_3 - 4x_4 &= 12 \end{aligned}$$

27. $$\begin{aligned} x - y + z + 2w &= 0 \\ x \quad + \quad w &= -1 \\ y - z - w &= 1 \\ x + 2y \quad &= -3 \end{aligned}$$

28. $$\begin{aligned} x \quad + 2z - w &= -1 \\ y + z + 2w &= 0 \\ x - y - 2z - w &= 4 \\ y + z \quad &= -2 \end{aligned}$$

29. $$\begin{aligned} 3x \quad + z + w &= -3 \\ 2y + 3z + 3w &= -14 \\ z - 2w &= 6 \\ w &= -4 \end{aligned}$$

30. $$\begin{aligned} x + y + z + w - t &= 1 \\ y \quad &= -1 \\ -2z - w + t &= -3 \\ w - 3t &= -1 \\ t &= 1 \end{aligned}$$

31. $$\begin{aligned} x \quad - t &= -2 \\ y \quad - z + t &= 5 \\ -y \quad + z - t &= -5 \\ y \quad - z + t &= 5 \\ -y - w \quad &= -1 \end{aligned}$$

32. $$\begin{aligned} x + 2y + 3z + 4w &= 0 \\ 2x + 2y + 3z + 4w &= 0 \\ 3x + 3y + 3z + 4w &= 0 \end{aligned}$$

33. $$\begin{aligned} x + 2y + 3z + 4w &= 0 \\ 2x + 2y + 3z + 4w &= 0 \\ 3x + 3y + 3z + 4w &= 0 \\ 4x + 4y + 4z + 4w &= 0 \end{aligned}$$

34. Resuelva el sistema para x, y y z.

$$\begin{aligned} x + y \quad &= a \\ x - y \quad &= b \\ x + y + z &= c \end{aligned}$$

35. Solucione el sistema para x, y y z.

$$\begin{aligned} x + y \quad &= a \\ x - y \quad &= b \\ x - y + z &= 0 \end{aligned}$$

36. Demuestre que los valores de λ, para los cuales el sistema

$$(a-\lambda)x + by = 0$$
$$cx + (d-\lambda)y = 0$$

tiene soluciones no triviales, deben satisfacer la ecuación cuadrática $\lambda^2 - (a+d)\lambda + ad - bc = 0$.

En los ejercicios 37 a 41 resuelva los sistemas con las matrices aumentadas dadas.

37. $\begin{bmatrix} 1 & -1 & -2 & 2 & -7 \\ -2 & 2 & 0 & 2 & -2 \\ -1 & 1 & 1 & -2 & 6 \\ -2 & 2 & -1 & 2 & -3 \end{bmatrix}$

38. $\begin{bmatrix} 0 & 2 & -2 & -2 & 10 \\ -2 & -2 & 2 & 2 & -2 \\ -1 & -1 & 0 & 0 & 3 \\ -1 & -1 & 1 & -2 & -1 \end{bmatrix}$

39. $\begin{bmatrix} -1 & 0 & 1 & 1 & -1 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & -1 & 1 & -6 \\ 0 & 1 & 1 & -1 & 0 & 3 \\ 1 & -1 & -1 & 0 & 1 & 2 \end{bmatrix}$

40. $\begin{bmatrix} 0 & -1 & -1 & -1 & -1 & 8 \\ 0 & 0 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 0 & -1 & -2 \\ 0 & 1 & 0 & 0 & 1 & -5 \\ -1 & 1 & -1 & 1 & 0 & -3 \end{bmatrix}$

41. $\begin{bmatrix} 1 & -1 & -1 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 & 1 & 0 \\ -1 & -1 & 1 & 1 & -1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

42. Resuelva el sistema homogéneo de cada matriz de coeficientes

a. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$

43. Verifique que los sistemas cuyas matrices aumentadas aparecen abajo son equivalentes.

$$A = \begin{bmatrix} -1 & -4 & 0 & -5 & 0 & -6 \\ 5 & 20 & 1 & 29 & 0 & 34 \\ 4 & 16 & 1 & 24 & 1 & 30 \end{bmatrix}$$

$$B = \begin{bmatrix} -1 & -4 & 0 & -5 & 0 & -6 \\ 2 & 8 & -1 & 6 & 0 & 8 \\ -2 & -8 & 3 & 2 & 1 & 2 \end{bmatrix}$$

44. La solución del sistema cuya matriz aumentada A codifica un mensaje puede obtenerse como sigue: cada letra del alfabeto se numera con su orden alfabético. ¿Cuál es el mensaje?

$$A = \begin{bmatrix} 1 & -1 & -1 & -1 & 0 & 0 & 0 & -27 \\ 0 & 0 & -1 & 0 & -1 & 0 & -1 & -27 \\ 0 & -1 & 0 & 1 & -1 & 0 & 0 & -21 \\ 1 & 1 & 1 & 0 & -1 & 0 & 0 & 27 \\ 1 & 1 & 0 & 1 & -1 & 0 & 0 & 16 \\ 0 & 0 & 1 & 1 & -1 & -1 & -1 & -8 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 13 \end{bmatrix}$$

En los ejercicios 45 y 46 considere sistemas lineales cuyas matrices aumentadas se reducen a las siguientes formas de escalón. ¿Qué puede decir acerca de los sistemas?

45. **a.** $\begin{bmatrix} 2 & a & b & d & f \\ 0 & 2 & c & e & g \\ 0 & 0 & 0 & 2 & h \end{bmatrix}$ **b.** $\begin{bmatrix} 2 & a & b & d \\ 0 & 2 & c & e \\ 0 & 0 & 2 & f \\ 0 & 0 & 0 & 2 \end{bmatrix}$

46. **a.** $\begin{bmatrix} 1 & a & b & c & d \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ **b.** $\begin{bmatrix} 2 & a & b & c \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 47 y 48, cada renglón de la tabla indica el tamaño y la cantidad de columnas pivote de la matriz aumentada de algún sistema. ¿Qué opina acerca del sistema?

47.

Tamaño	Cantidad de columnas pivote
3×5	3
4×4	4
4×4	3
5×3	3

48.

Tamaño	Cantidad de columnas pivote
6×4	4
5×5	5
5×5	4
4×6	4

49. Demuestre que si el tamaño de una matriz es $m\times n$, la cantidad de columnas pivote es menor o igual a m y menor o igual a n.

50. Compruebe que si el sistema lineal cuya matriz aumentada $[A : \mathbf{c}]$ es inconsistente, entonces aquél cuya matriz aumentada es $[A : \mathbf{b}]$ puede ser inconsistente, o presenta un número infinito de soluciones.

51. Verifique que si el sistema lineal cuya matriz aumentada es $[A : \mathbf{c}]$ tiene solamente una solución, entonces el que po-

sce la matriz aumentada $[A : \mathbf{b}]$ también presenta exactamente una solución.

En los ejercicios 52 y 53 calcule los valores de a tales que el sistema cuya matriz aumentada se indica tenga (a) exactamente una solución, (b) infinitas soluciones y (c) ninguna solución.

52. a. $\begin{bmatrix} 2 & 3 & 4 \\ 4 & a & 8 \end{bmatrix}$ b. $\begin{bmatrix} 2 & 3 & 4 \\ 4 & 6 & a \end{bmatrix}$

53. a. $\begin{bmatrix} 2 & 3 & 4 \\ a & 6 & 8 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 2 & 1 & 3 \\ 1 & 3 & -1 & 4 \\ 1 & 2 & a^2-8 & a \end{bmatrix}$

1.3 Soluciones numéricas

Objetivos del alumno para esta sección

1. Comprender algunas de las consideraciones numéricas en la solución de sistemas.
2. Ver cómo se resuelven con frecuencia los problemas en la práctica.

La mayor parte de los sistemas lineales que aparecen en las aplicaciones no pueden resolverse a mano, porque en forma característica consisten en cientos o miles de ecuaciones e incógnitas. Es absolutamente necesario usar computadoras con programas eficientes. Aunque las eliminaciones de Gauss y de Gauss-Jordan son métodos muy importantes para resolver un sistema, no siempre son los más eficientes, aun cuando se use una computadora. En esta sección describiremos dos métodos nuevos que se usan mucho en la práctica, así como algunas de las consideraciones numéricas que se presentan en la solución de sistemas particulares.

Primero compararemos nuestros dos métodos de eliminación. Obsérvese que éstos son **directos**, lo cual significa que siempre se obtiene la solución con una cantidad finita de pasos. De hecho, esa cantidad de pasos puede calcularse bien.

Elección entre las eliminaciones de Gauss y de Gauss-Jordan

En la sección 1.2 dedicamos un poco más de tiempo a estudiar la eliminación de Gauss, mientras que apenas se ilustró la de Gauss-Jordan. La razón es que la de Gauss-Jordan, aunque aparentemente más eficiente (porque no hay paso hacia atrás), requiere más operaciones aritméticas. De hecho, para un sistema de n ecuaciones con n incógnitas, cuando n es grande, puede demostrarse que la eliminación de Gauss necesita aproximadamente $2n^3/3$ operaciones aritméticas. Por otra parte, la eliminación de Gauss-Jordan requiere alrededor de n^3 operaciones. Son 50% más operaciones que en la eliminación de Gauss. Así, cuando se usa la eliminación de Gauss-Jordan para un sistema mediano, digamos de 500 ecuaciones con 500 incógnitas, se requieren casi 125 millones de operaciones, y sólo unas 83 millones con la de Gauss. Es la causa principal por la que se prefiere la eliminación de Gauss.

A pesar de lo anterior, uno no debe apresurarse a desechar la eliminación de Gauss-Jordan. De hecho, en el cómputo paralelo moderno, un algoritmo paralelo de Gauss-Jordan es ligeramente más eficiente que el correspondiente de Gauss.

Métodos iterativos

Además de los métodos directos, también se cuenta con **métodos iterativos**, en los que se trata de aproximar la solución de un sistema recurriendo a iteraciones que comienzan con un cálculo aproximado inicial. Si las iteraciones sucesivas se acercan a la solución, se dice que la iteración **converge**. En caso contrario, se dice que **diverge**. El procedimiento termina cuando en dos iteraciones sucesivas se obtiene la misma respuesta con una precisión especificada. A diferencia de los métodos directos, la cantidad de pasos necesarios no se conoce con anticipación. Describiremos dos métodos iterativos, el de Jacobi[7] y el de Gauss-Seidel.[8]

Iteración de Jacobi

La iteración de Jacobi se aplica a **sistemas cuadrados**, es decir, aquéllos que tienen tantas ecuaciones como incógnitas. Supongamos que hay un sistema de n ecuaciones con n incógnitas $x_1, \ldots, x_n$, como el siguiente:

$$\begin{aligned} 5x + y - z &= 14 \\ x - 5y + 2z &= -9 \\ x - 2y + 10z &= -30 \end{aligned} \tag{1.7}$$

Paso 1. Despeje x_i de la i-ésima ecuación del sistema.

$$\begin{aligned} x &= -0.2y + 0.2z + 2.8 \\ y &= 0.2x + 0.4z + 1.8 \\ z &= -0.1x + 0.2y - 3.0 \end{aligned} \tag{1.8}$$

Paso 2. Comience con unos valores iniciales $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$ de la solución. En ausencia de cualquier información, inicialice a cero todas las variables: $x_1^{(0)} = 0, x_2^{(0)} = 0, \ldots, x_n^{(0)} = 0$.

En nuestro ejemplo, sean $x^{(0)} = 0$, $y^{(0)} = 0$ y $z^{(0)} = 0$.

Paso 3. Sustituya los valores $x_1^{(k-1)}, x_2^{(k-1)}, \ldots, x_n^{(k-1)}$ determinados después de la $(k-1)$-ésima iteración, en el lado derecho de (1.8), para obtener los nuevos valores $x_1^{(k)}, x_2^{(k)}, \ldots x_n^{(k)}$.

En nuestro ejemplo, la sustitución $x = 0, y = 0, z = 0$ en el lado derecho de (1.8) produce $x = 2.8, y = 1.8, z = -3$. A continuación, al sustituir esos nuevos valores en el lado derecho de (1.8) se obtiene $x = 1.84, y = 1.16$ y $z = -2.92$. Repita el procedimiento.

Paso 4. Detenga el proceso cuando alcance la precisión deseada. Por lo general, esto ocurre cuando se obtienen los mismos valores con la precisión indicada en dos iteraciones consecutivas.

En nuestro ejemplo iteramos con una precisión de cuatro cifras decimales y nos detenemos cuando encontramos dos respuestas consecutivas iguales.

Las iteraciones sugieren que $x = 2, y = 1$ y $z = -3$ es la solución correcta del sistema, al menos con cuatro cifras decimales. En este caso en particular es la solución exacta.

---

[7] **Karl Gustav Jacobi** (1804–1851), eminente matemático alemán. Fue profesor en la Universidad de Könisberg, y aportó contribuciones fundamentales a la teoría de las funciones elípticas y a la teoría de las ecuaciones diferenciales.

[8] **Philipp Ludwig Seidel** (1821–1896), matemático alemán, profesor en Munich e investigador en análisis y en astronomía.

Iteración	x	y	z
Valor inicial	0.0000	0.0000	0.0000
1	2.8000	1.8000	−3.0000
2	1.8400	1.1600	−2.9200
3	1.9840	1.0000	−2.9520
4	2.0096	1.0160	−2.9984
5	1.9971	1.0026	−2.9978
6	1.9999	1.0003	−2.9992
7	2.0001	1.0003	−2.9999
8	2.0000	1.0001	−3.0000
9	2.0000	1.0000	−3.0000
10	2.0000	1.0000	−3.0000

Iteración de Gauss-Seidel

Esta iteración también se aplica a sistemas cuadrados, y sus pasos son los siguientes:

Paso 1. Igual que en la iteración de Jacobi.

Paso 2. Igual que en la iteración de Jacobi.

Paso 3. Sustituya la incógnita *calculada más recientemente* en el lado derecho de las ecuaciones obtenidas en el paso 1, para tener la nueva aproximación, $x_i^{(k)}$.

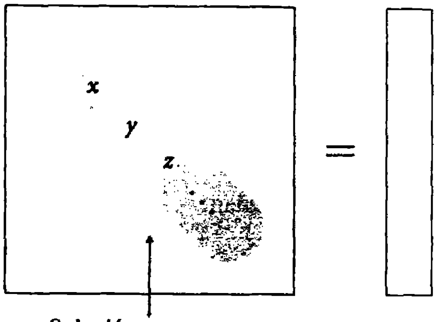
En nuestro ejemplo, la sustitución de $y = 0$, $z = 0$, en la primera ecuación da como resultado $x = 2.8$. En la segunda ecuación se reemplaza $z = 0$ y $x = 2.8$ (el valor más reciente de x) para obtener $y = 2.36$. En la tercera ecuación se sustituye $x = 2.8$ y $y = 2.36$ (las últimas x y y) para obtener $z = -2.808$. Continúe de esta manera.

Paso 4. Igual que en la iteración de Jacobi.

En nuestro ejemplo se obtiene lo siguiente.

Iteración	x	y	z
Valor inicial	0.0000	0.0000	0.0000
1	2.8000	2.3600	−2.8080
2	1.7664	1.0301	−2.9706
3	1.9999	1.0117	−2.9976
4	1.9981	1.0006	−2.9997
5	1.9999	1.0001	−3.0000
6	2.0000	1.0000	−3.0000
7	2.0000	1.0000	−3.0000

Observe que el método de Gauss-Seidel requirió menos iteraciones que el de Jacobi. Esto *parece* ser cierto en la mayor parte de los casos, pero no siempre. Desafortunadamente no se conoce de antemano cuál método es el más eficiente.



Solución con
Jacobi: **Actualice el sistema.**
Gauss-Seidel: **Actualice cada ecuación.**

Convergencia

Una condición suficiente para que converjan las iteraciones de Jacobi y de Gauss-Seidel es que la matriz de coeficientes del sistema sea **diagonalmente dominante**. Esto significa: (1) que la matriz es **cuadrada**, es decir, que tiene la misma cantidad de renglones y columnas, y (2) cada elemento (i, i), llamado elemento diagonal, tenga un valor absoluto mayor que la suma de los valores absolutos de los demás elementos del mismo renglón.

Por ejemplo, el sistema (1.7) tiene la matriz de coeficientes

$$\begin{bmatrix} 5 & 1 & -1 \\ 1 & -5 & 2 \\ 1 & -2 & 10 \end{bmatrix}$$

que es diagonalmente dominante, porque $|5| > |1| + |-1|$, $|-5| > |1| + |2|$, y $|10| > |1| + |-2|$. De este modo se garantiza que, en este caso, ambas iteraciones convergerán.

La matriz

$$\begin{bmatrix} 4 & 2 & -1 \\ 3 & -5 & 2 \\ 1 & -2 & 10 \end{bmatrix}$$

no es diagonalmente dominante, porque en el segundo renglón $|-5| = 5$ *no* es mayor que $|3| + |2| = 5$.

Observe que las iteraciones de Jacobi y de Gauss-Seidel pueden converger aun cuando la matriz de coeficientes del sistema no sea diagonalmente dominante (ejercicio 15).

A veces, un rearreglo de las ecuaciones dará como resultado una matriz de coeficientes diagonalmente dominante. Por ejemplo, el sistema

$$\begin{aligned} 2x + 4y - z &= 1 \\ x - 5y + 2z &= 2 \\ x - 2y + 10z &= 3 \end{aligned}$$

tiene una matriz de coeficientes que no es diagonalmente dominante. Sin embargo, si se intercambian las ecuaciones primera y segunda, la nueva matriz de coeficientes sí lo será.

$$\begin{bmatrix} 1 & -5 & 2 \\ 2 & 4 & -1 \\ 1 & -2 & 10 \end{bmatrix}$$

Elección entre la eliminación de Gauss y la iteración de Gauss-Seidel

Veamos cómo se comparan entre sí nuestros métodos aparentemente más eficientes, directo e iterativo. Puede demostrarse que para n grande, el método de Gauss-Seidel requiere unas $2n^2$ operaciones aritméticas por iteración. Si se necesitan menos que iteraciones $n/3$, la cantidad total de operaciones será menor que $2n^3/3$, y el de Gauss-Seidel será más eficiente que el método de Gauss. Por ejemplo, para que el método de Gauss-Seidel sea la mejor alternativa en un sistema cuadrado de 500 ecuaciones, se requiere que no haya más de 166 iteraciones.

En la práctica, con frecuencia se prefiere al método de Gauss-Seidel frente al de Gauss, aun cuando haya que efectuar más operaciones, porque durante el proceso de eliminación de este último se acumulan los errores de redondeo de la computadora en cada operación elemental de renglón, y afectan a la respuesta final. En la iteración de Gauss-Seidel hay sólo un error de redondeo, debido a la última iteración. En realidad, puede considerarse que la penúltima iteración es un valor inicial excelente.

Otra virtud de la iteración de Gauss-Seidel es el ser **autocorrectiva**. Si en cualquier etapa hubo un error de cálculo, puede seguirse usando la respuesta; tan sólo se considera como un nuevo valor inicial.

Por último, tanto la de Jacobi como la de Gauss-Seidel son excelentes opciones si la matriz de coeficientes es **dispersa**, es decir, tiene muchos elementos cero. Esto se debe a que en cada etapa se usan los mismos coeficientes, y los ceros se conservan en el proceso.

Consideraciones numéricas: Mal condicionado y pivoteo

Algunos sistemas (aun los pequeños) presentan un comportamiento que requiere un análisis numérico cuidadoso. Veamos los sistemas casi idénticos,

$$\begin{aligned} x + y &= 1 \\ 1.01x + y &= 2 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x + y &= 1 \\ 1.005x + y &= 2 \end{aligned}$$

La solución exacta del primero es $x = 100$, $y = -99$, mientras que la del segundo es $x = 200$, $y = -199$. Así, un cambio pequeño en los coeficientes ocasionó una modificación drástica en la solución. A este tipo de sistemas se les llama **mal condicionados**.

Por ejemplo, si usáramos aritmética de punto flotante con dos cifras decimales de precisión, nuestra aproximación a la solución del segundo sistema estaría equivocada un 50%. En este caso, la razón de este comportamiento es que las dos rectas definidas por el primer sistema son casi paralelas. Así, un cambio pequeño en la pendiente de 1 puede mover el punto de intersección hasta una distancia bastante apreciable (figura 1.5).

Cuando se usa la eliminación de Gauss o la de Gauss-Jordan con aritmética de punto flotante, se presenta otra clase de problemas y los elementos de la matriz aumentada de un sistema tienen tamaños considerablemente distintos.

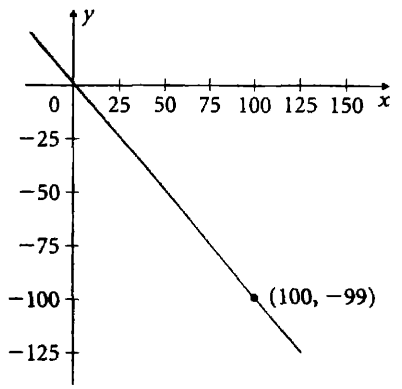


Figura 1.5 Líneas casi paralelas, en un sistema mal condicionado.

Por ejemplo, en el sistema

$$10^{-3}x + y = 2$$
$$2x - y = 0$$

Es fácil ver que la solución exacta es $x = 2\,000/2\,001$ y $y = 4\,000/2\,001$. Suponga que desea resolver numéricamente el sistema, pero que la aritmética de punto flotante sólo puede llevarse a cabo con tres cifras significativas.

SOLUCIÓN 1

$$\begin{bmatrix} 10^{-3} & 1 & 2 \\ 2 & -1 & 0 \end{bmatrix} \boxed{R_2 - 2 \cdot 10^3 R_1 \to R_2} \begin{bmatrix} 10^{-3} & 1 & 2 \\ 0 & \boxed{-2 \cdot 10^3} & -4 \cdot 10^3 \end{bmatrix}$$

El elemento (2, 2) real de la última matriz fue −2 001, que fue redondeado a −2 000 por estar trabajando con tres cifras significativas. El resto de la reducción es el acostumbrado:

$$\sim \begin{bmatrix} 10^{-3} & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 10^{-3} & 0 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

y así se obtiene $x = 0$, $y = 2$. Vemos que la aproximación para x es bastante mala.

SOLUCIÓN 2 Ahora supongamos que se intercambian las ecuaciones 1 y 2, y se escala el primer renglón para obtener un 1 delantero.

$$\begin{bmatrix} 2 & -1 & 0 \\ 10^{-3} & 1 & 2 \end{bmatrix} \boxed{\tfrac{1}{2}R_1 \to R_1} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 10^{-3} & 1 & 2 \end{bmatrix}$$

Entonces

$$\boxed{R_2 - 10^{-3}R_1 \to R_2} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & \boxed{1} & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

El elemento real (2, 2) de la primera matriz era $1 + (1/2) \cdot 10^{-3}$, que en esta aritmética se simplifica a 1. En consecuencia, $x = 1$, y $y = 2$ esta vez constituye una mejor aproximación. $\square$

¿Qué funcionó mal durante la solución 1? El coeficiente pequeño, 10^{-3}, en la primera posición pivote, forzó el manejo de coeficientes grandes en el segundo renglón y esto originó un pequeño error para y debido al redondeo. Y aunque en apariencia es mínimo, causó un error apreciable en la estimación de x, en $10^{-3}x + y = 2$, porque el coeficiente de x fue dominado por el de y.

La segunda solución no presentó este problema, porque el renglón con el elemento delantero de mayor tamaño se trasladó a la posición pivote. Así la eliminación no produjo coeficientes grandes. Y aun cuando mostró el mismo error de redondeo, el valor de x se afectó sólo muy ligeramente.

En la práctica, durante la eliminación de Gauss o de Gauss-Jordan, siempre se cambia el renglón que tiene el elemento delantero con *mayor valor absoluto* a la posición pivote previa al despeje. A esto se le llama **pivoteo parcial**, y ayuda a mantener bajo control al error de redondeo. También hay una variante en la que se elige como pivote al elemento del valor más alto en toda la matriz. Esto obliga al intercambio de *columnas*, además de renglones, lo cual significa que también hay que cambiar las variables. A esto se le llama **pivoteo total**, y produce mejores resultados numéricos, pero puede ser bastante lento. El pivoteo parcial es la modificación más usada de la eliminación de Gauss o de Gauss-Jordan.

Soluciones numéricas mediante sistemas algebraicos computacionales

Maple

```
> fsolve({3*x+2*y=1,2*x-y=-3},{x,y});
```

$$\{y = 1.571428572, x = -.7142857147\}$$

Mathematica

```
In[1]:=
     NSolve[{3x+2y==1,2x-y==-3},{x,y}]
Out[1]=
     {{x -> -0.714286, y -> 1.57143}}
```

MATLAB

```
>> rref([3 2 1; 2 -1 -3])
ans =
    1.0000         0   -0.7143
         0    1.0000    1.5714
```

Ejercicios 1.3

En los ejercicios 1 a 4 determine si la matriz es diagonalmente dominante.

1. $\begin{bmatrix} -4 & 1 \\ 0 & 5 \end{bmatrix}$

2. $\begin{bmatrix} 4 & 2 \\ -5 & 5 \end{bmatrix}$

3. $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 5 & 1 \\ 0 & 1 & 4 \end{bmatrix}$

4. $\begin{bmatrix} 5 & 1 & -1 \\ -1 & 5 & 1 \\ 1 & 1 & 5 \end{bmatrix}$

En los ejercicios 5 y 6 replantee el sistema para que su (nueva) matriz de coeficientes sea diagonalmente dominante.

5. $\begin{aligned} x - 2y &= -6 \\ 5x + y &= 14 \end{aligned}$

6. $\begin{aligned} x + y + 5z &= 15 \\ -x + 5y + z &= -9 \\ 5x + y - z &= 5 \end{aligned}$

7. Aplique el método de Jacobi con cuatro iteraciones y con los valores iniciales $x = 1, y = 1$, para aproximar la solución del sistema. Compare sus resultados con la solución exacta.

$$\begin{aligned} 5x + y &= 14 \\ x - 2y &= -6 \end{aligned}$$

8. Repita el ejercicio 7 con la iteración de Gauss-Seidel.

En los ejercicios 9 a 11 calcule soluciones aproximadas del sistema empleando el método de Jacobi con 4 iteraciones. Inicialice todas las variables en 0 .

9. $\begin{aligned} 7x \qquad - z &= 9 \\ -x + 4y \qquad &= 19 \\ y - 9z &= 23 \end{aligned}$

10. $\begin{aligned} 4x + y + z &= 17 \\ x + 4y + z &= 2 \\ x + y + 4z &= 11 \end{aligned}$

11. $\begin{aligned} 5x + y - z &= 5 \\ -x + 5y + z &= -9 \\ x + y + 5z &= 15 \end{aligned}$

En los ejercicios 12 a 14 calcule soluciones aproximadas del sistema con el método de Gauss-Seidel y cuatro iteraciones. Inicialice todas las variables en cero.

12. El sistema del ejercicio 9.

13. El sistema del ejercicio 10.

14. El sistema del ejercicio 11.

15. Las matrices de coeficientes de los sistemas de abajo no son diagonalmente dominantes. Aplique la iteración de Gauss-Seidel, inicializando $x = 0, y = 0$, e itere 5 veces. Demuestre que: (a) la iteración *diverge* en el primer sistema. (b) La iteración en el segundo sistema *converge* con dos cifras decimales. (La diferencia entre las dos últimas iteraciones de cada variable es < 0.005.)

a. $\begin{aligned} x - y &= 2 \\ x + y &= 0 \end{aligned}$ **b.** $\begin{aligned} 4x - y &= -3 \\ x + y &= 0 \end{aligned}$

En los ejercicios 16 a 19 utilice un pivoteo parcial en la eliminación de Gauss para resolver el sistema. Redondee a cuatro cifras significativas.

16. $\begin{aligned} x - 3y &= -11 \\ 10x + 5y &= 30 \end{aligned}$

17. $\begin{aligned} 1.2x - 4.5y &= -1.23 \\ -5.5x + y &= -15.95 \end{aligned}$

18. $\begin{aligned} x + 2y + 2z &= 6 \\ 2x + 4y + z &= 9 \\ 8x + 2y + z &= 19 \end{aligned}$

19. $\begin{aligned} 1.5x + 2.2y + 2.4z &= 3.2 \\ 2.5x + 4.2y + 1.5z &= 2.3 \\ -8.4x + 2.2y + 1.5z &= -17.5 \end{aligned}$

20. **(Escalamiento)** En el sistema siguiente todos los coeficientes de x son de distintos órdenes de magnitud que el resto de los coeficientes. En estos casos, los cálculos se simplifican si se escala la variable. Para este sistema, sea $x' = 0.001x$. Escriba el sistema con las variables x', y y z, y resuélvalo con la eliminación de Gauss. A continuación calcule x.

$$\begin{aligned} 0.004x + y - z &= 15.8 \\ 0.001x + 5y + z &= 14.2 \\ 0.001x + y + 5z &= -29.8 \end{aligned}$$

1.4 Aplicaciones

Objetivo del alumno para esta sección

Adquirir una idea de las cuantiosas aplicaciones de los sistemas lineales.

En esta sección describiremos algunas aplicaciones de los sistemas lineales a problemas antiguos y modernos. Al lector le agradará saber que aun con los pocos medios que ha aprendido hasta ahora, puede resolver —o abordar— una diversidad de problemas de la vida real en varios campos de la ciencia.

Ahora que el lector ya tiene bastante práctica en la solución de sistemas lineales, *nos saltaremos*, en la mayor parte de los casos, la descripción de la solución de un sistema lineal.

Asuntos de manufactura, sociales y financieros

■ EJEMPLO 24 (Manufactura) R.S.C.L.S y Asociados fabrica tres tipos de computadora personal: Ciclón, Cíclope y Cicloide. Para armar una Ciclón se necesitan 10 horas, otras 2 para probar sus componentes y 2 horas más para instalar sus programas. El tiempo requerido para la Cíclope es 12 horas en su ensamblado, 2.5 para probarla y 2 horas para instalarla. La Cicloide, la más sencilla de la línea, necesita 6 horas de armado, 1.5 horas de prueba y 1.5 horas de instalación. Si la fábrica de esta empresa dispone de 1 560 horas de trabajo por mes para armar, 340 horas para probar y 320 horas para instalar, ¿cuántas PC de cada tipo puede producir en un mes?

SOLUCIÓN Sean x, y, z las cantidades de Ciclones, Cíclopes y Cicloides producidas cada mes. Entonces se necesitan $10x + 12y + 6z$ horas para armar las computadoras. Por consiguiente, $10x + 12y + 6z = 1\,560$. En esta misma forma se obtienen ecuaciones para la prueba y la instalación. El sistema que resulta es

$$\begin{aligned} 10x + 12y + 6z &= 1560 \\ 2x + 2.5y + 1.5z &= 340 \\ 2x + 2y + 1.5z &= 320 \end{aligned}$$

cuya solución es $x = 60$, $y = 40$ y $z = 80$. Por consiguiente, cada mes se pueden fabricar 60 Ciclones, 40 Cíclopes y 80 Cicloides. □

■ EJEMPLO 25 (Cambio de moneda extranjera) Una empresaria internacional necesita, en promedio, cantidades fijas de yenes japoneses, libras inglesas y marcos alemanes durante cada viaje de negocios. Este año viajó 3 veces. La primera vez cambió un total de \$2 550 con las siguientes tasas: 100 yenes por dólar, 0.6 libras por dólar y 1.6 marcos por dólar. La segunda vez cambió \$2 840 en total con las tasas de 125 yenes, 0.5 libras y 1.2 marcos por dólar. La tercera vez, cambió un total de \$2 800 a 100 yenes, 0.6 libras y 1.2 marcos por dólar. ¿Cuántos yenes, libras y marcos compró cada vez?

SOLUCIÓN Sean x, y y z las cantidades fijas de yenes, libras y marcos que cambia en cada viaje. Entonces, la primera vez gastó $(1/100)x$ dólares comprando yenes, $(1/0.6)y$ dólares para comprar libras y $(1/1.6)z$ para comprar marcos. Por consiguiente, $(1/100)x + (1/0.6)y + (1/1.6)z = 2\,550$. El mismo razonamiento se aplica a las otras dos compras, y se obtiene el sistema

$$\begin{aligned}\tfrac{1}{100}x + \tfrac{1}{0.6}y + \tfrac{1}{1.6}z &= 2550\\ \tfrac{1}{125}x + \tfrac{1}{0.5}y + \tfrac{1}{1.2}z &= 2840\\ \tfrac{1}{100}x + \tfrac{1}{0.6}y + \tfrac{1}{1.2}z &= 2800\end{aligned}$$

Con eliminación de Gauss se obtiene $x = 80\,000$, $y = 600$ y $z = 1\,200$. En consecuencia, cada vez compró 80 000 yenes, 600 libras y 1 200 marcos para viajar. □

■ EJEMPLO 26 (Herencia) Un padre desea distribuir sus bienes raíces, cuyo valor es \$234 000, entre sus cuatro hijas de la manera siguiente: $\frac{2}{3}$ de las propiedades deben dividirse por igual entre las hijas. Para el resto, cada hija debe recibir \$3 000 cada año hasta su vigésimo primer cumpleaños. Como entre ellas se llevan 3 años, ¿cuánto recibiría cada una de los bienes de su padre? ¿Qué edad tienen ahora esas hijas?

SOLUCIÓN Sean x, y, z y w la cantidad de dinero que recibirá cada hija debido al $\frac{1}{3}$ de las propiedades, según la edad, comenzando con la de mayor edad. Entonces, $x + y + z + w = \frac{1}{3} \cdot 234\,000 = 78\,000$. Por otro lado, $w - z = 3 \cdot 3\,000$, $z - y = 3 \cdot 3\,000$ y $y - x = 3 \cdot 3\,000$. Así llegamos al sistema

$$\begin{aligned}x + y + z + w &= 78\,000\\ w - z &= 9\,000\\ z - y &= 9\,000\\ y - x &= 9\,000\end{aligned}$$

cuya solución es $x = 6\,000$, $y = 15\,000$, $z = 24\,000$, $w = 33\,000$. La cuarta parte de dos tercios de la herencia vale $\frac{1}{4} \cdot \frac{2}{3}\,(234\,000) = \$39\,000$. Así, la hija menor recibirá $33\,000 + 39\,000 = \$72\,000$; la siguiente, $24\,000 + 39\,000 = \$63\,000$; la siguiente, $15\,000 + 39\,000 = \$54\,000$, y la primera, $6\,000 + 39\,000 = \$45\,000$. La hija mayor recibirá $6\,000 = 2 \cdot 3\,000$, así que actualmente tiene $21 - 2 = 19$ años. La segunda tiene 16, la tercera 13 y la última tiene 10 años. □

■ EJEMPLO 27 (Clima) El promedio de las temperaturas en las ciudades de Nueva York, Washington, D. C. y Boston, fue 88°F durante cierto día de verano. En Washington fue 9° mayor que el promedio de las temperaturas de las otras dos ciudades. En Boston fue 9° menor que la temperatura promedio en las otras dos ciudades. ¿Cuál fue la temperatura en cada ciudad? □

SOLUCIÓN Sean x, y y z las temperaturas en Nueva York, Washington y Boston, respectivamente. La temperatura promedio en las tres ciudades es $(x + y + z)/3$, que es 88. Por otro lado, la temperatura en Washington es 9° mayor que el promedio de Nueva York y Boston, que es $(x + z)/2$. De modo que, $y = (x + z)/2 + 9$. En consecuencia, $z = (x + y)/2 - 9$. Así, el sistema es

$$\frac{x + y + z}{3} = 88$$

$$y = \frac{x + z}{2} + 9$$

$$z = \frac{x + y}{2} - 9$$

Después de replantear este sistema en forma canónica, aplicamos la eliminación de Gauss para obtener $x = 88°$, $y = 94°$ y $z = 82°$F. □

Economía

Una de las funciones más importantes en la manufactura, que concierne a los fabricantes, economistas, especialistas de mercado, etc., es la **función demanda**. Expresa la cantidad D de piezas de cierto producto que se venden en función de su demanda. La función demanda D (o Q_d para los economistas) depende de algunas variables, como el precio P del artículo, el ingreso I de los consumidores, el precio C de un artículo de la competencia, etc., con frecuencia la función demanda, D, y sus variables, forman una ecuación lineal. Por ejemplo, $D = -15P + 0.05I + 2.5C$. Obsérvese que en este caso en particular, a medida que una unidad aumenta el precio del artículo, la demanda disminuye 15 unidades. Del mismo modo, cuando se incrementan el ingreso del consumidor o el precio de un artículo de la competencia, aumenta la demanda.

■ EJEMPLO 28 (Cálculo de una función demanda) Bikey, Inc., quiere fabricar un nuevo tipo de zapato deportivo, poco costoso, e investiga el mercado de la demanda. Encuentra que si un par de zapatos nuevos cuesta \$20 en un área de ingreso familiar promedio de \$20 000, y que si su competidor, Triceps, Inc., vende cada par de zapatos a \$20, vendería 660 pares. Por otro lado, si el precio fuera igual y Triceps bajara su precio a \$10 el par, entonces, vendería 1 130 pares en un área de \$30 000 de ingreso. Por último, si el precio de los zapatos fuera \$15 el par, y la competencia se queda en \$20 el par, se venderían 1 010 pares en un área de \$25 000 de ingreso. Determine la función demanda, suponiendo que depende linealmente de sus variables.

SOLUCIÓN Sea $D = aP + bI + cC$. Deseamos conocer a, b y c. De acuerdo con el primer caso en la investigación, $20a + 20\,000b + 20c = 660$. De igual forma, al considerar los otros dos casos se obtiene el sistema lineal

$$\begin{aligned} 20a + 20\,000b + 20c &= 660 \\ 20a + 30\,000b + 10c &= 1130 \\ 15a + 25\,000b + 20c &= 1010 \end{aligned}$$

Mediante la eliminación de Gauss se obtiene $a = -20$, $b = 0.05$ y $c = 3$. Por consiguiente, la función demanda está expresada por $D = -20P + 0.05I + 3C$. □

■ EJEMPLO 29 (Soluciones químicas) Se necesitan tres ingredientes distintos, A, B y C, para producir determinada sustancia. Pero deben disolverse primero en agua, antes de ponerlos a reaccionar para producir la sustancia. La solución que contiene A con 1.5 gramos por centímetro cúbico (g/cm^3), combinada con la solución de B cuya concentración es de 3.6 g/cm^3 y con la solución de C con 5.3 g/cm^3 forma 25.07 g de la sustancia. Si las proporciones de A, B y C en esas soluciones se cambian a 2.5, 4.3 y 2.4 g/cm^3, respectivamente (permaneciendo iguales los volúmenes), se obtienen 22.36 g de la sustancia. Por último, si las proporciones se cambian a 2.7, 5.5 y 3.2 g/cm^3, respectivamente, se producen 28.14 g de la sustancia. ¿Cuáles son los volúmenes, en centímetros cúbicos, de las soluciones que contienen A, B y C?

SOLUCIÓN Sean x, y y z los centímetros cúbicos de volumen de las soluciones que contienen A, B y C. Entonces, en el primer caso $1.5x$ es la masa de A, $3.6y$ es la masa de B y $5.3z$ es la de C. Al sumarlas deben dar 25.07. Así, $1.5x + 3.6y + 5.3z = 25.07$. El mismo razonamiento se aplica a los otros dos casos, y se llega al sistema

$$\begin{aligned} 1.5x + 3.6y + 5.3z &= 25.07 \\ 2.5x + 4.3y + 2.4z &= 22.36 \\ 2.7x + 5.5y + 3.2z &= 28.14 \end{aligned}$$

cuya solución es $x = 1.5$, $y = 3.1$ y $z = 2.2$. Por consiguiente, los volúmenes correspondientes de las soluciones que contienen A, B y C son, 1.5 cm^3, 3.1 cm^3 y 2.2 cm^3. $\square$

Otra aplicación característica de los sistemas en química es el **balanceo de reacciones químicas**. Es preciso introducir coeficientes *enteros* frente a cada uno de los reactivos, para que la cantidad de átomos de cada elemento sea igual en ambos lados de la ecuación. Por ejemplo, en la combustión del metano:

$$a\, CH_4 + b\, O_2 \to c\, CO_2 + d\, H_2O \tag{1.9}$$

calcularemos los coeficientes a, b, c y d que balanceen la ecuación. Observe que en el siguiente caso es fácil resolver por aproximación, pero no es el caso general.

■ EJEMPLO 30 (Balanceo de reacciones químicas) Balancee la reacción (1.9).

SOLUCIÓN $a = c$, porque la cantidad de átomos de carbono debe ser igual en ambos lados. De igual manera llegamos a

$$\begin{aligned} a &= c \\ 4a &= 2d \\ 2b &= 2c + d \end{aligned}$$

La solución de este sistema homogéneo es $a = \frac{1}{2}d$, $b = d$, $c = \frac{1}{2}d$. Si $d = 2$, entonces $a = 1$, $b = 2$ y $c = 1$. $\square$

Física e ingeniería

Supongamos que hay una red eléctrica como la que se muestra en la figura 1.6(a). Las corrientes y las caídas de voltaje por el circuito se apegan a la primera ley de Kirchhoff:

TEOREMA 6

> **(Ley de corriente de Kirchhoff)**
>
> La suma algebraica de todas las corrientes en cualquier nodo es cero.

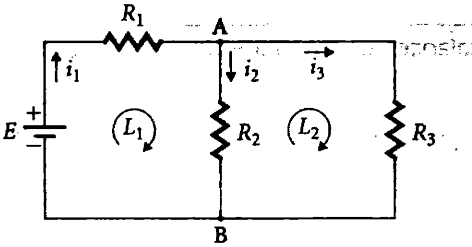
TEOREMA 7

> **(Ley de voltaje de Kirchhoff)**
>
> La suma algebraica de todos los cambios de voltaje en cualquier bucle (ciclo cerrado) es cero.

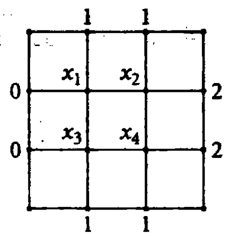
Una aplicación frecuente de esas leyes es cuando se especifica el voltaje de la fuerza electromotriz (que por lo general es una batería o un generador) y las resistencias de los resistores, y se pide calcular las corrientes.

Observe que para cada elemento de un circuito hay que elegir una dirección positiva para medir la corriente que pasa a través de dicho elemento. Las elecciones se indican con flechas. Para la fuente de voltaje se toma como positiva la dirección del signo negativo hacia el signo positivo. La fuente de voltaje agrega voltaje, y por consiguiente el cambio de voltaje es positivo, mientras que a través de los resistores es negativo, debido a que hay una caída de voltaje.

■ **EJEMPLO 31** **(Circuitos eléctricos)** Calcule las corrientes i_1, i_2 e i_3 en el circuito eléctrico de la figura 1.6(a) si el voltaje de la batería es $E = 6$ V y las resistencias son $R_1 = 2\ \Omega$, $R_2 = 2\ \Omega$ y $R_3 = 1\ \Omega$.



(a)



(b)

Figura 1.6 (a) Circuito eléctrico. (b) Transmisión de calor.

SOLUCIÓN De acuerdo con la primera ley, $i_1 - i_2 - i_3 = 0$ para el nodo A. Aplicando la segunda ley al bucle L_1 se obtiene $6 - i_1R_1 - i_2R_2 = 0$, por tanto, $2i_1 + 2i_2 = 6$. Del mismo modo, el bucle L_2 da como resultado $i_3R_3 - i_2R_2 = 0$, es decir, $-2i_2 + i_3 = 0$. Así,

$$\begin{aligned} i_1 - i_2 - i_3 &= 0 \\ 2i_1 + 2i_2 \quad &= 6 \\ -2i_2 + i_3 &= 0 \end{aligned}$$

y mediante la eliminación de Gauss puede obtenerse con facilidad $i_1 = 2.25$ A, $i_2 = 0.75$ A e $i_3 = 1.5$ A.

Otra aplicación frecuente de los sistemas lineales es en el área de los problemas de transmisión de calor, en física e ingeniería.

Supongamos que hay una placa metálica rectangular delgada, cuyas orillas se mantienen a temperaturas fijas. Por ejemplo, la orilla izquierda está a 0°C, la derecha a 2°C y las superior e inferior a 1°C, figura 1.6(b). Se desea calcular la temperatura en el interior de la placa.

Hay varias maneras de abordar los problemas de este tipo, y en algunas se requieren matemáticas de mayor nivel. El método que nos interesa será una aproximación como la siguiente: cubrir la placa con redes cada vez más finas, figura 1.6(b). Las intersecciones de las líneas del retículo se llaman *nodos*. Los nodos se dividen en puntos *en la frontera* y *en el interior*, dependiendo de si están en los bordes o en el interior de la placa. Podemos considerar que éstos son *elementos térmicos* porque cada uno influye sobre sus puntos adyacentes. Si conocemos la temperatura de los puntos en la frontera, podemos calcular la que prevalece en los puntos interiores. Es obvio que cuanto más fino sea el retículo, la aproximación a la distribución de temperaturas en la placa será mejor. Para calcular las temperaturas de los puntos en el interior se aplica el siguiente principio:

TEOREMA 8

(Propiedad promedio para la conducción de calor)

La temperatura en cualquier punto del interior es el promedio de las temperaturas de sus puntos adyacentes.

Para simplificar, supongamos que sólo se tienen cuatro puntos en el interior, cuyas temperaturas x_1, x_2, x_3 y x_4 se desconocen, y que en la frontera están los 12 puntos (sin designación) que se indican en la figura 1.6(b).

■ EJEMPLO 32 (Conducción de calor) Calcular x_1, x_2, x_3 y x_4.

SOLUCIÓN Según el teorema de la propiedad promedio,

$$\begin{aligned} x_1 &= \tfrac{1}{4}(x_2 + x_3 + 1) \\ x_2 &= \tfrac{1}{4}(x_1 + x_4 + 3) \\ x_3 &= \tfrac{1}{4}(x_1 + x_4 + 1) \\ x_4 &= \tfrac{1}{4}(x_2 + x_3 + 3) \end{aligned}$$

Podríamos emplear la eliminación de Gauss, pero para este sistema en particular resulta bastante tediosa. En su lugar, aprovecharemos la simetría de la gráfica y las ecuaciones para llegar a una solución rápida. Por simetría, sabemos que $x_1 = x_3$ y $x_2 = x_4$. Sea $r = x_1 + x_2 + x_3 + x_4 = 2x_1 + 2x_4$. Así, al sumar todas las ecuaciones, se obtiene $r = \frac{1}{4}(2r + 4)$, que implica que $r = 4$. Por consiguiente, $x_1 + x_4 = 2$. En consecuencia, para la segunda ecuación es necesario que $x_2 = \frac{5}{4}$. Entonces, la conclusión inmediata es que $x_1 = \frac{3}{4}$, $x_2 = \frac{5}{4}$, $x_3 = \frac{3}{4}$ y $x_4 = \frac{5}{4}$.

En este punto debe mencionarse que la solución los sistemas originados en los problemas de conducción de calor, como el anterior, es bastante tediosa. El uso de Maple, Mathematica o MATLAB puede ser de gran ayuda.

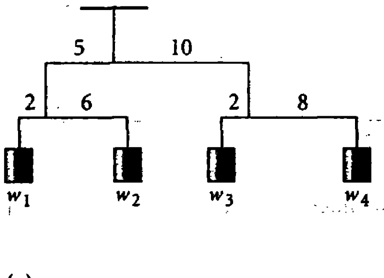
Estática y equilibrio de pesos

Ahora estudiaremos un problema característico de palancas en estática, el balanceo de pesos. Para ello, emplearemos el siguiente teorema.

TEOREMA 9 (Ley de la palanca de Arquímedes)[9]

Dos masas en una palanca se equilibran cuando sus pesos son inversamente proporcionales a sus distancias al punto de apoyo.

■ EJEMPLO 33 Calcule los pesos w_1, w_2, w_3 y w_4 para balancear las palancas de la figura 1.7(a).



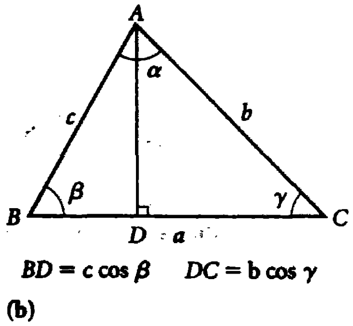


Figura 1.7 (a) Equilibrio de pesos. (b) La ley de los cosenos.

SOLUCIÓN Para balancear las dos palancas pequeñas, apegándose a la ley de Arquímedes, tenemos que $2w_1 = 6w_2$ para la palanca de la izquierda, y $2w_3 = 8w_4$ para la de la derecha. Para

[9] Aunque esta ley también se encuentra con anterioridad en los trabajos de Aristóteles, parece que Arquímedes fue el primero en basarla en la estática, y no en la cinemática. Es un caso especial del axioma de la simetría en un sistema en equilibrio, debido a Arquímedes.

equilibrar la palanca principal, se necesita que $5(w_1 + w_2) = 10(w_3 + w_4)$. De este modo llegamos al siguiente sistema homogéneo de tres ecuaciones con cuatro incógnitas:

$$\begin{aligned} 5w_1 + 5w_2 - 10w_3 - 10w_4 &= 0 \\ 2w_1 - 6w_2 \qquad\qquad\qquad &= 0 \\ 2w_3 - \ 8w_4 &= 0 \end{aligned}$$

El conjunto solución es monoparamétrico infinito, descrito por $w_1 = 7.5r$, $w_2 = 2.5r$, $w_3 = 4r$ y $w_4 = r$, $r \in \mathbf{R}$. Así, hay una cantidad infinita de pesos que pueden equilibrar este sistema, cosa que confirma nuestra experiencia, siempre y cuando los pesos, en el orden acostumbrado, sean múltiplos de los números 7.5, 2.5, 4 y 1. $\square$

Aplicaciones a la geometría

■ **EJEMPLO 34** (Ley de los cosenos) Demuestre la ley de los cosenos en geometría, es decir, que para el triángulo ABC, figura 1.7(b), se cumple

$$\cos\alpha = \frac{b^2 + c^2 - a^2}{2bc}, \qquad \cos\beta = \frac{a^2 + c^2 - b^2}{2ac}, \qquad \cos\gamma = \frac{a^2 + b^2 - c^2}{2ab}$$

SOLUCIÓN El lado a es la suma $BD + DC$. Pero $BD = c\cos\beta$ y $DC = b\cos\gamma$. De aquí que $c\cos\beta + b\cos\gamma = a$. En forma parecida obtenemos las otras dos ecuaciones del sistema:

$$\begin{aligned} c\cos\beta + b\cos\gamma &= a \\ c\cos\alpha + a\cos\gamma &= b \\ a\cos\beta + b\cos\alpha &= c \end{aligned}$$

Se trata de un sistema lineal en $\cos\alpha$, $\cos\beta$ y $\cos\gamma$. La matriz aumentada de este sistema se reduce como sigue:

$$\begin{bmatrix} 0 & c & b & : & a \\ c & 0 & a & : & b \\ b & a & 0 & : & c \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & : & (b^2 + c^2 - a^2)/2bc \\ 0 & 1 & 0 & : & (a^2 + c^2 - b^2)/2ac \\ 0 & 0 & 1 & : & (a^2 + b^2 - c^2)/2ab \end{bmatrix}$$

con lo cual se demuestra la ley. $\square$

■ **EJEMPLO 35** (Función cuadrática que pasa por tres puntos) Determine la ecuación de la parábola, con eje vertical y en el plano xy, que pasa por los puntos $P(1, 4)$, $Q(-1, 6)$ y $R(2, 9)$.

SOLUCIÓN Sea $y(x) = ax^2 + bx + c$ la ecuación de la parábola. Es necesario determinar los coeficientes a, b y c. Como el punto P pertenece a la parábola, debe cumplirse $4 = y(1) = a \cdot 1 + b \cdot 1 + c$. De manera similar, al usar los otros dos puntos se obtiene el sistema

$$\begin{aligned} a + b + c &= 4 \\ a - b + c &= 6 \\ 4a + 2b + c &= 9 \end{aligned}$$

Mediante la eliminación de Gauss se obtiene $a = 2$, $b = -1$ y $c = 3$. Por tanto, la ecuación de la parábola es $y = 2x^2 - x + 3$. □

■ EJEMPLO 36 (Plano que pasa por tres puntos) Deduzca la ecuación del plano, en el espacio xyz, que pasa por los puntos $P(1, 1, 2)$, $Q(1, 2, 0)$ y $R(2, 1, 5)$.

SOLUCIÓN Sea $ax + by + cz + d = 0$ la ecuación del plano. Es preciso determinar los coeficientes a, b, c y la constante d. Como el punto P pertenece al plano, debe cumplirse $a \cdot 1 + b \cdot 1 + c \cdot 2 + d = 0$. Siguiendo el mismo procedimiento con los otros dos puntos se obtiene el sistema homogéneo

$$\begin{aligned} a + b + 2c + d &= 0 \\ a + 2b + d &= 0 \\ 2a + b + 5c + d &= 0 \end{aligned}$$

Al resolverlo, encontraremos el conjunto infinito monoparamétrico $a = 3r$, $b = -2r$, $c = -r$ y $d = r$. Si se iguala $d = r = 1$, se obtiene la ecuación del plano $3x - 2y - z + 1 = 0$. (Cualquier otro valor de r produce un múltiplo constante de esta ecuación, que representa al mismo plano.) □

Álgebra

Los sistemas lineales se usan casi en todos los campos del álgebra, desde el estudio de polinomios y fracciones parciales hasta la demostración de identidades. A continuación revisaremos algunas de esas aplicaciones.

Recuérdese que un *polinomio* $f(x)$ en x es una expresión de la forma

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

en el que $a_n, a_{n-1}, \ldots, a_0$ son números fijos llamados *coeficientes* de $f(x)$ y x es una variable o indeterminado. Si $a_n \neq 0$, se dice que $f(x)$ es *de grado n*. Dos polinomios son *iguales* si sus coeficientes correspondientes son iguales. El polinomio cuyos coeficientes son todos cero es el *polinomio cero*. Los sistemas lineales se utilizan con frecuencia cuando se comparan los polinomios.

■ EJEMPLO 37 (Igualdad de polinomios) Calcule a, b y c tales que los polinomios $ax^2 + 3x^2 + 2ax - 2cx + 10x + 6c$ y $-2bx^2 - 3bx + 9 + a - 4b$ sean iguales.

SOLUCIÓN Los coeficientes de las potencias correspondientes de x deben ser iguales. Así,

$$(a + 3)x^2 + (2a - 2c + 10)x + 6c = -2bx^2 - 3bx + 9 + a - 4b$$

implica al sistema

$$\begin{aligned} a + 3 &= -2b \\ 2a - 2c + 10 &= -3b \\ 6c &= 9 + a - 4b \end{aligned}$$

El lector puede comprobar fácilmente que $a = 1$, $b = -2$ y $c = 3$, aplicando la eliminación de Gauss. $\square$

■ EJEMPLO 38 (Fracciones parciales) Calcule las constantes A y B tales que

$$\frac{1}{(x-1)(x-2)} = \frac{A}{x-1} + \frac{B}{x-2}$$

SOLUCIÓN Se debe cumplir

$$\frac{1}{(x-1)(x-2)} = \frac{A}{x-1} + \frac{B}{x-2} = \frac{A(x-2) + B(x-1)}{(x-1)(x-2)}$$

Por consiguiente, $1 = A(x-2) + B(x-1)$, o $(A+B)x - 2A - B = 1$, porque la primera y tercera fracciones iguales tienen los mismos denominadores. Así que es preciso resolver

$$\begin{aligned} A + B &= 0 \\ -2A - B &= 1 \end{aligned}$$

que dan como resultado $A = -1$ y $B = 1$. $\square$

■ EJEMPLO 39 (Suma de cuadrados) Deduzca una fórmula para la suma de cuadrados indicada abajo, suponiendo que la respuesta es un polinomio de grado 3 en n.

$$1^2 + 2^2 + \ldots + n^2$$

SOLUCIÓN Sea $f(x) = ax^3 + bx^2 + cx + d$ un polinomio con la siguiente propiedad: $f(n) = 1^2 + 2^2 + \cdots + n^2$. Como $1 = f(1) = a1^3 + b1^2 + c1 + d$, debe satisfacerse $a + b + c + d = 1$. Por otro lado,

$$n^2 = f(n) - f(n-1) = an^3 + bn^2 + cn + d - (a(n-1)^3 + b(n-1)^2 + c(n-1) + d)$$

Por consiguiente, desarrollando el lado izquierdo llegamos a

$$n^2 = a(3n^2 - 3n + 1) + b(2n - 1) + c \quad \text{es decir} \quad n^2 = 3an^2 + (-3a + 2b)n + a - b + c$$

Al comparar los coeficientes de las potencias de n en ambos lados, y teniendo en cuenta que $a + b + c + d = 1$, se obtiene

$$\begin{aligned} 3a &= 1 \\ -3a + 2b &= 0 \\ a - b + c &= 0 \\ a + b + c + d &= 1 \end{aligned}$$

Este sistema puede resolverse con facilidad: $a = \frac{1}{3}$, $b = \frac{1}{2}$, $c = \frac{1}{6}$, $d = 0$. Por consiguiente, $f(n) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$. Al analizar la última expresión llegamos a

$$1^2 + 2^2 + \cdots + n^2 = \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} \qquad \square$$

El problema de la pila de monedas de Fibonacci

Ahora regresemos a ese famoso problema de la pila de monedas, que resolvió Fibonacci hace algunos siglos.

■ **EJEMPLO 40** (El problema de Fibonacci) Tres hombres poseen una sola pila de dinero, y sus aportaciones son $\frac{1}{2}$, $\frac{1}{3}$ y $\frac{1}{6}$. Cada uno toma algo de dinero de la pila, hasta que no queda nada. A continuación el primer hombre regresa $\frac{1}{2}$ de lo que tomó, el segundo $\frac{1}{3}$ y el tercero $\frac{1}{6}$. Cuando el total que regresaron se divide por igual entre ellos, se descubre que cada hombre posee lo que le corresponde por su aportación. ¿Cuánto dinero había en la pila original, y cuánto tomó cada uno?

SOLUCIÓN Sean x, y y z la cantidad que tomaron los tres hombres de la pila de monedas, respectivamente, y sea w la cantidad de dinero original. Como no quedan monedas después de que los tres retiran, entonces

$$x + y + z = w$$

Los tres hombres reintegran un total de $x/2 + y/3 + z/6$, porque devolvieron respectivamente, $\frac{1}{2}$, $\frac{1}{3}$ y $\frac{1}{6}$ de lo que tomaron al principio. Esta cantidad se divide por igual, y así cada quien recibe $\frac{1}{3}(x/2 + y/3 + z/6)$.

El primer hombre tiene $x - x/2 = x/2$, que quedaron después de haber regresado $x/2$, y se le suma $\frac{1}{3}(x/2 + y/3 + z/6)$. El total debería ser lo que le corresponde, es decir, $w/2$. Así

$$\frac{x}{2} + \frac{1}{3}\left(\frac{x}{2} + \frac{y}{3} + \frac{z}{6}\right) = \frac{w}{2}$$

De igual manera, el segundo hombre tiene $y - y/3 = 2y/3$, que le quedaron después de regresar $y/3$, y además $\frac{1}{3}(x/2 + y/3 + z/6)$. El total debe ser igual a lo que le corresponde, que es $w/3$. Así,

$$\frac{2y}{3} + \frac{1}{3}\left(\frac{x}{2} + \frac{y}{3} + \frac{z}{6}\right) = \frac{w}{3}$$

Por último, el tercer hombre tiene $z - z/6 = 5z/6$, que le quedaron después de haber regresado $z/6$, y además $1/3\ (x/2 + y/3 + z/6)$. El total debe ser lo que le corresponde, que es $w/6$. Así,

$$\frac{5z}{6} + \frac{1}{3}\left(\frac{x}{2} + \frac{y}{3} + \frac{z}{6}\right) = \frac{w}{6}$$

y se tiene el sistema homogéneo

$$\begin{aligned}
x + y + z - w &= 0\\
\frac{2x}{3} + \frac{y}{9} + \frac{z}{18} - \frac{w}{2} &= 0\\
\frac{x}{6} + \frac{7y}{9} + \frac{z}{18} - \frac{w}{3} &= 0\\
\frac{x}{6} + \frac{y}{9} + \frac{16z}{18} - \frac{w}{6} &= 0
\end{aligned}$$

Vale la pena hacer notar que la suma de las tres últimas ecuaciones es igual a la primera. Esto quiere decir que si determináramos una solución simultánea de las tres últimas ecuaciones, automáticamente tendríamos una solución de la primera. Así, en esencia, se tienen tres ecuaciones con cuatro incógnitas. Por consiguiente, esperamos que haya un número infinito de soluciones.

La reducción de la matriz aumentada da como resultado

$$\begin{bmatrix} 1 & 0 & 0 & -\frac{33}{47} & : & 0\\ 0 & 1 & 0 & -\frac{13}{47} & : & 0\\ 0 & 0 & 1 & -\frac{1}{47} & : & 0\\ 0 & 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Por consiguiente, $x - \frac{33}{47}w = 0$, $y - \frac{13}{47}w = 0$ y $z - \frac{1}{47}w = 0$. Se tienen infinitas soluciones que se pueden expresar en la forma $x = \frac{33}{47}r$, $y = \frac{13}{47}r$, $z = \frac{1}{47}r$, $w = r$, $r \in \mathbf{R}$.

No se sabe si Fibonacci determinó el conjunto solución completo. Pero sí calculó la solución particular $w = 47$, $x = 33$, $y = 13$ y $z = 1$, que se obtiene al igualar $r = 47$. □

Cuadrados mágicos

Un **cuadrado mágico** de tamaño n es una matriz de $n \times n$ cuyos elementos consisten en todos los enteros entre 1 y n^2, en tal forma que las sumas de los elementos de cada columna, renglón o diagonal son iguales. La suma de los elementos de cualquier renglón, columna o diagonal de un cuadrado mágico de tamaño n es $n(n^2 + 1)/2$. (Para comprobarlo se usa la identidad $1 + 2 + \cdots + k = k(k + 1)/2$, donde $k = n^2$.)

■ **EJEMPLO 41** (Cuadrados mágicos de tamaño 2) Demuestre que no existen cuadrados mágicos de tamaño 2.

SOLUCIÓN Sea $\begin{bmatrix} a & b\\ c & d \end{bmatrix}$ un cuadrado mágico hipotético. Entonces

$$\begin{aligned}
a + b &= 5\\
c + d &= 5\\
a + c &= 5
\end{aligned}$$

$$b + d = 5$$
$$a + d = 5$$
$$b + c = 5$$

Es posible aplicar la eliminación de Gauss, o tan sólo observar que estas ecuaciones implican que $b = c$, de acuerdo con la primera y la tercera; por tanto, $2b = 5$, según la última ecuación. Esto es una contradicción, porque b debería ser un entero. De aquí se infiere que este sistema no tiene soluciones *enteras*. Por consiguiente, no hay cuadrados mágicos de tamaño 2.

■ EJEMPLO 42 (Cuadrados mágicos de tamaño 3) Determine el cuadrado mágico de tamaño 3 cuyo primer renglón es el vector (8, 1, 6).

SOLUCIÓN Ese cuadrado mágico tiene la forma

$$\begin{bmatrix} 8 & 1 & 6 \\ a & b & c \\ d & e & f \end{bmatrix}$$

para las incógnitas a, b, c, d, e y f. Según la definición, es un sistema de 7 ecuaciones con 6 incógnitas:

$$a + b + c = 15$$
$$d + e + f = 15$$
$$8 + a + d = 15$$
$$1 + b + e = 15$$
$$6 + c + f = 15$$
$$8 + b + f = 15$$
$$6 + b + d = 15$$

La eliminación de Gauss es muy tediosa, pero puede evitarse: si se despeja a de la primera y tercera ecuaciones, se obtiene $b + c = 8 + d$. Esto, combinado con la última ecuación, produce $2b + c = 17$. En consecuencia, de acuerdo con la primera ecuación, $b - a = 2$. Ahora veamos qué puede ser a. No puede ser 1, porque ese número ya se usó. Si $a = 2$, entonces $b = 4$ y $c = 9$, pero entonces la suma de la tercera columna sería mayor que 15. Si $a = 3$, $b = 5$ y $c = 7$; entonces, $d = 4$, $e = 9$ y $f = 2$, usando las tres columnas. Es fácil demostrar que estos valores forman una solución (la única) del sistema.

Por lo anterior, el cuadrado mágico es

$$\begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix}$$

Este cuadrado fue citado en un antiguo libro chino, *Nueve capítulos del arte matemático*.[10] ¿Puede el lector determinar otro cuadrado basado en éste?

[10] Vea *A History of Mathematics*, de Carl Boyer, p. 219 (Nueva York: Wiley).

Ejercicios 1.4

1. Suponga que las cantidades de bacterias de los tipos A y B son interdependientes y se apegan a la siguiente tabla experimental. ¿Hay alguna relación (ecuación) lineal entre A y B?

A	B
500	500
1 000	2 000
5 000	14 000
10 000	29 000

2. QuickInk Publisher edita tres calidades de libros: encuadernación rústica, con pasta dura y empastados en piel. Para los rústicos, la empresa gasta en promedio \$5 en papel, \$2 en ilustraciones y \$3 en las pastas. Para los de pasta dura, los gastos son \$10 en papel, \$4 en ilustraciones y \$8 en pastas; y para los de lujo empastados en piel, \$20 en papel, \$12 en ilustraciones y \$24 en pastas. Si el presupuesto permite \$235 000 en papel, \$110 000 en ilustraciones y \$205 000 en pastas, ¿cuántos libros de cada categoría pueden producirse?

3. Un negociante internacional necesita, en promedio, cantidades fijas de yenes japoneses, francos franceses y marcos alemanes para cada uno de sus viajes de negocios. Este año viajó 3 veces. La primera cambió un total de \$2 400 a la siguiente paridad: 100 yenes, 1.5 francos y 1.2 marcos por dólar. La segunda vez, cambió un total de \$2 350 con las tasas siguientes: 100 yenes, 1.2 francos y 1.5 marcos por dólar. La tercera vez cambió \$2 390 en total, a 125 yenes, 1.2 francos y 1.2 marcos por dólar, respectivamente. ¿Qué cantidades de yenes, francos y marcos compró cada vez?

4. Una madre desea distribuir sus propiedades, valuadas en \$400 000, entre sus 4 hijos como sigue: 3/4 de su herencia debe repartirse por igual entre los hijos. El resto deberá dividirlo de la forma siguiente: \$3 000 cada año a cada uno hasta que cumplan 25 años. Como los hijos sucesivos se llevan 4 años entre sí, ¿cuánto recibirá cada uno de herencia?

5. El promedio de temperaturas en las ciudades de Boston, Nueva York y Montreal fue 30°F durante cierto día de invierno. La temperatura en Nueva York fue 9° mayor que el promedio de temperaturas en las otras dos ciudades. En Montreal fue 9° menor que el promedio de temperaturas en las otras dos ciudades. ¿Cuál fue la temperatura en cada una de las ciudades?

6. Juguetes, S. A., desea fabricar un nuevo tren de juguete e investiga el mercado para determinar la demanda. Encuentra que si el tren cuesta \$40 en un área donde el ingreso familiar promedio es de \$30 000, y si al mismo tiempo Trenes Miniatura, S. A., su competidor, vende su producto equivalente a \$30, Juguetes podrá vender 1 160 trenes. Por otro lado, si el precio permanece igual y Trenes Miniatura eleva su precio a \$50 por tren, en un área de \$40 000 de ingreso familiar se venderían 1 700 trenes. Por último, si el precio del tren es de \$30 pero el de la competencia permanece en \$40, en una zona de ingreso familiar de \$35 000, se lograría una venta de 1 530 trenes. Calcule la función demanda, suponiendo que depende linealmente de sus variables.

7. Se necesitan tres ingredientes distintos, A, B y C, para producir determinada sustancia química. A, B y C deben disolverse en agua, por separado, antes de interactuar y formar la sustancia. La concentración de la solución que contiene A es 1.5 g/m^3 la que contiene B es de 1.8 g/m^3 y la de C es de 3.2 g/m^3, al combinarlas se produce 15.06 g de la sustancia. Si se modifican las concentraciones de A, B y C en esas soluciones a 2.0, 2.5 y 2.8 g/m^3, respectivamente (permaneciendo igual el volumen) se producen entonces 17.79 g de la sustancia. Por último, si las concentraciones se cambian a 1.2, 1.5 y 3.0 g/m^3, respectivamente, se producen 13.05 g de la sustancia. ¿Cuáles son los volúmenes, en centímetros cúbicos, de las soluciones que contienen A, B y C?

En los ejercicios 8 a 10 calcule las corrientes i_1, i_2 e i_3 de los respectivos circuitos, dado que en todos los casos el voltaje de la batería es $E = 6$ V.

8. De acuerdo con la figura 1.8, se sabe que $R_1 = 2\ \Omega$, $R_2 = 2\ \Omega$, $R_3 = 1\ \Omega$ y $R_4 = 2\ \Omega$.

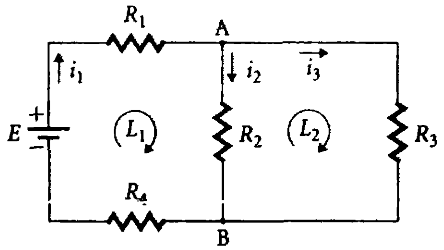


Figura 1.8

9. Con respecto a la figura 1.9, se sabe que $R_1 = 2\ \Omega$, $R_2 = 3\ \Omega$, $R_3 = 4\ \Omega$ y $R_4 = 2\ \Omega$.

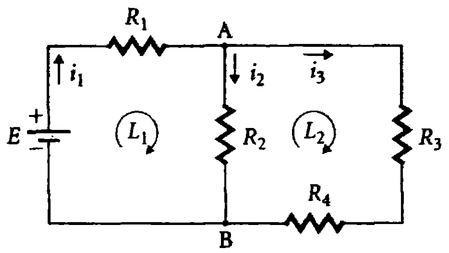


Figura 1.9

10. Con referencia a la figura 1.10, se sabe que $R_1 = 3\ \Omega$, $R_2 = 2\ \Omega$, $R_3 = 2\ \Omega$, $R_4 = 1\ \Omega$ y $R_5 = 2\ \Omega$.

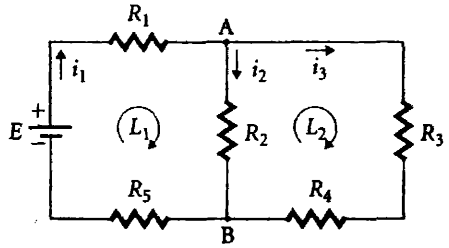


Figura 1.10

11. Calcule las temperaturas en los puntos x_1, x_2 y x_3 en la placa metálica triangular que se ilustra en la figura 1.11, si la temperatura en cada punto del interior es el promedio de las que prevalecen en sus cuatro puntos vecinos.

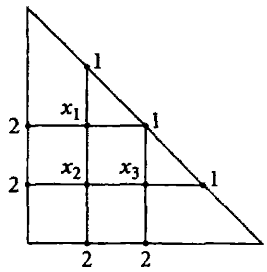


Figura 1.11

12. Determine las temperaturas en x_1, x_2, x_3 y x_4 de la placa metálica triangular que se ve en la figura 1.12, si la temperatura de cada punto del interior es igual al promedio de sus cuatro puntos vecinos.

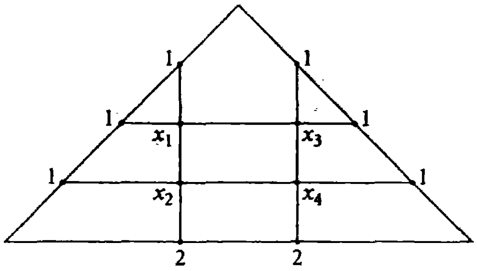


Figura 1.12

13. Balancee el sistema de pesos y palancas que se muestra en la figura 1.13.

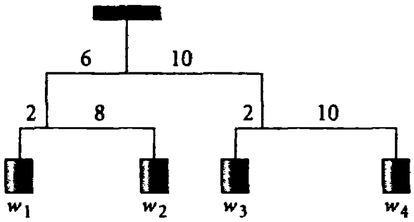


Figura 1.13

14. Deduzca la ecuación de la parábola en el plano xy que pasa por los puntos $P(1, 0)$, $Q(-1, 6)$ y $R(2, 0)$.

15. Encuentre la ecuación de la curva cúbica, en el plano xy, que pasa por los puntos $P(1, 1)$, $Q(-1, 5)$, $R(0, 1)$ y $S(-2, 7)$.

16. Deduzca la ecuación del plano, en el espacio xyz, que pasa por los puntos $P(1, 1, -1)$, $Q(2, 1, 2)$ y $R(1, 3, -5)$.

17. Calcule a, b y c, tales que las ecuaciones cuadráticas $(a - b)x^2 + (a - c)x + b + c$ y $(3 - c)x^2 - ax - b - 2c$ sean iguales.

18. Calcule a, b, c y d, tales que las ecuaciones cúbicas $(a + b + c)x^3 + (a + b)x^2 + 2bx$ y $(-d)x^3 + (2 - d)x^2 + (1 + a)x + b + c$ sean iguales.

19. Calcule las constantes A y B de modo que

$$\frac{1}{(x+1)(x-2)} = \frac{A}{x+1} + \frac{B}{x-2}$$

20. Determine los cuadrados mágicos de tamaño 3 que tengan la forma

$$\begin{bmatrix} 4 & a & b \\ c & 5 & d \\ e & f & 6 \end{bmatrix}$$

1.5 Miniproyectos

1 ■ Conjuntos de sistemas

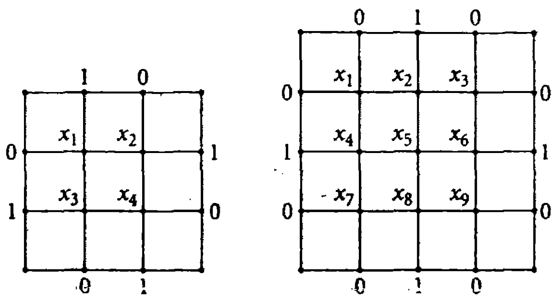
Suponga que tiene que solucionar varios sistemas con la misma matriz de coeficientes. Determine un método para resolver simultáneamente los sistemas, de tal modo que sólo se efectúe una vez la reducción de la matriz de coeficientes. Aplique su método en los cuatro sistemas lineales siguientes:

$$\begin{array}{rcccc}
x + 2y + 3z + 4w &= 1, & = 0, & = 0, & = 0 \\
2x + 2y + 3z + 4w &= 0, & = 1, & = 0, & = 0 \\
3x + 3y + 3z + 4w &= 0, & = 0, & = 1, & = 0 \\
4x + 4y + 4z + 4w &= 0, & = 0, & = 0, & = 1
\end{array}$$

2 ■ Inteligencia animal

Un conjunto de experimentos en psicología tiene por objeto estudiar la enseñanza de actividades a diversos animales, como cerdos, conejos, ratas, etc. Uno de esos experimentos implica la búsqueda de alimento. Un animal se coloca en alguna parte de una red cuadrada de corredores que pueden llevar al alimento (puntos identificados con 1), o a puntos de final ciego (puntos identificados con 0), figura 1.14. Se supone que la probabilidad de que un animal ocupe la posición x_i es el *promedio* de las probabilidades de que ocupe las posiciones vecinas directamente arriba, abajo, a la izquierda y a la derecha de x_i. Si algunos de esos puntos contiene alimento, se cuenta como éxito y su probabilidad es 100% = 1. Si una posición vecina es final ciego, se cuenta como fracaso y su probabilidad es 0% = 0. Por ejemplo, para (a) y (b) en la figura 1.14, se tiene, respectivamente:

$$\begin{aligned}
x_1 &= \tfrac{1}{4}(1 + 0 + x_3 + x_2) & \qquad x_1 &= \tfrac{1}{4}(0 + 0 + x_4 + x_2) \\
x_2 &= \tfrac{1}{4}(0 + x_1 + x_4 + 1) & \qquad x_2 &= \tfrac{1}{4}(1 + x_1 + x_5 + x_3) \\
x_3 &= \tfrac{1}{4}(x_1 + 1 + 0 + x_4) & \qquad x_3 &= \tfrac{1}{4}(0 + x_2 + x_6 + 0)
\end{aligned}$$



(a) (b)

Figura 1.14 Experimentos en inteligencia animal.

Calcule:

(a) Las probabilidades x_1, x_2, x_3 y x_4, para la figura 1.14(a).
(b) Las probabilidades $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ y x_9, para la figura 1.14(b).

Si se aprovecha cualquier *simetría* de los datos podrían evitarse cálculos tediosos.

3 ■ Aproximaciones a curvas; segmentos lineales

Una de las formas más elementales y básicas de aproximar la gráfica de una curva es conectando los puntos conocidos consecutivos de la curva con segmentos de recta. La poligonal que resulta es un *segmento lineal* de la curva. Con frecuencia y en la práctica, se da un conjunto de valores de una función incógnita (con frecuencia, algunos datos experimentales) y se ha de identificar la función. Una primera aproximación gráfica sería trazar el segmento lineal que une a esos puntos. Suponga, por ejemplo, que los datos son los dos conjuntos siguientes de puntos, que pertenecen a dos curvas distintas.

Conjunto 1:

x	−1	−0.5	0.	0.5	1.	1.5	2.	2.5	3
$f(x)$	−5	−0.875	1.	1.375	1.	0.625	1.	2.875	7

Conjunto 2:

x	−4.5	−3.5	−2.5	−1.5	−0.5
$f(x)$	3519.14	301.641	−32.4844	14.7656	−24.6094
x	0.5	1.5	2.5	3.5	4.5
$f(x)$	−24.6094	14.7656	−32.4844	301.641	3519.14

(a) Trace las gráficas de los dos segmentos lineales que unen los puntos consecutivos.
(b) Si la primera curva es una cúbica, es decir, que tiene la forma $ax^3 + bx^2 + cx + d$, y la segunda es un polinomio de sexto grado con la forma $ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx + g$, use las tablas de arriba para calcular los coeficientes de cada curva polinomial. Redondee las respuestas al entero más cercano.
(c) Trace las gráficas de las dos curvas.

4 ■ El problema del ganado de Arquímedes

Es un problema famoso que mandó Arquímedes de Siracusa,[11] antiguo matemático griego, a Eratóstenes en Alejandría. Su forma original era un conjunto de epigramas en griego antiguo. Por lo general se cree que Arquímedes resolvió este problema, pero se desconoce si es su autor. Describiremos la parte del problema que tiene importancia para nosotros en este momento. Para conocer una buena traducción al inglés, véase el libro *The Works of Archimedes*, de Sir Thomas L. Heath (Dover, 1953, pp. 319-326). Los epigramas comienzan como sigue:

[11] **Arquímedes** (287-212 a.C.) es considerado el matemático y físico más grande de la antigüedad, así como uno de los matemáticos más importantes en la historia de la humanidad. Creció en Siracusa, una población griega en Sicilia. Pheidias su padre, fue astrónomo. Después de estudiar matemáticas en Alejandría, Egipto, Arquímedes regresó a Siracusa, donde permaneció el resto de su vida. Fue asesinado por un soldado, cuando la ciudad cayó en poder de los romanos.

Calcula, ¡oh extranjero!, la cantidad de reses del Sol que alguna vez pastaban en los campos de la isla de Trinacia (Sicilia) que estaban divididos, según el color, en cuatro rebaños: uno blanco, uno negro, uno amarillo y uno manchado . . .

A continuación el manuscrito describe las relaciones entre las vacas y los toros de los cuatro rebaños. Sean W, B, D y Y las cantidades de toros en los rebaños blanco, negro, manchado y amarillo, respectivamente. En igual forma, sean w, b, d y y las cantidades de vacas en el mismo orden. Entonces $W + w$, $B + b$, $D + d$ y $Y + y$ son las cantidades de reses en los rebaños blanco, negro, manchado y amarillo, en forma correspondiente. El manuscrito nos comunica las siguientes relaciones (1) para los toros:

$$W = \left(\frac{1}{2} + \frac{1}{3}\right) B + Y$$

$$B = \left(\frac{1}{4} + \frac{1}{5}\right) D + Y$$

$$D = \left(\frac{1}{6} + \frac{1}{7}\right) W + Y$$

y (2) para las vacas:

$$w = \left(\frac{1}{3} + \frac{1}{4}\right) (B + b)$$

$$b = \left(\frac{1}{4} + \frac{1}{5}\right) (D + d)$$

$$d = \left(\frac{1}{5} + \frac{1}{6}\right) (Y + y)$$

$$y = \left(\frac{1}{6} + \frac{1}{7}\right) (W + w)$$

Resuelva este sistema de 7 ecuaciones con 8 incógnitas. Como el sistema es homogéneo, con más incógnitas que ecuaciones, es de esperarse llegar a un número infinito de soluciones. Demuestre que la solución entera mínima es $B = 7\,460\,514$. Calcule el resto de las incógnitas y la cantidad total de reses en este caso.

1.6 Ejercicios en computadora

Los primeros ejercicios en computadora tienen por objeto ayudar al lector a aprender los comandos básicos de los programas, en relación con el capítulo 1, y también revisar algo del material. Un ejercicio identificado con [S] requiere manipulación simbólica.[12] En todos los casos se supone que el lector use algún programa de matemáticas.

[12] Omítase si sólo se dispone de cálculo numérico.

1. Resuelva numéricamente el sistema. Llegue a los resultados con precisión por omisión y también con mayor exactitud. Si su programa admite aritmética racional, calcule la respuesta exacta. Por último, compruebe su respuesta.

$$\begin{aligned} \frac{1}{5}x + \frac{1}{6}y + \frac{1}{7}z &= \frac{241}{1260} \\ \frac{1}{6}x + \frac{1}{7}y + \frac{1}{8}z &= \frac{109}{672} \\ \frac{1}{7}x + \frac{1}{8}y + \frac{1}{9}z &= \frac{71}{504} \end{aligned} \tag{1.10}$$

2. [S] Resuelva este sistema para x y y.

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

3. Introduzca la matriz aumentada del sistema (1.10) y determine: (a) un escalonamiento de renglones (si hay disponible), (b) la forma de escalón reducida. ¿Cuál es la solución del sistema?
4. Sea A la matriz de coeficientes del sistema (1.10). ¿A es equivalente en renglones a B?

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix}$$

5. Para el siguiente sistema, use su programa para demostrar que si $c = -\frac{250}{3}$, tiene un número infinito de soluciones. Pero $c \neq -\frac{250}{3}$, no tiene soluciones.

$$\begin{aligned} \frac{1}{5}x - \frac{1}{6}y &= 100 \\ -\frac{1}{6}x - \frac{5}{36}y &= c \end{aligned}$$

6. Para la matriz B en el ejercicio 4 y con su programa informático, presente la primera columna de B, el segundo renglón, las dos primeras columnas los dos últimos renglones y la parte $\begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$
7. Si su programa admite números aleatorios, genere y resuelva un sistema aleatorio de tres ecuaciones con tres incógnitas. Si repite lo anterior varias veces, casi siempre ¿se obtienen sistemas consistentes o inconsistentes?
8. Use su programa para trazar las rectas definidas por un sistema de dos ecuaciones con dos incógnitas en la misma figura.
9. Use su programa para trazar los planos definidos por un sistema de tres ecuaciones con tres incógnitas en la misma figura.
10. Calcule las temperaturas de los nueve puntos interiores en una placa cuadrada que se ha subdividido con tres líneas verticales paralelas repartidas y tres horizontales, también repartidas. Suponga que los dos lados verticales del cuadrado se mantienen a 85°, y los dos verticales a 110°. Aplique la propiedad del valor promedio en la conducción de calor.

Algunas soluciones con Maple

función	Maple	función	Maple
Fin de comando	*; (punto y coma)*	*Continuar en el siguiente renglón*	*\ (diagonal inversa)*
petición de resultado	*: (dos puntos)*	*Resolver*	solve, linsolve
ayuda	?tema o bien ayuda (tema)	*Resolver numéricamente*	fsolve
comentarios	*# comentario hasta el final del renglón*	*Paquete de álgebra lineal*	with(linalg);
último resultado	*" (dobles comillas)*	*Forma de escalón*	gausselim, ffgausselim
asignación	*:= (dos puntos igual)*	*Forma de escalón reducida*	rref, gaussjordan
listas de argumentos	*( ) (paréntesis)*	*Sustitución hacia atrás*	backsub

```
# Ejercicio 1 —Parcial
sys: {1/5*x+1/6*y+1/7*z = 241/1260, 1/6*x+1/7*y+1/8*z = 109/672,
     1/7*x+1/8*y+1/9*z = 71/504};          # Asignación de un nombre al sistema.
solve(sys, {x,y,z});                        # Aritmética de números racionales.
evalf(");                                   # Evaluación del último resultado
                                            # con precisión por omisión.
evalf(" ",15):                              # Evaluación con mayor precisión.
fsolve(sys, {x,y,z});                       # Alternativa en una etapa.
solve({1./5*x+1/6*y+1/7*z=241/1260,         # También, forzamiento de aritmética de
1/6*x+1/7*y+1/8*z = 109/672,                # punto flotante con 1./5 .
1/7*x+1/8*y+1/9*z = 71/504},{x,y,z});
# Otra forma importante de resolver un sistema lineal es usando linsolve. Primero
con(linalg);                                # cargue el paquete de álgebra lineal, linalg. A con-
A: = matrix([[1/5, 1/6, 1/7] , [1/6,1/7,1/8] , [1/7, 1/8, 1/9]]);  # tinuación la matriz de
b: = vector([241/1260, 109/672,71/504]);                           # coeficientes y el vector de
linsolve(A,b);                                                     # constantes en linsolve.

# Ejercicio 2
solve({a1*x + b1*y = c1, a2x + b2*y = c2},{x,y});

# Ejercicio 3
m: =matrix([[1/5,1/6,1/7, 241/1260],        # Introducción de la matriz aumentada.
[1/6,1/7,1/8,109/672],[1/7,1/8,1/9,71/504]]);
gausselim(m);                               # Eliminación de Gauss; una forma de escalón.
rref(m);                                    # La forma de escalón reducida.
gaussjord(m);                               # Lo mismo. La última columna es la solución.

# Ejercicio 4
A: = matrix (3, 3, [1/5, 1/6, 1/7,          # Introducción de A en un formato
1/6, 1/7, 1/8, 1/7, 1/8, 1/9]);             # diferente.
B: = matrix (3, 3, [[1, 2, 3], [2, 2, 3], [3, 3, 3]]);  # Otro formato.
rref(A);                                    # Las 2 formas de escalón reducidas son
rref(B);                                    # iguales. Luego A y B son equivalentes.
```

```
# Ejercicio 6
col (B, 1); row(B, 2); col(B, 1..2); row(B, 2..3); submatrix(B, 1..2,1..2);
# Ejercicio 7 - Sugerencia
rin := rand(1..1000);               # genera un entero aleatorio entre 1 y 1 000.
a1 := evalf(rin()/1000);            # La división entre 1 000 genera un número real en [0,1].
# También relacionados: randmatrix y randvector.
# Ejercicio 8 - Sugerencia
plot(2*x-1, x=0..3);                # Gráfica 2x-1 variando x de 0 a 3.
plot({2*x-1, x+2}, x=0..3);         # Gráfica 2x-1 y x+2 en la misma figura.
# Ejercicio 9 - Sugerencia
plot3d(x-y, x=0..3, y=0..2);        # Gráfica x-y en [0, 3]x[0, 2], en tres dimensiones
plot3d({x-y, x+y}, x=0..3, y=0..2); # Gráficas en tres dimensiones en la misma figura.
```

Algunas soluciones con Mathematica

función	**Mathematica**	**función**	**Mathematica**
Fin de comando	*(cambio de renglón)*	*listas de argumentos*	[ ] *(paréntesis cuadrados)*
supresión de salida	; *(punto y coma)*	*pasar al siguiente renglón*	\ *(diagonal inversa)*
ayuda	?tema *o bien* ??tema	*multiplicación*	* *o bien (espacio)*
comentarios	(* *comentario* *)	*Resolver*	Solve, LinearSolve
último resultado	% *(signo de porcentaje)*	*Resolver numéricamente*	NSolve
asignación	= *(igual)*	*Forma de escalón reducida*	RowReduce
Un paquete de álgebra lineal:		<<LinearAlgebra 'MatrixManipulation'	

```
(* Ejercicio 1 - parcial *)
sys={1/5 x+1/6 y+1/7 z == 241/1260, 1/6 x+1/7 y+1/8 z == 109/672,
     1/7 x+1/8 y+1/9 z == 71/504}              (* Asignación de un nombre al sistema. *)
Solve[sys, {x, y, z}]                          (* Aritmética con números racionales.  *)
N [%]   (* Evaluación del último resultado con la precisión por omisión. *)
N [%%, 15]                                     (* Evaluación con mayor precisión.   *)
NSolve[sys, {x, y, z}]                         (* Alternativa en un paso.   *)
Solve[1./5 x+1/6 y+1/7 z == 241/1260,          (* También, forzamiento de aritmética  *)
1/6 x+1/7 y+1/8 z == 109/672,                  (* de punto flotante con 1./5 .      *)
1/7 x+1/8 y+1/9 z == 71/504}, {x, y, z}]
LinearSolve[{{1/5, 1/6, 1/7}, {1/6, 1/7, 1/8}, (* ¡También! Use LinearSolve con la matriz *)
{1/7, 1/8, 1/9}}, {241/1260, 109/672, 71/504}] (* de coeficientes y el vector de constantes. *)
(* Ejercicio 2 *)
Solve[{a1 x b1 y == c1, a2 x + b2 y == c2}, {x, y}]
Simplify [%]                                   (* Se necesita simplificar la solución.   *)
(* Ejercicio 3 *)
m={{1/5, 1/6, 1/7, 241/1260}, {1/6, 1/7, 1/8, 109/672}, {1/7, 1/8, 1/9, 71/504}}
RowReduce[m]      (* La forma de escalón reducida. La solución es la última columna.*)
```

```
(* Ejercicio 4 *)

A={{1/5, 1/6, 1/7}, {1/6, 1/7, 1/8}, {1/7, 1/8, 1/9}}
B={{1, 2, 3}, {2, 2, 3}, {3, 3, 3}}
RowReduce [A]                         (* Las 2 formas de escalón reducidas son *)
RowReduce [B]                         (* equivalentes, luego A y B son equivalentes.*)

(* Ejercicio 6 *)

<<LinearAlgebra 'MatrixManipulation'  (* Cargar un paquete de álgebra lineal. *)
TakeColumns [B, {1}]
TakeRows [B, {2}]
TakeColumns [B, {1, 2}]
TakeRows [B, {2, 3}]
TakeMatrix [B, {1, 1}, {2, 2}]

(* Ejercicio 7 - Sugerencia *)

Random[]                              (* Un número real aleatorio en [0,1]. *)

(* Ejercicio 8 - Sugerencia *)

Plot[2*x-1, {x, 0, 3}]                (* Gráfica 2x-1 variando x de 0 a 3. *)
Plot[{2*x-1, x+2}, {x, 0, 3}]         (* Gráfica 2x-1 y x+2 en la misma figura. *)

(* Ejercicio 9 - Sugerencia *)

p1=Plot3D[x-y, {x, 0, 3}, {y, 0, 2}]  (* Gráfica tridimensional de x-y en [0, 3]x[0, 2]. *)
p2=Plot3D[x+y, {x, 0, 3}, {y, 0, 2}]  (* Otra gráfica. *)
Show[{p1, p2}]                        (* Mostrarlas juntas. *)
```

Algunas soluciones con MATLAB

función	**MATLAB**	**función**	**MATLAB**
Fin de comando	*(cambio de renglón)*	*listas de argumento*	( ) *(paréntesis)*
supresión de salida	; *(punto y coma)*	*pasar al siguiente renglón*	. . . *(elipsis)*
ayuda	help tópico	*Resolver ecuación*	roots
comentarios	*% comentario hasta el final del renglón*	*Resolver el sistema Ax = b*	A\b
último resultado	ans	*Forma de escalón reducida*	rref
asignación	= *(igual)*	*Reducción en etapas*	rrefmovie

Nota: La notación (ST) significa que el Toolbox simbólico debe estar disponible.

```
% Ejercicio 1 - Parcial

A = [1/5 1/6 1/7; 1/6 1/7 1/8; 1/7 1/8 1/9]   % Para resolver un sistema cuadrado
b = [241/1260; 109/6t2; 71/504]   % se forma la matriz de coeficientes A, después
A/b                               % el vector constante b y se teclea A\b.
format long                       % Para mostrar mayor exactitud cambiar a
ans                               % formato largo y llamar al último resultado.
format short                      % Regreso al formato corto.
linsolve(A, b)                    % También se puede usar linsolve (ST).
```

```
% Ejercicio 2

solve('a1*x + b1*y = c1', 'a2*x + b2*y = c2', 'x,y')     % ST

% Ejercicio 3

m=[1/5 1/6 1/7 241/1260; 1/6 1/7 1/8 109/672; 1/7 1/8 1/9 71/504]
rref(m)         % La forma de escalón reducida. La última columna es la solución.

% Ejercicio 4

B=[1 2 3; 2 2 3; 3 3 3]         % La matriz A fue introducida en el ejercicio 1.
rref(A)                         % Las dos formas de escalón reducidas son
rref(B)                         % iguales, luego A y B son equivalentes.

% Ejercicio 6

B(:, 1)                         % Columna 1.
B(2, :)                         % Renglón 2.
B(:, 1:2)                       % Columnas 1 y 2.
B(2:3, :)                       % Renglones 2 y 3.
B(1:2, 1:2)                     % Bloque superior izquierdo.

% Ejercicio 7 – Sugerencia

rand                            % Un número real en [0 1].
% También relacionado: randn .

% Ejercicio 8 – Sugerencia

fplot('[2*x-1, x+2]', [0 3]   % Gráfica 2x-1 y x+2 en [0, 3] en una gráfica. También, un
x = 0:.1:3;                   % método más importante consiste en definir un vector x,
y1 = 2*x-1; y2 = x+2;         % a continuación aplique las funciones para obtener los
plot (x, y1, x, y2)           % vectores y y grafique.
                              % Vea también ezplot, de (ST).

% Ejercicio 9 – Sugerencia

x = 0:1/4:3;      % Para graficar x-y y x+y en [0, 3]x[0, 2] en la misma figura:
y = 0:1/6:2;      % Crear vectores para las coordenadas x y y de los puntos.
[X, Y]=meshgrid(x, y); % Forma un agrupamiento de x y y adecuado para gráficas en 3D.
Z=[X-Y, X+Y];          % Defina Z en función de las dos funciones de
mesh(Z);               % X y Y y use mesh para graficar.

% ADVERTENCIA: x y y deben tener las mismas dimensiones. Observe que x tiene
% (3-1)*4+1 = 13 componentes, y las componentes de y son (2-1)*6+1 = 13.
% Relacionado: ¡Explore la construcción linspace!
```

2

Vectores

Si la supervivencia es una medida de la calidad, los "Elementos" de Euclides y las Cónicas" de Apolonio fueron las mejores obras en sus campos.
Carl C. Boyer

Introducción

Apolonio y Descartes

Las cantidades como longitud, área, volumen, temperatura, masa y potencial se pueden determinar sólo con su magnitud. Sin embargo, pensemos en el desplazamiento, la velocidad y la fuerza. Con ellas necesitamos la magnitud y también la dirección para definirlas por completo. El desplazamiento, la velocidad y la fuerza son ejemplos de *vectores libres*. Nos interesan principalmente los vectores libres que comienzan en el *origen*. A ellos los llamaremos simplemente *vectores*.

En este capítulo estudiaremos los vectores, su aritmética y su geometría, porque desempeñan un papel importante en matemáticas, física, ingeniería, procesamiento de imágenes, gráficas computarizadas y en muchos otros campos de la ciencia y de la vida cotidiana.

En el plano y en el espacio los vectores tienen existencia doble: son a la vez objetos algebraicos y geométricos. Este tipo de dualidad nos permite estudiar la geometría con métodos algebraicos. Dos hombres, principalmente, tienen el crédito de haber realizado esto: Apolonio y Descartes.

Apolonio[1] posee un lugar prominente, junto con Euclides y Arquímedes, como uno de los tres grandes geómetras de la antigüedad. Demostró muchas propiedades geométricas de las secciones cónicas, usando líneas de referencia para medir distancias entre puntos de curvas. Éste es el primer uso de coordenadas sistemático conocido. En 1637 Descartes usó uno de los teoremas de Apolonio para ensayar su nueva geometría analítica.

[1] **Apolonio de Perga** nació en la ciudad griega de Perga, hoy el pueblo de Murtana, en Asia Menor. Pasó algún tiempo en Alejandría y Pérgamo. Se cree que vivió entre 262 y 200 a. C. La mayor parte de sus trabajos, como *Entrega rápida, Corte de una relación, Tangencias* e *Inclinaciones* se han perdido. Sobrevivieron 7 volúmenes de su grandioso tratado *Cónicas*. Sus teoremas de este libro se han usado en astronomía durante dos milenios.

Descartes[2] es el verdadero padre de la geometría analítica, y el primero en estudiar en forma sistemática la geometría con métodos totalmente algebraicos. A diferencia de Apolonio, Descartes introdujo sistemas de coordenadas independientes de las curvas. A él se debe el concepto de que una curva está *definida* por una ecuación, y no que la curva determina la ecuación. Su trabajo más famoso es un tratado de filosofía llamado *Discurso del método*. Uno de los apéndices de este libro se llama *Geometría*, y presenta lo que hoy se conoce como geometría analítica.

2.1 Operaciones vectoriales

Objetivos del alumno para esta sección

1. Ejercitarse en la aritmética vectorial básica y comprender su geometría.
2. Conocer qué es una combinación lineal de vectores.
3. Comprender la relación entre sistemas lineales y ecuaciones vectoriales.

Dedicaremos esta sección a los vectores y a su aritmética. También presentaremos el importante concepto de una *combinación lineal* de una sucesión de vectores.

Suma y multiplicación por escalar

Un **vector** es una matriz de una columna. Un **vector-n** o **n-vector** es una matriz de $n \times 1$. Por ejemplo,

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} 0.5 \\ 1 \\ 0 \\ -0.2 \end{bmatrix}$$

son vectores 2, 3 y 4, respectivamente. Algunas veces, al valor n suele llamársele **tamaño** del vector. Los elementos de un vector también se llaman **componentes**. Los componentes de $\mathbf{w}$ son 0.5, 1, 0 y −0.2. El conjunto de todos los vectores n se representa con $\mathbf{R}^n$.

$$\mathbf{R}^n = \{\mathbf{x}, \mathbf{x} \text{ es un vector-}n\}$$

$\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ son los elementos correspondientes de $\mathbf{R}^2$, $\mathbf{R}^3$ y $\mathbf{R}^4$.

Los vectores 2 y 3 pueden interpretarse geométricamente como puntos en el plano o en el espacio. Cualquier vector 2 por ejemplo, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, puede representarse gráficamente con el *punto* cuyas coordenadas son (x_1, x_2) en un plano de coordenadas cartesianas. Con frecuencia, $\mathbf{x}$ se considera como la *flecha* que comienza en el origen (0, 0) cuya punta tiene las coordenadas (x_1, x_2). La figura. 2.1(a) y 2.1(b) muestra los vectores $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, y

[2] **René Descartes** fue filósofo, militar y matemático. Nació en 1596 en Touraine, Francia, en una familia acomodada; fue educado en un colegio de jesuitas y estudió leyes. Viajó y participó en varias campañas militares. Entre sus trabajos están *Le Monde* y *Discours de la Méthode*. Murió en Estocolmo, en 1650, cuando asesoraba a la reina Cristina de Suecia.

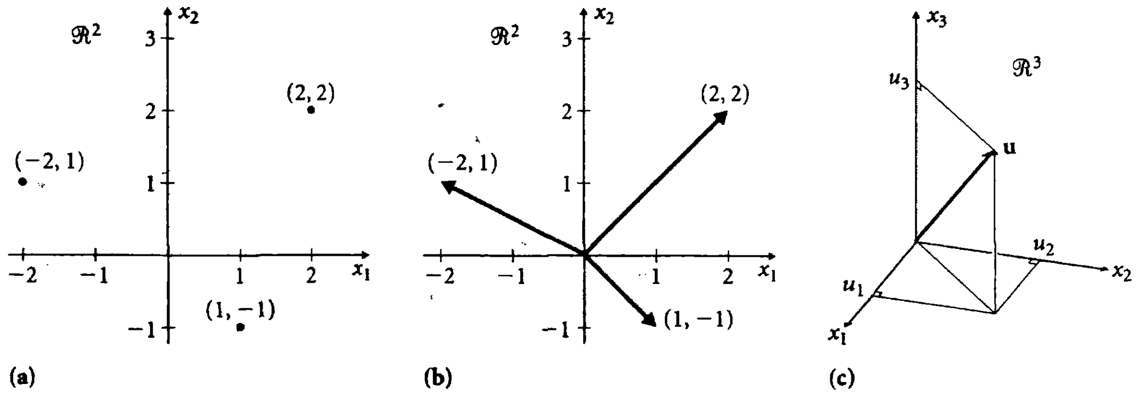


Figura 2.1 Representación geométrica de vectores 2 y vectores 3.

$\mathbf{w} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ en forma de puntos y de flechas que comienzan en el origen. Como es posible representar todos los vectores 2 de esta manera, $\mathbf{R}^2$ es el plano total. Los vectores 3 pueden graficarse en forma parecida, figura. 2.1(c), y $\mathbf{R}^3$ constituye entonces el espacio tridimensional total.

Se dice que dos vectores **u** y **v** del mismo tamaño son **iguales**, y se expresa $\mathbf{u} = \mathbf{v}$ si sus componentes respectivas son iguales. Pero los vectores de tamaños distintos nunca lo son.

■ EJEMPLO 1 La ecuación

$$\begin{bmatrix} 1 \\ a+b \end{bmatrix} = \begin{bmatrix} a \\ -1 \end{bmatrix}$$

sólo es válida si $a = 1$ y $b = -2$. □

Los vectores del mismo tamaño pueden sumarse componente por componente:

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ -2 \\ -7 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ -4 \end{bmatrix}$$

A esta operación se le llama **suma vectorial**.

La suma $\mathbf{u} + \mathbf{v}$ de dos vectores 2 o de dos vectores 3, **u** y **v**, se representa en forma geométrica como la flecha diagonal del paralelogramo cuyos lados son **u** y **v**, figura. 2.2(a). A esta regla se le denomina **ley del paralelogramo para la suma**.

Un vector n puede multiplicarse por un escalar, componente por componente:

$$7\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 14 \\ -7 \end{bmatrix}, \qquad -2\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 6 \end{bmatrix}$$

A esta operación se le llama **multiplicación por escalar**. El vector $(-1)\mathbf{v}$ se llama **opuesto** de **v** y se representa con $-\mathbf{v}$.

$$(-1)\mathbf{v} = -\mathbf{v}$$

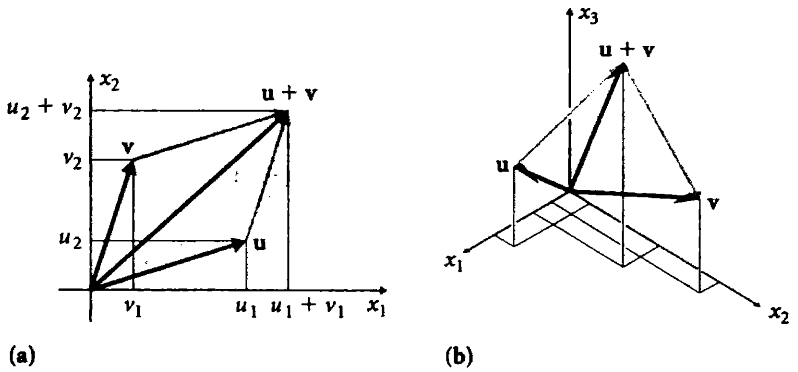


Figura 2.2 La ley del paralelogramo de la suma.

Se acostumbra escribir **u** – **v** para representar **u** + (–1)**v**, y al resultado de esta operación se le denomina **diferencia** entre **u** y v.

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-1)\mathbf{v}$$

Si todos los elementos de un vector son cero, se dice que es un **vector cero** y se representa por **0**.

$$\mathbf{0} = [0] \qquad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Geométricamente, el producto por escalar, c**u**, es la flecha **u** escalada por un factor de c. Si $c > 0$, entonces c**u** tiene la misma dirección que **u**. Si $c < 0$, c**u** tiene dirección contraria. Si $|c| > 1$, entonces c**u** se extiende en un factor de c. Si $|c| < 1$, entonces c**u** es una contracción de **u**, figura 2.3(a) y 2.3(b).

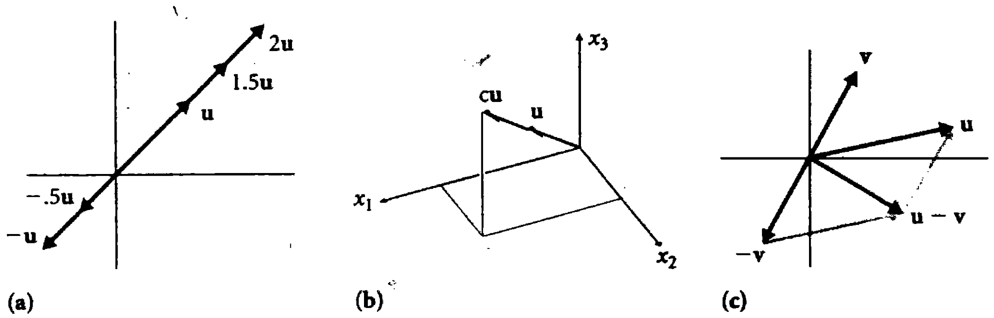


Figura 2.3 Productos por escalar: (a) y (b), (c) diferencia.

Observe que la diferencia **u** - v puede representarse como la suma **u** + (–1)**v**, figura 2.3(c).

Al vector n que tiene 1 como i-ésimo componente y todos los demás componentes 0 se denota mediante $\mathbf{e}_i$. Los vectores $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ se llaman **vectores de base estándar** de $\mathbf{R}^n$, o simplemente **la base de** $\mathbf{R}^n$. Por ejemplo, los vectores de base estándar de $\mathbf{R}^2$ son $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y

$\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, mientras que los de $\mathbf{R}^3$ son $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(Véase la figura 2.4)

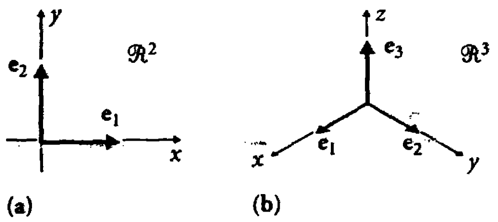


Figura 2.4 Los vectores de base normal.

La suma vectorial y la multiplicación por escalar tienen algunas propiedades fundamentales, que se resumen en el siguiente teorema.

TEOREMA 1

(Reglas para la suma vectorial y la multiplicación por escalar)

Sean **u**, **v** y **w** cualesquiera vectores n y sean a y b dos escalares cualquiera. Entonces se cumplen las siguientes igualdades entre los vectores n:

1. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ **Ley asociativa**
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ **Ley conmutativa**
3. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
4. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$
5. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ **Ley distributiva**
6. $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ **Ley distributiva**
7. $(ab)\mathbf{u} = a(b\mathbf{u}) = b(a\mathbf{u})$
8. $1\mathbf{u} = \mathbf{u}$
9. $0\mathbf{u} = \mathbf{0}$

DEMOSTRACIÓN Sólo comprobaremos la parte 1 y dejaremos las demás demostraciones como ejercicios. Como $\mathbf{u} + \mathbf{v}$ y $\mathbf{w}$ son vectores n, también lo es su suma $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$. Igualmente, $\mathbf{u} + (\mathbf{v} + \mathbf{w})$ es un vector n. Por consiguiente, $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$ y $\mathbf{u} + (\mathbf{v} + \mathbf{w})$ tienen el mismo tamaño. Sean u_i, v_i y w_i las i-ésimas componentes de **u**, **v** y **w**, respectivamente. Entonces, $u_i + v_i$ es la i-ésima componente de $\mathbf{u} + \mathbf{v}$, así que $(u_i + v_i) + w_i$ es la i-ésima componente de $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$. Del mismo modo, $u_i + (v_i + w_i)$ es la i-ésima componente de $\mathbf{u} + (\mathbf{v} + \mathbf{w})$. Como $(u_i + v_i) + w_i = u_i + (v_i + w_i)$, las componentes respectivas de $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$ y de $\mathbf{u} + (\mathbf{v} + \mathbf{w})$ son iguales. Esta afirmación es cierta para toda $i = 1, \ldots, n$. Se concluye entonces que $(\mathbf{u} + \mathbf{v}) + \mathbf{w}$ y $\mathbf{u} + (\mathbf{v} + \mathbf{w})$ son iguales. $\square$

El teorema 1 puede aplicarse para resolver ecuaciones vectoriales sencillas.

■ EJEMPLO 2 Determine el vector $\mathbf{x}$ tal que $2\mathbf{x} - 4\mathbf{v} = 3\mathbf{u}$.

SOLUCIÓN Se suma $4\mathbf{v}$ en ambos lados de la ecuación para obtener

$$\begin{aligned} &(2\mathbf{x} - 4\mathbf{v}) + 4\mathbf{v} &&= 3\mathbf{u} + 4\mathbf{v} && \\ \Leftrightarrow\ &2\mathbf{x} + (-4\mathbf{v} + 4\mathbf{v}) &&= 3\mathbf{u} + 4\mathbf{v} && \text{de acuerdo con 1 del teorema 1} \\ \Leftrightarrow\ &2\mathbf{x} + \mathbf{0} &&= 3\mathbf{u} + 4\mathbf{v} && \text{de acuerdo con 4} \\ \Leftrightarrow\ &2\mathbf{x} &&= 3\mathbf{u} + 4\mathbf{v} && \text{de acuerdo con 3} \end{aligned}$$

Se multiplican por $\frac{1}{2}$ ambos lados de la última ecuación:

$$\begin{aligned} \Leftrightarrow\ &\tfrac{1}{2}(2\mathbf{x}) &&= \tfrac{1}{2}(3\mathbf{u} + 4\mathbf{v}) && \\ \Leftrightarrow\ &\left(\tfrac{1}{2}\cdot 2\right)\mathbf{x} &&= \tfrac{1}{2}(3\mathbf{u}) + \tfrac{1}{2}(4\mathbf{v}) && \text{de acuerdo con 7, 5} \\ \Leftrightarrow\ &1\mathbf{x} &&= \left(\tfrac{1}{2}\cdot 3\right)\mathbf{u} + \left(\tfrac{1}{2}\cdot 4\right)\mathbf{v} && \text{de acuerdo con 7} \\ \Leftrightarrow\ &\mathbf{x} &&= \tfrac{3}{2}\mathbf{u} + 2\mathbf{v} && \text{de acuerdo con 8} \end{aligned}$$

Matrices como sucesiones de vectores

Con frecuencia se considera que las matrices son sucesiones de vectores. Por ejemplo, la matriz

$$\begin{bmatrix} 1 & 3 & 1 \\ 2 & 4 & 2 \end{bmatrix}$$

puede considerarse como igual a $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$, siendo $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Dijimos *sucesión* y no *conjunto* por dos motivos: a diferencia de los conjuntos, (1) los elementos de una sucesión tienen un orden definido, y (2) se permite que el mismo elemento se repita en posiciones distintas. Es claro que una matriz puede tener columnas repetidas, y el orden de ellas es importante. Si se toma en cuenta lo anterior, por lo general no hay problema en decir *conjunto* en lugar de *sucesión*.

Combinaciones lineales

Las leyes asociativa y conmutativa nos permiten eliminar los paréntesis en sumas múltiples, para simplificar la notación. Por ejemplo, las siguientes expresiones $(\mathbf{u} + \mathbf{v}) + (\mathbf{w} + \mathbf{r})$, $\mathbf{u} + ((\mathbf{v} + \mathbf{w}) + \mathbf{r})$, $\mathbf{u} + (\mathbf{v} + (\mathbf{w} + \mathbf{r}))$ y $\mathbf{v} + (\mathbf{u} + (\mathbf{w} + \mathbf{r}))$ se expresan como $\mathbf{u} + \mathbf{v} + \mathbf{w} + \mathbf{r}$. También, pueden escribirse sin ambigüedad expresiones como

$$\mathbf{v}_1 - 3\mathbf{v}_2 + 5\mathbf{v}_3 - 2\mathbf{v}_4$$

en la que $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$ son vectores del mismo tamaño. Estas sumas de múltiplos escalares de vectores se llaman *combinaciones lineales* de esos vectores.

(Combinación lineal)

Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vectores n, y sean $c_1, c_2, \ldots, c_k$ escalares. El vector n de la forma

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

se llama **combinación lineal** de $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Los escalares $c_1, \ldots, c_k$ se llaman **coeficientes** de la combinación lineal.

■ EJEMPLO 3 Calcular y dibujar la combinación lineal $\frac{1}{2}\mathbf{v}_1 - 3\mathbf{v}_2$, siendo

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

SOLUCIÓN $\frac{1}{2}\begin{bmatrix} 2 \\ 4 \end{bmatrix} - 3\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ -3 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$. Este proceso se muestra geométricamente en la figura 2.5. □

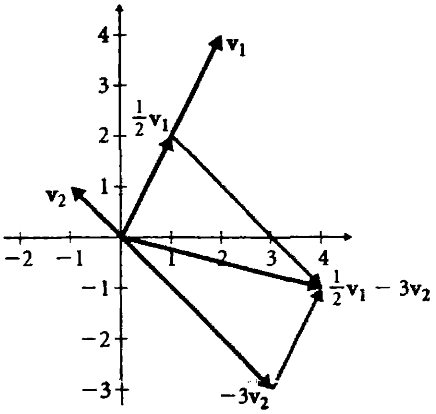


Figura 2.5 La combinación lineal $\frac{1}{2}\mathbf{v}_1 - 3\mathbf{v}_2$.

■ EJEMPLO 4 Determine si cada uno de los siguientes vectores

$$\mathbf{u} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

es una combinación lineal de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, donde

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

SOLUCIÓN Buscamos los escalares c_1, c_2 y c_3 tales que $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$.

$$\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Por consiguiente,

$$\begin{bmatrix} -c_1 + 2c_2 + c_3 \\ c_1 + c_3 \\ c_2 + c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \quad \text{o sea} \quad \begin{aligned} -c_1 + 2c_2 + c_3 &= 0 \\ c_1 + c_3 &= 2 \\ c_2 + c_3 &= 1 \end{aligned}$$

Este sistema es lineal, con las incógnitas c_1, c_2 y c_3. Su matriz aumentada tiene las columnas $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{u}$, y se reduce como sigue:

$$\begin{bmatrix} -1 & 2 & 1 & : & 0 \\ 1 & 0 & 1 & : & 2 \\ 0 & 1 & 1 & : & 1 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 1 & : & 2 \\ 0 & 1 & 1 & : & 1 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Por tanto, la solución del sistema es $c_1 = -r + 2$, $c_2 = -r + 1$, $c_3 = r$ para cualquier escalar r. En este caso hay una cantidad infinita de escalares tales que $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Por ejemplo, si $r = 0$, entonces $c_1 = 2$, $c_2 = 1$, $c_3 = 0$ y

$$\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

También buscamos los escalares d_1, d_2 y d_3 tales que $\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + d_3\mathbf{v}_3$, es decir

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = d_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + d_2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + d_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

La matriz aumentada del sistema correspondiente en d_1, d_2 y d_3 tiene las columnas $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}$, y se reduce como sigue:

$$\begin{bmatrix} -1 & 2 & 1 & : & 0 \\ 1 & 0 & 1 & : & 1 \\ 0 & 1 & 1 & : & 2 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 1 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix}$$

De aquí inferimos que el sistema es inconsistente. En vista de esto, $\mathbf{v}$ *no* es una combinación lineal de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

Sistemas lineales como ecuaciones vectoriales

En la solución del ejemplo 4, vimos que la ecuación vectorial $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{u}$ produjo un sistema cuya matriz aumentada tenía las columnas $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{u}$. A la inversa, el sistema que exhibe dicha modalidad de matriz representa la misma ecuación vectorial. Esta importante equivalencia de notación se aplica a cualquier sistema lineal.

Relación entre sistemas lineales y combinaciones lineales

Si $x_1, x_2, \ldots, x_n$ son las incógnitas de un sistema, la matriz de coeficientes es A, cuyos términos constantes son las componentes de un vector $\mathbf{b}$; por tanto, si $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ son las columnas de A, las notaciones siguientes son equivalentes:

$$[A : \mathbf{b}] \Leftrightarrow x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{2.1}$$

■ EJEMPLO 5 Escriba el sistema siguiente en forma de ecuación vectorial

$$\begin{aligned} x - 5y &= 1 \\ -x + 6y &= 3 \end{aligned}$$

SOLUCIÓN

$$x\begin{bmatrix} 1 \\ -1 \end{bmatrix} + y\begin{bmatrix} -5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

□

■ EJEMPLO 6 Escriba la ecuación vectorial

$$x_1\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

en forma de sistema lineal.

SOLUCIÓN

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

□

PREGUNTA ¿Qué expresión es equivalente a "el vector $\mathbf{b}$ es una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$"?

RESPUESTA El sistema cuya matriz aumentada es

$$[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n \quad : \quad \mathbf{b}]$$

es consistente. O también, "la última columna de esta matriz aumentada *no es* una columna pivote".

La equivalencia entre sistemas lineales y combinaciones lineales nos permite expresar las soluciones de los sistemas lineales en forma vectorial. Por ejemplo, la solución $x = 21$, $y = 4$ del sistema en el ejemplo 5 puede expresarse como $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 21 \\ 4 \end{bmatrix}$ o tan sólo como $\begin{bmatrix} 21 \\ 4 \end{bmatrix}$ si el orden de las variables está determinado.

PREGUNTA IMPORTANTE Con frecuencia, tanto en teoría como en la práctica (véase el ejemplo 11), se necesita responder la siguiente pregunta: ¿puede ser consistente para *toda* $\mathbf{b}$ el sistema cuya matriz aumentada es $[A: \mathbf{b}]$? La clave para contestarla es, una vez más, la reducción por renglones.

TEOREMA 2

Sea A una matriz de $m \times n$. Las afirmaciones siguientes son equivalentes.

1. El sistema lineal cuya matriz aumentada es $[A : \mathbf{b}]$ es consistente para todos los vectores $\mathbf{b} \in \mathbf{R}^m$.
2. Todo vector $\mathbf{b} \in \mathbf{R}^m$ es una combinación lineal de las columnas de A.
3. A tiene m posiciones pivote. (O bien, cada renglón tiene una posición pivote.)

DEMOSTRACIÓN De acuerdo con (2.1), las afirmaciones 1 y 2 son equivalentes, por lo que bastará demostrar dicha equivalencia entre 1 y 3. Suponga que el sistema es consistente para todo $\mathbf{b} \in \mathbf{R}^m$. Si un renglón de A no es renglón pivote, entonces cualquier forma de escalón de $[A : \mathbf{b}]$ tendrá un renglón de la forma $[0 \ \ 0 \ \ 0 \ \cdots \ 0 \ : \ b]$. Como $\mathbf{b}$ es cualquier vector n, podemos elegir componentes de $\mathbf{b}$ tales que $b \neq 0$. Pero entonces $[A : \mathbf{b}]$ será inconsistente para una $\mathbf{b}$ en particular. Esto contradice nuestra hipótesis de que $[A : \mathbf{b}]$ es consistente para todos los vectores m $\mathbf{b}$. En consecuencia, todos los renglones tienen una posición pivote. Por el contrario, supongamos que cada renglón de A tiene una posición pivote. Entonces, el elemento al final del último renglón de A es distinto de cero. Por consiguiente, la última columna de la matriz aumentada no puede ser una columna pivote, independientemente de lo que sea $\mathbf{b}$. Así, el sistema $[A : \mathbf{b}]$ es consistente para todos los vectores m $\mathbf{b}$. $\square$

NOTA

1. Decir que la matriz A de $m \times n$ tiene m posiciones pivote equivale a decir que $m \leq n$. (¿Por qué?)
2. Las partes 2 y 3 del teorema 2 se refieren a la matriz de *coeficientes* de un sistema, y *no* a la matriz aumentada. Por ejemplo, la matriz aumentada $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ tiene dos columnas pivote, pero el sistema correspondiente no tiene soluciones.

■ EJEMPLO 7 Se tienen los sistemas cuyas matrices aumentadas $[A : \mathbf{b}]$, $[B : \mathbf{b}]$ y $[C : \mathbf{b}]$, donde

$$A = \begin{bmatrix} -1 & 3 & 2 & 0 \\ 0 & 2 & -2 & 4 \\ 0 & -1 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 3 & 2 & 0 \\ 0 & 2 & -2 & 4 \\ 0 & -1 & 1 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & 3 \\ 0 & 2 \\ 0 & -1 \end{bmatrix}$$

¿Cuáles sistemas son consistentes para todos los vectores $\mathbf{b} \in \mathbf{R}^3$? ¿Qué puede decirse acerca de los sistemas que no lo son?

SOLUCIÓN Ya que

$$A \sim \begin{bmatrix} -1 & 3 & 2 & 0 \\ 0 & 2 & -2 & 4 \\ 0 & 0 & 0 & 4 \end{bmatrix} \qquad B \sim \begin{bmatrix} -1 & 3 & 2 & 0 \\ 0 & 2 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad C \sim \begin{bmatrix} -1 & 3 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$$

A tiene una posición pivote en cada renglón; por tanto, $[A : \mathbf{b}]$ se soluciona para todo $\mathbf{b} \in \mathbf{R}^3$ según el teorema 2. Los terceros renglones de B y C no tienen pivotes, y en consecuencia los sistemas $[B : \mathbf{b}]$ y $[C : \mathbf{b}]$ no son solubles para todo $\mathbf{b} \in \mathbf{R}^3$. De acuerdo con el ejercicio 50 de la sección 1.2, cada sistema de $[B : \mathbf{b}]$ y de $[C : \mathbf{b}]$ no tiene soluciones, o bien tiene un número infinito de ellas, según el valor que tenga $\mathbf{b}$. $\square$

La notación $(x_1, x_2, \ldots, x_n)$

A veces, para ahorrar espacio, se emplea la notación $(x_1, x_2, \ldots, x_n)$ para representar al vector cuyas componentes son $x_1, x_2, \ldots, x_n$. Esto no debe confundirse con la matriz $[x_1, x_2, \ldots, x_n]$, cuyo tamaño es $1 \times n$, y no es $n \times 1$. Así, $(1, 2)$ es lo mismo que $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, pero no que $[1 \quad 2]$.

Vectores libres

Un **vector libre** es una cantidad que puede determinarse por su magnitud y su dirección. Al igual que con los vectores, uno puede pensar geométricamente en los vectores libres como si fueran segmentos de recta dirigidos, o flechas, y pueden escribirse en la forma $\mathbf{a}, \mathbf{b}, \ldots$, o bien como $\overrightarrow{PQ}$, $\overrightarrow{RS}$, ... La longitud y la punta de la flecha describen la magnitud y la dirección del vector. En contraste con los vectores 2 y 3, los vectores libres pueden comenzar en cualquier punto del plano o del espacio. De hecho, todo vector libre cuyo punto inicial está en el origen del sistema de coordenadas es un vector.

Dos vectores libres $\mathbf{a}$ y $\mathbf{b}$ son **iguales** si tienen la misma magnitud y la misma dirección, y lo anterior se expresa como $\mathbf{a} = \mathbf{b}$, figura 2.6(a). *Un vector puede usarse para representar a todos los vectores libres iguales a él,* figura 2.6(b). Y aquél que corresponde a un conjunto de vectores libres iguales se llama **vector** del conjunto, o **vector de posición** del conjunto.

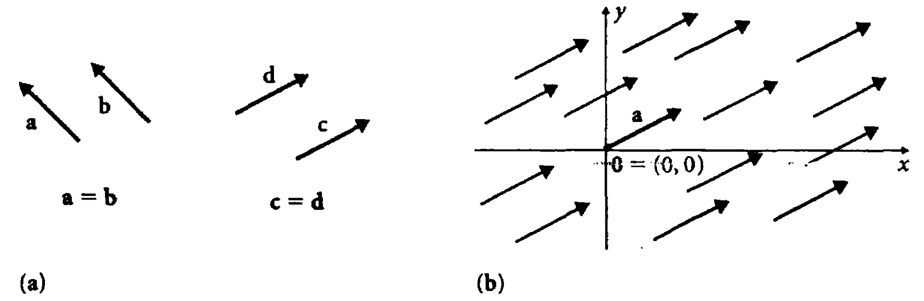


Figura 2.6 (a) Vectores libres iguales, (b) el vector de posición de los vectores libres iguales.

Dos vectores libres se suman determinando la suma de sus vectores correspondientes. De igual forma, la multiplicación por escalar se lleva a cabo en el vector correspondiente. Los componentes de un vector libre son los componentes de su vector. En general, *los vectores libres se estudian a través de sus vectores correspondientes.*

Un vector libre $\overrightarrow{PQ}$ en el plano, cuyo origen está en $P(p_1, p_2)$ y su punto terminal está en $Q(q_1, q_2)$, tiene las componentes

$$\overrightarrow{PQ} = (q_1 - p_1, q_2 - p_2)$$

porque

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = (q_1, q_2) - (p_1, p_2) = (q_1 - p_1, q_2 - p_2)$$

figura 2.7(a). La fórmula análoga es válida para los vectores libres en el espacio. Si $P(p_1, p_2, p_3)$ y $Q(q_1, q_2, q_3)$, los componentes de $\overrightarrow{PQ}$ se determinan con

$$\overrightarrow{PQ} = (q_1 - p_1, q_2 - p_2, q_3 - p_3)$$

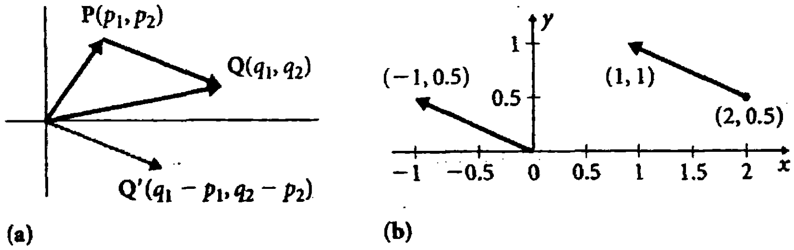


Figura 2.7 Componentes de vectores libres.

■ EJEMPLO 8 (Componentes de un vector libre). Encuentre las componentes de los siguientes vectores.

(a) $\overrightarrow{PQ}$ si $P(2, 0.5)$ y $Q(1, 1)$, figura 2.7(b).

(b) $\overrightarrow{RS}$ si $R(-2, -1, 4)$ y $S(-3, 1, 0)$.

SOLUCIÓN

$$\text{(a) } \overrightarrow{PQ} = \begin{bmatrix} 1-2 \\ 1-0.5 \end{bmatrix} = \begin{bmatrix} -1 \\ 0.5 \end{bmatrix}, \quad \text{(b) } \overrightarrow{RS} = \begin{bmatrix} -3-(-2) \\ 1-(-1) \\ 0-4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -4 \end{bmatrix}$$

Aplicaciones

■ EJEMPLO 9 (Equipo olímpico de clavados) El equipo olímpico de Estados Unidos, con 10 miembros, participó este año en tres competencias internacionales. Calcule el promedio de cada uno, si las calificaciones obtenidas en las competencias están representadas por los vectores **u**, **v** y **w**. ¿Cuál es la calificación promedio del décimo miembro del equipo?

$$\mathbf{u} = (8.5, 9.5, 8, 9.2, 9.9, 10, 8.8, 6.5, 9.4, 9.8)$$
$$\mathbf{v} = (9.5, 7.5, 8.2, 8.2, 8.9, 7.9, 7.8, 8.5, 9.4, 9.6)$$
$$\mathbf{w} = (8.5, 8.5, 8.9, 9.2, 8.6, 9.9, 9.8, 9.5, 9.1, 8.9)$$

SOLUCIÓN El vector promedio es

$$\frac{1}{3}(\mathbf{u} + \mathbf{v} + \mathbf{w}) = (8.83, 8.50, 8.36, 8.86, 9.13, 9.26, 8.80, 8.16, 9.3, 9.43)$$

Todos los elementos fueron truncados a dos cifras decimales. El décimo miembro del equipo tuvo un promedio de 9.43 puntos. □

Manufactura

■ EJEMPLO 10 Una empresa de artículos deportivos tiene dos fábricas, y en cada una se ensamblan bicicletas de montaña fabricadas en aluminio y titanio. La primera planta produce 150 bicicletas de aluminio y 15 de titanio por día. La segunda, 220 y 20, respectivamente. Si $\mathbf{v}_1 = \begin{bmatrix} 150 \\ 15 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 220 \\ 20 \end{bmatrix}$, calcule e interprete el significado de las expresiones (a) a (d).

(a) $\mathbf{v}_1 + \mathbf{v}_2$
(b) $\mathbf{v}_2 - \mathbf{v}_1$
(c) $10\mathbf{v}_1$
(d) $a\mathbf{v}_1 + b\mathbf{v}_2$, siendo $a, b > 0$.
(e) ¿Cuántos días debe trabajar cada fábrica para que la empresa entregue 2 600 bicicletas de aluminio y 250 de titanio?

SOLUCIÓN

(a) $\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 370 \\ 35 \end{bmatrix}$ es la cantidad total de bicicletas de aluminio (370) y de titanio (35) que producen las dos fábricas en un día.

(b) $\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 70 \\ 5 \end{bmatrix}$ indica cuántas bicicletas más produce la segunda fábrica en un día, en comparación con la primera fábrica.

(c) $10\mathbf{v}_1 = \begin{bmatrix} 1\,500 \\ 150 \end{bmatrix}$ representa cuántas bicicletas produce la primera fábrica en 10 días.

(d) $a\mathbf{v}_1 + b\mathbf{v}_2 = \begin{bmatrix} 150a + 220b \\ 15a + 20b \end{bmatrix}$ es la cantidad total de bicicletas producidas cuando la primera fábrica trabaja a días y la segunda trabaja b días.

(e) Sean x_1 y x_2 los días respectivos de trabajo. Entonces $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \begin{bmatrix} 2\,600 \\ 250 \end{bmatrix}$. Por tanto,

$$x_1 \begin{bmatrix} 150 \\ 15 \end{bmatrix} + x_2 \begin{bmatrix} 220 \\ 20 \end{bmatrix} = \begin{bmatrix} 2\,600 \\ 250 \end{bmatrix}$$

El sistema correspondiente tiene la matriz aumentada $\begin{bmatrix} 150 & 220 & 2\,600 \\ 15 & 20 & 250 \end{bmatrix}$ cuya forma de escalón reducida es $\begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & 5 \end{bmatrix}$. Por consiguiente, la primera fábrica debe trabajar 10 días, y la segunda 5 días. □

Transmisión de calor

Como ejemplo de la aplicación del teorema 2, regresaremos al problema de transferencia de calor que estudiamos en la sección 1.4. Una placa cuadrada tiene una distribución dada de temperaturas en sus bordes, $b_1, b_2, \ldots, b_8$ (figura 2.8). Se trata de calcular las temperaturas en el interior. Para simplificar, asignaremos elementos térmicos en ciertos nodos de la retícula. A continuación aplicaremos la *propiedad del valor promedio para conducción de calor*, que consiste en que la temperatura de un punto interior de la retícula es el promedio de las temperaturas de sus puntos vecinos en la retícula, y con ello se obtiene un sistema lineal. En este caso,

$$\begin{aligned} x_1 &= \tfrac{1}{4}(x_2 + x_3 + b_1 + b_3) \\ x_2 &= \tfrac{1}{4}(x_1 + x_4 + b_2 + b_4) \\ x_3 &= \tfrac{1}{4}(x_1 + x_4 + b_3 + b_7) \\ x_4 &= \tfrac{1}{4}(x_2 + x_3 + b_6 + b_8) \end{aligned} \tag{2.2}$$

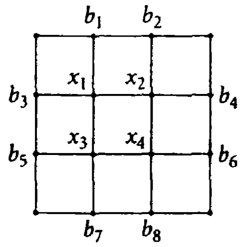


Figura 2.8 Transferencia de calor.

■ EJEMPLO 11 (Intuición de un ingeniero) Un ingeniero sabe (por intuición) que todos los valores de temperatura $b_1, \ldots, b_8$ en las orillas de la placa producen temperaturas únicas en $x_1, \ldots, x_4$. Demuestre matemáticamente lo anterior.

SOLUCIÓN Se necesita comprobar que el sistema (2.2) puede resolverse para toda $b_1, \ldots, b_8$. Primero se replantean las ecuaciones (2.2) en su forma canónica y se forma la matriz aumentada:

$$[A : \mathbf{b}] = \begin{bmatrix} 4 & -1 & -1 & 0 & : & b_1 + b_3 \\ -1 & 4 & 0 & -1 & : & b_2 + b_4 \\ -1 & 0 & 4 & -1 & : & b_5 + b_7 \\ 0 & -1 & -1 & 4 & : & b_6 + b_8 \end{bmatrix}$$

Ya que

$$A \sim \begin{bmatrix} 4 & -1 & -1 & 0 \\ 0 & -1 & -1 & 4 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & 0 & 12 \end{bmatrix}$$

A tiene cuatro pivotes. En consecuencia, $[A : \mathbf{c}]$ tiene una solución para todos los vectores 4 $\mathbf{c}$, según el teorema 2. En particular, si a $\mathbf{c}$ se le asignan los componentes $b_1 + b_3$, $b_2 + b_4$, $b_5 + b_7$ y $b_6 + b_8$, observaremos que el sistema (2.2) se resuelve para cualquier elección de temperaturas $b_1, \ldots, b_8$. □

Combinaciones lineales con sistemas algebraicos computacionales

Maple

```
> vector(2*[1,2,3]-5*[-3,2,-1]);
```

$$[17, -6, 11]$$

Mathematica

```
In[1]:=
     2*{1,2,3}-5*{-3,2,-1}
Out[1]=
     {17, -6, 11}
```

MATLAB

```
>> 2*[1 2 3]-5*[-3 2 -1]
ans =
     17    -6    11
```

Ejercicios 2.1

Vectores y aritmética vectorial

En los ejercicios 1 a 4 lleve a cabo, si es posible, las operaciones vectoriales indicadas.

1. a. $\begin{bmatrix} -1 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ -4 \end{bmatrix}$ b. $\begin{bmatrix} 5 \\ -3 \end{bmatrix} - \begin{bmatrix} 10 \\ 6 \end{bmatrix}$

2. a. $\frac{1}{3}\begin{bmatrix} 9 \\ -3 \end{bmatrix}$ b. $3\begin{bmatrix} 4 \\ 1 \end{bmatrix} + 2\begin{bmatrix} -6 \\ -2 \end{bmatrix}$

3. a. $\begin{bmatrix} 5 \\ -3 \\ 7 \end{bmatrix} - \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix}$ b. $2\begin{bmatrix} 5 \\ -3 \\ 7 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}$

4. a. $3\begin{bmatrix} a \\ 1 \end{bmatrix} + 2\begin{bmatrix} -1 \\ b \end{bmatrix}$ b. $2\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

5. Sean $\mathbf{a} = \begin{bmatrix} -5 \\ 8 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} -7 \\ 0 \end{bmatrix}$ Calcule:

a. $2\mathbf{a} - 4\mathbf{b} + 3\mathbf{c}$

b. $\mathbf{a} - 0\mathbf{b} - 6\mathbf{c} + 3\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

6. Determine un vector $\mathbf{x}$ tal que

a. $3\mathbf{x} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -2 \\ 4 \end{bmatrix}$

b. $\mathbf{x} + 2\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = -\begin{bmatrix} -7 \\ 4 \\ 4 \end{bmatrix}$

Sean

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

$$\mathbf{w}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$

En los ejercicios 7 a 9 trace los vectores.

7. $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 - \mathbf{v}_3, \mathbf{v}_3 - \mathbf{v}_1$

8. $2\mathbf{v}_1 - \mathbf{v}_2, 3\mathbf{v}_2 + 2\mathbf{v}_3, 3\mathbf{v}_1 - 2\mathbf{v}_3, -\mathbf{v}_1 - \mathbf{v}_2$

9. $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_1 + \mathbf{w}_2, 3\mathbf{w}_1, \mathbf{w}_2 - \mathbf{w}_1$

En los ejercicios 10 a 12 determine el o los valores (si es que los hay) de a y b que hacen válidas las igualdades.

10. a. $\begin{bmatrix} a-1 \\ a+b \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ b. $\begin{bmatrix} a-b \\ a+2b \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$

11. a. $\begin{bmatrix} a-b \\ a+b \end{bmatrix} - \frac{1}{2}\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

b. $\begin{bmatrix} b \\ a \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$

12. a. $\begin{bmatrix} a-b \\ a+b \end{bmatrix} - \frac{1}{2}\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

b. $\begin{bmatrix} b \\ a \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Sean **a** y **b** vectores 3.

13. Determine el vector 3, **x**, tal que $2\mathbf{x} - 4\mathbf{b} = 3\mathbf{a}$.
14. Calcule el vector 3, **x**, tal que $4\mathbf{x} + 3\mathbf{b} = 2\mathbf{a}$.
15. Encuentre los vectores 3, **x** y **y**, tales que

$$4\mathbf{x} - 3\mathbf{y} = 2\mathbf{a}$$
$$\mathbf{x} - \mathbf{y} = \mathbf{a} + \mathbf{b}$$

16. Determine los vectores 2, **x** y **y**, tales que

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$
$$2\mathbf{x} - 3\mathbf{y} = \begin{bmatrix} -10 \\ 15 \end{bmatrix}$$

17. Trace el vector libre $\vec{PQ}$ y calcule sus componentes.

a. $P = (1, 1), Q = (1, -1)$
b. $P = (2, -1, 1), Q = (-1, 1, 3)$

Combinaciones lineales

En los ejercicios 18 a 25 diga si el primer vector es una combinación lineal de los restantes.

18. $\begin{bmatrix} -a-2b \\ 4a+3b \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

19. $\begin{bmatrix} a+b+c \\ a+b-2c \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

20. $\begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

21. $\begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

22. $\begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

23. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

24. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

25. $\begin{bmatrix} -3 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

En los ejercicios 26 a 32 determine si **b** es una combinación lineal de las columnas de A.

26. $A = \begin{bmatrix} 1 & -1 \\ 0 & -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$

27. $A = \begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} x \\ y \end{bmatrix}$

28. $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

29. $A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 0 & -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix}$

30. $A = \begin{bmatrix} 1 & -1 & 6 \\ 2 & 1 & 3 \\ 0 & -2 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

31. $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

32. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

33. Demuestre que cualquier vector 3 $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ es una combinación lineal de $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, y $\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

34. Compruebe que cualquier vector 3 es una combinación lineal de las columnas de $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

35. Determine un vector 3 que *no sea* una combinación lineal de $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, y $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

36. Demuestre que cualquier vector 4 es una combinación lineal de los vectores de base estándar de $\mathbf{R}^4$.

En los ejercicios 37 a 41 describa el conjunto de todas las combinaciones lineales de los vectores dados:

37. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

38. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

39. $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

40. $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$

41. $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ en $\mathbf{R}^3$.

42. Calcule el o los valores de k tales que $\begin{bmatrix} k \\ 2 \\ -2k \end{bmatrix}$ sea una combinación lineal de $\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 0 \\ k \end{bmatrix}$

43. Calcule el o los valores de k tales que $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ sea una combinación lineal de las columnas de $\begin{bmatrix} 1 & -k & 1 \\ 0 & k & 0 \end{bmatrix}$

Ecuaciones vectoriales y sistemas lineales

44. Escriba los siguientes sistemas en forma de ecuaciones vectoriales.

a. $$\begin{aligned} x - 4y &= 1 \\ -2x + y &= 0 \end{aligned}$$

b. $$\begin{aligned} x - 4y + z &= 2x + 1 \\ -2x + y - z &= -y \end{aligned}$$

45. Exprese los siguientes sistemas en forma de ecuaciones vectoriales.

a. $$\begin{aligned} x - 2y - z &= 1 \\ -x + y \quad &= 0 \\ y - z &= 2 \end{aligned}$$

b. $$\begin{aligned} x - 2y - z &= 1 \\ -x \quad + 2z &= -1 \\ y - 2 &= 0 \end{aligned}$$

46. Plantee las ecuaciones vectoriales en forma de sistemas lineales.

a. $x\begin{bmatrix} 3 \\ -2 \end{bmatrix} - y\begin{bmatrix} -1 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

b. $x\begin{bmatrix} -3 \\ 2 \end{bmatrix} = y\begin{bmatrix} 4 \\ 5 \end{bmatrix}$

En los ejercicios 47 y 48 diga si el sistema $[A : \mathbf{b}]$ es consistente para todos los vectores $\mathbf{b} \in \mathbf{R}^3$, si A es

47. **a.** $\begin{bmatrix} 0 & 2 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$ **b.** $\begin{bmatrix} -1 & 0 & 2 \\ 1 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$

48. **a.** $\begin{bmatrix} -1 & 0 & 2 \\ 1 & 1 & -1 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 4 & -1 & 3 \end{bmatrix}$

Ejercicios de teoría

49. Demuestre las partes 2 a 5 del teorema 1.
50. Compruebe las partes 6 a 9 del teorema 1.
51. Sea A una matriz de $m \times n$. Suponga que el sistema $[A : \mathbf{b}]$ es consistente para todos los vectores n $\mathbf{b}$. ¿Es cierto o falso? (Explique por qué.)
    - **a.** $\mathbf{b}$ es una combinación lineal en las columnas de A.
    - **b.** A tiene m pivotes.
    - **c.** A tiene n pivotes.
    - **d.** Cada renglón de A tiene un pivote.
    - **e.** Cada columna de A tiene un pivote.
    - **f.** A puede tener más de m pivotes.
    - **g.** A puede tener más de n pivotes.
    - **h.** $m \leq n$.
52. Suponga que el sistema $[A : \mathbf{b}]$ es inconsistente.
    - **a.** ¿Qué puede decir acerca de la cantidad de pivotes de A?
    - **b.** ¿Puede ser $\mathbf{b}$ un múltiplo escalar de la primera columna de A?

Aplicaciones

53. Un bote de vela viaja a 10 mi/h hacia el este, y sopla un viento cruzado, de sur a norte, de 20 mi/h. ¿Cuál es el vector velocidad del bote?
54. Durante el despegue, un avión se eleva con rapidez de 20 mi/h. Su rapidez hacia el este es 150 mi/h, y su rapidez hacia el norte es 200 mi/h. Calcule el vector velocidad total.

55. Una aerolínea compra suministros para tres de sus aviones. El costo promedio por viaje, en dólares, se expresa con la matriz A cuyas columnas son $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$.

Clase	**Avión 1**	**Avión 2**	**Avión 3**
Primera	350	300	450
Negocios	500	600	700
Económica	800	700	900

Calcule y explique el significado de lo siguiente:

a. $\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3$ **b.** $\mathbf{a}_3 - \mathbf{a}_2$

c. $10\mathbf{a}_3$ **d.** $7\mathbf{a}_1 + 8\mathbf{a}_2 + 9\mathbf{a}_3$

56. Acerca del ejercicio 55, ¿cuántos viajes hizo cada avión, si la aerolínea gastó \$23 000 para primera clase, \$38 000 para la de negocios y \$49 000 para la económica? (*Sugerencia:* Si x_1, x_2 y x_3 son la cantidad de viajes que hizo cada avión, entonces $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \begin{bmatrix} 23{,}000 \\ 38{,}000 \\ 49{,}000 \end{bmatrix}$).

2.2 Producto punto

Objetivos del alumno para esta sección

1. Calcular la longitud de un vector, el producto punto y el ángulo entre vectores.
2. Calcular la proyección ortogonal de un vector sobre otro vector.

En esta sección definiremos al producto punto, o producto escalar, y la longitud en $\mathbf{R}^n$. Estos conceptos son básicos en la teoría y las aplicaciones de los vectores. En dos y tres dimensiones tienen interpretaciones geométricas familiares.

DEFINICIÓN

(Producto punto)

Sean $\mathbf{u} = (u_1, \ldots, u_n)$ y $\mathbf{v} = (v_1, \ldots, v_n)$ dos vectores n cualquiera. El **producto punto**, o **producto escalar** de $\mathbf{u}$ y $\mathbf{v}$ es el *número* siguiente:

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + \cdots + u_nv_n.$$

■ EJEMPLO 12 Sean

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -2 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} -2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

Calcular $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{u} \cdot \mathbf{w}$ y $\mathbf{v} \cdot \mathbf{w}$.

SOLUCIÓN

$$\mathbf{u} \cdot \mathbf{v} = (-1) \cdot (-2) + 2 \cdot 0 + 3 \cdot 2 + 0 \cdot 0 = 8$$
$$\mathbf{u} \cdot \mathbf{w} = (-1) \cdot (-2) + 2 \cdot 0 + 3 \cdot (-2) + 0 \cdot 1 = -4$$
$$\mathbf{v} \cdot \mathbf{w} = (-2) \cdot (-2) + 0 \cdot 0 + 2 \cdot (-2) + 0 \cdot 1 = 0$$

El ejemplo 12 demuestra que el producto punto puede ser menor que, mayor que o igual a 0. De hecho, si el producto punto de dos vectores es 0, se dice que son **ortogonales** entre sí. Así, $\mathbf{v}$ y $\mathbf{w}$ son ortogonales, pero $\mathbf{u}$ y $\mathbf{v}$ no lo son.

La **norma**, **longitud** o **magnitud** de un vector n, $\mathbf{u}$, es la raíz cuadrada positiva:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}} = \left(u_1^2 + \cdots + u_n^2\right)^{1/2}$$

La norma siempre está definida, porque

$$\|\mathbf{u}\|^2 = \mathbf{u}\cdot\mathbf{u} \tag{2.3}$$

es ≥ 0.

También, la **distancia (euclidiana)** entre $\mathbf{u}$ y $\mathbf{v}$ se define como sigue:

$$d = \|\mathbf{u} - \mathbf{v}\|$$

■ EJEMPLO 13 Calcule la norma de $\mathbf{w}$, que es $\|\mathbf{w}\|$, y la distancia d entre $\mathbf{u}$ y $\mathbf{v}$.

SOLUCIÓN La norma de $\mathbf{w}$ es

$$\|\mathbf{w}\| = \left((-2)^2 + 0^2 + (-2)^2 + 1^2\right)^{1/2} = \sqrt{9} = 3$$

La distancia entre $\mathbf{u}$ y $\mathbf{v}$ es

$$d = \|\mathbf{u} - \mathbf{v}\| = \|(1,2,1,0)\| = \sqrt{6}$$

Observe que la norma de un producto escalar, $c\mathbf{u}$, es

$$\|c\mathbf{u}\| = |c|\,\|\mathbf{u}\| \tag{2.4}$$

porque $\sqrt{c^2} = |c|$. Por ejemplo,

$$\|-5(1,2)\| = \|(-5,-10)\| = 5\sqrt{5} = |-5|\,\|(1,2)\|$$

La norma de un vector 2 o un vector 3 es exactamente lo que en geometría llamamos longitud. Si $\mathbf{u} = (u_1, u_2)$, entonces $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2} = \sqrt{(OP')^2 + (PP')^2} = OP$, de acuerdo con el teorema de Pitágoras, figura 2.9(a). Del mismo modo, si $\mathbf{u} = (u_1, u_2, u_3)$, entonces $\|\mathbf{u}\|$ es la longitud geométrica, aplicando el teorema de Pitágoras a OPP' y a $OP'P''$ en la figura 2.9(b).

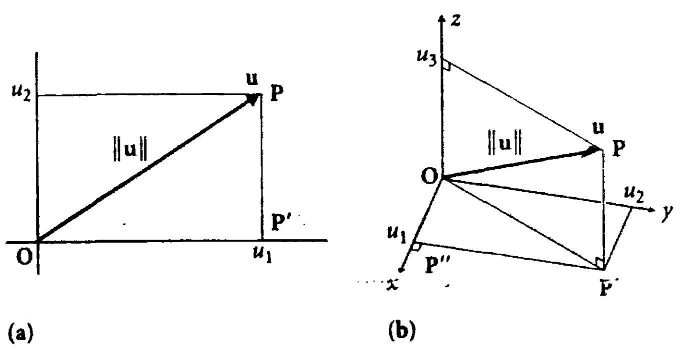


Figura 2.9 Magnitudes o normas de vectores en el plano y en el espacio.

Si $\overrightarrow{PQ}$ es un vector libre siendo $P(p_1, p_2, p_3)$ y $Q(q_1, q_2, q_3)$, en ese caso la longitud $\|\overrightarrow{PQ}\|$ de $\overrightarrow{PQ}$ es la longitud del vector correspondiente, es decir,

$$\left\|\overrightarrow{PQ}\right\| = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + (q_3 - p_3)^2}$$

Esta fórmula también expresa la **distancia** entre los puntos P y Q en el espacio. Entonces, la distancia entre dos vectores, tal como se definió, es la distancia geométrica en el espacio entre las puntas de los vectores. Si $p_3 = q_3 = 0$, se obtiene la distancia entre dos puntos en el plano.

■ EJEMPLO 14

(a) Calcule la longitud de $\mathbf{v} = (1, -2, 2)$.
(b) Determine la distancia entre $P(2, 3, -1)$ y $Q(-1, 0, -2)$.

SOLUCIÓN

$$\|\mathbf{v}\| = \sqrt{1^2 + (-2)^2 + 2^2} = 3$$
$$\left\|\overrightarrow{PQ}\right\| = \sqrt{(-3)^2 + (-3)^2 + (-1)^2} = \sqrt{19}$$

Un **vector** unitario el aquél cuya longitud es 1.

■ EJEMPLO 15 $\mathbf{u} = (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$ es un vector unitario, porque

$$\|\mathbf{u}\|^2 = \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 = 1$$

Por consiguiente, $\|\mathbf{u}\| = 1$.

Con frecuencia nos interesa obtener el vector unitario en la dirección de un vector dado.

TEOREMA 3

(Vector unitario en una dirección dada)

Sea $\mathbf{v} = (v_1, \ldots, v_n)$ un vector distinto de cero, y sea $\mathbf{u}$ el vector unitario en la dirección de $\mathbf{v}$. entonces

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \left(\frac{v_1}{\|\mathbf{v}\|}, \ldots, \frac{v_n}{\|\mathbf{v}\|}\right)$$

DEMOSTRACIÓN Como $\mathbf{u}$ es un múltiplo escalar positivo de $\mathbf{v}$, tiene la misma dirección que $\mathbf{v}$. Además, $\mathbf{u}$ tiene longitud unitaria, porque $\left\|\frac{1}{\|\mathbf{v}\|}\mathbf{v}\right\| = \frac{1}{\|\mathbf{v}\|}\|\mathbf{v}\| = 1$, de acuerdo con la ecuación (2.4).

■ EJEMPLO 16 Determine el vector unitario en la dirección de $\mathbf{v} = (1, -2, 1)$.

SOLUCIÓN Según el teorema 3,

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{\sqrt{6}}(1, -2, 1) = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

Los vectores 3 unitarios a lo largo de los ejes coordenados son $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$. También suele representárseles por **i**, **j** y **k**, respectivamente, figura. 2.10(b). Esta notación se acostumbra en física e ingeniería.

$$\mathbf{i} = \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{j} = \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{k} = \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

En el plano se tiene, figura 2.10(a),

$$\mathbf{i} = \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{j} = \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

□

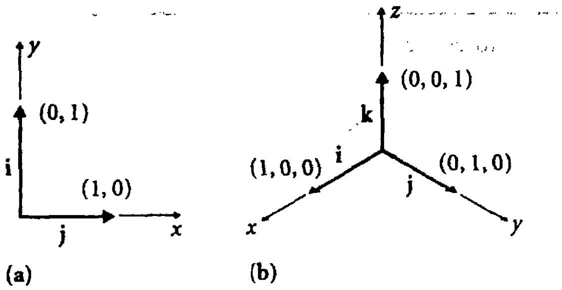


Figura 2.10 Vectores unitarios a lo largo de los ejes coordenados.

El producto punto tiene las propiedades que se resumen en el teorema 4.

TEOREMA 4

(Propiedades del producto punto)

Sean **u**, **v** y **w** vectores n, y c cualquier escalar.

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$	**Simetría**
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$	**Aditividad**
3. $c(\mathbf{u} \cdot \mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v})$	**Homogeneidad**
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$. Además, $\mathbf{u} \cdot \mathbf{u} = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.	**Definición positiva**

DEMOSTRACIÓN Únicamente comprobaremos las partes 2 y 4, y dejaremos los casos restantes como ejercicios. Sean $\mathbf{u} = (u_1, \ldots, u_n)$, $\mathbf{v} = (v_1, \ldots, v_n)$ y $\mathbf{w} = (w_1, \ldots, w_n)$. La parte 2 es consecuencia de

$$
\begin{aligned}
\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) &= \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \cdot \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix} \\
&= u_1(v_1 + w_1) + \cdots + u_n(v_n + w_n) \\
&= (u_1v_1 + \cdots + u_nv_n) + (u_1w_1 + \cdots + u_nw_n) \\
&= \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}
\end{aligned}
$$

En el caso de la parte 4,

$$\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2 \geq 0$$

y

$$
\begin{aligned}
\mathbf{u} \cdot \mathbf{u} = 0 &\Leftrightarrow u_1^2 + \cdots + u_n^2 = 0 \\
&\Leftrightarrow u_1 = \cdots = u_n = 0 \\
&\Leftrightarrow \mathbf{u} = \mathbf{0}
\end{aligned}
$$

Es posible combinar la ecuación (2.3) y el teorema 4 para formar muchas identidades nuevas. Por ejemplo, veamos el teorema 5.

TEOREMA 5

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores n cualquiera, se tiene

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v}$$

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u} \cdot \mathbf{v}$$

DEMOSTRACIÓN Como

$$
\begin{aligned}
\|\mathbf{u} + \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) && \text{de acuerdo con la ecuación (2.3)} \\
&= (\mathbf{u} + \mathbf{v}) \cdot \mathbf{u} + (\mathbf{u} + \mathbf{v}) \cdot \mathbf{v} && \text{según el teorema 4, parte 2} \\
&= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} && \text{tomando en cuenta las partes 1 y 2} \\
&= \mathbf{u} \cdot \mathbf{u} + 2\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} && \text{según la parte 1} \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v} && \text{de acuerdo con la ecuación (2.3)}
\end{aligned}
$$

Si se reemplaza $\mathbf{v}$ por $-\mathbf{v}$, se llega a la segunda identidad.

Una de las consecuencias más útiles del teorema 4 es la desigualdad de Cauchy-Schwarz.[3]

TEOREMA 6

(Desigualdad de Cauchy-Schwarz)

Para cualesquiera dos vectores n $\mathbf{u}$ y $\mathbf{v}$,

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$$

Además, la igualdad es válida si y sólo si $\mathbf{u}$ y $\mathbf{v}$ son múltiplos escalares entre sí.

[3] Los créditos y biografías pertinentes se describen en el capítulo 8.

DEMOSTRACIÓN Según el teorema 5,

$$0 \leq (x\mathbf{u} + \mathbf{v}) \cdot (x\mathbf{u} + \mathbf{v}) = x^2(\mathbf{u} \cdot \mathbf{u}) + x(2\mathbf{u} \cdot \mathbf{v}) + \mathbf{v} \cdot \mathbf{v} \tag{2.5}$$

para todos los escalares x. Es un polinomio cuadrático $p(x) = ax^2 + bx + c$, en donde $a = \mathbf{u} \cdot \mathbf{u}$, $b = 2\mathbf{u} \cdot \mathbf{v}$ y $c = \mathbf{v} \cdot \mathbf{v}$. Como $a \geq 0$ y $p(x) \geq 0$ para toda x, la gráfica de $p(x)$ es una parábola en el semiplano superior, cóncava hacia arriba. Por consiguiente, la parábola está arriba del eje x, en cuyo caso $p(x)$ tiene dos raíces complejas, o bien es tangente al eje x, y en esa circunstancia $p(x)$ tiene una raíz real repetida. En consecuencia, $b^2 - 2ac \leq 0$. Así

$$(2\mathbf{u} \cdot \mathbf{v})^2 - 4(\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}) \leq 0, \quad \text{o sea } 4(\mathbf{u} \cdot \mathbf{v})^2 - 4\|\mathbf{u}\|^2\|\mathbf{v}\|^2 \leq 0$$

lo cual implica la desigualdad de Cauchy-Schwarz.

La igualdad se aplica si y sólo si $b^2 - 2ac = 0$, o si y sólo si $p(x)$ tiene una raíz real doble, por ejemplo r. Por consiguiente, de acuerdo con la ecuación (2.5) con $x = r$,

$$\begin{aligned}(r\mathbf{u} + \mathbf{v}) \cdot (r\mathbf{u} + \mathbf{v}) = 0 &\Leftrightarrow \|r\mathbf{u} + \mathbf{v}\| = 0 \\ &\Leftrightarrow r\mathbf{u} + \mathbf{v} = \mathbf{0} \\ &\Leftrightarrow \mathbf{v} = -r\mathbf{u}\end{aligned}$$

Entonces $\mathbf{v}$ es un producto escalar de $\mathbf{u}$, lo cual demuestra la última afirmación del teorema. □

■ EJEMPLO 17 Compruebe la desigualdad Cauchy-Schwarz para $\mathbf{u} = (-1, 2, 0, -1)$ y $\mathbf{v} = (4, -2, -1, 1)$.

SOLUCIÓN

$$\begin{aligned}|\mathbf{u} \cdot \mathbf{v}| &= |(-1, 2, 0, -1) \cdot (4, -2, -1, 1)| = |-9| = 9 \\ \|\mathbf{u}\|\,\|\mathbf{v}\| &= \|(-1, 2, 0, -1)\|\,\|(4, -2, -1, 1)\| \\ &= \sqrt{6}\sqrt{22} \cong 11.489 \geq 9\end{aligned}$$

□

Como aplicación del teorema 4, y de la desigualdad de Cauchy-Schwarz, se obtiene el teorema siguiente útil.

TEOREMA 7

(La desigualdad del triángulo)

Para cualesquiera dos vectores *n* $\mathbf{u}$ y $\mathbf{v}$, tenemos

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

DEMOSTRACIÓN

$$\begin{aligned}\|\mathbf{u} + \mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v} && \text{según el teorema 5} \\ &\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| && \text{por la desigualdad de Cauchy-Schwarz} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2\end{aligned}$$

Por consiguiente, $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$. □

En dos y en tres dimensiones, la desigualdad del triángulo también puede visualizarse geométricamente: cada lado de un triángulo tiene su longitud menor que la suma de las longitudes de los otros dos lados (figura 2.11).

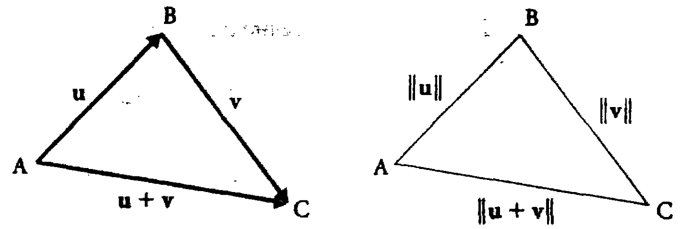


Figura 2.11 La desigualdad del triángulo en el plano.

Ángulo entre dos vectores

La desigualdad de Cauchy-Schwarz implica que

$$\frac{|\mathbf{u}\cdot\mathbf{v}|}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \le 1, \quad \text{o sea} \quad -1 \le \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \le 1$$

Como cualquier número entre -1 y 1 puede expresarse como $\cos\theta$ para $0 \le \theta \le \pi$ única, la última desigualdad permite definir el ángulo entre dos vectores n.

El **ángulo** entre **u** y **v**, dos vectores n distintos de cero, es el número θ único tal que

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}, \qquad 0 \le \theta \le \pi \tag{2.6}$$

También puede expresarse el producto punto en función del ángulo:

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta \tag{2.7}$$

■ **EJEMPLO 18** Calcule el ángulo θ entre los vectores

(a) $(1, 1)$ y $(3, 0)$
(b) $(1, 1, 1, 1)$ y $(1, 0, 0, 0)$

SOLUCIÓN

$$\text{(a)}\ \theta = \arccos\frac{(1,1)\cdot(3,0)}{\|(1,1)\|\,\|(3,0)\|} = \arccos\frac{1}{\sqrt{2}} = \frac{\pi}{4} = 45^\circ$$

$$\text{(b)}\ \theta = \arccos\frac{(1,1,1,1)\cdot(1,0,0,0)}{\|(1,1,1,1)\|\,\|(1,0,0,0)\|} = \arccos\frac{1}{2} = \frac{\pi}{3} = 60^\circ$$

□

Nuestra noción geométrica de ángulo concuerda con la definición anterior, en dos y tres dimensiones (figura 2.12). De hecho, puede comprobarse la ecuación (2.7) usando la ley de los cosenos como sigue:

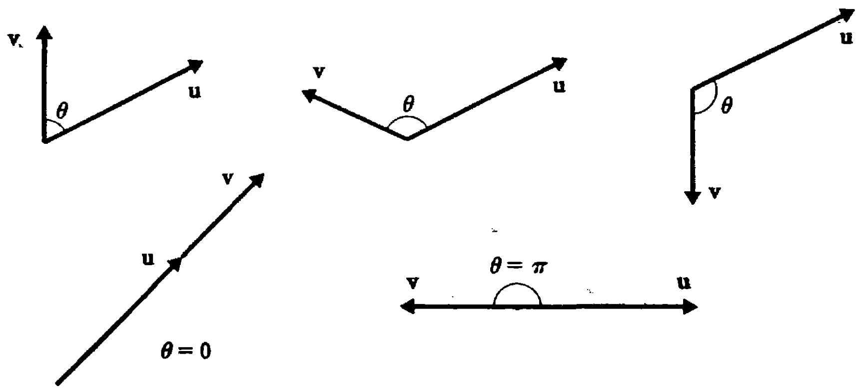


Figura 2.12 El ángulo θ formado por dos vectores es, en radianes, $0 \leq \theta \leq \pi$.

De acuerdo con la ley de los cosenos aplicada al triángulo OPQ (figura 2.13),

$$\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta = \frac{1}{2}\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|PQ\|^2\right)$$

Pero $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|PQ\|^2 + 2\mathbf{u}\cdot\mathbf{v}$, según el teorema 5. Por tanto,

$$\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta = \frac{1}{2}\left(\|PQ\|^2 + 2\mathbf{u}\cdot\mathbf{v} - \|PQ\|^2\right) = \mathbf{u}\cdot\mathbf{v}$$

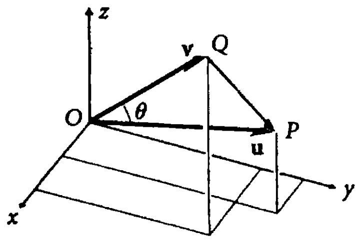


Figura 2.13

La ecuación (2.7) implica que $\mathbf{u}\cdot\mathbf{v}$ y $\cos\theta$ tienen el mismo signo, y uno es cero sólo si el otro lo es. Como el $\cos\theta$ es positivo para $0 \leq \theta \leq \frac{\pi}{2}$, negativo para $\frac{\pi}{2} < \theta \leq \pi$, y cero cuando $\theta = \frac{\pi}{2}$, entonces

1. θ es **agudo** si y sólo si $\mathbf{u}\cdot\mathbf{v} > 0$.
2. θ es **obtuso** si y sólo si $\mathbf{u}\cdot\mathbf{v} < 0$
3. θ es **recto** si y sólo si $\mathbf{u}\cdot\mathbf{v} = 0$.

Si $\theta = \frac{\pi}{2}$, entonces se dice que **u** y **v** son **perpendiculares**, lo cual se expresa con $\mathbf{u} \perp \mathbf{v}$. Como $\theta = \frac{\pi}{2}$ equivale a $\mathbf{u}\cdot\mathbf{v} = 0$ para vectores no cero, es evidente que **u** y **v** son *perpendiculares siempre y cuando sean ortogonales*, es decir,

$$\mathbf{u} \perp \mathbf{v} \Leftrightarrow \mathbf{u}\cdot\mathbf{v} = 0 \quad \text{cuando } \mathbf{u} \neq \mathbf{0} \text{ y } \mathbf{v} \neq \mathbf{0}$$

El teorema de Pitágoras también se puede generalizar (figura 2.14).

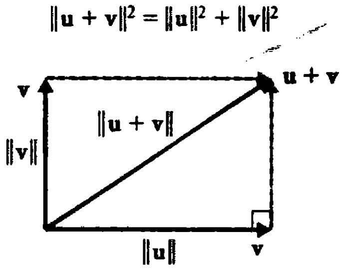


Figura 2.14 El teorema de Pitágoras.

TEOREMA 8

(Teorema de Pitágoras)

Los vectores *n* **u** y **v** son ortogonales si y sólo si

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

DEMOSTRACIÓN Se deja como ejercicio. (*Sugerencia:* Aplique el teorema 5.) □

Proyecciones ortogonales

Los productos punto pueden aplicarse para expresar cualquier vector *n* como la suma de vectores ortogonales. Sean **u** y **v** los vectores dados, distintos de cero, deseamos escribir **u** en la forma

$$\mathbf{u} = \mathbf{u}_{pr} + \mathbf{u}_c$$

en donde $\mathbf{u}_{pr}$ es un múltiplo escalar de **v** y $\mathbf{u}_c$, es ortogonal a $\mathbf{u}_{pr}$ (figura 2.15). Como vemos, siempre es posible esa descomposición de **u**, y es *única*. A $\mathbf{u}_{pr}$ se le llama **proyección ortogonal de u sobre v**, y a $\mathbf{u}_c$ se le denomina la **componente vectorial de u ortogonal a v**.

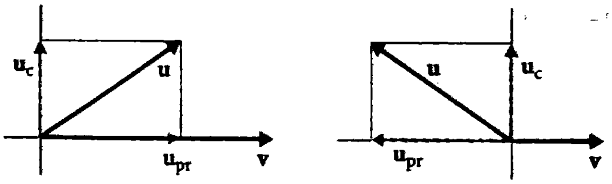


Figura 2.15 La proyección ortogonal de **u** sobre **v**.

Ahora calcularemos $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ en función de **u** y **v**. Como $\mathbf{u}_{pr}$ y **v** tienen la misma dirección, $\mathbf{u}_{pr} = c\mathbf{v}$ para algún escalar c. Además, como $\mathbf{u}_c$ y **v** son ortogonales, $\mathbf{u}_c \cdot \mathbf{v} = 0$. Por consiguiente, $\mathbf{u}_c \cdot \mathbf{v} = 0$. Por tanto,

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= (\mathbf{u}_{pr} + \mathbf{u}_c) \cdot \mathbf{v} \\ &= \mathbf{u}_{pr} \cdot \mathbf{v} + \mathbf{u}_c \cdot \mathbf{v} \\ &= (c\mathbf{v}) \cdot \mathbf{v} + 0 \\ &= c(\mathbf{v} \cdot \mathbf{v}) \\ &\Rightarrow c = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\end{aligned}$$

Y entonces

$$\mathbf{u}_{pr} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v} \qquad \text{proyección ortogonal de } \mathbf{u} \text{ sobre } \mathbf{v} \tag{2.8}$$

y

$$\mathbf{u}_c = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v} \qquad \text{componente vectorial de } \mathbf{u} \text{ ortogonal a } \mathbf{v} \tag{2.9}$$

■ EJEMPLO 19 (Proyección ortogonal) Si $\mathbf{u} = (1, 1, 1)$ y $\mathbf{v} = (2, 2, 0)$, determine la proyección ortogonal de $\mathbf{u}$ sobre $\mathbf{v}$ y la componente vectorial de $\mathbf{u}$ ortogonal a $\mathbf{v}$.

SOLUCIÓN

$$\mathbf{u}_{pr} = \frac{(1,1,1) \cdot (2,2,0)}{(2,2,0) \cdot (2,2,0)}(2,2,0) = \frac{4}{8}(2,2,0) = (1,1,0)$$

y

$$\mathbf{u}_c = \mathbf{u} - \mathbf{u}_{pr} = (1, 1, 1) - (1, 1, 0) = (0, 0, 1)$$

(Verifíquelo geométricamente.) $\square$

Producto punto con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> dotprod([1,2,3],[-3,2,-1]);
```

$$-2$$

Mathematica

```
In[1]:=
    Dot[{1,2,3},{-3,2,-1}]
Out[1]=
    -2
```

MATLAB

```
>> dot([1 2 3], [-3 2 -1])
ans =
    -2
```

Ejercicios 2.2

Sean $\mathbf{u} = (-1, 2, -2)$, $\mathbf{v} = (4, -3, 5)$, $\mathbf{w} = (-4, -2, 0)$ y $\mathbf{d} = (-1, -2, 1, \sqrt{3})$.

1. Encuentre las longitudes de los vectores siguientes:

$\mathbf{u}$	$\mathbf{v}$	$\mathbf{w}$
$\mathbf{u}+\mathbf{v}$	$\mathbf{u}-\mathbf{v}$	$\mathbf{u}-\mathbf{v}+\mathbf{w}$
$\mathbf{d}$	$10\mathbf{d}$	$\|\mathbf{d}\|\mathbf{d}$

2. Calcule las expresiones siguientes.

$\|\mathbf{u}+\mathbf{v}\|$	$\|\mathbf{u}\|+\|\mathbf{v}\|$
$\|\mathbf{u}\|-\|\mathbf{v}\|$	$\|\mathbf{u}-\mathbf{v}\|$
$\|\mathbf{v}\|\mathbf{v}+\|\mathbf{w}\|\mathbf{w}$	$(1/\|\mathbf{d}\|)\mathbf{d}$

3. Determine las expresiones siguientes.

$\mathbf{u}\cdot\mathbf{v}$	$\mathbf{w}\cdot\mathbf{u}$
$\mathbf{u}\cdot(\mathbf{v}+\mathbf{w})$	$\mathbf{v}\cdot\mathbf{u}+\mathbf{w}\cdot\mathbf{u}$
$\mathbf{d}\cdot\mathbf{d}$	$(\mathbf{d}\cdot\mathbf{d})\mathbf{d}$

4. ¿Cuáles de las siguientes expresiones son *indefinidas*, y por qué?

$\mathbf{u}\cdot\mathbf{v}\cdot\mathbf{w}$	$\mathbf{u}\cdot(\mathbf{v}\cdot\mathbf{w})$
$(\mathbf{u}\cdot\mathbf{v})\mathbf{w}$	$(\mathbf{u}\cdot\mathbf{v})(\mathbf{v}\cdot\mathbf{w})$
$\mathbf{u}\cdot(3\mathbf{v})$	$\mathbf{u}\cdot(3+\mathbf{v})$
$(\mathbf{d}\cdot\mathbf{d})^3$	$\mathbf{d}\cdot\mathbf{d}+2$

5. Determine el vector unitario en cada dirección.

   a. $\mathbf{u}$ b. $\mathbf{v}$ c. $\mathbf{w}$ d. $\mathbf{d}$

6. Encuentre un vector de longitud 2 con la dirección de $\mathbf{u}$.
7. Calcule un vector de longitud 2 en dirección opuesta a la de $\mathbf{u}$.
8. Determine un vector en la dirección de $\mathbf{d}$ cuya longitud sea 9 veces la de $\mathbf{d}$.
9. Encuentre un vector de longitud 9 en la dirección de $\mathbf{d}$.
10. Calcule la distancia entre los puntos P y Q.

    a. $P(1, 1, -1)$, $Q(-2, 3, 4)$
    b. $P(4, -3, 2, 0)$, $Q(0, 2, -6, 4)$

11. ¿Cuáles pares de vectores son ortogonales?

    a. $(1, 1)$, $(1, -1)$
    b. $(1, -1, -1)$, $(0, 1, -1)$
    c. $(4, -2, -1)$, $(-2, 3, 4)$
    d. $(-7, -3, 1, 0)$, $(0, 2, 6, 4)$

12. Encuentre dos vectores que sean, cada uno, ortogonales a $(1, -2, 4)$.
13. Determine un vector unitario que sea ortogonal a $(-2, 3, 1)$.

En los ejercicios 14 y 15 calcule el ángulo que forman $\mathbf{u}$ y $\mathbf{v}$. (Necesitará una calculadora).

14. a. $\mathbf{u} = (1, 1)$, $\mathbf{v} = (1, -1)$
    b. $\mathbf{u} = (-1, 1, 1)$, $\mathbf{v} = (2, -2, 1)$
    c. $\mathbf{u} = (1, -1, 1, -1)$, $\mathbf{v} = (0, 1, 1, 0)$
15. a. $\mathbf{u} = (3, 4)$, $\mathbf{v} = (5, 12)$
    b. $\mathbf{u} = (\sqrt{3}, 1, -2)$, $\mathbf{v} = (1, \sqrt{3}, -2)$
    c. $\mathbf{u} = (1, -1, -1, -1)$, $\mathbf{v} = (1, 1, 1, 1)$

16. Calcule la proyección ortogonal de $\mathbf{u}$ sobre $\mathbf{v}$.

    a. $\mathbf{u} = (2, 3)$, $\mathbf{v} = (-2, 1)$
    b. $\mathbf{u} = (0, -1, 6)$, $\mathbf{v} = (-1, -3, 5)$
    c. $\mathbf{u} = (-2, -1, 0, 1)$, $\mathbf{v} = (0, 0, -1, 3)$

17. Para el ejercicio 16, encuentre la componente vectorial de $\mathbf{u}$ ortogonal a $\mathbf{v}$.
18. Compruebe la desigualdad de Cauchy-Schwarz para los pares $\mathbf{u}$ y $\mathbf{v}$ del ejercicio 16.
19. Demuestre el teorema 8.
20. ¿Es cierto que si $\mathbf{u}\cdot\mathbf{v} = 0$, entonces $\mathbf{u} = \mathbf{0}$ o bien $\mathbf{v} = \mathbf{0}$?
21. ¿Es verdad que si $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{w}$ y $\mathbf{u} \neq \mathbf{0}$, entonces $\mathbf{v} = \mathbf{w}$?
22. **(Ley del paralelogramo)** Demuestre la identidad siguiente (figura 2.16)

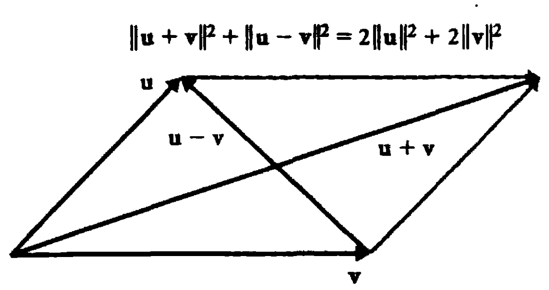


Figura 2.16 La ley del paralelogramo en el plano.

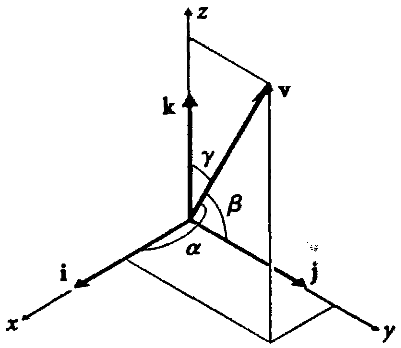


Figura 2.17 Ángulos directores.

23. **(Identidad de polarización)** Encuentre la identidad que expresa el producto punto en función de la norma.

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$$

24. Compruebe que si **u** es ortogonal a **v** y a **w**, entonces éste es ortogonal a cualquier combinación lineal $c\mathbf{v} + d\mathbf{w}$.

25. Describa geométricamente el conjunto **u** de todos los vectores 2 tales que $\|\mathbf{u}\| = 1$.

26. Exprese geométricamente el conjunto **v** de todos los vectores 3 tales que $\|\mathbf{v}\| = 1$.

Descomposición polar de un vector

Los ángulos α, β y γ que forma **v**, un vector 3 no cero y los vectores unitarios base, **i**, **j** y **k**, se llaman **ángulos directores** de **v** (figura 2.17). Los cosenos, $\cos\alpha$, $\cos\beta$ y $\cos\gamma$, de los ángulos directores, se llaman **cosenos directores** de **v**.

27. Compruebe que $(\cos\alpha, \cos\beta, \cos\gamma)$ es un vector unitario.

28. Demuestre que

$$\mathbf{v} = \|\mathbf{v}\|(\cos\alpha, \cos\beta, \cos\gamma) \qquad (2.10)$$

A esta ecuación se le llama **descomposición polar** de **u**.

29. Sea $\mathbf{v} = c\mathbf{u}$ para algún escalar $c > 0$, y algún vector unitario **u**. Demuestre que

$$c = \|\mathbf{v}\| \quad \text{y} \quad \mathbf{u} = (\cos\alpha, \cos\beta, \cos\gamma)$$

Para cada uno de los vectores dados en los ejercicios 30 y 31 determine lo siguiente:

(a) La descomposición polar
(b) Los cosenos directores
(c) Los ángulos directores (necesitará una calculadora)

30. $(1, 0, 1)$ $\quad(-1, 1, 0)$
$(1, -1, 1)$ $\quad(2, -1, 0)$

31. $(1, 1, 1)$ $\quad(-1, 1, -1)$
$(\sqrt{3}, 1, 0)$ $\quad(2, -2, 0)$

2.3 Espacio generado por un conjunto de vectores

Objetivo del alumno para esta sección
Comprender, algebraica y geométricamente, el espacio generado por un conjunto de vectores.

El material de esta sección es importante. Aprenderemos el *espacio generado* por una sucesión de vectores *n* y su interpretación geométrica en el caso de los vectores 2 y 3. Es una vasta generalización de las nociones de una línea o un plano que pasan por el origen.

DEFINICIÓN

(Espacio generado por un conjunto de vectores)

El conjunto de todas las combinaciones lineales de los vectores n $\mathbf{v}_1, \ldots, \mathbf{v}_k$, se llama **espacio generado por los vectores** $\mathbf{v}_1, \ldots, \mathbf{v}_k$, y se representa por Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Si $V =$ Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, se dice que $\mathbf{v}_1, \ldots, \mathbf{v}_k$ **generan a** V, y que $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un **conjunto generador** de V.

■ EJEMPLO 20 Demuestre que los siguientes vectores están en Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$.

$$\mathbf{0}, \quad \mathbf{v}_1, \quad \mathbf{v}_2, \quad \mathbf{v}_1 + \mathbf{v}_2, \quad 3\mathbf{v}_1, \quad 3\mathbf{v}_1 - 2.5\mathbf{v}_2$$

SOLUCIÓN Cada uno de esos vectores es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$, porque

$$\begin{aligned} \mathbf{0} &= 0\mathbf{v}_1 + 0\mathbf{v}_2 & \mathbf{v}_1 &= 1\mathbf{v}_1 + 0\mathbf{v}_2 & \mathbf{v}_2 &= 0\mathbf{v}_1 + 1\mathbf{v}_2 \\ \mathbf{v}_1 + \mathbf{v}_2 &= 1\mathbf{v}_1 + 1\mathbf{v}_2 & 3\mathbf{v}_1 &= 3\mathbf{v}_1 + 0\mathbf{v}_2 & 3\mathbf{v}_1 - 2.5\mathbf{v}_2 &= 3\mathbf{v}_1 + (-2.5)\mathbf{v}_2 \end{aligned}$$

■ EJEMPLO 21 Compruebe que Gen$\{\mathbf{v}\}$ es el conjunto de los múltiplos escalares de $\mathbf{v}$.

SOLUCIÓN Cualquier combinación lineal de $\mathbf{v}$ tiene la forma $c\mathbf{v}$ para algún escalar c. A la inversa, cualquier múltiplo escalar $c\mathbf{v}$ de $\mathbf{v}$ es una combinación lineal de $\mathbf{v}$. Por consiguiente, Gen$\{\mathbf{v}\}$ es el conjunto de los múltiplos escalares de $\mathbf{v}$.

■ EJEMPLO 22 ¿Está

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

en Gen $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \end{bmatrix} \right\}$?

SOLUCIÓN El vector está en el espacio generador si y sólo si hay escalares c_1 y c_2 tales que $\begin{bmatrix} 2 \\ 3 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 5 \end{bmatrix}$. Esto equivale a decir que el sistema con matriz aumentada $\begin{bmatrix} 1 & 3 & : & 2 \\ 2 & 5 & : & 3 \end{bmatrix}$, es consistente. Partiendo de la forma de escalón $\begin{bmatrix} 1 & 3 & : & 2 \\ 0 & -1 & : & -1 \end{bmatrix}$, vemos que la última columna *no es* columna pivote, de modo que el sistema es consistente, y el vector *sí está* en el generador.

Aunque no es necesario el sistema puede resolverse con facilidad, y se obtiene $c_1 = -1$, $c_2 = 1$. Entonces

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = -1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

■ EJEMPLO 23 Demuestre que Gen$\{\mathbf{e}_1, \mathbf{e}_2\} = \mathbf{R}^2$.

SOLUCIÓN Toda combinación lineal en $\mathbf{e}_1$, $\mathbf{e}_2$ es un vector de la forma $c_1\mathbf{e}_1 + c_2\mathbf{e}_2$ para los escalares c_1 y c_2. Como

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_2 = c_1\begin{bmatrix}1\\0\end{bmatrix} + c_2\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}c_1\\c_2\end{bmatrix}$$

se ve que $c_1\mathbf{e}_1 + c_2\mathbf{e}_2$ puede ser cualquier vector 2 y viceversa, cualquier vector 2 puede expresarse en forma de una combinación lineal $c_1\mathbf{e}_1 + c_2\mathbf{e}_2$. Por consiguiente, Gen$\{\mathbf{e}_1, \mathbf{e}_2\} = \mathbf{R}^2$. $\square$

■ EJEMPLO 24 Demuestre que Gen$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} = \mathbf{R}^n$.

SOLUCIÓN Esta comprobación se deja como ejercicio. $\square$

■ EJEMPLO 25 Determine Gen$\{\mathbf{e}_1, \mathbf{e}_3\}$ en $\mathbf{R}^3$.

SOLUCIÓN 1 Ya que

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_3 = c_1\begin{bmatrix}1\\0\\0\end{bmatrix} + c_2\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}c_1\\0\\c_2\end{bmatrix}$$

Gen$\{\mathbf{e}_1, \mathbf{e}_3\}$ es el conjunto de los vectores 3 cuya segunda componente es cero.

SOLUCIÓN 2 Necesitamos calcular todo $\mathbf{b}$ en $\mathbf{R}^3$, tal que $c_1\mathbf{e}_1 + c_2\mathbf{e}_3 = \mathbf{b}$ se cumpla para c_1 y c_2. También es posible determinar toda $\mathbf{b}$ de modo que el sistema $\begin{bmatrix}1 & 0 & : & b_1\\0 & 0 & : & b_3\\0 & 1 & : & b_2\end{bmatrix}$ sea consistente. Como esta matriz equivale a $\begin{bmatrix}1 & 0 & : & b_1\\0 & 1 & : & b_3\\0 & 0 & : & b_2\end{bmatrix}$, su última columna no es pivote si y sólo si $b_2 = 0$. En consecuencia, Gen$\{\mathbf{e}_1, \mathbf{e}_3\}$ incluye todos los vectores 3 cuyo segundo componente es cero. $\square$

Sean

$$\mathbf{v}_1 = \begin{bmatrix}1\\2\end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix}3\\5\end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix}2\\4\end{bmatrix}$$

■ EJEMPLO 26 Determine Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$.

SOLUCIÓN Será preciso encontrar todo $\mathbf{b} = \begin{bmatrix}b_1\\b_2\end{bmatrix}$ tal que el sistema $\begin{bmatrix}1 & 3 & : & b_1\\2 & 5 & : & b_2\end{bmatrix}$ sea consistente. Como $\begin{bmatrix}1 & 3\\2 & 5\end{bmatrix} \sim \begin{bmatrix}1 & 3\\0 & -1\end{bmatrix}$, la última columna de la *matriz aumentada* nunca es columna pivote. Así. el sistema es consistente para todo $\mathbf{b}$ en $\mathbf{R}^2$. Por tanto,

$$\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\} = \mathbf{R}^2$$ $\square$

■ EJEMPLO 27 Calcule Gen$\{\mathbf{v}_1, \mathbf{v}_3\}$.

SOLUCIÓN En vista de que

$$\begin{bmatrix} 1 & 2 & b_1 \\ 2 & 4 & b_2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & b_1 \\ 0 & 0 & b_2 - 2b_1 \end{bmatrix}$$

el sistema $\begin{bmatrix} 1 & 2 & : & b_1 \\ 2 & 4 & : & b_2 \end{bmatrix}$ es consistente si y sólo si $b_2 - 2b_1 = 0$, o si $b_2 = 2b_1$. Por consiguiente, **b** está en Gen$\{\mathbf{v}_1, \mathbf{v}_3\}$ si y sólo si $b_2 = 2b_1$, o si y sólo si **b** tiene la forma $r\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, para cualquier escalar r. En consecuencia, el generador consiste en todos los múltiplos escalares de $\mathbf{v}_1$.

$$\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\} = \{r\mathbf{v}_1, r \in \mathbf{R}\} = \text{Gen}\{\mathbf{v}_1\} \qquad \square$$

■ EJEMPLO 28 Determine un conjunto generador para

$$V = \left\{ \begin{bmatrix} 3a - b \\ a + 5b \\ a \end{bmatrix}, a, b \in \mathbf{R} \right\} \subseteq \mathbf{R}^3$$

SOLUCION Como $\begin{bmatrix} 3a - b \\ a + 5b \\ a \end{bmatrix} = a\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + b\begin{bmatrix} -1 \\ 5 \\ 0 \end{bmatrix}$ para todos los escalares a y b, V está generado por $\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 0 \end{bmatrix} \right\}$ $\square$

TEOREMA 9

(Reducción del conjunto generador)

Si uno de los vectores m $\mathbf{v}_1, \ldots, \mathbf{v}_k$ es una combinación lineal del resto, el generador permanece igual si se elimina ese vector.

DEMOSTRACIÓN Para comodidad en la notación podemos suponer que $\mathbf{v}_k$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ (si es necesario, renombrando los vectores). Entonces

$$\mathbf{v}_k = c_1\mathbf{v}_1 + \ldots + c_{k-1}\mathbf{v}_{k-1} \tag{2.11}$$

para algunos escalares $c_1, \ldots, c_{k-1}$. Sean V y V' los espacios generados correspondientes de $\mathbf{v}_1, \ldots, \mathbf{v}_k$ y de $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$, Es preciso comprobar que $V = V'$. Como cualquier combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_k$ (sumándole $0\mathbf{v}_k$), entonces $V' \subseteq V$. Basta demostrar que $V \subseteq V'$. Sea $\mathbf{u} \in V$. Entonces $\mathbf{u} = d_1\mathbf{v}_1 + \cdots + d_k\mathbf{v}_k$ para algunos escalares $d_1, \ldots, d_k$. Se tiene

$$\begin{aligned} \mathbf{u} &= d_1\mathbf{v}_1 + \cdots + d_{k-1}\mathbf{v}_{k-1} + d_k(c_1\mathbf{v}_1 + \cdots + c_{k-1}\mathbf{v}_{k-1}) \\ &= (d_1 + d_kc_1)\mathbf{v}_1 + \cdots + (d_{k-1} + d_kc_{k-1})\mathbf{v}_{k-1} \end{aligned}$$

que es una combinación lineal en $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$. Por consiguiente, $\mathbf{u} \in V'$, y entonces $V \subseteq V'$, como se afirmaba. $\square$

■ EJEMPLO 29 Compruebe que

$$\text{Gen}\left\{\begin{bmatrix}1\\-3\end{bmatrix},\begin{bmatrix}0\\2\end{bmatrix}\right\}=\text{Gen}\left\{\begin{bmatrix}1\\-3\end{bmatrix},\begin{bmatrix}0\\2\end{bmatrix},\begin{bmatrix}10\\-28\end{bmatrix}\right\}$$

SOLUCIÓN Puesto que $\begin{bmatrix}10\\-28\end{bmatrix}=10\begin{bmatrix}1\\-3\end{bmatrix}+\begin{bmatrix}0\\2\end{bmatrix}$, los dos generadores son iguales, de acuerdo con el teorema 9. $\square$

Quizá las dos propiedades más importantes del generador de un conjunto de vectores sean las que se describen en el teorema siguiente.

TEOREMA 10 Si $V = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Entonces, para cualesquiera $\mathbf{u}$ y $\mathbf{w}$ en V y cualquier escalar c:

1. $\mathbf{u} + \mathbf{w}$ está en V;
2. $c\mathbf{u}$ está en V.

DEMOSTRACIÓN Se deja como ejercicio. $\square$

El teorema 10 implica que cualquier combinación lineal de elementos en V está en V. Esto significa que el generador no sólo es el conjunto de combinaciones lineales de vectores, sino también que éste incluye cualquier combinación lineal de sus propios vectores.

Relación entre el generador y los sistemas lineales

Decir que un vector $\mathbf{b} \in \mathbf{R}^m$ es una combinación lineal de los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ equivale a decir que $\mathbf{b}$ está en el $\text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Por consiguiente,

1. El sistema cuya matriz aumentada es $[A : \mathbf{b}]$ es consistente si y sólo si $\mathbf{b}$ está en el $\text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.
2. $[A : \mathbf{b}]$ es consistente para todos los vectores $\mathbf{b} \in \mathbf{R}^m$ si y sólo si el $\text{Gen}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} = \mathbf{R}^m$. Esto es válido sólo si cada renglón de A tiene una posición pivote (teorema 2, sección 2.1).

■ EJEMPLO 30 ¿Es

$$\text{Gen}\left\{\begin{bmatrix}1\\-3\end{bmatrix},\begin{bmatrix}0\\2\end{bmatrix}\right\}=\mathbf{R}^2?$$

¿Qué puede decir acerca del sistema $\begin{bmatrix}1 & 0 & : & b_1\\-3 & 2 & : & b_2\end{bmatrix}$?

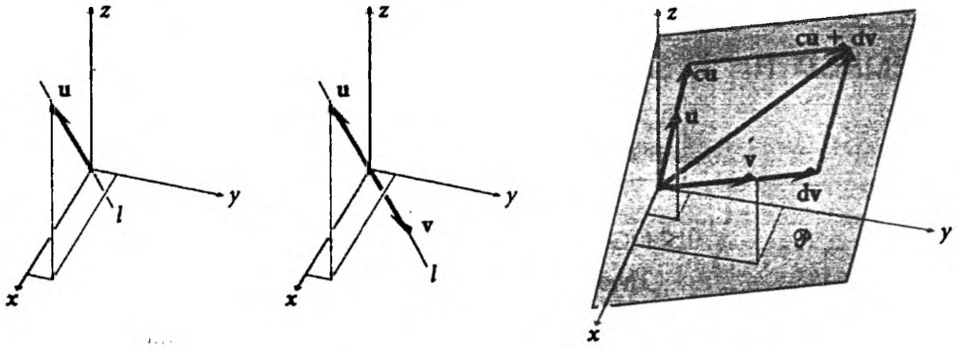
SOLUCIÓN Ya que cada renglón de $\begin{bmatrix}1 & 0\\-3 & 2\end{bmatrix}$ tiene un pivote (¿por qué?), $\begin{bmatrix}1 & 0 & : & b_1\\-3 & 2 & : & b_2\end{bmatrix}$ es consistente para todos los escalares b_1 y b_2, y $\text{Gen}\left\{\begin{bmatrix}1\\-3\end{bmatrix},\begin{bmatrix}0\\2\end{bmatrix}\right\}=\mathbf{R}^2$. $\square$

Interpretación geométrica de Gen{**u**} y de Gen{**u**, **v**}

Sean **u** y **v** vectores 2, o vectores 3.

Gen{u}, u ≠ 0

A medida que c asume cualquier valor escalar, c**u** pasa por todos los puntos de la recta que atraviesa el origen y tiene dirección **u**. Como, según el ejemplo 21, Gen{**u**} es el conjunto de todos los múltiplos escalares de **u**, llegamos a la conclusión de que, geométricamente, Gen{**u**} es la línea l que pasa por el origen y cuya dirección es **u**, figura 2.18(a). Por otro lado, cualquier línea recta que pasa por el origen puede escribirse como Gen{**u**}, si **u** es cualquier vector diferente de cero sobre la línea recta.



(a) Gen {u} = l (b) Gen {u, v} = l (c) Gen {u, v} = P

Figura 2.18 El generador de uno y dos vectores.

Gen{u, v}, u ≠ 0

Caso 1. Si **v** es un múltiplo escalar de **u**, entonces Gen{**u**, **v**} = Gen{**u**} = l, figura 2.18(b).
Caso 2. Si **v** no es un múltiplo escalar de **u**, entonces **v** ≠ **0**, y Gen{**u**, **v**} es el plano que pasa por el origen y contiene tanto a **u** como a **v**, figura 2.18(c).

De hecho, cualquier plano que pase por el origen puede representarse como Gen{**u**, **v**}, si **u** y **v** son dos vectores no colineales cualesquiera en ese plano.

Ejercicios 2.3

1. Sean $\mathbf{v}=\begin{bmatrix}9\\-3\end{bmatrix}$, $\mathbf{w}=\begin{bmatrix}-3\\1\end{bmatrix}$, y $S=\{\mathbf{w}\}$. ¿Es cierto o falso lo siguiente?

a. **v** está en S. b. **w** está en S.
c. **v** está en Gen(S). d. **w** está en Gen(S).

2. Si S es un conjunto finito de vectores n con al menos un vector distinto de cero. Explique por qué Gen(S) tiene una cantidad *infinita* de vectores.

Para los ejercicios 3 y 4, sean

$$\mathbf{a}=\begin{bmatrix}-1\\3\end{bmatrix}\quad \mathbf{b}=\begin{bmatrix}0\\-2\end{bmatrix}$$
$$\mathbf{c}=\begin{bmatrix}6\\3\end{bmatrix}\quad \mathbf{d}=\begin{bmatrix}-3\\9\end{bmatrix}$$

3. a. ¿**c** está en el Gen{**a**, **b**}? b. ¿**b** está en el Gen{**a**, **c**}?
c. ¿**a** está en el Gen{**b**, **c**}? d. ¿**d** está en el Gen{**a**}?

e. ¿a está en el Gen{**d**}? f. ¿**d** está en el Gen{**c**}?

g. ¿**d** está en el Gen{**b**, **c**}?

4. a. ¿Es cierto que el Gen{**a**, **b**} = $\mathbf{R}^2$?

   b. ¿Es verdad que el Gen{**a**, **d**} = $\mathbf{R}^2$?

   c. ¿Es cierto que el Gen{**a**, **c**} = Gen{**b**, **c**}?

   d. Compare el Gen{**a**} con el Gen{**d**}.

   e. ¿Cuál es el Gen{**a**, **b**, **c**}?

   f. ¿Cuál es el Gen{**a**, **c**, **d**}?

   g. ¿Cuál es el Gen{**a**, **b**, **c**, **d**}?

   h. ¿Cuál es el espacio generado de las columnas de $A = [\mathbf{a}\ \ \mathbf{a}\ \ \mathbf{b}\ \ \mathbf{d}]$?

   i. ¿Cuál es el espacio generado de las columnas de $B = [\mathbf{a}\ \ \mathbf{a}\ \ \mathbf{d}]$?

Para los ejercicios 5 y 6, sean

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$$

$$\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{v}_4 = \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix}$$

5. a. ¿$\mathbf{v}_4$ está en el Gen{$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$}?

   b. ¿$\mathbf{v}_4$ está en el Gen{$\mathbf{v}_1, \mathbf{v}_2$}?

   c. ¿$\mathbf{v}_4$ está en el Gen{$\mathbf{v}_2, \mathbf{v}_3$}?

6. a. Demuestre que el Gen{$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$} = $\mathbf{R}^3$.

   b. Compruebe que el Gen{$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4$} = $\mathbf{R}^3$.

   c. Demuestre que el Gen{$\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_4$} = el Gen{$\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$}.

En los ejercicios 7 al 13 determine si las columnas de la matriz dadas de $m \times n$ generan a $\mathbf{R}^m$.

7. a. $\begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix}$ b. $\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}$

8. a. $\begin{bmatrix} a & 2a \\ b & 2b \end{bmatrix}$ b. $\begin{bmatrix} a & b \\ 2a & 2b \end{bmatrix}$

9. a. $\begin{bmatrix} 1 & -1 & 0 \\ 7 & 3 & -1 \end{bmatrix}$ b. $\begin{bmatrix} 0 & 5 \\ 1 & -1 \\ 1 & 3 \end{bmatrix}$

10. a. $\begin{bmatrix} a & b & 0 \\ a & b & 0 \end{bmatrix}$ b. $\begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}$

11. a. $\begin{bmatrix} 1 & 2 & -3 & 2 & -5 & 5 \\ 0 & 0 & 0 & 2 & -5 & 1 \\ 0 & 0 & 0 & -2 & 5 & -1 \end{bmatrix}$

    b. $\begin{bmatrix} a & e & i \\ b & f & j \\ c & g & k \\ d & h & l \end{bmatrix}$

12. a. $\begin{bmatrix} 2 & 0 & 5 \\ 0 & -2 & 0 \\ -1 & 1 & 2 \end{bmatrix}$ b. $\begin{bmatrix} a & a & 0 \\ a & a & 0 \\ 1 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 2 & -3 & 2 & -5 & 5 \\ 0 & 0 & 0 & 2 & -5 & 1 \\ 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$

14. ¿Bajo qué restricciones de a y b, las columnas de $\begin{bmatrix} a & b & 0 \\ a & 0 & 0 \\ a & 0 & 1 \end{bmatrix}$ generan a $\mathbf{R}^3$?

15. ¿Cierto o falso?

    a. $\mathbf{R}^{10}$ puede generarse exactamente con 9 vectores 10.

    b. $\mathbf{R}^{10}$ puede generarse con al menos 9 vectores 10.

    c. $\mathbf{R}^{10}$ puede generarse con 10 vectores 9.

    d. $\mathbf{R}^{10}$ puede generarse con 10 vectores 10.

    e. $\mathbf{R}^{10}$ se puede generar con 11 vectores 10.

    f. 20 vectores 10 cualesquiera generan a $\mathbf{R}^{10}$.

    g. 20 vectores 10 pueden generar a $\mathbf{R}^{10}$.

16. Compruebe que

$$\text{Gen}\{\mathbf{u}, \mathbf{v}\} = \text{Gen}\{\mathbf{u} + \mathbf{v}, \mathbf{u} - \mathbf{v}\}$$

17. Confirme que

$$\text{Gen}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \text{Gen}\{\mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} + \mathbf{w}\}$$

18. Muestre que $S_1 = S_2$, siendo

$$S_1 = \text{Gen}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

$$S_2 = \text{Gen}\left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \end{bmatrix} \right\}$$

19. Demuestre el teorema 10.

20. Si A es una matriz cuyas columnas generan a $\mathbf{R}^{10}$. ¿Qué puede decir acerca de lo siguiente?

    a. El tamaño de A.

    b. El sistema $[A : \mathbf{b}]$.

21. Sea A una matriz de 10×9.
   a. ¿Las columnas de A pueden generar a $\mathbf{R}^{10}$?
   b. ¿Es cierto que hay un vector 10 **b**, en el que el sistema $[A : \mathbf{b}]$ es inconsistente?

22. Encuentre un conjunto generador finito para

$$V = \left\{ \begin{bmatrix} a-b \\ 2a+4b \end{bmatrix}, a, b \in \mathbf{R} \right\}$$

23. Determine un conjunto generador finito para

$$V = \left\{ \begin{bmatrix} 3a-b \\ 4b \\ -a \end{bmatrix}, a, b \in \mathbf{R} \right\}$$

24. Suponga que cada renglón de la matriz A, de $m \times n$, tiene una posición pivote. ¿Qué puede decir acerca del sistema $[A : \mathbf{b}]$?

25. Si **b** está en el generador de las columnas de A, ¿qué puede decir acerca del sistema $[A : \mathbf{b}]$?

26. Encuentre todos los valores de x tales que

$$\text{Gen}\left\{ \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} -10 \\ x \end{bmatrix} \right\} = \mathbf{R}^2$$

27. Determine todos los valores de x tales que

$$\text{Gen}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ x \end{bmatrix} \right\} = \mathbf{R}^3$$

28. Trace los siguientes conjuntos.
   a. $\text{Gen}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \end{bmatrix} \right\}$ b. $\text{Gen}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \end{bmatrix} \right\}$

29. Dibuje los siguientes conjuntos.
   a. $\text{Gen}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ b. $\text{Gen}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$

2.4 Independencia lineal

Objetivos del alumno para esta sección

1. Conocer qué son los vectores linealmente dependientes y linealmente independientes.
2. Comprender la interpretación geométrica de la independencia lineal.

El material en esta sección es muy importante. Aprenderemos lo que son los vectores n *linealmente dependientes* y *linealmente independientes*. Además, describiremos la interpretación geométrica de esas nociones para el caso de los vectores 2 y 3.

Una combinación lineal de vectores n $\mathbf{v}_1, \ldots, \mathbf{v}_k$ es

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

y se llama **no trivial** cuando algunas de las $c_1, \ldots, c_k$ son diferentes de cero. Pero si todos los coeficientes son cero se denomina **trivial**.

DEFINICIONES

Una *sucesión* de vectores n, $\mathbf{v}_1, \ldots, \mathbf{v}_k$, es **linealmente dependiente** si **0** es una combinación lineal no trivial de esos vectores. En otras palabras, cuando hay escalares $c_1, \ldots, c_k$ y no todos son cero, como

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0} \tag{2.12}$$

Un *conjunto* $\mathbf{v}_1, \ldots, \mathbf{v}_k$ de vectores n es linealmente dependiente si lo es como sucesión.[4] Una relación de la forma (2.12) cuando no todas las $c_1, \ldots, c_k$ son cero, se llama **relación de dependencia lineal**.

[4] Una vez más empleamos el término *sucesión* cuando se permiten repeticiones de vectores.

Un *conjunto* $\mathbf{v}_1, \ldots, \mathbf{v}_k$ de vectores n es **linealmente independiente** si no es linealmente dependiente. Es lo mismo que decir que la única combinación lineal de $\mathbf{0}$ en función de $\mathbf{v}_1, \ldots, \mathbf{v}_k$ es la trivial, o que la ecuación (2.12) implica que $c_1 = 0, \ldots, c_k = 0$.

■ EJEMPLO 31 El conjunto

$$\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$$

es linealmente dependiente, porque

$$1\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

□

■ EJEMPLO 32 La sucesión

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

es linealmente dependiente, porque

$$1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

□

■ EJEMPLO 33 ¿Es el conjunto

$$\left\{\begin{bmatrix} 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$$

linealmente dependiente? En caso afirmativo, determine una relación de dependencia.

SOLUCIÓN Sean c_1, c_2, c_3 tales que

$$c_1\begin{bmatrix} 0 \\ -2 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_3\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Por consiguiente,

$$\begin{aligned} c_2 + 2c_3 &= 0 \\ -2c_1 + 2c_2 + c_3 &= 0 \end{aligned}$$

Es necesario resolver este sistema lineal homogéneo para encontrar c_1, c_2 y c_3. Como la matriz aumentada $\begin{bmatrix} 0 & 1 & 2 & : & 0 \\ -2 & 2 & 1 & : & 0 \end{bmatrix}$ se reduce a $\begin{bmatrix} 1 & 0 & -\frac{3}{2} & : & 0 \\ 0 & 1 & 2 & : & 0 \end{bmatrix}$, obtenemos $c_1 = -\frac{3}{2}r$, $c_2 = -2r$ y $c_3 = r$. Así, hay soluciones no triviales y el conjunto es linealmente dependiente. Para conseguir una relación lineal particular de dependencia se asigna un valor al parámetro r. Por ejemplo, si $r = 2$,

$$-3\begin{bmatrix}0\\-2\end{bmatrix}-4\begin{bmatrix}1\\2\end{bmatrix}+2\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Ésta es una de las (infinitas) relaciones de dependencia lineal.

Obsérvese que si sólo tuviéramos que comprobar si hay dependencia lineal, bastaría una forma de escalón de la matriz aumentada. Por ejemplo, la forma de escalón $\begin{bmatrix}-2 & 2 & 1 & : & 0\\ 0 & 1 & 2 & : & 0\end{bmatrix}$ nos dice que el sistema *homogéneo* tiene variables libres. De aquí que haya soluciones no triviales y que el conjunto sea linealmente dependiente. □

TEOREMA 11

> Dados dos vectores n no cero, las siguientes afirmaciones son equivalentes:
>
> 1. Los vectores son linealmente dependientes.
> 2. Un vector es un múltiplo escalar del otro.
> 3. El ángulo que forman los vectores es 0 o π.

DEMOSTRACIÓN Se deja como ejercicio. □

■ EJEMPLO 34 Compruebe que $\{\mathbf{e}_1, \mathbf{e}_2\}$ es linealmente independiente en $\mathbf{R}^2$.

SOLUCIÓN Si c_1 y c_2 son escalares tales que $c_1\mathbf{e}_1 + c_2\mathbf{e}_2 = \mathbf{0}$. En otras palabras,

$$c_1\begin{bmatrix}1\\0\end{bmatrix}+c_2\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}.$$

Entonces $\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$. Por tanto, $c_1 = 0$ y $c_2 = 0$. En consecuencia, el conjunto es linealmente independiente. □

■ EJEMPLO 35 Demuestre que $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ en $\mathbf{R}^n$ es linealmente independiente.

SOLUCIÓN Se deja como ejercicio. □

■ EJEMPLO 36 Confirme que

$$\left\{\begin{bmatrix}2\\3\\2\end{bmatrix},\begin{bmatrix}8\\-6\\5\end{bmatrix},\begin{bmatrix}-4\\3\\1\end{bmatrix}\right\}$$

es linealmente independiente.

SOLUCIÓN Si $c_1\begin{bmatrix}2\\3\\2\end{bmatrix}+c_2\begin{bmatrix}8\\-6\\5\end{bmatrix}+c_3\begin{bmatrix}-4\\3\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$, entonces el sistema correspondiente tiene la matriz de coeficientes $\begin{bmatrix} 2 & 8 & -4 \\ 3 & -6 & 3 \\ 2 & 5 & 1 \end{bmatrix}$, que es equivalente a la forma de escalón $\begin{bmatrix} 2 & 8 & -4 \\ 0 & -3 & 5 \\ 0 & 0 & -21 \end{bmatrix}$. No hay variables libres, por lo que el sistema homogéneo sólo tiene la solución trivial, y el conjunto es linealmente independiente. $\square$

Como indicaron los ejemplos anteriores, la reducción por renglones es una forma efectiva de determinar la dependencia o independencia lineal.

TEOREMA 12

(Criterio de independencia lineal)

Las siguientes afirmaciones son equivalentes:

1. El conjunto de vectores m $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es linealmente independiente.
2. El sistema sólo tiene la solución trivial

$$[\,\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \cdots \;\; \mathbf{v}_n \;\; : \;\; \mathbf{0}\,]$$

3. La matriz tiene n posiciones pivote (es decir, cada columna es columna pivote):

$$[\,\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \cdots \;\; \mathbf{v}_n\,]$$

DEMOSTRACIÓN $1 \Leftrightarrow 2$, porque $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$, y $[\mathbf{v}_1 \cdots \mathbf{v}_n : \mathbf{0}]$ son expresiones equivalentes. Por otro lado, $2 \Leftrightarrow 3$, de acuerdo con el enunciado 2 del teorema 3, sección 1.2. $\square$

■ EJEMPLO 37 ¿Son linealmente independientes las columnas de A?

$$A = \begin{bmatrix} -3 & 3 & 3 \\ 2 & 2 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

SOLUCIÓN A tiene tres columnas pivote, porque $A \sim \begin{bmatrix} -3 & 3 & 3 \\ 0 & 4 & 4 \\ 0 & 0 & -1 \end{bmatrix}$. Por lo que sus columnas son linealmente independientes, según el teorema 12. $\square$

■ EJEMPLO 38 ¿Son linealmente independientes las columnas de A?

$$A = \begin{bmatrix} -5 & 5 & 5 \\ 9 & 0 & 9 \\ 4 & 6 & 16 \end{bmatrix}$$

SOLUCIÓN Ya que $A \sim \begin{bmatrix} -5 & 5 & 5 \\ 0 & 9 & 18 \\ 0 & 0 & 0 \end{bmatrix}$, la tercera columna de A no es una columna pivote. Por tanto, siguiendo el teorema 12, las columnas no son linealmente independientes. (Éstas son linealmente dependientes.) ☐

Una consecuencia útil del teorema 12 es el teorema 13.

TEOREMA 13

Si el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de vectores m es linealmente independiente, entonces $n \le m$.

DEMOSTRACIÓN Como el conjunto es independiente, la matriz $[\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \cdots \ \ \mathbf{v}_n]$ tiene n columnas pivote, según el teorema 12. Pero el número de pivotes no puede ser mayor que el número de columnas o que el número de renglones de una matriz, de acuerdo con el ejercicio 49, sección 1.2. En consecuencia, $n \le m$, como se afirmaba. ☐

El teorema 13 dice que si hay *más vectores que componentes, los vectores son linealmente dependientes.*

■ EJEMPLO 39 ¿Son linealmente dependientes las columnas de la matriz siguiente?

$$\begin{bmatrix} 45 & -80 & -93 & 92 \\ 43 & -62 & 77 & 66 \\ 54 & 55 & -99 & -61 \end{bmatrix}$$

SOLUCIÓN Sí; según con el teorema 13, cuatro vectores 3 no pueden ser linealmente independientes. ☐

El siguiente teorema es muy útil para demostrar que una sucesión o conjunto de vectores es linealmente dependiente.

TEOREMA 14

(Prueba de dependencia lineal)

Sea S un conjunto o sucesión finitos de vectores m. Entonces

1. Si S tiene un vector $\mathbf{v}$, entonces es linealmente dependiente si y sólo si $\mathbf{v} = \mathbf{0}$.
2. Si S consiste en dos o más vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$, entonces S es linealmente dependiente si y sólo si cuando menos un vector sea una combinación lineal de los demás vectores.
3. Si S incluye dos o más vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$ con $\mathbf{v}_1 \neq \mathbf{0}$, entonces S es linealmente dependiente si y sólo si al menos un vector, por ejemplo $\mathbf{v}_i$ ($i \ge 2$), es una combinación lineal de los vectores que le anteceden, es decir, de $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$.

DEMOSTRACIÓN Sólo comprobaremos la parte 3 y dejaremos las restantes como ejercicios. Si $\mathbf{v}_i$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$, entonces hay escalares $c_1, \ldots, c_{i-1}$ tales que

$$\mathbf{v}_i = c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1}$$

Por consiguiente,

$$c_1\mathbf{v}_1 + \ldots + c_{i-1}\mathbf{v}_{i-1} + (-1)\mathbf{v}_i + 0\mathbf{v}_{i+1} + \ldots + 0\mathbf{v}_k = \mathbf{0}$$

Ésta es una combinación lineal no trivial de $\mathbf{0}$ en función de $\mathbf{v}_1, \ldots, \mathbf{v}_k$, porque el coeficiente de $\mathbf{v}_i$ es $-1 \neq 0$. Así, $\mathbf{v}_1, \ldots, \mathbf{v}_k$ es linealmente dependiente.

Por el contrario, cuando $\mathbf{v}_1, \ldots, \mathbf{v}_k$ son linealmente dependientes, entonces hay escalares $c_1, \ldots, c_k$ no todos son cero, tales que

$$c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$$

Sea c_i el último escalar distinto de cero en la ecuación anterior. Entonces

$$c_1\mathbf{v}_1 + \ldots + c_i\mathbf{v}_i + 0\mathbf{v}_i + \ldots + 0\mathbf{v}_k = \mathbf{0}, \qquad \text{o sea} \qquad c_1\mathbf{v}_1 + \ldots + c_i\mathbf{v}_i = \mathbf{0}$$

siendo $c_i \neq 0$. Vemos que $i \geq 2$, porque si $i = 1$, la combinación no trivial $c_1\mathbf{v}_1$ o implicaría que $c_1\mathbf{v}_1 = \mathbf{0}$. Por lo que

$$c_i\mathbf{v}_i = (-c_1)\,\mathbf{v}_1 + \ldots + (-c_{i-1})\,\mathbf{v}_{i-1} \qquad \text{o sea} \qquad \mathbf{v}_i = \left(-\frac{c_1}{c_i}\right)\mathbf{v}_1 + \ldots + \left(-\frac{c_{i-1}}{c_i}\right)\mathbf{v}_{i-1}$$

En consecuencia, $\mathbf{v}_i$ es una combinación de $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$. $\square$

■ EJEMPLO 40 Los vectores

$$\begin{bmatrix}1\\2\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\5\\0\\0\end{bmatrix}, \begin{bmatrix}35\\67\\33\\88\end{bmatrix}$$

son linealmente dependientes, de acuerdo con el teorema 14, puesto que $\begin{bmatrix}1\\5\\0\\0\end{bmatrix} = 1\begin{bmatrix}1\\2\\0\\0\end{bmatrix} + 3\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$ $\square$

OBSERVACIÓN El teorema 14 *no* dice que *todo* vector sea una combinación lineal de los restantes (o de los anteriores). Por ejemplo $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$ es linealmente dependiente, porque $\begin{bmatrix}2\\0\end{bmatrix} = 2\begin{bmatrix}1\\0\end{bmatrix}$. Sin embargo, $\begin{bmatrix}0\\1\end{bmatrix}$ no es una combinación lineal de $\begin{bmatrix}1\\0\end{bmatrix}$ y de $\begin{bmatrix}2\\0\end{bmatrix}$.

Ahora reuniremos algunos hechos básicos, cuyas demostraciones se dejan como ejercicios.

TEOREMA 15

1. Cualquier conjunto o sucesión finita de vectores que contiene $\mathbf{0}$ es linealmente dependiente.
2. Una sucesión finita con vectores repetidos es linealmente dependiente.
3. Cualquier conjunto finito (sucesión) de vectores que contiene un conjunto o sucesión linealmente dependiente (sucesión) es en sí linealmente dependiente.
4. Cualquier subconjunto de un conjunto finito linealmente independiente es en sí linealmente independiente.

También contamos con el siguiente teorema.

TEOREMA 16

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente de vectores m.

1. Cualquier vector $\mathbf{v}$ en el generador de S es expresable en forma única como combinación lineal de vectores en S. Esto es, las relaciones
$$\mathbf{v} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n \quad \text{y} \quad \mathbf{v} = d_1\mathbf{v}_1 + \ldots + d_n\mathbf{v}_n$$
implican que
$$c_1 = d_1, \ldots, c_n = d_n$$
2. Si $\mathbf{v}$ no está en el generador de S, el conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}\}$ es linealmente independiente.

DEMOSTRACIÓN

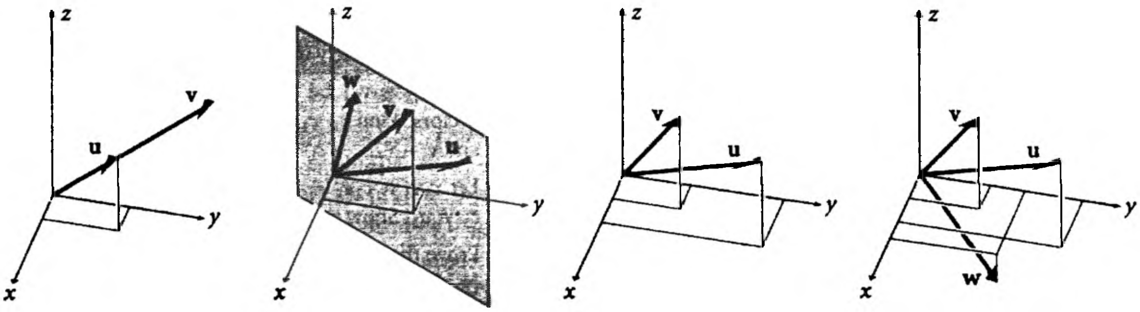
1. Como $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n$, entonces
$$(c_1 - d_1)\mathbf{v}_1 + \cdots + (c_n - d_n)\mathbf{v}_n = \mathbf{0}$$
De manera que $c_1 - d_1 = \cdots = c_n - d_n = 0$, porque S es linealmente independiente, en vista de lo cual $c_1 = d_1, \ldots, c_n = d_n$ comprueba lo que se afirmaba.

2. Supongamos que, por el contrario, $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}\}$ es linealmente dependiente. En ese caso hay una combinación lineal no trivial de $\mathbf{0}$,
$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n + c\mathbf{v} = \mathbf{0}$$
Si $c \neq 0$, podemos despejar $\mathbf{v}$ de la ecuación, y obtenemos $\mathbf{v} = (c^{-1}c_1)\mathbf{v}_1 + \cdots + (c^{-1}c_n)\mathbf{v}_n$. Pero entonces $\mathbf{v}$ sería una combinación lineal en S, y estaría en el generador de S, lo cual contradice nuestra hipótesis. Si, por otro lado, $c = 0$, la ecuación se reduce a
$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$
con al menos una $c_i \neq 0$. (¿Por qué?) Esto también contradice la hipótesis de que S es linealmente independiente. Entonces, llegamos a la conclusión que S' debe ser linealmente independiente.

Interpretación geométrica de la independencia lineal en $\mathbf{R}^2$ y $\mathbf{R}^3$

Al aplicar el teorema 15 y las definiciones geométricas de suma y multiplicación por escalar que describimos en la sección 2.1, podemos hacer las siguientes observaciones (figura 2.19):

1. Dos vectores no cero en $\mathbf{R}^2$ o en $\mathbf{R}^3$ son linealmente dependientes si y sólo si están en la misma línea que pasa por el origen.
2. Tres vectores en $\mathbf{R}^3$ son linealmente dependientes si y sólo si están en el mismo plano que pasa por el origen.
3. El generador de dos vectores en $\mathbf{R}^2$ o en $\mathbf{R}^3$ es una recta que pasa por el origen si los vectores son linealmente dependientes, o es el plano definido por éstos si son linealmente independientes.
4. El generador de dos vectores 2 linealmente independientes es $\mathbf{R}^2$.



Linealmente dependiente

Linealmente independiente

Figura 2.19

5. El generador de tres vectores 3 linealmente independientes es $\mathbf{R}^3$.
6. Cualquier conjunto linealmente independiente en $\mathbf{R}^3$ tiene tres vectores, cuando mucho.
7. Cualquier conjunto linealmente independiente en $\mathbf{R}^2$ tiene dos vectores, como máximo.
8. Cualquier conjunto que genere a $\mathbf{R}^2$ posee al menos dos vectores.
9. Cualquier conjunto que genere a $\mathbf{R}^3$ tiene tres vectores como mínimo.

Ejercicios 2.4

En los ejercicios 1 a 5 determine si los vectores son linealmente independientes.

1. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} a \\ 1 \end{bmatrix}, \begin{bmatrix} 10a \\ 100 \end{bmatrix}$, para $a \neq 0$

4. $\begin{bmatrix} a \\ a \\ 1 \end{bmatrix}, \begin{bmatrix} b \\ b \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

5. $\begin{bmatrix} a \\ a \\ 1 \end{bmatrix}, \begin{bmatrix} b \\ b \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, para $a \neq b$

En los ejercicios 6 a 9, determine si las columnas de la matriz son linealmente independientes.

6. $\begin{bmatrix} -1 & 2 & 1 \\ 3 & -3 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

7. $\begin{bmatrix} 15 & 0 \\ 20 & 0 \\ 25 & 0 \end{bmatrix}$

8. $\begin{bmatrix} -3 & 1 & 1 & 0 \\ 2 & 1 & 0 & -1 \\ 4 & 1 & -1 & 1 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Observe los ejercicios 10 a 18 y diga si los vectores son linealmente independientes.

10. $\begin{bmatrix} 555 \\ 123 \end{bmatrix}, \begin{bmatrix} 55{,}500 \\ 12{,}300 \end{bmatrix}$

11. $\begin{bmatrix} 555 \\ 123 \end{bmatrix}, \begin{bmatrix} 334 \\ 654 \end{bmatrix}, \begin{bmatrix} 446 \\ 667 \end{bmatrix}$

12. $\begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}$

13. $\begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 20 \\ 10 \\ 30 \\ -10 \end{bmatrix}$

14. $\begin{bmatrix} a \\ a \end{bmatrix}, \begin{bmatrix} b \\ b \end{bmatrix}$

15. $\begin{bmatrix} a \\ a \end{bmatrix}, \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} 0 \\ a-b \end{bmatrix}$

16. $\begin{bmatrix} a \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} a \\ 0 \\ 1 \end{bmatrix}$

17. $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} a \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} a \\ 0 \\ 1 \end{bmatrix}$

18. $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} a \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} a \\ a \\ 0 \end{bmatrix}$

19. ¿Para qué valores de a es linealmente dependiente el conjunto $\left\{ \begin{bmatrix} a \\ 1 \end{bmatrix}, \begin{bmatrix} a+2 \\ a \end{bmatrix} \right\}$?

20. ¿Para qué valores de a el conjunto $\left\{ \begin{bmatrix} a \\ 2 \end{bmatrix}, \begin{bmatrix} a-2 \\ a \end{bmatrix} \right\}$ es linealmente dependiente?

21. ¿Es cierto que el conjunto cuyos elementos son $\begin{bmatrix} a \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} a+2 \\ 2 \end{bmatrix}$ debe ser linealmente dependiente?

22. Las afirmaciones siguientes ¿son ciertas o falsas?
    a. Cualesquiera dos vectores n distintos son linealmente independientes.
    b. Cualesquiera tres vectores 2 diferentes generan $\mathbf{R}^2$.
    c. Cualesquiera dos vectores-2 linealmente independientes generan $\mathbf{R}^2$.
    d. Cualesquiera n vectores-n, linealmente independientes, generan $\mathbf{R}^n$.

23. Si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto linealmente independiente de vectores-n, determine c_1, c_2 y c_3 para que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ y $\mathbf{v} = (2c_2 - c_1)\mathbf{v}_1 + (c_3 - c_2)\mathbf{v}_2 + (c_2 - 1)\mathbf{v}_3$.

24. Sean $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + d_3\mathbf{v}_3$ y $c_3 \neq d_3$. Compruebe que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente.

25. Demuestre que si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente, también lo es $\{\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_2 - \mathbf{v}_3, \mathbf{v}_3 + \mathbf{v}_1\}$

26. Verifique que son ciertas las afirmaciones del ejemplo 35.

27. Demuestre el teorema 11.

28. Compruebe el teorema 15.

29. Suponga que las columnas de la matriz A, de $m \times n$, son linealmente independientes. Demuestre que para cualquier vector-m $\mathbf{b}$, el sistema $[A : \mathbf{b}]$ tiene cuando mucho una solución.

30. Sea A una matriz de $n \times n$ con columnas linealmente independientes. Verifique que para cualquier vector-n $\mathbf{b}$, el sistema $[A : \mathbf{b}]$ tiene solamente una solución.

31. Suponga que $S_1 = \{\mathbf{v}_1, \mathbf{v}_2\}$ y que $S_2 = \{\mathbf{w}_1, \mathbf{w}_2\}$ son subconjuntos linealmente independientes de $\mathbf{R}^3$. ¿Qué objeto geométrico es la intersección de $\mathrm{Gen}(S_1) \cap \mathrm{Gen}(S_2)$?

2.5 El producto $A\mathbf{x}$

Objetivos del alumno para esta sección

1. Aprender a calcular e interpretar la matriz producto $A\mathbf{x}$.
2. Saber cómo escribir un sistema lineal en función de $A\mathbf{x}$.

En esta sección definiremos el producto $A\mathbf{x}$ de una matriz por un vector y lo usaremos para presentar otra notación para sistemas lineales. A continuación estudiaremos el conjunto solución, o conjunto de soluciones, de un sistema homogéneo y sus propiedades.

El producto $A\mathbf{x}$

Definiremos el producto de una matriz por un vector. Si $A = \begin{bmatrix} -2 & 5 & -3 \\ 4 & 7 & 0 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} -3 \\ 2 \\ 5 \end{bmatrix}$, entonces el producto $A\mathbf{x}$ es la combinación lineal

$$\begin{bmatrix} -2 & 5 & -3 \\ 4 & 7 & 0 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \\ 5 \end{bmatrix} = -3\begin{bmatrix} -2 \\ 4 \end{bmatrix} + 2\begin{bmatrix} 5 \\ 7 \end{bmatrix} + 5\begin{bmatrix} -3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

En otras palabras, se forma la combinación lineal de las columnas de A y los coeficientes de los componentes de $\mathbf{x}$. Observe que esta operación tiene sentido *sólo si la cantidad de columnas de la matriz es igual al tamaño del vector.*

DEFINICIÓN

Sea A una matriz de $m \times n$, cuyas columnas son $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, y sea $\mathbf{x}$ un vector n cuyos componentes son $x_1, x_2, \ldots, x_n$. El producto matricial $A\mathbf{x}$ es el vector m expresado como la combinación lineal

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

■ EJEMPLO 41 Sean

$$A = \begin{bmatrix} 2 & -2 \\ -1 & 6 \\ 4 & -3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}, \quad \text{y} \quad \mathbf{c} = \begin{bmatrix} 10 \\ 20 \\ 30 \end{bmatrix}$$

Calcule, si es posible, los productos $A\mathbf{b}$ y $A\mathbf{c}$.

SOLUCIÓN

$$A\mathbf{b} = \begin{bmatrix} 2 & -2 \\ -1 & 6 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \end{bmatrix} = -3 \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} + 2 \begin{bmatrix} -2 \\ 6 \\ -3 \end{bmatrix} = \begin{bmatrix} -10 \\ 15 \\ -18 \end{bmatrix}$$

El producto $A\mathbf{c}$ es indefinido, porque A es de 3×2 y $\mathbf{c}$ no es un vector 2.

■ EJEMPLO 42 Determine $A\mathbf{x}$, donde

$$A = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$$

SOLUCIÓN

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = 2 \cdot 2 + (-1) \cdot 3 + 4 \cdot 1 = 5$$

Vemos que esta vez $A\mathbf{x}$ es justamente el *producto punto* $\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$.PÁGINA 105>

Observe que a medida que $\mathbf{x}$ varía en $\mathbf{R}^n$, el producto $A\mathbf{x}$ varía en $\mathbf{R}^m$. Así, la multiplicación por A define una **correspondencia** entre $\mathbf{R}^n$ y $\mathbf{R}^m$.

■ EJEMPLO 43 (Registro de calificaciones) Una profesora de álgebra lineal emplea matrices para registrar las calificaciones de su grupo. Si M es la matriz definida por

$$\begin{array}{c} \textbf{Calificación} \\ A \\ B \\ C \\ D \end{array} \quad \begin{array}{c} \begin{array}{ccc} \textbf{Otoño 1999} & \textbf{Primavera 2000} & \textbf{Otoño 2000} \end{array} \\ \begin{bmatrix} 2 & 3 & 1 \\ 10 & 15 & 12 \\ 13 & 15 & 15 \\ 8 & 11 & 12 \end{bmatrix} \end{array}$$

Calcule e interprete el producto $M \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

SOLUCIÓN

$$\begin{bmatrix} 2 & 3 & 1 \\ 10 & 15 & 12 \\ 13 & 15 & 15 \\ 8 & 11 & 12 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 10 \\ 13 \\ 8 \end{bmatrix} + 1 \begin{bmatrix} 3 \\ 15 \\ 15 \\ 11 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 12 \\ 15 \\ 12 \end{bmatrix} = \begin{bmatrix} 6 \\ 37 \\ 43 \\ 31 \end{bmatrix}$$

El producto es el vector cuyas componentes son las sumas de las columnas de M. Se obtiene la cantidad total de las calificaciones A, B, etcétera. □

La matriz de $n \times n$ cuyas columnas son $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ se llama **matriz identidad**, y se representa con I_n o con I.

■ EJEMPLO 44

$$I_2 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_3 = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad I_4 = I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

□

El teorema siguiente describe las propiedades básicas de la operación $A\mathbf{x}$.

TEOREMA 17

Sea A una matriz de $m \times n$, sean $\mathbf{x}$ y $\mathbf{y}$ vectores n y c cualquier escalar. Entonces

1. $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$
2. $A(c\mathbf{x}) = c(A\mathbf{x})$
3. $I_n\mathbf{x} = \mathbf{x}$

DEMOSTRACIÓN Comprobaremos sólo la parte 1 y dejaremos lo demás como ejercicios. Si $\mathbf{a}_1, \ldots, \mathbf{a}_n$ son las columnas de A y $x_1, \ldots, x_n$ y $y_1, \ldots, y_n$ son las componentes de $\mathbf{x}$ y $\mathbf{y}$, respectivamente. Entonces

$$A(\mathbf{x}+\mathbf{y}) = A\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}\right) = A\begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

De acuerdo con la definición del producto, esto es igual a

$$\begin{aligned} &(x_1+y_1)\mathbf{a}_1 + (x_2+y_2)\mathbf{a}_2 + \cdots + (x_n+y_n)\mathbf{a}_n \\ &= x_1\mathbf{a}_1 + y_1\mathbf{a}_1 + x_2\mathbf{a}_2 + y_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n + y_n\mathbf{a}_n && \text{Según el enunciado 6 del teorema 1} \\ &= (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n) + (y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + \cdots + y_n\mathbf{a}_n) && \text{Según los enunciados 1 y 2 del teorema 17} \\ &= A\mathbf{x} + A\mathbf{y} \end{aligned}$$

La ecuación $A\mathbf{x} = \mathbf{b}$

El producto $A\mathbf{x}$ permite contar con un método muy refinado y útil para representar un sistema lineal. Por ejemplo, para el sistema

$$\begin{aligned} 2x - y &= 1 \\ -x + 4y &= 3 \end{aligned}$$

que podemos expresar en notación vectorial como

$$x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Esta ecuación es equivalente a

$$\begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

de acuerdo con la definición del producto de una matriz por un vector.

Un sistema lineal se puede expresar en la forma $A\mathbf{x} = \mathbf{b}$, donde A representa la matriz de coeficientes, $\mathbf{b}$ el vector de términos constantes y $\mathbf{x}$ el vector de las incógnitas. De este modo tenemos las expresiones equivalentes

$$[A : \mathbf{b}] \Leftrightarrow A\mathbf{x} = \mathbf{b} \Leftrightarrow x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

en donde $\mathbf{a}_1, \ldots, \mathbf{a}_n$ son las columnas de A.

Afirmar que el sistema $[A : \mathbf{b}]$ es consistente equivale a decir que hay un vector $\mathbf{x}$ mediante el cual $\mathbf{b}$ puede expresarse como el producto $A\mathbf{x}$. Por ello, las siguientes afirmaciones son equivalentes:

1. $[A : \mathbf{b}]$ es consistente.
2. Hay un vector $\mathbf{x}$ tal que $A\mathbf{x} = \mathbf{b}$.
3. $\mathbf{b}$ es una combinación lineal de las columnas de A.
4. $\mathbf{b}$ es el generador de las columnas de A.

Tomando en cuenta la equivalencia de los puntos 2 y 4,

$$\mathrm{Gen}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\} = \{A\mathbf{x}, \text{ para todo } \mathbf{x} \text{ en } \mathbf{R}^n\} \tag{2.12}$$

■ EJEMPLO 45 Determine todos los vectores posibles de la forma

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

cuando x y y tienen cualquier valor real.

SOLUCIÓN De acuerdo con (2.12), $\left\{ \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, x, y \in \mathbf{R} \right\} = \text{Gen} \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$. Ya que $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ tiene dos pivotes y dos renglones; el generador y, en consecuencia, el conjunto dado es $\mathbf{R}^2$. ☐

El espacio nulo

El **espacio nulo,** $\nu(A)$, o **nulidad** de una matriz A de $m \times n$, contiene todos los vectores n en $\mathbf{x}$, tales que $A\mathbf{x} = 0$. Es el conjunto de soluciones del sistema homogéneo $A\mathbf{x} = \mathbf{0}$.

$$\nu(A) = \{\mathbf{x} \text{ en } \mathbf{R}^n \text{ tal que } A\mathbf{x} = \mathbf{0}\}$$

Observe que $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial si y sólo si $\nu(A) = \{\mathbf{0}\}$. Calcular $\nu(A)$ equivale a determinar todas las soluciones de $A\mathbf{x} = \mathbf{0}$.

■ EJEMPLO 46 ¿Cuál de

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

está en el espacio nulo de $\begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}$?

SOLUCIÓN En vista de que

$$\begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

es evidente que $\mathbf{u}$ está en el espacio nulo y $\mathbf{v}$ no lo está. ☐

Una de las propiedades más importantes del espacio nulo se describe en el teorema siguiente.

TEOREMA 18

Sea A una matriz de $m \times n$. Para cualesquiera $\mathbf{x}_1$, $\mathbf{x}_2$ en $\nu(A)$ y cualquier escalar c, tenemos que

1. $\mathbf{x}_1 + \mathbf{x}_2$ está en $\nu(A)$;
2. $c\mathbf{x}_1$ está en $\nu(A)$.

DEMOSTRACIÓN Como $A\mathbf{x}_1 = \mathbf{0}$ y $A\mathbf{x}_2 = \mathbf{0}$, entonces

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

$$A(c\mathbf{x}_1) = c(A\mathbf{x}_1) = c\mathbf{0} = \mathbf{0}$$

de acuerdo con el teorema 17. En consecuencia, $\mathbf{x}_1 + \mathbf{x}_2$ y $c\mathbf{x}_1$ están en $\nu(A)$. $\square$

Las soluciones de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$

En este párrafo describiremos la relación entre los conjuntos solución de un sistema lineal y de su sistema homogéneo asociado. Demostraremos que si $\mathbf{x}_h$ es la solución general de $A\mathbf{x} = \mathbf{0}$ y si $\mathbf{p}$ es una solución particular de $A\mathbf{x} = \mathbf{b}$, entonces

$$\mathbf{x} = \mathbf{p} + \mathbf{x}_h$$

es la solución general de $A\mathbf{x} = \mathbf{b}$.

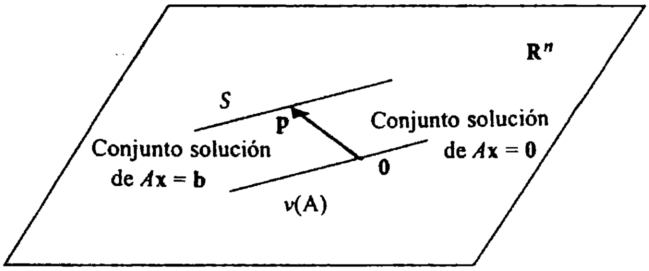
La relación entre los dos conjuntos solución puede describirse en forma refinada mediante la notación siguiente: si $\mathbf{p}$ es un vector n y S es un subconjunto de $\mathbf{R}^n$, representaremos con $\mathbf{p} + S$ al conjunto de todas las sumas $\mathbf{p} + \mathbf{x}$, donde $\mathbf{x}$ es un elemento de S.

$$\mathbf{p} + S = \{\mathbf{p} + \mathbf{x},\ \mathbf{x} \in S\}$$

TEOREMA 19

Si S es el conjunto de soluciones de $A\mathbf{x} = \mathbf{b}$, y $\mathbf{p}$ está en S. En ese caso

$$S = \mathbf{p} + \nu(A)$$



DEMOSTRACIÓN Si $\mathbf{p} + \mathbf{x}_h$ es cualquier elemento de $\mathbf{p} + \nu(A)$, entonces

$$A(\mathbf{x}_h + \mathbf{p}) = A\mathbf{x}_h + A\mathbf{p} = \mathbf{0} + \mathbf{b} = \mathbf{b}$$

De modo que $\mathbf{x}_h + \mathbf{p}$ es una solución de $A\mathbf{x} = \mathbf{b}$. Por tanto, $\mathbf{x}_h + \mathbf{p}$ está en S. Así, $\mathbf{p} + \nu(A) \subseteq S$.

Si $\mathbf{y}$ es cualquier elemento de S, entonces $A\mathbf{y} = \mathbf{b}$. La diferencia $\mathbf{y} - \mathbf{p}$ está en $\nu(A)$, porque

$$A(\mathbf{y} - \mathbf{p}) = A\mathbf{y} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Como consecuencia, $\mathbf{y} = \mathbf{p} + (\mathbf{y} - \mathbf{p})$ es un elemento de $\mathbf{p} + \nu(A)$, y así $S = \mathbf{p} + \nu(A)$, como se había dicho. $\square$

■ **EJEMPLO 47** Compruebe el teorema 19 para el sistema

$$\left[\begin{array}{rrcr} -2 & 4 & : & -6 \\ 1 & -2 & : & 3 \end{array}\right]$$

SOLUCIÓN El conjunto solución del sistema es $S = \left\{ \begin{bmatrix} 3+2r \\ r \end{bmatrix}, r \in \mathbf{R} \right\}$ y el espacio nulo de la matriz de coeficientes es $\nu(A) = \left\{ \begin{bmatrix} 2r \\ r \end{bmatrix}, r \in \mathbf{R} \right\}$. (Compruébelo.) Puede aplicar el teorema 19, porque $\begin{bmatrix} 3+2r \\ r \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 2r \\ r \end{bmatrix}$, y $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ es una solución particular del sistema.

El cálculo de $A\mathbf{x}$ con renglón-vector

A veces deseamos calcular sólo una componente, por ejemplo la i-ésima del producto $A\mathbf{x}$, sin tener que manejar toda la combinación lineal de la definición. Esto se hace fácilmente: tome el i-ésimo renglón de A y $\mathbf{x}$. Multiplique entre sí sus elementos correspondientes y sume todos los productos.

■ EJEMPLO 48 Calcule las componentes 1 y 3 del producto

$$\begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \\ 9 & 5 & 8 \\ 7 & 6 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$$

SOLUCIÓN

$$\begin{bmatrix} \boxed{2 \quad 0 \quad 1} \\ 2 \quad 1 \quad 2 \\ 9 \quad 5 \quad 8 \\ 7 \quad 6 \quad 4 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \qquad \begin{bmatrix} 2 \quad 0 \quad 1 \\ 2 \quad 1 \quad 2 \\ \boxed{9 \quad 5 \quad 8} \\ 7 \quad 6 \quad 4 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ 62 \\ \cdot \end{bmatrix}$$

$$2 \cdot 2 + 0 \cdot 4 + 1 \cdot 3 = 7 \qquad 9 \cdot 2 + 5 \cdot 4 + 8 \cdot 3 = 62$$

$A\mathbf{x}$ en términos del producto punto

Vemos que $A\mathbf{x}$ es el vector cuyas componentes son el *producto punto* de cada renglón de A por $\mathbf{x}$. Así, si $\mathbf{r}_1, \ldots, \mathbf{r}_m$ son los renglones de A, entonces

$$A\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{bmatrix}$$

$A\mathbf{x}$ con sistemas algebraicos computacionales

Maple

```
> with(linalg) :
> A:=matrix([[2,-3],[1,4]]):x:=vector([5,-6]);
> evalm(A&*x);
```

$$[28, -19]$$

Mathematica

```
In[1]:=
     A={{2,-3},{1,4}};x={{5},{-6}};
In[2]:= A . x
Out[2]=
     {{28}, {-19}}
```

MATLAB

```
>> A = [2 -3; 1 4]; x = [5; -6];
>> A*x
ans =
    28
   -19
```

Ejercicios 2.5

Sean

$$A = \begin{bmatrix} -3 & -2 \\ -1 & 0 \\ 5 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} -3 & 7 & -3 & -2 \\ 4 & 6 & -1 & 0 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 100 \\ 200 \\ 300 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

1. Calcule, si es posible, $A\mathbf{u}$, $A\mathbf{v}$ y $A\mathbf{w}$.
2. Determine, si es posible, $B\mathbf{u}$, $B\mathbf{v}$ y $B\mathbf{w}$.
3. Exprese $A\mathbf{x} = \mathbf{v}$ como un sistema lineal.
4. Convierta $B\mathbf{z} = \mathbf{u}$ a un sistema lineal.
5. Exprese el sistema siguiente en la forma $A\mathbf{x} = \mathbf{b}$.

$$\begin{aligned} x - 7y &= -5 \\ -2x + 4y &= 0 \end{aligned}$$

6. Escriba el sistema siguiente en la forma $A\mathbf{x} = \mathbf{b}$.

$$\begin{aligned} x - 4y &= -8 \\ -2y &= 7 \\ x + y &= 10 \end{aligned}$$

7. Encuentre todos los vectores n posibles que tengan la forma $\begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. ¿Cuál es el valor de n?

8. Determine todos los vectores n posibles que tengan la forma $\begin{bmatrix} -2 & 1 & -4 \\ 2 & -1 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$. ¿ Cuál es el valor de n?

9. Encuentre todos los vectores n posibles que tengan la forma $\begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 0 & -1 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$. ¿Cuál es el valor de n?

10. Demuestre las partes 2 y 3 del teorema 17.

11. Con referencia al ejemplo 43, calcule e interprete el producto siguiente.

$$\begin{bmatrix} 2 & 10 & 13 & 8 \\ 3 & 15 & 15 & 11 \\ 1 & 12 & 15 & 12 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

12. Una empresa de artículos deportivos vende bicicletas de los tipos 1, 2, 3 y 4 en tres tiendas. A continuación proporcionamos el orden de abastecimiento para las tiendas:

Tipo	Tienda 1	Tienda 2	Tienda 3
Bicicleta 1	25	15	35
Bicicleta 2	20	25	25
Bicicleta 3	15	35	20
Bicicleta 4	20	30	10

Si M es esta matriz, calcule e interprete los productos

$$M\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, M\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, M\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, M\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

13. Determine la primera y la última componentes del producto siguiente.

$$\begin{bmatrix} 2 & -2 & 2 \\ 3 & -3 & 3 \\ 4 & -4 & 4 \\ 5 & -5 & 5 \\ 6 & -6 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Sean $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

14. ¿Cuál(es) **u**, **v** y **w** están en el espacio nulo de

$$A = \begin{bmatrix} 2 & -2 & 2 \\ 0 & -3 & 3 \end{bmatrix}?$$

15. ¿Cuál(es) **u**, **v** y **w** se encuentran en el espacio nulo de

$$A = \begin{bmatrix} 2 & -2 & 2 \\ 1 & -4 & 4 \\ 0 & 1 & -1 \end{bmatrix}?$$

16. ¿Cuál(es) **u**, **v** y **w** están en el espacio nulo de

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ 3 & -3 & 3 \\ 4 & -4 & 5 \end{bmatrix}?$$

17. Confirme el teorema 19 para el sistema

$$\left[\begin{array}{rr:r} 2 & -6 & 20 \\ -1 & 3 & -10 \end{array}\right].$$

18. Compruebe el teorema 19 para el sistema

$$\left[\begin{array}{rrr:r} 1 & -1 & 2 & 0 \\ 2 & -1 & 6 & 1 \\ -2 & 4 & 0 & 2 \end{array}\right]$$

19. Demuestre que si $\mathbf{x}_1$ y $\mathbf{x}_2$ son soluciones del sistema $A\mathbf{x} = \mathbf{b}$, entonces $\mathbf{x}_2 - \mathbf{x}_1$ es una solución del sistema homogéneo asociado $A\mathbf{x} = \mathbf{0}$.

20. Si el conjunto de soluciones de $A\mathbf{x} = \mathbf{b}$ satisface lo afirmado por el teorema 18, ¿qué puede decirse acerca de **b**?

2.6 El producto cruz

Objetivos del alumno para esta sección

1. Calcular el producto cruz de dos vectores y comprender su geometría.
2. Comprender y aplicar las propiedades básicas del producto cruz.

En esta sección explicaremos el producto cruz. Aunque éste se limita a vectores 3, tiene muchas aplicaciones en ingeniería, física y matemáticas.

CONVENCIÓN

Todos los vectores en esta sección son vectores en el espacio.

Sistemas derechos e izquierdos

Hay dos tipos de sistemas coordenados en el espacio 3, el **derecho** y el **izquierdo**. Un sistema coordenado derecho es aquel cuyos semiejes positivos se identifican como sigue: cuando los dedos de la mano derecha se colocan a lo largo de la dirección positiva de las x, los dedos se cierran hacia la dirección positiva de las y, el pulgar debe apuntar hacia la dirección positiva de las z. Para un sistema izquierdo, la mano derecha se sustituye con la mano izquierda (figura 2.20).

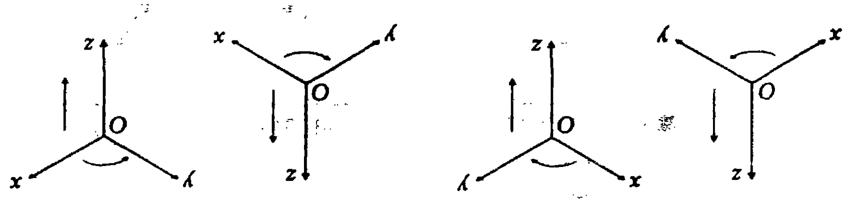


Figura 2.20

Notación para un determinante

La notación para un *determinante* que explicaremos a continuación puede ser muy útil para expresar algunas de las fórmulas básicas en esta sección. Primero, definiremos

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Por ejemplo,

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2$$

En segundo lugar, definiremos la notación

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

$$= a_1(b_2c_3 - c_2b_3) - b_1(a_2c_3 - c_2a_3) + c_1(a_2b_3 - b_2a_3)$$

Por ejemplo,

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & -2 & 1 \\ 4 & -1 & 0 \end{vmatrix} = 1 \begin{vmatrix} -2 & 1 \\ -1 & 0 \end{vmatrix} - 2 \begin{vmatrix} 2 & 1 \\ 4 & 0 \end{vmatrix} + 3 \begin{vmatrix} 2 & -2 \\ 4 & -1 \end{vmatrix} = 27$$

Producto cruz

DEFINICIÓN

(Producto cruz)

Sean $\mathbf{u} = (u_1, u_2, u_3)$ y $\mathbf{v} = (v_1, v_2, v_3)$. El **producto cruz**, $\mathbf{u} \times \mathbf{v}$ es el *vector* con componentes

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2,\ u_3v_1 - u_1v_3,\ u_1v_2 - u_2v_1)$$

Esta relación también puede expresarse en notación de determinantes:

$$\mathbf{u}\times\mathbf{v}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\u_1&u_2&u_3\\v_1&v_2&v_3\end{vmatrix}=\begin{vmatrix}u_2&u_3\\v_2&v_3\end{vmatrix}\mathbf{i}-\begin{vmatrix}u_1&u_3\\v_1&v_3\end{vmatrix}\mathbf{j}+\begin{vmatrix}u_1&u_2\\v_1&v_2\end{vmatrix}\mathbf{k}$$

la cual es lo mismo que

$$\mathbf{u}\times\mathbf{v}=\left(\begin{vmatrix}u_2&u_3\\v_2&v_3\end{vmatrix},-\begin{vmatrix}u_1&u_3\\v_1&v_3\end{vmatrix},\begin{vmatrix}u_1&u_2\\v_1&v_2\end{vmatrix}\right)$$

■ EJEMPLO 49 Sean $\mathbf{u}=(2,-1,3)$ y $\mathbf{v}=(1,-2,-1)$. Calcule $\mathbf{u}\times\mathbf{v}$.

SOLUCIÓN

$$\mathbf{u}\times\mathbf{v}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\2&-1&3\\1&-2&-1\end{vmatrix}=\begin{vmatrix}-1&3\\-2&-1\end{vmatrix}\mathbf{i}-\begin{vmatrix}2&3\\1&-1\end{vmatrix}\mathbf{j}+\begin{vmatrix}2&-1\\1&-2\end{vmatrix}\mathbf{k}$$

$$=7\mathbf{i}+5\mathbf{j}-3\mathbf{k}=(7,5,-3)$$

■ EJEMPLO 50 Calcule $\mathbf{i}\times\mathbf{j}$.

SOLUCIÓN

$$\mathbf{i}\times\mathbf{j}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\1&0&0\\0&1&0\end{vmatrix}=\begin{vmatrix}0&0\\1&0\end{vmatrix}\mathbf{i}-\begin{vmatrix}1&0\\0&0\end{vmatrix}\mathbf{j}+\begin{vmatrix}1&0\\0&1\end{vmatrix}\mathbf{k}=0\mathbf{i}-0\mathbf{j}+1\mathbf{k}=\mathbf{k}$$

Por consiguiente, $\mathbf{i}\times\mathbf{j}=\mathbf{k}$.

De hecho, todos los productos que se forman con **i**, **j** y **k** satisfacen

$$\begin{array}{ll}\mathbf{i}\times\mathbf{j}=\mathbf{k} & \mathbf{j}\times\mathbf{i}=-\mathbf{k}\\ \mathbf{j}\times\mathbf{k}=\mathbf{i} & \mathbf{k}\times\mathbf{j}=-\mathbf{i}\\ \mathbf{k}\times\mathbf{i}=\mathbf{j} & \mathbf{i}\times\mathbf{k}=-\mathbf{j}\end{array}$$

La figura 2.21 es un auxiliar mnemotécnico: al avanzar en el sentido de las manecillas del reloj, el producto cruz de dos vectores proporciona el tercero. Cuando el desplazamiento se da en sentido contrario al de las manecillas del reloj, el producto cruz de dos vectores obtiene el opuesto del tercero.

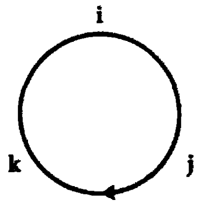


Figura 2.21

TEOREMA 20

(Propiedades del producto vectorial o producto cruz)[5]

Si $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ y $\mathbf{w} = (w_1, w_2, w_3)$ son vectores en el espacio y c es cualquier escalar. Entonces

1. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$;
2. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$;
3. $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$;
4. $c(\mathbf{u} \times \mathbf{v}) = (c\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (c\mathbf{v})$;
5. $\mathbf{0} \times \mathbf{u} = \mathbf{u} \times \mathbf{0} = \mathbf{0}$;
6. $\mathbf{u} \times \mathbf{u} = \mathbf{0}$;
7. $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$;
8. $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$

DEMOSTRACIÓN Demostraremos las partes 1, 7 y 8. El resto de las demostraciones se deja como ejercicio.

Para la parte 1,

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1) \\ &= -(v_2u_3 - v_3u_2, v_3u_1 - v_1u_3, v_1u_2 - v_2u_1) = -\mathbf{v} \times \mathbf{u} \end{aligned}$$

Para la parte 7, en vista de que

$$\mathbf{v} \times \mathbf{w} = (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1)$$

La primera componente de $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ es

$$u_2(v_1w_2 - v_2w_1) - u_3(v_3w_1 - v_1w_3) = u_2v_1w_2 - u_2v_2w_1 - u_3v_3w_1 + u_3v_1w_3$$

Por otro lado, la primera componente de $(\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$ es

$$(u_1w_1 + u_2w_2 + u_3w_3)v_1 - (u_1v_1 + u_2v_2 + u_3v_3)w_1 = u_2w_2v_1 + u_3w_3v_1 - u_2v_2w_1 - u_3v_3w_1$$

Por consiguiente, $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ y $(\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$ tienen sus primeras componentes iguales. La igualdad de las dos últimas componentes se demuestra en forma parecida.

Para la parte 8,

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= (u_1, u_2, u_3) \cdot \left(\begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix}, -\begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix}, \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \right) \\ &= u_1 \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} - u_2 \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} + u_3 \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \end{aligned}$$

$\square$

OBSERVACIÓN

1. De acuerdo con el enunciado 1 del teorema 20, el producto cruz **no** es conmutativo; es **anticonmutativo**.

[5] N. del T.: Al producto vectorial suele llamársele producto cruz.

2. El producto cruz **no es asociativo.**[6] Esto significa que en general $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \neq \mathbf{u} \times (\mathbf{v} \times \mathbf{w})$. Por ejemplo,

$$(\mathbf{i} \times \mathbf{j}) \times \mathbf{j} = \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \quad \text{mientras que} \quad \mathbf{i} \times (\mathbf{j} \times \mathbf{j}) = \mathbf{i} \times \mathbf{0} = \mathbf{0}$$

La mayor parte de las identidades donde intervienen los productos cruz pueden deducirse a partir del teorema 20. Por ejemplo, es fácil demostrar que la propiedad 8 implica que

$$\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0 \quad \text{y} \quad \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$$

De aquí que $\mathbf{u} \times \mathbf{v}$ sea *ortogonal* a $\mathbf{u}$ y a $\mathbf{v}$. Así,

$$\mathbf{u} \perp (\mathbf{u} \times \mathbf{v}) \quad \text{y} \quad \mathbf{v} \perp (\mathbf{u} \times \mathbf{v})$$

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores no cero, la dirección de $\mathbf{u} \times \mathbf{v}$ es perpendicular al plano definido por $\mathbf{u}$ y $\mathbf{v}$. Además, es posible comprobar que para un sistema coordenado derecho, los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} \times \mathbf{v}$ también forman un sistema derecho (figura 2.22). Esto determina la dirección del producto cruz. A continuación calcularemos su longitud.

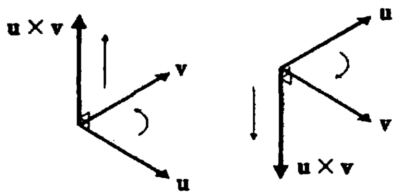


Figura 2.22 Dirección del producto cruz en un sistema derecho.

Al hacer el cálculo directo, la parte 8 del teorema 20 implica que

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) \tag{2.13}$$

(Esta propiedad también puede memorizarse con un diagrama de permutación cíclica, como el de la figura 2.21, reemplazando $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ por $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$.)

TEOREMA 21

(Longitud del producto cruz)

1. La siguiente identidad es válida:

$$\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \quad \textbf{Identidad de Lagrange}$$

2. Si θ es el ángulo que forman $\mathbf{u}$ y $\mathbf{v}$, entonces

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \operatorname{sen} \theta \tag{2.14}$$

[6] Para conocer una relación entre la no asociatividad del producto cruz y el problema de los cuatro colores, véase "Map Coloring and the Vector Cross Product", de Louis H. Kauffman, *Journal of Combinatorial Theory*, serie B, 48 (1990): 145-154.

DEMOSTRACIÓN

1. $\|\mathbf{u} \times \mathbf{v}\|^2 = (\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{u} \times \mathbf{v})$
$$= \mathbf{u} \cdot (\mathbf{v} \times (\mathbf{u} \times \mathbf{v})) \quad \text{de acuerdo con la ecuación (2.13)}$$
$$= \mathbf{u} \cdot ((\mathbf{v} \cdot \mathbf{v})\mathbf{u} - (\mathbf{v} \cdot \mathbf{u})\mathbf{v}) \quad \text{según el enunciado 7 del teorema 20}$$
$$= (\mathbf{v} \cdot \mathbf{v})(\mathbf{u} \cdot \mathbf{u}) - (\mathbf{v} \cdot \mathbf{u})(\mathbf{u} \cdot \mathbf{v})$$
$$= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$$

2. Según el anunciado 1 y de acuerdo con la ecuación (2.7), sección 2.2, se tiene

$$\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \cos^2 \theta$$
$$= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 (1 - \cos^2 \theta) = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \sin^2 \theta$$

En consecuencia, $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \operatorname{sen} \theta$.

El enunciado 2 del teorema 21 determina la longitud de $\mathbf{u} \times \mathbf{v}$. Geométricamente, esa longitud es igual al área del paralelogramo definida por $\mathbf{u}$ y $\mathbf{v}$ (figura 2.23). Por tanto, el área A del paralelogramo con $\mathbf{u}$ y $\mathbf{v}$ como lados adyacentes es

$$A = \|\mathbf{u} \times \mathbf{v}\|$$

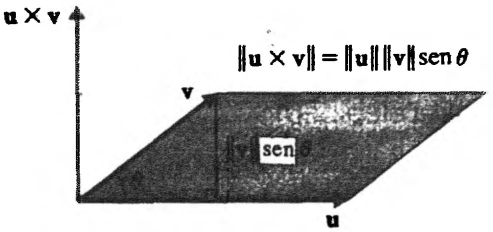


Figura 2.23

COROLARIO 22

$$\|\mathbf{u} \times \mathbf{v}\| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

DEMOSTRACIÓN Como $0 \leq \theta \leq \pi$, entonces $0 \leq \operatorname{sen} \theta \leq 1$. Por lo que la ecuación (2.14) implica que

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \operatorname{sen} \theta \leq \|\mathbf{u}\| \|\mathbf{v}\| \cdot 1 = \|\mathbf{u}\| \|\mathbf{v}\|$$

COROLARIO 23

(Criterio para que dos vectores sean paralelos)

Dos vectores $\mathbf{u}$ y $\mathbf{v}$, no cero, son paralelos si y sólo si $\mathbf{u} \times \mathbf{v} = \mathbf{0}$.

DEMOSTRACIÓN De acuerdo con la ecuación (2.14),

$$\mathbf{u} \times \mathbf{v} = \mathbf{0} \Leftrightarrow \|\mathbf{u} \times \mathbf{v}\| = 0 \Leftrightarrow \|\mathbf{u}\| \|\mathbf{v}\| \operatorname{sen} \theta = 0$$
$$\Leftrightarrow \operatorname{sen} \theta = 0 \Leftrightarrow \theta = 0, \pi \Leftrightarrow \mathbf{u} \text{ y } \mathbf{v} \text{ son paralelos}$$

Aplicaciones a geometría

■ EJEMPLO 51 (Área de un paralelogramo) Calcule el área del paralelogramo cuyos lados adyacentes son PQ y PR, siendo $P(2, 1, 0)$, $Q(1, -2, 1)$ y $R(-2, 2, 4)$.

SOLUCIÓN De acuerdo con la parte 2 del teorema 21, el área es $\|\overrightarrow{PQ} \times \overrightarrow{PR}\|$. Pero

$$\|\overrightarrow{PQ} \times \overrightarrow{PR}\| = \|(-1, -3, 1) \times (-4, 1, 4)\| = \|(-13, 0, -13)\| = 13\sqrt{2}$$

Por consiguiente, el área es $13\sqrt{2}$ unidades. □

■ EJEMPLO 52 (Área de un triángulo) Calcule el área del triángulo cuyos vértices están en las puntas de **i**, **j** y **k**.

SOLUCIÓN Dos lados del triángulo son $\mathbf{j} - \mathbf{i}$ y $\mathbf{k} - \mathbf{i}$. De manera que. $\|(\mathbf{j} - \mathbf{i}) \times (\mathbf{k} - \mathbf{i})\|$ es el área del paralelogramo definido por esos lados. La mitad de ésta es el área del triángulo.

$$\frac{1}{2}\|(\mathbf{j} - \mathbf{i}) \times (\mathbf{k} - \mathbf{i})\| = \frac{1}{2}\|(-1, 1, 0) \times (-1, 0, 1)\| = \frac{1}{2}\|(1, 1, 1)\| = \frac{1}{2}\sqrt{3}$$

□

TEOREMA 24

(Volumen de un paralelepípedo)

Demuestre que el volumen V del paralelepípedo cuyos lados adyacentes son los vectores de posición **u**, **v** y **w** (figura 2.24) está expresado por

$$V = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| \tag{2.15}$$

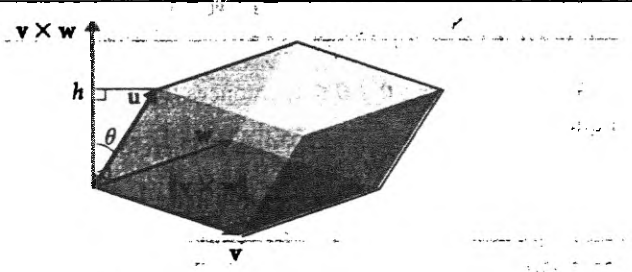


Figura 2.24

SOLUCIÓN Sea A el área de la base definida por **v** y **w**. Sea h la altura del paralelepípedo, y θ el ángulo que forman **u** y $\mathbf{v} \times \mathbf{w}$. Entonces, $h = \|\mathbf{u}\|\,|\cos\theta|$ y $A = \|\mathbf{v} \times \mathbf{w}\|$. Según la ecuación (2.7) en la sección 2.2,

$$V = Ah = \|\mathbf{v} \times \mathbf{w}\|\,\|\mathbf{u}\|\,|\cos\theta| = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$$

□

■ EJEMPLO 53 Calcule el volumen del paralelepípedo cuyos lados adyacentes son los vectores de posición $\mathbf{u} = (1, -1, 2)$, $\mathbf{v} = (0, 2, 1)$ y $\mathbf{w} = (3, -2, -1)$.

SOLUCIÓN Según el enunciado 8 del teorema 20,

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 3 & -2 & -1 \end{vmatrix} = -15$$

Tenemos que el volumen V del paralelepípedo es $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| = |-15| = 15$. □

La ecuación (2.15) define un criterio fácil mediante el cual podemos comprobar si tres puntos son coplanares. Si **u**, **v** y **w** son coplanares, el volumen del paralelepípedo cuyos lados adyacentes son **u**, **v** y **w** es cero (porque la altura es cero). Por el contrario, la única forma en que este volumen sea cero es que **u**, **v** y **w** sean coplanares.

TEOREMA 25

(Criterio para que tres vectores sean coplanares)

Los vectores **u**, **v** y **w** son coplanares si y sólo si $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0$.

Producto cruz con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> crossprod([1,2,3],[-3,2,-1]);
                              [-8  -8  8]
```

Mathematica

```
In[1]:=
    <<LinearAlgebra'CrossProduct'
In[2]:=
    Cross[{1,2,3},{-3,2,-1}]
Out[1]=
    {-8, -8, 8}
```

MATLAB

```
>> cross([1,2,3],[-3,2,-1])
ans =
   -8 -8 8
```

Ejercicios 2.6

1. Sean $\mathbf{u} = (-1, 2, -2)$, $\mathbf{v} = (4, -3, 5)$ y $\mathbf{w} = (-4, -2, 0)$. Efectúe las operaciones siguientes.

$$\mathbf{u} \times \mathbf{v} \qquad (\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$$
$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) \qquad (\mathbf{u} + \mathbf{v}) \times \mathbf{w}$$
$$\mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w} \qquad \mathbf{u} + (\mathbf{v} \times \mathbf{w})$$

2. Sea $\mathbf{u} \times \mathbf{v} = (2, 1, -5)$. Determine

$$\mathbf{v} \times \mathbf{u} \qquad -2(\mathbf{v} \times \mathbf{u})$$
$$\mathbf{u} \times (10\mathbf{v}) \qquad (-2\mathbf{u}) \times (10\mathbf{v})$$
$$\|\mathbf{u} \times \mathbf{v}\| \qquad \|(2\mathbf{v}) \times \mathbf{u}\|$$

3. Sea $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = -5$. Encuentre

$$(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u} \qquad \mathbf{u} \cdot (\mathbf{w} \times \mathbf{v})$$
$$\mathbf{v} \cdot (\mathbf{u} \times \mathbf{w}) \qquad \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$$

Sean u, v y w vectores 3.

4. ¿Cuáles de las expresiones siguientes son *indefinidas* y por qué?

$$\mathbf{u} \times \mathbf{u} \times \mathbf{u} \qquad \mathbf{u} \times \mathbf{v} \times \mathbf{w}$$
$$(\mathbf{u} \times \mathbf{u}) \times \mathbf{u} \qquad \mathbf{u} \times (\mathbf{v} \times \mathbf{w})$$
$$\mathbf{u} \cdot (\mathbf{u} \times \mathbf{w}) \qquad \mathbf{u} \times (\mathbf{u} \cdot \mathbf{w})$$
$$(\mathbf{u} \times \mathbf{u}) \times (\mathbf{v} \times \mathbf{v}) \qquad (\mathbf{u} \times \mathbf{w}) \cdot (\mathbf{v} \times \mathbf{w})$$

5. Compruebe la identidad de Lagrange para $\mathbf{u} = (-3, 4, 1)$ y $\mathbf{v} = (0, 5, -6)$.
6. Calcule el seno del ángulo formado por **u** y **v** para los valores siguientes.

   a. $\mathbf{u} = (6, 1, -2)$, $\mathbf{v} = (7, 5, -1)$
   b. $\mathbf{u} = (9, -7, 4)$, $\mathbf{v} = (0, -4, 3)$

7. Determine un vector unitario perpendicular al plano definido por $\mathbf{u} = (3, -4, 0)$ y $\mathbf{v} = (7, 5, -4.)$
8. Encuentre un vector de longitud 4 que sea perpendicular al plano definido por $\mathbf{u} = (1, -1, 1)$ y $\mathbf{v}(-1, 1, 0)$.
9. ¿Cuál es el área del triángulo cuyos vértices están en (1, 1, 1), (1, −1, −1) y (0, 1, −1)?
10. ¿Cuál es el área del paralelogramo cuyos lados adyacentes son $\overrightarrow{PQ}$ y $\overrightarrow{PR}$, siendo $P(1, 1, 1)$, $Q(1, -1, -1)$ y $R(0, 1, -1)$?
11. Calcule el volumen del paralelepípedo cuyos lados adyacentes son los vectores de posición (1, −2, 3), (2, 0, −5) y (0, 4, −1).
12. Aplique el producto cruz para demostrar que (1, 2, −1) y (−2, −4, 2) son paralelos.
13. Utilice el teorema 25 para determinar cuáles de los vectores **u**, **v** y **w** son coplanares.

    a. $\mathbf{u} = (-1, -1, 9)$, $\mathbf{v} = (0, 1, -3)$, $\mathbf{w} = (-1, 2, 0)$
    b. $\mathbf{u} = (1, -1, 1)$, $\mathbf{v} = (1, 0, 2)$, $\mathbf{w} = (1, -1, 0)$

14. ¿Es verdad que si $\mathbf{u} \times \mathbf{v} = 0$, entonces $\mathbf{u} = \mathbf{0}$ o $\mathbf{v} = \mathbf{0}$?
15. ¿Es cierto que si $\mathbf{u} \times \mathbf{v} = \mathbf{u} \times \mathbf{w}$, y también $\mathbf{u} \neq \mathbf{0}$, entonces $\mathbf{v} = \mathbf{w}$?
16. Termine la demostración del teorema 20.
17. Demuestre la identidad

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$$

18. **(Identidad de Jacobi)** Demuestre la identidad

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} + (\mathbf{v} \times \mathbf{w}) \times \mathbf{u} + (\mathbf{w} \times \mathbf{u}) \times \mathbf{v} = \mathbf{0}$$

19. **(Fórmula de Euler)** Si **u**, **v** y **w**, son tres lados adyacentes de un tetraedro con origen común (figura 2.25), demuestre que el volumen V se expresa con

$$V = \frac{1}{6} |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$$

[*Sugerencia:* Se sabe, de la geometría, que $V = \frac{1}{3}$ (área de la base)(altura).]

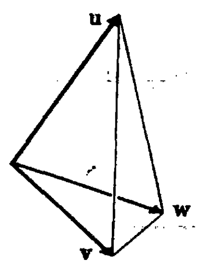


Figura 2.25

2.7 Líneas, planos e hiperplanos

Objetivos del alumno para esta sección

1. Formular las ecuaciones de: (a) una recta y (b) un plano, y determinar la normal de un plano.
2. Tener en cuenta que con frecuencia las cuestiones relacionadas con planos se reducen a cuestiones relacionadas con las normales a éstos.
3. Formular la ecuación de: (a) una recta, y (b) un hiperplano en $\mathbf{R}^n$.

Líneas

En esta sección describiremos la ecuación de una recta l que pasa por un punto dado $P(x_0, y_0, z_0)$ y es paralela a un vector $\mathbf{n} = (a, b, c)$ dado, *no cero*. Si $X(x, y, z)$ es cualquier punto de l, y $\mathbf{p} = (x_0, y_0, z_0)$ y $\mathbf{x}(x, y, z)$. Los múltiplos por escalar $t\mathbf{n}$ $(-\infty < t < \infty)$ representan a todos los vectores posibles paralelos a $\mathbf{n}$. Como $\mathbf{x} - \mathbf{p}$ es paralelo a $\mathbf{n}$, debe cumplirse $\mathbf{x} - \mathbf{p} = t\mathbf{n}$ (figura 2.26) para cierto escalar t. En consecuencia,

$$\mathbf{x} = \mathbf{p} + t\mathbf{n} \tag{2.16}$$

A esta vectorial se le llama una **ecuación paramétrica de la recta**, y t es el **parámetro** de la ecuación. Esta ecuación paramétrica también puede expresarse en función de sus componentes; $(x, y, z) = (x_0, y_0, z_0) + t(a, b, c)$ equivale a

$$\begin{aligned} x &= x_0 + ta \\ y &= y_0 + tb \\ z &= z_0 + tc \end{aligned} \tag{2.17}$$

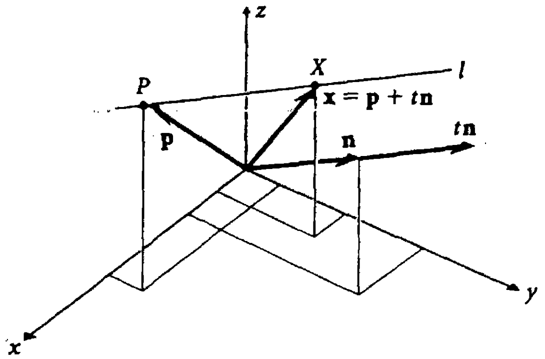


Figura 2.26

La ecuación (2.16) también es válida para líneas en el plano. Si $\mathbf{x} = (x, y)$, $\mathbf{p} = (x_0, y_0)$ y $\mathbf{n} = (a, b) \neq \mathbf{0}$, entonces $(x, y) = (x_0, y_0) + t(a, b)$, o bien

$$\begin{aligned} x &= x_0 + ta \\ y &= y_0 + tb \end{aligned} \tag{2.18}$$

■ EJEMPLO 54 Si l es la recta que pasa por $(1, -1, 2)$ en la dirección de $(1, 1, 1)$. Determine lo siguiente:

(a) Una ecuación paramétrica de l
(b) Dos puntos de l
(c) La intersección de l con los planos coordenados

SOLUCIÓN

(a) Como $\mathbf{n} = (1, 1, 1)$ y $\mathbf{p} = (1, -1, 2)$, una ecuación paramétrica de la recta es

$$\mathbf{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

o bien, en forma de componentes,

$$\begin{aligned} x &= 1 + t \\ y &= -1 + t \\ z &= 2 + t \end{aligned}$$

(b) Para determinar los puntos en l se necesita evaluar el parámetro t. Por ejemplo, $t = 1$ y $t = -1$ producen $(1, -1, 2) + 1(1, 1, 1) = (2, 0, 3)$ y $(1, -1, 2) + (-1)(1, 1, 1) = (0, -2, 1)$.

(c) Para calcular la intersección con el plano xy se iguala $z = 0$. De ahí que $z = 2 + t = 0$, es decir, $t = -2$. Al sustituir lo anterior en las dos primeras ecuaciones paramétricas tenemos, $x = -1, y = -3$. Así, $(-1, -3, 0)$ es la intersección con el plano xy. En forma parecida se obtienen las intersecciones con los planos xz y yz, que son $(2, 0, 3)$ y $(0, -2, 1)$, respectivamente. □

■ EJEMPLO 55 (Recta que pasa por dos puntos) Deduzca una ecuación paramétrica de la recta que pasa por los puntos $P(3, -1)$ y $Q(-1, 2)$.

SOLUCIÓN Como $\overrightarrow{PQ} = (-1, 2) - (3, -1) = (-4, 3)$ es paralelo a la recta, el vector de dirección es $\mathbf{n} = (-4, 3)$. Por tanto, una ecuación paramétrica de la recta es

$$\mathbf{x} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} + t \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$ □

■ EJEMPLO 56 (Rectas paralelas) Demuestre que las rectas cuyas ecuaciones paramétricas son

$$\begin{aligned} x &= 1 - 2t \\ y &= -1 + 4t \\ z &= 2 - 8t \end{aligned} \qquad \text{y} \qquad \begin{aligned} x &= -t \\ y &= 2 + 2t \\ z &= 7 - 4t \end{aligned}$$

son paralelas.

SOLUCIÓN Un vector de dirección para la primera recta es $(-2, 4, 8) = 2 \cdot (-1, 2, 4)$, que es un múltiplo escalar del vector de dirección $(-1, 2, 4)$ de la segunda recta. Ambas rectas tienen la misma dirección, así que son paralelas. $\square$

■ EJEMPLO 57 (Líneas perpendiculares) Demuestre que las rectas cuyas ecuaciones paramétricas

$$\begin{aligned} x &= 1 - 2t \\ y &= -1 + 4t \\ z &= 2 - 2t \end{aligned} \qquad \text{y} \qquad \begin{aligned} x &= -t \\ y &= 2 - 2t \\ z &= 7 - 3t \end{aligned}$$

son perpendiculares.

SOLUCIÓN Esto es cierto, porque los vectores de dirección $(-2, 4, -2)$ y $(-1, -2, -3)$ son ortogonales. $\square$

NOTA La ecuación paramétrica de una recta no es única. Puede emplearse cualquier punto de la recta o vector paralelo al vector de dirección.

Por el momento sólo consideraremos rectas en el plano. Como $\mathbf{n} \neq \mathbf{0}$, siempre es posible eliminar t del sistema (2.18) para obtener una ecuación de la recta en la forma conocida

$$Ax + By = C \tag{2.19}$$

como una relación entre y y x.

Por ejemplo, partiendo del conjunto paramétrico

$$\begin{aligned} x &= 2 - t \\ y &= 1 + 3t \end{aligned} \tag{2.20}$$

eliminamos t para obtener

$$y = -3x + 7 \tag{2.21}$$

A la inversa, la ecuación (2.20) puede determinarse a partir de la (2.21) haciendo $t = 2 - x$. Podemos obtener otro conjunto (más fácil) de ecuaciones paramétricas igualando $x = t$ y $y = -3t + 7$.

$$\begin{aligned} x &= t \\ y &= -3t + 7 \end{aligned} \tag{2.22}$$

Obsérvese que los dos conjuntos de ecuaciones (2.20) y (2.22) son equivalentes. Al reemplazar a t por $2 - t'$ en el sistema (2.20) se obtiene el sistema (2.22).

En el caso de rectas en el espacio, al despejar t no se obtendrá *una* ecuación, como con las rectas en el plano. Así, la *ecuación paramétrica es la única disponible para describir directamente una línea en el espacio.* Sin embargo, si $a \neq 0$, $b \neq 0$ y $c \neq 0$, podemos seguir eliminando t de las ecuaciones (2.17) para obtener *dos* ecuaciones:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

Las anteriores son **ecuaciones simétricas** de una recta. Describen *indirectamente* a la recta como una intersección de dos planos (que se explicará en la subsección siguiente).

■ EJEMPLO 58 (Ecuaciones simétricas de una recta) Deduzca las ecuaciones simétricas de la recta que pasa por $(-2, 3, 1)$ cuya dirección es $(-1, -2, 1)$.

SOLUCIÓN Como $(x_0, y_0, z_0) = (-2, 3, 1)$ y $(a, b, c) = (-1, -2, 1)$, entonces

$$\frac{x-(-2)}{-1} = \frac{y-3}{2} = \frac{z-1}{1} \qquad \square$$

Planos

Un vector $\mathbf{n} = (a, b, c)$ no cero se llama **normal** a un plano $\mathcal{P}$ si es perpendicular a $\mathcal{P}$ (figura 2.27). Sean $P(x_0, y_0, z_0)$ un punto dado de $\mathcal{P}$, y sea $X(x, y, z)$ cualquier otro punto. Si $\mathbf{p} = (x_0, y_0, z_0)$ y $\mathbf{x} = (x, y, z)$, entonces $\mathbf{x} - \mathbf{p}$ es paralelo a $\mathcal{P}$ y, en consecuencia, es ortogonal a la normal $\mathbf{n}$. Por consiguiente, el producto punto de $\mathbf{x} - \mathbf{p}$ por $\mathbf{n}$ es cero:

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0 \tag{2.23}$$

En función de componentes, esta ecuación puede expresarse como sigue:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{2.24}$$

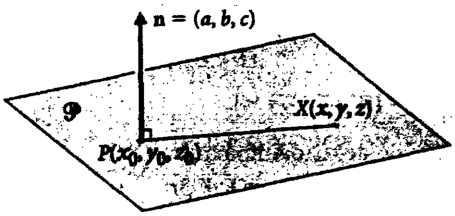


Figura 2.27

La ecuación (2.24) caracteriza a todos los puntos $\mathbf{x}$ de $\mathcal{P}$ en función de un vector normal $\mathbf{n}$ y de un punto $\mathbf{p}$ de $\mathcal{P}$. A esto se denomina forma **punto-normal** de la ecuación del plano $\mathcal{P}$.

Ésta es análoga a la ecuación simplificada[7] de la ecuación de una recta. Tanto la forma simple como la de punto-normal implican un punto dado y una inclinación. En el caso de la recta, la inclinación se expresa con la pendiente, mientras que en el caso del plano se define mediante la dirección de la normal.

■ EJEMPLO 59 (Ecuación punto-normal) ¿Cuál es la ecuación del plano que pasa por $(-1, 2, 3)$ y es perpendicular a $(-2, 1, 4)$? Determine otro punto en este plano.

SOLUCIÓN Como $\mathbf{p} = (-1, 2, 3)$ y $\mathbf{n} = (-2, 1, 4)$, con la ecuación (2.24) se obtiene

$$-2 \cdot (x + 1) + 1 \cdot (y - 2) + 4 \cdot (z - 3) = 0$$

[7] N. del T.: A la "forma simplificada" de la ecuación de la recta se le llama con frecuencia "forma punto-pendiente", porque su nombre en inglés es "point-slope form".

El punto en el plano puede obtenerse partiendo de cualquier solución de esta ecuación. Por ejemplo, con $x = 1, y = 2$, se obtiene $4z - 16 = 0$, es decir $z = 4$. Por lo que $(1, 2, 4)$ es otro punto. □

La ecuación (2.24) puede reformularse como sigue:

$$ax + by + cz + d = 0 \tag{2.25}$$

siendo $d = -ax_0 - by_0 - cz_0$. Ésta es la **ecuación general** del plano. Aun en esta forma, los coeficientes de x, y y z siguen definiendo una normal al plano. Por el contrario, la ecuación (2.25) es la ecuación de un plano.

TEOREMA 26

(Ecuación de un plano)

Si $(a, b, c) \neq \mathbf{0}$, la gráfica de la ecuación

$$ax + by + cz + d = 0$$

es un plano cuya normal es (a, b, c).

DEMOSTRACIÓN Si (x_0, y_0, z_0) satisface la ecuación, entonces

$$ax + by + cz + d = 0 \quad \text{y} \quad ax_0 + by_0 + cz_0 + d = 0$$

implican que

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad \text{o sea} \quad (a, b, c) \cdot (x - x_0, y - y_0, z - z_0) = 0$$

Como resultado (a, b, c) es una normal al plano que contiene al punto (x_0, y_0, z_0) y al vector $(x - x_0, y - y_0, z - z_0)$. □

■ EJEMPLO 60 (Planos paralelos) Determine la ecuación del plano que pasa por $(1, -2, 4)$ y que es paralelo al plano $2x - 5y + 2z - 1 = 0$.

SOLUCIÓN Como los dos planos son paralelos, tienen las mismas normales. La normal al plano dado es $(2, -5, 2)$ y por tanto

$$2 \cdot (x - 1) - 5 \cdot (y + 2) + 2 \cdot (z - 4) = 0$$

es la ecuación del plano desconocido. □

■ EJEMPLO 61 (Plano que pasa por tres puntos) Deduzca la ecuación del plano que pasa por los puntos $P(2, 0, 1)$, $Q(1, 2, 0)$ y $R(-3, 2, 1)$.

SOLUCIÓN El producto cruz $\overrightarrow{PQ} \times \overrightarrow{PR} = (-1, 2, -1) \times (-5, 2, 0) = (2, 5, 8)$ es una normal al plano. Por consiguiente, de acuerdo con la fórmula punto-normal, $2(x - 2) + 5y + 8(z - 1) = 0$, es decir

$$2x + 5y + 8z - 12 = 0$$

considerando que P es el punto en el plano. □

■ EJEMPLO 62 (Intersección de dos planos) Determine las ecuaciones paramétricas de la recta de intersección de los planos $x - y + z - 2 = 0$ y $2x + y + z + 1 = 0$.

SOLUCIÓN Sea $z = t$. Entonces, al despejar x y y del sistema

$$\begin{aligned} x - y + t - 2 &= 0 \\ 2x + y + t + 1 &= 0 \end{aligned}$$

se obtiene $x = -\frac{2}{3}t + \frac{1}{3}$ y $y = \frac{1}{3}t - \frac{5}{3}$. Así, las ecuaciones paramétricas son $x = -\frac{2}{3}t + \frac{1}{3}$, $y = \frac{1}{3}t - \frac{5}{3}$, $z = t$.

DEFINICIÓN

El **ángulo** que forman dos planos se define como el ángulo que forman dos normales a los planos.

■ EJEMPLO 63 (Ángulo entre dos planos) Calcule el coseno del ángulo que forman los planos $2x - y + z - 2 = 0$ y $x + 2y - z + 1 = 0$.

SOLUCIÓN Como $(2, -1, 1)$ y $(1, 2, -1)$ son las normales correspondientes, el coseno del ángulo que forman los planos es

$$\frac{(2,-1,1)\cdot(1,2,-1)}{\|(2,-1,1)\|\,\|(1,2,-1)\|} = -\frac{1}{6}$$

■ EJEMPLO 64 (Planos perpendiculares) Demuestre la perpendicularidad de los planos cuyas ecuaciones son $x + y + z = 0$ y $-x - y + 2z = 0$.

SOLUCIÓN Son perpendiculares, porque las normales $(1, 1, 1)$ y $(-1, -1, 2)$ son ortogonales.

Rectas e hiperplanos en $\mathbf{R}^n$

En $\mathbf{R}^n$ hay ecuaciones análogas de "rectas" y "planos", que se denominan *hiperplanos*. Un vector que va del punto $P(p_1, \ldots, p_n)$ a $Q(q_1, \ldots, q_n)$ tiene coordenadas

$$\overrightarrow{PQ} = (q_1 - p_1, \ldots, q_n - p_n)$$

Una **recta** que pasa por el punto $\mathbf{p} = (p_1, \ldots, p_n)$ y tiene la dirección $\mathbf{d} = (d_1, \ldots, d_n)$ es el conjunto de puntos $\mathbf{x} = (x_1, \ldots, x_n)$ tales que $\mathbf{x} - \mathbf{p}$ es paralelo a $\mathbf{d}$. Por consiguiente, $\mathbf{x} - \mathbf{p} = t\mathbf{d}$ para algún escalar t. Para $-\infty < t < \infty$, la ecuación

$$\mathbf{x} = \mathbf{p} + t\mathbf{d} \quad \text{es decir} \quad (x_1, \ldots, x_n) = (p_1, \ldots, p_n) + t(d_1, \ldots, d_n) \tag{2.26}$$

es la **ecuación paramétrica de una recta** en $\mathbf{R}^n$, con **parámetro** t.

■ EJEMPLO 65 La ecuación paramétrica de la recta en $\mathbf{R}^4$ que pasa por el punto $(1, 2, 3, 4)$ y que tiene la dirección $(1, 1, 1, 1)$ es

$$(x_1, x_2, x_3, x_4) = (1, 1, 1, 1) + t(1, 2, 3, 4)$$

■ EJEMPLO 66 Determine la ecuación paramétrica de la recta que pasa por los puntos $(-1, 1, -1, 1)$ y $(1, 2, 3, 4)$.

SOLUCIÓN La dirección de la recta es $(1, 2, 3, 4) - (-1, 1, -1, 1) = (2, 1, 4, 3)$. En consecuencia, la ecuación de la recta es $(x_1, x_2, x_3, x_4) = (-1, 1, -1, 1) + t(2, 1, 4, 3)$. $\square$

Ya vimos que los puntos $\mathbf{x}$ de un plano que pasa por un punto $\mathbf{p}$, y cuya normal es $\mathbf{n}$, deben satisfacer la ecuación

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0 \tag{2.27}$$

porque $\mathbf{n}$ y $\mathbf{x} - \mathbf{p}$ deben ser ortogonales. El análogo de esta ecuación en $\mathbf{R}^n$ es la **ecuación punto-normal de un hiperplano**. Si $\mathbf{n} = (a_1, \ldots, a_n)$, $\mathbf{x} = (x_1, \ldots, x_n)$ y $\mathbf{p} = (p_1, \ldots, p_n)$, la ecuación (2.27) implica que

$$a_1(x_1 - p_1) + \cdots + a_n(x_n - p_n) = 0 \tag{2.28}$$

Esta ecuación también puede escribirse como sigue:

$$a_1x_1 + \cdots + a_nx_n + d = 0 \tag{2.29}$$

que es lo que se llama **ecuación general de un hiperplano**.

■ EJEMPLO 67 Deduzca una ecuación del hiperplano en $\mathbf{R}^4$ que pasa por el punto $(1, 2, 3, 4)$ y es normal a la dirección $(-1, 2, -2, 1)$.

SOLUCIÓN De acuerdo con la ecuación (2.28),

$$\begin{aligned} -1(x_1 - 1) + 2(x_2 - 2) - 2(x_3 - 3) + 1(x_4 - 4) &= 0 \\ \Leftrightarrow -x_1 + 2x_2 - 2x_3 + x_4 - 1 &= 0 \end{aligned}$$

$\square$

Ejercicios 2.7

Para todos los ejercicios de esta sección, sean l_1, l_2, l_3 y l_4 las líneas cuyas respectivas ecuaciones paramétricas son:

$$l_1: \begin{aligned} x &= 5 - 3t \\ y &= -4 + 2t \\ z &= 2 - t \end{aligned} \qquad l_2: \begin{aligned} x &= 1 + 6t \\ y &= 6 - 4t \\ z &= 8 + 2t \end{aligned}$$

$$l_3: \begin{aligned} x &= 5 - s \\ y &= -7 - s \\ z &= 11 + s \end{aligned} \qquad l_4: \begin{aligned} x &= 14 + s \\ y &= -2 + 2s \\ z &= 13 + 3s \end{aligned}$$

y sean P, Q, R y S los puntos

$P(5, -4, 2)$ $\quad Q(2, -2, 1)$
$R(1, -2, 2)$ $\quad S(9, -12, -2)$

1. ¿Cuáles de los puntos P, Q, R y S están en l_1?
2. ¿Cuáles de los puntos P, Q, R y S están en l_3?
3. Determine 3 puntos en l_1.
4. Encuentre la intersección de l_3 con los planos coordenados.
5. Obtenga todos los pares de rectas paralelas que hay en el conjunto l_1, l_2, l_3 y l_4.
6. Encuentre todos los pares de rectas perpendiculares que hay en el conjunto l_1, l_2, l_3 y l_4.
7. Demuestre que l_1 y l_4 se intersecan. Determine su punto de intersección.
8. Compruebe que l_1 y l_3 son rectas *sesgadas*. (Esto es, que no son paralelas ni se intersecan.)

9. Deduzca las ecuaciones simétricas de las rectas l_1, l_2, l_3 y l_4.

10. Para cada recta de l_1, l_2, l_3 y l_4, obtenga las ecuaciones de dos planos cuya intersección esté definida por la recta dada. (*Sugerencia:* Emplee las ecuaciones simétricas.)

11. Determine las ecuaciones paramétricas de la recta que pasa por P y es paralela a $\mathbf{n} = (4, -3, 1)$.

12. Deduzca las ecuaciones paramétricas de la recta que pasa por $(0, 3, 2)$ y es paralela a $\mathbf{n} = (-1, 2, 4)$.

13. Obtenga las ecuaciones paramétricas de la recta que pasa por P y por Q.

14. Deduzca las ecuaciones paramétricas de la recta que pasa por $(2, -3, 5)$ y $(-7, 4, 1)$.

15. Formule las ecuaciones simétricas de la recta que pasa por $(3, -1, -2)$ y $(-1, 2, 5)$.

16. Encuentre las ecuaciones simétricas de la recta que pasa por P y tiene la dirección $\overrightarrow{SR}$.

17. ¿Cuál de los puntos P, Q, R y S pertenece al plano $x + 3y + 3z + 1 = 0$?

En los ejercicios 18 y 19 encuentre una forma punto-normal de la ecuación del plano que pasa por el punto X y cuya normal es $\mathbf{n}$.

18. $X = (-4, 2, 7)$ y $\mathbf{n} = (-3, 2, 1)$.

19. $X = P$ y $\mathbf{n} = (-6, 4, 5)$.

20. Obtenga una ecuación general del plano del ejercicio 18.

21. Determine una ecuación general del plano del ejercicio 19.

22. Deduzca una forma punto-normal de la ecuación del plano cuya ecuación general es
    a. $2x - 3y + z - 9 = 0$
    b. $x - 7y + 3 = 0$

23. Formule una ecuación del plano que pasa por P, Q y R.

24. Obtenga una ecuación del plano que pasa por $(2, -4, 1)$ y l_1.

25. Determine una ecuación del plano que contiene a las rectas l_1 y l_4.

26. Encuentre una ecuación del plano que pasa por $(2, 3, -1)$ y es perpendicular a $(-2, 4, 1)$.

27. Formule la ecuación del plano que pasa por $(-1, -2, 5)$ y es paralelo al plano $x - 6y + 2z - 3 = 0$.

28. Determine las ecuaciones paramétricas de la recta de intersección de los planos $x - y + z - 3 = 0$ y $-x + 5y + 3z + 4 = 0$.

29. Calcule el coseno del ángulo que forman los planos $6x + y + z - 1 = 0$ y $x + y - z + 1 = 0$.

30. Demuestre que los planos cuyas ecuaciones son $x - y + 2z + 3 = 0$ y $-x + 2y + \frac{3}{2}z = 0$ son perpendiculares.

31. Deduzca una ecuación del hiperplano en $\mathbf{R}^5$ que pasa por el punto $(1, 2, 0, -1, 0)$ y es normal a $(-1, 3, -2, 8, 4)$.

32. Demuestre que las rectas $l_1\!: \mathbf{x} = \mathbf{x}_1 + t\mathbf{n}_1$ y $l_2\!: \mathbf{x} = \mathbf{x}_2 + s\mathbf{n}_2$ se intersecan si y sólo si $\mathbf{x}_2 - \mathbf{x}_1$ está en $\mathrm{Gen}\{\mathbf{n}_1, \mathbf{n}_2\}$.

2.8 Aplicaciones

Objetivo del alumno para esta sección
Apreciar las numerosas aplicaciones de los vectores.

Hay un sinnúmero de aplicaciones de los vectores en casi todas las áreas de matemáticas, física e ingeniería. En esta sección describiremos algunas de las más importantes, si bien todavía elementales. Aunque el énfasis es hacia las aplicaciones geométricas de los vectores en la estática y en la ingeniería. También aplicaremos el producto $A\mathbf{x}$ para la obtención de mejores datos y así obtener gráficas más suaves y, más importante, a los sistemas dinámicos.

Selección de datos para suavizar gráficas

Al medir varias cantidades que dependen del tiempo, con frecuencia se recopilan datos que incluyen perturbaciones repentinas. Por ejemplo, supongamos que se miden velocidades del viento, y que se anotan algunos valores muy altos, por las rachas que sólo duran corto tiempo. Es posible tratar de minimizar el impacto de esas rachas breves que pudieran afectar la interpretación de los datos. Una forma de hacerlo es *suavizando*, uniformando o emparejando los datos. Una manera de suavizar es *promediando*. Si tenemos una sucesión de números

$$a, b, c, d, e, \ldots$$

puede transformarse en la sucesión de los *promedios* sucesivos

$$\frac{a}{2}, \frac{a+b}{2}, \frac{b+c}{2}, \frac{c+d}{2}, \frac{d+e}{2}, \ldots$$

que comienza con el promedio de a y 0 como primer número nuevo. (En la práctica se usan algunos otros esquemas.)

El promediar es, de hecho, una multiplicación por una matriz cuadrada finita de la forma

$$A = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & \cdots \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \cdots \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Por ejemplo, supongamos que cada hora se registran las siguientes velocidades de viento, en decenas de millas por hora:

$$2 \quad 1 \quad 3 \quad 3 \quad 4 \quad 5 \quad 3 \quad 4 \quad 3 \quad 2 \quad 1 \quad 2$$

Al graficar esos datos en función del tiempo se obtiene la figura 2.28.

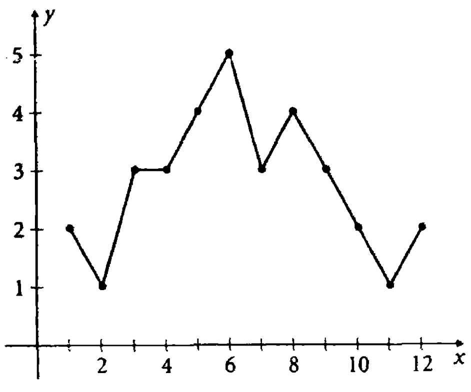


Figura 2.28 Gráfica de datos.

Al promediar se transforma esta sucesión en

$$1 \quad \frac{3}{2} \quad 2 \quad 3 \quad \frac{7}{2} \quad \frac{9}{2} \quad 4 \quad \frac{7}{2} \quad \frac{7}{2} \quad \frac{5}{2} \quad \frac{3}{2} \quad \frac{3}{2}$$

Al graficar estos datos se obtiene una gráfica más uniforme (figura 2.29).

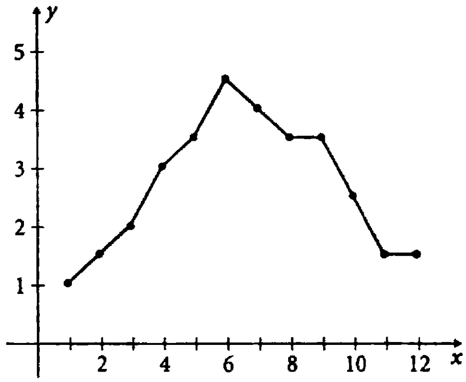


Figura 2.29 Datos promediados.

Si deseamos suavizar más la gráfica usando esta técnica, podemos promediar una vez más. También es posible hacerlo multiplicando la primera sucesión por una matriz de la forma B (compruébelo):

$$B = \begin{bmatrix} \frac{1}{4} & 0 & 0 & 0 & 0 & \cdots \\ \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & \cdots \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & \cdots \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & \cdots \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

La nueva sucesión es

$$\frac{1}{2} \quad \frac{5}{4} \quad \frac{7}{4} \quad \frac{5}{2} \quad \frac{13}{4} \quad 4 \quad \frac{17}{4} \quad \frac{15}{4} \quad \frac{7}{2} \quad 3 \quad 2 \quad \frac{3}{2}$$

Al graficarla se obtiene una gráfica aún más uniforme (figura 2.30).

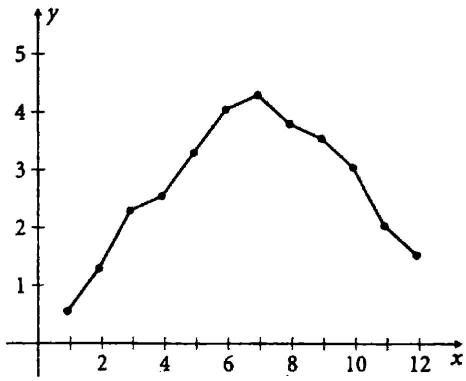


Figura 2.30 Datos promediados dos veces.

Para ahondar en éste y otros ejemplos interesantes de transformaciones matriciales, véase *The Mathematics of Matrices*, de Philip J. Davis (Nueva York: Blaisdell Publishing Co., 1965), pp. 253-264.

Sistemas dinámicos

En esta sección presentaremos los sistemas dinámicos discretos, como aplicaciones del producto $A\mathbf{x}$. Debido a su importancia, vamos a regresar varias veces a este tema, a medida que crezca nuestro acervo de álgebra lineal.

Para generalizar, diremos que un sistema dinámico es una ecuación o un sistema de ecuaciones que tienen por objeto estudiar cantidades que dependen del tiempo. Un ejemplo característico es la ecuación donde interviene el saldo, P_t, de una cuenta con intereses en el momento t. En un sistema dinámico *discreto*, la variable de tiempo sólo asume valores *enteros*.[8] Por ejemplo, si una cuenta gana 8% de interés compuesto anualmente, entonces P_0 representa al depósito inicial, P_1 el saldo al final del primer año, P_2 el saldo al final del segundo, y así sucesivamente. En este caso, aunque los valores no enteros de t, como por ejemplo $P_{1.5}$ tienen sentido, basta con conocer P en valores enteros, porque $P_{1.5} = P_1$, $P_{2.7} = P_2$, etc. Sea $P_0 = \$1\ 000$. Entonces, $P_1 = 0.08 \cdot P_0 + P_0 = 1\ 080$, $P_2 = 0.08P_1 + P_1 = 1\ 166.4, \ldots$ Al final del $(k + 1)$-ésimo año, el saldo es

$$P_{k+1} = 0.08P_k + P_k = 1.08P_k \tag{2.30}$$

A la ecuación (2.30) se le llama **sistema dinámico discreto**, o **ecuación en diferencias**, y da como resultado el valor siguiente de P en función del valor actual. P_k puede calcularse aplicando repetidamente la ecuación (2.30):

$$P_k = 1.08P_{k-1} = 1.08 \cdot 1.08P_{k-2} = (1.08)^2P_{k-2} = \cdots$$

Por lo que,

$$P_k = (1.08)^kP_0 \tag{2.31}$$

La ecuación (2.31) se denomina **solución** del sistema dinámico.

A veces, la cantidad que depende del tiempo tiene varios componentes, y puede representarse con un vector. En ese caso, se aplica la teoría de las matrices para estudiar el sistema dinámico correspondiente.

Un modelo de crecimiento de población

Supongamos que hay una población de insectos dividida en tres grupos de edad, A, B y C. El grupo A está formado por insectos de 0 a 1 semanas de edad, el grupo B por insectos de 1 a 2 semanas y el grupo C por insectos de 2 a 3 semanas. Supongamos que los grupos tienen A_k, B_k y C_k insectos al final de la k-ésima semana. Se desea estudiar cómo varían A, B y C al paso del tiempo, dadas las dos condiciones siguientes:

1. **(Tasa de supervivencia)** Al término de una semana, sólo sobrevive el 10% del grupo A. Por consiguiente,

$$B_{k+1} = \frac{1}{10}A_k \tag{2.32}$$

[8] Para conocer una excelente introducción a este tema, véase *Discrete Dynamical Systems, Theory and Applications*, de James T. Sandefur (Oxford: Clarendon Press, 1990).

Y al término de una semana sólo sobrevive el 40% del grupo B. Es decir,

$$C_{k+1} = \frac{2}{5}B_k \tag{2.33}$$

2. **(Tasa de natalidad)** Cada insecto del grupo A tiene un promedio de $\frac{2}{5}$ de descendientes, cada uno del grupo B tiene 4 descendientes y cada uno del grupo C tiene 5. En la semana $k + 1$, los insectos del grupo A son descendientes de los insectos en la semana k. En consecuencia,

$$A_{k+1} = \frac{2}{5}A_k + 4B_k + 5C_k \tag{2.34}$$

■ PROBLEMA Si la población de insectos se inicia con 1 000 en cada grupo de edad, ¿cuántos insectos hay en cada grupo al final de la tercera semana?

SOLUCIÓN Las ecuaciones (2.32), (2.33) y (2.34) se expresa en términos de vectores y matrices como sigue:

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \\ C_{k+1} \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & 4 & 5 \\ \frac{1}{10} & 0 & 0 \\ 0 & \frac{2}{5} & 0 \end{bmatrix} \begin{bmatrix} A_k \\ B_k \\ C_k \end{bmatrix}$$

Esta ecuación matricial es el *sistema dinámico* del problema.

La condición de la población inicial (**condición inicial**) es

$$\begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} = \begin{bmatrix} 1000 \\ 1000 \\ 1000 \end{bmatrix}$$

Al final de la primera semana se tiene:

$$\begin{bmatrix} A_1 \\ B_1 \\ C_1 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & 4 & 5 \\ \frac{1}{10} & 0 & 0 \\ 0 & \frac{2}{5} & 0 \end{bmatrix} \begin{bmatrix} 1000 \\ 1000 \\ 1000 \end{bmatrix} = \begin{bmatrix} 9400 \\ 100 \\ 400 \end{bmatrix}$$

Al terminar la segunda semana hay

$$\begin{bmatrix} A_2 \\ B_2 \\ C_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & 4 & 5 \\ \frac{1}{10} & 0 & 0 \\ 0 & \frac{2}{5} & 0 \end{bmatrix} \begin{bmatrix} 9400 \\ 100 \\ 400 \end{bmatrix} = \begin{bmatrix} 6160 \\ 940 \\ 40 \end{bmatrix}$$

y al final de la tercera semana,

$$\begin{bmatrix} A_3 \\ B_3 \\ C_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & 4 & 5 \\ \frac{1}{10} & 0 & 0 \\ 0 & \frac{2}{5} & 0 \end{bmatrix} \begin{bmatrix} 6160 \\ 940 \\ 40 \end{bmatrix} = \begin{bmatrix} 6424 \\ 616 \\ 376 \end{bmatrix}$$

De manera que después de 3 semanas, el grupo A tiene 6 424 insectos, 616 insectos el grupo B, y el C, 376 insectos. ☐

Geometría analítica

■ EJEMPLO 68 (Distancia de un punto a un plano) Deduzca una fórmula para determinar la distancia d más corta del punto $P(x_0, y_0, z_0)$ al plano $\mathcal{P}$ cuya ecuación es $ax + by + cz + d = 0$.

SOLUCIÓN La ecuación de $\mathcal{P}$ implica que $\mathbf{n} = (a, b, c)$ es un vector normal. Si $Q(x, y, z)$ es cualquier punto en $\mathcal{P}$. Entonces, d es la longitud de la proyección ortogonal, $\mathbf{d}$, de $\overrightarrow{QP}$ a lo largo de $\mathbf{n}$ (figura 2.31). Por consiguiente, la ecuación (2.8) de la sección 2.2 implica que

$$d = \|\mathbf{d}\| = \left\| \frac{\overrightarrow{QP} \cdot \mathbf{n}}{\mathbf{n} \cdot \mathbf{n}} \mathbf{n} \right\| = |\overrightarrow{QP} \cdot \mathbf{n}| \frac{1}{\|\mathbf{n}\|} = \frac{|a(x_0 - x) + b(y_0 - y) + c(z_0 - z)|}{\sqrt{a^2 + b^2 + c^2}}$$

$$= \frac{|ax_0 + by_0 + cz_0 - (ax + by + cz)|}{\sqrt{a^2 + b^2 + c^2}}$$

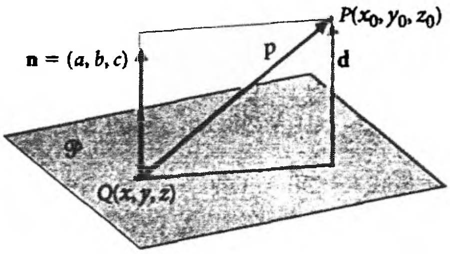


Figura 2.31 Distancia de un punto a un plano.

Las coordenadas (x, y, z) de Q satisfacen la ecuación del plano $ax + by + cz + d = 0$, porque Q está en $\mathcal{P}$. En consecuencia, $d = -ax - by - cz$, y así,

$$d = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$$

■ EJEMPLO 69 Calcule la distancia del punto $P(-1, 3, -2)$ al plano $2x - 3y + z - 1 = 0$.

SOLUCIÓN De acuerdo con el ejemplo 68,

$$d = \frac{|2 \cdot (-1) + (-3) \cdot 3 + 1 \cdot (-2) - 1|}{\sqrt{2^2 + (-3)^2 + 1^2}} = \sqrt{14} \cong 3.74$$

Las *líneas sesgadas* son rectas que no son paralelas ni se intersecan.

■ EJEMPLO 70 (Distancia entre rectas sesgadas) Determine la distancia d más corta entre las dos rectas sesgadas l_1 y l_2.

SOLUCIÓN Sean P, Q y R, S, dos pares de puntos en l_1 y l_2, respectivamente (figura 2.32). El producto cruz $\mathbf{n} = \overrightarrow{PQ} \times \overrightarrow{RS}$ es ortogonal a $\overrightarrow{PQ}$ y $\overrightarrow{RS}$. Sean $\mathcal{P}_1$ y $\mathcal{P}_2$ los planos que pasan por P y R, cuya normal es $\mathbf{n}$, entonces, $\mathcal{P}_1$ y $\mathcal{P}_2$ son paralelos y contienen a las líneas l_1 y l_2. Así, d es la distancia entre $\mathcal{P}_1$ y $\mathcal{P}_2$, y puede calculársele como longitud de la proyección ortogonal de $\overrightarrow{PR}$ a lo largo de la dirección de la normal $\mathbf{n}$. La ecuación (2.8) implica que

$$d = \left\| \frac{\overrightarrow{PR} \cdot \mathbf{n}}{\mathbf{n} \cdot \mathbf{n}} \mathbf{n} \right\| = \left| \overrightarrow{PR} \cdot \mathbf{n} \right| \frac{1}{\|\mathbf{n}\|}$$

Así, en términos de los puntos P, Q, R y S, tenemos

$$d = \left| \overrightarrow{PR} \cdot (\overrightarrow{PQ} \times \overrightarrow{RS}) \right| \frac{1}{\|\overrightarrow{PQ} \times \overrightarrow{RS}\|}$$

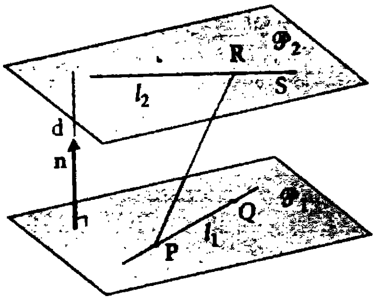


Figura 2.32 Distancia entre dos rectas sesgadas.

■ EJEMPLO 71 Determine la distancia más corta entre dos rectas sesgadas. l_1 pasa por $P(1, -2, -1)$ y $Q(0, -2, 1)$ y l_2 pasa por $R(-1, 2, 0)$ y $S(-1, 0, -2)$.

SOLUCIÓN $\overrightarrow{PQ} = (-1, 0, 2)$, $\overrightarrow{RS} = (0, -2, -2)$ y $\overrightarrow{PR} = (-2, 4, 1)$. Por consiguiente, $\overrightarrow{PQ} \times \overrightarrow{RS} = (4, -2, 2)$ y

$$d = |(-2, 4, 1) \cdot (4, -2, 2)| \frac{1}{\|(4, -2, 2)\|} = \frac{7}{\sqrt{6}} \cong 2.86$$

Geometría euclidiana

Para demostrar los teoremas de la geometría euclidiana pueden emplearse vectores, como se indica en los ejemplos siguientes.

■ EJEMPLO 72 Demuestre que el segmento de recta que bisecta dos lados de un triángulo tiene longitud igual a la mitad del tercer lado.

SOLUCIÓN De acuerdo con la figura 2.33(a), necesitamos demostrar que si $AP = PB$ y $AQ = QC$, entonces $PQ = \frac{1}{2}BC$. Se tiene

$$\overrightarrow{PQ} = \overrightarrow{PA} + \overrightarrow{AQ} = \frac{1}{2}\overrightarrow{BA} + \frac{1}{2}\overrightarrow{AC} = \frac{1}{2}(\overrightarrow{BA} + \overrightarrow{AC}) = \frac{1}{2}\overrightarrow{BC}$$

Por consiguiente, la longitud PQ de $\overrightarrow{PQ}$ es la mitad de la longitud BC. □

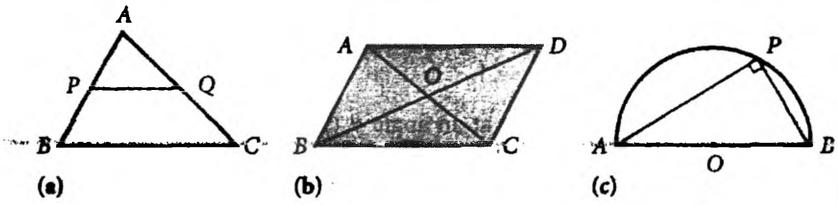


Figura 2.33

■ EJEMPLO 73 Demuestre que las diagonales de un paralelogramo se bisectan entre sí.

SOLUCIÓN Sea $ABCD$ el paralelogramo, y sea O el punto medio de la diagonal AC, figura 2.33(b). Entonces, $AO = OC$. Basta con demostrar que BOD es un segmento de recta, y que $BO = OD$. Puesto que $\overrightarrow{BC} = \overrightarrow{AD}$,

$$\overrightarrow{BO} = \overrightarrow{BC} + \overrightarrow{CO} = \overrightarrow{AD} + \overrightarrow{OA} = \overrightarrow{OA} + \overrightarrow{AD} = \overrightarrow{OD}$$

Por lo que $\overrightarrow{BO}$ y $\overrightarrow{OD}$ tienen la misma dirección, así como un punto común y la misma longitud. Por todo lo anterior, $BO = OD$, y B, O, D son colineales como se había afirmado. □

■ EJEMPLO 74 Demuestre que cualquier ángulo inscrito en un semicírculo es recto.

SOLUCIÓN De acuerdo con la figura 2.33(c), es suficiente con demostrar que $\overrightarrow{AP}$ y $\overrightarrow{BP}$ son perpendiculares, o que $\overrightarrow{AP} \cdot \overrightarrow{PB} = 0$. Sea r el radio. Entonces

$$\begin{aligned}\overrightarrow{AP} \cdot \overrightarrow{PB} &= (\overrightarrow{AO} + \overrightarrow{OP}) \cdot (\overrightarrow{PO} + \overrightarrow{OB}) \\ &= (\overrightarrow{OB} + \overrightarrow{OP}) \cdot (-\overrightarrow{OP} + \overrightarrow{OB}) = \overrightarrow{OB} \cdot \overrightarrow{OB} - \overrightarrow{OP} \cdot \overrightarrow{OP} \\ &= \|\overrightarrow{OB}\|^2 - \|\overrightarrow{OP}\|^2 = r^2 - r^2 = 0\end{aligned}$$

□

Física e ingeniería

Tanto el producto punto como el producto cruz tienen sorprendentes interpretaciones físicas. El producto punto puede considerarse como el trabajo efectuado por una fuerza constante, y el producto cruz como el vector momento de una fuerza.

Si una fuerza constante $\mathbf{F}$ hace mover un objeto una distancia $\mathbf{d}$ en la dirección de $\mathbf{F}$ figura 2.34(a), el trabajo efectuado es

$$W = \|\mathbf{F}\|\,\|\mathbf{d}\|$$

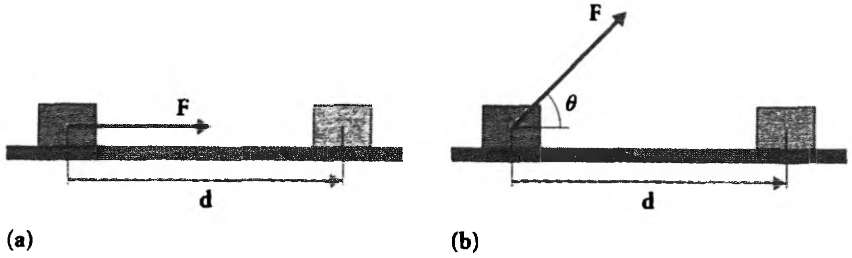


Figura 2.34

Si **F** y **d** forman un ángulo θ, figura 2.34(b), entonces se define a W en función de la componente numérica de **F** en la dirección de **d**. En otras palabras,

$$W = \|\mathbf{F}\|\,\|\mathbf{d}\| \operatorname{sen}\theta$$

Así, la ecuación (2.7) implica que el **trabajo** efectuado por **F** es

$$W = \mathbf{F} \cdot \mathbf{d} \tag{2.35}$$

■ EJEMPLO 75 (Trabajo efectuado por una fuerza constante) Calcule el trabajo efectuado por la fuerza constante $\mathbf{F} = 4\mathbf{i} - 2\mathbf{j} + \mathbf{k}$ si su punto de aplicación se mueve desde $P(0, 1, -2)$ hasta $Q(3, 0, 1)$.

SOLUCIÓN

$$\mathbf{d} = \overrightarrow{PQ} = (3, 0, 1) - (0, 1, -2) = (3, -1, 3)$$

Así, de acuerdo con la ecuación (2.35),

$$W = \mathbf{F} \cdot \mathbf{d} = (4, -2, 1) \cdot (3, -1, 3) = 17$$

Si la fuerza está en newtons y la distancia en metros, W se expresa en newton-metros ($\mathrm{N} \cdot \mathrm{m}$). □

Los vectores pueden emplearse en forma gráfica y algebraica para calcular las fuerzas resultantes, o las condiciones de equilibrio de un objeto.

■ EJEMPLO 76 (Plano inclinado) Calcule la fuerza **F** que se ejerce sobre la cuerda (figura 2.35) para equilibrar el objeto que pesa 500 lb, si la inclinación del plano es de 30°.

SOLUCIÓN Las fuerzas que actúan sobre el objeto son su peso, **W**, la fuerza sobre la cuerda, **F**, y la reacción del plano inclinado, **R**. Como el sistema está en equilibrio, la suma vectorial de esas fuerzas debe ser igual a cero. Conocemos la dirección de **F**, así que sólo se necesita conocer su magnitud. El peso **W** se puede expresar como la suma de dos vectores componentes, $\mathbf{W}_1$ y $\mathbf{W}_2$, en las direcciones de **F** y **R**, respectivamente. Como $\|\mathbf{F}\| = \|\mathbf{W}_1\|$, en ese caso

$$\|\mathbf{F}\| = \|\mathbf{W}_1\| = 500 \cdot \cos 60° = 500 \cdot \frac{1}{2} = 250 \text{ lb}$$ □

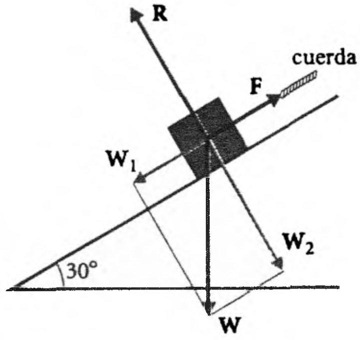


Figura 2.35

Ahora describiremos con más detalle el equilibrio de los cuerpos rígidos sometidos a fuerzas en el plano.

El **momento** m de una fuerza $\mathbf{F}$ respecto a un punto P es el producto $m = \|\mathbf{F}\|d$, donde d es la distancia más corta de P a la recta l definida por la dirección de $\mathbf{F}$ (figura 2.36). Sea Q cualquier punto en l y $\mathbf{r} = \overrightarrow{PQ}$. Entonces, $d = \|\mathbf{r}\|$ sen θ, siendo θ el ángulo que forman $\mathbf{r}$ y $\mathbf{F}$. Así,

$$m = \|\mathbf{r}\|\,\|\mathbf{F}\| \operatorname{sen} \theta$$

o bien, de acuerdo con la parte 2 del teorema 21,

$$m = \|\mathbf{r} \times \mathbf{F}\|$$

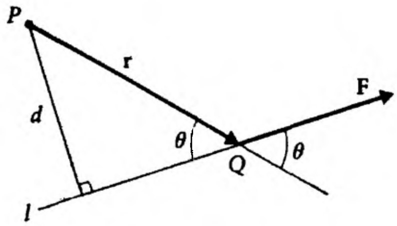


Figura 2.36

Definiremos al **vector momento**, o **momento m**, como sigue:

$$\mathbf{m} = \mathbf{r} \times \mathbf{F}$$

La magnitud de $\mathbf{m}$ es el momento m, y se desplaza a lo largo del eje de rotación que genera $\mathbf{F}$ con respecto a P.

■ EJEMPLO 77 (Relación entre el momento de fuerza y un punto) Se aplica una fuerza de 3 N formando un ángulo de 60° con el eje de x positiva, en el extremo del vector de posición $\mathbf{r} = (\sqrt{3}, 1)$. Calcule el momento de la fuerza en $\mathbf{r}$. ¿Cuál es el momento con respecto al origen?

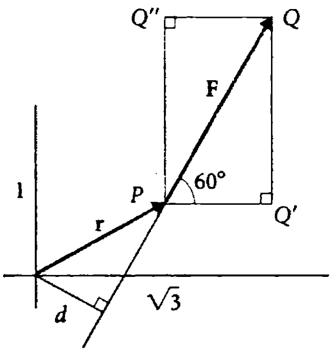


Figura 2.37

SOLUCIÓN En la figura 2.37 se ve que

$$PQ' = PQ \cos 60^\circ = 3 \cos 60^\circ = \frac{3}{2}, \qquad PQ'' = PQ \operatorname{sen} 60^\circ = 3 \operatorname{sen} 60^\circ = \frac{3\sqrt{3}}{2}$$

Así, la fuerza **F** es $(\frac{3}{2}, \frac{3\sqrt{3}}{2})$ como vector de posición, y $\mathbf{r} = (\sqrt{3}, 1)$. Para determinar el producto cruz $\mathbf{r} \times \mathbf{F}$ agregamos un tercer componente igual a cero a estos vectores. Por consiguiente,

$$\mathbf{m} = \mathbf{r} \times \mathbf{F} = (\sqrt{3}, 1, 0) \times \left(\frac{3}{2}, \frac{3\sqrt{3}}{2}, 0\right) = (0, 0, 3)$$

es el momento, y $\|(0, 0, 3)\| = 3$ es su magnitud.

También es posible calcular el momento como $3d$, determinando la distancia d del origen a la recta definida, mediante triángulos, por **F**. □

En la práctica, frecuentemente se adjudica un **signo** al momento. El signo de m es positivo si la fuerza tiende a producir rotación en sentido contrario a las manecillas del reloj, respecto al punto dado, y es negativo en caso contrario. Obsérvese que para un sistema derecho, el momento (con signo) es el tercer componente del momento, si el plano xy está en el plano definido por **F** y **r**.

Ya estamos preparados para describir las condiciones bajo las cuales un cuerpo rígido está en equilibrio, con fuerzas coplanares actuando sobre él.

Condiciones de equilibrio para fuerzas coplanares

Cuando las fuerzas coplanares actúan sobre un cuerpo rígido, éste se encontrará en equilibrio si se satisface lo siguiente:

1. La suma vectorial de todas las fuerzas es cero.
2. La suma algebraica de los momentos con signo de todas las fuerzas respecto a cualquier punto del plano es cero.

La segunda condición equivale a: la suma vectorial de los momentos de todas las fuerzas relacionadas con cualquier punto en el plano es igual a cero.

■ EJEMPLO 78 (Equilibrio) El extremo superior de una barra PQ uniforme, de 5 pies de longitud y que pesa 50 lb, descansa recargado en un muro vertical liso, figura 2.38(a). El extremo inferior descansa apoyado en un piso horizontal liso, a 3 pies del muro. Una cuerda OR sujeta al sistema en equilibrio. Si la distancia RQ es 1 pie, ¿cuál es la tensión de la cuerda?

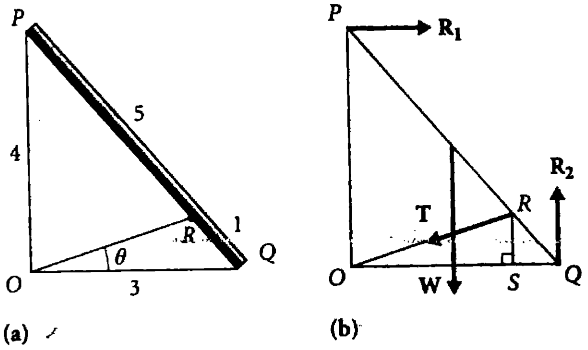


Figura 2.38

SOLUCIÓN Sean $\mathbf{R}_1$ y $\mathbf{R}_2$ las reacciones del muro y del piso, y sea $\mathbf{T}$ la tensión en la cuerda figura 2.38(b). Sean R_1, R_2 y T las magnitudes de esos vectores. La suma vectorial de $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{T}$ y el peso $\mathbf{W}$ debe ser cero. Por consiguiente, en términos de componentes,

$$R_1 - T\cos\theta = 0$$

$$-50 + R_2 - T\,\text{sen}\,\theta = 0$$

A partir de la segunda condición de equilibrio, podemos obtener una tercera ecuación. Igualando a cero la suma algebraica de los momentos respecto a O, se obtiene

$$-4R_1 + 3R_2 - \frac{3}{2}\cdot 50 = 0$$

A continuación se despeja T del sistema, para obtener

$$T = \frac{75}{4\cos\theta - 3\,\text{sen}\,\theta}$$

Ya que

$$\frac{RS}{4} = \frac{QS}{3} = \frac{1}{5}$$

se tiene $OS = 3 - QS = \frac{12}{5}$ y $OR = \sqrt{OS^2 + RS^2} = \sqrt{\left(\frac{12}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = \frac{4}{5}\sqrt{10}$. Por lo que

$$\cos\theta = \frac{OS}{OR} = \frac{3}{\sqrt{10}} \quad \text{y} \quad \text{sen}\,\theta = \frac{RS}{OR} = \frac{1}{\sqrt{10}}$$

Y entonces

$$T = \frac{75}{4(3/\sqrt{10}) - 3(1/\sqrt{10})} = \frac{25}{3}\sqrt{10} \cong 26.35 \text{ lb}$$

Ejercicios 2.8

Promediando

1. Grafique la sucesión y emplee matrices para promediarla dos veces. Trace cada vez que promedie.

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 3 & 7 & 2 & 3 & 9 & 1 & 10 \end{array}$$

2. Grafique la sucesión y promédiela dos veces con matrices; después grafique cada vez que promedie.

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 25 & 15 & 45 & 15 & 20 & 30 & 20 & 50 \end{array}$$

Sistemas dinámicos discretos

Suponga que una especie está formada por dos grupos de edades: los jóvenes y los adultos. Sean Y_k y A_k las cantidades de individuos después de k unidades de tiempo. Los jóvenes tienen una tasa de supervivencia s. La tasa de natalidad de este grupo es y (es decir, un individuo joven tiene y descendientes por unidad de tiempo), y la de los adultos es a. En los ejercicios 3 a 8 escriba, en forma matricial, el sistema dinámico que modela a la población total. Calcule la cantidad de individuos de cada grupo después de haber transcurrido 3 unidades de tiempo, para los valores dados de s, y, a, Y_0, y también para $A_0 = 100$.

3. $s = 4/5, y = 2, a = 10, Y_0 = 100.$
4. $s = 1/2, y = 2, a = 6, Y_0 = 100.$
5. $s = 1/2, y = 4, a = 10, Y_0 = 100.$
6. $s = 1/4, y = 2, a = 12, Y_0 = 100.$
7. $s = 1/3, y = 3, a = 12, Y_0 = 300.$
8. $s = 1/5, y = 5, a = 30, Y_0 = 100.$

9. Una población de moscas se divide en tres grupos de edades: A, B y C. En el grupo A se encuentran las moscas de 0 a 2 semanas de edad, en el B están las de 2 a 4 semanas, y en el C las de 4 a 6 semanas. Suponga que los grupos tienen A_k, B_k y C_k cantidades de moscas al final de la $2k$-ésima semana. La tasa de supervivencia del grupo A es 25%, mientras que la del grupo B es $33.\overline{33}\%$. Cada mosca del grupo A procrea 0.25 descendientes, cada una del grupo B tiene 2.5 descendientes, y cada una del grupo C, 1.5. Si la población original es de 4 800 moscas en cada grupo de edad, represente en forma matricial el sistema dinámico que modela a esta población. Calcule la cantidad de moscas en cada grupo al final de 6 semanas.

Geometría y física

10. Calcule la distancia del punto $P(8, 4, -5)$ al plano $2x - 2y + z - 6 = 0$.
11. Determine la distancia del punto $P(1, 1, 6)$ al plano $2x - 2y + z - 6 = 0$.
12. Calcule la distancia del punto $P(1, 2, -4)$ al plano que pasa por los puntos $(0, 0, 0)$, $(2, 1, 0)$ y $(3, 0, -1)$.

Si P, Q, R y S son los puntos

$$P(0,-1,2) \qquad Q(2,-2,1)$$
$$R(4,-2,1) \qquad S(0,2,-3)$$

En los ejercicios 13 a 15 calcule la distancia más corta entre las dos líneas sesgadas determinadas por los segmentos de recta especificados.

13. $\overrightarrow{PQ}$ y $\overrightarrow{RS}$
14. $\overrightarrow{PS}$ y $\overrightarrow{RQ}$
15. $\overrightarrow{PR}$ y $\overrightarrow{QS}$

16. Determine la distancia más corta entre las dos rectas cuyas ecuaciones paramétricas son:

$$l_1:\ \begin{aligned} x &= 5 - 3t \\ y &= -4 + 2t \\ z &= 2 - t \end{aligned} \qquad l_2:\ \begin{aligned} x &= 5 - s \\ y &= -2 - 2s \\ z &= 1 + s \end{aligned}$$

17. Sea AP el segmento de recta que biseca la hipotenusa BC del triángulo rectángulo ABC. Mediante productos vectoriales demuestre que $AP = \frac{1}{2} BC$.
18. **(Centroide)** El centroide de n puntos (de un polígono de n vértices) formado por las puntas de los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ está definido por

$$\frac{1}{n}(\mathbf{v}_1 + \cdots + \mathbf{v}_n)$$

(figura 2.39). Determine el centroide del triángulo PQR, donde $P(1, 2)$, $Q(2, -4)$ y $R(-1, 7)$.

19. De acuerdo con el ejercicio 18, determine el centroide de $P_1(0, 0, 0)$, $P_2(1, 1, 1)$, $P_3(1, 1, -1)$ y $P_4(-2, 1, 0)$.
20. Compruebe que el centroide de cualquier triángulo está en la intersección de sus tres medianas.
21. **(Centro de masa)** Si $m_1, \ldots, m_n$ n son las masas ubicadas en las puntas de los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y si $M = m_1 + \cdots + m_n$ es la masa total, entonces el centro de masa de esos sistemas se define con

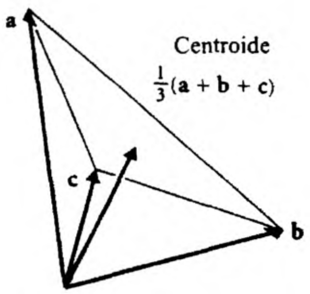


Figura 2.39

Determine el centro de masa del sistema cuyas masas de 1, 4, 5 y 2 kg están ubicadas, respectivamente, en $P_1(-1, 2, 0)$, $P_2(0, 5, -1)$, $P_3(1, 1, -3)$ y $P_4(-6, 1, -3)$.

22. A propósito de los ejercicios 18 y 21, demuestre que si todas las masas son iguales, su centro equivale al centroide.

En los ejercicios 23 y 24 calcule el trabajo efectuado por la fuerza constante **F**, cuando su punto de aplicación se mueve de X a Y.

23. $\mathbf{F} = \mathbf{i} + \mathbf{j} + \mathbf{k}, X = (0, -2, 5), Y = (1, 7, -2)$.

24. $\mathbf{F} = -2\mathbf{i} + 6\mathbf{j} + 8\mathbf{k}, X = (-2, 3, 0), Y = (-1, 6, -4)$

25. De acuerdo con la figura 2.40, calcule la fuerza **F** en la cuerda que equilibra el objeto de 25 lb.

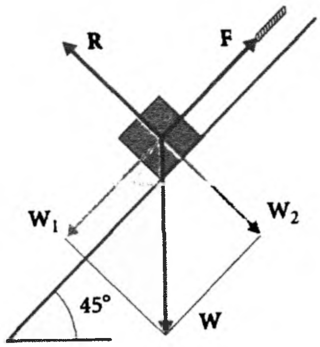


Figura 2.40

26. Calcule el momento de $\mathbf{F} = -2\mathbf{i} - 4\mathbf{j} + \mathbf{k}$ en $\mathbf{r} = (-2, 1, -3)$.

2.9 Miniproyectos

1 ■ Comportamiento de un sistema dinámico a largo plazo

En este proyecto veremos que, bajo condiciones iniciales especiales, es posible predecir una parte del comportamiento de un sistema dinámico discreto a largo plazo. En el capítulo 7 regresaremos a este importante tema.

Un modelo de población

Suponga que una especie incluye dos grupos de edad: los jóvenes y los adultos. A_k y B_k son las cantidades de individuos después de k unidades de tiempo. Los jóvenes tienen tasa de supervivencia igual a $\frac{6}{7}$. Su tasa de natalidad es de 3 (es decir, un individuo joven tiene 3 descendientes por unidad de tiempo) y la de los adultos es 21.

Problema A

1. Escriba en forma matricial el sistema dinámico que modela a esta población.
2. Deduzca una fórmula, en función de A_0 y B_0, que calcule la cantidad de individuos en cada grupo, después de 3 unidades de tiempo.
3. Evalúe su fórmula para $A_0 = 700$ y $B_0 = 700$.

Problema B

Refiérase al problema A.

1. Si $A_0 = 7$ y $B_0 = 1$, y p_k es la relación $A_k : B_k$, calcule el valor p a largo plazo, que es el de p_k cuando k se incrementa. Esto es, determine $p = \lim_{k \to \infty} p_k$. Justifique su respuesta. (No es necesario conocer lo relativo a límites.)
2. Supongamos ahora que $A_0 = 8$ y $B_0 = 2$; que q_k es la relación $A_k : B_k$, y q es el valor de q_k a largo plazo. ¿Es fácil predecir q esta vez? ¿Por qué si o por qué no?
3. Es un hecho que $p = q$. Encuentre el primer valor de k en q_k de manera que q_k esté dentro de 0.5 de p.

Problema C

Considere el sistema dinámico cuya forma matricial es

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \\ C_{k+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{5}{2} & \frac{3}{2} \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} A_k \\ B_k \\ C_k \end{bmatrix}$$

1. Formule un problema de crecimiento de población que se pueda modelarse con este sistema:
2. Determine las relaciones $A_k : B_k : C_k$ a largo plazo si $\begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix}$ es (a) $\begin{bmatrix} 24 \\ 6 \\ 2 \end{bmatrix}$ y (b) $\begin{bmatrix} 24 \\ 5 \\ 1 \end{bmatrix}$.
3. Explique qué significan estas relaciones a largo plazo en cada caso.

2 ■ Plano bisector de dos planos; recta bisectriz de dos rectas

En este proyecto debe lograr lo siguiente:

1. Desarrollar un método para calcular la ecuación del plano $\mathcal{P}$ que biseca la intersección de dos planos, $\mathcal{P}_1$ y $\mathcal{P}_2$.
2. Deducir una fórmula para calcular la distancia de un punto a una recta.
3. Obtener una fórmula para las rectas bisectrices interna y externa de dos rectas en el plano.

Con respecto al primer objetivo, si conocemos las ecuaciones de $\mathcal{P}_1$ y $\mathcal{P}_2$, podemos determinar un punto de intersección y, en consecuencia, un punto de $\mathcal{P}$. Así, con la fórmula punto-normal sólo necesitamos una normal a $\mathcal{P}$. Esta última debe bisecar los ángulos entre las normales de $\mathcal{P}_1$ y $\mathcal{P}_2$, figura 2.41(a), por lo que basta con determinar un vector que biseque el ángulo entre los dos vectores dados. De los muchos vectores que lo hacen, calcularemos uno con longitud unitaria.

Problema A

Si $\mathbf{n}_1$ y $\mathbf{n}_2$ son los dos vectores dados. Y si $\mathbf{u} = (u_1, u_2, u_3)$ es un vector unitario que biseca el ángulo entre $\mathbf{n}_1$ y $\mathbf{n}_2$. Demuestre que

$$\left(\frac{\mathbf{n}_1}{\|\mathbf{n}_1\|} - \frac{\mathbf{n}_2}{\|\mathbf{n}_2\|} \right) \cdot \mathbf{u} = \mathbf{0} \tag{2.36}$$

$$\mathbf{u} \cdot (\mathbf{n}_1 \times \mathbf{n}_2) = \mathbf{0} \tag{2.37}$$

$$u_1^2 + u_2^2 + u_3^2 = 1 \tag{2.38}$$

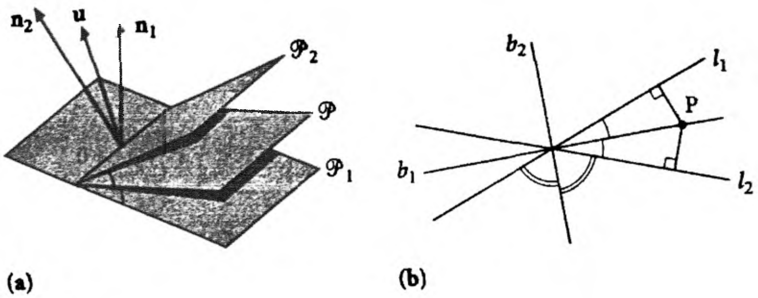


Figura 2.41

Problema B

Aplique el resultado del problema A para determinar un vector unitario que biseque el ángulo que forman $\mathbf{n}_1 = (1, 1, 1)$ y $\mathbf{n}_2 = (1, 0, 0)$. Encuentre una ecuación del plano que biseca los planos $x + y + z = 1$ y al plano yz.

Los dos problemas siguientes son independientes de los dos primeros, conducen a una fórmula para determinar las bisectrices interna y externa, b_1 y b_2, de dos rectas en el plano, l_1 y l_2, figura 2.41(b).

Problema C

Si $P(x_0, y_0)$ es un punto y $ax + by + c = 0$ es la ecuación de una recta l en el plano xy. Use el producto punto para encontrar una fórmula para la distancia más corta de P a l. ¿Cuál es la distancia más corta de $P(1, 1)$ a $3x - y + 2 = 0$?

Problema D

Sean $a_1x + b_1y + c_1 = 0$ y $a_2x + b_2y + c_2 = 0$ las ecuaciones de dos rectas, l_1 y l_2, en el plano. Utilice el resultado del problema C para llegar a una fórmula de las dos rectas que bisecan los ángulos entre l_1 y l_2. (*Sugerencia:* Escoja un punto P en una bisectriz, figura 2.41(b), y trace las distancias perpendiculares de P a l_1 y l_2.) Determine las dos bisectrices de las rectas $3x - y + 2 = 0$ y $2x - 3y + 1 = 0$.

3 ■ Criterio de colinealidad y un teorema de Papo

En este proyecto se usarán los productos cruz para demostrar una condición necesaria, suficiente y sencilla, para que tres vectores coplanares sean colineales. A continuación aplicaremos este criterio para resolver un problema llamado *caso paralelo de un teorema de Papo*.[9]

---

[9] **Papo de Alejandría** fue el último de los grandes geómetras griegos. (320 d.C.) Escribió *Synagogue* (colección), un gran tratado de geometría en 10 volúmenes. Además de la riqueza de información acerca de los descubrimientos de Arquímedes, en *Synagogue* aparecen nombres y trabajos importantes de otros matemáticos griegos, que de otra manera se desconocerían. Muchos de los teoremas en este tratado se deben al mismo Papo.

TEOREMA 27

(Criterio de colinealidad)

Tres vectores coplanares $\mathbf{a} = (a_1, a_2, 0)$, $\mathbf{b} = (b_1, b_2, 0)$ y $\mathbf{c} = (c_1, c_2, 0)$ son colineales si y sólo si

$$\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a} = \mathbf{0}$$

Problema A

Demuestre el criterio de colinealidad y demuestre que equivale a la fórmula

$$\begin{vmatrix} a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \\ c_1 & c_2 & 1 \end{vmatrix} = 0$$

Problema B

Aplique el criterio de colinealidad para demostrar lo siguiente.

TEOREMA 28

(Caso de las paralelas del teorema de Papo)[10]

Si P, Q, R y P', Q', R', son triadas de puntos colineales, si QR' es paralela a $Q'R$ y si RP' es paralela a $R'P$, la recta PQ' es paralela a la recta $P'Q$ (figura 2.42).

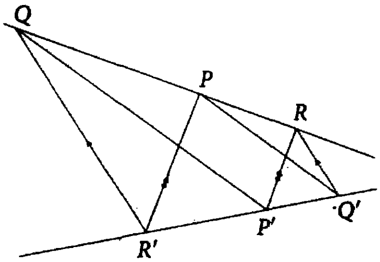


Figura 2.42

Sugerencia: Elija un origen O y, usando la notación $OP = \mathbf{p}$, $OP' = \mathbf{p}'$, etc., demuestre que

$$(\mathbf{p} - \mathbf{q}') \times (\mathbf{p}' - \mathbf{q}) + (\mathbf{q} - \mathbf{r}') \times (\mathbf{q}' - \mathbf{r}) + (\mathbf{r} - \mathbf{p}') \times (\mathbf{r}' - \mathbf{p}) = \mathbf{0}$$

4 ■ Teorema de Varignon

Sean $\mathbf{F}_1$ y $\mathbf{F}_2$ dos fuerzas con el mismo origen P, y sea O cualquier punto en el plano definido por $\mathbf{F}_1$ y $\mathbf{F}_2$ (figura 2.43).

[10] Véase *Geometry, a Comprehensive Course*, de Dan Pedoe (Nueva York: Dover, 1988), p. 47.

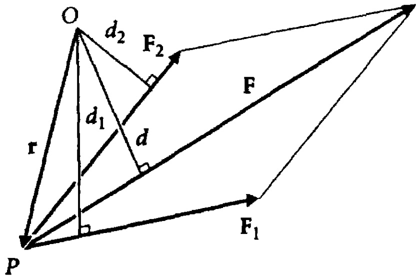


Figura 2.43

Problema A

1. Aplique las propiedades del producto cruz para demostrar el teorema de Varignon,[11] que establece: la suma algebraica de los momentos con signo de $\mathbf{F}_1$ y $\mathbf{F}_2$ respecto a O es igual al momento con signo de la resultante $\mathbf{F}_1 + \mathbf{F}_2$ respecto a O.
2. Vuelva a demostrar el teorema de Varignon, esta vez geométricamente.

Problema B

Se tiene la barra uniforme PQ de peso W que toca el piso y un plano, figura 2.44(a) y (b). Demuestre que la fuerza horizontal F que debe aplicarse en P para equilibrar la barra es

$$F = \frac{1}{2} W \cot \theta \quad \text{o bien} \quad F = \frac{1}{2} \frac{W}{\tan \theta + \cot \phi}$$

dependiendo de si el plano es vertical o forma un ángulo ϕ con la horizontal.

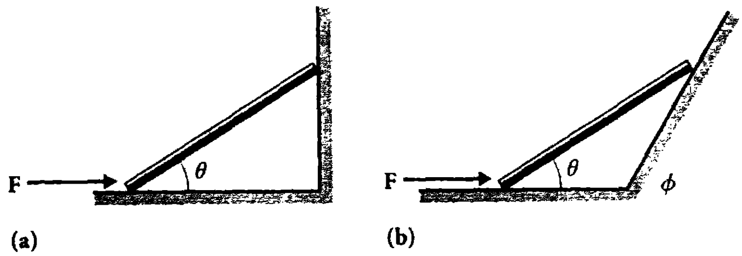


Figura 2.44

[11] **Pierre Varignon** (1654-1722), francés, era contemporáneo de Newton. Descubrió el principio de los momentos. Girwin, en su *A Historical Appraisal of Mechanics*, le acredita haber sido el primero en deducir las ecuaciones del movimiento.

2.10 Ejercicios en computadora

Esta sección para computadora ayudará al estudiante en el aprendizaje de la manipulación de vectores y matrices, y a efectuar operaciones vectoriales básicas con sus programas de matemáticas. También le auxiliará en el repaso de algunos temas del material básico de este capítulo.

Sean

$$\mathbf{u} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} 2 \\ -3 \\ 1 \\ -4 \end{bmatrix}$$

$$M = \begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 2 \\ 4 & 6 & 8 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$$

1. Determine

   a. $\mathbf{u} + \mathbf{v}$ b. $\mathbf{u} - \mathbf{v}$ c. $10\mathbf{u}$ d. $\mathbf{u} - 2\mathbf{v} + 3\mathbf{w}$

2. Compruebe las identidades

   a. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ b. $10(\mathbf{u} + \mathbf{v}) = 10\mathbf{u} + 10\mathbf{v}$

3. Si es posible, escriba $\mathbf{v}$ como una combinación lineal de las columnas de M.
4. Si es posible, exprese $\mathbf{w}$ como una combinación lineal de las columnas de N.
5. Grafique $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ por separado, y también juntas en la misma figura.
6. Determine

   a. $\|\mathbf{u}\| + \|\mathbf{v}\|$ b. $\|\mathbf{u} + \mathbf{v}\|$ c. $\mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}$

7. Encuentre el ángulo que forman $\mathbf{u}$ y $\mathbf{v}$.
8. Calcule la proyección ortogonal de $\mathbf{v}$ sobre $\mathbf{w}$. Compruebe su respuesta.
9. ¿Cuáles de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ están en el generador de las columnas de M?
10. ¿Cuáles de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ están en el generador de las columnas de N?
11. ¿Es linealmente independiente el conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$? ¿Genera a $\mathbf{R}^3$?
12. Las columnas de M, ¿generan a $\mathbf{R}^3$? ¿Son linealmente independientes?
13. Las columnas de N, ¿generan a $\mathbf{R}^3$? ¿Son linealmente independientes?
14. ¿Cierto o falso?

    a. $\{M\mathbf{x}, \mathbf{x} \in \mathbf{R}^3\} = \mathbf{R}^3$ b. $\{N\mathbf{x}, \mathbf{x} \in \mathbf{R}^4\} = \mathbf{R}^3$

15. Escriba la cuarta columna de N en forma de una combinación lineal de las primeras tres.
16. Demuestre que el sistema de abajo es consistente para todos los valores de b_1, b_2 y b_3.

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 &= b_1 \\ x_1 + 2x_2 + 2x_3 + x_4 &= b_2 \\ x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= b_3 \end{aligned}$$

17. Dadas las ecuaciones anteriores, encuentre una solución para $b_1 = 1$, $b_2 = -1$ y $b_3 = 1$, y compruebe su respuesta revisando la ecuación vectorial correspondiente.

18. Demuestre que las columnas de N son linealmente independientes y generan a $\mathbf{R}^3$. ¿Cómo se afecta esta doble propiedad si el lector agrega una columna de su elección? ¿Qué sucede si omite una columna de su elección?

19. Si es posible, calcule e interprete los productos $M\mathbf{u}$, $M\mathbf{r}$, $N\mathbf{u}$ y $N\mathbf{r}$.

20. Compruebe que $\mathbf{u}$ y $\mathbf{u} \times \mathbf{v}$ son ortogonales.

21. Verifique la **identidad de Jacobi**: $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} + (\mathbf{v} \times \mathbf{w}) \times \mathbf{u} + (\mathbf{w} \times \mathbf{u}) \times \mathbf{v} = \mathbf{0}$.

22. Sea l la recta cuya ecuación paramétrica es

$$\mathbf{x} = (-1, 2, 1) + t(-4, 1, 5)$$

¿Cuál de $R(-9, 4, 11)$ y $S(7, 0, -10)$ está en l? Grafique l desde $t = -2$ hasta $t = 3$. Grafique l desde $P(-13, 5, 16)$ hasta $Q(-21, 7, 26)$.

23. Encuentre una forma normal-punto de la ecuación para el plano que pasa por $P(1, 2, -3)$, $Q(-2, 4, 5)$ y $R(3, 3, 3)$. Grafique este plano. Determine dos puntos: uno en el plano y el otro fuera de él.

24. Escriba y pruebe el código para una función que calcule la distancia de un punto a un plano.

25. Exprese y pruebe el código para una función que calcule la distancia entre dos líneas sesgadas, dados dos puntos en cada recta.

Soluciones seleccionadas con Maple

```
# DATOS  y observaciones.

With(linalg):                 # Cargar el paquete linalg.
u := vector([1, 3, 2]);       (# Definición de u, v, w como vectores.
v := vector ([-1, 1, 2]);     # O bien  u := array([1, 2, 3]); , etc..
w := vector([2, 1, -4]);      # Los vectores no son matrices columna. Un elemento de vector
                              # se determina con un número; un elemento de matriz con dos.
u;                            # Muestra sólo el nombre y no el valor.
evalm(u);                     # Ahora muestra el valor. También eval(u), pero es limitado.
M := matrix(3, 3, [1,3,5, 7,9,2, 4,6,8]);         # M.
# O bien M := matrix([[1,3,5],[7,9,2],[4,6,8]]);
# O bien M := matrix(3,3,[[1,3,5],[7,9,2],[4,6,8]]);
M;                            # Igual que con los vectores, muestra el nombre de la matriz.
eval(M)                       # Ahora se evalúa M.
N := matrix(3,4 [1,2,3,4,2,3,4,5,3,4,5,6]);

# Ejercicios 1, 2.

evalm(u+v);                   # También add(u,v);
evalm (u-v);                  # También add(u,-v); pero eval(u-v); no hará la operación aritmética.
evalm(10*u);                  # También scalarmul(u,10);
evalm (u-2*v+3*w);            # Combinación lineal.
equal((u+v)+w, u+(v+w));      # Comprobación de la igualdad.
equal(10*(u+v), 10*u+10*v);   # Igual.

# Ejercicios 2, 3.

am := augment (M, v);               # La matriz aumentada [M:v].
rm := rref (am);                    # Reducción: la última columna no es pivote.
                                    # El sistema es consistente, entonces v es comb. lin.
                                    # en las columnas de M. Los coeficientes de
```

```
evalm(rm[1,4]*col(M, 1)+rm[2, 4]*\  # la comb. lin. son elementos de la última columna
col(M, 2)+rm[3,4]*col(M,3));        # de rm. En realidad, al calcular la comb. lin. se obtiene v.
an := augment(N, w);                # [N:w].
rref(an);                           # La última columna es pivote. No hay soluciones.
                                    # w no es una comb. lin. en las columnas de N.

# Ejercicios 5-8

with(plots):
o:=vector([0, 0, 0]):               # El origen.
p1:=polygonplot3d([o, u]):          # Definir y nombrar los segmentos de recta, pero
p2:=polygonplot3d([o, v]):          # todavía no mostrar las gráficas. Observe que
p3:=polygonplot3d([o, w]):          # polygonplot3d([o,u],axes=boxed); no muestra a u.
display(p1, axes=boxed);            # Mostrar p1 en ejes entre [ ]. Repetir con p2 y p3.
display([p1, p2, p3] axes=boxed)    # Mostrar todo.
with(linalg):                       # La longitud de (1,2,3) es
norm(u, 2)+norm(v, 2);              # norm([1,2,3],2) o bien norm(vector([1,2,3]),2);
evalf(");                           # o bien norm([1,2,3],frobenius);
norm(u+v, 2); evalf(");             # norm(vector[1,2,3],frobenius);
dotprod(u, v)-dotprod(u, w);        # dotprod(u,v) es el producto punto de u por v.
angle(u, v); evalf(");
pr:=evalm(dotprod(v, w)/dotprod(w, w) * w);   # Proyección ortogonal de v sobre w.
vc:=evalm(v-pr);                    # The vector component of v orthogonal to w.
dotprod(pr, vc);                    # El producto punto es cero, y la
evalm(pr+vc);                       # suma es v, como se esperaba.

# Ejercicio 9 – Parcial.

rref (augment (M, [1, 3, 2]));      # La última columna no es pivote, así que u está en el espacio genera-
                                    do por un conjunto de vectores.

# Ejercicio 11

concat(u, v, w); rref (");          # [u v w] tiene 3 pivotes, así que los vectores son independientes.

# Ejercicio 15

rref (N);                           # De la última columna: (-2)xcol1+3xcol2 = col4.

# Ejercicio 16 – Sugerencia: Reducir la matriz de coeficientes con operaciones de renglón para ver que
# su última columna es pivote y que la última columna de la matriz aumentada no sea columna pivote.

# Ejercicio 19.

evalm(M &* vector([1, 3, 2]));      # Etc.. Mr y Nu son indefinidas.

# Ejercicios 20, 21.

u:=[1, 2, 3]: v:=[-1, -1, 1]: w:=[2, 1, -4]:
with(linalg):
uv := crossprod(u, v);              # En lugar de listas [,,,] se podrían
dotprod(u, uv);                     # usar vectores, vector([,,,]).
evalm(crossprod(crossprod(u, v), w)+crossprod(crossprod(v, w), u)+
crossprod(crossprod(w, u), v));

# Ejercicio 22.

solve({-1, -4*t=-9, 2+t=4, 1+5*t=11}, t);     # t=2 para R, de modo que Q está en 1.
solve({-1-3*t=7, -2+3*t=0, 2-5*t=-10}, t)     # El sistema no tiene solución, así que S no está en l.
with(plots):                                  # Cargar paquete de gráficas para usar
spacecurve([-1-4*t, 2+t, 1+5*t], t=-2..3);    # el comando spacecurve. Etc. .
```

```
# Ejercicio 25.

LineToLine := proc (p, q, r, s) local u, v, w, cr;             (* Code. *)
u:=q-p; v:=s-r; w:=r-p; cr := linalg [crossprod] (u, v);
abs(linalg[dotprod] (w, cr)) / linalg[norm] (cr, 2) end:
LineToLine([1, -2, -1], [0, -2, 1], [-1, 2, 0], [-1, 0, -2]);  (* Testing. *)
```

Soluciones seleccionadas con Mathematica

```
(* DATOS y observaciones *)
<<LinearAlgebra 'MatrixManipulation';   (* Cargar el paquete con funciones matriciales.*)
(* Problem A.*)                (* Definición de u, v, w como vectores columna, o bien *)
u = {{1}, {3}, {2}}            (* como matrices columna. u = {1,2,3}, etc. las*)
v = {{-1}, {1}, {-2}}          (* definiría como "vectores renglón". *)
w = {{2}, {1}, {-4}}
M = {{1, 3, 5}, {7, 9, 2}, {4, 6, 8}}     (*M.*)
n = {{1, 2, 3, 4}, {2, 3, 4, 5}, {3, 4, 5, 6}}
                               (* N. El símbolo N se reserva para evaluación numérica*)
(* Para mostrar una matriz m en forma renglón-columna usar MatrixForm[m]. *)

(* Ejercicios 1,2. *)

u + v                          (* Suma. *)
u - v                          (* Diferencia. *)
10 u                           (* Producto por escalar. *)
u - 2 v + 3 w                  (* Combinación lineal. *)
(u+v)+w===u+(v+w)    (* Comprobación de igualdad. También, igualQ[(u+v)+w,u+(v+w)] *)
10 (u+v) === 10 u + 10 v       (* Igual. *)

(* Ejercicios 2, 3. *)

am = AppendRows [M, v)             (* La matriz aumentada [M:v].    *)
rm = RowReduce [am]                (* Reducción: La última columna no es pivote.   *)
                                   (* El sistema es consistente, así que v es comb. lin. *)
                                   (* en las columnas de M. Los coeficientes de   *)
rm[[1, 4]] TakeColumns [M, {1}]+   (* la comb. lin. son elementos de la última columna   *)
rm[[2, 4]] TakeColumns [M, {2}]+   (* de rm. En realidad, al calcular la comb. lin. se obtiene v. *)
rm[[3, 4]] TakeColumns [M, {3}]
an = AppendRows [n, w]             (* [N:w].    *)
RowReduce[an]                      (* La última columna es pivote. No hay soluciones.   *)
                                   (* w no es comb. lin. de las columnas de N.    *)
(* Ejercicios 5-8. *)

o={0, 0, 0}; u ={1, 3, 2}; v={-1, 1, 2}; w={2, 1, -4};
p1=Line[{o, u}];           (* Definir y nombrar los segmentos de recta, pero *)
p2=Line[{o, v}];           (* todavía no mostrar las gráficas.    *)
p3=Line[{o, w}];                      (* A continuación mostrar p1 con ejes identificados. *)
Show[Graphics3D [p1, Axes->True]]     (* A continuación repetir con p2 y p3. . .   *)
Show[Graphics3D[{p1, p2, p3}, Axes->True]]  (* Mostrar todo.        *)
(* A continuación hay que definir la función norma-longitud-magnitud propia. Comenzando *)
(* con una lista-vector a, a^2 es la lista cuyos elementos son cuadrados de los elementos de a.   *)
(*Entonces aplicar (@@ o Apply) Plus a a^2 para sumar todos sus *)
(* componentes y por último sacar raíz cuadrada de la suma.  *)
Norm[a_]:=Sqrt[Plus@@(a^2)]
Norm[u] + Norm[v]
N[%]
```

```
Norm[u+v]
N[%]
u . v - u . w              (* El producto punto de u por v es u.v, o punto[u,v].   *)
(* Definir la propia función de ángulo.     *)
Angle[a_, b_] :=ArcCos[(a. b)/(Norm[a] Norm[b])]
N[Angle[u, v]]
pr = (v.w / w.w) w       (* Usar la fórmula para la proyección ortogonal, *)
(* o bien usar la función Projection incorporada en Mathematica en el   *)
(* paquete LinearAlgebra 'Orthogonalization'.        *)
<<LinearAlgebra 'Orthogonalization';
pr=Projection[v, w]
vc = v - pr                        (* La componente vectorial de v ortogonal a w.   *)
pr . vc                            (* El producto punto es cero y la    *)
pr + vc                            (* suma es v, como se esperaba.     *)

(* Ejercicio 9 - parcial.  *)

RowReduce [AppendRows [M, {{1}, {3}, {2}}]]   (* La última columna no es pivote, así que u *)
                                              (* está en el generador. *)

(* Ejercicio 11.  *)

AppendRows [{{1}, {3}, {2}}, {{-1}}, {1}, {2}}, {{2}, {1}, {-4}}]
RowReduce [%]          (* [u v w] tiene 3 pivotes, así que los vectores son Independiente. *)

(* Ejercicio 15. *)

RowReduce [n]                   (* De la última columna: (-2)xcol1+3xcol2 = col4. *)

(* Ejercicio 16 - Sugerencia: Reducir la matriz de coeficientes con operaciones de renglón para verificar *)

(* que su última columna sea pivote, para que la última columna de la matriz aumentada *)
(* no sea columna pivote.  *)

(* Ejercicio 19.  *)

M . {{1}, {3}, {2}}                         (*. Etc. .   Mr y Nu son indefinidas. *)

(* Ejercicios 20, 21. *)

u={1, 2, 3}; v={-1, -1, 1}; w={2, 1, -4};
<<LinearAlgebra 'CrossProduct';
uv = Cross [u, v]
u . uv
Cross [Cross [u, v], w]+Cross [Cross [v, w], u]+Cross [w, u], v]

(* Ejercicio 22. *)

Solve [{-1-4*t==-9, 2+t==4, 1+5*t==11}, t]   (* t=2, de modo que R está en 1. *)
Solve [{-1-4*t==7, 2+t==0, 1+5*t==-11}, t]   (* No hay solución. S no está en 1. *)
ParametricPlot3D [{-1+3*t, -2+3*t, 2-5*t}, {t, -6, 7}]

(* Ejercicio 25. *)

LineToLine[p_, q_, r_, s_] := Module [{u, v, w, cr},
                     u=q-p; v=s-r; w=r-p; cr = Cross[u, v];
                     Abs[w.cr] / Sqrt [cr[[1]]^2+cr[[2]]^2+cr[[3]]^2] ]
LineToLine[{1, -2, -1}, {0, -2, 1}, {-1, 2, 0}, {-1, 0, -2}]       (* Prueba *)
```

Soluciones seleccionadas con MATLAB

```
% DATOS
u = [1; 3; 2]                          % Definición de u, v, w.
v = [-1; 1; 2]
w = [2; 1; -4]
M = [1 3 5; 7 9 2; 4 6 8]              % M.
N = [1 2 3 4; 2 3 4 5; 3 4 5 6]        % N.

% Ejercicios 1, 2

u + v                                  % Suma
u - v                                  % Diferencia.
10 * u                                 % Producto por escalar.
u-2*v+3*w                              % Combinación lineal.
(u+v)+w==u+(v+w)                       Comprobación de igualdad. Muestra
10*(u+v)==10*u+10*v                    % % 1 (= CIERTO) para cada elemento.
% Ejercicios 3, 4.
am = [M v]                        % La mátriz aumentada [M:v].
rm = rref (am)                    % Reducción: la última columna no es pivote.
                                  % El sistema es consistente, de modo que v es comb. lin.
                                  % de las columnas de M. Los coeficientes de
rm(1, 4)*M(:,1)+rm(2, 4)*. . .    % la comb. lin. son elementos de la última columna
M(:,2)+rm(3, 4)*M(:,3)            % de rm. En realidad, al calcular la comb. lin. se obtiene v.
an = [N w]                        % [N:w].
rref (an)                         % La última columna es pivote. No hay soluciones.
                                  % w no es comb. lin. de las columnas de N.

% Ejercicios 5-8.

o=[0 0 0]; u=[1 3 2]; v=[-1 1 2]; w=[2, 1 -4]
plot3[0 1], [0 3], [0 2]                       % Vectores de posición u, v, w.
plot3[0 -1], [0 1], [0 2]                      %
plot3[0 2], [0 1], [0 -4], grid                % grid agrega un retículo a la gráfica.
plot3[0 1 0 -1 0 2], [0 3 0 1 0 1], [0 2 0 2 0 -4]   % u, v, w juntas.
norm(u) + norm(v)
norm(u+v)
dot(u, v) - dot(u, w)                          % dot(a,b) es a.b .
acos(dot(u, v) / norm(u) / norm(v))            % Ángulo que forman u y v.
pr = (dot(v, w) / dot(w, w))*w                 % Fórmula de la proyección ortogonal.
vc = v - pr                                    % Componente vectorial de v ortogonal a w.
dot (pr, vc)                                   % El producto punto es (casi) cero y
pr + vc                                        % la suma es v, como se esperaba.

% Ejercicio 9 - Parcial.

rref ([M [1;3;2]]        % La última columna no es pivote, así que u está en el generador.

% Ejercicio 11.

rref ([u; v; w])         % [u v w] tiene 3 pivotes, así que los vectores son independientes.

% Ejercicio 15.

rref (N)                 % De la última columna: (-2)xcol1+3xcol2 = col4.

% Ejercicio 16 - Sugerencia: La matriz de coeficientes con operaciones de renglón para verificar que su
última

% columna sea pivote, y que la última columna de la matriz aumentada no sea columna pivote.
```

```
% Ejercicio 19.
M*[1;3;2]            % También: M*[2;-3;1;-4], etc. . Mr y Nu son indefinidas.
% Ejercicios 20 y 21.
u=[1 2 3]; v=[-1 -1 1]; w=[2 1 -4];
uv = cross (u, v)
cross(cross(u, v), w)+cross(cross(v, w), u)+cross(cross(w, u), v)
% Ejercicio 22.
% R está en l, porque el sistema -1-4*t=-9,2+t=4,1+5*t=11 es consistente porque
[roots([-4 -1+9]) roots([1 2-4]) roots([5 1-11])]  % muestra [2 2 2].
[roots([-4 -1-7]) roots([1 2-0]) roots([5 1-10])]  % [.2,.2,-2.2] por lo que el
t = -2:.25:3;                  % sistema no tiene solución y S no está en l.
plot3(-1-4*t, 2+t, 1+5*t)      % Graficar la recta.
% Ejercicio 25.
function [A] = LnToLn(p,q,r,s)                   % En un archivo.
        A = abs(dot(r-p, cross(q-p, s-r))) / norm(cross(q-p, s-r));
        end
LnToLn([1 -2 -1], [0 -2 1], [-1 2 0], [-1, 0 -2])  % En la sesión.
```

3

Matrices

Por lo que he logrado debo agradecer a mi industria,
mucho más que a algún talento sobresaliente.
(Julius Wilhelm) Richard Dedekind (1831-1916)

Introducción

Sylvester y Cayley

Lord Cayley[1] es uno de los fundadores de la teoría de las matrices, aunque su amigo Sylvester[2] fue quien acuñó el término *matriz*. Tanto Sylvester como Cayley son considerados entre los mejores matemáticos de sus tiempos. Sylvester fue el primer profesor del Departamento de Matemáticas en la Universidad Johns Hopkins, y fundó la prestigiosa revista *American Journal of Mathematics*.

A continuación veamos un ejemplo de lo que ocupaba a Cayley. Tres sistemas coordenados (x, y), (x', y') y (x'', y'') están conectados mediante las siguientes transformaciones

$$\begin{aligned} x'' &= x' - y' \\ y'' &= x' + y' \end{aligned} \qquad \text{y} \qquad \begin{aligned} x' &= x + 2y \\ y' &= 2x - y \end{aligned}$$

La relación entre (x, y) y (x'', y'') se describe con la sustitución

$$\begin{aligned} x'' &= x' - y' = (x + 2y) - (2x - y) = -x + 3y \\ y'' &= x' + y' = (x + 2y) + (2x - y) = 3x + y \end{aligned}$$

[1] (Sir) **Arthur Cayley** (1821–1895) nació en Surrey, Inglaterra, y estudió en la Universidad de Cambridge. Ejerció la abogacía y al mismo tiempo escribía aportaciones en matemáticas. Pocos años después de encontrar a su colega Sylvester, otro licenciado y matemático, dejó la abogacía y se dedicó de tiempo completo a las matemáticas.

[2] **James Joseph Sylvester** (1814–1897) nació en Londres, de padres judíos. Entró a la Universidad de Cambridge, pero por su religión no pudo obtener un diploma, sino hasta varios años después de haber terminado sus estudios. También ejerció la abogacía y al mismo tiempo hacía investigaciones en el campo de las matemáticas. Él y Cayley tuvieron una larga y fructífera colaboración en la *teoría de los invariantes*, campo relacionado con el álgebra lineal.

Esta transformación también puede obtenerse como sigue: si abreviamos los tres cambios de coordenadas mediante las *matrices* de los coeficientes, tenemos

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \qquad C = \begin{bmatrix} -1 & 3 \\ 3 & 1 \end{bmatrix}$$

Ahora es posible calcular C directamente de A y B por medio de $[A\mathbf{b}_1 \;\; A\mathbf{b}_2]$; es decir, C es la matriz cuyas columnas son $A\mathbf{b}_1$ y $A\mathbf{b}_2$, y donde $\mathbf{b}_1$ y $\mathbf{b}_2$ son las columnas de B. Este procedimiento se llama *multiplicación matricial* y se describe en la sección 3.1.

3.1 Operaciones matriciales

Objetivos del estudiante para esta sección

1. Aprender a efectuar la suma, la multiplicación por escalar y la multiplicación matricial.
2. Comprender las propiedades básicas de esas operaciones.
3. Resolver ecuaciones matriciales sencillas.

En esta sección presentaremos las operaciones matriciales básicas: suma, multiplicación por escalar y multiplicación matricial.

Recuérdese que una **matriz** A de $m \times n$ es un arreglo rectangular de $m \cdot n$ números en forma de m **renglones** horizontales y n **columnas** verticales.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

El número a_{ij} es el **(i,j)-ésimo elemento** de A. El **i-ésimo renglón** y la **j-ésima columna** de A son

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

respectivamente. También podemos considerar que la matriz A es una secuencia de sus columnas $\mathbf{a}_1, \ldots, \mathbf{a}_n$:

$$A = [\mathbf{a}_1 \; \mathbf{a}_2 \cdots \mathbf{a}_n]$$

Una matriz de $1 \times n$ se llama **matriz renglón**. Una matriz de $m \times 1$ se denomina **matriz columna**, o **vector**. Una matriz de $n \times n$ se llama **matriz cuadrada** (que también se definió en la sección 1.3). Esta última tiene igual cantidad de renglones que de columnas. Los ejemplos respectivos son:

$$\begin{bmatrix} 1 & -2 & \sqrt{3} \end{bmatrix} \qquad \begin{bmatrix} 3.5 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Una matriz cuya totalidad de elementos es cero se llama **matriz cero** y se representa con **0**.

$$\mathbf{0} = [0] \qquad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathbf{0} = \begin{bmatrix} 0 & 0 \end{bmatrix} \qquad \mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \mathbf{0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Sea una matriz *cuadrada* A cuyos elementos son a_{ij}; su **diagonal principal** estará formada por a_{ii}. La matriz A es **triangular superior** si todos sus elementos debajo de la diagonal principal son cero. A es **triangular inferior** si todos sus elementos arriba de la diagonal principal son cero. A es **diagonal** si todos los elementos arriba y debajo de la diagonal principal son cero. A es **escalar** si es diagonal y todos los elementos diagonales son iguales.

■ EJEMPLO 1 Sean

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 5 & -4 \\ 0 & 0 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

A, C y D son triangulares superiores. B, C y D son triangulares inferiores. C y D son diagonales. D es escalar. La diagonal principal de A es 1, 5, 9, mientras que la de C es 1, −2, 2.

Dos matrices son **iguales** si tienen el mismo tamaño y sus elementos correspondientes son iguales. Es lo mismo que decir que las columnas correspondientes son iguales, consideradas como vectores. Por ejemplo, $\begin{bmatrix} a & 5 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ -1 & b \end{bmatrix}$ sólo si $a = 1$ y $b = 0$. Las matrices de *distintos tamaños nunca son iguales.*

Suma de matrices y multiplicación por escalar

La **suma**, $A + B$, dos matrices A y B del mismo tamaño se obtiene sumando los elementos de ambas matrices. Para la **diferencia**, $A - B$, se restan los elementos correspondientes. Las matrices de distintos tamaños *no se pueden* sumar ni restar.

■ EJEMPLO 2

$$\begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 1 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & 3 \\ 5 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 0 & -3 \end{bmatrix} - \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 2 \\ 3 & -2 & 0 \end{bmatrix}$$

Sean A cualquier matriz y c cualquier escalar. El **producto por escalar**, cA, es la matriz que se obtiene al multiplicar cada elemento de A por c. Si $c = -1$, a $(-1)A$ se le llama **opuesta**[3] de A y se representa por $-A$.

■ EJEMPLO 3 Calcule $2A$ y $(-1)A = -A$, si

$$A = \begin{bmatrix} 1 & -3 \\ 0 & 1 \\ -2 & 2 \end{bmatrix}$$

SOLUCIÓN

$$2A = \begin{bmatrix} 2 & -6 \\ 0 & 2 \\ -4 & 4 \end{bmatrix}, \qquad (-1)A = -A = \begin{bmatrix} -1 & 3 \\ 0 & -1 \\ 2 & -2 \end{bmatrix}$$

Observe que

$$A - B = A + (-B)$$

$$\begin{bmatrix} a & \cdots \\ \vdots & \end{bmatrix} + \begin{bmatrix} b & \cdots \\ \vdots & \end{bmatrix} = \begin{bmatrix} a+b & \cdots \\ \vdots & \end{bmatrix}$$

$$c\begin{bmatrix} a & \cdots \\ \vdots & \end{bmatrix} = \begin{bmatrix} ca & \cdots \\ \vdots & \end{bmatrix}$$

La suma, resta y multiplicación por escalar de matrices se apegan a unas cuantas reglas básicas, que se resumen en el siguiente teorema, análogo al teorema 1 de la sección 2.1.

TEOREMA 1

(Leyes para suma de matrices y multiplicación por escalar)

Sean A, B y C matrices de $m \times n$ cualesquiera, y sean a, b y c escalares cualesquiera. Es válido lo siguiente:

1. $(A + B) + C = A + (B + C)$
2. $A + B = B + A$
3. $A + \mathbf{0} = \mathbf{0} + A = A$
4. $A + (-A) = (-A) + A = \mathbf{0}$
5. $c(A + B) = cA + cB$
6. $(a + b)C = aC + bC$
7. $(ab)C = a(bC) = b(aC)$
8. $1A = A$
9. $0A = \mathbf{0}$

OBSERVACIONES

1. La propiedad 1 se denomina **ley asociativa** para la suma, mientras que la propiedad 2 es la **ley conmutativa**. Las propiedades 5 y 6 se conocen como **leyes distributivas**.

---

[3] N. del T.: También se le llama **negativa de** A o **inversa aditiva de** A.

2. Como puede considerarse que las matrices son sucesiones de vectores, el teorema 1 se deduce con facilidad del teorema 1 de la sección 2.1. Sin embargo, también ayuda el considerar que una matriz es un arreglo rectangular de sus elementos. Practicaremos esto último demostrando la propiedad 5.

DEMOSTRACIÓN DE 5 Es claro que $c(A + b)$ y $cA + cB$ tienen el mismo tamaño. Sea p_{ij} el (i, j)-ésimo elemento de $c(A + B)$, y q_{ij} el (i, j)-ésimo elemento de $cA + cB$. Necesitamos demostrar que $p_{ij} = q_{ij}$. Si a_{ij} y b_{ij} son los (i, j)-ésimos elementos de A y de B, entonces $p_{ij} = c(a_{ij} + b_{ij})$, y $q_{ij} = ca_{ij} + cb_{ij}$. Pero $c(a_{ij} + b_{ij}) = ca_{ij} + cb_{ij}$. En consecuencia, $p_{ij} = q_{ij}$, de modo que todos los elementos correspondientes son iguales. La conclusión es que $c(A + B) = cA + cB$. $\square$

NOTA La ley asociativa permite eliminar los paréntesis en sumas múltiples. Así, todas las expresiones siguientes son iguales: $(A + B) + (C + D)$, $A + ((B + C) + D)$, $A + (B + (C + D))$, y pueden escribirse simplemente $A + B + C + D$.

Ecuaciones matriciales simples

El teorema 1 puede emplearse para resolver ecuaciones matriciales sencillas. Sean

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & -1 \end{bmatrix}$$

■ EJEMPLO 4 Determine la matriz X tal que $2X - 4B = 3A$.

SOLUCIÓN Se suma $4B$ en ambos lados de la ecuación para obtener

$$\begin{aligned} (2X - 4B) + 4B &= 3A + 4B \\ 2X + (-4B + 4B) &= 3A + 4B && \text{de acuerdo con 1, teorema 1} \\ 2X + \mathbf{0} &= 3A + 4B && \text{según la propiedad 4, del teorema 1} \\ 2X &= 3A + 4B && \text{según la propiedad 3, del teorema 1} \end{aligned}$$

Ambos lados de la última ecuación se multiplican por $\frac{1}{2}$:

$$\frac{1}{2}(2X) = \frac{1}{2}(3A + 4B) \quad \text{es decir} \quad X = \frac{3}{2}A + 2B \quad \text{según las propiedades 8, 5 y 7, del teorema 1}$$

Por consiguiente,

$$X = \frac{3}{2}\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} + 2\begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & 5 & 6 \\ 2 & 0 & -\frac{7}{2} \end{bmatrix}$$ $\square$

Multiplicación matricial

La multiplicación de matrices es la operación matricial más importante. Definir el producto de dos matrices como la matriz formada por los productos de cada uno de sus elementos correspondientes *no resulta* de mucha utilidad en las aplicaciones. La definición siguiente es mucho más útil.

DEFINICIÓN

(Multiplicación matricial)

Si A es una matriz de $m \times k$ y B una matriz de $k \times n$. El **producto** AB es la matriz de $m \times n$ cuyas columnas son $A\mathbf{b}_1, \ldots, A\mathbf{b}_n$, en la que $\mathbf{b}_1, \ldots, \mathbf{b}_n$ son las columnas de B.

Por ejemplo, si

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & 2 & 4 \\ -2 & 4 & 5 \\ 0 & 3 & -2 \end{bmatrix}$$

entonces

$$\begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 14 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \end{bmatrix}$$

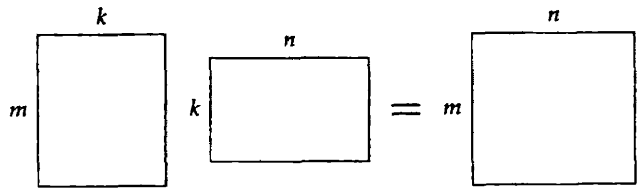
Por consiguiente,

$$AB = [\, A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad A\mathbf{b}_3 \,] = \begin{bmatrix} 6 & 7 & 6 \\ 4 & 14 & 9 \end{bmatrix}$$

También,

$$[\,5 \quad 1 \quad -1\,] \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} = [-2], \qquad \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} [\,5 \quad 1 \quad -1\,] = \begin{bmatrix} 5 & 1 & -1 \\ -15 & -3 & 3 \\ 20 & 4 & -4 \end{bmatrix}$$

PRECAUCIÓN La multiplicación matricial sólo es posible si la cantidad de columnas de la primera matriz es igual a la cantidad de renglones de la segunda. (En cualquier otro caso, $A\mathbf{b}_1, \ldots$, sería indefinida.)



Como en el caso de $A\mathbf{x}$, con frecuencia es útil obtener AB elemento por elemento. El (i, j)-ésimo elemento de AB puede calcularse como sigue: tome el i-ésimo renglón de A y la j-ésima columna de B. Multiplique entre sí sus elementos correspondientes y sume todos los productos.

■ EJEMPLO 5 Calcule los elementos (1, 2) y (2, 3) de AB.

SOLUCIÓN

$$\begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 2 & 4 \\ -2 & 4 & 5 \\ 0 & 3 & -2 \end{bmatrix} = \begin{bmatrix} \cdot & 7 & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

$$2 \cdot 2 + 0 \cdot 4 + 1 \cdot 3 = 7$$

$$\begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 2 & 4 \\ -2 & 4 & 5 \\ 0 & 3 & -2 \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & 9 \end{bmatrix}$$

Cálculo renglón-columna de AB

En general, si $C = [c_{ij}] = AB$, los elementos c_{ij} se determinan como sigue:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ik} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{k1} & \cdots & b_{kj} & \cdots & b_{kn} \end{bmatrix}$$

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}$$

En otras palabras, el (i, j)-ésimo elemento de AB es el **producto punto** del i-ésimo renglón de A por la j-ésima columna de B.

Con frecuencia sólo es necesario calcular determinado renglón o columna de un producto de matrices.

■ EJEMPLO 6 (Columna del producto) Calcule la tercera columna de AB.

SOLUCIÓN Es igual que el producto de A por la tercera columna de B.

$$\begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \end{bmatrix}$$

■ EJEMPLO 7 (Renglón del producto) Determine el segundo renglón de AB.

SOLUCIÓN Es lo mismo que el producto del segundo renglón de A por la matriz B.

$$[2 \quad 1 \quad 2]\begin{bmatrix} 3 & 2 & 4 \\ -2 & 4 & 5 \\ 0 & 3 & -2 \end{bmatrix} = [4 \quad 14 \quad 9]$$

En general,

El i-ésimo renglón de AB es el producto del i-ésimo renglón de A por B.

La j-ésima columna de AB es el producto de A por la j-ésima columna de B.

Las propiedades básicas de la multiplicación de matrices se resume en el teorema siguiente.

TEOREMA 2

(Leyes de la multiplicación matricial)

Si A es una matriz $m \times n$, y B y C tienen tamaños tales que las operaciones siguientes puedan llevarse a cabo. Y si a es cualquier escalar.

1. $(AB)C = A(BC)$	**Ley asociativa**
2. $A(B + C) = AB + AC$	**Ley distributiva izquierda**
3. $(B + C)A = BA + CA$	**Ley distributiva derecha**
4. $a(BC) = (aB)C = B(aC)$	
5. $I_m A = AI_n = A$	**Identidad multiplicativa**
6. $\mathbf{0}A = \mathbf{0}$ y $A\mathbf{0} = \mathbf{0}$	

DEMOSTRACIÓN DE 1 Comprobaremos la *asociatividad*, y dejaremos las demás leyes como ejercicios. Primero demostraremos el caso especial en el que $C = \mathbf{v}$ es un vector (una matriz de una columna) cuyos componentes son $(v_1, \ldots, v_k)$. Entonces

$$\begin{aligned}(AB)\mathbf{v} &= [A\mathbf{b}_1 \cdots A\mathbf{b}_k]\mathbf{v} = v_1(A\mathbf{b}_1) + \cdots + v_k(A\mathbf{b}_k) = A(v_1\mathbf{b}_1 + \cdots + v_k\mathbf{b}_k) \\ &= A(B\mathbf{v}) \end{aligned} \tag{3.1}$$

Ahora, si C tiene l columnas. De acuerdo con (3.1),

$$(AB)C = [(AB)\mathbf{c}_1 \cdots (AB)\mathbf{c}_l] = [A(B\mathbf{c}_1) \cdots A(B\mathbf{c}_l)] = A\,[B\mathbf{c}_1 \cdots B\mathbf{c}_l] = A(BC)$$

□

OBSERVACIÓN Al igual que en la suma, esta ley asociativa permite eliminar paréntesis de productos múltiples. Por ejemplo,

$$(AB)(CD) = A((BC)D) = A(B(CD)) = ABCD$$

■ EJEMPLO 8 Compruebe la ley distributiva de la multiplicación de matrices si

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

SOLUCIÓN

$$A(B + C) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 7 \\ 4 & 5 \\ 2 & 2 \end{bmatrix}$$

$$AB + AC = \begin{bmatrix} 4 & 6 \\ 3 & 4 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 7 \\ 4 & 5 \\ 2 & 2 \end{bmatrix}$$

□

Potencias de una matriz cuadrada

Sea A una matriz cuadrada. El producto AA también se representa por A^2. Igualmente, $AAA = A^3$ y $AA \cdots A = A^n$ cuando hay n factores de A. A también se escribe A^1. Si A es diferente de cero, también se escribe A^0 en vez de I.

$$A^n = \underbrace{AA\cdots A}_{n \text{ factores}}, \qquad A^1 = A, \qquad A^0 = I$$

■ EJEMPLO 9 Si

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad \text{y} \quad C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Entonces

$$A^1 = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 3 & -4 \\ -8 & 11 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 11 & -15 \\ -30 & 41 \end{bmatrix} \quad \cdots$$

$$B^1 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad B^2 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad B^3 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

$$C^1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad C^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad C^3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \cdots$$

□

TEOREMA 3

Las relaciones que se muestran a continuación son válidas para cualesquiera enteros positivos n y m.

$$A^nA^m = A^{n+m}, \quad (A^n)^m = A^{nm}, \quad (cA)^n = c^nA^n$$

Problemas que se presentan en la multiplicación de matrices

El teorema 2 describe propiedades que se comparten tanto en la multiplicación matricial como en la ordinaria (con I en lugar de 1). Sin embargo, hay diferencias importantes que hacen que la multiplicación de matrices sea algo más difícil, pero mucho más interesante que la multiplicación ordinaria. La diferencia más notable es que la multiplicación de matrices *no es conmutativa*, lo que significa que la propiedad $ab = ba$ válida para todos los números, no lo es para las matrices. Pero seamos más específicos.

1. Si AB está definida, esto no implica necesariamente que BA lo esté (por ejemplo, si A es 2×2 y B es 2×3).
2. Si AB y BA son definidos, no tienen por qué ser del mismo tamaño (por ejemplo, si A es 3×2 y B es 2×3).
3. Si AB y BA están definidas, y son del mismo tamaño, no tienen por qué ser iguales. Por ejemplo,

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \text{mientras que} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

4. $AB = \mathbf{0}$ *no siempre* implica que A o B sean $\mathbf{0}$ (ni siquiera cuando $A = B$). Por ejemplo,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

5. $CA = CB$ *no* necesariamente implica que $A = B$. Por ejemplo,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

6. $AC = BC$ *no* siempre implica que $A = B$. (Encuentre un ejemplo.)
7. $A^2 = I$ *no* necesariamente implica que $A = \pm I$. (Vea el ejercicio 20.)
8. En general, $(AB)^n \neq A^nB^n$. (Vea los ejercicios 21 y 23.)

Si dos matrices A y B satisfacen $AB = BA$, se dice que **conmutan**.

■ EJEMPLO 10

$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$ conmutan. Compruébelo. $\square$

Una **submatriz** es la matriz que se obtiene eliminando cualesquiera renglones o columnas de la primera matriz. Por ejemplo, $\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$, $[4\ \ 5\ \ 6]$ son submatrices de $M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$.

Con frecuencia se divide o secciona una matriz grande que no puede manipularse con facilidad y se forman "bloques" o submatrices que son más fáciles de manejar. Una matriz se secciona separando renglones y columnas mediante rectas horizontales y verticales. La matriz de submatrices que resulta se llama matriz **de bloques**.

Por ejemplo, seccionemos una matriz A de 3×6 como sigue:

$$A = \left[\begin{array}{ccc:cc:c} 1 & 2 & 0 & 1 & -1 & 1 \\ \hdashline 1 & 3 & 5 & 0 & 1 & 2 \\ 2 & 4 & 6 & 1 & 0 & 0 \end{array}\right]$$

Se puede considerar que A es la matriz de bloques de 2×3

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

en donde $A_{11} = [1\ \ 2\ \ 0]$, ..., $A_{23} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

La razón principal para usar matrices de bloques es que con ellas podemos efectuar operaciones matriciales exactamente como si fueran matrices "simples", siempre y cuando los tamaños de los bloques sean compatibles.[4] Por ejemplo, si

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ \cdots & \cdots \\ 1 & 1 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & 2 \\ -3 & 4 \\ \cdots & \cdots \\ -5 & 6 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

es fácil de comprobar que

$$A + B = \begin{bmatrix} A_1 + B_1 \\ A_2 + B_2 \end{bmatrix}$$

También, si

$$C = \left[\begin{array}{ccc:cc} 1 & 3 & -4 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 2 & 4 & 1 & 0 \\ 3 & 5 & 7 & 0 & 1 \end{array}\right] = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

y

$$D = \begin{bmatrix} -1 & 0 \\ 3 & 1 \\ 2 & 5 \\ \cdots & \cdots \\ -4 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$$

se ve que

$$CD = \begin{bmatrix} C_{11}D_1 + C_{12}D_2 \\ C_{21}D_1 + C_{22}D_2 \end{bmatrix}$$

Aplicaciones

Las matrices y las operaciones matriciales pueden emplearse para registrar, actualizar y escalar datos tabulares.

■ EJEMPLO 11 (Actualización de existencias) Una empresa tiene tres librerías, y cada una de ellas tiene libros de ficción, de viajes y de deportes. Las cantidades de libros se tabulan como sigue:

$$\begin{array}{c|ccc} \text{Librería} & \text{Ficción} & \text{Viajes} & \text{Deportes} \\ 1 & \multicolumn{3}{c}{\multirow{3}{*}{$\begin{bmatrix} 300 & 300 & 100 \\ 300 & 100 & 240 \\ 50 & 150 & 200 \end{bmatrix}$}} \\ 2 & & & \\ 3 & & & \end{array}$$

[4] Las operaciones con matrices de bloques se usan principalmente para manipular matrices muy grandes, que no caben en la memoria de una computadora.

Suponga que las entregas a cada librería están representadas por D. Calcule las existencias actualizadas.

$$D = \begin{bmatrix} 60 & 40 & 20 \\ 60 & 40 & 30 \\ 60 & 40 & 30 \end{bmatrix}$$

SOLUCIÓN Para conocer las existencias actualizadas es necesario sumar las dos matrices.

$$\begin{bmatrix} 360 & 340 & 120 \\ 360 & 140 & 270 \\ 110 & 190 & 230 \end{bmatrix}$$

Así, la librería 3 tiene ahora 190 libros de viajes. $\square$

■ EJEMPLO 12 (Escalamiento) Suponga que las distancias, en millas, entre Annapolis, Baltimore y Washington, D. C., se expresan como sigue:

$$\begin{array}{l|ccc} \text{Ciudad} & \text{Annapolis} & \text{Baltimore} & \text{Washington} \\ \hline \text{Annapolis} & 0 & 30 & 25 \\ \text{Baltimore} & 30 & 0 & 18 \\ \text{Washington} & 25 & 18 & 0 \end{array}$$

Si deseamos trazar un mapa cuya escala sea tal que 1 pulgada en el papel corresponda a 5 mi de distancia real, ¿cuál es la matriz de las distancias del mapa?

SOLUCIÓN Por ejemplo, una distancia de 30 millas se representa con 30 mi$\cdot \frac{1}{5} \frac{\text{in.}}{\text{mi}}$ = 6 pulg. De modo, que necesitamos multiplicar todos los elementos de la matriz por $\frac{1}{5} = 0.2$. Éste es el producto de la matriz por el escalar 0.2.

$$0.2 \begin{bmatrix} 0 & 30 & 25 \\ 30 & 0 & 18 \\ 25 & 18 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 6 & 5 \\ 6 & 0 & 3.6 \\ 5 & 3.6 & 0 \end{bmatrix}$$ $\square$

■ EJEMPLO 13 (Ingresos procedentes de diversas fuentes) Cada una de las tiendas, o_1 y o_2, reciben diariamente televisores (t) y videocaseteras (v) de dos fabricantes, f_1 y f_2. Las recepciones o ventas se representan como sigue:

$$\begin{array}{cc} & \begin{array}{cc} t & v \end{array} \\ \begin{array}{c} f_1 \\ f_2 \end{array} & \begin{bmatrix} 40 & 50 \\ 70 & 80 \end{bmatrix} \end{array}$$

El precio en dólares, por aparato en cada tienda, es como sigue:

$$\begin{array}{cc} & \begin{array}{cc} o_1 & o_2 \end{array} \\ \begin{array}{c} t \\ v \end{array} & \begin{bmatrix} 200 & 250 \\ 300 & 280 \end{bmatrix} \end{array}$$

Si A y B son las matrices de las tablas anteriores, calcule e interprete el producto AB.

SOLUCIÓN

$$AB = \begin{bmatrix} 40 & 50 \\ 70 & 80 \end{bmatrix} \begin{bmatrix} 200 & 250 \\ 300 & 280 \end{bmatrix} = \begin{bmatrix} 23{,}000 & 24{,}000 \\ 38{,}000 & 39{,}900 \end{bmatrix}$$

El elemento (1, 1) es $40 \cdot 200 + 50 \cdot 300 = 23\,000$, y representa los ingresos de la primera tienda por vender todos los electrodomésticos que provienen de la primera fábrica. De igual forma,

$$AB = \begin{bmatrix} \$ \text{ en } o_1 \text{ de } f_1 & \$ \text{ en } o_2 \text{ de } f_1 \\ \$ \text{ en } o_1 \text{ de } f_2 & \$ \text{ en } o_2 \text{ de } f_2 \end{bmatrix}$$

Algunas notas útiles

Multiplicación por una matriz escalar

Si A es una matriz de $m \times n$, y Q y R son matrices escalares $m \times m$ y $n \times n$, respectivamente (con el escalar c en sus diagonales principales), pueden combinarse las propiedades 4 y 5 del teorema 2 para calcular $QA = (cI_m)A = c(I_mA) = cA$. Igualmente, $AR = cA$. Por consiguiente,

$$QA = AR = cA$$

En otras palabras, *la multiplicación por una matriz escalar es igual que la multiplicación por un escalar.* Además, la suma, diferencia y producto de las matrices escalares que poseen el mismo tamaño dan como resultado una matriz escalar.

■ EJEMPLO 14

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ -6 & -4 \\ -2 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = -2 \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 1 & 1 \end{bmatrix}$$

Cálculo rápido de A^k mediante la elevación al cuadrado

Es muy tedioso calcular las potencias de una matriz cuadrada A. Por ejemplo, digamos que necesitamos calcular A^8. Si tratamos de seguir la definición

$$A^8 = ((((((AA)A)A)A)A)A)A$$

Sería preciso efectuar 7 multiplicaciones de matrices. Sin embargo, si primero se calcula A^2, después se eleva al cuadrado, $A^2A^2 = A^4$, y este resultado también se eleva a la segunda potencia, al cuadrado, $A^4A^4 = A^8$, sólo se necesitan 3 multiplicaciones de matrices.

$$A^8 = \left((A^2)^2\right)^2$$

Este método se aplica a cualquier potencia de matriz, A^n. Por ejemplo, A^{13} puede determinarse como sigue: encontramos A^2, después se eleva al cuadrado para obtener A^4 y el resultado se eleva al cuadrado para tener A^8. Ahora bien, $A^{13} = A^8A^4A$. Este proceso sólo necesita 5 multiplicaciones de matrices, en lugar de las 12 que serían necesarias si optáramos por la definición.

$$A^{13} = \left((A^2)^2\right)^2 A^4 A$$

Cálculo rápido de $A^k\mathbf{x}$ sin potencias de matrices

Si se desea calcular el producto $A^k\mathbf{x}$ para una matriz A de $n \times n$, por $\mathbf{x}$, un vector n, es posible e ingenioso evitar el cálculo de A^k. Primero necesitamos calcular $A\mathbf{x}$, el vector n. Después determinamos $A(A\mathbf{x})$, cuyo resultado es un vector n, y continuamos en esta forma hasta terminar. En otras palabras, empleamos

$$A(\ldots(A(A\mathbf{x}))\ldots) = A^k\mathbf{x}$$

donde podemos observar que el lado izquierdo de la ecuación tiene $k - 1$ pares de paréntesis. La ventaja de este método es que siempre se calcula el producto de una matriz por un vector, y nunca el producto de dos matrices.

Para ejemplificarlo, sean $n = 3$ y $k = 2$. Si queremos determinar primero A^2 se requieren 27 multiplicaciones; para $A^2\mathbf{x}$ se necesitan otras 9, lo cual haría un total de 36 multiplicaciones. Si se calcula primero $A\mathbf{x}$, se requieren 9 multiplicaciones. A continuación, para $A(A\mathbf{x})$ son necesarias 9 multiplicaciones más, lo cual significa un total de sólo 18 multiplicaciones.

Operaciones matriciales con los sistemas algebraicos computacionales

Se puede calcular $AB - B^2$ como sigue, siendo

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 7 \\ -5 & 2 \end{bmatrix}$$

Maple

```
> with(linalg):
> A:=matrix(2,2,[1,-2,3,4]):B:=matrix(2,2,[0,7,-5,2]):
> evalm(A&*B-B^2);
```

$$\begin{bmatrix} 45 & -11 \\ -10 & 60 \end{bmatrix}$$

Mathematica

```
In[1]:=
     A={{1,-2},{3,4}};B={{0,7},{-5,2}};
In[2]:=
     A . B-MatrixPower[B, 2]
Out[2]=
     {{45, -11}, {-10, 60}}
```

MATLAB

```
>> A=[1 -2; 3 4]; B=[0 7; -5 2];
>> A*B-B^2
ans =
    45    -11
   -10     60
```

Ejercicios 3.1

Matrices; igualdad de matrices

1. Identifique los renglones, las columnas, los tamaños y los elementos (2, 2) de las matrices. Determine el elemento (3, 1) de A y el elemento (2, 3) de B.

$$A = \begin{bmatrix} -1 & 0 \\ 2 & 3 \\ -2 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & 0 & -2 \\ 2 & 2 & 1 \end{bmatrix}$$

2. Determine los valores de x, y y z tales que las matrices siguientes sean iguales.

a. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & -3 \end{bmatrix} = \begin{bmatrix} x & 0 & 1 \\ x+y & 2 & x+z \end{bmatrix}$

b. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & -3 \end{bmatrix} = \begin{bmatrix} x+z & 0 & 1 \\ -y & 2 & -z \end{bmatrix}$

3. Demuestre lo siguiente.

a. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \neq \begin{bmatrix} x & y-2 \\ x-y & 2 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \neq \begin{bmatrix} x+y & x+y \\ -y+z & x+z \end{bmatrix}$

4. ¿Cuál(es) de las matrices

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

$$G = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, \qquad H = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

es o son de los siguientes tipos?

a. Triangular superior
b. Triangular inferior
c. Diagonal
d. Escalar
e. Nada de lo anterior

Suma, multiplicación por escalar

5. Calcule, si es posible, lo siguiente. Si las operaciones no pueden efectuarse, explique por qué.

a. $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$

b. $-3\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$

c. $\begin{bmatrix} 2 & 2 \\ -2 & -2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

d. $\begin{bmatrix} -2 & 3 \\ 4 & -5 \\ -6 & 7 \end{bmatrix} - \begin{bmatrix} 7 & -6 \\ -5 & 4 \\ 3 & -2 \end{bmatrix}$

e. $-[1 \quad -2] + \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$

f. $3\begin{bmatrix} 0 & 2 \\ -4 & -6 \end{bmatrix} - 4\begin{bmatrix} 3 & -5 \\ 7 & 0 \end{bmatrix}$

Sean $A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 5 & -2 \end{bmatrix}$, y

$C = \begin{bmatrix} 8 & -2 \\ -6 & 4 \end{bmatrix}$.

En los ejercicios 6 y 7 determine la matriz X que satisfaga las ecuaciones.

6. a. $2X + B = -3A + C$
   b. $\frac{1}{2}A - \frac{3}{2}X = B + 2C$
   c. $0C + X + 5B = 2A$
7. a. $2X + B + C = \mathbf{0}$
   b. $\frac{1}{2}A - X = B - C$
   c. $0C + \frac{1}{2}X = 2A - 4B$
8. Demuestre las propiedades 1 a 4 del teorema 1.
9. Demuestre las propiedades 6 a 9 del teorema 1.

Multiplicación de matrices

10. Si es posible, calcule

a. $\begin{bmatrix} 3 \\ 4 \end{bmatrix}[1 \quad 2]$

b. $[1 \quad 2]\begin{bmatrix} 3 \\ 4 \end{bmatrix}$

c. $\begin{bmatrix} 1 & -2 \\ 4 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 & 4 \\ -2 & -4 & 3 & 0 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ -2 & -4 & 3 & 0 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 4 & 0 \end{bmatrix}$

e. $\begin{bmatrix} -3 & 0 \\ 2 & -5 \\ -7 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

11. Obtenga el tercer renglón de AB si

$$A = \begin{bmatrix} 3 & 4 \\ 4 & 3 \\ 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 5 & 6 \\ 6 & 5 & 2 & 1 \end{bmatrix}$$

12. Calcule la segunda columna de AB si

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 6 & 5 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & 4 \\ 4 & 3 \\ 1 & 2 \end{bmatrix}$$

13. Encuentre el elemento (2, 2) de

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -3 & 1 \end{bmatrix}$$

14. Determine $(2A)^3$ si

$$A^3 = \begin{bmatrix} 1 & 1 \\ -5 & -2 \end{bmatrix}$$

15. Calcule A^8 si

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

¿puede adivinar cuál será A^n?

16. Compruebe el teorema 2 para $A = \begin{bmatrix} -2 & 3 \\ 4 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 5 \\ 0 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 0 \\ 1 & -2 \end{bmatrix}$, y $a = -3$.

17. Demuestre las partes 2 a 6 del teorema 2.

18. Encuentre A, una matriz de 2×2 tal que $A^2 = \mathbf{0}$.

19. Determine $A \neq I$, una matriz de 2×2 tal que $A^2 = A$.

20. Recuerde que para los números reales la ecuación $a^2 = 1$ sólo tiene dos soluciones, que son $a = \pm 1$. Pero una afirmación análoga no es válida para las matrices. Cite cuatro matrices A, de 2×2, tales que $A^2 = I$.

21. Encuentre las matrices A y B de 2×2 tales que $(AB)^2 \neq A^2B^2$.

22. Determine las matrices A y B de 2×2 que conmuten.

23. Compruebe que si A y B conmutan, entonces $(AB)^2 = A^2B^2$.

24. **a.** Demuestre que si A y B son matrices cuadradas del mismo tamaño, entonces

$$(A + B)^2 = A^2 + 2AB + B^2 \Leftrightarrow AB = BA$$

b. Forme dos matrices, A y B de 2×2, tales que

$$(A + B)^2 \neq A^2 + 2AB + B^2$$

25. Compruebe el teorema 3 para $A = \begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix}$, $n = 2$, $m = 1$, y $c = -3$.

26. Demuestre el teorema 3.

27. Forme todas las submatrices 2×2 de M.

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

28. Si

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 5 \\ \cdots & \cdots \\ 1 & 1 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

y

$$B = \begin{bmatrix} -1 & 2 \\ 8 & 9 \\ \cdots & \cdots \\ 5 & -6 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

compruebe que

$$A + B = \begin{bmatrix} A_1 + B_1 \\ A_2 + B_2 \end{bmatrix}$$

29. Si

$$C = \left[\begin{array}{ccc:cc} 7 & 3 & -4 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 2 & 3 & 1 & 0 \\ 3 & 1 & -8 & 0 & 1 \end{array}\right] = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

y

$$D = \begin{bmatrix} 1 & 0 \\ -3 & 1 \\ 5 & -5 \\ \cdots & \cdots \\ -3 & 0 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$$

demuestre que

$$CD = \begin{bmatrix} C_{11}D_1 + C_{12}D_2 \\ C_{21}D_1 + C_{22}D_2 \end{bmatrix}$$

Aplicaciones

30. Una empresa embotelladora de agua entrega tres tipos de agua embotellada: Fuente, Montaña y Polar, a cuatro tiendas. Esas tiendas son abastecidas, en centenares de botellas, de acuerdo con la siguiente tabla

$$\begin{array}{c c} \text{Tienda} & \begin{array}{cccc} \text{Fuente} & \text{Montaña} & \text{Polar} \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 8 & 4 & 3 \\ 10 & 5 & 4 \\ 6 & 3 & 2 \\ 6.5 & 3.5 & 2.5 \end{bmatrix} \end{array}$$

Actualice el inventario cuando se hacen las entregas D.

$$D = \begin{bmatrix} 1 & 2 & 0.5 \\ 0.5 & 2 & 1 \\ 2 & 0.5 & 1 \\ 2.5 & 1.5 & 1.5 \end{bmatrix}$$

31. Una empresa tiene tres tiendas de departamentos. Cada una vende sacos, pantalones y calzado, cuyos precios son los de la siguiente tabla. Calcule los valores de los artículos, si en las 3 tiendas se ofrece una barata del 20%.

$$\begin{array}{c c} \text{Tienda} & \begin{array}{ccc} \text{Sacos} & \text{Pantalones} & \text{Zapatos} \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 80 & 40 & 30 \\ 100 & 50 & 40 \\ 60 & 30 & 20 \end{bmatrix} \end{array}$$

32. Dos tiendas de departamentos, d_1 y d_2, son abastecidas semanalmente por dos fábricas de ropa, f_1 y f_2, con pantalones (p) y sacos (s), de acuerdo con

$$\begin{array}{c c} & \begin{array}{cc} s & j \end{array} \\ \begin{array}{c} f_1 \\ f_2 \end{array} & \begin{bmatrix} 50 & 20 \\ 60 & 30 \end{bmatrix} \end{array}$$

El precio por artículo para cada tienda, en dólares, es el siguiente:

$$\begin{array}{c c} & \begin{array}{cc} d_1 & d_2 \end{array} \\ \begin{array}{c} s \\ j \end{array} & \begin{bmatrix} 100 & 85 \\ 350 & 400 \end{bmatrix} \end{array}$$

Si A y B son las matrices de esas tablas, calcule e interprete el producto AB.

3.2 Matriz inversa

Objetivos del estudiante para esta sección

1. Definir la matriz inversa y sus propiedades.
2. Identificar y calcular la matriz inversa, si ésta existe.
3. Resolver algunos sistemas cuadrados invirtiendo la matriz de coeficientes.

La inversión de matrices es la última operación matricial básica. Sólo se aplica a matrices *cuadradas*. Cuando existe una inversa, ésta es análoga al recíproco de un número distinto de cero.

DEFINICIÓN

(Inversa)

Se dice que la matriz A de $n \times n$ es **invertible**, o **no singular**, si existe una matriz B, llamada la **inversa** de A tal que

$$AB = I \quad \text{y} \quad BA = I$$

Obsérvese que la definición obliga a que B tenga el tamaño de $n \times n$.

Una matriz invertible sólo tiene una inversa, es decir, **la inversa es única**. Si C fuera otra inversa, entonces

$$B = BI_n = B(AC) = (BA)C = I_nC = C$$

Por consiguiente, $B = C$. La inversa única de una matriz invertible A se representa por A^{-1}. Así,

$$AA^{-1} = I \quad \text{y} \quad A^{-1}A = I$$

Una matriz cuadrada que no tiene inversa se llama **no invertible** o **singular**.

■ EJEMPLO 15 Demuestre que

$$\begin{bmatrix} 2 & 7 \\ 1 & 4 \end{bmatrix} \text{ es la inversa de } \begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix}$$

SOLUCIÓN

$$\begin{bmatrix} 2 & 7 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 4 & -7 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 7 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

□

En el caso de las matrices de 2×2 podemos determinar exactamente cuáles son invertibles y citar una fórmula explícita de la inversa. Esto es más difícil de predecir para tamaños mayores a 2.

TEOREMA 4

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es invertible si y sólo si $ad - bc \neq 0$, en cuyo caso,

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

DEMOSTRACIÓN Se deja como ejercicio. □

■ EJEMPLO 16 Demuestre que

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

es invertible, y determine su inversa.

SOLUCIÓN $1 \cdot 4 - 2 \cdot 3 \neq 0$. Por consiguiente, según el teorema 4, la matriz A es invertible. Además,

$$A^{-1} = \frac{1}{1 \cdot 4 - 2 \cdot 3}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

□

■ EJEMPLO 17 Demuestre que

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

no es invertible.

SOLUCIÓN Esto es consecuencia del teorema 4, porque $1 \cdot 0 - 0 \cdot 0 = 0$. □

TEOREMA 5

Si la matriz A, $n \times n$, puede invertirse, el sistema $A\mathbf{x} = \mathbf{b}$ tiene exactamente una solución para cada vector n $\mathbf{b}$. Esta solución única es

$$\mathbf{x} = A^{-1}\mathbf{b}$$

DEMOSTRACIÓN $\mathbf{x} = A^{-1}\mathbf{b}$ es una solución porque al sustituirla en el sistema se obtiene $A(A^{-1}\mathbf{b}) = \mathbf{b}$, es decir $(AA^{-1})\mathbf{b} = \mathbf{b}$, o $I\mathbf{b} = \mathbf{b}$, o $\mathbf{b} = \mathbf{b}$, lo cual es cierto. Además, la solución es única, porque si $\mathbf{y}$ también lo fuera entonces

$$A\mathbf{y} = \mathbf{b} \Rightarrow A^{-1}A\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow \mathbf{y} = A^{-1}\mathbf{b}$$

En este caso, todas las soluciones se expresan con la misma fórmula: $A^{-1}\mathbf{b}$. ☐

CASO ESPECIAL Si A es invertible, entonces $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial.

Observe que los sistemas que se describen aquí tienen *la misma cantidad de ecuaciones y de incógnitas*, porque A es una matriz cuadrada.

■ EJEMPLO 18 Aplique la inversión matricial para resolver el sistema

$$\begin{aligned} x - 4y &= 2 \\ x - 3y &= 1 \end{aligned}$$

SOLUCIÓN Como la forma matricial del sistema es

$$\begin{bmatrix} 1 & -4 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

entonces

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -4 \\ 1 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

Por consiguiente, $x = -2$ y $y = -1$. ☐

El teorema siguiente describe las propiedades básicas de las matrices inversas.

TEOREMA 6

(Propiedades de las matrices inversas)

1. El producto de dos matrices inversas es invertible. Su inversa es el producto de las inversas de los factores en orden inverso. Así, si A y B son matrices $n \times n$ invertibles, también AB lo será, y
$$(AB)^{-1} = B^{-1}A^{-1}$$

2. La inversa de una matriz invertible también es invertible. Su inversa es la matriz original. Por consiguiente, si A es invertible, también lo es A^{-1}, y
$$(A^{-1})^{-1} = A$$

3. Cualquier producto de un escalar distinto de cero por una matriz invertible es invertible. Su inversa es el producto del recíproco del escalar por la inversa de la matriz. Por consiguiente, si A es invertible y c es un escalar distinto de cero, entonces cA es invertible y
$$(cA)^{-1} = \frac{1}{c}A^{-1}$$

DEMOSTRACIÓN

1. Necesitamos comprobar que $(AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB)$. Tenemos que

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$
$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

2. Como A es invertible, $AA^{-1} = I = A^{-1}A$. Esto demuestra que A^{-1} también lo es, y $(A^{-1})^{-1} = A$.
3. Esta demostración se deja como ejercicio. □

OBSERVACIÓN Si $A_1, A_2, \ldots, A_{n-1}, A_n$ son invertibles y del mismo tamaño, también lo es el producto $A_1A_2 \cdots A_{n-1}A_n$. Su inversa es

$$(A_1A_2 \ldots A_{n-1}A_n)^{-1} = A_n^{-1}A_{n-1}^{-1} \ldots A_2^{-1}A_1^{-1}$$

PRECAUCIÓN En general (véanse los ejercicios 25 y 26).

$$(A + B)^{-1} \neq A^{-1} + B^{-1} \quad \text{y} \quad (AB)^{-1} \neq A^{-1}B^{-1}$$

Potencias de matrices con exponentes negativos

Si A es invertible, se definen las potencias de A con exponentes negativos en la siguiente forma. Para $n > 0$ y n entero,

$$A^{-n} = (A^{-1})^n = \underbrace{A^{-1}A^{-1} \cdots A^{-1}}_{n \text{ factores}}$$

TEOREMA 7

(Propiedades de las potencias)

Sea A una matriz invertible $k \times k$; sean m y n dos enteros cualesquiera (> 0, < 0 o $= 0$) y sea c un escalar diferente de cero. Entonces, A^n es invertible y

1. $(A^n)-1 = (A^{-1})^n$
2. $A^nA^m = A^{n+m}$
3. $(A^n)^m = A^{nm}$
4. $(cA)^n = c^nA^n$

DEMOSTRACIÓN Se deja como ejercicio. □

Recuérdese que en la sección 3.1, en general, para una ecuación matricial $CA = CB$ no se pueden simplificar las C. Sin embargo, esa anulación es posible cuando C es invertible.

TEOREMA 8

(Leyes de la simplificación)

Si C es invertible, entonces

1. $CA = CB \Rightarrow A = B$
2. $AC = BC \Rightarrow A = B$

DEMOSTRACIÓN 1. C^{-1} existe; por consiguiente,

$$C = CB \Rightarrow C^{-1}(CA) = C^{-1}(CB) \Rightarrow (C^{-1}C)A = (C^{-1}C)B \Rightarrow IA = IB \Rightarrow A = B$$

La parte 2 se demuestra en forma parecida.

Ecuaciones sencillas con productos de matrices

Las propiedades básicas de las operaciones matriciales nos permiten resolver algunas ecuaciones matriciales.

PRECAUCIÓN Cuando se multiplica ambos lados de una ecuación matricial por una matriz, se debe usar la multiplicación por la izquierda o por la derecha, pero no ambas a la vez. Así,

$$A = B \Rightarrow CA = CB, \qquad A = B \Rightarrow AD = BD$$

mientras que

$$A = B \text{ no implica que } CA = BC.$$

■ EJEMPLO 19 Despeje X de la ecuación matricial $CXA - B = \mathbf{0}$, suponiendo que A y C son invertibles y que todos los tamaños son compatibles.

SOLUCIÓN Aplicando las propiedades de las operaciones matriciales se puede despejar X al pasarlo al lado izquierdo de la ecuación.

$$\begin{aligned} AXC - B &= \mathbf{0} \\ AXC &= B \\ A^{-1}AXC &= A^{-1}B \\ XC &= A^{-1}B \\ XCC^{-1} &= A^{-1}BC^{-1} \\ X &= A^{-1}BC^{-1} \end{aligned}$$

Los pasos 3 y 5 son posibles, porque existen A^{-1} y C^{-1}. Si especificamos A, B y C, como por ejemplo

$$A = \begin{bmatrix} 2 & 1 \\ 7 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & -1 \\ -2 & -1 \end{bmatrix}$$

entonces

$$X = A^{-1}BC^{-1} = \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ \frac{3}{2} & \frac{3}{2} \end{bmatrix}$$

ERROR FRECUENTE La expresión $A^{-1}BA$ por lo general *no* se simplifica para obtener B. (Sólo si A y B conmutan.)

Cálculo de A^{-1}

Ahora veremos cómo aplicar la reducción por operaciones de renglón para calcular la inversa. El ejemplo siguiente ilustra y justifica el método que se presentó por primera vez en el proyecto 1 del capítulo 1.

Sea $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ y sea $A^{-1} = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$. Como $AA^{-1} = I$,

$$\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{o sea} \quad \begin{bmatrix} 2x+3z & 2y+3w \\ x+2z & y+2w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Igualando las columnas se obtienen dos sistemas

$$\begin{aligned} 2x + 3z &= 1 \\ x + 2z &= 0 \end{aligned} \quad \text{y} \quad \begin{aligned} 2y + 3w &= 0 \\ y + 2w &= 1 \end{aligned}$$

cuyas matrices aumentadas son

$$\begin{bmatrix} 2 & 3 & : & 1 \\ 1 & 2 & : & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & 3 & : & 0 \\ 1 & 2 & : & 1 \end{bmatrix}$$

y cuyas formas de escalón reducidas son

$$\begin{bmatrix} 1 & 0 & : & 2 \\ 0 & 1 & : & -1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 & : & -3 \\ 0 & 1 & : & 2 \end{bmatrix}$$

Por consiguiente, $x = 2$, $z = -1$, para el primer sistema, y $y = -3$, $w = 2$ para el segundo. En consecuencia, tenemos que $A^{-1} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$. Como los dos sistemas tienen la misma matriz de coeficientes, podemos ahorrar escritura si los combinamos en la forma $\begin{bmatrix} 2 & 3 & : & 1 & 0 \\ 1 & 2 & : & 0 & 1 \end{bmatrix}$ y después los reducimos con operaciones de renglón a $\begin{bmatrix} 1 & 0 & : & 2 & -3 \\ 0 & 1 & : & -1 & 2 \end{bmatrix}$

Observe que comenzamos con $[A : I]$ y después de la reducción obtuvimos $[I : A^{-1}]$, de donde podemos obtener A^{-1}. Es todo el método. A continuación lo practicaremos.

■ EJEMPLO 20 Calcule A^{-1} si

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 4 & -2 \\ 3 & 5 & -2 \end{bmatrix}$$

SOLUCIÓN $[A : I]$ se reduce mediante operaciones de renglón.

$$\begin{bmatrix} 1 & 0 & -1 & : & 1 & 0 & 0 \\ 3 & 4 & -2 & : & 0 & 1 & 0 \\ 3 & 5 & -2 & : & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & : & 1 & 0 & 0 \\ 0 & 4 & 1 & : & -3 & 1 & 0 \\ 0 & 5 & 1 & : & -3 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & -1 & : & 1 & 0 & 0 \\ 0 & 4 & 1 & : & -3 & 1 & 0 \\ 0 & 0 & -\frac{1}{4} & : & \frac{3}{4} & -\frac{5}{4} & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & : & 1 & 0 & 0 \\ 0 & 4 & 1 & : & -3 & 1 & 0 \\ 0 & 0 & 1 & : & -3 & 5 & -4 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & : & -2 & 5 & -4 \\ 0 & 4 & 0 & : & 0 & -4 & 4 \\ 0 & 0 & 1 & : & -3 & 5 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & : & -2 & 5 & -4 \\ 0 & 1 & 0 & : & 0 & -1 & 1 \\ 0 & 0 & 1 & : & -3 & 5 & -4 \end{bmatrix}$$

Por consiguiente,

$$A^{-1} = \begin{bmatrix} -2 & 5 & -4 \\ 0 & -1 & 1 \\ -3 & 5 & -4 \end{bmatrix} \qquad \square$$

Con este método no sólo se calculan matrices inversas, sino también se detectan matrices no invertibles. Por ejemplo, si $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, entonces

$$\left[\begin{array}{cc:cc} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc:cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right]$$

Es claro que no podemos obtener la matriz identidad en la izquierda, porque el segundo elemento del segundo renglón es 0. Al reinterpretar la respuesta en términos de sistemas, notaremos que el segundo renglón implica las ecuaciones incorrectas $0 = -2$ y $0 = 1$. Por tanto, no existe la inversa.

El algoritmo siguiente nos muestra cuándo es invertible una matriz. Además, si la respuesta es positiva, nos permite calcular la inversa.

Algoritmo 1

(Algoritmo para invertir una matriz)

Para determinar A^{-1}, si existe, haga lo siguiente:

1. Obtenga la forma de escalón reducida de la matriz $[A : I]$, digamos $[B : C]$.
2. Si B tiene un renglón 0, deténgase. A no es invertible. De lo contrario, pase a 3.
3. La matriz reducida se encuentra ahora en la forma $[I : A^{-1}]$. Anote la inversa A^{-1}.

En la sección 3.3 analizaremos el algoritmo de inversión de una matriz y explicaremos su funcionamiento.

Transposición y operaciones con matrices

La **transpuesta**, A^T, de una matriz A $m \times n$, es la matriz $n \times m$ cuyas columnas son los renglones de A en el mismo orden. Por ejemplo,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad B^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

La operación empleada para determinar la transpuesta de una matriz se llama transposición. Una matriz que es igual a su transpuesta se denomina **simétrica**. Por consiguiente, B es simétrica. Veamos cómo las operaciones matriciales básicas afectan a la transposición.

TEOREMA 9

1. $(A^T)^T = A$.
2. $(A + B)^T = A^T + B^T$.
3. $(cA)^T = cA^T$.
4. $(AB)^T = B^T A^T$.
5. Si A es invertible, también lo es A^T. En este caso $(A^T)^{-1} = (A^{-1})^T$.

Obsérvese que las matrices mencionadas en las partes 1 a 4 no son necesariamente cuadradas. Bosquejaremos la demostración de la parte 4 y dejaremos como ejercicios las restantes.

DEMOSTRACIÓN PARCIAL DE 4 Primero, el caso especial en el que A es de $1 \times n$ y B es de $n \times 1$, implicado en

$$[a_1 \ldots a_n]\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = a_1b_1 + \cdots + a_nb_n = b_1a_1 + \cdots + b_na_n = [b_1 \ldots b_n]\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

A continuación demostraremos el caso especial en el que A es $m \times n$ y B es un vector n, por ejemplo $\mathbf{v}$. El caso general que se deja como ejercicio, se basa en éste y se demuestra en forma parecida. Se tiene que

$$\begin{aligned}
j\text{-ésima columna de } (A\mathbf{v})^T &= (j\text{-ésimo renglón de } A\mathbf{v})^T \\
&= (j\text{-ésimo renglón de } A \text{ por } \mathbf{v})^T \\
&= \mathbf{v}^T \text{ por el } j\text{-ésimo renglón de } A^T \quad (\text{de acuerdo con el caso } 1 \times n \text{ por } n \times 1). \\
&= \mathbf{v}^T \text{ por la } j\text{-ésima columna de } A^T \\
&= j\text{-ésima columna de } \mathbf{v}^T A^T
\end{aligned}$$

□

■ EJEMPLO 21 Compruebe el enunciado 5 del teorema 9 para

$$A = \begin{bmatrix} 1 & -2 \\ 3 & -7 \end{bmatrix}$$

SOLUCIÓN

$$(A^T)^{-1} = \begin{bmatrix} 1 & 3 \\ -2 & -7 \end{bmatrix}^{-1} = \begin{bmatrix} 7 & 3 \\ -2 & -1 \end{bmatrix}, \quad (A^{-1})^T = \begin{bmatrix} 7 & -2 \\ 3 & -1 \end{bmatrix}^T = \begin{bmatrix} 7 & 3 \\ -2 & -1 \end{bmatrix}$$

□

Aplicación a la ingeniería

Como ejemplo de inversión de matrices, regresemos al problema de transmisión de calor que estudiamos en la sección 2.1. Ahí se proporcionó la distribución de temperaturas $b_1, b_2, \ldots, b_8$, para una placa cuadrada, y se pidió determinar las temperaturas $x_1, \ldots, x_4$ en el interior (figura 2.8). El sistema lineal resultante fue

$$\begin{aligned}
x_1 &= \frac{1}{4}(x_2 + x_3 + b_1 + b_3) \\
x_2 &= \frac{1}{4}(x_1 + x_4 + b_2 + b_4) \\
x_3 &= \frac{1}{4}(x_1 + x_4 + b_5 + b_7) \\
x_4 &= \frac{1}{4}(x_2 + x_3 + b_6 + b_8)
\end{aligned}$$

■ EJEMPLO 22 (Conducción de calor) Demuestre que el problema de transferencia de calor para el sistema anterior siempre tiene una solución única para cualesquiera valores de temperatura $b_1, b_2, \ldots, b_8$ en la frontera. Calcule esta solución.

SOLUCIÓN Para simplificar la notación, sean $B_1 = b_1 + b_3$, $B_2 = b_2 + b_4$; $B_3 = b_5 + b_7$ y $B_4 = b_6 + b_8$. El sistema puede escribirse en la forma canónica $A\mathbf{x} = \mathbf{b}$ siendo

$$A = \begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix}$$

Si calculamos A^{-1} con el algoritmo de inversión de matrices, nos da

$$\frac{1}{24}\begin{bmatrix} 7 & 2 & 2 & 1 \\ 2 & 7 & 1 & 2 \\ 2 & 1 & 7 & 2 \\ 1 & 2 & 2 & 7 \end{bmatrix}$$

Por tanto, el sistema tiene una solución única, $\mathbf{x} = A^{-1}\mathbf{b}$ para cualquier elección de términos constantes, de acuerdo con el teorema 5. Así,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = A^{-1}\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 7B_1 + 2B_2 + 2B_3 + B_4 \\ 2B_1 + 7B_2 + B_3 + 2B_4 \\ 2B_1 + B_2 + 7B_3 + 2B_4 \\ B_1 + 2B_2 + 2B_3 + 7B_4 \end{bmatrix}$$

Aplicación a la economía

Con frecuencia, los economistas estudian las condiciones para los *equilibrios de mercados*, es decir, las condiciones bajo las que se relacionan los precios de diversos artículos.

Por ejemplo, examinemos los mercados relacionados de bolígrafos y lápices. Sean P_l y P_b los precios, en dólares, de un lápiz o un bolígrafo, respectivamente. Algunas condiciones del mercado obligan a que los dos precios satisfagan la relación $P_l + P_b = 1.5$, mientras que otras circunstancias requieren de la relación $P_l - P_b = 0.5$. El *precio de equilibrio* para cada mercado es aquel que cumple ambas condiciones. Ésta es la solución del sistema de las dos ecuaciones. Así, los precios de equilibrio son $P_l = 0.5$ y $P_b = 1$.

■ EJEMPLO 23 (Equilibrio en mercados relacionados) Las condiciones de equilibrio entre tres mercados relacionados (carne de pollo, de cerdo y de res) se expresan como sigue:

$$\begin{aligned} 5P_p - P_c - 2P_r &= 1 \\ -2P_p + 6P_c - 3P_r &= 3 \\ -2P_p - P_c + 4P_r &= 10 \end{aligned}$$

Calcule el precio de equilibrio, en dólares, para cada mercado.

SOLUCIÓN El sistema $AP = B$, con

$$A = \begin{bmatrix} 5 & -1 & -2 \\ -2 & 6 & -3 \\ -2 & -1 & 4 \end{bmatrix}, \qquad P = \begin{bmatrix} P_c \\ P_p \\ P_b \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 3 \\ 10 \end{bmatrix}$$

puede resolverse invirtiendo A. Si empleamos el algoritmo de inversión de matrices obtenemos

$$P = \begin{bmatrix} P_c \\ P_p \\ P_b \end{bmatrix} = A^{-1}B = \frac{1}{63}\begin{bmatrix} 21 & 6 & 15 \\ 14 & 16 & 19 \\ 14 & 7 & 28 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 10 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$$

Por consiguiente, los precios de equilibrio son \$3 para la carne de pollo, \$4 para la de cerdo y \$5 para la de res. ☐

La inversa con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> inverse(matrix([[1,1],[3,4]]));
```

$$\begin{bmatrix} 4 & -1 \\ -3 & 1 \end{bmatrix}$$

Mathematica

```
In[1]:=
    Inverse[{{1,1},{3,4}}]
Out[1]=
    {{4, -1}, {-3, 1}}
```

MATLAB

```
>> inv([1 1; 3 4])
ans =
    4.0000   -1.0000
   -3.0000    1.0000
```

Ejercicios 3.2

1. Aplique el teorema 4 para determinar las inversas de las matrices siguientes.

$$\begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 7 & 5 \\ 4 & 4 \end{bmatrix}$$

2. Utilice el teorema 4 para explicar por qué las matrices siguientes no son invertibles.

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \qquad \begin{bmatrix} -10 & 20 \\ 20 & -40 \end{bmatrix}$$

3. Obtenga una matriz cuya inversa sea

$$\begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix}$$

4. Determine la inversa de $10A$, si la inversa de A es

$$\begin{bmatrix} 4 & 4 \\ 8 & 6 \end{bmatrix}$$

5. Si $a^2 + b^2 = 1$, aplique el teorema 4 para demostrar que A es invertible. Calcule A^{-1}.

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

6. Encuentre la inversa de

$$\begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

7. Determine $(2A)^{-3}$ si

$$A^3 = \begin{bmatrix} 1 & 1 \\ -5 & -2 \end{bmatrix}$$

8. Encuentre $A^2, A^{-2}, A^3 - A, A^{-3}$ si

$$A = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}$$

9. Obtenga las inversas de las siguientes matrices mediante observación.

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

10. Sin hacer cálculos, demuestre que A no es invertible.

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

[*Sugerencia:* Si B es una inversa hipotética, examine el elemento (3, 3) de AB.]

11. Sin hacer cálculos, compruebe que A no es invertible.

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

[*Sugerencia:* Si B fuera una inversa hipotética, examine el elemento (1, 1) de AB].

En los ejercicios 12 a 16 aplique el algoritmo de inversión de matrices para, en caso de ser posible, calcular la inversa de la matriz dada.

12. **a.** $\begin{bmatrix} -3 & -2 \\ -5 & -3 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}$ **d.** $\begin{bmatrix} -\frac{1}{3} & \frac{1}{3} \\ -\frac{2}{3} & -\frac{1}{3} \end{bmatrix}$

13. **a.** $\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **b.** $\begin{bmatrix} -2 & 0 & -1 \\ 0 & -1 & -2 \\ 0 & 0 & -2 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 2 & -1 \\ 1 & -2 & -1 \\ 1 & 6 & -1 \end{bmatrix}$ **d.** $\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix}$

14. **a.** $\begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix}$ **b.** $\begin{bmatrix} -1 & 0 & 1 \\ 1 & \frac{2}{3} & -\frac{4}{3} \\ 0 & -1 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & -1 \\ 3 & 4 & -3 \end{bmatrix}$ **d.** $\begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & -2 \\ 2 & -1 & 0 \end{bmatrix}$

15. $\begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$

16. $\begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ -1 & 1 & 1 & 2 \\ -1 & 1 & 1 & 1 \end{bmatrix}$

En los ejercicios 17 y 18, resuelva los sistemas calculando primero la inversa de la matriz de coeficientes.

17. **a.** $\begin{aligned} x + y - z &= 1 \\ x \quad\;\; - z &= 2 \\ -x + y \quad\;\; &= 3 \end{aligned}$ **b.** $\begin{aligned} x - \;\; y - z &= 2 \\ x - \;\; y \quad\;\; &= 4 \\ x - 2y - z &= -1 \end{aligned}$

18. **a.** $\begin{aligned} -2x + y \quad\;\; &= a \\ x - y - z &= b \\ x + y - z &= c \end{aligned}$ **b.** $\begin{aligned} -x + y + z &= a \\ -x + \quad\;\; z &= a^2 \\ y - z &= a^3 \end{aligned}$

19. Determine A, si

$$A^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

20. Calcule todos los valores de a para los cuales existe A^{-1}.

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ -1 & -2 & a \end{bmatrix}$$

21. Si c_1, c_2, c_3 y c_4 son escalares distintos de cero, calcule A^{-1} y B^{-1}.

$$A = \begin{bmatrix} 0 & 0 & c_1 \\ 0 & c_2 & 0 \\ c_3 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 0 & c_1 \\ 0 & 0 & c_2 & 0 \\ 0 & c_3 & 0 & 0 \\ c_4 & 0 & 0 & 0 \end{bmatrix}$$

22. Determine $(ABC)^{-1}$ si $A^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$, $B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$, y $C^{-1} = \begin{bmatrix} 3 & 2 \\ -1 & -1 \end{bmatrix}$

23. Calcule $A^{-1}, A^{-2}, A^{-3}, A^{-24}$ y A^{-25} siendo

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

24. Compruebe la identidad $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
25. Determine A y B tales que $(A + B)^{-1} \neq A^{-1} + B^{-1}$.
26. Obtenga A y B tales que $(AB)^{-1} \neq A^{-1}B^{-1}$.
27. Demuestre que una matriz con un renglón de ceros no es invertible.
28. Compruebe que una matriz con una columna de ceros no es invertible.
29. Demuestre que una matriz diagonal A es invertible si y sólo si cada elemento de la diagonal principal es distinto de cero. ¿Qué es A^{-1} en este caso?
30. Si A es invertible y $AB = \mathbf{0}$, demuestre que $B = \mathbf{0}$.
31. Suponga que $AB = \mathbf{0}$ y que $B \neq \mathbf{0}$. Demuestre que A no puede invertirse.
32. Si A es una matriz cuadrada con la propiedad $A^2 = \mathbf{0}$, demuestre que existe la inversa de $I - A$ y es igual a $A + I$.
33. Si A es una matriz cuadrada con la propiedad $A^3 = \mathbf{0}$, compruebe que la inversa de $I - A$ existe y es igual a $A^2 + A + I$.
34. Suponga que $A^2 + 2A - I = \mathbf{0}$. Demuestre que $A^{-1} = A + 2I$.
35. Compruebe el teorema 4.
36. Demuestre la parte 3 del teorema 6.
37. Compruebe el teorema 7.
38. Demuestre la parte 2 del teorema 8.
39. Para el teorema 9:
   a. Demuestre las partes 1 a 3 y la parte 5.
   b. Termine de comprobar la parte 4.
40. Use la inversión de matrices para calcular el precio de equilibrio, en dólares, de cada uno de los artículos relacionados (cereal, galletas y pastelillos) si las condiciones de equilibrio se expresan como sigue:

$$\begin{aligned} -P_{ce} + P_{cr} \phantom{{}- P_{co}} &= -1 \\ P_{ce} \phantom{{}+ P_{cr}} - P_{co} &= 1 \\ P_{ce} + P_{cr} - P_{co} &= 3 \end{aligned}$$

41. Demuestre que AA^T siempre es simétrica.

3.3 Matrices elementales e invertibles

Objetivos del estudiante para esta sección

1. Comprender la relación entre las matrices elementales y las operaciones elementales de renglón.
2. Conocer las diferentes caracterizaciones de matrices invertibles.
3. Entender por qué funciona el algoritmo de inversión de matrices.

En esta sección estudiaremos las matrices elementales y las usaremos para caracterizar a las matrices invertibles en muchas formas interesantes. Además, justificaremos el algoritmo de inversión de matrices.

Matrices elementales

DEFINICIÓN

(Matriz elemental)

Una matriz de $n \times n$ se llama **elemental** si puede obtenerse a partir de la matriz identidad I_n usando una y sólo una operación elemental de renglón (eliminación, escalamiento o intercambio). Así, las matrices elementales siempre son equivalentes a I_n.

Por ejemplo, las matrices

$$E_1 = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \qquad E_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

son elementales, porque cada una se obtiene de I_3 aplicando $R_1 - 3R_2 \to R_1$, $4R_2$ R_2 y $R_1 \Leftrightarrow R_2$, respectivamente. Las matrices elementales, como la E_3, que se obtienen de I intercambiando dos renglones, se llaman **matrices elementales de permutación.**

La razón principal para estudiar las matrices elementales se debe a que multiplica una matriz A por la izquierda por una matriz elemental E; el producto EA es la matriz que se obtiene de A usando la misma operación de renglón que produjo a E partiendo de I_n. Por ejemplo:

$$\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} = \begin{bmatrix} a-3e & b-3f & c-3g & d-3h \\ e & f & g & h \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ 4e & 4f & 4g & 4h \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} = \begin{bmatrix} e & f & g & h \\ a & b & c & d \end{bmatrix}$$

A continuación se observará que las operaciones elementales de renglón son reversibles o **invertibles**; es decir, para cada una de ellas hay otra operación elemental de renglón que invierte el efecto de la primera. (Véase también el ejercicio 6 de la sección 1.2.) Por ejemplo, $\frac{1}{4}R_2 \to R_2$ anula el efecto de $4R_2 \to R_2$. En general,

Operación elemental de renglón	**Operación inversa correspondiente**
$R_i \leftrightarrow R_j$	$R_i \leftrightarrow R_j$
$cR_i \to R_i$	$(1/c)R_i \to R_i$
$R_i + cR_j \to R_i$	$R_i - cR_j \to R_i$

Como las operaciones elementales de renglón son invertibles, es posible recuperar a I_n de una matriz elemental llevando a cabo la operación inversa. Por ejemplo, I_3 se obtiene de E_1, E_2 o E_3 aplicando $R_1 + 3R_2 \to R_1$, $\frac{1}{4}R_2 \to R_2$, o $R_1 \Leftrightarrow R_2$, respectivamente. Esto, a su vez, implica que las matrices elementales son invertibles, porque si E se obtuvo a partir de I_n mediante una operación elemental, y E' es la matriz elemental obtenida de I_n aplicándole la operación inversa, entonces $EE' = I_n$. De igual manera, $E'E = I_n$. Por tanto, E puede invertirse y $E^{-1} = E'$. Por ejemplo, E_1, E_2 y E_3 son invertibles, y

$$E_1^{-1} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \qquad E_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix}, \qquad E_3^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

TEOREMA 10

Toda matriz elemental E tiene una inversa que también es una matriz elemental. E^{-1} se obtiene al invertir la operación elemental de renglón que produjo E a partir de I.

Si las matrices A y B son equivalentes en renglones, entonces podemos determinar B a partir de A mediante una sucesión finita de operaciones elementales de renglón, digamos $O_1, \ldots, O_k$. Sean $E_1, \ldots, E_k$ las matrices elementales que corresponden a esas operaciones. El efecto de la operación O_1 sobre A es igual al producto E_1A. De igual manera, el efecto de O_2 es $E_2(E_1A) = E_2E_1A$. Con el mismo procedimiento se obtiene $B = E_k \cdots E_2E_1A$.

TEOREMA 11

$A \sim B$ si y sólo si hay matrices elementales $E_1, \ldots, E_k$ tales que

$$B = E_k \cdots E_1A$$

En particular, el teorema 11 se aplica a una matriz A y a una forma de escalón U de A. Si U se obtiene de A con operaciones de renglón cuyas matrices elementales correspondientes son $E_1, \ldots, E_k$, entonces

$$U = E_k \cdots E_1A \tag{3.2}$$

Al despejar A se obtiene $A = (E_k \cdots E_1)^{-1}U$, es decir

$$A = E_1^{-1} \cdots E_k^{-1}U \tag{3.3}$$

En cualquiera de las ecuaciones (3.2) y (3.3) *está implícita la reducción de A mediante operaciones de renglón*. El estudio minucioso del ejemplo siguiente le permitirá comprender cómo funciona.

■ EJEMPLO 24 Reduzca la matriz A con operaciones elementales de renglón y exprese U como un producto de matrices elementales a partir de A. Después escriba A en forma de un producto de matrices elementales por U.

$$A = \begin{bmatrix} 1 & 3 & 7 \\ 2 & 6 & 8 \\ 0 & 4 & 3 \end{bmatrix}$$

SOLUCIÓN Si U es la forma de escalón de la matriz en la reducción

$$\begin{bmatrix} 1 & 3 & 7 \\ 2 & 6 & 8 \\ 0 & 4 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 7 \\ 0 & 0 & -6 \\ 0 & 4 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 7 \\ 0 & 4 & 3 \\ 0 & 0 & -6 \end{bmatrix} = U$$

Las operaciones que condujeron a U fueron $-2R_1 + R_2 \to R_2$ y $R_2 \Leftrightarrow R_3$. Así, las matrices elementales correspondientes son

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Por consiguiente y según nuestro análisis, $U = E_2E_1A$:

$$\begin{bmatrix} 1 & 3 & 7 \\ 0 & 4 & 3 \\ 0 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 7 \\ 2 & 6 & 8 \\ 0 & 4 & 3 \end{bmatrix}$$

(Coincide.) De este modo *factorizamos* a U como el producto E_2E_1A. Porque $A = E_1^{-1}E_2^{-1}U$,

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{y} \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

entonces

$$\begin{bmatrix} 1 & 3 & 7 \\ 2 & 6 & 8 \\ 0 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 7 \\ 0 & 4 & 3 \\ 0 & 0 & -6 \end{bmatrix}$$

(Coincide.) Es una factorización de A en términos de una de sus formas de escalón y de las matrices elementales de las operaciones inversas que la produjeron. □

Caracterización de las matrices invertibles

En esta sección caracterizaremos una matriz invertible con varios métodos. Primero, se observará lo siguiente.

TEOREMA 12

> La forma de escalón reducida R de una matriz cuadrada A, por operaciones elementales de renglón, es I, o bien tiene un renglón de ceros.

DEMOSTRACIÓN Si cada renglón de A tiene un pivote, entonces $R = I$, porque A es cuadrada. Por otro lado, si alguno de esos renglones no tiene pivote, el que corresponde a R es cero. □

TEOREMA 13

> Las afirmaciones siguientes son equivalentes:
>
> 1. A es invertible.
> 2. $A \sim I$.
> 3. A es un producto de matrices elementales.

DEMOSTRACIÓN

$1 \Rightarrow 2$. Sea A invertible. Entonces A es cuadrada, digamos de $n \times n$, y $A\mathbf{x} = \mathbf{b}$ es consistente para todos los vectores n $\mathbf{b}$, de acuerdo con el teorema 5. Así, cada renglón de A tiene un pivote, según el teorema 2 de la sección 2.1. Por consiguiente, la forma de escalón reducida por operaciones de renglón de A es I_n (porque A es cuadrada). En conclusión, $A \sim I_n$.

$2 \Rightarrow 3$. Sea $A \sim I_n$ para un n. Entonces A es cuadrada, de $n \times n$, y hay matrices elementales $E_1, \ldots, E_k$ tales que $A = E_k \cdots E_1 I_n = E_k \cdots E_1$, de acuerdo con el teorema 11. Así, A es un producto de matrices elementales.

$3 \Rightarrow 1$. Sea A un producto de matrices elementales, por ejemplo, $A = E_k \cdots E_1$. Entonces $E_1, \ldots, E_k$ son invertibles según el teorema 10. Así, A es invertible por ser un producto de matrices invertibles. □

El algoritmo de inversión de matrices

Ahora emplearemos matrices elementales para explicar por qué funciona el algoritmo de inversión de matrices. Si A es una matriz de $n \times n$ cuya forma de escalón reducida mediante operaciones de renglón es R, entonces existen matrices elementales $E_1, \ldots, E_k$ tales que

$$A = E_k \cdots E_1 R$$

La reducción de A llevada a cabo por medio de operaciones de renglón que corresponden a esas matrices elementales puede describirse con

$$E_k^{-1}A = E_{k-1} \cdots E_1 R, \qquad E_{k-1}^{-1}E_k^{-1}A = E_{k-2} \cdots E_1 R, \quad \ldots, \qquad E_1^{-1} \cdots E_k^{-1}A = R$$

Así, en la reducción de $[A : I]$, la matriz que se obtiene colocando $I_n = I$ junto a A da como resultado

$$[E_1^{-1} \cdots E_k^{-1}A : E_1^{-1} \cdots E_k^{-1}] \text{ o sea } [R : E_1^{-1} \cdots E_k^{-1}]$$

Si R tiene un renglón de ceros, A no es invertible. En caso contrario, R es I y A es invertible. Por consiguiente, $E_1^{-1} \cdots E_k^{-1}A = R = I$, lo cual implica que $A^{-1} = E_1^{-1} \cdots E_k^{-1}$. Así que en este caso

$$[R : E_1^{-1} \cdots E_k^{-1}] = [I : A^{-1}]$$

Llegamos a la conclusión que la reducción de $[A : I]$, que es el algoritmo de inversión de matrices, detecta una matriz no invertible, o calcula su inversa, que es lo que se afirmó en la sección 3.2.

■ EJEMPLO 25 Escriba

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

y A^{-1} como productos de matrices elementales.

SOLUCIÓN A^{-1} se determina como siempre:

$$\begin{bmatrix} 1 & 1 & : & 1 & 0 \\ 1 & 2 & : & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & : & 1 & 0 \\ 0 & 1 & : & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & : & 2 & -1 \\ 0 & 1 & : & -1 & 1 \end{bmatrix}$$

Las matrices elementales que corresponden a las operaciones de renglón $-R_1 + R_2 \to R_2$ y $R_1 - R_2 \to R_1$ son $E_1 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ y $E_2 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$. Por consiguiente,

$$E_2E_1A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = I$$

De este modo,

$$A = (E_2E_1)^{-1} = E_1^{-1}E_2^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

y

$$A^{-1} = E_2E_1 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

son los productos requeridos. □

Una implicación interesante de los teoremas 12 y 13 es el teorema siguiente, que indica que para matrices cuadradas del mismo tamaño, basta *sólo una* de las dos condiciones $AB = I$ y $BA = I$ para asegurar que tanto A (y por consiguiente B) es invertible.

TEOREMA 14

Sean A y B matrices $n \times n$. Si $AB = I$, entonces A y B son invertibles y $A^{-1} = B$, $B^{-1} = A$. En particular, $AB = I$ si y sólo si $BA = I$.

DEMOSTRACIÓN Si R es la forma de escalón reducida mediante operaciones de renglón de A, encontraremos que hay matrices elementales $E_1, \ldots, E_k$ tales que $R = E_n \cdots E_1 A$, de acuerdo con el teorema 11. En consecuencia, $RB = E_n \cdots E_1 AB = E_n \cdots E_1$, puesto que $AB = I$. De aquí se deduce que RB es invertible porque es el producto de matrices invertibles. Así, R no puede tener un renglón de ceros. Por consiguiente, según el teorema 12, $R = I$. De modo que A^{-1} existe y $AB = I$ implica que $A^{-1}AB = A^{-1}I$, es decir $A^{-1} = B$. De aquí, $B^{-1} = A$. $\square$

Ahora caracterizaremos las matrices invertibles en varias formas básicas.

TEOREMA 15

Sea A una matriz de $n \times n$. Las siguientes afirmaciones son equivalentes.

1. A es invertible.
2. $A \sim I_n$.
3. A es un producto de matrices elementales.
4. Hay una matriz B, $n \times n$, tal que $AB = I_n$.
5. Hay una matriz C, $n \times n$, tal que $CA = I_n$.
6. Cada columna de A es columna pivote.
7. Cada renglón de A tiene un pivote.
8. Las columnas de A son linealmente independientes.
9. Los renglones de A son linealmente independientes.
10. Las columnas de A generan a $\mathbf{R}^n$.
11. Los renglones de A generan a $\mathbf{R}^n$.
12. El sistema $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para todo vector n $\mathbf{b}$.
13. El sistema $A\mathbf{x} = \mathbf{b}$ tiene solamente una solución para todo vector n $\mathbf{b}$.
14. El sistema homogéneo $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial.

DEMOSTRACIÓN Según el teorema 13, $1 \Leftrightarrow 2 \Leftrightarrow 3$. Es claro que $2 \Leftrightarrow 4$ y que $2 \Leftrightarrow 5$. También $4 \Rightarrow 2$ y $5 \Rightarrow 2$, por el teorema 14. Por tanto, los enunciados 1 al 5 son equivalentes. También hay equivalencia en las afirmaciones 6 a 12 y 14, de acuerdo con los teoremas 2 y 12 (secciones 2.1 y 2.4), y con los enunciados similares para $m = n$ anteriores al ejemplo 46 (sección 2.5). Además, $2 \Leftrightarrow 6$, porque A es cuadrada. Es claro que $13 \Rightarrow 12$. Por último, demostraremos que $12 \Rightarrow 13$. Supongamos que 12 es cierto. Así, para cada vector n $\mathbf{b}$, el sistema $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución, digamos $\mathbf{v}_1$. Si $\mathbf{v}_2$ es otra solución, entonces $A\mathbf{v}_1 = \mathbf{b} = A\mathbf{v}_2$. Por consiguiente, $A(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$. Así, $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, es decir $\mathbf{v}_1 = \mathbf{v}_2$, porque $12 \Leftrightarrow 14$. Así, la solución de $A\mathbf{x} = \mathbf{b}$ es única. Esto demuestra la afirmación 13. $\square$

UN CONCEPTO QUE CON FRECUENCIA SE MALINTERPRETA Las afirmaciones 12 y 13 del teorema 15 pueden dar la impresión *equivocada* de que si un sistema cuadrado tiene una solución, ésta debe ser única. Eso no es cierto. El sistema $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{b}$ tiene infinitas soluciones para $\mathbf{b}=\begin{bmatrix}1\\1\end{bmatrix}$, pero no las tiene para $\mathbf{b}=\begin{bmatrix}1\\2\end{bmatrix}$. En este caso falla la frase clave para todo vector n $\mathbf{b}$.

El teorema 15 tiene una consecuencia interesante.

TEOREMA 16

Sean A y B matrices $n \times n$. Si A o B no son invertibles, AB tampoco es invertible.

DEMOSTRACIÓN

1. Si B no tiene inversa, entonces $B\mathbf{x} = \mathbf{0}$ posee una solución no trivial, de acuerdo con el teorema 15. Por tanto, $AB\mathbf{x} = \mathbf{0}$ tiene una solución no trivial, lo cual implica que AB es no invertible.
2. Si B es invertible, entonces suponemos que A no tiene inversa. Así que, $A\mathbf{x} = \mathbf{0}$ tiene la solución no trivial, digamos, $\mathbf{v} \neq \mathbf{0}$ por el teorema 15. Por consiguiente, $A\mathbf{v} = \mathbf{0}$. El vector $B^{-1}\mathbf{v}$ no es igual a $\mathbf{0}$ (¿por qué?) y $(AB)B^{-1}\mathbf{v} = \mathbf{0}$. Por consiguiente, $AB\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales. Entonces, AB no es invertible. ☐

Ejercicios 3.3

En los ejercicios 1 y 2, indique cuáles de las matrices son elementales. En cada una de ellas identifique la operación elemental de renglón que la produjo a partir de la matriz identidad.

1. $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$, $B=\begin{bmatrix}-1&-1\\0&1\end{bmatrix}$,

$C=\begin{bmatrix}1&0&-2\\0&1&0\\0&0&1\end{bmatrix}$, $D=\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix}$

2. $E=\begin{bmatrix}2&0\\0&2\end{bmatrix}$, $F=\begin{bmatrix}1&-1\\0&1\end{bmatrix}$,

$G=\begin{bmatrix}1&0&-1\\0&0&1\\0&1&0\end{bmatrix}$, $H=\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&-1&0&0\end{bmatrix}$

3. ¿Qué operación elemental de renglón produce cada una de las matrices elementales siguientes a partir de la matriz identidad del mismo tamaño?

$J=\begin{bmatrix}0&1\\1&0\end{bmatrix}$, $K=\begin{bmatrix}1&0&0\\0&2&0\\0&0&1\end{bmatrix}$

$L=\begin{bmatrix}1&0&-5&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$, $M=\begin{bmatrix}1&0&0\\0&1&0\\-1&0&1\end{bmatrix}$

4. ¿Cuál operación de renglón produce una matriz identidad a partir de J, K, L y M?
5. Escriba las inversas de las operaciones matriciales elementales siguientes.

   a. $R_1 \leftrightarrow R_3$ b. $R_1 + 5R_3 \to R_1$
   c. $(1/2)R_4 \to R_4$ d. $10R_1 + R_3 \to R_3$

6. Para efectuar las operaciones matriciales siguientes, multiplique el lado izquierdo de $A=\begin{bmatrix}1&2&3\\-1&-1&-1\\0&1&0\end{bmatrix}$ por una matriz elemental adecuada:

   a. $R_1 \leftrightarrow R_3$ b. $R_1 - 2R_3 \to R_1$
   c. $-2R_2 \to R_2$ d. $5R_1 + R_3 \to R_3$

7. Escriba A y A^{-1} en forma de productos de matrices elementales.

$$A=\begin{bmatrix}2&1\\-1&0\end{bmatrix}$$

8. Demuestre que la descomposición de una matriz invertible en forma de un producto de matrices elementales *no es*

única: hágalo mediante una segunda factorización de la matriz A en el ejercicio 7.

9. Exprese cada una de las matrices siguientes en forma de productos de matrices elementales.

$$A = \begin{bmatrix} 3 & -6 \\ 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

10. Las matrices A y B son equivalentes. Determine las matrices equivalentes E_1 y E_2 tales que $A = E_2E_1B$.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -4 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 & 3 \\ -1 & -4 & -1 \\ 1 & 2 & 3 \end{bmatrix}$$

11. Explique por qué la matriz siguiente no puede escribirse en forma de un producto de matrices elementales.

$$D = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

12. Escriba la matriz D del ejercicio 11 como un producto de dos matrices elementales por una matriz no invertible en forma de escalón reducida con operaciones de renglón.

13. Aplique el teorema 15 para demostrar que el sistema siguiente sólo tiene una solución para cualquier elección de b_1 y b_2.

$$\begin{aligned} x - y &= b_1 \\ x + 2y &= b_2 \end{aligned}$$

14. Use el teorema 15 para demostrar que el sistema siguiente tiene un número infinito de soluciones.

$$\begin{aligned} x + y + z &= 0 \\ y &= 0 \\ x + z &= 0 \end{aligned}$$

15. Sean c_1, c_2 y c_3 escalares distintos de cero. Exprese A y A^{-1} como productos de matrices elementales.

$$A = \begin{bmatrix} 0 & 0 & c_1 \\ 0 & c_2 & 0 \\ c_3 & 0 & 0 \end{bmatrix}$$

16. Sean c_1 y c_2 escalares distintos de cero. Escriba A en forma de un producto de matrices elementales por una matriz no invertible en forma de escalón reducida.

$$A = \begin{bmatrix} 0 & 0 & c_1 \\ 0 & c_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

17. Compruebe las afirmaciones 1 a la 9 del teorema 15 para

$$\begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 0 \\ -2 & 0 & 3 \end{bmatrix}$$

18. Sea A una matriz elemental de permutación. Demuestre que $A^2 = I$. Deduzca que $A^{-1} = A$.

19. ¿Cierto o falso? Justifique sus elecciones. La matriz A es invertible si

    a. Todas sus columnas son linealmente independientes.
    b. Su forma de escalón reducida es I.
    c. Hay una matriz B tal que $AB = I$.
    d. Hay una matriz C tal que $CA = I$.
    e. Cada renglón y cada columna de A tienen un pivote.
    f. A^T es un producto de matrices elementales.

20. Si A es una matriz, y U es una de sus formas de escalón, ¿puede determinar si A es invertible viendo a U?

Inversas derecha a izquierda

Sea A una matriz $m \times n$. Una matriz B se llama **inversa derecha** de A si $AB = I$. En forma parecida, C es una **inversa izquierda** de A si $CA = I$. Por ejemplo, si $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ y $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, entonces P es una inversa izquierda de Q y Q es una inversa derecha de P, porque

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (3.4)$$

21. Si una matriz A $m \times n$ tiene una inversa derecha B, ¿cuál es el tamaño de B?

22. Si una matriz A $m \times n$ tiene una inversa izquierda C, ¿cuál es el tamaño de C?

23. Demuestre que si A tiene una inversa derecha, entonces A^T tiene una inversa izquierda.

24. Compruebe que si A tiene una inversa izquierda, entonces A^T posee una inversa derecha.

25. Sea A una matriz $m \times n$. Demuestre que las afirmaciones siguientes son equivalentes:

    a. A tiene una inversa derecha.
    b. El sistema $A\mathbf{x} = \mathbf{b}$ es consistente para todos los vectores m $\mathbf{b}$.

c. Cada renglón de A tiene un pivote.
d. Las columnas de A generan a $\mathbf{R}^m$.

(*Sugerencia:* Para demostrar (b) $\Rightarrow$ (a), considere a la matriz $B = [\mathbf{b}_1 \cdots \mathbf{b}_n]$, si $\mathbf{b}_i$ es una solución de $A\mathbf{x} = \mathbf{e}_i$.)

26. Sea A una matriz $m \times n$. Demuestre las equivalencias de las afirmaciones siguientes:

a. A tiene una inversa izquierda.
b. El sistema $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial.
c. Cada columna de A es columna pivote.
d. Las columnas de A son linealmente independientes.

27. Demuestre que si una matriz A $m \times n$ tiene una inversa derecha B y también una inversa izquierda C, entonces es válido lo siguiente.

a. $m = n$.
b. $B = C$.
c. A es invertible.

3.4 Factorización LU

Objetivos del estudiante para esta sección
1. Saber cómo factorizar una matriz como producto de una matriz triangular inferior y una matriz triangular superior.
2. Conocer cómo usar esta factorización para resolver sistemas lineales.

En la sección 3.3 vimos cómo factorizar una matriz como un producto de matrices elementales por una de sus formas de escalón. En general, esta factorización puede ser de mucha utilidad para comprender las propiedades de la matriz. Por ejemplo, suponga que se conoce cómo factorizar una matriz A, $m \times n$ en la forma

$$A = LU$$

donde L es una matriz triangular inferior $m \times m$ y U es una matriz escalonada $m \times n$. Entonces el sistema

$$A\mathbf{x} = \mathbf{b} \tag{3.5}$$

puede resolverse en dos pasos más fáciles. Primero, se despeja $\mathbf{y}$ de la ecuación (3.6):

$$L\mathbf{y} = \mathbf{b} \tag{3.6}$$

y a continuación se despeja $\mathbf{x}$ de la ecuación (3.7):

$$U\mathbf{x} = \mathbf{y} \tag{3.7}$$

De hecho, la solución de estos dos sistemas equivale a resolver el sistema original, porque

$$LU\mathbf{x} = L(U\mathbf{x}) = L\mathbf{y} = \mathbf{b}$$

La ventaja de no resolverlo en forma directa es que la matriz L en (3.6) es un sistema triangular inferior que puede determinarse con facilidad mediante una sustitución directa, y la matriz U en la ecuación (3.7) es triangular superior, que puede resolverse fácilmente con una solución por sustitución hacia atrás.

Una factorización de A como la indicada, es decir, como el producto de una matriz triangular inferior L y una triangular superior U, si existe, se llama **factorización LU**, o **descomposición LU**.

■ EJEMPLO 26 Use la factorización LU de A,

$$A = \begin{bmatrix} 4 & -2 & 1 \\ 20 & -7 & 12 \\ -8 & 13 & 17 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 & 1 \\ 0 & 3 & 7 \\ 0 & 0 & -2 \end{bmatrix} = LU \qquad (3.8)$$

para despejar $\mathbf{x} = (x_1, x_2, x_3)$ en el sistema

$$A\mathbf{x} = \begin{bmatrix} 11 \\ 70 \\ 17 \end{bmatrix} = \mathbf{b}$$

SOLUCIÓN Sea $\mathbf{y} = (y_1, y_2, y_3)$ un nuevo vector de incógnitas. Primero resolveremos el sistema triangular inferior $L\mathbf{y} = \mathbf{b}$,

$$\begin{aligned} y_1 \qquad\qquad &= 11 \\ 5y_1 + \ y_2 \qquad &= 70 \\ -2y_1 + 3y_2 + y_3 &= 17 \end{aligned}$$

por eliminación directa. Como $y_1 = 11$, de la segunda ecuación se obtiene $y_2 = 70 - 5y_1 = 15$, y de la tercera, $y_3 = 17 + 2y_1 - 3y_2 = -6$.

A continuación se resuelve el sistema triangular superior $U\mathbf{x} = \mathbf{y}$,

$$\begin{aligned} 4x_1 - 2x_2 + \ x_3 &= 11 \\ 3x_2 + 7x_3 &= 15 \\ -2x_3 &= -6 \end{aligned}$$

por sustitución hacia atrás se obtiene $x_3 = 3$ de la tercera ecuación, $x_2 = -2$ de la segunda y $x_1 = 1$ de la primera. Por consiguiente, la solución del sistema original es $(1, -2, 3)$.

□

De acuerdo con el ejemplo 26, es claro que una vez que se tiene una factorización LU de A, es muy fácil resolver el sistema $A\mathbf{x} = \mathbf{b}$. Esto es muy útil cuando hay que solucionar *varios* sistemas con la misma matriz A de coeficientes. Así, por ejemplo, si tenemos el sistema $A\mathbf{x} = (5, 21, -20)$, y A corresponde al ejemplo 26, tan sólo resuelve

$$\begin{aligned} y_1 \qquad\qquad &= 5 \\ 5y_1 + \ y_2 \qquad &= 21 \\ -2y_1 + 3y_2 + y_3 &= -20 \end{aligned}$$

para obtener $y_1 = 5, y_2 = -4, y_3 = 2$. A continuación solucionamos

$$\begin{aligned} 4x_1 - 2x_2 + \ x_3 &= 5 \\ 3x_2 + 7x_3 &= -4 \\ -2x_3 &= 2 \end{aligned}$$

y tenemos que $x_1 = 2, x_2 = 1, x_3 = -1$.

Si en verdad las factorizaciones LU son tan útiles, ¿cómo se calculan? La respuesta es sorprendentemente sencilla. De la sección 3.3, recordamos que cualquier matriz A se puede factorizar como

$$A = E_1^{-1} \cdots E_k^{-1} U$$

donde U es una forma de escalón de A, y $E_1, \ldots, E_k$ son las matrices elementales que corresponden a las operaciones elementales de renglón usadas para reducir A a U.

Es evidente que cualquier matriz A puede reducirse a la forma de escalón *sin operaciones de escalamiento*, sólo con intercambios y eliminaciones. Si nuestra matriz A se puede reducir empleando sólo eliminaciones (lo cual no siempre es posible), entonces las matrices $E_1, \ldots, E_k$ son *triangulares inferiores* y contienen al número 1 en la diagonal principal. A esas matrices se les llama diagonales inferiores **unitarias**. Es fácil demostrar que $E_1^{-1} \cdots E_k^{-1}$ son también triangulares inferiores (vea los ejercicios 25 y 28). Si se define

$$L = E_1^{-1} \cdots E_k^{-1}$$

entonces $A = LU$ es una descomposición LU de A.

Por ejemplo, la reducción

$$\begin{bmatrix} 4 & -2 & 1 \\ 20 & -7 & 12 \\ -8 & 13 & 17 \end{bmatrix} \sim \begin{bmatrix} 4 & -2 & 1 \\ 0 & 3 & 7 \\ 0 & 9 & 19 \end{bmatrix} \sim \begin{bmatrix} 4 & -2 & 1 \\ 0 & 3 & 7 \\ 0 & 0 & -2 \end{bmatrix}$$

produce la matriz U de la ecuación (3.8), y con operaciones elementales de renglón se obtiene

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix}$$

Por consiguiente,

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$$

Así,

$$L = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & 3 & 1 \end{bmatrix}$$

En consecuencia,

$$A = \begin{bmatrix} 4 & -2 & 1 \\ 20 & -7 & 12 \\ -8 & 13 & 17 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 & 1 \\ 0 & 3 & 7 \\ 0 & 0 & -2 \end{bmatrix} = LU$$

Un examen más detallado de L muestra que no hay necesidad de calcular inversas ni productos. De hecho, podemos determinarla directamente de las eliminaciones. En primer lugar, es una matriz triangular inferior de tamaño $m \times m$, cuando A es de tamaño $m \times n$. Después, su diagonal está formada por números 1. El elemento $(1, 2)$ es 5, y puede obtenerse con la operación $R_2 - 5R_1 \to R_2$, que se usa para obtener un cero en $(1, 2)$. En L usamos 5 en lugar de -5, porque fue necesario invertir E_1. De forma parecida, -2 se obtiene con la operación $R_3 + 2R_1 \to R_3$, y 3 se obtiene con $R_3 - 3R_1 \to R_3$ en la segunda etapa de la reducción.

■ EJEMPLO 27 Determine una factorización LU de A.

$$A = \begin{bmatrix} 2 & 3 & -1 & 4 & 1 \\ -6 & -6 & 5 & -11 & -4 \\ 4 & 18 & 6 & 14 & -1 \\ -2 & -9 & -3 & 4 & 9 \end{bmatrix}$$

SOLUCIÓN El tamaño de L es 4×4, y entonces

$$A \sim \begin{bmatrix} 2 & 3 & -1 & 4 & 1 \\ 0 & 3 & 2 & 1 & -1 \\ 0 & 12 & 8 & 6 & -3 \\ 0 & -6 & -4 & 8 & 10 \end{bmatrix}, \qquad \text{así } L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & ? & 1 & 0 \\ -1 & ? & ? & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 3 & -1 & 4 & 1 \\ 0 & 3 & 2 & 1 & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 10 & 8 \end{bmatrix}, \qquad \text{así } L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 \\ -1 & -2 & ? & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 3 & -1 & 4 & 1 \\ 0 & 3 & 2 & 1 & -1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} = U, \qquad \text{así } L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 \\ -1 & -2 & 5 & 1 \end{bmatrix}$$

■ EJEMPLO 28 Encuentre una factorización LU de A.

$$A = \begin{bmatrix} 2 & 3 & -1 \\ -6 & -6 & 5 \\ 4 & 18 & 6 \\ -2 & -9 & -3 \end{bmatrix}$$

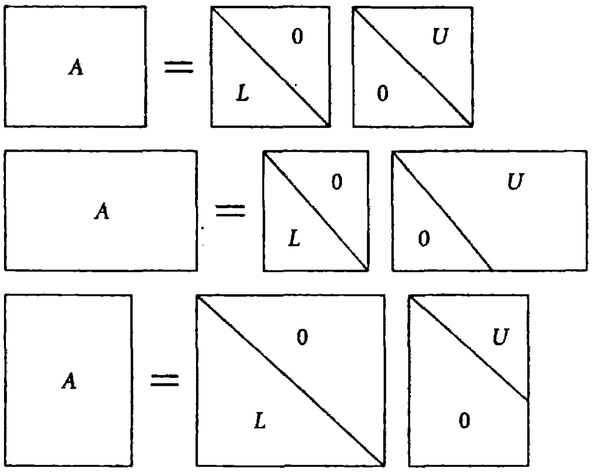
SOLUCIÓN El tamaño de L es 4×4, y

$$A \sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 3 & 2 \\ 0 & 12 & 8 \\ 0 & -6 & -4 \end{bmatrix}, \qquad \text{así } L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & ? & 1 & 0 \\ -1 & ? & ? & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = U, \qquad \text{así } L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 \\ -1 & -2 & ? & 1 \end{bmatrix}$$

En este caso no queda más eliminación, pero como el elemento (4, 3) de L corresponde a la operación $R_4 - 0R_3 \to R_4$, entonces

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{bmatrix}$$



Nuestro análisis, hasta ahora, indica lo siguiente.

TEOREMA 17

Si A es una matriz $m \times n$ que puede reducirse a la forma de escalón U sólo con eliminaciones, entonces A tiene una factorización LU. En particular, A puede factorizarse como

$$A = LU$$

Aquí, L es una matriz triangular inferior de $m \times m$ donde sólo hay números 1 en la diagonal principal. El (i, j)-ésimo elemento l_{ij} $(i > j)$ de L proviene de la operación $R_i - l_{ij}R_j \to R_i$, que se usó para obtener 0 en esta posición durante la reducción.

OBSERVACIONES

1. A veces, a los elementos de L situados debajo de la diagonal principal se les llama **multiplicadores de Gauss.**
2. Si A es cuadrada, usamos un tipo de factorización particular denominada **de Doolittle**. Hay otra versión normal, en la que la matriz triangular *superior* U tiene números 1 en su diagonal principal, y se llama **factorización de Crout.**
3. Los programas de cómputo que determinan las factorizaciones LU emplean la **sobreescritura**. Calculan L y U en forma simultánea, y sobreescriben en la matriz original, de modo que la parte de A situada abajo de la diagonal se transforma en L, y arriba de la diagonal en U. Es muy importante la sobreescritura en las matrices grandes, porque ahorra espacio de memoria. Este ahorro puede incrementarse cuando no se almacenan explícitamente los unos en la diagonal principal. A continuación presentamos un ejemplo de reducción LU con sobreescritura gradual en la matriz original. Los números en cuadro son los elementos de L debajo de la diagonal. El resto de ellos corresponden a U en la diagonal y arriba de ella.

$$\begin{bmatrix} 4 & -2 & 1 \\ 20 & -7 & 12 \\ -8 & 13 & 17 \end{bmatrix} \to \begin{bmatrix} 4 & -2 & 1 \\ \boxed{5} & 3 & 7 \\ \boxed{-2} & 9 & 19 \end{bmatrix} \to \begin{bmatrix} 4 & -2 & 1 \\ \boxed{5} & 3 & 7 \\ \boxed{-2} & \boxed{3} & -2 \end{bmatrix}$$

Eficiencia de cómputo con LU

Si A es una matriz cuadrada, digamos $n \times n$, puede demostrarse que para resolver el sistema lineal $A\mathbf{x} = \mathbf{b}$ con factorización LU es preciso efectuar aproximadamente $2n^3/3$ operaciones, cuando n es grande.[5] De esas operaciones, $2n^2$ se llevan a cabo durante las eliminaciones directa y hacia atrás. Para tener una idea de la utilidad que tiene LU, supongamos que necesitamos solucionar dos sistemas con 500 ecuaciones y 500 incógnitas, y con la misma matriz A de coeficientes. Si se usara la eliminación de Gauss, se requerirían $2n^3/3 = 2 \cdot 500^3/3$ operaciones por sistema, un total de unas 160 millones de operaciones. Sin embargo, si se empleara una factorización LU de A para resolver el primer sistema ($2 \cdot 500^3/3$ operaciones), el segundo sistema sólo requeriría las eliminaciones directa y hacia atrás, otras $2n^2 = 2 \cdot 500^2$ operaciones, un total de aproximadamente 83 millones de ellas. Es evidente que si hay más sistemas o son más grandes, los ahorros en cómputo son verdaderamente enormes.

Cuándo son necesarios los intercambios

Hasta ahora, nuestra factorización LU existe sólo si la matriz A puede reducirse a la forma de escalón sólo con eliminaciones. ¿Y si esto no es posible? Por ejemplo para reducir $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ con operaciones de renglón se necesita un intercambio. Como LU es demasiado útil como para ignorarla aun en estos casos, ofreceremos una descripción *breve* para ilustrar los *conceptos* que intervienen.[6]

Recordará el lector que el intercambio de dos renglones de una matriz A puede expresarse como P_iA, donde P_i es la matriz elemental de permutación que corresponde a ese intercambio. Si durante una reducción de A con operaciones de renglón primero se llevan a cabo todos los intercambios $P_1, \ldots, P_k$, entonces es posible reducir la matriz $P_k \cdots P_1A$ con operaciones de renglón que sean sólo eliminaciones, de modo que se permita la factorización LU. La matriz $P = P_k \cdots P_1$, que es un producto de matrices elementales de permutación, se llama **matriz de permutación.** Lo primero que debe hacerse es calcular P y a continuación determinar una factorización LU para PA.

$$PA = LU:$$

Para ilustrarlo, a continuación veremos una eliminación de Gauss que requiere dos intercambios.

$$A = \begin{bmatrix} 0 & 0 & 4 \\ 1 & 2 & 3 \\ 1 & 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 1 & 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 2 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & -2 \\ 0 & 0 & 4 \end{bmatrix}$$

Las matrices elementales de permutación para los intercambios son

$$P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

---

[5] Es exactamente la cantidad de operaciones en la eliminación de Gauss.

[6] Los detalles teóricos y las implementaciones eficientes se pueden encontrar en textos más avanzados sobre álgebra lineal numérica, o de análisis numérico.

Primero se calcula la matriz P de permutación:

$$P = P_2P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Después se obtiene PA:

$$PA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 4 \\ 1 & 2 & 3 \\ 1 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

Por último, se determina la descomposición LU de PA:

$$PA = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & -2 \\ 0 & 0 & 4 \end{bmatrix} = LU$$

Aun cuando A por sí misma no tenía factorización LU, la factorización LU de PA es *igualmente útil.*

■ EJEMPLO 29 Emplee $PA = LU$ como antes para resolver el sistema

$$\begin{bmatrix} 0 & 0 & 4 \\ 1 & 2 & 3 \\ 1 & 4 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 12 \\ 14 \\ 12 \end{bmatrix} = \mathbf{b}$$

SOLUCIÓN Primero se multiplica $A\mathbf{x} = \mathbf{b}$ por P y por la izquierda.

$$PA\mathbf{x} = P\mathbf{b} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ 14 \\ 12 \end{bmatrix} = \begin{bmatrix} 14 \\ 12 \\ 12 \end{bmatrix}$$

A continuación se aplica la factorización LU de PA para solucionar el *nuevo* sistema. $PA\mathbf{x} = P\mathbf{b}$. Se resuelve el sistema triangular inferior $L\mathbf{y} = P\mathbf{b}$,

$$\begin{aligned} y_1 &= 14 \\ y_1 + y_2 &= 12 \\ y_3 &= 12 \end{aligned}$$

para obtener $y_1 = 14$, $y_2 = -2$ y $y_3 = 12$. A continuación se resuelve el sistema triangular superior, $U\mathbf{x} = \mathbf{y}$,

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 14 \\ 2x_2 - 2x_3 &= -2 \\ 4x_3 &= 12 \end{aligned}$$

con el que se obtiene $x_1 = 1$, $x_2 = 2$ y $x_3 = 3$, que es la solución del sistema original.

OBSERVACIÓN En algunos paquetes informáticos de matemáticas hay rutinas que determinan la factorización LU de una matriz A. Por lo general, lo que se calcula es una matriz de permutación P junto con los factores L y U de PA. Observe que si A se puede reducir sin intercambios, P es justamente la matriz identidad.

LU con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> LUdecomp(matrix([[4,-2,1],[20,-7,12],
>          [-8,13,17]]),L='l',U='u'):
> evalm(l);evalm(u);
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 & 1 \\ 0 & 3 & 7 \\ 0 & 0 & -2 \end{bmatrix}$$

Mathematica

```
In[1]:=
    LUDecomposition[{{4,-2,1},{20,-7,12},{-8,13,17}}]
Out[1]=    1    1              3              6
    {{{4, -, -(-)}, {20, 7, -(-)}, {-8, 19, -}}, {1, 3, 3}}
           2    4              7              7
```

MATLAB

```
>> [L,U] = lu([4 -2 1; 20 -7 12; -8 13 17])
 L =
    0.2000    -0.0588    1.0000
    1.0000          0         0
   -0.4000     1.0000         0
 U =
   20.0000    -7.0000   12.0000
         0    10.2000   21.8000
         0          0   -0.1176
```

Ejercicios 3.4

En los ejercicios 1 a 5 determine la solución del sistema $A\mathbf{x} = \mathbf{b}$, donde A está factorizada como LU. No hay necesidad de calcular A en forma explícita.

1. $\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 0 & -1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -11 \\ 32 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & -7 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 12 \\ 46 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & -2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 7 \\ -3 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -7 & 3 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & -5 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 3 \\ 9 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -4 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 1 \\ 0 & 5 & -1 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 6 \\ 22 \\ -13 \end{bmatrix}$

En los ejercicios 6 a 13 obtenga una factorización LU de la matriz.

6. $\begin{bmatrix} 4 & 1 \\ 12 & 2 \end{bmatrix}$

7. $\begin{bmatrix} 2 & 1 \\ -10 & -12 \end{bmatrix}$

8. $\begin{bmatrix} 2 & -2 & 1 \\ -8 & 11 & -5 \\ 4 & -13 & 3 \end{bmatrix}$

9. $\begin{bmatrix} -1 & 2 & 1 \\ 4 & -5 & -5 \\ -7 & 5 & 5 \end{bmatrix}$

10. $\begin{bmatrix} 4 & 1 & 1 \\ -12 & 2 & -4 \\ 20 & -5 & 10 \end{bmatrix}$

11. $\begin{bmatrix} 4 & 1 & 1 & 2 \\ -12 & -1 & -4 & -4 \\ 0 & -4 & 5 & -2 \\ 20 & 3 & 6 & 7 \end{bmatrix}$

12. $\begin{bmatrix} 4 & 1 & 1 & 2 & 1 \\ -12 & -1 & -4 & -4 & -1 \\ 4 & -3 & 3 & 0 & -4 \end{bmatrix}$

13. $\begin{bmatrix} 4 & 1 & 1 \\ -12 & -1 & -4 \\ 0 & -4 & 5 \\ 20 & 3 & 6 \end{bmatrix}$

En los ejercicios 14 a 17 determine la solución del sistema $A\mathbf{x} = \mathbf{b}$, empleando una factorización LU de la matriz de coeficientes A.

14. $\begin{bmatrix} 5 & 1 \\ -10 & -3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

15. $\begin{bmatrix} 2 & 1 \\ 14 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 6 \\ -8 \end{bmatrix}$

16. $\begin{bmatrix} 2 & 1 & 1 \\ 12 & 11 & 5 \\ -2 & 9 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 17 \\ 18 \end{bmatrix}$

17. $\begin{bmatrix} 2 & 1 & 1 & 2 \\ 0 & 1 & -1 & 2 \\ -4 & 0 & -1 & 2 \\ 0 & 5 & -5 & 12 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ 1 \\ -4 \\ 5 \end{bmatrix}$

En los ejercicios 18 a 20 encuentre una matriz de permutación P y una factorización LU de PA.

18. $A = \begin{bmatrix} 0 & 3 \\ -5 & 4 \end{bmatrix}$

19. $A = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 2 & -4 \\ 2 & -5 & 1 \end{bmatrix}$

20. $A = \begin{bmatrix} 0 & 0 & 2 \\ -1 & 5 & -2 \\ 3 & 6 & 7 \end{bmatrix}$

En los ejercicios 21 a 23 resuelva el sistema $A\mathbf{x} = \mathbf{b}$ empleando una factorización $PA = LU$.

21. $\begin{bmatrix} 0 & 3 & -1 \\ 2 & 0 & 1 \\ 2 & -6 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ -1 \\ -1 \end{bmatrix}$

22. $\begin{bmatrix} 0 & 3 & -1 \\ 0 & 0 & 1 \\ 2 & -6 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ -10 \end{bmatrix}$

23. $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & -4 \\ 2 & -5 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 4 \\ -8 \end{bmatrix}$

24. Compruebe que el producto de dos matrices triangulares inferiores es triangular inferior.

25. Demuestre que el producto de dos matrices triangulares inferiores unitarias es una matriz triangular inferior unitaria.

26. Compruebe que una matriz triangular inferior es invertible si y sólo si todos sus elementos diagonales son distintos de cero.

27. Demuestre que la inversa de una matriz triangular inferior invertible también es triangular inferior.

28. Demuestre que la inversa de una matriz triangular inferior unitaria también es triangular inferior unitaria.

29. **(Unicidad)** Suponga que A es invertible con dos factorizaciones LU, LU y $L'U'$ (L y L' son triangulares inferiores *unitarias*). Demuestre que $L = L'$ y que $U = U'$.

3.5 Aplicaciones

Objetivos del estudiante para esta sección.
1. Comprender las aplicaciones que aquí se presentan.
2. Darse cuenta de la utilidad que pueden tener las matrices.

Teoría de las gráficas

Matrices de adyacencia de gráficas y diagráficas

El álgebra de matrices tiene aplicaciones importantes en la teoría de las gráficas.[7] En la actualidad, las gráficas constituyen los métodos principales en la investigación de operaciones, ingeniería eléctrica, programación de computadoras y su conexión en redes, administración de empresas, sociología, economía, mercadotecnia y redes de comunicaciones. Esta lista de aplicaciones podría ser infinita.

Una **gráfica** es un conjunto de puntos llamados **vértices** o **nodos**, junto con un conjunto de "líneas" llamadas **aristas** o **ramas** que unen algunos pares de nodos. Dos nodos P y Q conectados con una rama e se llaman **adyacentes** o **vecinos**, y se dice que P y Q son **incidentes** a e. Una rama que parte de un nodo y regresa a este mismo se llama **lazo** o **bucle**. Los nodos pueden estar conectados con más de una rama, en cuyo caso se dice que tienen **rama múltiple**.

Un ejemplo de gráfica de rama múltiple, llamada también **multigráfica**, es una red de comunicaciones en la que dos puntos pueden estar conectados con varias líneas telefónicas.

■ EJEMPLO 30 La figura 3.1 muestra tres gráficas, G_1, G_2 y G_3. G_1 es una muligráfica con bucles. Los nodos A y B son adyacentes, mientras que A y E no lo son. El nodo D tampoco está conectado en algún otro. El nodo P de G_2 es incidente a las ramas e_1, e_5 y e_4. El nodo Q no es incidente a la rama e_5. El nodo casa es adyacente al resto de los nodos de G_3. □

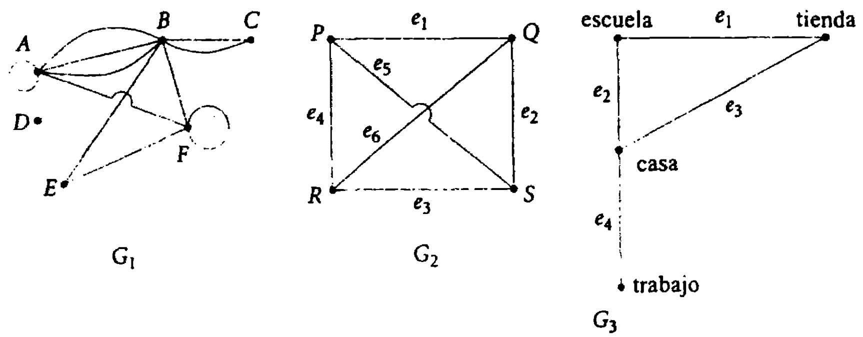


Figura 3.1 Gráficas.

[7] N. del T.: En los libros menos recientes puede aparecer la palabra "redes" en lugar de "gráficas."

Las matrices se emplean en el estudio de las gráficas. La fácil programación de operaciones matriciales en computadora permite estudiar el comportamiento de gráficas muy grandes. (Por ejemplo, en su aplicación a redes telefónicas, una gráfica puede tener decenas de miles de nodos.)[8]

La **matriz de una gráfica** es de tamaño $n \times n$, cuyo (i, j)-ésimo elemento es la cantidad de ramas que unen al i-ésimo con el j-ésimo nodo.

■ EJEMPLO 31 Ordenaremos los nodos de G_1, G_2 y G_3 de la figura 3.1 como A, B, C, D, E, F, P, Q, R, S y escuela, tienda, casa y trabajo, respectivamente. Entonces, las matrices correspondientes son

$$\begin{bmatrix} 1 & 3 & 0 & 0 & 0 & 1 \\ 3 & 0 & 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

□

La matriz de una gráfica no se usa con tanta frecuencia como la matriz de adyacencia de la gráfica, cuyos elementos sólo son unos y ceros. (Los teóricos de las gráficas las denominan a menudo *matriz* (0, 1).)

DEFINICIÓN

La **matriz de adyacencia,** $A(G)$, de una gráfica G es aquella cuyo (i, j)-ésimo elemento es 1 si los nodos i y j son adyacentes, y cero si no lo son.

■ EJEMPLO 32 Las matrices de adyacencia de las gráficas G_1, G_2 y G_3 son

$$A(G_1) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad A(G_2) = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

$$A(G_3) = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

□

Cuando las gráficas son grandes con frecuencia se facilita más extraer información de la matriz de adyacencia que de la gráfica misma, la cual puede ser visualmente complicada.

---

[8] Para conocer más información, véase el artículo de Frederic Bien: "Construction of Telephone Networks by Group Representations", *Notices of American Mathematical Society*, 36 (enero de 1989): 5-22.

Supongamos que hay una gráfica sin bucles y sin ramas múltiples, como G_2 o G_3. Fijemos dos vértices, P y Q. Una **caminata (o recorrido) de longitud** m de P a Q es una sucesión de vértices o nodos

$$P = V_1, V_2, \ldots, V_m, V_{m+1} = Q$$

de manera que V_i, V_{i+1} son adyacentes para toda i entre 1 y m.

Por ejemplo, P, Q, S, R, P es una caminata de longitud 4 en G_2. También casa, escuela y tienda es una caminata, pero su longitud es 2 en G_3.

El resultado siguiente es útil en aplicaciones a la física e ingeniería, en especial las redes.

TEOREMA 18

> La cantidad de caminatas de longitud m del nodo i al nodo j en una gráfica G es igual al (i, j)- ésimo elemento de $A(G)^m$.

■ EJEMPLO 33 Las cantidades de caminatas de longitud 2 y 3 en G_3 pueden observarse en los elementos de $A(G_3)^2$ y $A(G_3)^3$, respectivamente.

$$A(G_3)^2 = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}, \qquad A(G_3)^3 = \begin{bmatrix} 2 & 3 & 4 & 1 \\ 3 & 2 & 4 & 1 \\ 4 & 4 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix} \qquad \square$$

La cantidad de caminatas de longitud 2 de la casa a la tienda es 1, porque la única forma de llegar allí es pasando por la escuela. Esto también se comprueba en el elemento (3, 2) de $A(G_3)^2$, que es 1. En forma parecida, la cantidad de caminatas de longitud 2 que empiezan en la escuela y regresan a este mismo punto es 3, porque se puede ir a la tienda y regresar, a la escuela y regresar o al trabajo y regresar. Esta cantidad es el elemento (3, 3) de $A(G_3)^2$.

Las únicas caminatas de longitud 3 de la escuela a la tienda (a pesar de lo complicado que parezcan) son: escuela-tienda-escuela-tienda, escuela-casa-escuela-tienda y escuela-tienda-casa-tienda, un total de 3 caminos. Esta cantidad es el elemento (1, 2) de $A(G_3)^3$.

COROLARIO 19

> La cantidad de caminatas de longitud 1, o 2, . . ., o m de P a Q es igual al (i, j)-ésimo elemento de la matriz $A(G) + A(G)^2 + \cdots + A(G)^m$. Esto también se interpreta como la cantidad de accesos del j-ésimo nodo desde el i-ésimo, en 1, o 2, . . ., o m etapas.

■ EJEMPLO 34 La cantidad de accesos en 1 o 2 etapas al j-ésimo nodo, desde el i-ésimo en G_3, es el (i, j)-ésimo elemento de $A(G_3) + A(G_3)^2$:

$$A(G_3) + A(G_3)^2 = \begin{bmatrix} 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 2 & 2 & 3 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Por ejemplo, podemos caminar de la escuela a casa con longitudes menores o iguales a 2 en dos formas: directo o pasando por la tienda. Este número es el elemento (1, 3) de $A(G_3) + A(G_3)^2$.

DEFINICIÓN

Una **diagráfica**, o **gráfica dirigida**, es una gráfica cuyas ramas son segmentos de recta *dirigidos*.

■ EJEMPLO 35 La figura 3.2 muestra 3 diagráficas, D_1, D_2 y D_3. (La diagráfica D_3 podría representar calles en uno y dos sentidos en un área del centro de una ciudad.)

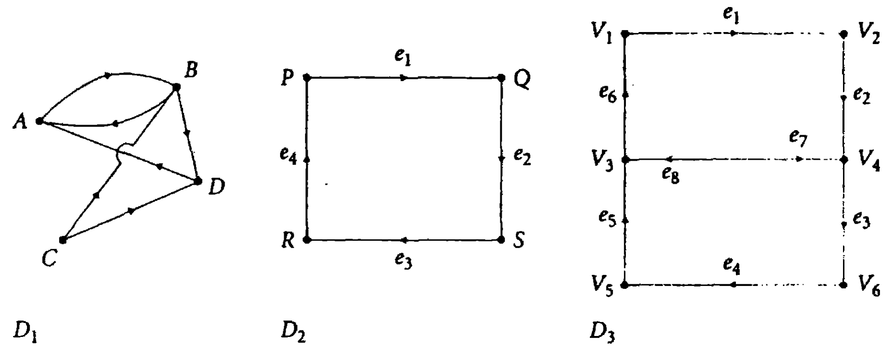


Figura 3.2 Diagráficas.

DEFINICIÓN

La **matriz de adyacencia** $A(D)$ de una diagráfica D es la matriz cuyo (i, j)-ésimo elemento es 1 si hay cuando menos una rama dirigida que conecte al i-ésimo con el j-ésimo vértice, y será cero si están conectados.

■ EJEMPLO 36 Ordenaremos los nodos de D_1, D_2 y D_3 de la figura 3.2 como A, B, C, D, P, Q, R, S y V_1, V_2, V_3, V_4, V_5 y V_6, respectivamente. Entonces las matrices de adyacencia correspondientes son

$$A(D_1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \qquad A(D_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$A(D_3) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Aquí también es posible definir una *caminata* de longitud m desde el i-ésimo hasta el j-ésimo nodo en las diagráficas. La única diferencia es que las ramas son dirigidas. De nuevo se cuenta con un teorema sobre la cantidad de caminatas de longitud m.

TEOREMA 20

La cantidad de caminatas de longitud m del nodo i al j en una diagráfica D es igual al (i, j)-ésimo elemento de $A(D)^m$.

■ EJEMPLO 37 Demuestre que la cantidad de caminatas de longitud 4 desde cualquier nodo de D_2 para regresar al mismo es 1.

SOLUCIÓN Debido a la dirección de las ramas, se necesita caminar en las cuatro ramas para rodear una vez y regresar al mismo vértice. Esto también se confirma con el cálculo de $A(D_2)^4$, que en este caso es la matriz identidad I_4:

$$A(D_2)^4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$ □

Matrices de incidencia de gráficas y diagráficas; gráficas de línea

Sea G una gráfica con el conjunto de nodos $\{V_1, V_2, \ldots, V_m\}$, y el conjunto de ramas $\{e_1, e_2, \ldots, e_n\}$. La **matriz de incidencia** $I(G)$ de G es la matriz $m \times n$ $I(G) = [a_{ij}]$ cuyos elementos son

$$a_{ij} = \begin{cases} 1, & \text{si } V_i \text{ y } e_j \text{ son incidentes} \\ 0, & \text{en cualquier otro caso} \end{cases}$$

de modo que $I(G)$ tiene la forma

$$I(G) = \begin{array}{c} V_1 \\ V_2 \\ \vdots \\ V_m \end{array} \left\{ \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{e_1 \quad e_2 \quad \cdots \quad e_n} \right.$$

cuyos elementos a_{ij} son 0 o 1.

■ EJEMPLO 38 Para G_2 y G_3 de la figura 3.1, tenemos

$$I(G_2) = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}, \qquad I(G_3) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$ □

DEFINICIÓN

La *gráfica de línea* de una gráfica G es una gráfica $L(G)$ cuyos nodos tienen correspondencia biunívoca con las ramas de G. Dos nodos de $L(G)$ son adyacentes sólo cuando las ramas correspondientes de G son incidentes a un vértice común.

■ EJEMPLO 39 La figura 3.3 muestra tres gráficas y sus gráficas de línea.

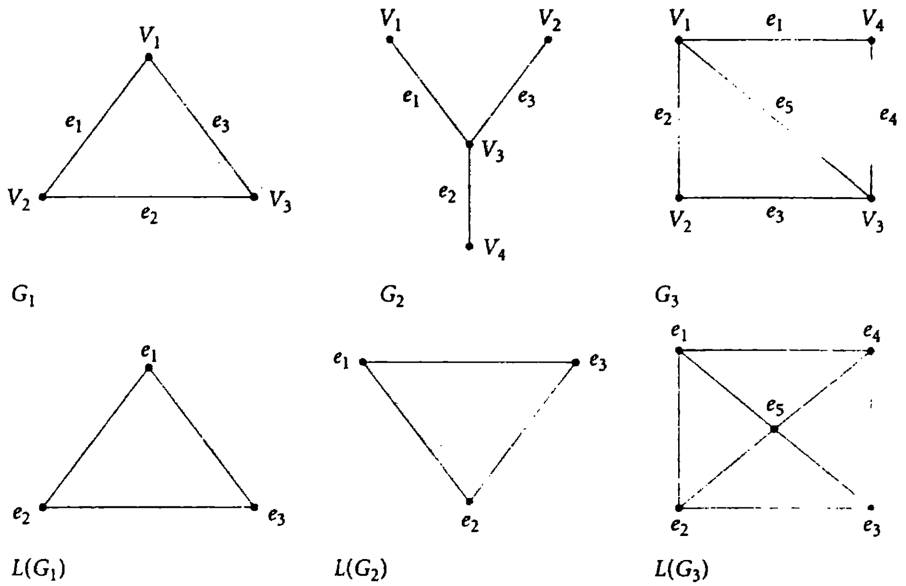


Figura 3.3 Gráficas de línea o gráficas.

Hay una relación interesante entre la matriz de incidencia de una gráfica G y la matriz de adyacencia de su gráfica de línea.

TEOREMA 21

Sea G una gráfica con n ramas. Entonces

$$A(L(G)) = I(G)^T I(G) - 2I_n$$

■ EJEMPLO 40 Compruebe el teorema para G_1 de la figura 3.3.

SOLUCIÓN En este caso,

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = A(L(G_1))$$

También podemos definir las matrices de incidencia de las diagráficas. Sea D una diagráfica con el conjunto de nodos $\{V_1, V_2, \ldots, V_m\}$ y el conjunto de ramas dirigidas $\{e_1, e_2, \ldots, e_n\}$. La **matriz de incidencia** $I(D)$ de D es la matriz $I(D) = [d_{ij}]$ $m \times n$, con elementos:

$$d_{ij} = \begin{cases} 1, \text{ si } e_j \text{ apunta a } V_i \\ -1, \text{ si } e_j \text{ comienza en } V_i \\ 0, \text{ en cualquier otro caso} \end{cases}$$

de modo que $I(D)$ tiene la forma

$$I(D) = \begin{array}{c} V_1 \\ V_2 \\ \vdots \\ V_m \end{array} \left\{ \underbrace{\begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \\ d_{m1} & d_{m2} & \cdots & d_{mn} \end{bmatrix}}_{\begin{array}{cccc} e_1 & e_2 & \cdots & e_n \end{array}} \right.$$

y cuyos elementos d_{ij} son 0, 1 o -1.

■ EJEMPLO 41 La matriz de incidencia de la diagráfica D_4 de 3 nodos (figura 3.4) es

$$I(D_4) = \begin{array}{c} V_1 \\ V_2 \\ V_3 \end{array} \underbrace{\begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}}_{\begin{array}{ccc} e_1 & e_2 & e_3 \end{array}}$$

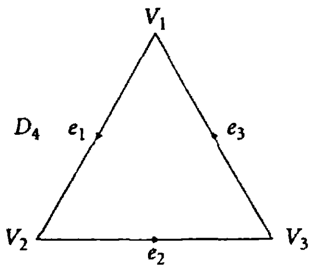


Figura 3.4 Una diagráfica con 3 nodos.

Sociología y psicología

Gráficas de dominancia

Los sociólogos y los psicólogos emplean gráficas para determinar los diversos tipos de relaciones, como la influencia, la dominancia y la comunicación en grupos. Supongamos que en un grupo, para cada par de miembros V_i y V_j, V_i influye en (o domina a) V_j, o V_j influye en V_i, o bien no hay influencia directa entre V_i y V_j. Esta situación puede describirse con una diagráfica D que tiene *cuando mucho una rama dirigida* que une a dos nodos cualesquiera. A esta gráfica se le llama diagráfica de **dominancia**.

■ EJEMPLO 42 La figura 3.5 muestra las relaciones de dominancia entre 7 individuos, $V_1, \ldots, V_7$. ☐

La matriz de adyacencia de una diagráfica de dominancia muestra la información acerca de las relaciones de influencia en un grupo. Los renglones con más unos representan a los miembros del grupo con mayor influencia. Las caminatas de longitud 1 simbolizan una influencia directa, mientras que las de longitud mayor que 1 representan una influencia indirecta. Entonces, la n-ésima potencia de la matriz de adyacencia expresa la influencia indirecta de un miembro sobre otro en n etapas, o en la n-ésima etapa. Examinando $A(D)$ de la diagráfica en la figura 3.5 se ve que V_1 es el miembro con mayor influencia directa, porque el primer renglón es el que tiene más cantidad de unos. Sin embargo, al observar detenidamente $A(D)^2$ nos percataremos que V_5 tiene más influencia en 2 etapas que V_1. Esto también es evidente en la gráfica, porque V_5 influye en V_2, V_3 y V_7 en dos etapas, mientras que V_1 sólo tiene influencia en V_4 y V_6.

$$A(D) = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \qquad A(D)^2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

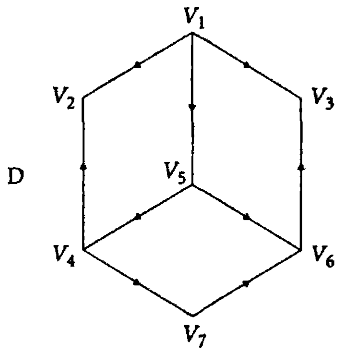


Figura 3.5 Una gráfica de dominancia.

Matrices estocásticas y doblemente estocásticas; procesos de Markov

Las matrices estocásticas corresponden a un tipo especial y se usan con frecuencia en el estudio de fenómenos aleatorios, en teoría de la probabilidad y estadística.

DEFINICIÓN

Una **matriz estocástica** es una matriz cuadrada con elementos reales no negativos, en las cuales cada columna[9] suma 1. Y se dice que es **doblemente estocástica** cuando también cada uno de sus renglones suma 1.

[9] Algunos autores consideran a las matrices estocásticas como aquellas en las que sus *renglones* suman 1.

Observe que una matriz estocástica tiene como elementos a números entre 0 y 1. (¿Por qué?) Asimismo, la transpuesta de una matriz doblemente estocástica también es doblemente estocástica.

■ EJEMPLO 43 Las matrices siguientes son estocásticas. Además, las matrices C y D son doblemente estocásticas.

$$A = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{3}{4} & \frac{1}{6} \\ \frac{1}{4} & \frac{5}{6} \end{bmatrix} \qquad C = \begin{bmatrix} 0.25 & 0.75 \\ 0.75 & 0.25 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{3}{4} & 0 & \frac{1}{4} \end{bmatrix}$$

La matriz $\begin{bmatrix} 0 & \frac{1}{3} \\ 1 & \frac{1}{2} \end{bmatrix}$ no es estocástica, porque para la segunda columna $\frac{1}{3} + \frac{1}{2} \neq 1$. La matriz $\begin{bmatrix} 2 & -1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ no es estocástica, porque el elemento 2 es mayor que 1. La matriz A no es doblemente estocástica, puesto que para el primer renglón $1 + \frac{1}{2} \neq 1$.

TEOREMA 22

> Si A y B son dos matrices estocásticas $n \times n$, entonces el producto AB es estocástico.

DEMOSTRACIÓN Comprobaremos el teorema sólo para $n = 2$. La idea de la demostración se generaliza con facilidad a cualquier tamaño $n \times n$. Sean

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$$

donde $a + c = 1$, $b + d = 1$, $a' + c' = 1$, $b' + d' = 1$. Por tanto,

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{bmatrix}$$

Pero $(aa' + bc') + (ca' + dc') = a'(a + c) + c'(b + d) = a' + c' = 1$. De igual manera, la segunda columna suma 1. Por consiguiente, AB es estocástica.

COROLARIO 23

> Sea A una matriz estocástica $n \times n$. Entonces, para cualquier entero positivo k, la potencia A^k es estocástica.

NOTA En el teorema y corolario anteriores puede reemplazarse *estocástica* por *doblemente estocástica*. La demostración de esta afirmación se deja como ejercicio.

Probabilidad Ahora presentaremos la noción de la *probabilidad* de la ocurrencia de un evento, y veremos cómo se relaciona con las matrices estocásticas.

Si existe la certeza de que acontecerá un evento, se dice que su probabilidad de ocurrencia es 1; si no hay certeza alguna, su probabilidad es de 0. Otros valores de probabilidades son números entre 0 y 1. Por ejemplo, se dice que la probabilidad de lluvia es 70% = 0.7 en determinado día. Mientras más alta es la probabilidad de que ocurra un evento, hay mayor certeza de que ocurra. Si este evento tiene n resultados igualmente probables de ocurrir, de los cuales nos interesa m, la probabilidad de que ocurra uno de los interesantes es m/n.

Por ejemplo, si tiramos un dado, los resultados posibles son 6 (se puede tener un 1 o un 2, o un 3, o un 4, o un 5 o un 6, cada uno con la misma posibilidad). La probabilidad de obtener 2 (uno de los resultados) es 1 de 6, o $\frac{1}{6}$. La probabilidad de que salga un número par es 3, porque el dado tiene un 2, un 4 y un 6, es decir, $\frac{3}{6} = \frac{1}{2}$.

Como los elementos de una matriz estocástica son números entre 0 y 1, éstos pueden considerarse como probabilidades de resultados de eventos.

Veamos el estudio siguiente acerca de los hábitos de fumar en un grupo de personas. Supongamos que la probabilidad de que un fumador continúe fumando un año después es de 65% (por lo que hay un 35% de probabilidades de que deje de hacerlo), mientras que la probabilidad de que un no fumador continúe sin este hábito es de 85% (por lo que hay 15% de probabilidades de que se convierta en fumador). Esta información se puede tabular con la siguiente matriz estocástica de probabilidades, que se llama **matriz de probabilidades de transición** (o de **transición de probabilidades).**

$$\begin{array}{ll} & \overbrace{\begin{array}{cc}\text{Fumador} & \text{No fumador}\end{array}}^{\text{Estado inicial}} \\ \text{Estado final} \begin{cases}\text{Fumador} \\ \text{No fumador}\end{cases} & \begin{bmatrix} 0.65 & \qquad 0.15 \\ 0.35 & \qquad 0.85 \end{bmatrix} \end{array}$$

■ EJEMPLO 44 Supongamos que en 1960, cuando comenzó el estudio, 70% de los miembros del grupo eran fumadores y 30% eran no fumadores. Si la matriz anterior de probabilidades de transición es válida durante los 10 años siguientes, ¿cuáles serán los porcentajes de fumadores y no fumadores en 1961? ¿En 1962? ¿En 1964?

SOLUCIÓN En 1961, el porcentaje de fumadores está formado por quienes lo eran en 1960, es decir, 70% · 65% = 0.455; a este valor se suman quienes adquirieron el hábito después de 1960, es decir, 30% · 15% = 0.045, un total de 0.5 = 50%. De igual manera, el porcentaje de los no fumadores en 1961 es $0.7 \cdot 0.35 + 0.3 \cdot 0.85 = 0.5 = 50\%$. Ambos números pueden calcularsse como el producto de las matrices

$$\begin{bmatrix} 0.65 & 0.15 \\ .035 & 0.85 \end{bmatrix} \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

Para 1962 se usa la misma matriz estocástica, comenzando con el nuevo vector (0.5, 0.5):

$$\begin{bmatrix} 0.65 & 0.15 \\ 0.35 & 0.85 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.65 & 0.15 \\ 0.35 & 0.85 \end{bmatrix}^2 \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

Para 1964,

$$\begin{bmatrix} 0.65 & 0.15 \\ 0.35 & 0.85 \end{bmatrix}^4 \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.325 \\ 0.675 \end{bmatrix}$$

Por consiguiente, en 1964 hay 32.5% de fumadores y 67.5% de no fumadores. En general, en k años a partir de 1960, los porcentajes pueden determinarse como el vector inicial (0.7, 0.3) multiplicado por la k-ésima potencia de la matriz de probabilidades de transición:

$$\begin{bmatrix} 0.65 & 0.15 \\ 0.35 & 0.85 \end{bmatrix}^k \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$$

En este caso, $k \leq 10$. $\square$

Al sustituir k cada vez mayores, es posible observar que aun cuando se incrementan los porcentajes de los no fumadores, nunca rebasan 0.7 = 70%, aunque k adquiera valores muy grandes. Se dice que este proceso *tiende* al vector final (0.3, 0.7). Por consiguiente, a largo plazo, el 30% de las personas serán fumadores, y el 70% no fumadores, siempre y cuando se apliquen las hipótesis actuales acerca de las probabilidades de transición.

El proceso que acabamos de describir es un ejemplo de **proceso de Markov**, o de **cadena de Markov**. En una cadena de Markov[10] el siguiente estado de un sistema sólo depende de su estado actual. En nuestro caso, los porcentajes de fumadores y no fumadores sólo dependen de los porcentajes del año anterior. Cada vez el vector actual de porcentajes se multiplicó por la matriz de probabilidades de transición, que es fija.

Para citar otro ejemplo, veamos el clásico del fútbol estudiantil más conocido en Estados Unidos: el juego anual entre el Ejército y la Marina (*Army-Navy*). Suponga que la probabilidad de que el Ejército gane un año y que la Marina gane en el siguiente es de 70%. Por tanto, la probabilidad de que el Ejército obtenga la victoria en dos años consecutivos es 30%. Considere también la que la probabilidad de que la Marina resulte triunfadora un año y el Ejército lo sea en el siguiente es de 30%. Por consiguiente, la probabilidad de que la Marina gane dos años seguidos es de 70%. Esta situación puede expresarse con la siguiente matriz *doblemente estocástica* de probabilidades de transición:

$$\begin{array}{lll} & & \overbrace{\text{Gana la Marina} \quad \text{Gana el Ejército}}^{\text{Este año}} \\ \text{Año próximo} & \left\{ \begin{array}{l} \text{Gana la Marina} \\ \text{Gana el Ejército} \end{array} \right. & \begin{bmatrix} 0.7 & \qquad\qquad 0.3 \\ 0.3 & \qquad\qquad 0.7 \end{bmatrix} \end{array}$$

■ EJEMPLO 45 (Juegos Ejército-Marina) Como la Marina triunfó este año, ¿cuál es la probabilidad de que gane dentro de 2 años?

SOLUCIÓN Como la Marina ganó este año, su probabilidad de ganar es 1, mientras que la probabilidad de que triunfe el Ejército es 0. Por consiguiente, el vector del estado inicial de probabilidades es (1, 0). Según nuestro análisis del último ejemplo, las probabilidades de que [Marina Ejército] gane el año próximo se expresan con el vector

[10] **Andrei Andreevitch Markov** (o Markoff) (1856-1922) nació en San Petersburgo (Leningrado), Rusia. Fue alumno de Chebyshev, cuyos trabajos coeditó. Llevó a cabo investigaciones importantes en análisis matemático y fue profesor de matemáticas en la Universidad de San Petersburgo. Se le conoce más por sus contribuciones a los fundamentos de teoría de la probabilidad, en particular el estudio de las propiedades de procesos que ahora se conocen como cadenas de Markov.

$$\begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$$

que comprueba nuestras hipótesis. Para dentro de 2 años, las probabilidades son

$$\begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.58 & 0.42 \\ 0.42 & 0.58 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.58 \\ 0.42 \end{bmatrix}$$

Por consiguiente, la probabilidad de que gane la Marina es 50%. En general, si la Marina obtiene la victoria este año, su probabilidad de ganar de nuevo dentro de k años se expresa con el primer componente del vector $\begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Esto también puede escribirse como sigue:

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Cuando k es grande, la inspección de este número indica que tiende a 0.5 = 50%.

Economía: modelos de entrada y salida de Leontief

Describiremos ahora una economía formada por n industrias, cada una produciendo sólo un artículo necesario para las demás, y posiblemente para sí misma. Por ejemplo, supongamos que se tuviera carbón, acero y automóviles, y que su interrelación estuviera descrita por una matriz de 3 × 3 como sigue: Si c_{ij} es el valor en dólares del i-ésimo artículo necesario para producir la cantidad de j-ésimo producto valuada en 1 dólar, digamos que se necesitan 0.30 dólares de carbón para producir 1 dólar de acero. El valor 0.30 es el elemento (1, 2) de la matriz:

$$\begin{array}{l} \\ \text{Carbón} \\ \text{Acero} \\ \text{Automóvil} \end{array} \begin{array}{c} \begin{array}{ccc} \textbf{Carbón} & \textbf{Acero} & \textbf{Automóvil} \end{array} \\ \begin{bmatrix} 0.10 & 0.30 & 0.25 \\ 0.25 & 0.20 & 0.45 \\ 0.05 & 0.15 & 0.10 \end{bmatrix} \end{array}$$

Según esta matriz, también se necesitan 0.45 dólares de acero para producir la parte de un automóvil valuada en 1 dólar. Observe que el automóvil es el mayor consumidor de acero, y el acero es el mayor consumidor de carbón. La industria del acero depende más de la automovilística para sobrevivir.

Éste es un ejemplo de matriz de **entrada y salida**, o **de consumo**, que describe la interdependencia de los sectores económicos. Los elementos de esa matriz son positivos y menores que 1. Además, la suma de los elementos de cada columna debe ser menor que 1 si cada sector debe producir más de lo que consume. Las matrices de consumo fueron introducidas y estudiadas por Wassily W. Leontief,[11] economista de Harvard, en la década de los treinta.

Supongamos que n sectores económicos se interrelacionan en una forma descrita por una matriz de consumo $C = [c_{ij}]$. Sea x_i la cantidad total de producción necesaria para que el i-ésimo sector satisfaga las demandas de todos los demás. Entonces, $c_{ij}x_j$ es la cantidad nece-

[11] En 1973, Leontief recibió el Premio Nobel en Economía por este trabajo.

saria del artículo i para producir x_j unidades del artículo j. Como la producción total de la industria i es igual a la suma de las demandas de todos los sectores,

$$\begin{aligned} x_1 &= c_{11}x_1 + \cdots + c_{1n}x_n \\ &\vdots \\ x_n &= c_{n1}x_1 + \cdots + c_{nn}x_n \end{aligned}$$

Si $\mathbf{x}$ es el vector cuyos componentes son $x_1, \ldots, x_n$, estas relaciones pueden expresarse con la ecuación matricial

$$\mathbf{x} = C\mathbf{x}$$

Hasta ahora sólo hemos considerado demanda de artículos por parte de los sectores económicos productivos. Este caso se llama **modelo cerrado de Leontief**. En realidad, también hay demanda por parte de sectores no productivos, como los consumidores, los gobiernos, etc. Por ejemplo, el gobierno puede tener demanda carbón, acero y automóviles. Todos los sectores no productivos forman el **sector abierto**. Suponga que d_i es la demanda del sector abierto, hacia el i-ésimo sector productivo. Entonces, $x_i = c_{i1}x_1 + \cdots + c_{in}x_n + d_i$. Si $\mathbf{d}$ es el vector con componentes $d_1, \ldots, d_n$ no negativos, entonces

$$\mathbf{x} = C\mathbf{x} + \mathbf{d}$$

Esta ecuación matricial describe un **modelo abierto de Leontief**, y toma en cuenta el sector abierto; $\mathbf{x}$ es el **vector salida** y $\mathbf{d}$ es el **vector demanda**. Por lo general, a los economistas les interesa calcular el vector de salida $\mathbf{x}$ dado el vector demanda $\mathbf{d}$. Esto se puede hacer despejando a $\mathbf{x}$ como sigue:

$$\mathbf{x} = C\mathbf{x} + \mathbf{d} \Rightarrow (I - C)\mathbf{x} = \mathbf{d} \Rightarrow \mathbf{x} = (I - C)^{-1}\mathbf{d}$$

siempre que la matriz $I - C$ sea invertible. Si además $(I - C)^{-1}$ tiene elementos no negativos, los elementos de $\mathbf{x}$ son no negativos, y por consiguiente se pueden aceptar como valores de producción. En general, una matriz C se llama **productiva** si existe $(I - C)^{-1}$ y tiene elementos no negativos.

■ EJEMPLO 46 Sea C la matriz de consumo y $\mathbf{d}$ el vector demanda, en millones de dólares, para una economía de sector abierto con tres industrias interdependientes. Calcule la producción demandada por las industrias y por el sector abierto.

$$C = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 10 \\ 20 \\ 30 \end{bmatrix}$$

SOLUCIÓN

$$I - C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & 0 \\ 0 & -\frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

Por consiguiente,

$$\mathbf{x} = (I - C)^{-1}\mathbf{d} = \begin{bmatrix} \frac{9}{4} & \frac{1}{2} & \frac{3}{4} \\ \frac{3}{4} & \frac{3}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 10 \\ 20 \\ 30 \end{bmatrix} = \begin{bmatrix} 55 \\ 45 \\ 70 \end{bmatrix}$$

Llegamos a la conclusión de que los niveles productivos de las tres industrias deberían ser \$55, \$45 y \$70 millones para satisfacer las demandas. □

Con frecuencia, los analistas suponen que los niveles de producción se conocen, y desean calcular la demanda que puede esperarse de los sectores productivos. En esos casos, $\mathbf{x}$ es un dato y $\mathbf{d}$ es incógnita; entonces, $\mathbf{d}$ se calcula con mucha facilidad como

$$\mathbf{d} = \mathbf{x} - C\mathbf{x}$$

Las matrices de entrada y salida se usan para analizar la economía de un país, o hasta de toda una región geográfica. Los sectores productivos son por lo general ciertas industrias clave, como la agrícola, la del acero, la química, la del carbón, la ganadera, etc. Para la matriz nacional de entrada y salida en Estados Unidos, los sectores abiertos son los gobiernos federal, estatal y local.

Ejercicios 3.5

Teoría de las gráficas

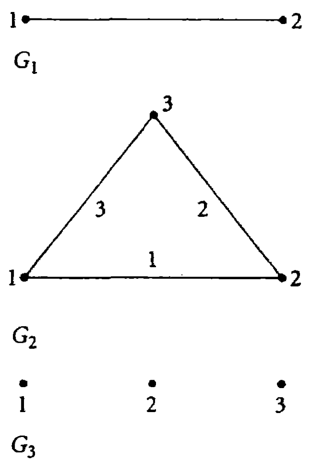


Figura 3.6

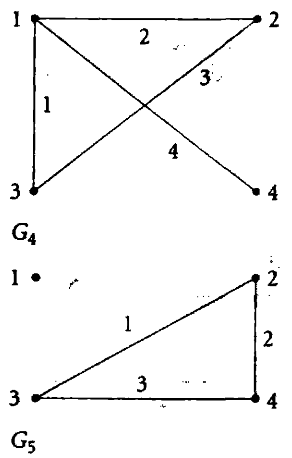


Figura 3.7

1. Determine las matrices M_1, M_2 y M_3 de las gráficas G_1, G_2 y G_3, respectivamente (figura 3.6).
2. Calcule las matrices M_4, M_5 y M_6 de las gráficas G_4, G_5 y G_6, respectivamente (figuras 3.7 y 3.8).
3. Determine las matrices de adyacencia para las gráficas G_1, G_2 y G_3.
4. Obtenga las matrices de adyacencia para las gráficas G_4, G_5 y G_6.
5. Calcule $A(G_7)^2$. Cuente la cantidad de caminatas de longitud 2 de (a) 1 a 1, (b) 1 a 2 y (c) 1 a 3. ¿Concuerdan estas respuestas con los (1, 1), (1, 2) y (1, 3)-ésimo elementos de $A(G_7)^2$, figura 3.8?

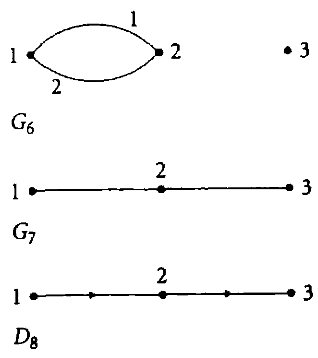


Figura 3.8

6. Conteste el ejercicio 5 para la diagráfica D_8 (figura 3.8).
7. Determine las matrices de incidencia para las gráficas G_1, G_2 y G_3.
8. Calcule las matrices de incidencia para las gráficas G_4, G_5 y G_6.
9. Trace las gráficas de línea de G_1, G_2 y G_3.
10. Trace las gráficas de línea de G_4, G_5 y G_6.
11. Sea $L(G_7)$ la gráfica de línea de G_7. Compruebe la relación $A(L(G_7)) = I(G_7)^T I(G_7) - 2I_2$.

Procesos de Markov

En los ejercicios 12 a 15 investigue si las matrices son estocásticas y doblemente estocásticas.

12. $A = \begin{bmatrix} \frac{1}{3} & 0 \\ \frac{2}{3} & 1 \end{bmatrix}, B = \begin{bmatrix} -0.1 & 1.1 \\ 1.1 & -0.1 \end{bmatrix}$

13. $C = \begin{bmatrix} \frac{1}{3} & 0 \\ \frac{2}{3} & 1 \end{bmatrix}, D = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{3} & 1 \\ \frac{1}{6} & 0 \end{bmatrix}$

14. $E = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, F = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{2} \end{bmatrix}$

15. $G = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix}, H = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{bmatrix}$

16. ¿Es estocástica $\begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{2} \end{bmatrix}^{100}$?

17. ¿Es doblemente estocástica $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}^{100}$?

18. Vea el ejemplo 45 del texto.
    a. Como el Ejército fue el ganador este año, ¿cuál es la probabilidad de que también lo sea dentro de 2 años?
    b. Dado que la Marina ganó este año, ¿cuál es la probabilidad de que triunfe dentro de 3 años?

19. Determine, si es posible, x y y tales que las matrices siguientes sean doblemente estocásticas.

$$A = \begin{bmatrix} x & 0.2 \\ 0.2 & y \end{bmatrix}, \qquad B = \begin{bmatrix} \frac{1}{4} & x \\ \frac{3}{4} & y \end{bmatrix}$$

20. Determine, si es posible, x, y y z tales que la matriz siguiente sea estocástica.

$$A = \begin{bmatrix} x & 0 & 1 \\ \frac{1}{4} & y & 0 \\ \frac{1}{4} & \frac{1}{4} & z \end{bmatrix}$$

21. Sea C una matriz de consumo:

$$C = \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.6 \end{bmatrix}$$

    a. ¿Es productiva C?
    b. En caso afirmativo, sea D un vector demanda:

$$D = \begin{bmatrix} 10 \\ 20 \end{bmatrix}$$

Determine el vector salida de producción.

3.6 Miniproyectos

1 ■ Códigos

Con frecuencia los gobiernos, las agencias nacionales de seguridad y las empresas se interesan en la transmisión de mensajes codificados que sean difíciles de descifrar por otros, en caso de ser interceptados, pero que se *decodifiquen* con facilidad por quienes lo reciben. Hay muchas formas interesantes de cifrar o *codificar* un mensaje, y en su mayor parte usan la teoría de los números o el álgebra lineal. Describiremos uno que es eficaz, en especial cuando se usa una matriz invertible de gran tamaño.

Comenzaremos con una matriz M invertible, que sólo la conocen quienes la transmiten y quienes la reciben. Por ejemplo,

$$M = \begin{bmatrix} -3 & 4 \\ -1 & 2 \end{bmatrix}$$

Supongamos que se desea codificar el mensaje

ATTACK NOW

Reemplazamos cada letra con el número que le corresponde a su posición en el alfabeto. Un espacio se representa por 0.

$$\begin{array}{cccccccccc} \text{A} & \text{T} & \text{T} & \text{A} & \text{C} & \text{K} & & \text{N} & \text{O} & \text{W} \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 1 & 20 & 20 & 1 & 3 & 11 & 0 & 14 & 15 & 23 \end{array}$$

El mensaje se ha convertido a la sucesión de números 1, 20, 20, 1, 3, 11, 0, 14, 15, 23, que agrupamos como una sucesión de vectores columna,

$$\begin{bmatrix} 1 \\ 20 \end{bmatrix} \begin{bmatrix} 20 \\ 1 \end{bmatrix} \begin{bmatrix} 3 \\ 11 \end{bmatrix} \begin{bmatrix} 0 \\ 14 \end{bmatrix} \begin{bmatrix} 15 \\ 23 \end{bmatrix}$$

y multiplicamos por la izquierda a M:

$$M\begin{bmatrix} 1 \\ 20 \end{bmatrix} = \begin{bmatrix} 77 \\ 39 \end{bmatrix}, \quad M\begin{bmatrix} 20 \\ 1 \end{bmatrix} = \begin{bmatrix} -56 \\ -18 \end{bmatrix}, \quad M\begin{bmatrix} 3 \\ 11 \end{bmatrix} = \begin{bmatrix} 35 \\ 19 \end{bmatrix}$$

$$M\begin{bmatrix} 0 \\ 14 \end{bmatrix} = \begin{bmatrix} 56 \\ 28 \end{bmatrix}, \quad M\begin{bmatrix} 15 \\ 23 \end{bmatrix} = \begin{bmatrix} 47 \\ 31 \end{bmatrix}$$

con lo que se obtiene la sucesión de números 77, 39, −56, −18, 35, 19, 56, 28, 47, 31. Éste es el mensaje cifrado. Para decodificarlo, quien lo recibe necesita calcular M^{-1},

$$M^{-1} = \begin{bmatrix} -1 & 2 \\ -\frac{1}{2} & \frac{3}{2} \end{bmatrix}$$

y multiplicarlo por los vectores $\begin{bmatrix} 77 \\ 39 \end{bmatrix}, \begin{bmatrix} -56 \\ -18 \end{bmatrix}, \begin{bmatrix} 35 \\ 19 \end{bmatrix}, \begin{bmatrix} 56 \\ 28 \end{bmatrix}, \begin{bmatrix} 47 \\ 31 \end{bmatrix}$ para obtener los números originales.

$$M^{-1}\begin{bmatrix} 77 \\ 39 \end{bmatrix} = \begin{bmatrix} 1 \\ 20 \end{bmatrix}, \quad M^{-1}\begin{bmatrix} -56 \\ -18 \end{bmatrix} = \begin{bmatrix} 20 \\ 1 \end{bmatrix}, \ldots$$

Problema A (descifrar un mensaje)

Basado en el método anterior, decodifique el mensaje expresado por los números 17, 15, 29, 15, 17, 29, 16, 31, 47, 6, 19, 20, 35, 24, 39, 14, 19, 19, si

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Problema B (Rompiendo un código)

Suponga que interceptó el siguiente mensaje cifrado sobre el mercado accionario: 1 156, −203, 624, −84, −228, 95, 1 100, −165, 60, 19. Sus fuentes le informan que fue codificado con una matriz simétrica de 2 × 2. Su intuición le dice que es muy probable que la primera palabra del mensaje sea *sell* (venda) o *buy* (compre). ¿Puede descifrar el mensaje?

2 ■ Los números de Fibonacci

A principios de la Edad Media, Fibonacci estudió la sucesión de números $f_0, f_1, f_2, \ldots$ que se produce cuando se cuentan pares macho-hembra de conejos que se reproducen mensualmente y crean otro par macho-hembra. El proceso comienza con un par inicial, de modo que $f_0 = 1$, y se supone que los conejos se vuelven reproductivos a partir de su segundo mes.

Comenzando con un par, $f_0 = 1$, al final del primer mes todavía se tiene un par, $f_1 = 1$, el cual se reproduce, y al final del segundo mes, cuando tenemos $f_2 = 2$ pares. Al final del tercer mes se reproduce el primer par, pero el par más reciente es demasiado joven para reproducirse. Así, al final del tercer mes se tienen $f_3 = f_2 + f_1 = 2 + 1 = 3$ pares. De igual manera, al final del k-ésimo mes, se tienen los pares que había al final del mes anterior, f_{k-1}, más la cantidad de descendientes de los pares que se reprodujeron hace 2 meses, f_{k-2}. Así, $f_k = f_{k-1} + f_{k-2}$. Por consiguiente, la sucesión producida puede describirse con las relaciones recursivas:

$$f_0 = 1, \quad f_1 = 1, \quad f_k = f_{k-1} + f_{k-2}, \quad k \geq 2$$

Los primeros términos de la sucesión son 1, 1, 2, 3, 5, 8, 13, 21, . . .
En este proyecto se pide lo siguiente:

1. Calcular $f_0, f_1, \ldots, f_{15}$.
2. Determinar una matriz A 2×2 que tenga la siguiente propiedad:

$$A\begin{bmatrix} f_k \\ f_{k-1} \end{bmatrix} = \begin{bmatrix} f_{k+1} \\ f_k \end{bmatrix} \text{ para } k = 1, 2, 3, \ldots$$

Por ejemplo, si

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

entonces

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad A\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \ldots$$

3. Calcular

$$A^2\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad A^3\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad A^4\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad A^5\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

4. ¿Cómo se relaciona el producto $A^k\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ con el cálculo del $(k + 1)$-ésimo número de Fibonacci?

3 ■ Probabilidades de transición

Problema A

Un grupo de personas compra automóviles, cada 4 años, con algunos de los tres fabricantes A, B y C. Las probabilidades de cambiar de un fabricante a otro (probabilidades de transición) se expresa con la siguiente matriz:

$$\begin{bmatrix} 0.5 & 0.4 & 0.6 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.1 \end{bmatrix}$$

Suponga que en el año 1995 el fabricante A vendió 100 automóviles, el B 80 y el C 40.

1. ¿Cuántos automóviles se venderían en 1999?
2. ¿Cuántos automóviles se vendieron en 1991?
3. A largo plazo, ¿alguno de los fabricantes dominará el mercado?

Problema B

Se tiene la matriz estocástica de probabilidades de transición

$$T = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix}$$

que expresa el flujo de clientes desde y hacia los mercados A y B después de una compra. Suponga que la primera vez, aproximadamente $\frac{2}{3}$ de los compradores compran en A y $\frac{1}{3}$ en B.

1. ¿Cuáles son las cuotas de mercado después de la primera compra?
2. ¿Cuáles son las cuotas de mercado después de la segunda compra?
3. ¿Hay un *equilibrio de mercado* (es decir, un vector de las cuotas (a, b) que permanece sin cambio de una compra a la siguiente)? En caso afirmativo, determínelo. (Tenga en cuenta que los clientes sólo pueden comprar en A o en B. Esto significa que $a + b = 1$.)

4 ■ Caminatas en diagráficas

La empresa Helados Regios entrega a cuatro tiendas. Las tiendas y las rutas de entrega (algunas de un solo sentido) forman una diagráfica cuya matriz de adyacencia es

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

1. Trace la diagráfica D.
2. Determine las matrices que representan la cantidad de rutas que pueden ser recorridas de una a otra tienda, de modo que un camión de entrega pase

   (a) Sólo por una tienda.
   (b) Sólo por dos tiendas.

(c) Cuando mucho por una tienda.
(d) Cuando mucho por dos tiendas.
(e) ¿De cuántas formas puede uno ir de la tienda 3 a la 4 pasando sólo por otra tienda más?

3. ¿Puede ser A la matriz de adyacencia de una gráfica (a diferencia de una diagráfica)?

5 ■ Un problema teórico

Problema

Si A, B y C son matrices $n \times n$, y r es cualquier escalar distinto de cero.[12] Si

$$A + B + rAB = 0$$

$$B + C + rBC = 0$$

$$C + A + rCA = 0$$

demuestre que $A = B = C$. (*Sugerencia:* Demuestre que cada matriz $I + rA$, $I + rB$, $I + rC$ es invertible, usando el ejercicio 27 de la sección 3.3.)

3.7 Ejercicios en computadora

Esta sección de cómputo le ayudará a familiarizarse con los comandos básicos de manipulación de matrices en su *software*. Es importante saber cómo introducir los elementos de columnas, renglones y submatrices de una matriz, y cómo crear nuevas matrices partiendo de otras. También es preciso saber cómo llevar a cabo operaciones matriciales y cómo usarlas para ilustrar las propiedades conocidas, y quizá investigar nuevas. Todo esto se hace con un repaso simultáneo del material básico. Observe que un ejercicio identificado con [S] requiere manipulación simbólica.[13]

Sean

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & -2 & -3 & -4 \\ 5 & 6 & 7 & 8 \\ -5 & -6 & -7 & -8 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & -2 & 3 \\ 3 & -3 & 3 \end{bmatrix},$$

$$B = \begin{bmatrix} 2 & 3 & 6 \\ 3 & 3 & 6 \\ 6 & 6 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 4 \\ 0 & 4 \\ 1 & 4 \end{bmatrix}$$

Primero capture esas matrices y nómbrelas como se describe. Si su programa ya usó una de las letras, cámbiela. Use su programa para resolver los siguientes ejercicios:

1. Muestre el cuarto renglón, la tercera columna y el elemento (2, 3) de M.
2. Muestre la matriz obtenida de M usando las tres primeras columnas.

---

[12] Este interesante problema, con $r = 1996$, fue tomado de los Exámenes Nacionales de Admisión a Profesional en Grecia.
[13] Omita esos ejercicios si no dispone de manipulación de símbolos.

3. Muestre la matriz que se obtiene de M empleando sólo los dos últimos renglones.
4. Muestre la parte de M que se obtiene omitiendo el primer renglón y las dos primeras columnas.
5. Muestre la matriz que se obtiene de M agregando los números 4, 3, 2, 1 como (a) un primer renglón, y (b) última columna.
6. Muestre la matriz que se obtiene poniendo juntas A y B.
7. Muestre la matriz que se obtiene poniendo A arriba de B.
8. Muestre una matriz diagonal cuyos elementos diagonales sean 1, 2, 3, 4.

Sea T la matriz obtenida invirtiendo los renglones de M. Es decir, el último renglón se transforma en el primero, el cuarto en el segundo, y así sucesivamente. Calcule lo siguiente.

9. $M - T$.
10. $15M - 35T$.
11. Despeje X de la ecuación matricial $17X - 51M = 62T$.
12. Calcule: (a) AB, (b) BA, (c) $(AB)C$, (d) $A(BC)$.
13. Calcule: (a) $(AB)^2$, (b) A^2B^2, (c) $(A^3)^4$, (d) A^{12}.
14. Calcule: (a) $(A + B)^2$, (b) $A^2 + 2AB + B^2$.
15. Calcule A^{-1} con: (a) el comando inversión de su programa; (b) reduciendo con operaciones de renglón $[A : I]$. Compare las respuestas.
16. [S] Sea $S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Calcule S^3, S^{-3} y S^3S^{-3}.
17. Verifique las identidades: (a) $(A^{-1})^{-1} = A$, (b) $(10AB)^{-1} = \frac{1}{10}B^{-1}A^{-1}$, (c) $(ABA)^{-1} = A^{-1}B^{-1}A^{-1}$, (d) $A^5A^{-2} = A^3$.
18. Verifique las identidades: (a) $(A^T)^T = A$, (b) $(A^T)^{-1} = (A^{-1})^T$, (c) $(ABC)^T = C^TB^TA^T$, (d) $(A^{-2})^T = (A^T)^{-2}$.
19. Despeje a la matriz X de 3×3, de las ecuaciones matriciales: (a) $AX = B$, (b) $XA = B$.

Considere la sucesión de matrices de tamaño mayor que o igual a 3, con ceros en la diagonal y unos en los demás elementos:

$$A_3 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}, \ldots$$

20. Escriba una función de un argumento, llamada **diagzero**, que muestre esas matrices de acuerdo con el tamaño. Por ejemplo, **diagzero**, (3) es A_3, y así sucesivamente.
21. Use **diagzero**, para mostrar A_3, A_4 y A_5, y calcule A_3^{-1}, A_4^{-1} y A_5^{-1}.
22. Trate de determinar la fórmula de A_n^{-1}.
23. Escriba y haga la prueba del código de tres funciones que produzca matrices elementales de determinado tamaño que se obtienen a partir de las operaciones elementales de renglón.
24. Suponga que una gráfica tiene la matriz de adyacencia A_4, como la que aparece arriba. Determine la matriz que produzca: (a) la cantidad de caminatas de longitud 4; (b) la cantidad de caminatas de longitudes 1 o 2, o 3 o 4.
25. Escriba el código de una función, llamada **sumpower**, que toma dos argumentos, una matriz cuadrada A y un entero positivo n. El valor de la función es la matriz

$$A + A^2 + \cdots + A^n$$

Aplique sumpower con $A = A(G)$ y $n = 4$ para comprobar su respuesta en la segunda parte del ejercicio anterior.

Operaciones con matrices en Maple, Mathematica y MATLAB

Para matrices A y B de tamaños compatibles:

función	Maple	Mathematica	MATLAB
$A + B$	evalm($A + B$);	$A + B$	$A + B$
$2A - 3B$	evalm(2 * A – 3 * B);	$2A - 3B$	2 * A – 3 * B
AB	evalm(A& * B);	$A.B$	A * B
Rango $m, m + 1, \ldots, n$	m..n	Range[m,n]	m:n
A^T	transpose(A);	Transpose[A]	A.′ ver también A′

Soluciones seleccionadas con Maple

```
with(linalg);                                      # Carga del paquete linalg.
M := matrix(4, 4, [1, 2, 3, 4, -1, -2, -3, -4, 5, 6, 7, 8, -5, -6, -7, -8] )   # Definiciones de matrices.
A := matrix(3, 3, [1, -2, 3, 2, -2, 3, 3, -3, 3] );
B := matrix(3, 3, [2, 3, 6, 3, 3, 6, 6, 6, 6] );
C := matrix(3, 2, [2, 4, 0, 4, 1, 4])
A := matrix(3, 3, [ [1, -2, 3], [2, -2, 3], [3, -3, 3] ] );     # Forma alternativa.
A := matrix( [ [1, -2, 3], [2, -2, 3], [3, -3, 3] ] );          # Otra forma alternativa.

# Ejercicios 1 a 5.

row(M, 4);                       # Cuarto renglón.
submatrix(M, 4..4, 1..4);        # De nuevo cuarto renglón.
submatrix(M, 1..4, 3..3);        # Tercera columna como matriz columna.
col(M, 3);                       # Tercera columna como vector.
matrix(4, 1, col(M, 3));         # Otra manera: tercera columna como matriz columna.
M[2, 3];                         # Elemento (2, 3).
submatrix(M, 1..4, 1..3);        # Tres primeras columnas como matriz.
col(M, 1..3);                    # Tres primeras columnas como sucesión de vectores.
delcols(M, 4..4);                # Otra forma: tres primeras columnas por eliminación.
submatrix(M, 3..4, 1..4);        # Dos últimos renglones como matriz.
row(M, 3..4);                    # Dos últimos renglones como sucesión de vectores.
delrows(M, 1..2);                # Otra forma: dos últimos renglones por eliminación.
submatrix(M, 2..4, 3..4);        # Eliminación de primer renglón y dos primeras columnas.
delcols(delrows(M, 1..1), 1..2); # Otra forma.
v:=vector([4, 3, 2, 1]);         # Definir el vector (4, 3, 2, 1).
stack(v, M);                     # Agregar v a M como primer renglón.
augment (M, v);                  # Agregar v a M como última columna.
concat (M, v);                   # concat es lo mismo que augment.
# Otro comando útil para agrandar una matriz es extend.

# Ejercicios 6 y 7.

concat (A, B);                   # A y B juntos.
```

```
stack(A, B);                     # A sobre B.

# Ejercicio 8.

diag(1, 2, 3, 4);

# Ejercicios 9 a 11.

T:=matrix(4, 4, (i, j)->M[5-i, j] );   # Los elementos de renglón se intercambian,
                                        # dejando intactos los elementos de columna.
evalm(M-T);
evalm(15*M-35*T);
evalm(1/17*(51*M+62*T));

# Ejercicios 12 a 14.

evalm(A&*B);                # La multiplicación de matrices se indica con &*.
evalm(B&*B);
evalm((A&*B)&*C);
evalm(A&*(B&*C));
evalm(A&*B)^2);             # La elevación a potencias de matrices se indica con ^.
evalm((A&*B)&*(A&*B));      # Igual que arriba.
evalm(A^2&*B^2);
evalm((A^3)^4);
evalm(A^12);
evalm((A+B)^2);
evalm(A^2+2*A&*B+B^2);

# Ejercicio 15.

inverse(A);                                    # Calcula A ^ (-1).
evalm(A^(-1));                                 # Otra forma.
evalm(1/A);                                    # Otra forma.
augment(A, [1, 0, 0], [0, 1, 0], [0, 0, 1];    # [A : I].
rref(");                                       # La mitad derecha es A ^ (-1).

# Ejercicio 16.

S := matrix(2, 2, [a, b, c, d]);    # Matriz simbólica S.
evalm(S^3)                          # S ^ 3.
s:=map(simplify, ");                # Se simplifica. Se necesita "transformar"
evalm(S^(-3)):                      # el comando simplify a través de los
ss :=map(simplify, ");              # elementos de la matriz
evalm(s &* ss)                      # resultante.
map(simplify, ");

# Ejercicios 17 y 18.

inverse(inverse(A));
inverse(10*A&*B);
evalm((1/10)*inverse(B)&*inverse(A));
inverse(A&*B&*A);
evalm(inverse(A) &* inverse(B) &* inverse(A));
evalm(A^3);
evalm(A^5 &* A^(-2));
transpose(transpose(A));
transpose(inverse(A));
```

```
inverse(transpose(A));
transpose(A&*B&*C);
evalm(transpose(C) &* transpose (B) &* transpose(A));
transpose(A^(-2));
evalm(transpose(A)^(-2));

# Ejercicio 19.

evalm(inverse(A) &* B);
evalm(B &* inverse(A));

# Ejercicios 20 a 22.

diagzero := proc(n) local i; evalm(matrix(n, n, [seq(1, i=1..n^2)]) - &*()); end:
# Matriz de unos menos I_n.
# También es útil para conocer cómo usar códigos más completos como:
diagzero := proc(n) local i, j, a;
            a := array(1..n, 1..n):
            for i from 1 to n do
              for j from 1 to n do
                if i=j then a[i, j] :=0 else a[i, j] :=1 fi od: od:
            evalm(a);
            end:
diagzero(3);
diagzero(4);
diagzero(5);
inverse(diagzero(3));
inverse(diagzero(4));
inverse(diagzero(5));
# A_n^(-1) tiene -(n-2)/(n-1) en la diagonal principal y 1/(n-1) en las demás posiciones.

# Ejercicio 24.

AA := diagzero (4);     # A_4 Usando diagzero, o teclearlo directamente.
evalm(AA^4);            # Matriz que produce la cantidad de caminatas de longitud 4.
evalm(AA+AA^2+AA^3+AA^4);

# Ejercicio 25.

sumpower := proc(A,n) evalm(sum('A^i', 'i'=1. .n)) end:
sumpower (AA,4);
```

Soluciones seleccionadas con Mathematica

```
M = {{1, 2, 3, 4}, {-1, -2, -3, -4}, {5, 6, 7, 8}, {-5, -6, -7, -8}}   (* Definiciones *)
A = {{1, -2, 3}, {2, -2, 3}, {3, -3, 3}}                                 (* de matrices. *)
B = {{2, 3, 6}, {3, 3, 6}, {6, 6, 6}}
C1 = {{2, 4}, {0, 4}, {1, 4}}   (* C se reserva para constantes diferenciales. *)
MatrixForm[A]                   (* Muestra A como una matriz. *)

(* Ejercicios 1 al 5. *)

(* Primer método. Usando LinearAlgebra `MatrixManipulation`. *)
<<LinearAlgebra `MatrixManipulation`   (* Carga del paquete. *)
```

```
TakeRows[M, {4}]                    (* El cuarto renglón. *)
TakeColumns[M, {3}]                 (* La tercera columna. *)
TakeColums[M, {1, 3}]               (* Las tres primeras columnas. *)
TakeRows[M, {3, 4}]                 (* Los dos últimos renglones.  *)
TakeMatrix[M, {2, 3}, {4, 4}]       (* Eliminación del primer renglón y las dos primeras columnas. *)
AppendColumns[{v}, M]               (* Agrega v a M como primer RENGLÓN.*)
AppendRows[M, {{4}, {2}, {3}, {1}}] (* Agrega v a M como última COLUMNA. *)
(* Segundo método. Con manipulaciones de listas. Más útil en la corrida larga. *)
M[[4]]  (* El cuarto renglón. En realidad, el cuarto elemento de la lista M. *)
M[[Range[4, 4], Range[1, 4]]]   (* Otra forma usando rangos de columna y de renglón. *)
(* Precaución: las dos respuestas {13,14,15,16} y {{13,14,15,16}} en realidad *)
(* son diferentes. La primera lista tiene 4 elementos, y la segunda sólo uno. *)
M[[Range[1, 4], Range[3, 3]]]     (* Tercera columna usando rangos de columna y de renglón. *)
Map[#[[3]]&, M]                   (* Otra forma: la tercera parte de cada renglón. *)
(* De nuevo, los números son iguales, pero las listas no. *)
#[[3]]& /@ M                      (* Igual que lo último, pero con notación distinta. *)
Transpose[M] [[3]]                (* Otra forma. *)
M[[2, 3]]                         (* Elemento (2, 3). (El tercer elemento del segundo renglón.) *)
M[[2]] [[3]]                      (* Igual que el anterior, en distinta notación. *)
M[[Range[1, 4], Range[1, 3]]]     (* Primeras tres columnas. *)
Map[#[[Range[1, 3]]]&, M]         (* Otra forma.  *)
M[[Range[3, 4], Range[1, 4]]]     (* Últimos dos renglones. *)
M[[Range[3, 4]]]                  (*  Igual que arriba. *)
M[[Range[2, 4], Range[3, 4]]]     (* Eliminación del primer renglón y dos primeras columnas. *)
v={4, 3, 2, 1}                    (Definir el vector (4,3,2,1). *)
Prepend[M, v]                     (* Agregar v a M como primer renglón. *)
Join[{v}, M]                      (* Otra forma. *)
(* Append[M,v] agregaría v a M como último renglón. *)
Transpose[Append[Transpose[M], v]]    (* Agrega v a M como última columna. *)
(* Esto se hizo indirectamente. Primero cambiar columnas a renglones y agregar v como *)
(* último renglón, y después invertir el cambio. *)
```

```
(* Ejercicios 6 y 7. *)

<<LinearAlgebra 'MatrixManipulation'
AppendRows[A, B]          (* A y B juntas. *)
AppendColumns[A, B]       (* A sobre B. *)

(* Ejercicio 8. *)

DiagonalMatrix[{1, 2, 3, 4}]

(* Ejercicios 9 a 11. *)

T=Reverse[M]              (* Invierte el orden de los elementos (= los renglones) de la lista. *)
M-T
15M-35T                   (* Se puede usar 15M en lugar de 15 M, porque el primer factor   *)
                          (* es un número y el segundo un símbolo.   *)
1/17 (51M+62T)

(* Ejercicios 12 a 14. *)

A.B                    (* La multiplicación de matrices se representa por un punto. *)
B.A
```

```
(A.B).C1
A.(B.C1)
MatrixPower[A.B, 2]              (* A^n se indica con MatrixPower[A,n]. *)
MatrixPower[A, 2].MatrixPower[B, 2]
(A.A).(B.B)                                              (* Igual que arriba.*)
(* Cuidado: al teclear A^2 se obtendrán los cuadrados de los elementos de la lista de A, y no de la matriz
A^2. *)
MatrixPower[MatrixPower[A, 3], 4]
MatrixPower[A, 12]
(A+B).(A+B)
A.A+2A.B+B.B

(* Ejercicio 15. *)

Inverse[A]                              (* Calcula A(-1). *)
<<LinearAlgebra `MatrixManipulation`;   (* Se necesita AppendRows del paquete. *)
AppendRows[A, {{1, 0, 0}, {0, 1, 0}, {0, 0, 1}}]
RowReduce[%]                            (* La mitad derecha es A^(-1). *)

(* Ejercicio 16. *)

S = {a, b}, {c, d}}                     (* Matriz simbólica. *)
MatrixPower[S, 3]                       (* S^3. *)
s = Simplify[%]                         (* Necesita simplificarse. *)
MatrixPower[S, -3]                      (* S^(-3), *)
ss = Simplify[%]                        (* simplificado. *)
Simplify[s.ss]                          (* El producto es la identidad. *)

(* Ejercicios 17 y 18. *)

Inverse[Inverse[A]]
Inverse[10 A . B]
1/10 Inverse[B] . Inverse[A]
Inverse[A . B . A]
Inverse[A] . Inverse[B] . Inverse[A]
MatrixPower[A, 3]
MatrixPower[A, 5] . MatrixPower[A, -2]
Transpose[Transpose[A]]
Transpose[Inverse[A]]
Inverse[Transpose[A]]
Transpose[A.B.C1]
Transpose[C1].Transpose[B].Transpose[A]
Transpose[MatrixPower[A, -2]]
MatrixPower[Transpose[A], -2]

(* Ejercicio 19. *)

Inverse[A] . B
A . Inverse[B]

(* Ejercicios 20 a 22. *)

diagzero[n_] := Table [If[i==j, 0, 1], {i, 1, n}, {j, 1, n}]
(* La tabla de 2 dimensiones = matriz con elementos 0 en la    *)
(* diagonal y 1 los demás. *)
(* También se puede usar la matriz de nxn de unos menos I_n.  *)
```

```
diagzero[n_] := Table[1, {n}, {n}] - IdentityMatrix[n]
diagzero[3]
diagzero[4]
diagzero[5]
Inverse[diagzero[3]]
Inverse[diagzero[4]]
Inverse[diagzero[5]]
(* A_n^(-1) tiene -(n-2)/(n-1) en la diagonal principal y 1/(n-1) fuera de ella. *)

(* Ejercicio 24. *)

AA = diagzero[4]        (* Usar diagzero o teclear directamente la matriz. *)
MatrixPower[AA, 4]      (* Matriz que obtiene la cantidad de caminatas de longitud 4. *)
AA+MatrixPower[AA, 2]+MatrixPower[AA, 3]+MatrixPower[AA, 4]

(* Ejercicio 25. *)

sumpower[A_, n_] := Sum[MatrixPower[A,i], {i,1,n}]
sumpower[AA,4]
```

Soluciones seleccionadas con MATLAB

```
M = [1 2 3 4; -1 -2 -3 -4; 5 6 7 8; -5 -6 -7 -8]  % Definiciones
A = [1 -2 3; 2 -2 3; 3 -3 3]                      % de matrices.
B = [2 3 6; 3 3 6; 6 6 6]
C = [2 4; 0 4; 1 4]

% Ejercicios 1 a 5.

M(4, :)             % Cuarto renglón.
M(:, 3)             % Tercera columna.
M(2, 3)             % Elemento (2, 3).
M(:, 1:3)           % Tres primeras columnas.
M(3:4, :)           % Dos últimos renglones.
M(2:4, 3:4)         % Eliminación del primer renglón, dos primeras columnas.
v = [4 3 2 1]       % Definir el vector (4, 3, 2, 1).
[v ; M]             % Agrega v a M como primer renglón.
[M v.']             % Agrega v a M como última columna. v.' es v en forma de columna.
[M [4; 3; 2; 1]]    % Lo mismo, pero tecleando v directamente como vector columna.

% Ejercicios 6 y 7.

[A, B]                  % A y B juntas.
[A; B]                  % A sobre B.

% Ejercicio 8.

diag([1, 2, 3, 4])

% Ejercicios 9 a 11.

T=flipud(M)         % Flips Voltea a M de cabeza. fliplr flips la voltea de izquierda a derecha.
M=T
15*M-35*T
1/17*(51*M+62*T)
```

```
% Ejercicios 12 a 14.

A*B                      % La multiplicación de matrices se indica con *.
B*A
(A*B)*C
A*(B*C)
(A*B)^2                  % Las potencias de matrices se indican con ^.
(A*B)*(A*B)              % Igual que arriba.
A^2*B^2
(A^3)^4
A^12
(A+B) ^2
A^2+2*A*B+B^2

% Ejercicio 15.

inv(A)                             % Calcula A^(-1).
A^(-1)                             % Otra forma.
[A, [1 0, 0; 0 1 0; 0 0 1]]        % [A:I]
rref(ans)    % La mitad derecha es A^(-1).

% Ejercicio 16.  (ST) Requiere symbolic toolbox.

S = sym('[a b; c d]')              % Matriz simbólica.
s=sympow(S, 3)                     % S^3.
ss=sympow(S, -3)                   % S^(-1).
symmul(s, ss)                      % El producto es la matriz identidad.

% Ejercicios 17 y 18.

inv(inv(A))
inv(10*A*B)
(1/10)*inv(B)*inv(A)
inv(A*B*A)
inv(A) * inv(B) * inv(A)
A^3
A^5 * A^(-2)
A.'.'
inv(A).'
(inv(A)).'
(A*B*C).'
C.' * B.' * A.'
(A^(-2)).'
(A.')^(-2)

% Ejercicio 19.

inv(A)*B
B*inv(A)

% Ejercicios 20 a 22.

function [A] = diagzero(n)         % Escribir el código a la izquierda en un archivo m.
          A = ones(n)-eye(n);      % Matriz con unos menos I_n.
          end
% También es útil conocer cómo usar un código más completo como:
```

```
función [A] = diagzero (n)
            para i = 1 : n, para j = 1:n,
   if i==j A(i, j)=0; else A(i, j)=1; end; end; end;
            A;
            end
% Después en la sesión MATLAB en el mismo directorio llamar las funciones:
diagzero(3)
diagzero(4)
diagzero(5)
inv(diagzero(3))
inv(diagzero(4))
inv(diagzero(5))
% A_n^(-1) tiene -(n-2)/(n-1) en la diagonal principal y 1/(n-1) fuera de ella.

% Ejercicio 24.

AA = diagzero(4)         % Usar diagzero o teclear directamente la matriz.
AA^4                     % Matriz que presenta la cantidad de caminatas de longitud 4.
AA+AA^2+AA^3+AA^4

% Ejercicio 25.

function [B] = sumpower(A,n)                  % Código en un archivo.
            B = A;
            for i=1: (n-1), B = B*A + A; end  % A ->A^2+A->A^3+A^2+A->. . .
            end
sumpower(AA,4)                                % sesión de teclear.
```

4

Espacios vectoriales

Matemáticas: el fundamento inconmovible de la ciencia. . .
Isaac Barrow (1630-1677)

Introducción

En este capítulo generalizaremos los conceptos básicos del capítulo 2: vectores, generadores e independencia lineal. Las propiedades comunes de la aritmética matricial y vectorial (teorema 1 de las secciones 2.1 y 3.1) se transforman en propiedades definitorias para un conjunto de vectores abstractos o generalizados, llamado *espacio vectorial.* Los conjuntos de matrices y de vectores ordinarios son ejemplos de espacios vectoriales. También lo son una gran variedad de otros conjuntos.

La ventaja principal de estas generalizaciones estriba en los inmensos ahorros de trabajo, porque las propiedades de los vectores abstractos se aplican a todos los ejemplos particulares. Además, las demostraciones se tornan claras y fáciles, porque no tienen la notación de algún ejemplo específico.

Grassmann y Peano

El matemático a quien se acredita la introducción y la primera aplicación de estas ideas es **Hermann Grassmann.**[1] Según los historiadores (Bourbaki, van der Waerden), Grassmann fue el primero en definir un espacio vectorial n dimensional (al que llamó "sistema de números hipercomplejos") y la independencia lineal.

Giuseppe Peano,[2] matemático italiano, aclaró el trabajo de Grassmann. Según Bourbaki, también a él se debe la definición actual (sin coordenadas) de una transformación lineal. Peano

[1] **Hermann Günther Grassmann** (1809–1877) nació y murió en Stettin, Alemania (hoy Polonia). Dio clases en la escuela superior de su pueblo natal, y en 1844 publicó un libro donde presentaba varias ideas nuevas de geometría n dimensional y espacios vectoriales. Como Grassmann no estaba adiestrado como matemático investigador, su libro es de difícil lectura, y no recibió tanto reconocimiento como ahora.

[2] **Giuseppe Peano** (1858-1932) fue criado desde los 11 años por su tío en Turín, Italia. Se graduó en la Universidad de Turín, donde enseñó durante el resto de su carrera. También dio clases en una academia militar cercana, de la cual fue forzado a renunciar cuando comenzó a explicar su "nuevo simbolismo". Peano es bien conocido por su curva "llenadora del espacio" y su definición axiomática de los números naturales. Aparte de las matemáticas, también intervino en el mejoramiento de la educación en escuelas secundarias y en lingüística.

vivió durante la "era axiomática" de las matemáticas y llegó a ser uno de los principales representantes de esa corriente. La definición de un espacio vectorial tuvo sus orígenes en la lectura que hizo Peano de los trabajos de Grassmann. Introdujo algunas notaciones matemáticas que se usan hoy como el símbolo $\in$, que significa *pertenece a*, *está en*, o *es un miembro de*.

4.1 Subespacios de $\mathbf{R}^n$

Objetivos del estudiante para esta sección
1. Saber qué es un subespacio de $\mathbf{R}^n$.
2. Conocer qué son las bases y cómo investigarlas.

En esta sección describiremos los conceptos fundamentales de *subespacio* y *base* en $\mathbf{R}^n$. Ésta es una preparación para los conceptos abstractos correspondientes de las secciones 4.2 y 4.3. Estos temas deben dominarse mediante un estudio cuidadoso y con la práctica.

Subespacios de $\mathbf{R}^n$

DEFINICIÓN

(Subespacio de $\mathbf{R}^n$)

Un subconjunto V no vacío de $\mathbf{R}^n$ se llama **subespacio (vectorial o lineal)** de $\mathbf{R}^n$ si satisface las siguientes propiedades.

1. Si $\mathbf{u}$ y $\mathbf{v}$ están en V, entonces $\mathbf{u} + \mathbf{v}$ está en V.
2. Si c es cualquier escalar y $\mathbf{u}$ está en V, entonces $c\mathbf{u}$ está en V.

Las propiedades 1 y 2 implican que cualquier combinación lineal de elementos de V también está en V. Si un conjunto S no vacío de $\mathbf{R}^n$ satisface la parte 1 de la definición, se dice que S es **cerrado bajo (o respecto a) la suma (vectorial)**. Si S cumple la parte 2, se dice que S es **cerrado bajo (o respecto a) la multiplicación por escalares**. Así, un subespacio de $\mathbf{R}^n$ es un subconjunto cerrado bajo la suma vectorial y la multiplicación por escalares.

Todo subespacio V de $\mathbf{R}^n$ contiene al vector cero $\mathbf{0}$. (V es no vacío, de modo que tiene al menos un elemento, por ejemplo $\mathbf{u}$. Pero entonces $0\mathbf{u} = \mathbf{0}$ está en V, según la parte 2 de la definición.)

■ EJEMPLO 1 $\{\mathbf{0}\}$ y $\mathbf{R}^n$ son subespacios de $\mathbf{R}^n$.

EXPLICACIÓN $\{\mathbf{0}\}$ es un subespacio de $\mathbf{R}^n$ porque

$$\mathbf{0} + \mathbf{0} = \mathbf{0}, \quad \text{y} \quad c\mathbf{0} = \mathbf{0} \text{ para toda } c \in \mathbf{R}$$

$\mathbf{R}^n$ es un subespacio de $\mathbf{R}^n$ porque la suma de dos vectores n cualesquiera es un vector n, y cualquier múltiplo escalar de un vector n es de nuevo un vector n. □

$\{\mathbf{0}\}$ también se llama **subespacio cero** de $\mathbf{R}^n$. $\{\mathbf{0}\}$ y $\mathbf{R}^n$ son **subespacios triviales** de $\mathbf{R}^n$.

■ EJEMPLO 2

$$V = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}, \ x, y \in \mathbf{R} \right\}$$

es un subespacio de $\mathbf{R}^3$.

EXPLICACIÓN V es no vacío, porque contiene al vector cero (suponiendo que $x = y = 0$). La suma de dos vectores en V,

$$\begin{bmatrix} x_1 \\ y_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ 0 \end{bmatrix}$$

también está en V. Así, se aplica la parte 1 de la definición. Cualquier múltiplo escalar de un vector en V

$$c \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} cx \\ cy \\ 0 \end{bmatrix}$$

también está en V. Entonces se aplica la parte 2 de la definición. Por consiguiente, es un subespacio de $\mathbf{R}^3$.

■ EJEMPLO 3 $V = \{(x, y, x + y), x, y \in \mathbf{R}\}$ es un subespacio de $\mathbf{R}^3$.

EXPLICACIÓN V es no vacío. (¿Por qué?) Sean $\mathbf{v}_1 = (x_1, y_1, x_1 + y_1)$ y $\mathbf{v}_2 = (x_2, y_2, x_2 + y_2)$ cualesquiera elementos de V y sea c cualquier escalar. Entonces

$$(x_1, y_1, x_1 + y_1) + (x_2, y_2, x_2 + y_2) = (x_1 + x_2, y_1 + y_2, (x_1 + x_2) + (y_1 + y_2))$$

y

$$c(x_1, y_1, x_1 + y_1) = (cx_1, cy_1, (cx_1) + (cy_1))$$

Por consiguiente, $\mathbf{v}_1 + \mathbf{v}_2, c\mathbf{v}_1 \in V$. Entonces, V es un subespacio de $\mathbf{R}^3$.

■ EJEMPLO 4 ¿Es el conjunto

$$T = \left\{ \begin{bmatrix} x \\ x + 1 \end{bmatrix}, \ x \in \mathbf{R} \right\}$$

un subespacio de $\mathbf{R}^2$?

RESPUESTA No, el vector cero no está en T, porque si $\begin{bmatrix} x \\ x+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, entonces $x = 0$ y $x + 1 = 0$, lo cual es un sistema inconsistente. (También puede demostrarse con facilidad que T no es cerrado bajo la suma o la multiplicación por escalar.) ☐

■ EJEMPLO 5 Si $\mathbf{v}_1$ y $\mathbf{v}_2$ están en $\mathbf{R}^n$. Demuestre que Gen$\{\mathbf{v}_1, \mathbf{v}_2\}$ es un subespacio de $\mathbf{R}^n$.

SOLUCIÓN Sea $V = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\} = \{c_1\mathbf{v}_1 + c_2, \mathbf{v}_2, c_1, c_2 \in \mathbf{R}\}$. Por tanto, V es no vacío, porque contiene a $\mathbf{v}_1$ y $\mathbf{v}_2$. Sean $\mathbf{u}, \mathbf{v} \in V$. Entonces hay escalares c_1, c_2 y d_1, d_2 tales que $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ y $\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2$. Así,

$$\mathbf{u} + \mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + d_1\mathbf{v}_1 + d_2\mathbf{v}_2 = (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2$$

Por consiguiente $\mathbf{u} + \mathbf{v} \in V$, y V es cerrado bajo la suma. También, para cualquier escalar c,

$$c\mathbf{u} = c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2$$

Esto demuestra que $c\mathbf{u} \in V$. Por tanto, V es cerrado bajo la multiplicación por escalar, y entonces es un subespacio de $\mathbf{R}^n$. ☐

El teorema siguiente nos permitirá generalizar, de acuerdo con el teorema 10 de la sección 2.3.

TEOREMA 1

> Si $\mathbf{v}_1, \ldots, \mathbf{v}_k$ son vectores n, entonces Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un subespacio de $\mathbf{R}^n$.

■ EJEMPLO 6 Toda línea l que pase por el origen en $\mathbf{R}^2$ ($\mathbf{R}^3$) es un subespacio de $\mathbf{R}^2$ ($\mathbf{R}^3$) (figura 4.1).

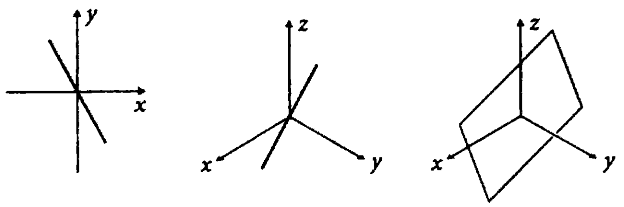


Figura 4.1 Las rectas y los planos que pasan por el origen son subespacios.

SOLUCIÓN Sea $\mathbf{u}$ un vector 2 no cero en l. Entonces $l = \text{Gen}\{\mathbf{u}\}$, como vimos en la sección 2.3. Pero Gen$\{\mathbf{u}\}$ es un subespacio de $\mathbf{R}^2$, de acuerdo con el teorema 1. De igual manera puede demostrarse que una línea l en $\mathbf{R}^3$ es un subespacio de $\mathbf{R}^3$. ☐

■ EJEMPLO 7 Todo plano $\mathcal{P}$ que pasa por el origen es un subespacio de $\mathbf{R}^3$ (figura. 4.1).

SOLUCIÓN Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores de $\mathcal{P}$ tales que no sean colineales. Entonces $\mathcal{P} = \text{Gen}\{\mathbf{u}, \mathbf{v}\}$, de acuerdo con la sección 2.3. En consecuencia, $\mathcal{P}$ es un subespacio de $\mathbf{R}^3$, según el teorema 1.

Observe que las rectas y los planos que no pasan por el origen (figura 4.2) **no** son subespacios. (¿Por qué?) □

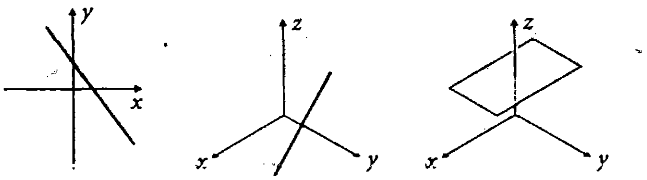


Figura 4.2 Las rectas y los planos que no pasan por el origen no son subespacios.

Con frecuencia se demuestra que V es un subespacio de $\mathbf{R}^n$ escribiéndolo en forma del generador de un conjunto de vectores.

■ EJEMPLO 8 ¿Es el conjunto

$$V = \left\{ \begin{bmatrix} a - 3b \\ 2a + b \\ a \end{bmatrix}, \ a, b \in \mathbf{R} \right\}$$

un subespacio de $\mathbf{R}^3$?

RESPUESTA Sí, porque

$$\begin{bmatrix} a - 3b \\ 2a + b \\ a \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + b \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$$

vemos que $V = \text{Gen}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} \right\}$. Por consiguiente, V es un subespacio de $\mathbf{R}^3$, según el teorema 1. □

Base de un subespacio de $\mathbf{R}^n$

En este párrafo presentaremos el concepto fundamental de una base de un subespacio de $\mathbf{R}^n$. Las bases son conjuntos muy privilegiados de vectores. Son linealmente independientes y al mismo tiempo conjuntos generadores. La elección y uso de una base de un subespacio se parece a la elección y uso de un marco de coordenadas en el plano o en el espacio. Con frecuencia, una elección ingeniosa de las coordenadas (o de la base) ayuda a simplificar un cálculo que de otro modo sería complicado.

DEFINICIÓN

(Base)

Un subconjunto no vacío $\mathcal{B}$ de un subespacio V no cero de $\mathbf{R}^n$ es una **base** de V si

1. $\mathcal{B}$ es linealmente independiente;
2. $\mathcal{B}$ genera a V.

Véanse las figuras 4.3 y 4.4.

También se acostumbra decir que el *conjunto vacío* es la única base del subespacio cero $\{\mathbf{0}\}$.

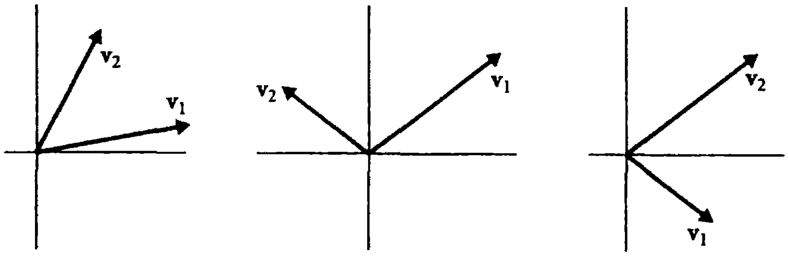


Figura 4.3 Algunas bases del plano.

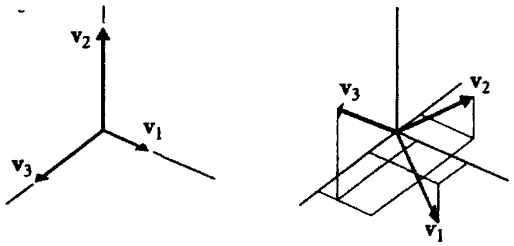


Figura 4.4 Bases del espacio.

■ EJEMPLO 9 Demuestre que $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ es una base de $\mathbf{R}^n$, a la cual se llama **base estándar** o **base normal** de $\mathbf{R}^n$.

SOLUCIÓN De acuerdo con los ejemplos 24 (sección 2.3) y 35 (sección 2.4), $\mathcal{B}$ genera a $\mathbf{R}^n$ y es linealmente independiente. Por consiguiente, es una base de $\mathbf{R}^n$. □

■ EJEMPLO 10 Compruebe que

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$$

es una base de $\mathbf{R}^3$.

SOLUCIÓN Primero es necesario demostrar que $\mathcal{B}$ es linealmente independiente y que genera a $\mathbf{R}^3$. Sea A la matriz cuyas columnas son los vectores de $\mathcal{B}$. Como

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 1 & 1 \\ -1 & 2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & -8 \end{bmatrix}$$

vemos que cada columna de A es una columna pivote; y por tanto, las columnas de A son linealmente independientes, de acuerdo con el teorema 12 de la sección 2.4. También, cada renglón de A tiene una posición pivote. Por consiguiente, las columnas de A generan a $\mathbf{R}^3$, de acuerdo con el teorema 2 de la sección 2.1. ▭

Los ejemplos 9 y 10 demuestran que un subespacio puede tener varias bases.

■ EJEMPLO 11 ¿Es el conjunto $S = \{(1, 0, 0), (0, 1, 0)\}$ una base en $\mathbf{R}^3$? ¿Es S una base del subespacio $V = \{(x, y, 0), x, y \in \mathbf{R}\}$ de $\mathbf{R}^3$?

RESPUESTA S es linealmente independiente, porque todas las columnas de $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ son pivote. S no es una base de $\mathbf{R}^3$ porque A tiene 3 renglones y sólo 2 pivotes, de modo que S no genera a $\mathbf{R}^3$. Por otra parte, S genera al menor espacio V porque

$$(x, y, 0) = x(1, 0, 0) + y(0, 1, 0)$$

Por consiguiente, S *sí es* una base de V. ▭

■ EJEMPLO 12 ¿Es el conjunto $T = \{(1, 1, 1), (2, 1, -1), (1, 0, -2)\}$ una base de $\mathbf{R}^3$?

RESPUESTA No, puesto que $\begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & -1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$, T no es linealmente independiente (3 columnas y sólo 2 pivotes). En realidad T ni siquiera genera a $\mathbf{R}^3$ (3 renglones, sólo 2 pivotes). ▭

■ EJEMPLO 13 ¿Es el conjunto $S = \{(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 1)\}$ una base en $\mathbf{R}^3$?

RESPUESTA No, porque S es linealmente dependiente. ▭

■ EJEMPLO 14 Demuestre que el conjunto $\mathcal{B} = \{(2, 1, -1, 1), (1, 0, -2, 1), (0, 0, 0, 1)\}$ es una base del subespacio $V = \{(2x + y, x, -x - 2y, x + y + z), x, y, z \in \mathbf{R}\}$ de $\mathbf{R}^4$.

SOLUCIÓN Puesto que

$$(2x + y, x, -x - 2y, x + y + z) = x(2, 1, -1, 1) + y(1, 0, -2, 1) + z(0, 0, 0, 1)$$

vemos que V es generado por $\mathcal{B}$. Dejamos al lector la demostración de que $\mathcal{B}$ es linealmente independiente. En consecuencia, $\mathcal{B}$ es una base de V. ▭

Coordenadas con respecto a la base

En este párrafo, V es un subespacio no cero de $\mathbf{R}^n$.

TEOREMA 2

> **(Unicidad de coordenadas)**
>
> Un subconjunto $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ de V es una base de V si y sólo si para cada vector $\mathbf{v}$ en V hay escalares **únicos** $c_1, \ldots, c_k$ tales que
>
> $$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

DEMOSTRACIÓN Sea $\mathcal{B}$ una base de V. Por consiguiente, $\mathcal{B}$ genera a V, y cada vector $\mathbf{v}$ es una combinación lineal de vectores en $\mathcal{B}$. También, $\mathcal{B}$ es linealmente independiente. En consecuencia, la representación de $\mathbf{v}$ como combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_k$ es única, de acuerdo con el teorema 16 de la sección 2.4.

De manera inversa, sea $\mathcal{B}$ tal que todo vector $\mathbf{v}$ se escribe en la única forma como $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$. Por tanto, $\mathcal{B}$ genera a V. Ahora, sea $d_1\mathbf{v}_1 + \cdots + d_k\mathbf{v}_k = \mathbf{0}$ para algunos escalares $d_1, \ldots, d_k$. Como $0\mathbf{v}_1 + \cdots + 0\mathbf{v}_k = \mathbf{0}$, la unicidad supuesta de la representación implica que $d_1 = \cdots = d_k = 0$. Por consiguiente, $\mathcal{B}$ es linealmente independiente, y entonces $\mathcal{B}$ es una base de V. ▭

Los escalares únicos $c_1, \ldots, c_k$ que expresan al vector $\mathbf{v} \in V$ como la combinación lineal de una base $\mathcal{B}$ de V se llaman **coordenadas de v respecto a $\mathcal{B}$**. El vector cuyos componentes son $c_1, \ldots, c_k$ se llama **vector de coordenadas de v con respecto a $\mathcal{B}$** y se representa con $[\mathbf{v}]_{\mathcal{B}}$.

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} \qquad \left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$$

■ EJEMPLO 15 Determine las coordenadas y el vector de coordenadas de $\mathbf{v} = (4, 0, -4)$ con respecto a la base $\mathcal{B} = \{(1, 1, -1), (0, 1, 2), (-2, 1, 0)\}$ de $\mathbf{R}^3$.

SOLUCIÓN El hecho de que $\mathcal{B}$ sea una base de $\mathbf{R}^3$ se verificó en el ejemplo 10. Ahora se necesitan escalares c_1, c_2 y c_3 tales que $c_1 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + c^3 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ -4 \end{bmatrix}$. Al resolver el sistema correspondiente se obtienen $c_1 = 2$, $c_2 = -1$ y $c_3 = -1$. Por consiguiente, $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$ ▭

■ EJEMPLO 16 Encuentre las coordenadas de $\mathbf{v} = (3, -2, 0)$ en el subespacio $V = \{(x, y, 0), x, y \in \mathbf{R}\}$ con respecto a la base $\mathcal{B} = \{(1, 0, 0), (0, 1, 0)\}$ de V.

SOLUCIÓN Como

$$\mathbf{v} = \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (-2)\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

entonces

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

■ EJEMPLO 17 Si $\mathbf{v}$ es cualquier vector n y $\mathcal{B}$ es la base estándar de $\mathbf{R}^n$, compruebe que

$$\mathbf{v} = [\mathbf{v}]_{\mathcal{B}}$$

SOLUCIÓN Se deja como ejercicio.

Ejercicios 4.1

Subespacios

En los ejercicios 1 a 10, demuestre que los conjuntos dados de vectores n son subespacios de $\mathbf{R}^n$.

1. $V = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix}, a \in \mathbf{R} \right\}$
2. $V = \left\{ \begin{bmatrix} a-b \\ 2a+b \end{bmatrix}; a, b \in \mathbf{R} \right\}$
3. $V = \left\{ \begin{bmatrix} a \\ 0 \\ -2a \end{bmatrix}, a \in \mathbf{R} \right\}$
4. $V = \left\{ \begin{bmatrix} a-c \\ b+c \\ 5c \end{bmatrix}, a, b, c \in \mathbf{R} \right\}$
5. $V = \left\{ \begin{bmatrix} a-b \\ b-c \\ c-d \\ d-a \end{bmatrix}, a, b, c, d \in \mathbf{R} \right\}$
6. El conjunto de todos los vectores 3 cuyos primero y último componente son cero.
7. El conjunto de todos los vectores 4 cuyos primeros tres componentes son cero.
8. El conjunto de todos los múltipos de $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ por un escalar.
9. El conjunto de todas las combinaciones lineales de $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$
10. Gen $\left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \end{bmatrix} \right\}$

En los ejercicios 11 a 29 determine si los conjuntos dados de vectores n son subespacios de $\mathbf{R}^n$.

11. $\left\{ \begin{bmatrix} a \\ a \end{bmatrix}, a \in \mathbf{R} \right\}$
12. $\left\{ \begin{bmatrix} a+1 \\ a \end{bmatrix}, a \in \mathbf{R} \right\}$
13. $\left\{ \begin{bmatrix} a \\ b \end{bmatrix}, a + b \leq 2, a, b \in \mathbf{R} \right\}$
14. $\left\{ \begin{bmatrix} a \\ b \end{bmatrix}, a^2 + b^2 = 1, a, b \in \mathbf{R} \right\}$
15. $\left\{ \begin{bmatrix} a \\ b \end{bmatrix}, a \leq 1, b \geq 1, a, b \in \mathbf{R} \right\}$
16. $\left\{ \begin{bmatrix} a \\ b \end{bmatrix}, a \leq b, a, b \in \mathbf{R} \right\}$
17. $\left\{ \begin{bmatrix} a \\ b \end{bmatrix}, a = -3b, a, b \in \mathbf{R} \right\}$

18. $\left\{\begin{bmatrix} a+1 \\ a+b \end{bmatrix}, a,b \in \mathbf{R}\right\}$

19. $\left\{\begin{bmatrix} a-1 \\ 2a+1 \end{bmatrix}, a,b \in \mathbf{R}\right\}$

20. El conjunto de todos los vectores 3 cuyos dos primeros componentes son iguales.
21. El conjunto de todos los vectores 4 cuyos dos últimos componentes son iguales.
22. El conjunto de todos los vectores 3 cuyo primer componente es el doble del segundo.
23. El conjunto de todos los vectores 2 en el primer cuadrante.
24. El conjunto de todos los vectores 2 en el cuarto cuadrante.
25. El conjunto de todos los vectores 3 en el primer octante.
26. La recta determinada por los puntos cuyas coordenadas son (1, 0), (0, 1).
27. La recta determinada por los puntos cuyas coordenadas son (−1, −1), (1, 1).
28. El plano determinado por los puntos cuyas coordenadas son (1, 0, 0), (0, 0, 1), (0, 0, 0).
29. El plano determinado por los puntos cuyas coordenadas son (1, 0, 0), (0, 1, 0), (1, 1, 1).
30. Sean V_1 y V_2 subespacios de $\mathbf{R}^n$. Demuestre que la intersección $V_1 \cap V_2$ también es un subespacio de $\mathbf{R}^n$.
31. Sean V_1 y V_2 dos planos que pasan por el origen en $\mathbf{R}^3$. Explique, geométricamente, por qué $V_1 \cap V_2$ es un subespacio de $\mathbf{R}^3$.
32. Sean $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vectores n. Demuestre que Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es el mínimo subespacio de $\mathbf{R}^n$ que contiene al conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. (*Sugerencia:* Si V es un subespacio de $\mathbf{R}^n$ que contiene a $\mathbf{v}_1, \ldots, \mathbf{v}_k$, demuestre que Gen$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq V$.)

Bases

En los ejercicios 33 a 39 determine si los conjuntos dados de vectores n son bases de $\mathbf{R}^n$.

33. $\left\{\begin{bmatrix} 0 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}$

34. $\left\{\begin{bmatrix} 20 \\ 10 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}$

35. $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$

36. $\left\{\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}\right\}$

37. $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$

38. $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}\right\}$

39. $\left\{\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \end{bmatrix}\right\}$

En los ejercicios 40 a 44 determine una base para los subespacios descritos.

40. El subespacio V de $\mathbf{R}^2$ del ejercicio 1

41. El subespacio V de $\mathbf{R}^2$ del ejercicio 2

42. El subespacio V de $\mathbf{R}^3$ del ejercicio 3

43. El subespacio V de $\mathbf{R}^3$ del ejercicio 4

44. El subespacio V de $\mathbf{R}^4$ del ejercicio 5

En los ejercicios 45 a 49 determine *mediante inspección* si los conjuntos dados de vectores n son bases de $\mathbf{R}^n$.

45. $\left\{\begin{bmatrix} 71 \\ 0 \end{bmatrix}, \begin{bmatrix} 25 \\ 34 \end{bmatrix}\right\}$

46. $\left\{\begin{bmatrix} 75 \\ -45 \end{bmatrix}, \begin{bmatrix} -150 \\ 90 \end{bmatrix}\right\}$

47. $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}\right\}$

48. $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}\right\}$

49. $\left\{\begin{bmatrix} -3 \\ 0 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 4 \\ 0 \end{bmatrix}\right\}$

En los ejercicios 50 a 53 determine si las columnas de las matrices de $m \times n$ forman bases de $\mathbf{R}^m$.

50. a. $\begin{bmatrix} 1 & 2 \\ 0 & 5 \end{bmatrix}$ b. $\begin{bmatrix} 2 & 0 \\ -1 & 4 \end{bmatrix}$

51. a. $\begin{bmatrix} 1 & 3 & 8 \\ 0 & 1 & 7 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 3 \\ -1 & 1 \\ 0 & 1 \end{bmatrix}$

52. a. $\begin{bmatrix} 1 & 3 & 8 \\ -5 & 1 & 7 \\ 0 & 0 & 4 \end{bmatrix}$ b. $\begin{bmatrix} 2 & 5 & 6 & 7 \\ 0 & 1 & 3 & 8 \\ 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

53. a. $\begin{bmatrix} 2 & 2 & 2 & 2 \\ 0 & 3 & 3 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 5 & 5 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 3 & 0 & 3 \\ 4 & 0 & 4 & 0 \\ 0 & 5 & 0 & 5 \end{bmatrix}$

54. Calcule una base para $V = \text{Gen}\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -3 \\ 8 \end{bmatrix}\right\}$.

55. Encuentre una base para $V = \text{Gen}\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 5 \end{bmatrix}\right\}$.

56. Determine una base para

$V = \text{Gen}$

57. Obtenga una base para

$$V = \text{Gen}\left\{\begin{bmatrix} -5 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 7 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 8 \end{bmatrix}\right\}$$

58. Determine una base para

$$V = \text{Gen}\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}\right\}$$

59. Compruebe que todo subespacio V de $\mathbf{R}^n$ tiene una base $\mathcal{B}$ formada por una cantidad finita de vectores.

(*Sugerencia:* Sea $\mathbf{v}_1 \neq \mathbf{0}$ en V. Si $\text{Gen}\{\mathbf{v}_1\} = V$, haga que $\mathcal{B} = \{\mathbf{v}_1\}$ y deténgase. En caso contrario, escoja un $\mathbf{v}_2 \in V$ que no esté en $\text{Gen}\{\mathbf{v}_1\}$. Entonces, demuestre que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente independiente, y continúe.)

Coordenadas

En los ejercicios 60 a 62 determine el vector n, $\mathbf{x}$, dando una base $\mathcal{B}$ de $\mathbf{R}^n$ y el vector n de coordenadas $[\mathbf{x}]_{\mathcal{B}}$.

60. $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix}\right\}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$

61. $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right\}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \\ 4 \end{bmatrix}$

62. $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -3 \end{bmatrix}\right\}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} a \\ b \end{bmatrix}$

En los ejercicios 63 a 67 determine el vector n de coordenadas $[\mathbf{x}]_{\mathcal{B}}$, dando una base $\mathcal{B}$ de $\mathbf{R}^n$ y $\mathbf{x}$.

63. $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}, \mathbf{x} = \begin{bmatrix} 17 \\ 4 \end{bmatrix}$

64. $\mathcal{B} = \left\{\begin{bmatrix} 4 \\ -7 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \end{bmatrix}\right\}, \mathbf{x} = \begin{bmatrix} -6 \\ 17 \end{bmatrix}$

65. $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ -1 \end{bmatrix}\right\}, \mathbf{x} = \begin{bmatrix} -8 \\ 6 \\ -1 \end{bmatrix}$

66. $\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}\right\},$

$\mathbf{x} = \begin{bmatrix} 4 \\ -1 \\ -4 \\ -4 \end{bmatrix}$

67. $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix}\right\}, \mathbf{x} = \begin{bmatrix} 7a + 3b \\ a - b \end{bmatrix}$

68. Si $\mathcal{B}$ es una base de un subespacio V de $\mathbf{R}^n$ y están $\mathbf{v}_1, \ldots, \mathbf{v}_k$ en V, demuestre cada una de las siguientes afirmaciones.

a. $\mathbf{v}$ en V es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_k$ si y sólo si $[\mathbf{v}]_{\mathcal{B}}$ es una combinación lineal de $[\mathbf{v}_1]_{\mathrm{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}}$.

b. $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es linealmente dependiente si y sólo si $\{[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}}\}$ es linealmente dependiente.

c. $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es linealmente independiente si y sólo si $\{[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}}\}$ lo es también.

4.2 Espacios vectoriales

Objetivos del estudiante para esta sección

1. Estudiar y practicar las definiciones de espacio y subespacio vectorial.
2. Elaborar los detalles de los ejemplos en la sección.

En esta sección estudiaremos ampliamente a $\mathbf{R}^n$ y sus operaciones. Consideraremos conjuntos generales en los que se pueden definir las operaciones abstractas de suma y multiplicación por escalar, no a partir de determinadas reglas directas y específicas, sino requiriendo que los conjuntos satisfagan las propiedades básicas de la suma vectorial y la multiplicación por escalar, expresadas en los teoremas 1 de las secciones 2.1 y 3.1.

DEFINICIÓN

(Espacio vectorial)

Sea V un conjunto equipado con dos operaciones, llamadas **suma** y **multiplicación por escalar**. La suma es una regla que asocia dos elementos cualesquiera, $\mathbf{u}$ y $\mathbf{v}$, de V con un tercero, la **suma** de $\mathbf{u}$ y $\mathbf{v}$, representada por $\mathbf{u} + \mathbf{v}$. La multiplicación por escalar es una regla que asocia cualquier escalar (real) c y cualquier elemento $\mathbf{u}$ de V con otro de V, el **múltiplo escalar de u** por c, el cual está representado por $c\mathbf{u}$. Ese conjunto V se denomina **espacio vectorial** (real) si las dos operaciones cumplen las propiedades siguientes, llamadas **axiomas**, de un espacio vectorial.

Suma:

(A1) $\mathbf{u} + \mathbf{v}$ pertenece a V para todas $\mathbf{u}, \mathbf{v} \in V$.
(A2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ para todas $\mathbf{u}, \mathbf{v} \in V$.
(A3) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ para todas $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
(A4) Para toda $\mathbf{u}$ en V existe un elemento único $\mathbf{0} \in V$, llamado **cero** de V, tal que para toda $\mathbf{u}$ en V,
$$\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$$
(A5) Para cada $\mathbf{u} \in V$ existe un elemento único $-\mathbf{u} \in V$, llamado **negativo** u **opuesto** de $\mathbf{u}$, tal que
$$\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$$

Multiplicación por escalar:

(M1) $a\mathbf{u}$ pertenece a V para toda $\mathbf{u} \in V$ y toda $a \in \mathbf{R}$.
(M2) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ para todas $\mathbf{u}, \mathbf{v} \in V$ y toda $a \in \mathbf{R}$.
(M3) $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ para toda $\mathbf{u} \in V$ y todas $a, b \in \mathbf{R}$.
(M4) $a(b\mathbf{u}) = (ab)\mathbf{u}$ para toda $\mathbf{u} \in V$ y todas $a, b \in \mathbf{R}$.
(M5) $1\mathbf{u} = \mathbf{u}$ para toda $\mathbf{u} \in V$.

Los elementos de un espacio vectorial se llaman **vectores.** Los axiomas (A1) y (M1) también se expresan diciendo que V es **cerrado bajo la suma** y **la multiplicación por escalar**. (A2) y (A3) son la **ley conmutativa** y la **ley asociativa**, respectivamente, y (M2) y (M3) son las **leyes distributivas**. Observe que *un espacio vectorial es un conjunto no vacío*, porque contiene un cero, de acuerdo con (A4).

OBSERVACIÓN Debe tenerse en cuenta que en la definición de un espacio vectorial no se especifican ni los vectores ni las operaciones. Pero se aceptan *cualesquiera* de ellas que satisfagan los axiomas. Algunos autores emplean diferente notación para representarlas, como por ejemplo $\oplus$ y $\odot$, para diferenciar de la suma y multiplicación de vectores n por escalar. No lo haremos aquí, porque puede llegar a confundirnos después de cierto tiempo.

Se acostumbra escribir $\mathbf{u} - \mathbf{v}$ para representar la suma $\mathbf{u} + (-\mathbf{v})$:

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

Los axiomas (A1), (A2), (A3) y (M1) permiten sumar varios múltiplos escalares de vectores, sin ocuparnos del orden o del agrupamiento de los términos. De hecho, es posible sumar cualquier conjunto *finito* de múltiplos escalares de vectores. Si $\mathbf{v}_1, \ldots, \mathbf{v}_n$ son vectores y $c_1, \ldots, c_n$ son escalares, entonces la expresión

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$$

está bien definida y se llama **combinación lineal** de $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Por ejemplo, $2\mathbf{u} - 4\mathbf{v} + 3\mathbf{w} + 0.5\mathbf{z}$ es una combinación lineal de $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ y $\mathbf{z}$.

TEOREMA 3

Sea V un espacio vectorial. Sean $\mathbf{u} \in V$ y $c \in \mathbf{R}$. Entonces

1. $0\mathbf{u} = \mathbf{0}$;
2. $c\mathbf{0} = \mathbf{0}$;
3. Si $c\mathbf{u} = \mathbf{0}$, entonces $c = 0$ o $\mathbf{u} = \mathbf{0}$;
4. $(-c)\mathbf{u} = -(c\mathbf{u})$.

Demostraremos las partes 1 y 4, y dejaremos las restantes como ejercicios.

DEMOSTRACIÓN DE 1

$$\begin{array}{llll}
 & 0\mathbf{u} + 0\mathbf{u} = (0+0)\mathbf{u} = 0\mathbf{u} & & \text{De acuerdo con (M3)} \\
\Rightarrow & (0\mathbf{u} + 0\mathbf{u}) + (-0\mathbf{u}) & = 0\mathbf{u} + (-0\mathbf{u}) & \text{Sumando } -0\mathbf{u} \\
\Rightarrow & 0\mathbf{u} + (0\mathbf{u} + (-0\mathbf{u})) & = \mathbf{0} & \text{De acuerdo con (A3) y (A5)} \\
\Rightarrow & 0\mathbf{u} + \mathbf{0} & = \mathbf{0} & \text{De acuerdo con (A5)} \\
\Rightarrow & 0\mathbf{u} & = \mathbf{0} & \text{De acuerdo con (A4)}
\end{array}$$

DEMOSTRACIÓN DE 4 Según (M3) y la parte 1,

$$c\mathbf{u} + (-c)\mathbf{u} = (c + (-c))\mathbf{u} = 0\mathbf{u} = \mathbf{0}$$

Así, $c\mathbf{u} + (-c)\mathbf{u} = \mathbf{0}$ y por tanto $(-c)\mathbf{u} + c\mathbf{u} = \mathbf{0}$, de acuerdo con (A2). Por consiguiente, $(-c)\mathbf{u} = -c\mathbf{u}$, según (A5). ☐

Los axiomas permiten desarrollar aritmética semejante a la de vectores n en un espacio vectorial. Por ejemplo, para demostrar que $\mathbf{u} + \mathbf{u} = 2\mathbf{u}$,

$$\mathbf{u} + \mathbf{u} = 1\mathbf{u} + 1\mathbf{u} = (1+1)\mathbf{u} = 2\mathbf{u}$$

Ejemplos de espacios vectoriales

Para comprobar que determinado conjunto es un espacio vectorial es preciso *definir*, o *especificar, explícitamente*

1. Las dos operaciones, suma y multiplicación por escalar.
2. El elemento que actuará como cero.

3. El negativo de cada elemento.[3]
4. A continuación, comprobar la vigencia de los axiomas.

Los siguientes son ejemplos importantes de espacios vectoriales. Usted deberá desarrollar todos los detalles.

■ EJEMPLO 18 $\mathbf{R}^n$ (Casos especiales: $\mathbf{R}$, $\mathbf{R}^2$, $\mathbf{R}^3$):

1. *Operaciones:* La suma vectorial normal y la multiplicación por escalar (véase la sección 2.1).
2. *Cero:* El vector n cero, $\mathbf{0}$.
3. *Negativo:* El vector negativo $-\mathbf{u}$ de cada vector n $\mathbf{u}$.
4. *Axiomas:* Todos los axiomas se cumplen, de acuerdo con el teorema 1 de la sección 2.1.

■ EJEMPLO 19 El conjunto M_{mn} de todas las matrices $m \times n$ con elementos reales:

1. *Operaciones:* La suma de matrices normal y la multiplicación por escalar (véase la sección 3.1).
2. *Cero:* La matriz $\mathbf{0}$ $m \times n$.
3. *Negativo:* La matriz negativa $-A$ de cada matriz A.
4. *Axiomas:* Se satisfacen (A1) y (M1), porque la suma de dos matrices $m \times n$ es también una matriz $m \times n$, y cualquier múltiplo de una matriz $m \times n$ por un escalar es una matriz de tamaño $m \times n$.

   El resto de los axiomas son consecuencia del teorema 1, sección 3.1.

■ EJEMPLO 20 El conjunto P de todos los polinomios con coeficientes reales:

1. *Operaciones:* Sea x la variable independiente de los polinomios. (a) Suma: Cuando son dos polinomios, esta operación se lleva a cabo sumando los coeficientes de las mismas potencias de x de los polinomios. (b) La multiplicación por escalar es la multiplicación de todo el polinomio por una constante.
2. *Cero:* El polinomio cero, $\mathbf{0}$, es aquel cuya totalidad de coeficientes es cero.
3. *Negativo:* El negativo $-p$ de un polinomio p tiene por coeficientes los opuestos de los coeficientes de p.
4. *Axiomas:* Se deja como ejercicio la verificación de los axiomas.

■ EJEMPLO 21 El conjunto $F(\mathbf{R})$ de todas las funciones de valor real definidas en $\mathbf{R}$:

1. *Operaciones:* Sean f y g dos funciones de valor real con dominio $\mathbf{R}$ y sea c cualquier escalar. (a) Suma: Se define la suma $f + g$ de f y g como la función cuyos valores están expresados por

$$(f + g)(x) = f(x) + g(x) \quad \text{para toda } x \in \mathbf{R}$$

Igualmente, el producto por escalar cf se define como sigue:

$$(cf)(x) = c\,f(x) \quad \text{para toda } x \in \mathbf{R}$$

(Véase figura 4.5.)

[3] Una vez definida una multiplicación por escalar, el negativo de $\mathbf{v}$ puede definirse como $(-1)\mathbf{v}$.

2. *Cero:* La función cero, $\mathbf{0}$, es aquella cuyos valores son todos cero.

$$\mathbf{0}(x) = 0 \quad \text{para toda } x \in \mathbf{R}$$

3. *Negativo:* La negativa $-f$ de f es la función $(-1)f$.
4. *Axiomas:* La comprobación de los axiomas se deja como ejercicio.

De manera más general, el conjunto $F(X)$ de todas las funciones de valor real definidas en un conjunto X es un espacio vectorial. Las operaciones, el cero y el negativo se definen en la misma forma. La única deferencia es que x se encuentra en el conjunto X, en lugar de estar en $\mathbf{R}$. $\square$

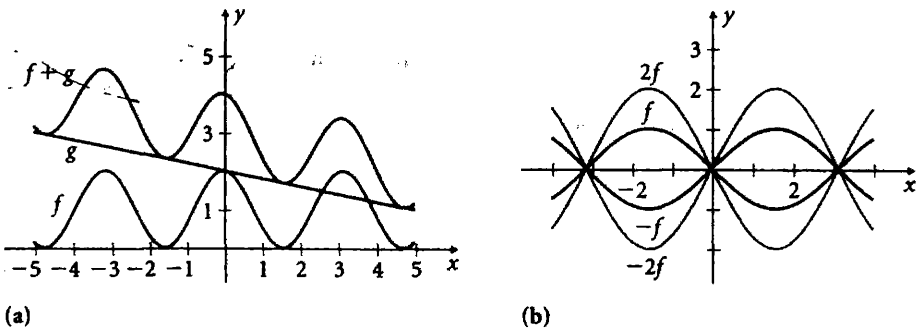


Figura 4.5 (a) Gráficas de f, g y $f + g$, (b) f y algunos de sus múltiplos por escalar.

Subespacios

Ahora explicaremos la noción importante de un subespacio. Generaliza nuestra noción familiar de un subespacio de $\mathbf{R}^n$.

DEFINICIÓN

> **(Subespacio)**
>
> Un subconjunto W de un espacio vectorial V se llama **subespacio** de V si W por sí mismo es un espacio vectorial bajo la suma y multiplicación por escalar tal como se definen en V.

Para demostrar que W es un subespacio de V no hay necesidad de comprobar los 10 axiomas. W toma sus operaciones y sus propiedades de V, para el cual ya se hizo la mayor parte de las verificaciones. El teorema siguiente aclara un poco más el concepto.

TEOREMA 4

> **(Criterio para un subespacio)**
>
> Para que un subconjunto no vacío W de un espacio vectorial V sea un subespacio, debe cumplirse lo siguiente:
>
> 1. Si $\mathbf{u}$ y $\mathbf{v}$ están en W, entonces $\mathbf{u} + \mathbf{v}$ está en W.
> 2. Si c es cualquier escalar y $\mathbf{u}$ está en W, entonces $c\mathbf{u}$ está en W.

DEMOSTRACIÓN Si W es un subespacio de V, todos los axiomas son válidos, en particular (A1) y (M1), los cuales corresponden exactamente a las condiciones 1 y 2.

A la inversa, si W es un subconjunto que satisface las condiciones 1 y 2, entonces (A1) y (M1) son válidas. (A2), (A3), (M2), (M3), (M4) y (M5) también se cumplen porque son válidas en V. Se necesita verificar (A4) y (A5). La condición 2 implica que $0\mathbf{u} = \mathbf{0}$ está en W cuando $\mathbf{u}$ está en W, igualando $c = 0$. Del mismo modo $(-1)\mathbf{u} = -\mathbf{u}$ está en W para toda $\mathbf{u}$ en W, eligiendo $c = -1$. Las ecuaciones de (A4) y (A5) son consecuencia de lo anterior. ☐

■ EJEMPLO 22 (Los subespacios triviales) $\{\mathbf{0}\}$ y V son subespacios de V, llamados **subespacios triviales** de V. $\{\mathbf{0}\}$ es el **subespacio cero** de V.

SOLUCIÓN Es claro que V es un subespacio de sí mismo. $\{\mathbf{0}\}$ también es un subespacio, porque se aplican las condiciones 1 y 2 del teorema 4:

$$\mathbf{0} + \mathbf{0} = \mathbf{0} \quad \text{y} \quad c\mathbf{0} = \mathbf{0} \text{ para toda } c \in \mathbf{R}$$ ☐

La afirmación del teorema 4 para $V = \mathbf{R}^n$ sirvió como definición de un subespacio de $\mathbf{R}^n$ en la sección 4.1.

■ EJEMPLO 23 Todo subespacio de $\mathbf{R}^n$, en términos de la sección 4.1, es un subespacio de $\mathbf{R}^n$. ☐

En particular se tienen los ejemplos siguientes.

■ EJEMPLO 24 Toda recta que pasa por el origen en $\mathbf{R}^3$ ($\mathbf{R}^2$) es un subespacio de $\mathbf{R}^3$ ($\mathbf{R}^2$). ☐

■ EJEMPLO 25 Todo plano que pasa por el origen es un subespacio de $\mathbf{R}^3$. ☐

NOTA Un subespacio contiene siempre el vector cero $\mathbf{0}$. Una recta que no pasa por el origen en $\mathbf{R}^3$ no es un subespacio de $\mathbf{R}^3$. $\{(x, 1), x \in \mathbf{R}\}$ no es un subespacio de $\mathbf{R}^2$, porque (0, 0) no está en el conjunto.

Recuérdese que el **grado** de un polinomio no cero está dado por la mayor potencia de x cuyo coeficiente sea distinto de cero. Por ejemplo, el grado de $1 - 2x^2 + x^6$ es 6. El grado de cualquier polinomio constante distinto de cero es cero. *El grado del polinomio cero está indefinido.*

■ EJEMPLO 26 Demuestre que el conjunto P_n formado por todos los polinomios de grado menor que, o igual a n, y el polinomio cero, es un subespacio de P.

SOLUCIÓN Un polinomio en P_n tiene la forma

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

La suma de dos polinomios de este tipo da como resultado un polinomio de grado menor que o igual a n, o es cero. También, un múltiplo constante de uno de esos polinomios es un polinomio de grado menor que o igual a n, o a cero. Así se satisfacen las condiciones 1 y 2 del teorema 4, y P_n es un subespacio de P. □

■ EJEMPLO 27 Compruebe que el conjunto $W = \{c\mathbf{v}, c \in \mathbf{R}\}$ de todos los múltiplos escalares del vector fijo $\mathbf{v}$ de un espacio vectorial V es un subespacio de V.

SOLUCIÓN W no es vacío. Contiene a $\mathbf{v}$ (¿Por qué?) Porque

$$c_1\mathbf{v} + c_2\mathbf{v} = (c_1 + c_2)\mathbf{v} \qquad \text{y} \qquad r(c\mathbf{v}) = (rc)\mathbf{v}$$

W es cerrado bajo la suma y la multiplicación por escalar. Por tanto, W es un subespacio de V. □

■ EJEMPLO 28 Si $\mathbf{a}$ es un vector fijo en $\mathbf{R}^3$ y W es el conjunto de todos los vectores ortogonales en $\mathbf{a}$, demuestre que W es un subespacio de $\mathbf{R}^3$.

SOLUCIÓN W no es vacío, porque contiene $\mathbf{0}$. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en W. Entonces $\mathbf{a} \cdot \mathbf{u} = 0$ y $\mathbf{a} \cdot \mathbf{v} = 0$. Así,

$$\mathbf{a} \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{a} \cdot \mathbf{u} + \mathbf{a} \cdot \mathbf{v} = 0 + 0 = 0$$

Por consiguiente, $\mathbf{u} + \mathbf{v}$ es ortogonal a $\mathbf{a}$, de modo que $\mathbf{u} + \mathbf{v}$ es un elemento de W. Si c es cualquier escalar, entonces

$$\mathbf{a} \cdot (c\mathbf{u}) = c(\mathbf{a} \cdot \mathbf{u}) = c0 = 0$$

implica que $c\mathbf{u}$ está en W. Por lo anterior, W es cerrado bajo la suma y la multiplicación por escalar y, en consecuencia, es un subespacio de $\mathbf{R}^3$. □

■ EJEMPLO 29 (Se requiere cálculo) El conjunto $C(\mathbf{R})$ de todas las funciones continuas de valor real, definido en $\mathbf{R}$, es un subespacio de $F(\mathbf{R})$. (Compruébelo.) □

■ EJEMPLO 30 Tanto P_n como P son subespacios de $F(\mathbf{R})$. (Compruébelo.) □

■ EJEMPLO 31 El conjunto D_n de todas las matrices diagonales de tamaño n es un subespacio de M_{nn}. (Compruébelo.) □

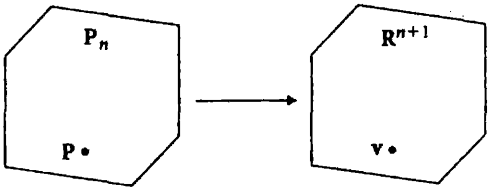
Correspondencia con $\mathbf{R}^n$

Se puede considerar que los elementos de P_n son vectores $(n + 1)$ si se opta por eliminar la variable y sus potencias, y anotar sólo los coeficientes en un orden definido. (Considere el análogo entre un sistema lineal y su matriz aumentada.)

Por ejemplo, $p(x) = 1 + 4x - 2x^2$ y $q(x) = -3 + x^2$ en P_3 pueden considerarse como los vectores 3

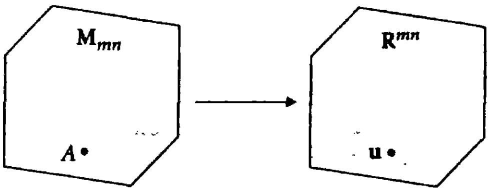
$$p' = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} \quad \text{y} \quad q' = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

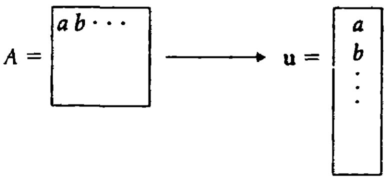
Observe que la suma de p y q corresponde a sumar p' y q'. También, todo múltiplo escalar cp corresponde a cp'. Así, en cierto sentido, es posible identificar a P_2 con $\mathbf{R}^3$ (y, en general, P_n con $\mathbf{R}^{n+1}$) como espacios vectoriales, lo cual simplifica sustancialmente la notación. Pero aun considerando que P_n equivale a $\mathbf{R}^{n+1}$, deben mantenerse vigentes ambas notaciones. Si elimináramos por completo a P_n, perderíamos perspectiva. Por ejemplo, cuando multiplicamos polinomios, el mantener las potencias de x ayuda a calcular el producto en forma correcta.



$$p = a_0 + a_1x + \cdots + a_nx^n \longrightarrow \mathbf{v} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}$$

En forma análoga podemos considerar los elementos de M_{mn} como vectores (mn). Por ejemplo, puede decirse que $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ es el vector 4 $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ y $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ como el vector 6 $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$





Ejercicios 4.2

NOTA A menos que se diga otra cosa, todos los subconjuntos de $\mathbf{R}^n$, P, M_{mn} o $F(\mathbf{R})$ tomados en cuenta obedecen a la suma y multiplicación por escalar definidas en la sección 4.2.

1. Demuestre que P_n es un espacio vectorial.
2. Compruebe que P es un espacio vectorial.
3. Si p y q están en P, demuestre que el conjunto $\{ap + bq$, donde $a, b \in \mathbf{R}\}$ es un espacio vectorial.
4. Con las operaciones siguientes, ¿es $\mathbf{R}^2$ un espacio vectorial?
$$(x,y)+(x',y') = (x+x',y+y'), \qquad c(x,y) = (cx,y)$$
5. Con las operaciones siguientes, ¿es $\mathbf{R}^2$ un espacio vectorial?
$$(x,y)+(x',y') = (x+x',y+y'), \qquad c(x,y) = (0,0)$$
6. Demuestre que el conjunto de todos los polinomios reales de grado n *no es* un espacio vectorial.
7. Compruebe que el conjunto de todas las matrices invertibles 2×2 *no es* un espacio vectorial.

En los ejercicios 8 a 15, determine si el conjunto dado es un subespacio de P. Use la notación

$$p = a_0 + a_1x + \cdots + a_nx^n$$

8. $\{p \in P,\ a_0 = 0\}$
9. $\{p \in P,\ a_n = 1\}$
10. $\{p \in P,\ a_0 > 0\}$
11. $\{p \in P,\ p(1) = 1\}$
12. $\{p \in P,\ p(1) = 0\}$
13. $\{p \in P,\ p(0) = 0\}$
14. $\{p \in P,\ a_0 + a_1 + \cdots + a_n = 0\}$
15. $\{p \in P,\ a_1a_0 = 0,\ n > 1\}$

En los ejercicios 16 a 20, determine si el subconjunto dado de M_{22} es un subespacio de M_{22}.

16. El conjunto de matrices de la forma $\begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}$
17. El conjunto de matrices de la forma $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$
18. Todas las matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tales que $a + d = 0$.
19. Todas las matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tales que $a + d = 1$.
20. El conjunto de matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tales que $ad = 0$.

En los ejercicios 21 a 24, determine si el conjunto dado puede transformarse en un espacio vectorial con las reglas normales de suma y multiplicación por escalar para matrices y polinomios (x es la variable independiente de todos los polinomios que se consideran).

21. $\left\{\begin{bmatrix} p & c \\ 0 & q \end{bmatrix}, \text{donde}, p,q \in P_1,\ c \in \mathbf{R}\right\}$
22. $\left\{\begin{bmatrix} p & c \\ q & 0 \end{bmatrix}, \text{donde}, p \in P_1,\ q \in P_2,\ c \in \mathbf{R}\right\}$
23. $\{[xp \quad q], \text{donde}, p,q \in P\}$
24. $\{[x \quad p], \text{donde}, p \in P\}$

Considere los conjuntos V_p y V_i de las funciones **pares** e **impares**, respectivamente.

$$V_p = \{f \in F(\mathbf{R}) \text{ tal que } f(x) = f(-x) \text{ para toda } x \in \mathbf{R}\}$$
$$V_i = \{f \in F(\mathbf{R}) \text{ tal que } f(x) = -f(-x) \text{ para toda } x \in \mathbf{R}\}$$

25. ¿Es V_p un subespacio de $F(\mathbf{R})$?
26. ¿Es V_i un subespacio de $F(\mathbf{R})$?
27. Demuestre las partes 2 y 3 del teorema 3.
28. Sea V un espacio vectorial, y sean $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ tales que $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$. Mediante los axiomas demuestre que $\mathbf{u} = \mathbf{v}$.
29. Sea V un espacio vectorial, y sean $\mathbf{u}, \mathbf{v} \in V$ tales que $r\mathbf{v} = r\mathbf{u}$ para algún escalar r. Demuestre mediante los axiomas que $r = 0$, o bien $\mathbf{u} = \mathbf{v}$.
30. Compruebe que el producto cartesiano $V_1 \times V_2$ de dos espacios vectoriales V_1 y V_2
$$V_1 \times V_2 = \{(\mathbf{v}_1, \mathbf{v}_2), \text{ donde } \mathbf{v}_1 \in V_1,\ \mathbf{v}_2 \in V_2\}$$
es un espacio vectorial que cuenta con suma y multiplicación por escalares de componente por componente.
$$(\mathbf{u}_1,\mathbf{u}_2) + (\mathbf{v}_1,\mathbf{v}_2) = (\mathbf{u}_1 + \mathbf{v}_1, \mathbf{u}_2 + \mathbf{v}_2)$$
$$c(\mathbf{v}_1,\mathbf{v}_2) = (c\mathbf{v}_1, c\mathbf{v}_2)$$
31. Sean U y W subespacios de un espacio vectorial V. Demuestre que la intersección $U \cap W$ es un subespacio de V.

4.3 Independencia lineal y bases

Objetivos del estudiante para esta sección

1. Repasar las secciones 2.1, 2.3, 2.4 y 4.1.
2. Comprender los conceptos de generador, independencia lineal y base.

Las nociones fundamentales de conjunto generador, independencia lineal y base que se estudiaron para vectores n se extienden con facilidad a los vectores abstractos. Esta sección incluye una gran cantidad de ejemplos para ayudar a comprender detalladamente estos conceptos. Los teoremas generalizan a los de las secciones 2.3, 2.4 y 4.1, y sus demostraciones son esencialmente las mismas.

El generador de $\mathbf{v}_1, \ldots, \mathbf{v}_k$

DEFINICIÓN

(Generador)

Sea V un espacio vectorial, y $\mathbf{v}_1, \ldots, \mathbf{v}_k$ los vectores en V. El conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_k$ se llama **generador** de $\mathbf{v}_1, \ldots, \mathbf{v}_k$, y se representa con $\mathrm{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Si $V = \mathrm{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ se dice que $\mathbf{v}_1, \ldots, \mathbf{v}_k$ **genera** a V y que $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un **conjunto generador de** V.

■ EJEMPLO 32 Demuestre que los vectores siguientes están en $\mathrm{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$.

$$\mathbf{0}, \qquad \mathbf{v}_1, \qquad \mathbf{v}_2, \qquad \mathbf{v}_1 + \mathbf{v}_2, \qquad -2\mathbf{v}_1, \qquad 3\mathbf{v}_1 - 2\mathbf{v}_2$$

SOLUCIÓN Cada uno de esos vectores es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$, porque se puede escribir

$$\begin{aligned} \mathbf{0} &= 0\mathbf{v}_1 + 0\mathbf{v}_2 & \mathbf{v}_1 &= 1\mathbf{v}_1 + 0\mathbf{v}_2 & \mathbf{v}_2 &= 0\mathbf{v}_1 + 1\mathbf{v}_2 \\ \mathbf{v}_1 + \mathbf{v}_2 &= 1\mathbf{v}_1 + 1\mathbf{v}_2 & -2\mathbf{v}_1 &= -2\mathbf{v}_1 + 0\mathbf{v}_2 & 3\mathbf{v}_1 - 2\mathbf{v}_2 &= 3\mathbf{v}_1 + (-2)\mathbf{v}_2 \end{aligned}$$

■ EJEMPLO 33 Compruebe que $\mathrm{Gen}\{\mathbf{v}\}$ es el conjunto de todos los múltiplos escalares de $\mathbf{v}$.

$$\mathrm{Gen}\{\mathbf{v}\} = \{c\mathbf{v}, c \in \mathbf{R}\}$$

SOLUCIÓN Cualquier combinación lineal de $\mathbf{v}$ tiene la forma $c\mathbf{v}$ para algún escalar c. A la inversa, cualquier múltiplo escalar $c\mathbf{U}$ de $\mathbf{v}$ es una combinación lineal de $\mathbf{v}$.

Por ejemplo, el generador de $\{(1, -1)\}$ en $\mathbf{R}^2$ es la recta que pasa por el origen en dirección de $(1, -1)$, como en la figura 4.6(a). De igual manera, el generador de x^2 en P_2 consiste en todas las funciones que son múltiplos escalares de x^2 (los vectores x^2, $2x^2$, $4x^2$, $-1.5x^2$, $-3.5x^2$ del generador se muestran en la figura 4.6(b)).

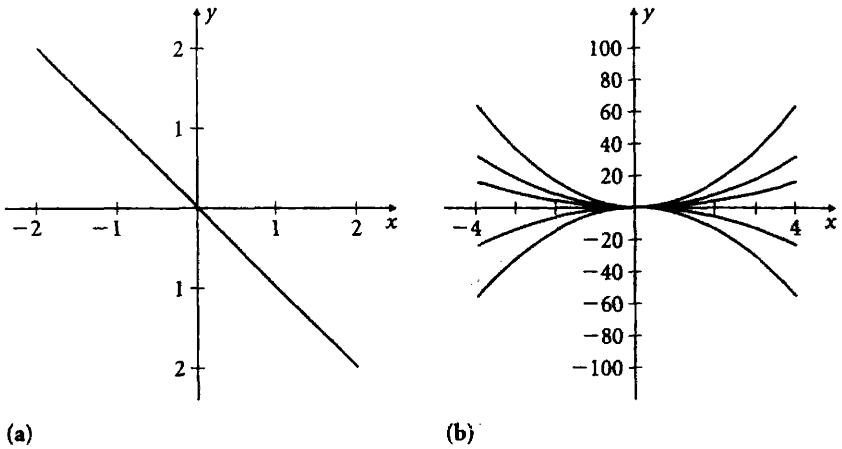


Figura 4.6 (a) Gen{(1, −1)}; (b) Gen{x^2}.

■ EJEMPLO 34 ¿Está $-1 + x^2$ en el generador de $p = 1 + x + x^3$ y $q = -x - x^2 - x^3$ en P_3?

SOLUCIÓN Sean a y b escalares tales que $-1 + x^2 = a(1 + x + x^3) + b(-x - x^2 - x^3)$. Entonces

$$-1 + x^2 = a + (a - b)x - bx^2 + (a - b)x^3$$

Igualando los coeficientes de las mismas potencias de x se obtiene el sistema

$$a - b = 0, \qquad -b = 1, \qquad a - b = 0, \qquad a = -1$$

cuya solución es $a = b = -1$. Por consiguiente, $-1 + x^2 = -p - q$. Como $-1 + x^2$ es una combinación lineal de p y q, está en Gen{p, q}. □

■ EJEMPLO 35 Demuestre que Gen{$1, x, x^2, x^3$} $= P_3$.

SOLUCIÓN Cualquier elemento $p = a_0 + a_1x + a_2x^2 + a_3x^3$ de P_3 ya es una combinación lineal de {$1, x, x^2, x^3$}. Además, a_0, a_1, a_2 y a_3 son los coeficientes de $1, x, x^2$ y x^3, respectivamente. □

Sea E_{ij} la matriz en M_{mn} cuyo (i, j)-ésimo elemento es 1 y el resto de ellos son cero. Por ejemplo, en M_{32},

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$E_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad E_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad E_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

■ EJEMPLO 36 Compruebe que $\{E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}\}$ genera a M_{32}.

SOLUCIÓN Es cierto porque

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = aE_{11} + bE_{12} + cE_{13} + dE_{21} + eE_{22} + fE_{23}$$

□

■ EJEMPLO 37 Demuestre que $\{(1, 2, -1), (-1, 1, -2), (1, 1, 1)\}$ genera a $\mathbf{R}^3$.

SOLUCIÓN Como

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ -1 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

A tiene 3 pivotes. Por consiguiente, los vectores generan a $\mathbf{R}^3$. □

■ EJEMPLO 38 Determine el generador de $\{A,B\}$ en M_{22}, donde

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

SOLUCIÓN Cualquier combinación de A y B es una matriz diagonal:

$$aA + bB = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ 0 & -b \end{bmatrix}$$

Por el contrario, cualquier matriz diagonal puede expresarse como una combinación lineal de A y B, porque

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = (a+b)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - b\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Por tanto, $\mathrm{Gen}\{A, B\} = D_2$ es el conjunto de todas las matrices diagonales 2×2. □

OBSERVACIÓN Las afirmaciones siguientes pueden comprobarse con facilidad, y se dejan como ejercicios.

1. $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ genera a $\mathbf{R}^n$.
2. $\{1, x, x^2, \ldots, x^n]$ genera a P_n.
3. $\{1, x, x^2, \ldots, x^n, \ldots\}$ genera a P.
4. $\{E_{11}, E_{12}, E_{13}, \ldots, E_{mn}]$ genera a M_{mn}.
5. $\{E_{11}, E_{22}, E_{33}, \ldots, E_{nn}\}$ genera a D_n.

TEOREMA 5

Si S es un subconjunto de un espacio vectorial V, entonces

1. $\mathrm{Gen}(S)$ es un subespacio de V;
2. $\mathrm{Gen}(S)$ es el subespacio más pequeño de V que contiene a S.

DEMOSTRACIÓN

1. Sean $\mathbf{u}_1, \ldots, \mathbf{u}_n$ y $\mathbf{v}_1, \ldots, \mathbf{v}_m$ vectores en S y sean $c_1, \ldots, c_n$ y $d_1, \ldots, d_n$ escalares. Considerando las dos combinaciones lineales de vectores en S:

$$c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n \qquad \text{y} \qquad d_1 \mathbf{v}_1 + \cdots + d_m \mathbf{v}_m$$

La suma

$$c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n + d_1 \mathbf{v}_1 + \cdots + d_m \mathbf{v}_m$$

está bien definida en V y de nuevo es una combinación lineal de vectores en S. Si c es cualquier escalar, entonces

$$c(c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n) = c(c_1 \mathbf{u}_1) + \cdots + c(c_n \mathbf{u}_n) = (cc_1)\,\mathbf{u}_1 + \cdots + (cc_n)\,\mathbf{u}_n$$

también es una combinación lineal de vectores en S. En consecuencia, Gen(S) es cerrada bajo la suma y la multiplicación de V por escalar. Por consiguiente, es un subespacio de V.

2. Si W es un subespacio que contiene a S, entonces también incluye todas las combinaciones lineales de sus elementos. En particular, todas las combinaciones lineales de los elementos de S. Pero éstos son los elementos de Gen(S). Por consiguiente, Gen (S) $\subseteq W$. Por lo anterior, Gen(S) es el subespacio contenido en cualquier subespacio W que incluya a S. Con esto se demuestra la segunda afirmación. □

TEOREMA 6

(Reducción del conjunto generador)

Si uno de los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$ del espacio vectorial V es una combinación lineal del resto, el generador permanece igual si se elimina ese vector.

DEMOSTRACIÓN Vea la comprobación del teorema 9, de la sección 2.3. □

Independencia lineal

DEFINICIÓN

Se dice que un conjunto de vectores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ de un espacio vectorial V es **linealmente dependiente** si hay escalares $c_1, \ldots, c_n$, *no todos cero*, tales que

$$c_1 \mathbf{v}_1 + \cdots + c_n \mathbf{v}_n = \mathbf{0} \tag{4.1}$$

Se dice que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ es **linealmente independiente** si no es linealmente dependiente. En otras palabras, la ecuación (4.1) implica que $c_1 = \cdots = c_n = 0$. Si S es cualquier subconjunto de V (posiblemente infinito), sólo lo llamaremos linealmente dependiente cuando contenga un subconjunto finito linealmente dependiente. En cualquier otro caso, S es linealmente independiente.

■ EJEMPLO 39 El conjunto $\{2 - x + x^2, 2x + x^2, 4 - 4x + x^2\}$ es linealmente dependiente en P_3, porque $4 - 4x + x^2 = 2(2 - x + x^2) - (2x + x^2)$. □

■ EJEMPLO 40 El conjunto $\{A, B, C\}$ es linealmente dependiente en M_n porque $A = B + C$.

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & -1 \\ 2 & 2 \end{bmatrix}$$

■ EJEMPLO 41 Los conjuntos $\{1, \cos 2x, \cos^2 x\}$ y $\{\operatorname{sen} x \cos x, \operatorname{sen} 2x\}$ son linealmente dependientes en $F(\mathbf{R})$, porque

$$\cos^2 x = \frac{1}{2} \cdot 1 + \frac{1}{2} \cos 2x \quad \text{y} \quad \operatorname{sen} 2x = 2 \operatorname{sen} x \cos x \quad \text{para toda } x \in \mathbf{R}$$

■ EJEMPLO 42 Demuestre que el conjunto $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ es linealmente independiente en M_{22}.

SOLUCIÓN Sean

$$c_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

o

$$\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Por consiguiente, $c_1 = c_2 = c_3 = c_4 = 0$.

■ EJEMPLO 43 Compruebe que el conjunto $\{1, x, \ldots, x^n\}$ es linealmente independiente en P_n.

SOLUCIÓN Si una combinación lineal $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ en $\{1, x, \ldots, x^n\}$ es el polinomio cero, $\mathbf{0}$, entonces

$$a_0 + a_1 r + \cdots + a_n r^n = 0 \quad \text{para toda } r \in \mathbf{R}$$

Del álgebra recordemos ahora el hecho básico que un *polinomio distinto de cero, de grado n, tiene como máximo n raíces*. Puesto que p tiene más de n raíces, debe ser el polinomio cero. Con esto se demuestra la afirmación.

■ EJEMPLO 44 Demuestre que el conjunto $\{x^2, 1 + x, -1 + x\}$ es linealmente independiente en P_3.

SOLUCIÓN Si una combinación $p(x) = ax^2 + b(1 + x) + c(-1 + x)$ es $\mathbf{0}$, el polinomio cero, entonces $p(x) = (b - c) + (b + c)x + ax^2 = 0$ para toda $x \in \mathbf{R}$. Por consiguiente, $b - c = 0$, $b + c = 0$ y $a = 0$. De lo anterior, $a = b = c = 0$, y el conjunto es linealmente independiente.

Notación de operaciones matriciales

En el ejemplo 44 podemos aprovechar la identificación de P_2 con $\mathbf{R}^3$ (que se explicó en la sección 4.1), para demostrar que $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$ es linealmente dependiente, en lugar de $\{x^2, 1 + x, -1 + x\}$. Esto es cierto porque

$$\begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \tag{4.2}$$

tiene tres columnas pivote.

■ EJEMPLO 45 Compruebe que el conjunto $\{\cos x, \operatorname{sen} x\}$ es linealmente independiente en $F(\mathbf{R})$.

SOLUCIÓN Si una combinación $a \cos x + b \operatorname{sen} x$ es la función cero, entonces

$$a \cos y + b \operatorname{sen} y = 0 \quad \text{para toda } y \in \mathbf{R}$$

Si y se iguala a 0 y a $\frac{\pi}{2}$, se obtiene $a = 0$ y $b = 0$, respectivamente. Así, $\{\cos x, \operatorname{sen} x\}$ es linealmente independiente en $F(\mathbf{R})$. ☐

OBSERVACIÓN Los siguientes conjuntos son linealmente independientes.

1. $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \subseteq \mathbf{R}^n$
2. $\{1, x, x^2, \ldots, x^n, \ldots\} \subseteq P$
3. $\{E_{11}, E_{12}, E_{13}, \ldots, E_{mn}\} \subseteq M_{mn}$
4. $\{1, \cos x, \cos 2x\} \subseteq F(\mathbf{R})$
5. $\{e^x, e^{2x}\} \subseteq F(\mathbf{R})$

TEOREMA 7

(Pruebas para la dependencia lineal)

Sea S un subconjunto de un espacio vectorial V.

1. Si S contiene un vector $\mathbf{v}$, se dice que es linealmente dependiente si y sólo si $\mathbf{v} = \mathbf{0}$.
2. Si S consiste en dos o más vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$, entonces S es linealmente dependiente, si y sólo si cuando menos un vector sea una combinación lineal de los vectores restantes.
3. Si S contiene dos o más vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$ con $\mathbf{v}_1 \neq \mathbf{0}$, entonces S es linealmente dependiente si y sólo si al menos un vector, por ejemplo $\mathbf{v}_i$ $(i \geq 2)$, es una combinación lineal de los vectores que le preceden, es decir, de $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$.

DEMOSTRACIÓN Revise la comprobación del teorema 14 en la sección 2.4. ☐

El teorema siguiente es semejante al teorema 15 de la sección 2.4. La demostración se deja como ejercicio.

TEOREMA 8

1. Cualquier conjunto de vectores que contiene a $\mathbf{0}$ es linealmente dependiente.
2. Dos vectores son linealmente dependientes si y sólo si uno es múltiplo escalar del otro.
3. Cualquier conjunto de vectores que contiene a un conjunto linealmente dependiente es a su vez linealmente dependiente.
4. Cualquier subconjunto de un conjunto linealmente independiente también es linealmente independiente.

TEOREMA 9

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente de vectores de un espacio vectorial V.

1. Cualquier vector $\mathbf{v}$ en el generador de S es expresable de manera única como combinación lineal de vectores en S; es decir, las relaciones

$$\mathbf{v} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n \quad \text{y} \quad \mathbf{v} = d_1\mathbf{v}_1 + \ldots d_n\mathbf{v}_n$$

implican que

$$c_1 = d_1, \ldots, c_n = d_n$$

2. Si $\mathbf{v}$ no está en el generador de S, entonces el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}\}$ es linealmente independiente.

DEMOSTRACIÓN Revise la demostración del teorema 16 de la sección 2.4. □

Base de un espacio vectorial

En este párrafo presentaremos el concepto fundamental de una base de un espacio vectorial. Las bases, al igual que en el caso de los subespacios de $\mathbf{R}^n$, son conjuntos generadores que también son linealmente independientes. Conocer la base de un espacio vectorial puede ser muy útil para comprender al espacio y sus propiedades.

DEFINICIÓN

(Base)

Un subconjunto $\mathcal{B}$ no vacío de un espacio vectorial V distinto de cero es una base de V si

1. $\mathcal{B}$ es linealmente independiente, y si
2. $\mathcal{B}$ genera a V.

El conjunto vacío es, por convención, la única base del espacio vectorial cero $\{\mathbf{0}\}$.

Para demostrar el siguiente e importante teorema se requiere teoría de conjuntos, lo cual está fuera del propósito de este libro y por tanto se omite.[4]

TEOREMA 10

(Existencia de la base)

Todo espacio vectorial tiene al menos una base.

A continuación citaremos algunos ejemplos de bases. Hemos visto que todos los conjuntos mencionados son linealmente independientes y son generadores.

■ EJEMPLO 46 $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es una base de $\mathbf{R}^n$, y se llama **base estándar** de $\mathbf{R}^n$. □

[4] Véase el teorema 2 de la sección 7, capítulo 2 en *Algebra I*, de N. Bourbaki (París: Hermann/Addison-Wesley, 1974).

■ EJEMPLO 47 $\{1, x, x^2, \ldots, x^n\}$ es una base de P_n, y se denomina **base estándar** de P_n (figura 4.7).

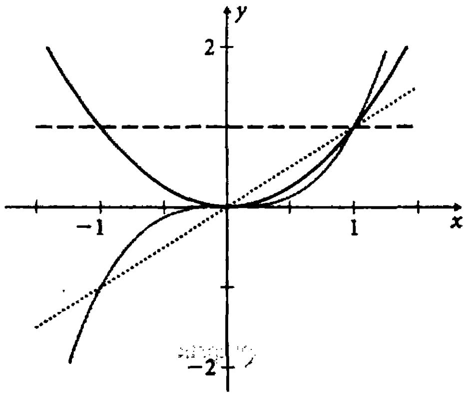


Figura 4.7 La base estándar $\{1, x, x^2, x^3\}$ de P_3.

■ EJEMPLO 48 $\{1, x, x^2, \ldots, x^n, \ldots\}$ es una base de P y se llama **base estándar** de P.

■ EJEMPLO 49 $\{E_{11}, E_{12}, E_{13}, \ldots, E_{mn}\}$ es una base de M_{mn} y se llama **base estándar** de M_{mn}.

Además de las bases estándar, $\mathbf{R}^n$, P_n, P y M_{mn} tienen muchas otras bases.

■ EJEMPLO 50 Demostrar que $\mathcal{B} = \{1 + x, -1 + x, x^2\}$ es una base de P_2.

SOLUCIÓN Necesitamos comprobar que $\mathcal{B}$ es linealmente independiente y que genera a P_2. En otras palabras, se necesita demostrar que

$$a(1 + x) + b(-1 + x) + cx^2 = \mathbf{0} \quad \text{o sea} \quad (a - b) + (a + b)x + cx^2 = \mathbf{0}$$

implica que $a = b = c = 0$, y que si $A + Bx + Cx^2$ es cualquier polinomio en P_2, entonces existen escalares a, b, c tales que

$$a(1 + x) + b(-1 + x) + cx^2 = A + Bx + Cx^2 \quad \text{o sea} \quad (a - b) + (a + b)x + cx^2 = A + Bx + Cx^2$$

Los dos sistemas resultantes

$$\begin{aligned} a - b &= 0 \\ a + b &= 0 \\ c &= 0 \end{aligned} \qquad \begin{aligned} a - b &= A \\ a + b &= B \\ c &= C \end{aligned}$$

tienen matrices invertibles de coeficientes. Por consiguiente, el primer sistema sólo tiene la solución trivial, y el segundo es consistente para toda elección de A, B y C. ☐

■ EJEMPLO 51 Demuestre que ni $X = \{1, 1 + x, -1 + x\}$ ni $Y = \{1 + x, -1 + x\}$ constituyen una base de P_2.

SOLUCIÓN X no es linealmente independiente porque $-2 \cdot 1 + (1 + x) - (-1 + x) = 0$ para toda x en $\mathbf{R}$. Por tanto,

$$-2 \cdot 1 + (1 + x) - (-1 + x) = \mathbf{0}$$

es una combinación lineal no trivial de $\mathbf{0}$, es decir, el polinomio cero.

Por otro lado, Y es linealmente independiente, pero no genera a P_2 (¿Por qué?) Aunque Y no es una base de P_2, sí lo es de P_1. (Compruébelo.) ☐

■ EJEMPLO 52 Para el conjunto $W = \{a \cos x + b \operatorname{sen} x, a, b \in \mathbf{R}\}$, demuestre que:

1. W es un espacio vectorial.
2. $\{\cos x, \operatorname{sen} x\}$ es una base de W.

SOLUCIÓN W es el conjunto de todas las combinaciones lineales de $\{\cos x, \operatorname{sen} x\}$ en $F(\mathbf{R})$; en consecuencia, es un subespacio de $F(\mathbf{R})$, de acuerdo con el teorema 5. En particular, W es un espacio vectorial generado por el conjunto $\{\cos x, \operatorname{sen} x\}$; por definición, y como se mostró en el ejemplo 45, éste es linealmente independiente en $F(\mathbf{R})$. Así, éste es linealmente independiente en el pequeño espacio W. Por consiguiente, forma una base de W.

Los dos vectores base, $\cos x$, $\operatorname{sen} x$, de W, junto con los vectores $2 \operatorname{sen} x + 3 \cos x$ y $3 \cos x - 2 \operatorname{sen} x$, se ilustran en la figura 4.8. ☐

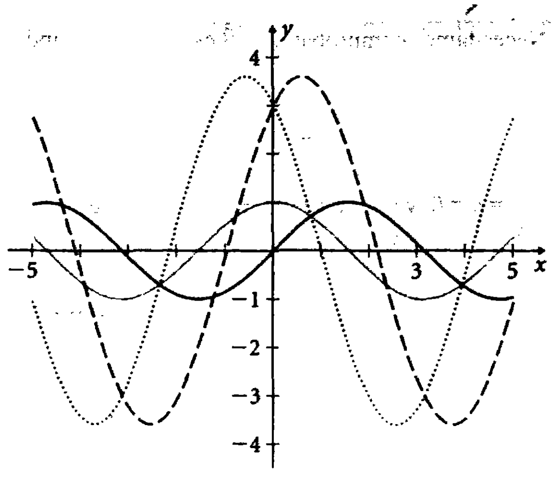


Figura 4.8 Algunos vectores de Gen({cos x, sen x}).

Una de las principales caracterizaciones de una base se describe en el teorema siguiente.

TEOREMA 11

(Unicidad de representación)

Un subconjunto $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de un espacio vectorial V es una base de V si y sólo si para cada vector $\mathbf{v}$ en V hay escalares $c_1, \ldots, c_n$ **únicos**, tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

DEMOSTRACIÓN Revise la comprobación del teorema 2 de la sección 4.1. $\square$

Ejercicios 4.3

Generador

Sean p_1, p_2, p_3 y p_4 en P_2, donde

$$p_1 = 1 + 3x - x^2 \qquad p_2 = -3x + 2x^2$$
$$p_3 = x^2 \qquad p_4 = 4 - x^2$$

1. a. ¿Está p_4 en $\text{Gen}\{p_1, p_2, p_3\}$?
   b. ¿Está p_4 en $\text{Gen}\{p_1, p_2\}$?
   c. ¿Está p_4 en $\text{Gen}\{p_2, p_3\}$?

2. a. Demuestre que $\text{Gen}\{p_1, p_2, p_3\} = P_2$.
   b. Compruebe que $\text{Gen}\{p_1, p_2, p_4\} = P_2$.
   c. Demuestre que

$$\text{Gen}\{p_1, p_3, p_4\} = \text{Gen}\{p_2, p_3, p_4\}$$

En los ejercicios 3 a 7 determine si el conjunto dado genera a P_2.

3. $\{1 + x + x^2,\ 1 + 2x + x^2,\ x\}$
4. $\{1 - x + x^2,\ 1 + x - x^2,\ 1\}$
5. $\{1 + x, -1 + x, 2 + x + x^2\}$
6. $\{1 + x + x^2,\ 1 + x,\ 1\}$
7. $\{-1 + x + x^2,\ 1 - x + x^2,\ 1 + x - x^2\}$
8. ¿Genera $\{4, 1 + x, -1 + x^2, -1 + x^3\}$ a P_3?

En los ejercicios 9 a 13 determine si el conjunto dado genera a M_{22}.

9. $\left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 4 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 5 \\ 0 & 5 \end{bmatrix}\right\}$

10. $\left\{\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 3 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 4 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}\right\}$

11. $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}\right\}$

12. $\{E_{11}, E_{12}, -E_{21}, E_{11} + E_{12}\}$
13. $\{E_{11}, E_{11} + E_{12}, E_{11} + E_{12} + E_{21}, E_{11} + E_{12} + E_{21} + E_{22}\}$
14. ¿Bajo qué restricción o restricciones para a y b el conjunto $\{a + ax + ax^2, bx^2, 1\}$ genera a P_2?
15. ¿Cierto o falso?
    a. P_9 puede ser generado exactamente por 9 polinomios en él.
    b. P_9 puede ser generado al menos por 9 polinomios en él.
    c. P_9 puede ser generado por 10 polinomios de P_8.
    d. P_9 puede ser generado por 10 polinomios en él.
    e. P_9 puede ser generado por 11 polinomios en él.
    f. 20 polinomios cualesquiera de P_9 generan a P_9.
    g. Veinte polinomios de P_9 pueden generan a P_9.
16. Demuestre que para cualesquiera vectores $\mathbf{u}$, $\mathbf{v}$ de un espacio vectorial V,

$$\text{Gen}\{\mathbf{u}, \mathbf{v}\} = \text{Gen}\{\mathbf{u} + \mathbf{v}, \mathbf{u} - \mathbf{v}\}$$

17. Compruebe que para cualesquiera vectores $\mathbf{u}$, $\mathbf{v}$, de un espacio vectorial V,

$$\text{Gen}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \text{Gen}\{\mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} + \mathbf{w}\}$$

Independencia lineal

En los ejercicios 18 a 21 determine si el conjunto es linealmente independiente.

18. $\{-2x + x^2, 1 + x + x^2, 1 - x\}$
19. $\{1 + ax + ax^2, 1 + bx + bx^2, 1\}$
20. $\{1 + ax + ax^2, 1 + bx + bx^2, x^2\}$, cuando a y b son constantes desiguales.

21. $\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}1&1\\1&0\end{bmatrix},\begin{bmatrix}2&1\\0&0\end{bmatrix},\begin{bmatrix}3&0\\-1&1\end{bmatrix}\right\}$

22. ¿Para qué valores de a el conjunto $\{1+ax, a+(a+2)x\} \subseteq P_1$ es linealmente dependiente?

23. ¿Cierto o falso?

   a. Dos vectores distintos cualesquiera de un espacio vectorial son linealmente independientes.

   b. Tres polinomios distintos cualesquiera de P_1 generan a P_1.

   c. Tres polinomios linealmente independientes cualesquiera de P_2 generan a P_2.

   d. Cualesquiera n polinomios linealmente independientes de P_n generan a P_n.

24. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ un conjunto de vectores linealmente independientes de un espacio vectorial V. Determine c_1, c_2 y c_3, para que $\mathbf{v}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+c_3\mathbf{v}_3$ y $\mathbf{v}=(2c_2-c_1)\mathbf{v}_1+(c_3-c_2)\mathbf{v}_2+(c_2-1)\mathbf{v}_3$.

25. Si V es un espacio vectorial, y $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ están en V, demuestre que si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente, también lo es $\{\mathbf{v}_1-\mathbf{v}_2, \mathbf{v}_2-\mathbf{v}_3, \mathbf{v}_3+\mathbf{v}_1\}$.

26. Compruebe el teorema 8.

27. Sean p, q y r polinomios P_2. Suponga que $\{p, q\}$ y $\{q, r\}$ son conjuntos linealmente independientes. ¿Implica lo anterior que $\{p, r\}$ es linealmente independiente? Explique por qué.

Bases

En los ejercicios 28 a 30 determine si los conjuntos dados son bases de M_{22}.

28. $\left\{\begin{bmatrix}1&2\\1&2\end{bmatrix},\begin{bmatrix}3&4\\3&4\end{bmatrix},\begin{bmatrix}5&6\\5&6\end{bmatrix},\begin{bmatrix}7&8\\7&8\end{bmatrix}\right\}$

29. $\left\{\begin{bmatrix}1&0\\0&1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix},\begin{bmatrix}1&0\\1&0\end{bmatrix},\begin{bmatrix}0&1\\0&1\end{bmatrix}\right\}$

30. $\left\{\begin{bmatrix}1&2\\3&4\end{bmatrix},\begin{bmatrix}2&2\\3&4\end{bmatrix},\begin{bmatrix}3&3\\3&4\end{bmatrix},\begin{bmatrix}4&4\\4&4\end{bmatrix}\right\}$

31. Sea $V \subseteq M_{22}$ el conjunto de todas las matrices de la forma $\begin{bmatrix}a&b\\c&-a\end{bmatrix}$. Demuestre que $\mathcal{B}=\{E_{11}-E_{22}, E_{12}, E_{21}\}$ es una base de V.

32. Demuestre que $\{x^2, (1+x)^2, (-1+x)^2\}$ es una base de P_2.

33. Compruebe que $\{1, x, 2x^2, 3-3x+x^3\}$ es una base de P_3. (Éstos son los primeros cuatro **polinomios de Chebyshev del primer tipo**. Se presentan en forma natural en varias áreas de matemáticas y física.)

En los ejercicios 34 a 39:

(a) Determine una base para V.

(b) En cada caso, verifique si $V = P_2$.

34. $V = \text{Gen}\,\{1+x+x^2, 1, -1-x^2, x^2\}$
35. $V = \text{Gen}\,\{2+x+2x^2, x^2, 1-x-x^2, 1\}$
36. $V = \text{Gen}\,\{x+x^2, 1+x, -1+x^2\}$
37. $V = \text{Gen}\,\{x^2, 1+x, -1+x^2\}$
38. $V = \text{Gen}\,\{1-x-5x^2, 7+x+4x^2, 8-x^2\}$
39. $V = \text{Gen}\,\{-x+x^2, -5+x, -x^2, 3+x^2\}$

En los ejercicios 40 a 45 amplíe el conjunto linealmente independiente dado para formar una base de P_2.

40. $\{1+x+x^2, 1\}$
41. $\{-x+x^2, x+x^2\}$
42. $\{x+x^2, 1+x\}$
43. $\{1+x, -1+x^2\}$
44. $\{1-x+x^2, 2-x^2\}$
45. $\{-x+x^2, -5+x\}$

4.4 Dimensión

Objetivo del estudiante para esta sección

Comprender el concepto básico de la dimensión de un espacio vectorial.

Hasta ahora hemos hablado libremente de vectores en 2 o 3 dimensiones, con lo que en realidad nos referimos a vectores 2 o vectores 3. En esta sección precisaremos la noción intuitiva de dimensión. Aprovecharemos este concepto para clasificar a todos los subespacios de $\mathbf{R}^2$ y

$\mathbf{R}^3$. Para quien esté orientado hacia las aplicaciones, lo que sigue le parecerá bastante teórico. Sin embargo, nos ayudará a comprender los fundamentos que hemos manejado, y también el importante concepto del *rango*, presentado en la sección 4.6.

El teorema siguiente es fundamental para demostrar que la dimensión de un espacio vectorial es un número bien definido, el cual se debe a **Steinitz**, y su demostración se presentará al final de la sección.

TEOREMA 12

(Teorema del intercambio)

Si un espacio vectorial V es generado por n vectores, entonces cualquier subconjunto que éste contenga con más de n vectores es linealmente dependiente. En otras palabras, todo subconjunto linealmente independiente de V tiene cuando mucho n vectores.

En consecuencia, se obtiene el teorema 13.

TEOREMA 13

Si un espacio vectorial V tiene una base con n elementos, entonces toda base de V tiene n elementos (figura 4.9).

DEMOSTRACIÓN Sea B una base con n vectores, y sea $\mathcal{B}'$ otra base. Si $\mathcal{B}'$ tuviera además de n otros elementos, sería un conjunto linealmente dependiente, según el teorema 12, porque $\mathcal{B}$ es un conjunto generador. En consecuencia, $\mathcal{B}'$ es finito y si m es su cantidad de elementos, entonces $m \leq n$. De acuerdo con el mismo argumento, si $\mathcal{B}$ y $\mathcal{B}'$ se intercambian, se ve que $n \leq m$. Por tanto, $n = m$. $\square$

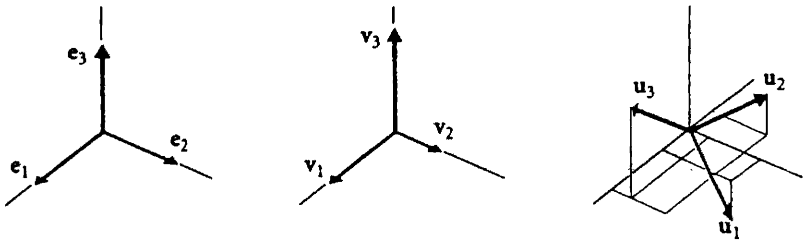


Figura 4.9 Todas las bases en un espacio vectorial tienen la misma cantidad de vectores.

DEFINICIÓN

(Dimensión)

Si un espacio vectorial V tiene una base con n elementos, entonces se dice que V es **dimensional finito,** y n es la **dimensión** de V. Se expresa

$$\dim(V) = n$$

De acuerdo con el teorema 13, *la dimensión es un número bien definido y no depende de la elección de la base*. La dimensión del espacio cero $\{\mathbf{0}\}$, *se define* como cero. Por consiguiente, $\{\mathbf{0}\}$ es dimensional finito. Un espacio vectorial que no tenga una base finita se llama **dimensional infinito**.

Un espacio vectorial que contiene un conjunto infinito linealmente independiente es dimensional infinito, porque si tuviera una base finita sería un conjunto generador finito. Por consiguiente, de acuerdo con el teorema 12, cualquier conjunto con más elementos sería linealmente dependiente.

Al contar la cantidad de elementos de las bases estándar se llega a la conclusión que $\mathbf{R}^n$, P_n y M_{mn} son dimensionales finitos. Además,

1. $\dim(\mathbf{R}^n) = n$ (figura 4.10).
2. $\dim(P_n) = n + 1$.
3. $\dim(M_{mn}) = m \cdot n$.
4. P es dimensional *infinito.*
5. $F(\mathbf{R})$ es dimensional *infinito.*

Las afirmaciones 4 y 5 son válidas, porque P y $F(\mathbf{R})$ contienen al conjunto infinito y linealmente independiente $\{1, x, x^2, \ldots, x^n, \ldots\}$.

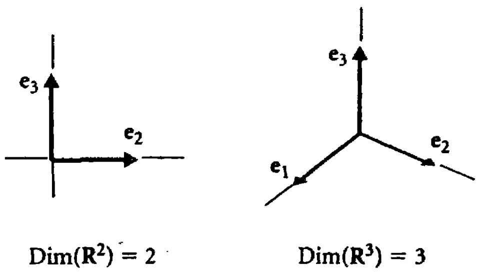


Figura 4.10 Las dimensiones de $\mathbf{R}^2$ y $\mathbf{R}^3$.

Observe que, como un subespacio de un espacio vectorial es en sí un espacio vectorial, tiene sentido de hablar de la *dimensión de un subespacio.*

■ EJEMPLO 53 Determine la dimensión del subespacio $V = \{(2x + y, x, -x - 2y, x + y + z), x, y, z \in \mathbf{R}\}$ de $\mathbf{R}^4$.

SOLUCIÓN En el ejemplo 14 de la sección 4.1, se determinó una base de V con 3 elementos. Por consiguiente, $\dim(V) = 3$. $\square$

■ EJEMPLO 54 Encuentre la dimensión del subespacio $V = \mathrm{Gen}\{(1, 1, 1), (2, 1, -1)\ (1, 0, -2)\}$ de $\mathbf{R}^3$.

SOLUCIÓN Debido a que $(2, 1, -1) - (1, 1, 1) = (1, 0, -2)$, entonces $V = \mathrm{Gen}\{(1, 1, 1), (2, 1, -1)\}$, según el teorema 6 de la sección 4.3. Por tanto, $\{(1, 1, 1), (2, 1, -1)\}$ generan a V; como es linealmente independiente (compruébelo), es una base. Así, $\dim(V) = 2$. La figura 4.11 muestra las dimensiones posibles del generador de dos vectores. $\square$

El teorema siguiente dice que cualquier conjunto linealmente independiente no posee más elementos que la dimensión, y que un conjunto generador no contiene menos elementos que la dimensión.

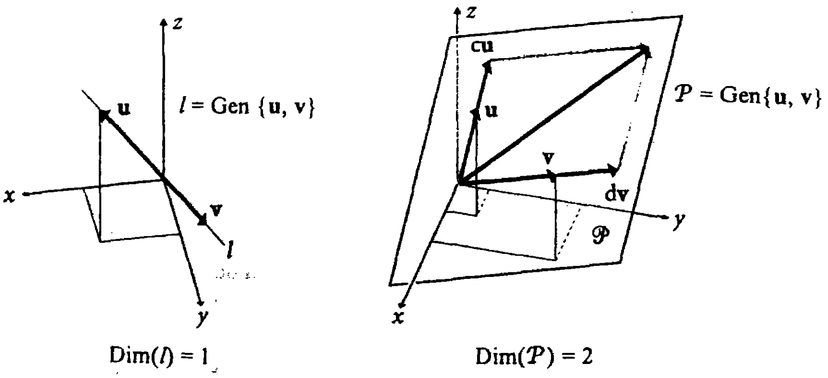


Figura 4.11 Dimensiones de los generadores.

TEOREMA 14

> Sea V un espacio vectorial de n dimensiones, y S un conjunto con m elementos.
>
> 1. Si S es linealmente independiente, entonces $m \leq n$.
> 2. Si S genera a V, entonces $m \geq n$.

DEMOSTRACIÓN Sea $\mathcal{B}$ una base de V. Como dim(V) = n, $\mathcal{B}$ es un conjunto linealmente independiente, generador, con n elementos. De acuerdo con el teorema 12, cualquier conjunto independiente tiene cuando mucho n elementos. Esto demuestra la parte 1. Por lo anterior, si un conjunto generador tiene menos de n elementos, entonces $\mathcal{B}$ es linealmente dependiente, lo cual es una contradicción. En consecuencia, un conjunto generador debe tener n o más elementos. Con esto se demuestra el enunciado 2. ▭

■ EJEMPLO 55 Sea S un conjunto de 10 vectores en $\mathbf{R}^k$. ¿Qué puede decirse acerca de k si S (a) es linealmente independiente? (b) ¿Genera a $\mathbf{R}^k$? (c) ¿Es una base de $\mathbf{R}^k$?

SOLUCIÓN Con base en el teorema 14, (a) $k \geq 10$, (b) $k \leq 10$, y (c) $k = 10$. ▭

El teorema siguiente dice que un conjunto con muchos elementos como dimensión del espacio vectorial al que genera *o bien* al que es linealmente independiente, es una base. Por tanto, no es necesario comprobar ambos requisitos. Esto reduce el trabajo (con frecuencia es fácil demostrar que un conjunto es linealmente independiente). El costo es que se necesita conocer de antemano la dimensión del espacio.

TEOREMA 15

> Sea V un espacio vectorial n dimensional, y sea S un conjunto con n elementos.
>
> 1. Si S es linealmente independiente, entonces S es una base.
> 2. Si S genera a V, entonces S es una base.

DEMOSTRACIÓN

1. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un subconjunto linealmente independiente de V. Si S no genera a V, hay un elemento $\mathbf{v}$ de V que no está en Gen(S). En ese caso el conjunto $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}\}$ sería linealmente independiente, según la afirmación 2 del teorema 9 de la sección 4.3. Esto contradice el enunciado 1 del teorema 14, porque S' tiene $n + 1 > n$ elementos. Por consiguiente, S genera a V, como se afirmó, y como es linealmente independiente, es una base.

2. S es un conjunto generador con n elementos. Si S es linealmente dependiente, entonces un elemento es la combinación lineal del resto, de acuerdo con el teorema 7 de la sección 4.3. Si eliminamos este elemento, el conjunto resultante y S tendrían el mismo generador, basándonos el teorema 6 en la sección 4.3. Pero entonces V estaría generado por menos de n elementos, lo cual contradice el enunciado 2 del teorema 14. Llegamos a la conclusión que S es linealmente independiente. Así, S es una base, porque genera a V. $\square$

■ EJEMPLO 56 Demuestre que $S = \{(1, -1), (0, 1)\}$ es una base para $\mathbf{R}^2$.

SOLUCIÓN S es linealmente independiente (¿por qué?) y tiene exactamente dos elementos. Como la dimensión de $\mathbf{R}^2$ también es dos, S es una base según el teorema 15. $\square$

El teorema siguiente puede ser muy útil en la práctica. En él se afirma que podemos obtener una base agregando elementos a un conjunto linealmente independiente, o quitándolos de un conjunto generador, en ambos casos debe hacerse en forma adecuada.

TEOREMA 16

Sea V un espacio vectorial n dimensional, y S un conjunto con m elementos.

1. Si S es linealmente independiente y $m < n$, entonces S se puede ampliarse a una base.
2. Si S envuelve a V, entonces S contiene una base.

DEMOSTRACIÓN

1. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ un subconjunto linealmente independiente de V, con $m < n$. Según el teorema 13, S no puede generar a V. Por tanto, hay un elemento $\mathbf{v}_{m+1}$ que no está en el generador de S. Así, el conjunto $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}\}$ es linealmente independiente, de acuerdo con el enunciado 2 del teorema 9 de la sección 4.3. Ahora repetimos este proceso con S' en lugar de S y continuamos hasta que el conjunto final tenga n elementos linealmente independientes. Según el teorema 15, este conjunto sería una base y contiene a S. Por consiguiente, S puede ampliarse a una base.

2. El enunciado 2 del teorema 14 afirma que $m \geq n$, porque S genera a V. Si $m = n$, entonces S es una base, de acuerdo con el teorema 15. Si $m > n$, S es linealmente dependiente, según el teorema 14. Si S' es el conjunto que se obtiene eliminando un elemento de S, y el cual es una combinación lineal de los elementos restantes, entonces S' tiene $m - 1$ elementos y sigue generando a S, basándonos en el teorema 6 de la sección 4.3. Este proceso se repite con S'. La eliminación subsecuente de elementos de S debe hacerse de manera que los que aún continúan en el conjunto sigan generando a V. Este proceso termina cuando se llega a un subconjunto generador con la cantidad mínima de elementos, es decir, n elementos exactamente, en concordancia con el teorema 14. Este conjunto final es una base, según el teorema 15. $\square$

■ EJEMPLO 57 Amplíe el conjunto linealmente independiente $S = \{-1 + x^2, 3 - 2x\}$ hasta que sea una base de P_3.

SOLUCIÓN Primero ampliamos S hasta S', el cual generará a P_3, agregando la base estándar de P_3.

$$S' = \{-1 + x^2, 3 - 2x, 1, x, x^2, x^3\}$$

A partir de lo que se afirma en el teorema 7 de la sección 4.3, S' es linealmente dependiente, porque la base estándar genera a P_3. En concordancia con el mismo teorema, un elemento es una combinación lineal del anterior. Como S es linealmente independiente, se comienza con 1. Pero 1 no es una combinación lineal en S. (¿Por qué?) Sin embargo, tanto x como x^2 son combinaciones lineales de $-1 + x^2$, $3 - 2x$, 1, de modo que los eliminamos de S'. Por último, x^3 no es combinación lineal de $-1 + x^2$, $3 - 2x$, 1 y se conservan. De manera que, $\{-1 + x^2, 3 - 2x, 1, x^3\}$ es linealmente independiente y sigue generando a P_3. Por consiguiente, es una base que contiene a S. □

El teorema que expondremos a continuación afirma que la dimensión de un subespacio no puede ser mayor que la dimensión del espacio vectorial.

TEOREMA 17

Sea W un subespacio de un espacio vectorial V de n dimensiones. Entonces

1. $\dim(W) \leq n$;
2. $\dim(W) = n$ si y sólo si $W = V$.

DEMOSTRACIÓN

1. Como cualquier base de W está en V y es linealmente independiente, tiene cuando mucho n elementos, de acuerdo con el teorema 12. Por consiguiente, $\dim(W) \leq n$.

2. Sea $\dim(W) = n$. Entonces, cualquier base $\mathcal{B}$ de W tiene n elementos linealmente independientes. Por tanto, $\mathcal{B}$ es una base de V, según el teorema 15. Así, $\text{Gen}(\mathcal{B}) = W$. En consecuencia, lo inverso es trivial. □

Los subespacios de $\mathbf{R}^2$ y $\mathbf{R}^3$

El teorema 17 ayuda a clasificar a todos los subespacios de $\mathbf{R}^2$ y $\mathbf{R}^3$.

■ EJEMPLO 58 Determine todos los subespacios de $\mathbf{R}^2$.

SOLUCIÓN Con referencia al teorema 17, un subespacio puede ser 0, 1 o 2 dimensional. El subespacio cero es el único subespacio 0 dimensional. También de acuerdo con dicho teorema, $\mathbf{R}^2$ es el único subespacio bidimensional. Así, sólo necesitamos identificar los subespacios unidimensionales. Si V es un subespacio unidimensional, y $\{\mathbf{w}\}$ es una base de V, entonces $V = \text{Gen}(\{\mathbf{w}\}) = \{r\,\mathbf{w}, r \in \mathbf{R}\}$. Por tanto, V es el conjunto de todos los múltiplos escalares de $\mathbf{w}$, que es una recta que pasa por el origen en la dirección de $\mathbf{w}$. A la inversa, cualquier recta que atraviesa el origen es un subespacio, porque es el generador de cualquier vector distinto de cero en la recta. Hemos demostrado que los subespacios unidimensionales incluyen a las rectas que cruzan el origen.

Para resumir, los subespacios de $\mathbf{R}^2$ son

- Subespacios cero dimensionales: $\{\mathbf{0}\}$.
- Subespacios unidimensionales: todas las rectas que pasan por el origen.
- Subespacios bidimensionales: $\mathbf{R}^2$. $\square$

■ EJEMPLO 59 Determine todos los subespacios de $\mathbf{R}^3$.

SOLUCIÓN Tenemos la clasificación siguiente. Los detalles se dejan como ejercicios.

Los subespacios de $\mathbf{R}^3$ son (figura 4.12)

- Subespacios cero dimensionales: $\{\mathbf{0}\}$.
- Subespacios unidimensionales: todas las rectas que pasan por el origen.
- Subespacios bidimensionales: todos los planos que pasan por el origen.
- Subespacios tridimensionales: $\mathbf{R}^3$. $\square$

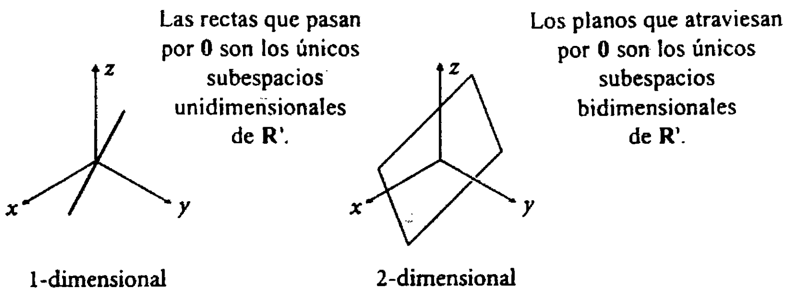


Figura 4.12 Los subespacios uni y bidimensional de $\mathbf{R}^3$.

Dos demostraciones del teorema de intercambio

DEMOSTRACIÓN 1 Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ el conjunto generador, y sea T cualquier conjunto con al menos $n + 1$ vectores. Comprobaremos que T es linealmente dependiente. T' es un subconjunto de T formado exactamente por $n + 1$ vectores, por ejemplo $T' = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}\}$. Si demostramos que T' es linealmente dependiente habremos terminado, de acuerdo con el enunciado 3 del teorema 8, sección 4.3.

Como S genera a V, todos los elementos de T' pueden expresarse como combinaciones lineales en S. Por consiguiente, existen escalares a_{ij} tales que

$$\mathbf{u}_1 = a_{11}\mathbf{v}_1 + \cdots + a_{1n}\mathbf{v}_n$$

$$\mathbf{u}_2 = a_{21}\mathbf{v}_1 + \cdots + a_{2n}\mathbf{v}_n$$

$$\mathbf{u}_{n+1} = a_{n+1,1}\mathbf{v}_1 + \cdots + a_{n+1,n}\mathbf{v}_n$$

Veamos la matriz A' $n \times (n+1)$ cuyo elemento (i, j) es a_{ji} $(1 \leq i \leq n,\ 1 \leq j \leq n+1)$. El sistema homogéneo $[A' : \mathbf{0}]$ tiene soluciones no triviales, porque contiene más incógnitas que ecuaciones ($n + 1$ en comparación con n), según el teorema 5 de la sección 1.2. Podemos inferir entonces que existe un vector $(n + 1)$ distinto de cero y que $\mathbf{c} = (c_1, \ldots, c_{n+1})$, lo cual resuelve el sistema. Ahora consideraremos la suma

$$c_1\mathbf{u}_1 + \cdots + c_{n+1}\mathbf{u}_{n+1} = c_1(a_{11}\mathbf{v}_1 + \cdots + a_{1n}\mathbf{v}_n) + c_2(a_{21}\mathbf{v}_1 + \cdots + a_{2n}\mathbf{v}_n)$$
$$+ \cdots + c_{n+1}(a_{n+1,1}\mathbf{v}_1 + \cdots + a_{n+1,n}\mathbf{v}_n)$$

Si multiplicamos y agrupamos los términos, veremos que el coeficiente de cada $\mathbf{v}_i$ es $c_1a_{1i} + \cdots + c_{n+1}a_{n+1,i}$, que es cero, porque $\mathbf{c}$ es una solución del sistema. Entonces, $c_1\mathbf{u}_1 + \cdots + c_{n+1}\mathbf{u}_{n+1} = \mathbf{0}$. Ésta es una combinación lineal no trivial del vector cero. Por consiguiente, como se afirmó T' es linealmente dependiente.

DEMOSTRACIÓN 2 Si $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es el conjunto generador y $T = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ un conjunto linealmente independiente, bastará con demostrar que $m \leq n$. El conjunto

$$S' = \{\mathbf{u}_m, \mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

es linealmente dependiente, de acuerdo con el teorema 7 de la sección 4.3, porque S es un conjunto generador. En consecuencia, según el mismo teorema, una de las $\mathbf{v}$, digamos $\mathbf{v}_i$, es una combinación lineal de los vectores precedentes. Así, el conjunto S'' formado a partir de S' eliminando $\mathbf{v}_i$ sigue siendo un conjunto generador.

$$S'' = \{\mathbf{u}_m, \mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$$

Ahora agregamos $\mathbf{u}_{m-1}$ a S'' para obtener S''',

$$S''' = \{\mathbf{u}_{m-1}, \mathbf{u}_m, \mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n\}$$

y aplicamos el mismo argumento para demostrar que S''' es linealmente independiente y generador. Ninguna de las $\mathbf{u}$ es una combinación lineal de los vectores precedentes, porque las $\mathbf{u}$ son linealmente independientes; por esa causa se puede eliminar una de las $\mathbf{v}$ como antes. Si continuamos con este proceso finito observaremos que las $\mathbf{u}$ se agotan antes que las $\mathbf{v}$. De no ser así, las $\mathbf{u}$ restantes serían combinaciones lineales de las $\mathbf{u}$ ya incluidas en el conjunto. Esto es imposible, porque las $\mathbf{u}$ son linealmente independientes. En consecuencia, $m \leq n$, como se afirmó. $\square$

OBSERVACIÓN El nombre *teorema de intercambio* proviene de la demostración 2, en la que se intercambiaron los vectores generadores por otros linealmente independientes.

Ejercicios 4.4

En los ejercicios 1 a 8 determine la dimensión de V.

1. $V = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix},\ a \in \mathbf{R} \right\}$

2. $V = \left\{ \begin{bmatrix} a - b \\ 2a + b \end{bmatrix},\ a, b \in \mathbf{R} \right\}$

3. $V = \left\{ \begin{bmatrix} a \\ 0 \\ -2a \end{bmatrix},\ a \in \mathbf{R} \right\}$

4. $V = \left\{ \begin{bmatrix} a - c \\ b + c \\ 5c \end{bmatrix},\ a, b, c \in \mathbf{R} \right\}$

5. V es el conjunto de todos los vectores 3 cuyos primeros componentes son cero.

6. V es el conjunto de todos los vectores 4 cuyos primeros y últimos componentes son cero.

7. V es el conjunto de todos los vectores 4 cuyos tres primeros componentes son cero.

8. $V = \text{Gen}\left\{\begin{bmatrix}-5\\-1\\1\end{bmatrix}, \begin{bmatrix}4\\1\\7\end{bmatrix}, \begin{bmatrix}-1\\0\\8\end{bmatrix}\right\}$

En los ejercicios 9 a 11 determine la dimensión del generador de los conjuntos dados en M_{22}.

9. $\left\{\begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}, \begin{bmatrix}3 & 4\\3 & 4\end{bmatrix}, \begin{bmatrix}5 & 6\\5 & 6\end{bmatrix}\right\}$

10. $\left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\right\}$

11. $\left\{\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}, \begin{bmatrix}2 & 2\\3 & 4\end{bmatrix}, \begin{bmatrix}3 & 3\\3 & 4\end{bmatrix}\right\}$

12. Obtenga la dimensión del conjunto V de todas las matrices de la forma $\begin{bmatrix}a & b\\c & -a\end{bmatrix}$

En los ejercicios 13 a 18 determine la dimensión de $V \subseteq P_2$.

13. $V = \text{Gen}\{1 + x + x^2, 1, -1 - x^2, x^2\}$
14. $V = \text{Gen}\{2 + x + x^2, x^2, 1 - x - x^2, 1\}$
15. $V = \text{Gen}\{x + x^2, 1 + x, -1 + x^2\}$
16. $V = \text{Gen}\{x^2, 1 + x, -1 + x^2\}$
17. $V = \text{Gen}\{1 - x - 5x^2, 7 + x + 4x^2, 8 - x^2\}$
18. $V = \text{Gen}\{-x + x^2, -5 + x, -x^2, 3 + x^2\}$
19. ¿Cierto o falso?

   a. $\mathbf{R}^{10}$ tiene una base con 11 elementos.
   b. $\mathbf{R}^{10}$ tiene una base con sólo 10 elementos.
   c. $\mathbf{R}^{10}$ sólo tiene 10 elementos.
   d. $\mathbf{R}^{10}$ tiene un subespacio 9 dimensional.
   e. $\mathbf{R}^{10}$ sólo tiene un subespacio 9 dimensional.
   f. $\mathbf{R}^{10}$ es el único subespacio 10 dimensional de $\mathbf{R}^{10}$.
   g. $\mathbf{R}^2$ es un subespacio bidimensional de $\mathbf{R}^{10}$.

20. ¿Cierto o falso?

   a. Un subespacio V distinto de cero, de $\mathbf{R}^{10}$, puede tener dos bases distintas.
   b. Un subespacio V distinto de cero, de $\mathbf{R}^{10}$, puede tener dos bases con diferentes cantidades de elementos.
   c. Un subespacio V distinto de cero, de $\mathbf{R}^{10}$, puede tener una base con 10 elementos.
   d. Un subespacio V distinto de cero, de $\mathbf{R}^{10}$, sólo puede tener 10 elementos.
   e. Un subespacio V distinto de cero, de $\mathbf{R}^{10}$, puede tener una base con 11 elementos.
   f. Un subespacio V distinto de cero, de $\mathbf{R}^{10}$, puede tener una base con 9 elementos.
   g. La dimensión de un subespacio V de $\mathbf{R}^{10}$, es cero si y sólo si $V = \{\mathbf{0}\}$.

21. Obtenga la dimensión del Gen$\{e^x, e^{2x}, 2e^x\}$ en $F(R)$.
22. Determine la dimensión del Gen$\{\cos(x)\,\text{sen}(x), \text{sen}(2x)\}$ en $F(R)$.
23. Calcule la dimensión del Gen$\{\cos^2(x), \text{sen}^2(x), 1\}$ en $F(R)$.
24. Demuestre que para cualesquiera vectores $\mathbf{u}$, $\mathbf{v}$ de un espacio vectorial V,

$$\dim(\text{Gen}\{\mathbf{u}, \mathbf{v}\}) = \dim(\text{Gen}\{\mathbf{u} + \mathbf{v}, \mathbf{u} - \mathbf{v}\})$$

25. Describa geométricamente todos los subespacios de $\mathbf{R}^4$.

4.5 Vectores de coordenadas y cambio de base

Objetivos del estudiante para esta sección

1. Calcular el vector de coordenadas con respecto a una base determinada.
2. Obtener la matriz de transición y aplicarla para cambiar de una base a otra.

Muchos problemas de física e ingeniería pueden simplificarse considerablemente si se elige el sistema adecuado de coordenadas. Lo mismo puede hacerse con los problemas en espacios vectoriales, cuando se escoge la base adecuada. Primero estudiaremos las coordenadas de un vector general con respecto a una base fija. Después indicaremos cómo cambiar las coordenadas de una base anterior a una nueva. Como las coordenadas son números, muchos de los cálculos pueden llevarse a cabo con computadora.

Vectores de coordenadas

DEFINICIÓN

Sea V un espacio vectorial de dimensiones finitas con base $\mathcal{B} = \{\mathbf{v}_1 \ldots, \mathbf{v}_n\}$. Según el teorema 11 de la sección 4.3, para cada $\mathbf{v} \in V$, existen escalares únicos $c_1, \ldots, c_n$ tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

El vector cuyos componentes son los coeficientes de $\mathbf{v}$, expresado como $[\mathbf{v}]_{\mathcal{B}}$, se llama **(vector de coordenadas** o **vector coordenado) de v con respecto a** $\mathcal{B}$.

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$[\mathbf{v}]_{\mathcal{B}}$ se modifica cuando cambia la base $\mathcal{B}$ (figura 4.13). También $[\mathbf{v}]_{\mathcal{B}}$ depende del *orden* de los elementos de $\mathcal{B}$. Mantendremos fijo este orden usando siempre una **base ordenada**.

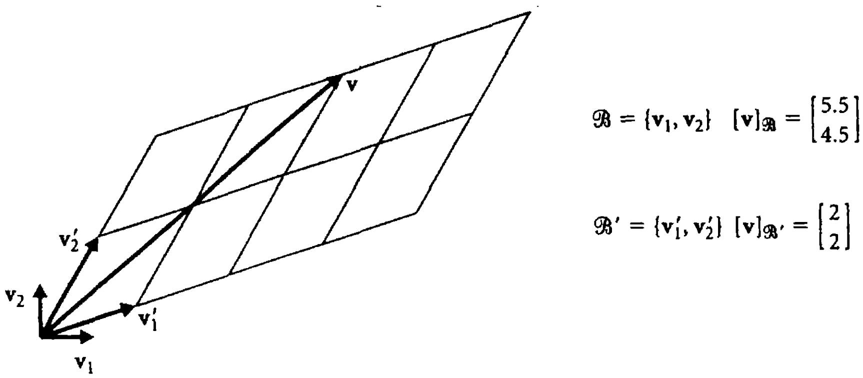


Figura 4.13 Coordenadas con respecto a diferentes bases.

OBSERVACIÓN Si $\mathbf{a} = (a_1, \ldots, a_n)$ es un vector n y $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ es la base estándar de $\mathbf{R}^n$, entonces

$$[\mathbf{a}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \mathbf{a}$$

porque $\mathbf{a} = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n$.

■ EJEMPLO 60 Se tiene la base $\mathcal{B} = \{(1, 0, -1), (-1, 1, 0), (1, 1, 1)\}$ de $\mathbf{R}^3$ y el vector $\mathbf{v} = (2, -3, 4)$.

(a) Determine $[\mathbf{v}]_{\mathcal{B}}$.

(b) Calcule el vector $\mathbf{w}$ si $[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} 6 \\ -3 \\ 2 \end{bmatrix}$.

SOLUCIÓN (a) $[\mathbf{v}]_{\mathcal{B}}$ tiene como componentes a los escalares c_1, c_2, c_3 tales que

$$(2, -3, 4) = c_1(1, 0, -1) + c_2(-1, 1, 0) + c_3(1, 1, 1)$$

lo cual implica que $c_1 = -3$, $c_2 = -4$, $c_3 = 1$. De aquí,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ -4 \\ 1 \end{bmatrix}$$

(b) Debido a que los componentes de $[\mathbf{w}]_{\mathcal{B}}$ son 6, –3, 2, $\mathbf{w}$ se expresa como

$$\mathbf{w} = 6(1, 0, -1) - 3(-1, 1, 0) + 2(1, 1, 1) = (11, -1, -4)$$

■ EJEMPLO 61 Determine el vector coordenado de $\mathbf{v} = (a, b, c)$ en $\mathbf{R}^3$ con respecto a $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, donde

$$\mathbf{v}_1 = \mathbf{e}_3, \qquad \mathbf{v}_2 = \mathbf{e}_1, \qquad \mathbf{v}_3 = \mathbf{e}_2$$

SOLUCIÓN Los componentes de $[\mathbf{v}]_{\mathcal{B}'}$ son escalares c_1, c_2, c_3 tales que

$$(a, b, c) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = c_1\mathbf{e}_3 + c_2\mathbf{e}_1 + c_3\mathbf{e}_2 = (c_2, c_3, c_1)$$

Por consiguiente, $c_1 = c$, $c_2 = a$ y $c_3 = c$. Así,

$$[\mathbf{v}]_{\mathcal{B}'} = \begin{bmatrix} c \\ a \\ b \end{bmatrix}$$

■ EJEMPLO 62 Obtenga el vector de coordenadas de $p = 1 + 2x + 3x^2$ en P_2 con respecto a cada una de las bases siguientes:

(a) La base (estándar) $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, siendo

$$\mathbf{v}_1 = 1, \qquad \mathbf{v}_2 = x, \qquad \mathbf{v}_3 = x^2$$

(b) La base $\mathcal{B}' = \{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3'\}$, siendo

$$\mathbf{v}_1' = 1 + x, \quad \mathbf{v}_2' = 1 - x^2, \quad \mathbf{v}_3' = 1 + x + x^2$$

SOLUCIÓN (a) Ya que $p = 1 \cdot 1 + 2 \cdot x + 3 \cdot x^2$, entonces

$$[p]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

(b) Los componentes de $[p]_{\mathcal{B}'}$ son escalares c_1, c_2, c_3, tales que

$$p = c_1\mathbf{v}_1' + c_2\mathbf{v}_2' + c_3\mathbf{v}_3' = c_1(1 + x) + c_2(1 - x^2) + c_3(1 + x + x^2)$$
$$\Rightarrow 1 + 2x + 3x^2 = (c_1 + c_2 + c_3) + (c_1 + c_3)x + (-c_2 + c_3)x^2$$

Así, $1 = c_1 + c_2 + c_3$, $2 = c_1 + c_3$, $3 = -c_2 + c_3$ es decir $c_1 = 0$, $c_2 = -1$, $c_3 = 2$. Por consiguiente,

$$[p]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$$

■ EJEMPLO 63 Determine el vector de coordenadas de

$$A = \begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix}$$

en M_{22} con respecto a cada una de las bases siguientes:

(a) La base (estándar) $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, donde

$$\mathbf{v}_1 = E_{11}, \qquad \mathbf{v}_2 = E_{12}, \qquad \mathbf{v}_3 = E_{21}, \qquad \mathbf{v}_4 = E_{22}$$

(b) La base $\mathcal{B}' = \{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3', \mathbf{v}_4'\}$, en la que

$$\mathbf{v}_1' = -E_{21}, \qquad \mathbf{v}_2' = E_{22}, \qquad \mathbf{v}_3' = -E_{12}, \qquad \mathbf{v}_4' = E_{11}$$

SOLUCIÓN (a) Los componentes de $[A]_{\mathcal{B}}$ son escalares c_1, c_2, c_3 y c_4 tales que

$$A = \begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix} = c_1E_{11} + c_2E_{12} + c_3E_{21} + c_4E_{22} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$$

Por consiguiente, $c_1 = 2$, $c_2 = -3$, $c_3 = 1$ y $c_4 = 4$. Entonces

$$[A]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \\ 1 \\ 4 \end{bmatrix}$$

(b) Los componentes de $[A]_{\mathcal{B}'}$ son escalares c_1, c_2, c_3 y c_4 tales que

$$A = \begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix} = -c_1E_{21} + c_2E_{22} - c_3E_{12} + c_4E_{11} = \begin{bmatrix} c_4 & -c_3 \\ -c_1 & c_2 \end{bmatrix}$$

Por tanto, $c_1 = -1$, $c_2 = 4$, $c_3 = 3$ y $c_4 = 2$. En consecuencia,

$$[A]_{\mathcal{B}'} = \begin{bmatrix} -1 \\ 4 \\ 3 \\ 2 \end{bmatrix}$$

OBSERVACIÓN Si $\mathcal{B} = \{4\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de un espacio vectorial V de dimensión finita, en ese caso

$$[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i, \qquad i = 1, 2, \ldots, n$$

TEOREMA 18

Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base de un espacio vectorial V de dimensión finita. Sean $\mathbf{u}, \mathbf{u}_1, \ldots, \mathbf{u}_m$ vectores en V. Entonces $\mathbf{u}$ es una combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_m$ en V, si y sólo si $[\mathbf{u}]_{\mathcal{B}}$ es una combinación lineal de $[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}$ en $\mathbf{R}^n$. Además, para los escalares $c_1, \ldots, c_m$,

$$\mathbf{u} = c_1\,\mathbf{u}_1 + \cdots + c_m\,\mathbf{u}_m \tag{4.3}$$

si y sólo si

$$[\mathbf{u}]_{\mathcal{B}} = c_1\,[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_m\,[\mathbf{u}_m]_{\mathcal{B}} \tag{4.4}$$

DEMOSTRACIÓN Sean

$$[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \quad \text{y} \quad [\mathbf{u}_i]_{\mathcal{B}} = \begin{bmatrix} u_{i1} \\ \vdots \\ u_{in} \end{bmatrix}, \qquad i = 1, 2, \ldots, m$$

Suponiendo la ecuación (4.3), entonces

$$\begin{aligned}\mathbf{u} &= c_1(u_{11}\mathbf{v}_1 + \cdots + u_{1n}\mathbf{v}_n) + \cdots + c_m(u_{m1}\mathbf{v}_1 + \cdots + u_{mn}\mathbf{v}_n) \\ &= (c_1u_{11} + \cdots + c_mu_{m1})\mathbf{v}_1 + \cdots + (c_1u_{1n} + \cdots + c_mu_{mn})\mathbf{v}_n\end{aligned}$$

Por consiguiente,

$$\begin{aligned}[\mathbf{u}]_{\mathcal{B}} &= \begin{bmatrix} c_1u_{11} + \cdots + c_mu_{m1} \\ \vdots \\ c_1u_{1n} + \cdots + c_mu_{mn} \end{bmatrix} = c_1\begin{bmatrix} u_{11} \\ \vdots \\ u_{1n} \end{bmatrix} + \cdots + c_m\begin{bmatrix} u_{m1} \\ \vdots \\ u_{mn} \end{bmatrix} \\ &= c_1\,[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_m\,[\mathbf{u}_m]_{\mathcal{B}}\end{aligned}$$

con lo cual se demuestra la ecuación (4.4). Todos estos pasos se pueden invertir para completar la demostración. □

Suponiendo que $\mathbf{u} = \mathbf{0}$, el teorema 18 tiene el siguiente y útil corolario:

TEOREMA 19

Sea $\mathcal{B}$ una base de un espacio vectorial n dimensional V. Entonces $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ es linealmente independiente en V si y sólo si $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}\}$ es linealmente independiente en $\mathbf{R}^n$.

■ EJEMPLO 64 Demuestre que $p_1(x) = 1 - x^2, p_2(x) = -1 + x, p_3(x) = 1 + x + x^2$ son linealmente independientes en P_2.

SOLUCIÓN De acuerdo con el teorema 19, basta con demostrar la independencia de los vectores de coordenadas con respecto a la base estándar $\mathcal{B}$,

$$[p_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad [p_2]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \qquad [p_3]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

lo cual se comprueba con facilidad. □

Cambio de base

Sea **v** un vector en un espacio vectorial V de dimensiones finitas, y sean $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}' = \{\mathbf{v}_1', \ldots, \mathbf{v}_n'\}$ dos bases. A continuación definiremos una relación entre $[\mathbf{v}]_{\mathcal{B}}$ y $[\mathbf{v}]_{\mathcal{B}'}$.

Como $\mathcal{B}'$ es una base, los elementos de $\mathcal{B}$ son combinaciones lineales de los elementos de $\mathcal{B}'$. Entonces, hay escalares $a_{11}, a_{12}, \ldots, a_{nn}$ tales que

$$\mathbf{v}_i = a_{1i}\mathbf{v}_1' + \cdots + a_{ni}\mathbf{v}_n' \qquad i = 1, 2, \ldots, n \tag{4.5}$$

P es la matriz cuyo elemento (i, j) es a_{ij}.

Como $\mathcal{B}$ genera a V, hay escalares $c_1, \ldots, c_n$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. Por consiguiente,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

y según el teorema 18,

$$[\mathbf{v}]_{\mathcal{B}'} = c_1\,[\mathbf{v}_1]_{\mathcal{B}'} + \cdots + c_n\,[\mathbf{v}_n]_{\mathcal{B}'}$$

Empleando la ecuación (4.5), esto puede escribirse en la forma

$$\begin{aligned}
[\mathbf{v}]_{\mathcal{B}'} &= c_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + c_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + c_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix} \\
&= P \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = P\,[\mathbf{v}]_{\mathcal{B}}
\end{aligned}$$

Por consiguiente,

$$[\mathbf{v}]_{\mathcal{B}'} = P\,[\mathbf{v}]_{\mathcal{B}}$$

Así, $[\mathbf{v}]_{\mathcal{B}'}$ es el producto de la matriz P que tiene como columnas a los vectores coordenados de la base "anterior" $\mathcal{B}$, con respecto a la base "nueva" $\mathcal{B}'$ y $[\mathbf{v}]_{\mathcal{B}}$.

Además, podemos comprobar que la matriz P es *invertible* si se demuestra que el sistema

$$P\mathbf{x} = \mathbf{b}$$

tiene una solución para todo vector **b** en $\mathbf{R}^n$. En realidad, sea

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Considere al vector $\mathbf{v} = b_1\mathbf{v}_1' + \cdots + b_n\mathbf{v}_n'$. Entonces

$$\mathbf{b} = [\mathbf{v}]_{\mathcal{B}'} = P\,[\mathbf{v}]_{\mathcal{B}}$$

De aquí inferimos que $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$ es la solución de $P\mathbf{x} = \mathbf{b}$ para el $\mathbf{b}$ dado. Con esto demostramos que P es invertible.

Por último, mediante la propiedad $[\mathbf{v}]_{\mathcal{B}'} = P[\mathbf{v}]_{\mathcal{B}}$, podemos comprobar que P es la *única* matriz. Porque si P' es otra matriz, entonces $[\mathbf{v}]_{\mathcal{B}'} = P'[\mathbf{v}]_{\mathcal{B}}$. Igualando $\mathbf{v} = \mathbf{v}_i$, se obtiene

$$[\mathbf{v}_i]_{\mathcal{B}'} = P\,[\mathbf{v}_i]_{\mathcal{B}} = P\mathbf{e}_i \qquad \text{y} \qquad [\mathbf{v}_i]_{\mathcal{B}'} = P'\,[\mathbf{v}_i]_{\mathcal{B}} = P'\mathbf{e}_i$$

porque $[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i$ como vectores. De aquí que $P\mathbf{e}_i = P'\mathbf{e}_i$. En consecuencia, las i-ésimas columnas de P y de P' son iguales para cada $i = 1, 2, \ldots, n$. Por tanto, $P = P'$.

Hemos contestado totalmente la pregunta acerca de la relación entre $[\mathbf{v}]_{\mathcal{B}}$ y $[\mathbf{v}]_{\mathcal{B}'}$. A continuación haremos un resumen y aplicaremos los conceptos aprendidos.

TEOREMA 20

(Cambio de base)

Sean $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ dos bases de un espacio vectorial de dimensión finita. Sea P la matriz $n \times n$ cuyas columnas son $[\mathbf{v}_1]_{\mathcal{B}'}, \ldots, [\mathbf{v}_n]_{\mathcal{B}'}$.

$$P = [[\mathbf{v}_1]_{\mathcal{B}'}\ [\mathbf{v}_2]_{\mathcal{B}'} \cdots [\mathbf{v}_n]_{\mathcal{B}'}]$$

Entonces P es invertible y ésta es la única matriz en la que para todo $\mathbf{v} \in V$,

$$[\mathbf{v}]_{\mathcal{B}'} = P\,[\mathbf{v}]_{\mathcal{B}}$$

DEFINICIÓN

(Matriz de transición)

La matriz P del teorema 20 se denomina **matriz de transición** (o **matriz de cambio de base**) de $\mathcal{B}$ a $\mathcal{B}'$ (figura 4.14).

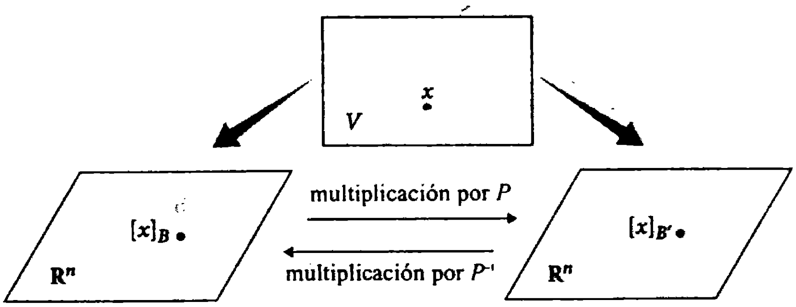


Figura 4.14 La acción de la matriz de transición y de su inversa.

COROLARIO 21

Si P es la matriz de transición de $\mathcal{B}$ a $\mathcal{B}'$, entonces P^{-1} es la matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$ (figura 4.14).

DEMOSTRACIÓN Según el teorema 20, P^{-1} existe y $[\mathbf{v}]_{\mathcal{B}'} = P[\mathbf{v}]_{\mathcal{B}}$ para todo $\mathbf{v} \in V$. De aquí que

$$[\mathbf{v}]_{\mathcal{B}} = P^{-1}[\mathbf{v}]_{\mathcal{B}'} \quad \text{para toda } \mathbf{v} \in V$$

Por consiguiente, de nuevo, el teorema 20 implica que P^{-1} es la (única) matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$. ☐

■ EJEMPLO 65 Sea $\mathcal{B}$ la base estándar de $\mathbf{R}^2$, y $\mathcal{B}'$ la base $\mathcal{B}' = \{(1, 1), (-1, 1)\}$.

(a) Calcule la matriz de transición P de $\mathcal{B}$ a $\mathcal{B}'$.
(b) Determine la matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$.
(c) Compruebe la relación $[\mathbf{v}]_{\mathcal{B}'} = P[\mathbf{v}]_{\mathcal{B}}$ para $\mathbf{v} = (4, -2)$.

SOLUCIÓN (a) P es la matriz cuyas columnas son $[\mathbf{e}_1]_{\mathcal{B}'}$, $[\mathbf{e}_2]_{\mathcal{B}'}$. Para $[\mathbf{e}_1]_{\mathcal{B}'}$ se necesitan escalares c_1, c_2 tales que

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

La solución del sistema que resulta es $c_1 = \frac{1}{2}$, $c_2 = -\frac{1}{2}$. De aquí, $[\mathbf{e}_1]_{\mathcal{B}'} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$

De igual manera, para $[\mathbf{e}_2]_{\mathcal{B}'}$ necesitamos los escalares c_1, c_2 tales que

$$\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Al resolver el sistema resultante se obtiene $c_1 = \frac{1}{2}$, $c_2 = \frac{1}{2}$. Por consiguiente, $[\mathbf{e}_2]_{\mathcal{B}'} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$

En vista de lo anterior, la matriz de transición es

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

(b) La matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$ es P^{-1}, según el corolario 21. Entonces,

$$P^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

(c) El vector coordenado $[\mathbf{v}]_{\mathcal{B}'}$ se puede determinar en dos formas distintas; usando P,

$$P[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

o directamente a partir de $\mathcal{B}'$, calculando c_1 y c_2 tales que

$$\begin{bmatrix} 4 \\ -2 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Al resolver el sistema se obtiene $c_1 = 1$, $c_2 = -3$. Por tanto, $[\mathbf{v}]_{\mathcal{B}'} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ en cada caso. ☐

■ EJEMPLO 66 Calcule la matriz de transición P de la base estándar $\mathcal{B}$ de $\mathbf{R}^4$ a la base $\mathcal{B}' = \{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3', \mathbf{v}_4'\}$, donde

$$\mathbf{v}_1' = \mathbf{e}_4, \qquad \mathbf{v}_2' = \mathbf{e}_3, \qquad \mathbf{v}_3' = \mathbf{e}_2, \qquad \mathbf{v}_4' = \mathbf{e}_1$$

Si $\mathbf{v} = (a, b, c, d)$ emplee P para deducir una fórmula para $[\mathbf{v}]_{\mathcal{B}'}$.

SOLUCIÓN Como

$$\mathbf{e}_1 = 0\mathbf{v}_1' + 0\mathbf{v}_2' + 0\mathbf{v}_3' + 1\mathbf{v}_4', \qquad \mathbf{e}_2 = 0\mathbf{v}_1' + 0\mathbf{v}_2' + 1\mathbf{v}_3' + 0\mathbf{v}_4'$$
$$\mathbf{e}_3 = 0\mathbf{v}_1' + 1\mathbf{v}_2' + 0\mathbf{v}_3' + 0\mathbf{v}_4', \qquad \mathbf{e}_4 = 1\mathbf{v}_1' + 0\mathbf{v}_2' + 0\mathbf{v}_3' + 0\mathbf{v}_4'$$

tenemos que

$$[\mathbf{e}_1]_{\mathcal{B}'} = \mathbf{e}_4, \qquad [\mathbf{e}_2]_{\mathcal{B}'} = \mathbf{e}_3, \qquad [\mathbf{e}_3]_{\mathcal{B}'} = \mathbf{e}_2, \qquad [\mathbf{e}_4]_{\mathcal{B}'} = \mathbf{e}_1$$

Se ve que,

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Por consiguiente,

$$[\mathbf{v}]_{\mathcal{B}'} = P[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} d \\ c \\ b \\ a \end{bmatrix}$$

□

■ EJEMPLO 67 Determine la matriz P de transición de la base $\mathcal{B}$ estándar de $\mathbf{R}^2$ a la base $\mathcal{B}'$, la cual se obtiene al girar $\mathcal{B}$ 45° en sentido contrario al de las manecillas del reloj en torno al origen. Obtenga las nuevas coordenadas del vector (1, 1).

SOLUCIÓN Ya que sen 45° = cos 45° = $\frac{\sqrt{2}}{2}$, entonces

$$\mathcal{B}' = \{\mathbf{e}_1', \mathbf{e}_2'\} = \left\{\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)\right\}$$

(figura 4.15). Así,

$$\mathbf{e}_1 = \frac{\sqrt{2}}{2}\mathbf{e}_1' - \frac{\sqrt{2}}{2}\mathbf{e}_2' \quad \text{y} \quad \mathbf{e}_2 = \frac{\sqrt{2}}{2}\mathbf{e}_1' + \frac{\sqrt{2}}{2}\mathbf{e}_2'$$

Por tanto,

$$P = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

Las nuevas coordenadas de (1, 1) se calculan como sigue:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{B}'} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}$$

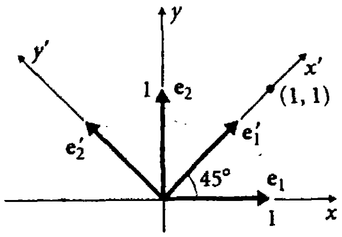


Figura 4.15 Rotación de 45° con respecto al origen.

Ejercicios 4.5

Coordenadas

En los ejercicios 1 a 3 deduzca el polinomio p, a partir de una base $\mathcal{B}$ de P_n y el vector coordenado $[p]_{\mathcal{B}}$.

1. $\mathcal{B} = \{1 + 2x, 5x\}$, $[p]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$

2. $\mathcal{B} = \{1 + x + 2x^2, -x^2, 1 + 2x\}$, $[p]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 3 \\ -2 \end{bmatrix}$

3. $\mathcal{B} = \{2 + 2x, -3 + 3x\}$, $[p]_{\mathcal{B}} = \begin{bmatrix} a \\ b \end{bmatrix}$

En los ejercicios 4 a 7 calcule el vector coordenado $[p]_{\mathcal{B}}$, partiendo de una base $\mathcal{B}$ de P_n y p.

4. $\mathcal{B} = \{1 + 2x, 1 - x\}$, $p = 4 + 17x$
5. $\mathcal{B} = \{-7 + 4x, 2 - 3x\}$, $p = 17 - 6x$
6. $\mathcal{B} = \{1 + 2x + 2x^2, 2x - x^2, -1 - 2x\}$, $p = -1 + 6x - 8x^2$
7. $\mathcal{B} = \{1 + 2x, 5x\}$, $p = (a - b) + (7a + 3b)x$

Sea $\mathcal{B}$ la siguiente base de M_{22}.

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right\}$$

8. Calcule M si $[M]_{\mathcal{B}} = (4, -3, 8, 10)$.

9. Determine el vector coordenado $\left[\begin{bmatrix} 4 & -1 \\ -4 & -4 \end{bmatrix} \right]_{\mathcal{B}}$

Matriz de transición

10. Obtenga la matriz de transición de $\{\mathbf{v}_1, \mathbf{v}_2\}$ a $\{\mathbf{v}_1', \mathbf{v}_2'\}$, donde

$$\mathbf{v}_1 = (1, 1) \qquad \mathbf{v}_2 = (1, 2)$$
$$\mathbf{v}_1' = (1, 3) \qquad \mathbf{v}_2' = (1, 4)$$

11. Determine la matriz de transición de $\{\mathbf{v}_1, \mathbf{v}_2\}$ a $\{\mathbf{v}_1', \mathbf{v}_2'\}$, siendo

$$\mathbf{v}_1 = (1, 0) \qquad \mathbf{v}_2 = (0, 1)$$
$$\mathbf{v}_1' = (0, 1) \qquad \mathbf{v}_2' = (1, 0)$$

12. Calcule la matriz de transición de $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a $\{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3'\}$, donde

$$\mathbf{v}_1 = \mathbf{e}_3 \qquad \mathbf{v}_2 = \mathbf{e}_1 \qquad \mathbf{v}_3 = \mathbf{e}_2$$
$$\mathbf{v}_1' = \mathbf{e}_1 \qquad \mathbf{v}_2' = \mathbf{e}_2 \qquad \mathbf{v}_3' = \mathbf{e}_3$$

13. Determine la matriz de transición de $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a $\{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3'\}$, en la que

$$\mathbf{v}_1 = \mathbf{e}_1 \qquad \mathbf{v}_2 = \mathbf{e}_2 \qquad \mathbf{v}_3 = \mathbf{e}_3$$
$$\mathbf{v}_1' = \mathbf{e}_3 \qquad \mathbf{v}_2' = \mathbf{e}_1 \qquad \mathbf{v}_3' = \mathbf{e}_2$$

En los ejercicios 14 a 19 determine la matriz de transición de la base estándar $\mathcal{B}_1 = \{1, x, x^2, x^3\}$ de P_3 a la base dada $\mathcal{B}_2$. (En cada caso, los polinomios de $\mathcal{B}_2$ son epónimos y surgen en forma natural en varios campos de las matemáticas y la física.)

14. **(Polinomios de Chebyshev, primer tipo)** $\mathcal{B}_2 = \{1, x, -1 + 2x^2, -3x + 4x^3\}$
15. **(Polinomios de Chebyshev, segundo tipo)** $\mathcal{B}_2 = \{1, 2x, -1 + 4x^2, -4x + 8x^3\}$
16. **(Polinomios de Laguerre)** $\mathcal{B}_2 = \{1, 1 - x, 1 - 2x + (1/2)x^2, 1 - 3x + (3/2)x^2 - (1/6)x^3\}$
17. **(Polinomios de Hermite)** $\mathcal{B}_2 = \{1, 2x, -2 + 4x^2, -12x + 8x^3\}$
18. **(Polinomios de Legendre)** $\mathcal{B}_2 = \{1, x, -\frac{1}{2} + \frac{3}{2}x^2, -\frac{3}{2}x + \frac{5}{2}x^3\}$.
19. **(Polinomios de Euler)** $\mathcal{B}_2 = \{1, -\frac{1}{2} + x, -x + x^2, \frac{1}{4} - \frac{3}{2}x^2 + x^3\}$.
20. Calcule la matriz P de transición de la base estándar $\mathcal{B}$ de $\mathbf{R}^2$ a la base $\mathcal{B}'$, obtenida girando $\mathcal{B}$ 45° en sentido de las manecillas del reloj en torno al origen. Determine las nuevas coordenadas del vector $(1, 1)$.
21. Calcule la matriz P de transición de la base estándar $\mathcal{B}$ de $\mathbf{R}^2$ a la base $\mathcal{B}'$, la cual se obtiene al reflejar $\mathcal{B}$ cerca de la línea $y = -x$. Encuentre las nuevas coordenadas del vector $(1, 1)$.
22. Obtenga la matriz P de transición de la base estándar $\mathcal{B}$ de $\mathbf{R}^3$ a la base $\mathcal{B}'$, que se obtiene al girar $\mathcal{B}$ 90° alrededor del eje z, en sentido contrario a las manecillas del reloj. Determine las nuevas coordenadas del vector $(1, 1, 1)$.

4.6 Rango y nulidad

Objetivos del estudiante para esta sección

1. Aprender a determinar las bases para los espacios nulo de renglones y de columnas de una matriz.
2. Calcular una base para el generador de vectores en $\mathbf{R}^n$.
3. Comprender y aplicar el teorema del rango.

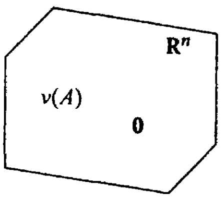
En esta sección estudiaremos tres espacios vectoriales importantes asociados con cualquier matriz: el *espacio nulo*, el *espacio de columnas* y el *espacio de renglones*. El espacio nulo fue presentado en la sección 2.5. Los otros dos espacios son los generadores de las columnas y los renglones de la matriz. Los subproductos de este estudio serán dos métodos que nos servirán para calcular las bases de un conjunto de vectores.

Nulidad

De acuerdo con la sección 2.5, recordamos que el **espacio nulo** $\nu(A)$ de una matriz A $m \times n$ contiene todos los vectores n $\mathbf{x}$, tales que $A\mathbf{x} = \mathbf{0}$. Éste es el conjunto de todas las soluciones del sistema homogéneo $A\mathbf{x} = \mathbf{0}$.

$$\nu(A) = \{\mathbf{x} \text{ en } \mathbf{R}^n \text{ tal que } A\mathbf{x} = \mathbf{0}\}$$

De acuerdo con el teorema 18 de la sección 2.5, el espacio nulo de A es un *subespacio* de $\mathbf{R}^n$. La dimensión de $\nu(A)$ se llama **nulidad** de A.



Puesto que $\nu(A)$ es un subespacio, cabría preguntarse cómo determinar una base para él. Esto se explica en el algoritmo siguiente que también es un ejemplo.

Algoritmo

(Determinación de una base para el espacio nulo)

Para obtener una base para $\nu(A)$ es preciso:

1. Calcular el vector solución general del sistema $A\mathbf{x} = \mathbf{0}$.
2. Escribir el vector solución como combinación lineal con los parámetros (variables libres) como coeficientes.
3. Los vectores de la combinación lineal forman una base para $\nu(A)$.

■ EJEMPLO 68 Determine una base para el espacio nulo de A. ¿Cuál es la nulidad de A?

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 & 0 \\ -1 & 0 & -4 & 3 & -1 \\ 2 & -1 & 6 & 0 & 1 \\ -1 & 2 & 0 & -1 & 1 \end{bmatrix}$$

SOLUCIÓN La matriz aumentada $[A : \mathbf{0}]$ del sistema $A\mathbf{x} = \mathbf{0}$ tiene una forma escalonada reducida

$$\begin{bmatrix} 1 & 0 & 4 & 0 & 1 & : & 0 \\ 0 & 1 & 2 & 0 & 1 & : & 0 \\ 0 & 0 & 0 & 1 & 0 & : & 0 \\ 0 & 0 & 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Por consiguiente, el sistema original equivale a

$$\begin{aligned} x_1 \qquad + 4x_3 \qquad + x_5 &= 0 \\ x_2 + 2x_3 \qquad + x_5 &= 0 \\ x_4 \qquad &= 0 \end{aligned}$$

Si $x_5 = r$ y $x_3 = s$ son escalares cualesquiera, la solución general es

$$\begin{aligned} x_1 &= -4s - r \\ x_2 &= -2s - r \\ x_3 &= s \\ x_4 &= 0 \\ x_5 &= r \end{aligned}$$

Ya que

$$\begin{bmatrix} -4s - r \\ -2s - r \\ s \\ 0 \\ r \end{bmatrix} = r \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + s \begin{bmatrix} -4 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

el espacio nulo de A queda generado por el conjunto

$$B = \left\{ \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

que se ve con facilidad es linealmente independiente. En consecuencia, B es una base de $\nu(A)$, y como B tiene dos elementos, la nulidad de A es 2.

NOTA Cuando la solución general de $A\mathbf{x} = \mathbf{0}$ se escribe como una combinación lineal cuyos parámetros son coeficientes, los vectores de la combinación lineal no sólo generan al espacio nulo, sino que también son *linealmente independientes*, porque los parámetros se presentan en distintos componentes de la combinación. Por ejemplo, sea $r\mathbf{v}_1 + s\mathbf{v}_2 = \mathbf{0}$, siendo $\mathbf{v}_1$ y $\mathbf{v}_2$ los dos vectores de B en el último ejemplo. Entonces $r = 0$, porque r corresponde a la variable libre x_5, así que el quinto componente del vector solución es $r \cdot 1$. De igual forma, $s = 0$. Una consecuencia inmediata de esto es que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes. Es la razón por la que nuestro algoritmo produce una *base* para el espacio nulo.

Como la cantidad de parámetros determina la cantidad de vectores en la base de $\nu(A)$, llegamos al teorema siguiente.

TEOREMA 22

La nulidad de una matriz A es igual a la cantidad de variables libres de $A\mathbf{x} = \mathbf{0}$.

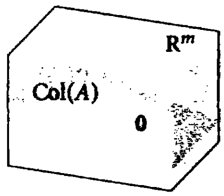
El espacio de columnas

A continuación seccionaremos una matriz en columnas y renglones para estudiar los subespacios generados por ellos.

El **espacio de columnas**, Col(A), de una matriz A es el generador de sus columnas. Por ejemplo, si

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 2 & 0 & 2 \end{bmatrix} \quad \text{donde} \quad \mathrm{Col}(A) = \mathrm{Gen}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix} \right\}$$

El espacio de columnas de una matriz $m \times n$ es un *subespacio* de $\mathbf{R}^m$, porque es el generador de vectores m.



Puesto que un sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente si y sólo si $\mathbf{b}$ es el generador de las columnas de A, llegamos al teorema siguiente.

TEOREMA 23

> Un sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente, si y sólo si, $\mathbf{b}$ está en el Col(A).

■ EJEMPLO 69 ¿Cuál de

$$\mathbf{u} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

está en el espacio de columnas de $A = \begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}$?

SOLUCIÓN Ya que

$$\begin{bmatrix} 1 & -2 & -2 \\ 2 & -4 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -2 \\ 0 & 0 & 6 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

el sistema $A\mathbf{x} = \mathbf{u}$ es inconsistente, mientras que $A\mathbf{x} = \mathbf{v}$ es consistente. Por tanto, $\mathbf{v}$ está en Col(A) y $\mathbf{u}$ no está. ☐

Una base para el espacio de columnas

Nuestra siguiente tarea es demostrar que las columnas pivote de cualquier matriz forman una base para su espacio de columnas. Este hecho puede aprovecharse para determinar una base del generador de un conjunto o sucesión finita de vectores. Comencemos con un ejemplo en el que la matriz está en forma escalonada.

■ EJEMPLO 70 Determine una base para Col(B).

$$B = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

SOLUCIÓN Las columnas 1, 3 y 5 son columnas pivote linealmente independientes. Las que no lo son pueden escribirse como combinaciones lineales de las columnas pivote. Por ejemplo, es fácil ver que $\mathbf{b}_2 = -2\mathbf{b}_1$, y que $\mathbf{b}_4 = -\mathbf{b}_1 + \mathbf{b}_3$. Así,

$$\text{Col}(B) = \text{Gen}\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4, \mathbf{b}_5\} = \text{Gen}\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$$

basándonos el teorema 9 de la sección 2.3 (o el teorema 6 de la sección 4.3). Por consiguiente, las columnas pivote $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ forman una base para el Col(B). ☐

■ EJEMPLO 71 Encuentre una base para Col(A).

$$A = \begin{bmatrix} 1 & -2 & 2 & 1 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 2 & -4 & 6 & 4 & 0 \\ 3 & -6 & 8 & 5 & 1 \end{bmatrix}$$

SOLUCIÓN No es difícil comprobar que A se reduce a la matriz B del ejemplo 70. En consecuencia, las columnas pivote de A son 1, 3 y 5. Es algo más complicado demostrar que son linealmente independientes. Para las columnas no pivote se tiene $\mathbf{a}_2 = -2\mathbf{a}_1$ y $\mathbf{a}_4 = -\mathbf{a}_1 + \mathbf{a}_3$ (las mismas relaciones de dependencia lineal que antes). Por consiguiente, las columnas pivote $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ forman una base para Col(A). □

Los ejemplos 70 y 71 sugieren que las columnas de una matriz y las de una forma escalonada equivalente satisfacen las mismas relaciones de dependencia lineal.

TEOREMA 24

Si $A \sim B$, las columnas de A y B cumplen en las mismas relaciones de dependencia lineal, es decir,

$$c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n = \mathbf{0} \Leftrightarrow c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = \mathbf{0} \tag{4.6}$$

DEMOSTRACIÓN Como $A \sim B$, los sistemas $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen las mismas soluciones. Entonces

$$A\mathbf{c} = \mathbf{0} \Leftrightarrow B\mathbf{c} = \mathbf{0} \tag{4.7}$$

Si $\mathbf{c}$ tiene componentes $c_1, \ldots, c_n$, se obtiene la ecuación (4.6), de acuerdo con la definición de $A\mathbf{c}$ y $B\mathbf{c}$. □

El teorema 24 implica que cualquier conjunto de columnas de A es linealmente dependiente (o independiente) si y sólo si el conjunto correspondiente de columnas de B es linealmente dependiente (o independiente).

TEOREMA 25

Las columnas pivote de cualquier matriz forman una base para su espacio de columnas.

DEMOSTRACIÓN Si A es una matriz $m \times n$ y B es su forma escalonada reducida, demostraremos que las columnas pivote de A son linealmente independientes, y que las columnas no pivote son combinaciones lineales de las primeras. De acuerdo con el teorema 24, basta con demostrar estas afirmaciones para B.

Supongamos que B tiene k columnas pivote, digamos $\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_k}$. Como B está en su forma escalonada reducida, $\mathbf{b}_{i_1} = \mathbf{e}_1, \ldots, \mathbf{b}_{i_k} = \mathbf{e}_k$, y cada $\mathbf{e}_j$ está en $\mathbf{R}^m$. De aquí, $\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_k}$ son linealmente independientes. Ya que B está en forma escalonada, los últimos $m - k$ componentes de sus columnas son cero, así que $\mathrm{Col}(B) \subseteq \mathrm{Gen}\{\mathbf{e}_1, \ldots, \mathbf{e}_k\} \subseteq \mathbf{R}^m$. Por consiguiente, $\dim \mathrm{Col}(B) \leq k$, según el teorema 17 de la sección 4.4. Como ya tenemos k columnas pivote linealmente independientes, $\dim \mathrm{Col}(B) \geq k$, como se planteó en el teorema 14 de la sección 4.4. Por lo anterior, $\dim \mathrm{Col}(B) = k$, y las columnas pivote forman una base, de acuerdo con el teorema 15 de la sección 4.4. □

ADVERTENCIA Las *operaciones elementales de renglón pueden cambiar el espacio de columnas de una matriz.* Regresando a los ejemplos 70 y 71, vemos que como el último elemento de las columnas pivote de B es cero, la última columna de A no está en el Col(B). Así, aunque $A \sim B$, el $\mathrm{Col}(A) \neq \mathrm{Col}(B)$ en este caso. Cuando se determina una base para el Col(A), hay que asegurarse de usar las columnas pivote de la matriz A *dada*, y *no* las columnas de su forma escalonada B.

■ EJEMPLO 72 (Selección de una base a partir de un conjunto generador) Determine una base de S para Gen(S), donde

$$S = \{(1, -1, 2, 3), (-2, 2, -4, -6,), (2, -1, 6, 8), (1, 0, 4, 5), (0, 0, 0, 1)\}$$

SOLUCIÓN Basta con calcular una base para el espacio de columnas de la matriz cuyas columnas son los vectores de S. Esta matriz A corresponde al ejemplo 71, donde las columnas pivote fueron la 1, 3 y 5. Por tanto,

$$\{(1, -1, 2, 3), (2, -1, 6, 8), (0, 0, 0, 1)\}$$

es una base para Gen(S), de acuerdo con el teorema 25. ☐

Algoritmo A

(Cálculo de una base para Gen(S))

Sea $S = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subseteq \mathbf{R}^m$. Se puede determinar una base para Gen(S) como sigue:

1. Forme la matriz A $m \times n$ con las columnas $\mathbf{a}_1, \ldots, \mathbf{a}_n$.
2. Reduzca A con operaciones de renglón hasta llegar a una forma escalonada B e identifique las columnas pivote de A.
3. Una base para Gen(S) es el conjunto de las columnas pivote de A.

- **Ventaja del algoritmo A:** Todos los vectores de la base están en S.
- **Desventaja:** Los vectores de base pueden no tener muchos componentes cero.

El teorema 25 también puede aplicarse para ampliar un conjunto linealmente independiente de $\mathbf{R}^n$ a una base, como veremos en el ejemplo siguiente:

■ EJEMPLO 73 (Ampliación de un conjunto linealmente independiente a una base) Extienda el conjunto linealmente independiente $S = \{(1, 0, -1, 0), (-1, 1, 0, 0)\}$ a una base en $\mathbf{R}^4$.

SOLUCIÓN Ampliaremos S a un conjunto generador S', agregando la base estándar de $\mathbf{R}^4$. En notación de columnas tenemos,

$$S' = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

A continuación se reduce con operaciones elementales de renglón la matriz cuyas columnas son los elementos de S', para obtener

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Por consiguiente, las columnas 1, 2, 3 y 6 de S' forman una base para Gen(S) = $\mathbf{R}^4$, según el teorema 25. Entonces

$$\{(1, 0, -1, 0), (-1, 1, 0, 0), (1, 0, 0, 0), (0, 0, 0, 1)\}$$

es una base de $\mathbf{R}^4$ que amplía el conjunto linealmente independiente S. ☐

NOTA El método del ejemplo 73 se aplica a *todo*[5] subconjunto S linealmente independiente de $\mathbf{R}^n$.

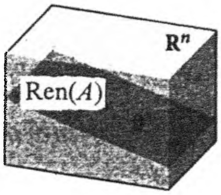
El espacio de renglones

DEFINICIÓN

El **espacio de renglones**, Ren(A), de una matriz A es el generador de sus renglones.

$$\text{Si A} = \begin{bmatrix} 1 & 0 \\ 1 & -2 \\ 2 & 0 \end{bmatrix}, \text{ entonces Ren}(A) = \text{Gen}\left\{\begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\} = \text{Gen }\{(2, 0), (1, -2)\}.$$

El espacio de renglones de una matriz $m \times n$ es un *subespacio* de $\mathbf{R}^n$, porque es el generador de vectores n.



A diferencia de los espacios de columnas, *los espacios de renglones no se afectan con las operaciones elementales de renglón*. En realidad, B se obtiene de A por una operación elemental de renglón $\mathcal{O}$. Si $\mathcal{O}$ es $R_i \leftrightarrow R_j$, el conjunto de renglones permanece igual. Si $\mathcal{O}$ es $cR_i \to R_i$ o $R_i + cR_j \to R_i$, el nuevo i-ésimo renglón es una combinación lineal de los anteriores. Por consiguiente, Ren(B) $\subseteq$ Ren(A). Como $\mathcal{O}$ es reversible, también Ren(A) $\subseteq$ Ren(B). Así, en todos los casos, Ren(A) = Ren(B).

TEOREMA 26

Si $A \sim B$, entonces Ren(A) = Ren(B).

DEMOSTRACIÓN B se obtiene de A con un conjunto finito de operaciones elementales de renglón, y todos los espacios de renglones de las matrices intermedias son iguales. ☐

TEOREMA 27

Los renglones no cero de una matriz A en forma escalonada de renglón son linealmente independientes.

DEMOSTRACIÓN Sea $c_1\mathbf{r}_{i_1} + \cdots c_k\mathbf{r}_{i_k} = \mathbf{0}$, donde $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_k}$ son los renglones no cero de A. Como esta matriz es escalonada, todos los elementos debajo del elemento delantero de $\mathbf{r}_{i_1}$ son 0. Así, $c_1 = 0$. Entonces podemos eliminar el término $c_1\mathbf{r}_{i_1}$ y repetir el argumento. Al final, todas las c_i serán cero. Por consiguiente, $\{\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_k}\}$ es linealmente independiente. ☐

[5] Si S tiene k elementos y S' se obtuvo de S agregándole la base estándar de $\mathbf{R}^n$. Y como S es linealmente independiente, las k primeras columnas de la matriz reducida S' son columnas pivote. Por consiguiente, todos los vectores de S se toman, por la reducción con operaciones de renglón, como parte de la base.

Los teoremas 26 y 27 conducen al siguiente.

TEOREMA 28

> Los renglones no cero de cualquier forma escalonada de una matriz A forman una base para Ren(A).

■ **EJEMPLO 74** Determine una base para Ren(A)

$$A = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 1 & 3 & 1 & -2 \\ 1 & 1 & 3 & 0 \\ 0 & 1 & -1 & -1 \\ 1 & 2 & 2 & -1 \end{bmatrix}$$

SOLUCIÓN A se reduce a la matriz de forma escalonada de renglón

$$B = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

El conjunto $\{(1, 2, 2, -1), (0, 1, -1, -1)\}$ de renglones de B distintos a cero forma una base para Ren(A) = Ren(B).

Observe que esta base no consiste exclusivamente en renglones de A. □

El método del ejemplo 74 ofrece una forma alternativa para determinar una base del generador de un conjunto finito de vectores n. Primero se forma la matriz cuyos renglones son los vectores dados y después se calcula una base para su espacio de renglones. Es posible que esta base temporal *no* contenga totalmente los vectores dados.

■ **EJEMPLO 75** **(Base para el generador)** Obtenga una base para Gen(S), siendo

$$S = \{(1, -1, 2, 3), (-2, 2, -4, -6), (2, -1, 6, 8), (1, 0, 4, 5), (0, 0, 0, 1)\}$$

SOLUCIÓN En el ejemplo 72 contestamos esta pregunta. A continuación mostraremos otra manera de calcular el espacio de renglones de la matriz con los elementos de S como *renglones*. Puesto que

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 \\ -2 & 2 & -4 & -6 \\ 2 & -1 & 6 & 8 \\ 1 & 0 & 4 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\{(1, -1, 2, 3), (0, 1, 2, 2), (0, 0, 0, 1)\}$ es una base para Gen(S). Ésta es un poco distinta de la que se determinó en el ejemplo 72. En particular $(0, 1, 2, 2)$ *no está* en el conjunto original S.

Otra base es $\{(1, 0, 4, 0), (0, 1, 2, 0), (0, 0, 0, 1)\}$, la cual se obtiene de la forma reducida de A por operaciones de renglón. En general, esa base tiene *más componentes cero*, por lo que muchas veces es más fácil de usar. □

Algoritmo B

(Cálculo de la base para Gen(S))

Sea $S = \{\mathbf{r}_1, \ldots, \mathbf{r}_m\} \subseteq \mathbf{R}^n$. Se puede determinar una base para Gen(S) como sigue:

1. Forme la matriz A $m \times n$ cuyos renglones son $\mathbf{r}_1, \ldots, \mathbf{r}_m$.
2. Reduzca A mediante operaciones de renglón hasta una forma escalonada B.
3. Una base para Gen(S) es el conjunto de renglones de B distintos de cero.

- **Ventaja del algoritmo B**: Pueden obtenerse bases fáciles (varios ceros en los componentes de vector).
- **Desventaja**: Los vectores de la base pueden no estar en S.

ADVERTENCIA *Las operaciones elementales de renglón no conservan las relaciones de dependencia lineal entre los renglones*. Por ejemplo, considere $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ y $\mathbf{r}_2 = 2\mathbf{r}_1$ para la primera matriz, y $\mathbf{r}_2 \neq 2\mathbf{r}_1$ para la segunda.

Rango

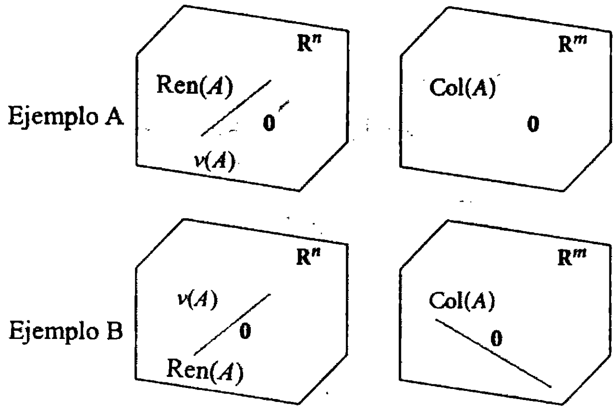
Como la dimensión del Col(A) es la cantidad de pivotes de A, la cual también es igual que la cantidad de renglones no cero de una forma escalonada de A, entonces:

TEOREMA 29

Para cualquier matriz A,

$$\dim \mathrm{Col}(A) = \dim \mathrm{Ren}(A)$$

En forma gráfica:



DEFINICIÓN

(Rango)

La dimensión común de los espacios de columnas y de renglones de A se llama **rango** de A y se representa por Rango(A).

El rango es la cantidad de pivotes de A. Para determinarlo, se reduce A a la forma escalonada y se cuenta la cantidad de renglones distintos de cero, o la de columnas pivote.

■ EJEMPLO 76 En el ejemplo 74, el rango de A es 2, porque la forma escalonada de renglón B tiene dos renglones no cero. □

NOTA El rango de una matriz $m \times n$ es menor que o igual a m, n. (¿Por qué?)

■ EJEMPLO 77 Una matriz de 5×9 ¿puede tener el rango 6?

RESPUESTA No, su rango no puede ser mayor que 5. □

A continuación presentaremos una consecuencia importante del teorema 29.

COROLARIO 30

> A y A^T tienen el mismo rango.

DEMOSTRACIÓN El espacio de columnas de A es el mismo que el de A^T. □

El teorema del rango

El resultado siguiente es uno de los *teoremas más importantes* del álgebra lineal.

TEOREMA 31

> (El teorema del rango)
>
> Para toda matriz A,
>
> $$\text{Rango}(A) + \text{Nulidad}(A) = \text{cantidad de columnas de } A.$$

DEMOSTRACIÓN El rango de A es la cantidad de columnas pivote de A. Por otro lado, la nulidad de A es la cantidad de variables libres de $A\mathbf{x} = \mathbf{0}$, según el teorema 22. Como se tienen tantas variables libres como columnas no pivote, la nulidad es igual a la cantidad de columnas no pivote. El teorema es consecuencia de

$$\text{Núm. de columnas pivote} + \text{Núm. de columnas no pivote} = \text{Núm. de columnas}$$ □

■ EJEMPLO 78 Compruebe el teorema del rango para

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 & 0 \\ -1 & 0 & -4 & 3 & -1 \\ 2 & -1 & 6 & 0 & 1 \\ -1 & 2 & 0 & -1 & 1 \end{bmatrix}$$

SOLUCIÓN En el ejemplo 68 demostramos que la nulidad de A es 2. Por otro lado, la forma de escalón reducida de A es

$$\begin{bmatrix} 1 & 0 & 4 & 0 & 1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

En consecuencia, el rango de A es 3. Sumando $2 + 3 = 5$ se obtiene la cantidad de columnas de A, tal como dice el teorema del rango. □

■ EJEMPLO 79 Suponga que el sistema $A\mathbf{x} = \mathbf{0}$ tiene 20 incógnitas, y que su espacio de soluciones está generado por 6 vectores linealmente independientes.

(a) ¿Cuál es el rango de A?
(b) ¿A puede tener el tamaño 13×20?

SOLUCIÓN (a) La cantidad de columnas de A es 20, y la nulidad es 6. Por consiguiente, el rango de A es $20 - 6 = 14$, de acuerdo con el teorema del rango.

(b) No. El rango no puede ser mayor que la cantidad de renglones, de modo que A debe tener al menos 14 renglones. □

■ EJEMPLO 80 Sea $A\mathbf{x} = \mathbf{b}$ un sistema de 20 ecuaciones con 24 incógnitas. Si el espacio nulo de A es generado por cuatro vectores linealmente independientes, ¿podemos tener la certeza de que el sistema es consistente para cualquier elección de $\mathbf{b}$?

SOLUCIÓN Sí. La matriz A tiene 24 columnas y su nulidad es 4, de modo que su rango es 20. (¿Por qué?) Así, la dimensión del espacio de columnas es 20 en $\mathbf{R}^{20}$. Por tanto, todo el espacio de columnas consiste en $\mathbf{R}^{20}$, según el teorema 17 de la sección 4.4. Entonces, cualquier vector $\mathbf{b}$ en $\mathbf{R}^{20}$ está generado por las columnas de A. En vista de lo anterior, $A\mathbf{x} = \mathbf{b}$ es consistente para todo $\mathbf{b}$ en $\mathbf{R}^{20}$. □

Rango y sistemas lineales

La teoría y los métodos que se desarrollaron en esta sección están muy relacionados con los sistemas lineales.

Como un sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente si y sólo si $\mathbf{b}$ está en el Col(A), de acuerdo con el teorema 23, llegamos al teorema siguiente.

TEOREMA 32

> El sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente si y sólo si
>
> $$\text{Rango}(A) = \text{Rango}([A : \mathbf{b}])$$

El interés de este teorema es principalmente teórico. En la práctica, hay que reducir $[A : \mathbf{b}]$ de todos modos, sea para determinar su rango o para ver si la última columna es pivote. En consecuencia, no se gana ventaja alguna.

Por último, veamos lo que sucede en los casos extremos en los que el rango de una matriz $m \times n$ es m o n. Los dos teoremas siguientes resumen los principales resultados del capítulo 2 hasta este punto. Las demostraciones se han hecho parcialmente en otros apartados, y se dejan como ejercicios.

TEOREMA 33

Sea A una matriz $m \times n$. Las afirmaciones siguientes son equivalentes.

1. A tiene rango m.
2. A tiene m pivotes.
3. Cada renglón de A tiene un pivote.
4. El sistema $A\mathbf{x} = \mathbf{b}$ es consistente para todos los vectores m $\mathbf{b}$.
5. Las columnas de A generan a $\mathbf{R}^m$.
6. $\mathrm{Col}(A) = \mathbf{R}^m$.
7. $\dim \mathrm{Col}(A) = m$.
8. $\dim \mathrm{Ren}(A) = m$.
9. Nulidad $(A) = n - m$.
10. A^T tiene rango m.

TEOREMA 34

Sea A una matriz $m \times n$. Las afirmaciones siguientes son equivalentes.

1. A tiene rango n.
2. A tiene n pivotes.
3. Cada columna de A es columna pivote.
4. Las columnas de A son linealmente independientes.
5. El sistema homogéneo $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial.
6. $\nu(A) = \{\mathbf{0}\}$.
7. Nulidad$(A) = 0$.
8. $\dim \mathrm{Col}(A) = n$.
9. $\dim \mathrm{Ren}(A) = n$.
10. A^T tiene rango n.

Unicidad de la forma escalonada reducida (opcional)

Como aplicación de los conceptos explicados en esta sección, demostraremos que la forma escalonada reducida de cualquier matriz es única.

Teorema 1 de la sección 1.2

Cada matriz equivale en renglones a una y sólo una matriz en forma escalonada reducida con operaciones elementales de renglón.

DEMOSTRACIÓN Si A es cualquier matriz $m \times n$, tenemos la garantía de que tiene al menos una forma escalonada reducida, ya que el proceso de eliminación de Gauss-Jordan determina una.

Si N es otra forma escalonada reducida de A, demostraremos que $M = N$. Primero, $M \sim N$, porque $M \sim A$ y $A \sim N$. Así, las columnas de M y N satisfacen las mismas relaciones de dependencia, según el teorema 24. Si M tiene k columnas pivote, éstas son precisamente $\mathbf{e}_1, \ldots, \mathbf{e}_k$, para cada $\mathbf{e}_i$ en $\mathbf{R}^m$, porque M está en su forma escalonada reducida. Además, una columna de M (y de N) es columna pivote si y sólo si *no es* una combinación lineal de las columnas a su izquierda. Suponga que $\mathbf{m}_i$ es la i-ésima columna de M.

Caso 1: Sea $\mathbf{m}_i$ una columna pivote. Entonces $\mathbf{m}_i = \mathbf{e}_j$, para alguna j, y $\mathbf{m}_i$ no es combinación lineal de las columnas anteriores. Y lo mismo aplica para la i-ésima columna, $\mathbf{n}_i$ de N, porque las columnas de M y N satisfacen las mismas relaciones de dependencia. De manera que $\mathbf{n}_i$ es una columna pivote de N; puesto que es la j-ésima columna pivote, $\mathbf{n}_i = \mathbf{e}_j$. En consecuencia, $\mathbf{m}_i = \mathbf{n}_i$.

Caso 2: Sea $\mathbf{m}_i$ una columna no pivote. Entonces $\mathbf{m}_i$ es una combinación lineal de las columnas pivote anteriores, de acuerdo con el teorema 25. Así, lo mismo vale para la i-ésima columna, $\mathbf{n}_i$ de N, porque las columnas de M y N cumplen las mismas relaciones de dependencia. Pero las columnas pivote de M y N son las mismas, y por esa causa, $\mathbf{m}_i = \mathbf{n}_i$.

Nuestra conclusión es que M y N tienen las mismas columnas, por lo que la igualdad $M = N$ es acertada. $\square$

Col(A) y $\nu(A)$ con sistemas algebraicos computacionales

Determinación de bases para Col(A) y $\nu(A)$:

Maple

```
> with(linalg):
> A:=matrix([[1,2,3,4],[2,3,4,5],[3,4,5,6]]);
```

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$$

```
> colspace(A);nullspace(A);
```

$$\{[1,0,-1],[0,1,2]\}$$
$$\{[1,-2,1,0],[2,-3,0,1]\}$$

Mathematica

```
    A={{1,2,3,4},{2,3,4,5},{3,4,5,6}};
In[2]:=
    RowReduce[A]
    NullSpace[A]
Out[2]=
    {{1, 0, -1, -2}, {0, 1, 2, 3}, {0, 0, 0, 0}}
Out[3]=
    {{2, -3, 0, 1}, {1, -2, 1, 0}}
```

Col(A): cálculo indirecto

MATLAB

```
>> A=[1 2 3 4; 2 3 4 5; 3 4 5 6];
>> rref(A), null(A)
ans =                          ans =
    1    0   -1   -2             0.1507   -0.5266
    0    1    2    3             0.1916    0.8144
    0    0    0    0            -0.8352   -0.0491
                                 0.4929   -0.2388
```

Col(A): cálculo indirecto

Ejercicios 4.6

Espacio nulo

En los ejercicios 1 a 7 determine una base para el espacio nulo y la nulidad de la matriz dada. (Recuerde que el subespacio cero tiene dimensión 0 y es base del conjunto vacío.)

1. **a.** $\begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}$ **b.** $\begin{bmatrix} 2 & -2 & 2 \\ 3 & -3 & 3 \\ 4 & -4 & 5 \end{bmatrix}$

2. **a.** $\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 8 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ 3 & -3 & 3 \\ 4 & -4 & 4 \end{bmatrix}$

3. **a.** $\begin{bmatrix} 1 & 2 & -1 & -3 & 0 & 6 \\ 0 & 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 9 \end{bmatrix}$

b. $\begin{bmatrix} -1 & 1 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

4. **a.** $\begin{bmatrix} -1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 4 \\ 0 & -3 & 3 & 9 \end{bmatrix}$

b. $\begin{bmatrix} -1 & 1 & 1 & 2 \\ 2 & -2 & -2 & -4 \\ 0 & -3 & 3 & 9 \end{bmatrix}$

5. **a.** $\begin{bmatrix} 1 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \\ 2 & -4 & 6 & 0 \end{bmatrix}$

b. $\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \\ -4 & 4 \\ 0 & 0 \end{bmatrix}$

6. $\begin{bmatrix} 1 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \\ 2 & -4 & 6 & 0 \\ 3 & 3 & 0 & -1 \\ 0 & -1 & 1 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 1 & -1 & 2 & 3 & 0 \\ 2 & -1 & 6 & 0 & 1 \\ -1 & 2 & 0 & -1 & 1 \end{bmatrix}$

En los ejercicios 8 a 10 sume la nulidad a la cantidad de columnas pivote de la matriz. ¿Cómo se relaciona la suma con el tamaño de la matriz?

8. **a.** $\begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & -1 \\ 0 & 7 \end{bmatrix}$

9. **a.** $\begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 2 & 2 & -2 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 2 & -1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

11. Sea A una matriz de $m \times n$. Si el conjunto de soluciones del sistema $A\mathbf{x} = \mathbf{b}$ forma un subespacio vectorial de $\mathbf{R}^n$, ¿qué puede decir acerca de $\mathbf{b}$?

Espacio de columnas

Sean

$$\mathbf{a} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ -6 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -4 \\ -8 \\ 1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$$

En los ejercicios 12 a 17 determine cuál(es) de **a**, **b**, **c**, **u**, **v** y **w** están en el espacio de columnas de la matriz dada.

12. $\begin{bmatrix} 1 & -5 \\ -2 & 10 \end{bmatrix}$ **13.** $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \end{bmatrix}$ **15.** $\begin{bmatrix} 1 & -2 & 3 \\ 0 & -4 & 5 \\ 0 & 0 & 0 \end{bmatrix}$

16. $\begin{bmatrix} -2 & 1 & 4 & -6 \\ 0 & 4 & -5 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

17. $\begin{bmatrix} -2 & 1 & 4 & -6 & 1 \\ 0 & 4 & -5 & 0 & 2 \\ 0 & -8 & 10 & 0 & -4 \end{bmatrix}$

En los ejercicios 18 a 24 obtenga una base para el Col(A).

18. $A = \begin{bmatrix} 2 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

19. $A = \begin{bmatrix} 2 & 0 & 0 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

20. $A = \begin{bmatrix} 5 & 0 & 2 & -1 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix}$

21. $A = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 2 & 2 & 4 & 6 & -5 \\ 0 & 2 & -4 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 2 & 1 & 2 \end{bmatrix}$

22. $A = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 1 & -2 & 4 & 2 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 8 & 7 \\ 1 & -2 & -8 & -7 \end{bmatrix}$

23. $A = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 2 & 2 & 4 & 6 & -5 \\ 0 & 0 & 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 2 & 1 & 2 \end{bmatrix}$

24. $A = \begin{bmatrix} 4 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & -3 \\ 4 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

25. Haga un esquema del Col(A) para $A = \begin{bmatrix} 2 & 1 & 0 & -1 \\ -2 & 1 & 1 & 1 \end{bmatrix}$

26. Haga un esquema del Col(A) para $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$

27. Haga un esquema del Col(A) y $\nu(A)$ para $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{bmatrix}$

28. Haga un esquema del Col(A) y $\nu(A)$ para $A = \begin{bmatrix} 3 & -1 \\ -3 & 1 \\ 6 & -2 \end{bmatrix}$

29. Calcule las matrices A y B tales que $A \sim B$ y $\text{Col}(A) \neq \text{Col}(B)$.

En los ejercicios 30 a 33 determine una base para el generador del conjunto dado de vectores.

30. $\left\{ \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

31. $\left\{ \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ -9 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ -8 \end{bmatrix} \right\}$

32. $\left\{ \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -4 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$

33. $\left\{ \begin{bmatrix} 1 \\ -2 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ -1 \\ 3 \end{bmatrix} \right\}$

En los ejercicios 34 a 38 amplíe el conjunto dado linealmente independiente de vectores n, hasta una base de $\mathbf{R}^n$.

34. $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ **35.** $\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$

36. $\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\}$ **37.** $\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \right\}$

38. $\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$

Espacio de renglones

En los ejercicios 39 a 45 determine una base para el Ren(A) y determine el Rango(A).

39. $A = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 0 & -1 & 2 & 3 \\ 1 & 1 & 4 & 2 \end{bmatrix}$ 40. $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

41. $A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & -1 & 2 \\ 1 & 1 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}$. 42. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 2 \\ 1 & 1 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \\ 1 & 4 & -8 \end{bmatrix}$

En los ejercicios 43 a 45, aplique el algoritmo B para determinar una base para el generador del conjunto dado de vectores.

43. $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \end{bmatrix} \right\}$

44. $\left\{ \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \right\}$

45. $\left\{ \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \right\}$.

Rango

Sean

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & -1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

46. Compruebe el corolario 30 para A.

47. Demuestre el corolario 30 para B.

48. Compruebe el teorema del rango para A.

49. Verifique el teorema del rango para B.

50. Demuestre el teorema del rango para B^T.

51. Use el teorema 32 para comprobar que el sistema $[B : \mathbf{b}]$ con $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, es consistente.

52. Suponga que el sistema $A\mathbf{x} = \mathbf{0}$ tiene 250 incógnitas y que su espacio solución está generado por 50 vectores linealmente independientes.

    a. ¿Cuál es el rango de A?

    b. ¿A puede tener el tamaño 150×250?

    c. ¿A puede tener el tamaño 200×200?

    d. ¿A puede tener el tamaño 250×150?

    e. ¿A puede tener el tamaño 200×250?

    f. ¿A puede tener el tamaño 250×250?

53. Sea $A\mathbf{x} = \mathbf{b}$ un sistema con 400 ecuaciones y 450 incógnitas. Suponga que el espacio nulo de A está generado por 50 vectores linealmente independientes. Ese sistema, ¿es consistente para todos los vectores 400 $\mathbf{b}$?

54. Compruebe el teorema 33.

55. Demuestre el teorema 34.

4.7 Aplicaciones a la teoría de la codificación

Objetivo del estudiante para esta sección

Aprender la interesante aplicación de los espacios vectoriales a la teoría de la codificación.

Casi todos los mensajes transmitidos, desde la voz humana hasta los datos que se reciben de un satélite, están sujetos al ruido. En consecuencia, es importante poder codificar un mensaje de tal modo que después de haberse mezclado con el ruido, puede decodificarse a su forma original (figura 4.16). Esto se lleva a cabo a veces repitiendo el mensaje dos o tres veces, algo muy común en la conversación humana. Sin embargo, la repetición no siempre es eficiente: para copiar los datos almacenados en un disco flexible o duro deteriorado se requiere una gran cantidad de espacio adicional de almacenamiento.

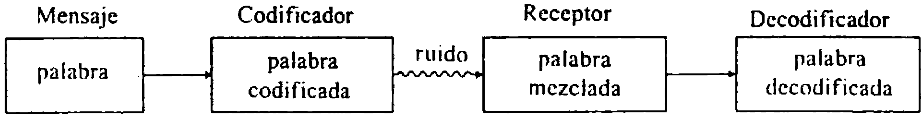


Figura 4.16 Proceso de codificación.

En esta sección examinaremos las formas de codificar un mensaje y decodificarlo después de que el ruido lo ha distorsionado. Al proceso se le llama **codificación**. Un código o clave que detecta errores en un mensaje mezclado se llama **de detección de error**. Si además puede corregir el error, se llama **de corrección de error**. Es mucho más difícil encontrar códigos correctores de error que detectores de éste. Describiremos algunos ejemplos.

La mayor parte de los mensajes es *digital*: secuencias de ceros y unos, como 10101 o 1010011; supongamos que se desea mandar el mensaje 1011. Esta "palabra" binaria puede representar una palabra real, como comprar, o una frase, como comprar acciones de canciones de los Beetles. Codificar 1011 significa agregar una cola binaria a la secuencia, para que si el mensaje se distorsiona, por ejemplo a 0011, pueda detectarse el error. Una técnica sencilla es agregar un 1 o un 0, dependiendo de si se tiene una cantidad impar o par de unos en la palabra. En esta forma, todas las palabras codificadas tendrán un número par de unos. Así, 1011 se codifica 10111. Ahora, si esto se distorsiona a 00111, se sabrá que ha ocurrido un error, porque sólo se recibió una cantidad impar de unos. A esta clave de detección de error se le llama **comprobación de paridad** (figura 4.17) y es demasiado sencilla para ser de mucha utilidad. Por ejemplo, si se cambiaran dos dígitos, nuestro esquema no detectaría el error. Aun cuando sólo hubiera un error, no sabríamos dónde está para arreglarlo. Definitivamente, ésta no es una clave correctora de error. Otro método sería codificar el mensaje, repitiéndolo, como 10111011. Así, si se recibiera 00111011, sabríamos que una de las dos mitades iguales se ha distorsionado. Si sólo hubiera ocurrido un error, claramente estaría en la posición 1. Este esquema de codificación también da malos resultados y no se usa con frecuencia. Podríamos tener mejores resultados repitiendo varias veces el mensaje, pero esto toma espacio y tiempo.

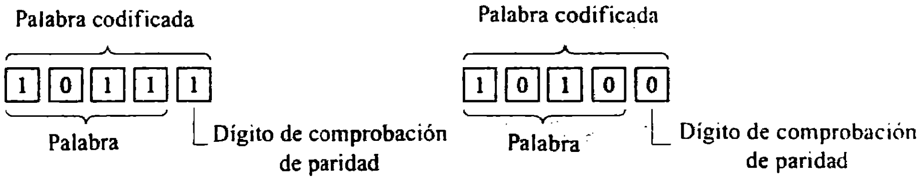


Figura 4.17 Codificación con una verificación de paridad.

Código de Hamming

Espacios vectoriales sobre Z_2

Ahora describiremos un código muy interesante que corrige errores únicos, introducido por R. H. Hamming en la década de los cincuenta. Antes de llegar a los detalles, notaremos que la definición de espacio vectorial puede ampliarse para poder usar escalares que no sean números reales, por ejemplo, números racionales o hasta números complejos. Nos interesa el conjunto de escalares $Z_2 = \{0, 1\}$ que son los *enteros mod* 2. La suma y la multiplicación en Z_2 se definen como sigue:

$$0+0=0, \quad 1+0=1, \quad 0+1=1, \quad 1+1=0$$
$$0\cdot 0=0, \quad 1\cdot 0=0, \quad 0\cdot 1=0, \quad 1\cdot 1=1$$

Como $1 + 1 = 0$, el opuesto de 1 también es 1, y así $-1 = 1$. En consecuencia, *la resta es idéntica a la suma*. Estas operaciones tienen las propiedades adecuadas de adición y multiplicación. Por ejemplo,

$$(1+1)+(1\cdot 0+1)+1\cdot(0+1)=0+1+1=0$$

Sea Z_2^n el conjunto de vectores n cuyos componentes son los elementos de Z_2. Si $n = 3$, Z_2^3 consiste de los 8 vectores siguientes:

$$Z_2^3 = \{(0,0,0),(1,0,0),(0,1,0),(0,0,1),(1,1,0),(1,0,1),(0,1,1),(1,1,1)\}$$

En general, Z_2^n tiene 2^n elementos.

Igual que hicimos con $\mathbf{R}^n$, haremos equivalente a Z_2^n mediante la suma y multiplicación escalar componente por componente con las operaciones en Z_2. En Z_2^4,

$$\begin{aligned}(1,1,0,1)+(0,1,1,0)&=(1,0,1,1)\\ 1(1,0,0,1)&=(1,0,0,1)\\ 0(1,1,1,0)&=(0,0,0,0)\end{aligned}$$

Bajo estas operaciones, Z_2^n cumple con todos los axiomas de un espacio vectorial, con la excepción de que los escalares son de Z_2. Se dice que Z_2^n **es un espacio vectorial sobre** Z_2. Todos los conceptos y propiedades básicos, como subespacios, bases, vectores linealmente independientes, conjuntos generados, reducción de matrices con operaciones de renglón, espacio de columnas, espacio de renglones, rango y nulidad, se aplican a espacios vectoriales sobre Z_2 y a matrices cuyos elementos son de Z_2.

TEOREMA 35

Si V es un subespacio vectorial sobre Z_2 de dimensión n, entonces V tiene 2^n elementos.

DEMOSTRACIÓN Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de V, entonces V está formada por todas las combinaciones lineales diferentes

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \text{ siendo } c_1, \ldots, c_n \text{ ya sea 0 o 1.}$$

Para cada coeficiente hay dos alternativas, de modo que hay un total de 2^n combinaciones distintas. ☐

■ EJEMPLO 81 Determine las bases y la cantidad de elementos del Col(A) y $\nu(A)$ sobre Z_2. ¿Son iguales rango y nulidad, como en el caso en que A se considera como una matriz con elementos reales?

$$A=\begin{bmatrix}1&1&1&0\\1&0&0&1\\0&1&1&1\end{bmatrix}$$

SOLUCIÓN Sobre Z_2 (tenga en cuenta que la reducción se hace con aritmética Z_2)

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Por consiguiente, el rango es 2 y las dos primeras columnas,

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

de A forman una base para su espacio de columnas. El espacio nulo se obtiene igualando la forma escalonada reducida a (0, 0, 0) y despejando las variables delanteras. Si $x_4 = r$ y $x_3 = s$, entonces $x_1 = -r = r$ y $x_2 = -r - s = r + s$, donde $r, s \in \{0, 1\}$. Así, el espacio nulo sobre Z_2 es generado por los vectores

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

y la nulidad es 2 sobre Z_2. Puesto que la dimensión de Col(A) es 2, tiene $2^2 = 4$ elementos. De igual manera, $\nu(A)$ tiene 4 elementos.

Sin embargo, cuando utilizamos **R**, no llegamos a las mismas respuestas. Una forma escalonada reducida de A es

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

de modo que el rango de A es 3 sobre **R**. Por consiguiente, la nulidad sobre **R** es 1, de acuerdo con el teorema del rango. □

La clave (7, 4) de Hamming

Ya estamos listos para definir el interesante código corrector de un solo error de Hamming.[6] Un **código lineal** (n, k) es un subespacio de Z_2^n de dimensión k. Todos los vectores de un código lineal se llaman **palabras (de) código, o palabras codificadas**.

Se tiene la matriz H sobre Z_2,

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Observe que las columnas $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_7$ de H son los vectores de Z_2^3 distintos de cero.

El espacio nulo de H se llama **código (7, 4) de Hamming**. Sea $\nu(H) = \mathrm{N}_H$. Aquí H es la **matriz de comprobación de paridad** para el código N_H. Igual que en el ejemplo 81, puede calcularse fácilmente una base $\mathcal{B}$ para N_H,

[6] Véanse (1) *Error Correcting Codes*, de Wesley Peterson (Cambridge, MA.: MIT Press, 1961; y (2) *Introduction to the Theory of Error Correcting Codes*, de Vera Pless (Nueva York: Wiley, 1982).

$$\mathcal{B} = \{(1,0,0,0,0,1,1),(0,1,0,0,1,0,1),(0,0,1,0,1,1,0),(0,0,0,1,1,1,1)\}$$

Así, N_H es una clave lineal (7, 4) y tiene $2^4 = 16$ vectores. Como $H(\mathbf{e}_i) = \mathbf{h}_i$ para $i = 1, \ldots, 7$, se ve que ninguno de los vectores de base estándar $\mathbf{e}_1, \ldots, \mathbf{e}_7$ de Z_2^7 está en N_H.

La matriz G cuyos renglones son los elementos de $\mathcal{B}$,

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

se llama **matriz generadora** del código (7, 4) de Hamming.

TEOREMA 36

> Sea $\mathbf{v} = (v_1, \ldots, v_7)$ en Z_2^7.
>
> 1. Si $\mathbf{v} \in N_H$, entonces $\mathbf{v} + \mathbf{e}_i \notin N_H$ para $i = 1, \ldots, 7$.
> 2. Si $H\mathbf{v} = \mathbf{h}_j$, entonces $\mathbf{v} + \mathbf{e}_j \in N_H$. Además, $\mathbf{v} + \mathbf{e}_i \notin N_H$, para $i \neq j$.

En otras palabras, si se modifica alguna coordenada de un vector en N_H, el nuevo vector ya no estará en N_H. También, si $H\mathbf{v}$ es la j-ésima columna de H, entonces al cambiar la j-ésima coordenada de $\mathbf{v}$ *sólo* se colocará el nuevo vector en N_H.

DEMOSTRACIÓN

1. como $H\mathbf{v} = \mathbf{0}$,

$$H(\mathbf{v} + \mathbf{e}_i) = H(\mathbf{v}) + H(\mathbf{e}_i) = \mathbf{0} + \mathbf{h}_i = \mathbf{h}_i \neq \mathbf{0}$$

2.

$$H(\mathbf{v} + \mathbf{e}_j) = H(\mathbf{v}) + H(\mathbf{e}_j) = \mathbf{h}_j + \mathbf{h}_j = \mathbf{0}$$

de modo que $\mathbf{v} + \mathbf{e}_i \in N_H$. También,

$$H(\mathbf{v} + \mathbf{e}_i) = H(\mathbf{v}) + H(\mathbf{e}_i) = \mathbf{h}_j + \mathbf{h}_i \neq \mathbf{0}, \qquad i \neq j$$

$\square$

Codificación y decodificación

Veamos ahora cómo codificar un mensaje y decodificar su recepción distorsionada. Se supone que la palabra por codificar es binaria y tiene longitud 4, por ejemplo, 1011, y que el ruido sólo alteró un dígito binario de ella.

Para codificar 1011, se forma la combinación lineal $\mathbf{v}$ en la base $\mathcal{B}$ del código (7, 4) de Hamming, con los coeficientes 1, 0, 1, 1 de nuestro mensaje.

$$\begin{aligned} \mathbf{v} &= 1(1,0,0,0,0,1,1) + 0(0,1,0,0,1,0,1) + 1(0,0,1,0,1,1,0) + 1(0,0,0,1,1,1,1) \\ &= (1,0,1,1,0,1,0) \end{aligned}$$

Es igual a la multiplicación matricial por la derecha por G sobre Z_2.

$$\mathbf{v}^T G = \begin{bmatrix} 1 & 0 & 1 & 1 \end{bmatrix} G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

La palabra codificada $\mathbf{v}$ está en N_H, por construcción. Contiene el mensaje original en los primeros cuatro componentes y al final agrega una especie de comprobación de paridad 0, 1, 0. Suponga que la cadena 1011010 se transmite y se recibe como 0011010. Sea $\mathbf{u} = (0, 0, 1, 1, 0, 1, 0)$. Para corregir el mensaje recibido se calcula el producto $H\mathbf{u}$.

$$H\mathbf{u} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Como $H\mathbf{u}$ es la primera columna de H, enunciado 2 del teorema 36, implica que $\mathbf{u} + \mathbf{e}_1$ está en N_H y nada de $\mathbf{u} + \mathbf{e}_i$, $i \neq 1$, está en N_H. Por consiguiente, $\mathbf{u} + \mathbf{e}_1 = \mathbf{v}$ es el único mensaje codificado y corregido, y se recupera el mensaje original 1011.

■ EJEMPLO 82 Supongamos que se recibieron los mensajes 1010101 y 1100111 codificados con Hamming. Si cuando mucho hay un error en cada transmisión, ¿cuáles fueron los mensajes originales?

SOLUCIÓN Sean $\mathbf{v}_1 = (1, 0, 1, 0, 1, 0, 1)$ y $\mathbf{v}_2 = (1, 1, 0, 0, 1, 1, 1)$.

1. $H\mathbf{v}_1 = (0, 0, 0)$. Entonces $\mathbf{v}_1 \in N_H$. Como el mensaje codificado original ya estaba en N_H, un solo error sacaría a $\mathbf{v}_1$ de N_H, de acuerdo con el enunciado 1 del teorema 36. Así, no hubo error en la transmisión del primer mensaje, que fue 1010.

2. $H\mathbf{v}_2 = (1, 1, 1)$. Así, el séptimo componente de $\mathbf{v}_2$ necesita corregirse a 0, de acuerdo con el enunciado 2 del teorema 36. Entonces, el mensaje original fue 1100. Esta vez el ruido afectó la parte de comprobación de paridad y el mensaje original no se alteró. □

Haremos un resumen del método.

Algoritmo para corrección de error con el código (7, 4) de Hamming

Supongamos que se codifica una palabra $\mathbf{w}$ de 4 dígitos binarios, de modo que $\mathbf{u}$ está en N_H. Si $\mathbf{u}$ se distorsiona a $\mathbf{v} = (v_1, \ldots, v_7)$ cambiando cuando mucho un componente, para recuperar el mensaje original se hace lo siguiente:

ENTRADA: $\mathbf{v}$.

1. Calcule $H\mathbf{v}$.
2. Si $H\mathbf{v} = \mathbf{0}$, sea $\mathbf{w} = v_1 v_2 v_3 v_4$. Deténgase.
3. Si $H\mathbf{v} = \mathbf{h}_i$, cambie el i-ésimo componente de $\mathbf{v}$ para obtener un nuevo vector $\mathbf{v}' = (v_1', \ldots, v_7')$.
4. Sea $\mathbf{w} = v_1' v_2' v_3' v_4'$.

SALIDA: $\mathbf{w}$

Otros tipos de códigos

Nuestra presentación del código (7, 4) de Hamming tenía por objeto ilustrar algunas de las muchas y fructíferas ideas de C. E. Shannon, R. H. Hamming y otros, a finales de la década de los cuarenta y principios de los cincuenta, en los campos de ingeniería eléctrica y teoría de la información. No entraremos en detalles. El código de Hamming sólo es bueno para codificar palabras binarias de longitud 4, de las cuales sólo hay $2^4 = 16$. Si se desea un "alfabeto" mayor, o corregir cuando menos dos errores en un mensaje mezclado, se necesitan otras clases de códigos.

Actualmente, en la práctica se usa una diversidad de técnicas de codificación que permiten codificar más palabras, y en consecuencia mensajes más grandes. También, muchos códigos permiten más errores por ruido que el de Hamming. Una gran parte de estos interesantes códigos son **no lineales**. Las definiciones y los ejemplos se encuentran en los textos sobre el tema.

En el estudio de los códigos, los principales aportes son de matemáticas, y en primera línea están el álgebra lineal, la teoría de los números y la teoría de campos.

Ejercicios 4.7

Aritmética Z_2^n

Sean

$$\mathbf{u} = \begin{bmatrix}1\\0\\1\end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix}0\\1\\1\end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix}1\\1\\0\end{bmatrix}$$

$$A = [\mathbf{u}\ \mathbf{v}\ \mathbf{w}] = \begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\\1 & 1 & 0\end{bmatrix}$$

1. Efectúe las operaciones indicadas en Z_2^3.

 a. $\mathbf{u} + \mathbf{v}$ b. $-\mathbf{v}$

 c. $1\mathbf{u} + 0\mathbf{v} - 1\mathbf{w}$ d. $\mathbf{u} + \mathbf{v} + \mathbf{w}$

2. Resuelva la ecuación para $\mathbf{x}$ sobre Z_2.

$$\mathbf{x} - \mathbf{u} + \mathbf{v} = \mathbf{w} + \mathbf{u}$$

3. Calcule A^2 y A^3 sobre Z_2.
4. ¿Es $\{\mathbf{u}, \mathbf{v}\}$ linealmente independiente sobre Z_2? ¿Y qué pasa con $\{\mathbf{u}, \mathbf{w}\}$?
5. ¿Es $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ linealmente independiente sobre Z_2?
6. Agregue un vector a $\{\mathbf{u}, \mathbf{v}\}$ para que el conjunto resultante sea una base de Z_2^3.
7. Determine una base y los vectores del espacio nulo de A sobre Z_2. Repita sobre $\mathbf{R}$.
8. Calcule la inversa de $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 1\\1 & 1 & 0\end{bmatrix}$ sobre Z_2 y compruebe su respuesta.
9. Sean A y B matrices binarias conmutativas 2×2. Demuestre que sobre Z_2:

$$(A + B)^2 = A^2 + B^2.$$

Códigos

10. Codifique el mensaje 1110.
11. Codifique el mensaje 0101.
12. Codifique el mensaje 0010.

En los ejercicios 13 a 17 suponga que se codificó una palabra de mensaje con el método de Hamming. Durante la transmisión se altera cuando mucho una coordenada. Recupere el mensaje original a partir del vector binario que se recibió en cada caso.

13. a. (1, 1, 1, 1, 0, 1, 1) b. (1, 1, 1, 1, 1, 0, 0)
14. a. (0, 1, 1, 1, 1, 0, 1) b. (0, 1, 1, 1, 1, 0, 0)
15. a. (0, 1, 1, 0, 0, 0, 1) b. (0, 1, 0, 0, 0, 1, 1)
16. a. (0, 1, 1, 0, 0, 1, 1) b. (1, 1, 1, 1, 0, 0, 0)
17. a. (1, 1, 1, 0, 0, 1, 0) b. (1, 1, 1, 0, 0, 0, 0)
18. Escriba los 16 elementos del código (7, 4) de Hamming N_H.

El **peso** $w(\mathbf{v})$ de un vector $\mathbf{v}$ en Z_2^n es igual a la cantidad de sus elementos distintos de cero. Por ejemplo,

$$w(0, 1, 1, 0) = 2 \quad \text{y} \quad w(1, 0, 1, 1, 1) = 4$$

La **distancia** $d(\mathbf{u}, \mathbf{v})$ entre dos vectores $\mathbf{u}$ y $\mathbf{v}$ en Z_2^n es igual a la cantidad de elementos en los que difieren ambos. En consecuencia,

$$d(\mathbf{u}, \mathbf{v}) = w(\mathbf{u} - \mathbf{v}) = w(\mathbf{u} + \mathbf{v})$$

19. Demuestre que

$$d(\mathbf{u}, \mathbf{v}) = d(\mathbf{0}, \mathbf{u} - \mathbf{v})$$

20. Compruebe que $w(\mathbf{v}) \geq 3$ para todos los vectores $\mathbf{v}$ distintos de cero en N_H. (*Sugerencia:* Use el ejercicio 18.)

21. Demuestre que $d(\mathbf{u} - \mathbf{v}) \geq 3$ para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ distintos en N_H. (*Sugerencia:* Use el ejercicio 20.)

También se pueden definir códigos de detección de error en términos de la función distancia d. Un código lineal $V \subseteq Z_2^n$ es **detector de un solo error** si para toda palabra de código $\mathbf{v} \in V$ y para cualquier vector que está $\mathbf{u}$ en Z_2^n, la relación $d(\mathbf{u}, \mathbf{v}) \leq 1$ implica que $\mathbf{u}$ no es palabra de código a menos que $\mathbf{v} = \mathbf{u}$.

22. Aplique el ejercicio 21 para verificar que N_H es detectora de un solo error, de acuerdo con la definición anterior. También, demuestre que esta afirmación equivale al enunciado 1 del teorema 36.

4.8 Miniproyectos

El enfoque de esta sección de proyectos es describir y generalizar todavía más a los espacios vectoriales. En la sección 4.7 definimos los espacios vectoriales sobre Z_2. Ahora permitiremos tipos más generales de escalares, que sean elementos de un *campo*.

1 ■ Campos

Definición

Un **campo** F es un conjunto de elementos llamados **escalares**, para el cual se definen dos operaciones, suma $(a + b)$ y multiplicación (ab) que satisfacen las propiedades siguientes:

Suma:

(A1) $a + b$ pertenece a F para todas $a, b \in F$.
(A2) $a + b = b + a$ para todas $a, b \in F$.
(A3) $(a + b) + c = a + (b + c)$ para todas $a, b, c \in F$.
(A4) Existe un escalar único $0 \in F$, llamado el **cero de** F, tal que para toda a en F,

$$a + 0 = a$$

(A5) Para cada $a \in F$, existe un escalar único $-a$, llamado **negativo** u **opuesto de** a, tal que

$$a + (-a) = 0$$

Multiplicación:

(M1) ab pertenece a F para todas $a, b \in F$.
(M2) $(a + b)c = ac + bc$ para todas $a, b, c \in F$.
(M3) $ab = ba$ para todas $a, b \in F$.

(M4) $(ab)c = a(bc)$ para todas $a, b, c \in F$.
(M5) Existe un escalar único *distinto de cero*, $1 \in F$, llamado **uno**, tal que para toda a en F,

$$a1 = a$$

(M6) Para toda $a \in F$, $a \neq 0$, existe un escalar único a^{-1} (o $1/a$), llamado el **inverso** o **recíproco** de a, tal que

$$aa^{-1} = 1$$

Se acostumbra escribir $a - b$ para indicar la suma $a + (-b)$:

$$a - b = a + (-b)$$

Problema A

Compruebe que en un campo F,

$$\text{si } ab = 0, \text{ entonces } a = 0 \text{ o } b = 0$$

Problema B

Demuestre que los siguientes son campos. En cada caso use la suma, multiplicación y el recíproco usuales.

1. El conjunto **R** de los números reales
2. El conjunto **Q** de los números racionales
3. El conjunto **C** de los números complejos
4. El conjunto Z_2 de los enteros mod 2
5. El conjunto $\mathbf{Q}(\sqrt{2})$ de todos los números de la forma $a + b\sqrt{2}$, siendo a y b números racionales

Sugerencia: Para $\mathbf{Q}(\sqrt{2})$, el recíproco de $a + b\sqrt{2}$ puede escribirse en la forma $A + B\sqrt{2}$, multiplicando y dividiendo $1/(a + b\sqrt{2})$ por el **conjugado** $a - b\sqrt{2}$. Por ejemplo, el inverso de $1 - 3\sqrt{2}$ es

$$\frac{1}{1-3\sqrt{2}} = \frac{1+3\sqrt{2}}{(1-3\sqrt{2})(1+3\sqrt{2})} = \frac{1+3\sqrt{2}}{-17} = -\frac{1}{17} - \frac{3}{17}\sqrt{2}$$

Problema C

Explique por qué los conjuntos siguientes **no son** campos. En cada caso, use la suma y multiplicación usuales.

1. El conjunto **Z** de los enteros
2. El conjunto **N** de los enteros positivos
3. El conjunto $\mathbf{R}^2$ con la suma acostumbrada, componente por componente $(a, b) + (a', b') = (a + a', b + b')$, y la multiplicación componente por componente $(a, b)(a', b') = (aa', bb')$

2 ■ Espacios vectoriales generales

Un espacio vectorial V sobre un campo F es un conjunto no vacío con dos operaciones: suma y multiplicación por escalar, que satisfacen todos los axiomas de un espacio vectorial, definidos en la sección 4.2, con la excepción que todos los escalares provienen del campo F en lugar de los números reales **R**. Los elementos también se llaman **vectores**. Si el campo F es **R**, se dice que V es un espacio vectorial **real**. Si F = **Q**, el conjunto de números racionales, se dice que V es un espacio vectorial **racional**. Si F = **C**, el conjunto de los números complejos, se dice que V es un espacio vectorial **complejo**.

Representaremos al conjunto de todos los pares ordenados (a, b) mediante F^2, donde a y b son cualesquiera elementos de F. En general, denotaremos con F^n el conjunto de todos los arreglos ordenados de orden n $(a_1, \ldots, a_n)$, siendo $a_1, \ldots, a_n$ cualesquiera elementos de F. F^n tiene definida a la suma y a la multiplicación por escalar componente por componente.

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$
$$c(a_1, \ldots, a_n) = (ca_1, \ldots, ca_n)$$

Problema A

Demuestre que los siguientes conjuntos son espacios vectoriales sobre el campo especificado F.

1. **Q** sobre **Q**
2. $\mathbf{Q}^n$ sobre **Q**
3. **C** sobre **C**
4. $\mathbf{C}^n$ sobre **C**
5. Cualquier campo F sobre F
6. F^n sobre F

Problema B

Compruebe que los conjuntos siguientes son espacios vectoriales sobre el campo F especificado.

1. Los números reales **R** sobre el conjunto de los números racionales **Q**. La suma es la acostumbrada $r_1 + r_2$, $r_1, r_2 \in \mathbf{R}$. La multiplicación por escalar tiene la forma qr, siendo q un número racional y r real.
2. Los números complejos **C** sobre el conjunto **R** de los números reales. La suma es la $z_1 + z_2$ acostumbrada, $z_1, z_2 \in \mathbf{C}$. La multiplicación por escalar tiene la forma rz, en la que r es un número real y z es complejo.
3. El conjunto $\mathbf{Q}(\sqrt{2})$ de todos los números reales de la forma $a + b\sqrt{2}$, $a, b \in \mathbf{Q}$ sobre **Q**.

Problema C

Determine la *dimensión* de los espacios vectoriales dados sobre el campo F especificado.

1. **C** sobre **C**
2. $\mathbf{C}^2$ sobre **C**
3. **C** sobre **R**
4. F^n sobre F
5. $\mathbf{Q}(\sqrt{2})$ sobre **Q**

3 ■ Espacios vectoriales sobre campos finitos

En este párrafo definiremos algunos campos interesantes formados por una cantidad finita de elementos y algunos espacios vectoriales definidos sobre ellos.

Z_p: Los enteros mod p

Un número primo p es un entero positivo cuyos únicos divisores son 1 y p. Por ejemplo,

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29$$

son números primos. Sin embargo,

$$4, 6, 8, 9, 10, 12, 14, 15, 16, 18$$

no son números primos, porque cada uno tienen otros divisores además de ellos mismos y de 1.

Sea p un número primo. El conjunto Z_p de los enteros mod p está formado por los p elementos $\{0, 1, \ldots, p-1\}$. En Z_p se definen la suma y la multiplicación como sigue.

Si a y b están en Z_p, entonces $a + b$ es el residuo mínimo positivo que se obtiene si se divide el entero $a + b$ entre p. Por ejemplo, si $p = 5$, $Z_5 = \{0, 1, 2, 3, 4\}$, y $2 + 3$ produce el residuo 0 cuando se divide entre 5, de modo que $2 + 3 = 0$ en Z_5. También, $3 + 4 = 2$ en Z_5, porque $7 = 5 \cdot 1 + 2$ tiene un residuo igual a 2 al ser dividido entre 5.

Si a y b están en Z_p, entonces ab se define en la misma forma, es decir, el residuo mínimo positivo que se obtiene al dividir el entero ab entre p. Por ejemplo, en Z_5, $2 \cdot 3 = 1$, porque $6 = 5 \cdot 1 + 1$. De igual manera, $3 \cdot 3 \cdot 3 \cdot 3 = 1$, porque $81 = 16 \cdot 5 + 1$.

Las operaciones que acabamos de definir se llaman operaciones **mod p** de los enteros.

Ahora practicaremos la aritmética Z_p determinando los opuestos de todos los elementos de Z_5 y los recíprocos de todos sus elementos distintos de cero. Es claro que el opuesto de 0, -0, es 0, porque $0 + 0 = 0$. El opuesto de -1 es 4, porque $1 + 4 = 0$ en Z_5. Puede escribirse $-1 = 4$. De igual manera, $-2 = 3$, $-3 = 2$, $-4 = 1$. El recíproco de 1 es 1, porque $1 \cdot 1 = 1$. El recíproco de 2 es 3, porque $2 \cdot 3 = 1$ en Z_5. También se escribe $\frac{1}{2} = 3$. De igual forma, $\frac{1}{3} = 2$ y $\frac{1}{4}4 = 4$.

Problema A

1. Determine -1 en Z_7.
2. Determine -10 en Z_{17}.
3. Determine $\frac{1}{3}$ en Z_{17}.
4. Determine $\frac{1}{6}$ en Z_7.
5. Determine $\frac{1}{10}$ en Z_{11}.
6. Determine $\frac{1}{p-1}$ en Z_p (p es primo).

Si m es cualquier entero positivo, $Z_m = \{0, \ldots, m-1\}$, los enteros mod m, se define en la misma forma que Z_p y se le asignan las mismas operaciones mod m.

Problema B

Con respecto a las operaciones mod p, demuestre cada una de las siguientes afirmaciones:

1. Z_3 es un campo.
2. Z_7 es un campo.
3. Z_p es un campo (p es primo).
4. Z_4 **no** es un campo. (*Sugerencia:* ¿Existe $\frac{1}{2}$?)
5. Si m no es entero primo, Z_m no es un campo.

Como Z_p es un campo para cualquier primo p, se puede hablar de espacios vectoriales sobre Z_p. Como F^n es un espacio vectorial sobre F, según lo que se determinó en el proyecto 2, Z_p^n, el conjunto de los arreglos ordenados de orden n $(a_1, \ldots, a_n)$ con $a_i \in Z_p$, es un espacio vectorial sobre Z_p, de dimensión n.

Problema C

Para el espacio vectorial Z_p^n sobre Z_p:

1. Demuestre que Z_p^n tiene p^n elementos.
2. Si V es un subespacio de Z_p^n de dimensión m, entonces V tiene p^m elementos.
3. Determine una base para Z_3^2.
4. Sea

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 4 & 3 & 2 \\ 3 & 3 & 1 \end{bmatrix}$$

una matriz con elementos en Z_5. Reduzca A con operaciones elementales de renglón empleando sólo mod 5.

5. Determine las bases para los espacios nulo y de columnas de A sobre Z_5. Compruebe el teorema del rango.

4.9 Ejercicios en computadora

Esta sesión en la computadora ayudará al lector a dominar los comandos de sus programas que se relacionan con los temas de este capítulo. Además, permitirá repasar algunos de los conceptos básicos.

$$M = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \mathbf{v}_4\ \mathbf{v}_5\ \mathbf{v}_6] = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix},$$

$$N = [\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3\ \mathbf{u}_4\ \mathbf{u}_5] = \begin{bmatrix} 1 & 2 & 1 & 1 & 1 \\ 2 & 4 & 3 & 4 & 5 \\ 3 & 6 & 5 & 7 & 9 \\ 4 & 8 & 7 & 10 & 13 \end{bmatrix}$$

y sean

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\}, \qquad T = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}, \qquad \mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{u}_3, \mathbf{u}_4\}$$

1. ¿Es S una base de $\mathbf{R}^3$?
2. ¿Es T una base de $\mathbf{R}^4$?
3. Demuestre que $\mathcal{B}$ es una base de $\mathbf{R}^4$.
4. Determine $[\mathbf{u}_1]_{\mathcal{B}}$ y $[\mathbf{u}_2]_{\mathcal{B}}$.
5. Si $[\mathbf{x}]_{\mathcal{B}} = \mathbf{u}_5$, ¿qué es $\mathbf{x}$?
6. Determine una base para Gen$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.
7. Calcule una base para Gen$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$
8. Determine las bases para el Col(M), el Ren(M) y $\nu(M)$.
9. Calcule el rango y la nulidad de M. Compruebe el teorema del rango.

10. Verifique que Rango(M) = Rango(M^T).
11. Determine bases para el Col(N) Ren(N)y $\nu(N)$.
12. Calcule el rango y la nulidad de N. Compruebe el teorema del rango.
13. Determine dos bases para Gen$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_1, \mathbf{e}_2\}$ empleando los algoritmos A y B de la sección 4.6.
14. Calcule las columnas pivote de N y demuestre que son linealmente independientes. Demuestre que la cuarta columna es una combinación lineal de las columnas pivote precedentes.
15. Amplíe el conjunto linealmente independiente $\{\mathbf{u}_2, \mathbf{u}_3\}$ a una base de $\mathbf{R}^4$.
16. Sean

$$a = -7, \qquad b = 2, \qquad \mathbf{u} = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix},$$

$$\mathbf{v} = \begin{bmatrix} -1 & 1 \\ 0 & 8 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} 2 & 1 \\ 0 & -9 \end{bmatrix}$$

Verifique la validez de los axiomas (A2), (A3), (M2), (M3) y (M4) para a, b $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$. ¿Por qué el conjunto de todas las matrices de la forma $\begin{bmatrix} c_1 & c_2 \\ 0 & c_3 \end{bmatrix}$ es un subespacio de M_{22}?

17. Defina la función de tres variables, $f(a, b, x) = a\cos(3x) + b\,\text{sen}(2x)$ que representa las combinaciones lineales de $\cos(3x)$ y $\text{sen}(2x)$ en $F(\mathbf{R})$. Use f para trazar en una gráfica las combinaciones lineales con $\{a = 1, b = 1\}$ $\{a = 3, b = 0\}$, $\{a = 0, b = -3\}$, $\{a = 3, b = -4\}$ y $\{a = -3, b = 4\}$.
18. Demuestre que el conjunto V de los polinomios de la forma $ax^3 + bx^2$, $a, b \in \mathbf{R}$ es un subespacio de P_3.
19. Compruebe que los conjuntos de polinomios siguientes forman bases de P_3.

$$\mathcal{B}_1 = \{1, -x + 1, x^2 - x, -x^3 + x^2 - 1\}$$

$$\mathcal{B}_2 = \{x + 2, 4x^2 - x, x^3 - x, x^2 + 1\}$$

20. Determine la matriz de transición P de $\mathcal{B}_1$ a $\mathcal{B}_2$.
21. Calcule la matriz Q de transición de $\mathcal{B}_2$ a $\mathcal{B}_1$.
22. Demuestre que $P = Q^{-1}$.
23. Compruebe que $[-x^3 + 5x + 1]_{\mathcal{B}_2} = P[-x^3 + 5x + 1]_{\mathcal{B}_1}$.
24. Determine una base para el conjunto V de polinomios de la forma $ax^3 + bx^2 + cx + d$ tal que

$$a + 2c = 0, \qquad 3b - d = 0$$

25. Determine a tal que los polinomios formen una base de P_2.

$$x + a, \qquad ax^2 + ax + 1, \qquad ax^2 + x + 1 - a$$

Soluciones seleccionadas con Maple

Comandos augment, col, collect, colspace, evalm, expand, gausselim, matrix, nullspace, plot, proc, rowspace, rref, stack, submatrix, transpose, vector.

```
# Datos.
With (linalg):
v1 : = vector ([1 , 2 , 3]) ; v2 : = vector ([2 , 3 , 4]) ; v3 : = vector ([3 , 4 , 5]) ;
v4 : = vector ([4 , 5 , 6]) ; v5 : = vector ([5 , 6 , 7]) ; v6 : = vector ([6 , 7 , 8]) ;
```

```
u1 := vector ([1, 2, 3, 4]); u2 := vector ([2, 4, 6, 8]);
u3 := vector ([1, 3, 5, 7]); u4 := vector ([1, 4, 7, 10]);
u5 := vector ([1, 5, 9, 13]);
e1 := vector ([1, 0, 0, 0]); e2 := vector ([0, 1, 0, 0]);  # Vectores de base estándar.
e3 := vector ([0, 0, 1, 0]); e4 := vector ([0, 0, 0,  1]);
M := augment (v1, v2, v3, v4, v5, v6);
N := augment (u1, u2, u3, u4, u5);
B := augment (e1, e2, u3, u4);  # Matriz cuyas columnas son los vectores de B.

# Ejercicios 1 a 7.

rref (M);     # 2 pivotes, 3 renglones: no genera a R^3. No es base.
rref (N);     # 2 pivotes, 4 renglones: no genera a R^4. No es base.
rref (B);     # 4 pivotes, 4 renglones, 4 columnas: generador y
              # linealmente independiente: Base de R^4.
rref (augment (B, u1);     # Se resolverá el sistema [B:u1] por reducción. La última
col (", 5);                # columna expresa las coordenadas de u1. Repetir con u2.
evalm (B &* u5):           # Ejercicio 5: x es justamente Bu5.
rref (augment (v1, v2, v3));          # Pivotea en (1, 1), (2, 2). Así, {v1, v2} es una base.
rref (augment (u1, u2, u3, u4) );     # Pivotea en (1,1),  (2,3). Así, {u1,u3} es una base.

# Ejercicios 8 a 12.

colspace (M);              # Base para espacio de columnas. 2 vectores, así que rango = 2.
colspace (M, 'r'); r;      # Otra forma que produce el mismo rango.
rowspace (M);              # Base para espacio de renglones. También rowspace(M,'r'); r;
nullspace (M);             # Base para el espacio nulo. 4 vectores, así que nulidad = 4.
nullspace (M, 'n ' ); n;   # Otra forma que produce la nulidad.
n + r;                     # Es igual al número de columnas. Teorema del rango.
colspace (transpose (M), 'r1'   # El rango de la transpuesta también es 2.
# Repetir con N.

# Ejercicio 13.

augment (u1, u2, e1, e2);        # Algoritmo A: La matriz [u1 u2 e1 e2].
rref (");                        # Pivotea en (1,1), (2,3), (3,3). {u1,e1,e2} es una base.
stack (u1, u2, e1, e2);          # Algoritmo B. Los vectores como renglones de una matriz.
gausselim (")                    # Los renglones no cero (los primeros 3) forman una base.
rref (")                         # La reducción completa obtiene todavía más ceros.

# Ejercicio 14.

rref (N);                                    # Las columnas pivote son 1 y 3.
pcol = augment (col (N, 1) col (N , 3));     # Reducción 2 pivotes 2 columnas,
rref (pcol);                                 # así que las columnas pivote son independientes.
rref (augment (pcol, col (N, 4)));           # Reducir [col1, col3, col4] para obtener (-1,2,0,0)
evalm (-1*col (N, 1) + 2*col (N, 3) - col (N, 4));  # como última columna para que -1*col1+2*col3=col4.

# Ejercicio 15.

m := augment (u1, u2, e1, e2, e3, e4);       # Columnas pivote 1, 3, 4, 5. Así, base:
rref (m);                                    # {u1, e1, e2, e3}.

# Ejercicio 16.

a := - 7 : b : = 2 : u : matrix (2, 2, [1, 3, 0, 2] ):               # Introducción de escalares
v := matrix (2, 2, [- 1, 1, 0, 8] : w := matrix (2 , 2, [2, 1, 0, - 9] :   # y matrices.
evalm ( (u+v) - (v+u) );                                            # A(2),
```

```
evalm ( (u+v) +w- (u+ (v+w)));                          # A(3),
evalm (a* (u+v) - (a*u+a*v));                           # M(2),
evalm ( (a+b) *u - (a*u+b*u))                           # M(3) y
evalm ( (a*b) *u - a*(b*u));                            # M(4) se aplican.
```

```
# Ejercicio 17.
f := proc (a, b, x) a*cos (3*x) + b*sin (2*x) end:      # Definición de f.
Plot ({f (1, 1, x), f(3, 0, x), f(0, -3, x), f (3, -4 , x) , f(4 , -3 , x)}, x=0. .pi) :  # Graficación.
```

```
# Ejercicio 18.
p1 := a1*x^3 + b1*x^2 : p2 := a2*x^3 + b2*x:        # p general.
collect (p1+p2, x);               # La suma p1 + p2 está en V.
collect (expand (c*p1), x);       # c*p1 está en V. Entonces V es un subespacio.
```

```
# Ejercicio 19.
B1 .= matrix ([[0, 0, 0, 1], [0, 0, -1, 1], [0, 1, -1, 0], [-1, 1, 0, -1]]);
B2 := matrix ( [ [0, 0, 1, 2], [0, 4, -1, 0], [1, 0, -1, 0], [0, 1, 0, 1]]);
rref (B1);      # Todas las matrices tienen 4 renglones pivote, de modo que los polinomios son
rref (B2)       # linealmente independientes, luego forman una base, porque dim(P_3) = 4.
```

```
# Ejercicios 20 a 23.
P1 := rref (augment (B2 , B1));                # Reducir [B1:B2] y conservar las últimas 4
P : submatrix (P , 1..4 , 5 . . 8);            # columnas para obtener P.
Q1 := rref(augment (B1, B2));                  # Repetir con Q.
Q : = submatrix (Q1, 1.. 4, 5.. 8);
evalm (P &* Q) ;                               #PQ = 1 tienen el mismo tamaño, de modo que P^(-1) = Q.
p : = vector ([-1, 0, 5, 1]);                  # En un vector de la forma x^3+5x+1.
pb1 : = col (rref (augment (B1, p)) , 5) ;     # [p]_B1.
pb2 : = col (rref (augment (B2, p)) , 5) ;     # [p]_B2.
evalm (P &* pb1) ;                             # P [p]_B2 produce [p]_B1 como se esperaba.
```

Soluciones seleccionadas con Mathematica

Comandos AppendRows, Clear, Collect, Expand, Flatten, NullSpace, Plot, RowReduce, TakeColumns.

```
(* Datos. *)
<<LinearAlgebra 'MatrixManipulation';
v1 = {{1}, {2}, {3}}; v2 = {{2}, {3}, {4}};, v3 = {{3}, {4}, {5}};
v4 = {{4}, {5}, {6}}; v5 = {{5}, {6}, {7}}; v6 = {{6}, {7}, {8}};
u1 = {{1}, {2}, {3}, {4}}; u2 = {{2}, {4}, {6}, {8}};
u3 = {{1}, {3}, {5}, {7}}; u4 = {{1}, {4}, {7}, {10}};
u5 = {{1}, {5}, {9}, {13}};
e1 = {{1}, {0}, {0}, {0}}; e2 = {{0}, {1}, {0}, {0}};   (* Vectores de base estándar. *)
e3 = {{0}, {0}, {1}, {0}}; e4 = {{0}, {0}, {0}, {1}};
M  = AppendRows[v1, v2, v3, v4, v5, v6]
n  = AppendRows[u1, u2, u3, u4, u5]    (* N ya la usa el programa. *)
B  = AppendRows[e1, e2, u3, u4]        (* La matriz cuyas columnas son los vectores de B. *)
(* Ejercicios 1 a 7. *)
RowReduce[M]     (* 2 pivotes, 3 renglones: no genera a R^3; no es base. *)
```

```
RowReduce[n]              (* 2 pivotes, 4 renglones; no genera a R^4. No es base. *)
RowReduce[B]              (* 4 pivotes, 4 renglones, 4 columnas: generador y *)
                          (* linealmente independiente: base de R^4. *)
RowReduce[AppendRows[B, u1]]  (* Resolveremos el sistema {B:u1} por reducción. *)
TakeColumns[%, {5}]       (* La última columna produce las coordenadas de u1. Repetir con u2. *)
B . u5                    (* Ejercicio 5: x es justamente Bu5. *)
RowReduce[AppendRows[v1, v2, v3]]       (* Pivotea en (1, 1), (2, 2). Así, {v1, v2} es una base. *)
RowReduce[AppendRows[u1, u2, u3, u4]]   (* Pivotea en (1, 2), (2, 3). Así, {u1, u3} es una base. *)
(* Ejercicios 8 a 12. *)
RowReduce[M]              (* Cols. pivote 1, 2, así {v1, v2} es una base. Rango = 2. *)
RowReduce[Transpose[M]]   (* Los dos primeros renglones forman una base para el espacio de renglones. *)
NullSpace[M]              (* Base para el espacio nulo. 4 vectores, así nulidad=4. *)
(* Nulidad + rango = 4 + 2 es igual al número de columnas, 6. El teorema del rango se comprueba. *)
(* De nuevo para el ejercicio 11: RowReduce[Transpose[M]]. *)
(* Repetir con N. *)
(* Ejercicio 13. *)
AppendRows[u1, u2, e1, e2]  (* Algoritmo A: la matriz {u1, u2, e11, e2}. *)
RowReduce[%]                (* Pivotes en (1, 1), (2, 3), (3, 3). {u1, e1, e2} es una base. *)
{Flatten[u1], Flatten[u2], Flatten[e1], Flatten[e2]}   (* Algoritmo B. Los vectores *)
RowReduce[%]   (* como renglones de una matriz. Los renglones no cero (los primeros 3) forman una
base. *)
(* Ejercicio 14. *)
RowReduce[n]                          (* Las columnas pivote son 1 y 3. *)
pcol = AppendRows[TakeColumns[n, {1}], TakeColumns[n, {3}]]      (* Reducción 2 pivotes *)
RowReduce[pcol]               (* 2 columnas, así, las columnas pivote son independientes. *)
RowReduce[AppendRows[TakeColumns[n, {1}],\
                                      (* Reducir [col1, col3, col4] para obtener (-1, 2, 0, 0) *)
TakeColumns[n, {3}], TakeColumns[n, {4}]]]   (* como úlltima columna. Así -1* col 1+2* col3 = col4. *)
-1 TakeColumns[n, {1}]+2 TakeColumns[n, {3}]-TakeColumns[n, {4}]   (* Verificación *)
(* Ejercicio 15. *)
m = AppendRows[1, u2, e1, e2, e3, e4]      (* Columnas pivote 1, 3, 4, 5. Así, la base es: *)
RowReduce[m]                               (* {u1, e1, e2, e3}. *)
(* Ejercicio 16. *)
a=-7; b=2; u={{1, 3}, {0, 2}};           (* Captura de escalares *)
v={{-1, 1}, {0, 8}}; w={{2, 1}, {0, -9}};  (* y matrices. *)
(u+v) - (v+u)                           (* A(2), *)
(u+v)+w - (u+(v+w))                     (* A(3), *)
a(u+v) - (a u+a v)                      (* M(2), *)
(a+b) u - (a u+b u)                     (* M(3), y *)
(a b) u - a (b u)                       (* M(4) son válidas. *)
(* Ejercicio 17. *)
Clear[a, b]                          (* Borrar los valores de a y b. *)
f[a_, b_, x_] = a Cos[3 x] + b Sin[2 x]  (* Definición de f. *)
Plot[{f[1, 1, x], f[3, 0, x], f[0, -3, x], f[3, -4, x], f[4, -3, x]}, {x, 0, Pi}]  (* Graficado. *)
(* Ejercicio 18. *)
p1 = a1 x^3 + b1 x^2;     p2 = a2 x^3 + b2 x^2;    (* Polinomios generales de V. *)
Collect[p1+p2, x]         (* La suma p1 + p2 está en V. *)
```

```
Collect[Expand[c p1], x]    (* c* p1 está en V, Así, V es un subespacio. *)
(* Ejercicio 19. *)
B1 = {{0, 0, 0, 1}, {0, 0, -1, 1}, {0, 1, -1, 0}, {-1, 1, 0, -1}}
B2 = {{0, 0, 1, 2}, {0, 4, -1, 0}, {1, 0, -1, 0}, {0, 1, 0, 1}}
RowReduce[B1]   (* Todas las matrices tienen 4 renglones pivote, así los polinomios son linealmente *)
RowReduce[B2]   (* independientes; por tanto forman una base, ya que dim(P_3)=4. *)
(* Ejercicios 20 y 23. *)
P1=RowReduce[AppendRows[B2, B1]]   (* Reducir [B1:B2] y mantener las *)
P=TakeColumns[P1, {5, 8}]          (* últimas 4 columnas para obtener P. *)
Q1=RowReduce[AppendRows[B1, B2]]   (* Repetir con Q. *)
Q=TakeColumns[Q1, {5, 8}]
P . Q                              (* PQ = I y del mismo tamaño, así que P^(-1) = Q. *)
p={{-1}, {0}, {5}, {1}}            (* x^3+5x+1 en forma vectorial. *)
pb1=TakeColumns[RowReduce[AppendRows[B1, p]], {5}]   (* [p]_B1. *)
pb2=TakeColumns[RowReduce[AppendRows[B2, p]], {5}]   (* [p]_B2. *)
P . pb1   (* P[p]_B2 produce [p]_B1 como se esperaba. *)
```

Soluciones seleccionadas con MATLAB

NOTA La indicación (ST) quiere decir que el comando está en "symbolic toolbox".

Comandos Colspace, fplot function, null, nullspace, rref.

```
% Datos.
v1 = [1; 2; 3]; v2 = [2; 3; 4]; v3 = [3; 4; 5];
v4 = [4; 5; 6]; v5 = [5; 6; 7]; v6 = [6; 7; 8];
u1 = [1; 2; 3; 4]; u2 = [2; 4; 6; 8];
u3 = [1; 3; 5; 7]; u4 = [1; 4; 7; 10];
u5 = [1; 5; 9; 13];
e1 = [1; 0; 0; 0]; e2 = [0; 1; 0; 0];    % Vectores de base estándar.
e3 = [0; 0; 1; 0]; e4 = [0; 0; 0; 1];
M = [v1 v2 v3 v4 v5 v6]
N = [u1 u2 u3 u4 u5]
B = [e1 e2 u3 u4];    % La matriz cuyas columnas son los vectores de B.
% Ejercicios 1 a 7.
rref (M)              % 2 pivotes, 3 renglones: no envuelve R^3. No es base.
rref (N)              % 2 pivotes, 4 renglones: no envuelve R^4. No es base.
rref (B)              % 4 pivotes, 4 renglones, 4 columnas: generador y
                      % linealmente independiente: Base de R^4.
rref ([B u1])         % Resolveremos el sistema [B:u1] por reducción. La última
ans (: , 5)           % columna expresa las coordenadas de u1. Repetir con u2.
B * u5                % Ejercicio 5: x es tan sólo Bu5.
rref ([v1 v2 v3]      % Pivotes en (1,1),(2,2). Así, {v1,v2} es una base.
rref ([u1 u2 u3 u4])  % Pivotes en (1,1),(2,3). Así, {u1,u3} es una base.
% Ejercicios 8 a 12.
rref (M)              % Columnas pivote 1,2, así {v1,v2} es base. Rango = 2.
colspace (M)          % (ST) Otra forma. Base de espacio de columnas con ST.
```

```
rank (M)                 % El rango es 2.
rref (M')                % Los dos primeros renglones forman una base para el espacio de columnas.
null (M)                 % Base para el espacio nulo. 4 vectores, así la nulidad = 4.
nullspace (M)            % (ST) Otra base usando ST.
% Nulidad + rango = 4 + 2 es igual a 6, el número de columnas. El teorema del rango está bien.
rank (M')                % El rango de la transpuesta también es 2.
% Repetir con N.

% Ejercicio 13.

[u1 u2 e1 e2]                      % Algoritmo A: La matriz [u1 u2 e1 e2].
rref (ans)                         % Pivotes en (1,1), (2, 3), (3,3). {u1,e1,e2} es una base.
[u1'; u2'; e1'; e2']               % Algoritmo B. Los vectores como renglones de una matriz.
rref (ans)                         % Los renglones no cero (los primeros 3) forman una base.

% Ejercicio 14.

rref (N)                           % Las columnas pivote son 1 y 3.
pcol = [N (; , 1), N (; , 3)]      % Reducción 2 pivotes 2 columnas,
rref (pcol)                        % así las columnas pivote son independientes.
rref ([pcol N(: , 4)]              % Reducir [col1,col3,col4] para obtener (-1,2,0,0)
- N (:, 1) + 2 * N (: , 3) - N (:, 4)   % como última columna, así -1*col1+2*col3=col4.

%Ejercicio 15.

m = [u1 u2 e1 e2 e3 e4]            % Columnas pivote 1, 3, 3, 5. Entonces la base:
rref (m)                           % {u1,e1,e2,e3}.

% Ejercicio 16.

a = -7; b = 2; u = [1 3; 0 2];     % Introducción de escalares
v = [-1 1; 0 8]; w = [2 1; 0 -9];  % y matrices.
(u + v) - (v + u)                  % A(2),
(u + v) + w - (u + (v + w))        % A(3),
a* (u + v) . (a*u + a*v)           % M(2),
(a + b) * u - (a*u + b*u)          % M(3), y
(a*b) *u - a*(b*u)                 % M(4) se aplican.

% Ejercicio 17.

function [A] = f (a, b, x)         % Definición de f en un archivo m. Teclear el
  A = a* cos (3*x) + b* sin (2*x): % código a la izquierda en un archivo llamado f.m .
  end                              % A continuación, en sesión de MATLAB teclear:
fplot ('[f (1, 1,  x) f (3, 0, x) f (0, -3, x) f (4, -3, x) f (-3, 4, x) ]', [0 pi]) % Graficación.

% Ejercicio 19.

B1 = [0 0 0 1; 0 0 -1 1; 0 1 -1 0; -1 1 0 -1]
B2 = [0 0 1 2; 0 4 -1 0; 1 0 -1 0; 1 0 ]
rref (B1)     % Todas las matrices tienen 46 renglones pivote, así que los polinomios son linealmente
rref (B2)                % independientes, por lo que forman una base, ya que dim(P_3) = 4.
s% Ejercicios 20 a 23.

P1 = rref ([B1:B2])      % Reducir [B1:B2] y guardar las últimas
P = P1 (: , 5:8)         % 4 columnas para obtener P.
Q1 = rref ([B1:B2])      % Repetir con Q.
Q = Q1 (: , 5: 8)
P * Q                    % PQ=I y del mismo tamaño, así P^(-1)=Q.
p = [-1; 0; 5; 1]        % x^3+5x+1 en forma vectorial.
```

```
rref([B1, p])            % [p]_B1.
pb1=ans(:, 5)
rref([B2, p])            % [p]_B2.
pb2=ans(:, 5)            % P[p]_B2 produce [p]_B1, como se esperaba.
p * pb1
```

5

Transformaciones lineales

Carezco de un don especial. Sólo soy apasionadamente curioso.
Albert Einstein (1879-1955)

Introducción

Los vectores y las matrices se relacionan en forma íntima a través de la multiplicación matricial. Para una matriz fija A de $m \times n$, cualquier vector n $\mathbf{x}$ corresponde al vector m $A\mathbf{x}$. Esta correspondencia definida por el producto matricial $A\mathbf{x}$ es el principal ejemplo de una transformación lineal, cuya definición actual se debe a Peano. Las transformaciones lineales desempeñan un papel muy importante en matemáticas, física, ingeniería, procesamiento de imágenes, gráficas en computadora y muchas otras áreas de la ciencia y de la vida diaria. Ahora dirigiremos nuestra atención a un tipo contemporáneo de aplicaciones.

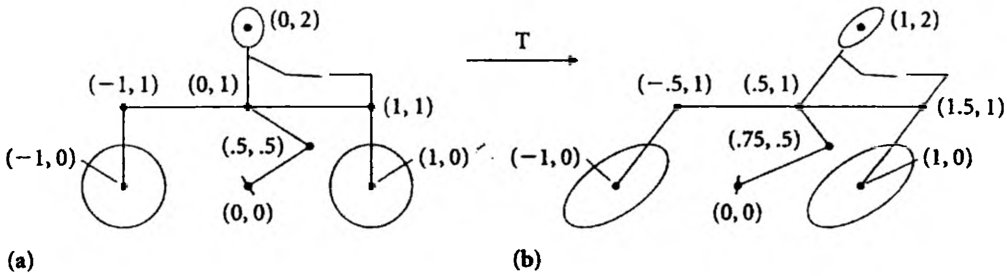


Figura 5.1 Transformación lineal de una imagen.

El dilema de un caricaturista

Un caricaturista moderno emplea computadoras y álgebra lineal para transformar las imágenes que dibuja. Supongamos que trata de dar la sensación de movimiento a la imagen de la figura 5.1(a), inclinándola y estirándola (horizontalmente) en forma gradual para llegar a la de la figura 5.1(b). Si el estiramiento gradual necesario, por ejemplo, a lo largo del eje x es 50%, ¿cómo puede modelarlo matemáticamente y hacer que la computadora trace la imagen inclinada? El método debería ser independiente de la imagen (cuadro) inicial para poder aplicarlo a otros cuadros. Como veremos en la sección 5.1, en la respuesta interviene una sencilla multiplicación de matriz por vector. De hecho, necesitamos multiplicar por la izquierda el vector coordenado de cualquier punto en el plano que deseemos transformar, por la matriz $\begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}$.

5.1 Transformaciones matriciales

Objetivos del estudiante para esta sección

1. Conocer qué es una transformación matricial y cómo determinar su rango.
2. Comprender geométricamente las transformaciones matriciales básicas de $\mathbf{R}^2$.

En esta sección presentaremos las transformaciones de matrices y estudiaremos algunas transformaciones matriciales del plano que desempeñan un papel importante en las gráficas en computadora. También describiremos un ejemplo de la física: las transformaciones de Galileo.

Transformaciones generales

Con frecuencia se necesita conocer cómo se relacionan los elementos de un conjunto con los de otro. A veces se usa una *regla* que describa esta relación. A continuación presentamos algunos ejemplos de esas reglas.

(R_1) Para cada vector-2 (x, y) se le asigna el vector 3 $(x \quad y, 0, y)$.

(R_2) Para cada vector-2 $\begin{bmatrix} x \\ y \end{bmatrix}$ se le asigna el vector 3 definido por $\begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

(R_3) Para cada $x > 0$ se le asigna la solución real para y de $y^2 - x = 0$.

(R_4) Para cada x real se le asigna la solución real de $x^2 + 1 - = 0$.

Hay varias diferencias agudas entre algunas de estas reglas. La regla (R_4) no tiene sentido, porque $x^2 + 1 = 0$ no tiene solución real. (R_3) es ambigua, porque $y^2 - x = 0$ implica a $\pm\sqrt{x}$. Así, a cada x le corresponden *dos* números, y no uno. Por otro lado, las reglas (R_1) y (R_2) no tienen estos problemas. A cada vector 2 se le asigna exactamente un vector 3 definido por la regla correspondiente. (R_1) y (R_2) están *bien definidas*, y constituyen ejemplos de *transformaciones*.

Una **transformación** T (**mapeo** o **función**) de un conjunto A a un conjunto B, representada por $T: A \to B$, es una regla que asocia a *cada* elemento de A un elemento b, *único*, de B, llamado **imagen** de a bajo T. Se escribe $T(a) = b$ y se dice que a se **mapea** a $T(a)$. A se llama **dominio** de T. B es **codominio** de T. El subconjunto de B formado por todas las imágenes de los elementos de A se llama **rango**[1] **o contradominio** de T y se representa por $R(T)$ o por $T(A)$.

---

[1] N. del T.: El uso de la palabra "rango" de una transformación es cada vez más general y se debe a la traducción indiscriminada de las palabras "range" y "rank"; pero para tratar de conservar la diferencia entre "rango" como dimensión del contradominio, y el contradominio mismo, usaremos aquí las palabras "rango" y "contradominio", respectivamente. También hay que recordar que no es lo mismo "range" (contradominio) que "rank" (rango) de una matriz. Además veremos más adelante que es indispensable usar las dos palabras, porque a veces aparecerán juntas en una sola frase, cosa que no sucede en otros libros de álgebra lineal.

Es posible que dos o más elementos de A tengan la misma imagen (figura 5.2). Dos transformaciones T_1, T_2: $A \to B$ son **iguales** (se escribe $T_1 = T_2$) si sus imágenes correspondientes son iguales, es decir, si

$$T_1(a) = T_2(a) \quad \text{para toda } a \text{ en } A$$

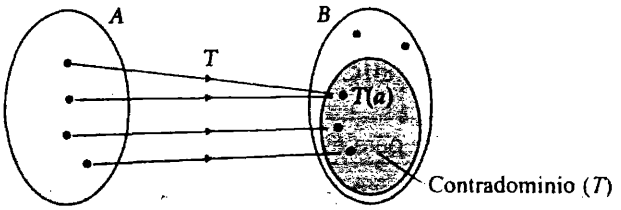


Figura 5.2 Dominio, codominio y contradominio.

Las reglas (R_1)y (R_2) definen transformaciones iguales, ya que

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ 0 \\ y \end{bmatrix}$$

■ EJEMPLO 1 Sea T: $\mathbf{R}^3 \to \mathbf{R}^2$ la transformación expresada por

$$T(x,y,z) = (x-y+z, x+y-z)$$

(a) ¿Por qué T es una transformación? ¿Cuál es su dominio? ¿Cuál es su codominio?
(b) ¿Cuál(es) de los vectores $(1,-2,3)$, $(1,2,-3)$ y $(1,0,5)$ tienen la misma imagen bajo T?
(c) Determine todos los vectores-3 que se aplican a $(0, 0)$.
(d) Describa el contradominio de T.

SOLUCIÓN

(a) T es una transformación porque *cada* vector 3, como, (x, y, z) corresponde *exactamente* a un vector 2, que es $(x-y+z, x+y-z)$. El dominio de T es $\mathbf{R}^3$. El codominio es $\mathbf{R}^2$.
(*b*) $T(1,-2,3) = (1-(-2)+3, 1+(-2)-3) = (6,-4)$. De igual manera, $T(1,2,-3) = (-4, 6)$ y $T(1,0,5) = (6,-4)$. Por consiguiente, $(1, -2, 3)$ y $(1, 0, 5)$ tienen la misma imagen.
(c) Es preciso conocer todos los vectores (x, y, z) tales que $T(x, y, z) = (x-y+z, x+y-z) = (0, 0)$. Entonces

$$x - y + z = 0$$
$$x + y - z = 0$$

cuya solución general es $(0, r, r)$, $r \in \mathbf{R}$. Todos ellos son los vectores 3 que aplican a $(0, 0)$.
(d) Para determinar el contradominio se necesitan todos los vectores 2, (a, b), para los cuales existen números x, y y z tales que $T(x, y, z) = (a, b)$. Entonces, se necesitan todos los (a, b) que hacen consistente al sistema

$$x - y + z = a$$
$$x + y - z = b$$

Puesto que la matriz de coeficientes tiene dos pivotes (compruébelo), el sistema es consistente para todas a, b. En consecuencia, el contradominio de T es $\mathbf{R}^2$. ☐

Transformaciones de matrices

Consideremos por ejemplo la matriz determinada $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$. Si tomamos los vectores 3 $\mathbf{x}$ y formamos los productos $A\mathbf{x}$, obtenemos vectores 2 únicos. Por ejemplo,

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \end{bmatrix}$$

y en general,

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ x + y - z \end{bmatrix}$$

Es claro que podemos definir una transformación $T: \mathbf{R}^3 \to \mathbf{R}^2$ por la regla $T(\mathbf{x}) = A\mathbf{x}$. De hecho, T es el mapeo del ejemplo 1. Éste es un ejemplo de *transformación matricial*, o de matrices. Y esta operación es la transformación más importante del álgebra lineal.

DEFINICIÓN

(Transformación matricial)

Una transformación matricial T se expresa mediante $T: \mathbf{R}^n \to \mathbf{R}^m$ y a ésta le corresponde una matriz A $m \times n$ tal que

$$T(\mathbf{x}) = A(\mathbf{x})$$

para todo $\mathbf{x} \in \mathbf{R}^n$. A se llama **matriz (estándar)** de T.

Por ejemplo, (R_1) y (R_2) definen la transformación matricial $T: \mathbf{R}^2 \to \mathbf{R}^3$, $T(\mathbf{x}) = A\mathbf{x}$, con $\begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$

■ EJEMPLO 2 Para la A y T anteriores, compare $R(T)$ y $\mathrm{Col}(A)$. Determine una descripción explícita de $R(T)$.

SOLUCIÓN Un vector 3 $\mathbf{w} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ está en $R(T)$ si y sólo si hay un vector 2 $\begin{bmatrix} x \\ y \end{bmatrix}$ tal que

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ 0 \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Esto equivale a decir que el sistema cuya matriz aumentada es $[A : \mathbf{w}]$ es consistente, o también que $\mathbf{w}$ está en $\mathrm{Col}(A)$. Por consiguiente, $R(T) = \mathrm{Col}(A)$.

Como $T(\mathbf{x}) = \mathbf{w}$ implica que $a = x - y$, $b = 0$ y $c = y$, se tiene que $x = a + c$, $b = 0$ y $y = c$. Así, a y c pueden ser escalares cualquiera y $b = 0$. En consecuencia, $R(T) = \left\{ \begin{bmatrix} a \\ 0 \\ c \end{bmatrix}, a, c \in \mathbf{R} \right\}$

TEOREMA 1

Si $T(\mathbf{x}) = A\mathbf{x}$ es cualquier transformación matricial, entonces

$$R(T) = \mathrm{Col}(A).$$

DEMOSTRACIÓN

$$\begin{aligned} \mathbf{w} \in R(T) &\Leftrightarrow T(\mathbf{x}) = \mathbf{w} \text{ para algún } \mathbf{x} \\ &\Leftrightarrow A\mathbf{x} = \mathbf{w} \text{ para algún } \mathbf{x} \\ &\Leftrightarrow [A : \mathbf{w}] \text{ es consistente} \\ &\Leftrightarrow \mathbf{w} \in \mathrm{Col}(A) \end{aligned}$$

El teorema siguiente describe las dos propiedades más importantes de una transformación matricial. Su demostración es consecuencia del teorema 17, sección 2.5.

TEOREMA 2

Cualquier transformación matricial $T: \mathbf{R}^n \to \mathbf{R}^m$, $T(\mathbf{x}) = A\mathbf{x}$ satisface lo siguiente:

1. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ para todos $\mathbf{x}$, $\mathbf{y}$ en $\mathbf{R}^n$.
2. $T(c\mathbf{x}) = cT(\mathbf{x})$ para todo $\mathbf{x}$ en $\mathbf{R}^n$ y todos los escalares c.

Algunas transformaciones matriciales del plano

Ahora estudiaremos algunas transformaciones matriciales geométricas del plano que son muy interesantes ($\mathbf{R}^2 \to \mathbf{R}^2$): reflexiones, compresiones-expansiones, cortes, rotaciones y proyecciones.

Reflexiones

Las reflexiones se definen respecto a cualquier recta en el plano. Nos interesan aquellas que están vinculadas con una recta que atraviesa el origen, en especial respecto a los ejes coordenados (R_x y R_y) y a la recta diagonal $y = x$ (R_d). Véase la figura 5.3. Esas reflexiones se definen con las fórmulas

$$R_y(x, y) = (-x, y) \qquad R_x(x, y) = (x, -y) \qquad R_d(x, y) = (y, x)$$

y todas éstas son transformaciones matriciales; sus matrices correspondientes son

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Por ejemplo, $R_y\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}$, y así sucesivamente.

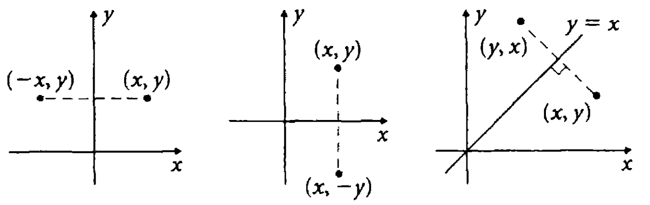


Figura 5.3 Reflexiones básicas.

Además existe la reflexión básica respecto al origen, cuya fórmula y matriz son

$$R_O(x, y) = (-x, -y) \qquad \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Esto también puede considerarse como rotación de 180° en torno al origen.

Compresiones-expansiones

Las compresiones y expansiones son escalamientos a lo largo de los ejes coordenados. Con más precisión: para $c > 0$, la transformación $C_x(x, y) = (cx, y)$ escala las coordenadas x en un factor de c, dejando inalteradas a las coordenadas y. Si $0 < c < 1$, se trata de una **compresión** en la dirección del eje x positivo. Si $c > 1$, se refiere a una **expansión** (figura 5.4). También se tienen compresiones y expansiones a lo largo del eje y, expresadas por $C_y(x, y) = (x, cy)$ para $c > 0$.

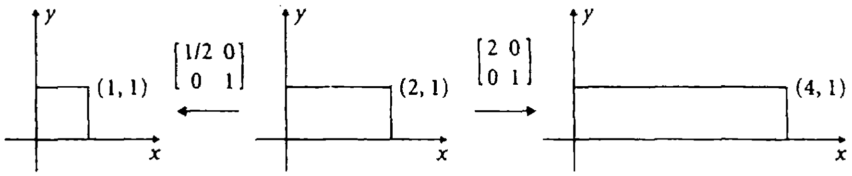


Figura 5.4 Compresión y estiramiento a lo largo del eje x.

Otro tipo son los escalamientos simultáneos a lo largo de los ejes x y y, como $C_{xy}(x, y) = (cx, dy)$ con factores de escala $c > 0$ y $d > 0$ a lo largo de las direcciones x y y (figura 5.5).

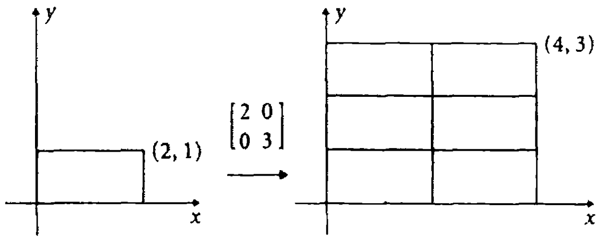


Figura 5.5 Escalamiento a lo largo de los ejes x y y.

Tanto C_x como C_y y C_{xy} son transformaciones matriciales, con sus respectivas matrices

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix} \qquad \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix}$$

Cortes

Un **corte** o **deslizamiento**[2] **a lo largo del eje** x es una transformación de la forma

$$S_x(x, y) = (x + cy, y)$$

En otras palabras, cada punto se mueve a lo largo de la dirección x una cantidad proporcional a la distancia al eje x (figura 5.6). También hay cortes a lo largo del eje y:

$$S_y(x, y) = (x, cx + y)$$

S_x y S_y son transformaciones matriciales cuyas matrices son

$$\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$$

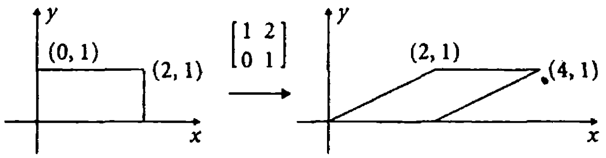


Figura 5.6 Deslizamiento a lo largo del eje x.

Observe que la constante c en la fórmula para un corte puede ser negativa. La figura 5.7 ilustra este caso para $S_y(x, y) = (x, -2x + y)$.

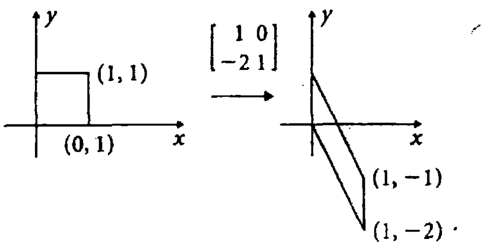


Figura 5.7 Deslizamiento a lo largo de la dirección negativa de y.

Quienes trabajan con gráficas en computadora aplican las transformaciones matriciales de corte y de otros tipos para modificar las imágenes. Estas operaciones se adaptan bien a los cálculos en computadora, porque se implantan con facilidad los productos $A\mathbf{x}$.

[2] N. del T.: Este nombre de transformación da idea de las *deformaciones por esfuerzo cortante* y de los *deslizamientos* de capas cristalinas en los materiales sometidos a esfuerzo cortante.

■ EJEMPLO 3 (Aplicación a gráficas en computadora) Determine la transformación S_x que se desliza en dirección positiva de x en un factor de 0.5. Para la figura 5.8, obtenga las imágenes de los puntos (0, 2), (0, 1), (0.5, 0.5), (0, 0), (1, 0), (1, 1), (–1, 1), (–1, 0).[3]

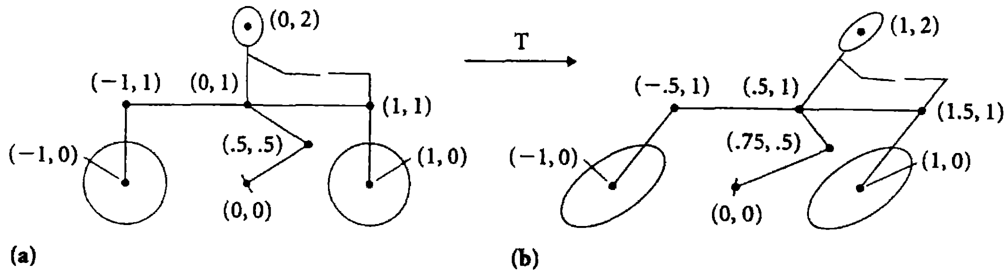


Figura 5.8 Deslizamiento en transformaciones de imagen.

SOLUCIÓN S_x se expresa como $S_x(x, y) = (x + 0.5y, y)$. Por consiguiente,

$$S_x\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \tag{5.1}$$

Las imágenes de los puntos identificados pueden calcularse sustituyendo sus coordenadas en la ecuación (5.1). Sin embargo, se ahorra espacio escribiendo los productos matriz por vector en forma de un producto de matrices:

$$\begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0.5 & 0 & 1 & 1 & -1 & -1 \\ 2 & 1 & 0.5 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0.5 & 0.75 & 0 & 1 & 1.5 & -0.5 & -1 \\ 2 & 1 & 0.5 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Las coordenadas de las imágenes son (1, 2), (0.5, 1), (0.75. 0.5), (0, 0), (1, 0), (1.5, 1), (–0.5, 1), (–1, 0). □

Rotaciones

Otro tipo común de transformación en el plano es la rotación o giro en torno a cualquier punto en el plano. Nos interesan principalmente las rotaciones en torno al origen.

■ EJEMPLO 4 (Rotación en el plano) La transformación $R_\theta : \mathbf{R}^2 \to \mathbf{R}^2$ se define por

$$R_\theta\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

y hace girar cada vector θ rad en sentido contrario al de las manecillas del reloj en torno al origen.

[3] Para generar la parte derecha de la figura 5.8, habría que calcular las imágenes de varios puntos de la parte izquierda.

SOLUCIÓN De acuerdo con la figura 5.9(a), $\overrightarrow{OB}$ es la rotación de $\overrightarrow{OA}$ un ángulo θ. Entonces

$$x = r\cos\phi \qquad y = r\,\mathrm{sen}\,\phi$$
$$x' = r\cos(\phi+\theta) \qquad y' = r\,\mathrm{sen}(\phi+\theta)$$

Mediante identidades trigonométricas se llega a

$$x' = r\cos\phi\cos\theta - r\,\mathrm{sen}\,\phi\,\mathrm{sen}\,\theta \qquad y' = r\,\mathrm{sen}\,\phi\cos\theta + r\cos\phi\,\mathrm{sen}\,\theta$$
$$x' = x\cos\theta - y\,\mathrm{sen}\,\theta \qquad y' = y\cos\theta + x\,\mathrm{sen}\,\theta$$

Por consiguiente

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\,\mathrm{sen}\,\theta \\ x\,\mathrm{sen}\,\theta + y\cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = R_\theta\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$$

que demuestran que R_θ es una rotación de θ rad en torno al origen.

Por ejemplo, calcularemos la imagen de (1, 1) para $\theta = \frac{\pi}{2}$ (figura 5.9(b)).

$$R_{\pi/2}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \cos\frac{\pi}{2} & -\mathrm{sen}\frac{\pi}{2} \\ \mathrm{sen}\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \square$$

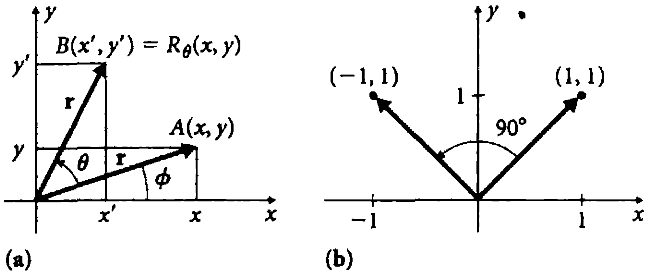


Figura 5.9 Rotación en torno al origen.

NOTA Al justificar la afirmación del ejemplo 4, en realidad hemos demostrado que *toda* rotación en el plano en torno al origen se expresa con la transformación matricial anterior para alguna θ.

Proyecciones

Las proyecciones del plano sobre una recta también son transformaciones del plano. Nos interesan las proyecciones ortogonales sobre las rectas que pasan por el origen, en especial sobre los ejes coordenados.

Las **proyecciones** P_x sobre el eje x y P_y sobre el eje y (figura 5.10) se expresan por

$$P_x(x, y) = (x, 0) \qquad P_y(x, y) = (0, y)$$

las cuales son transformaciones matriciales con matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

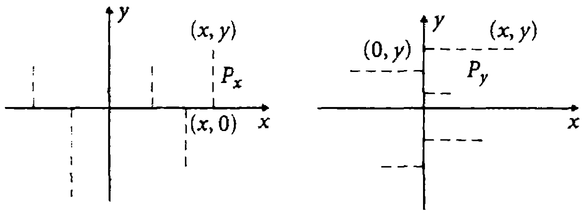


Figura 5.10 Proyecciones ortogonales sobre los ejes.

Un ejemplo de la física (transformaciones de Galileo)

Sean F y F' dos marcos de referencia con ejes coordenados paralelos, x, y, z y x', y', z'. Supongamos que el marco F' se aleja del marco F a una velocidad relativa constante v en una dirección a lo largo de los ejes x, x' (figura 5.11). Además, supongamos que hay dos observadores en F y en F', por ejemplo, un matemático y un físico, respectivamente, y que ambos se comunican sus mediciones (de distancias, velocidades, etcétera).[4]

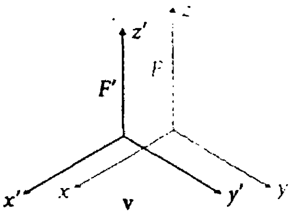


Figura 5.11 Marcos de referencia moviéndose con velocidad relativa constante.

Si (x, y, z, t) son las coordenadas del espacio tiempo que usa el matemático para sus mediciones, y (x', y', z', t') son las que emplea el físico, ¿cuál es la relación entre los dos marcos de coordenadas?

Como no hay movimiento a lo largo de los ejes y, y' y z, z', se debe cumplir que $y = y'$ y $z = z'$. También, en la mecánica newtoniana, $t = t'$ para ambos marcos. Para mediciones a lo largo de los ejes x, x' sucede lo siguiente: una medición x hecha en el momento t en el marco F por el matemático, es mayor que la medición correspondiente x' hecha por el físico al mismo tiempo, y la diferencia es vt. Por consiguiente, $x' = x - vt$. Así,

$$
\begin{aligned}
x' &= x - vt \\
y' &= y \\
z' &= z \\
t' &= t
\end{aligned}
$$

[4] Un ejemplo de este caso sería cuando el matemático y el físico van en dos trenes distintos, de Nueva York a Baltimore, cuando ambos trenes se mueven a velocidades constantes por un tramo recto, pero el físico va en el *MetroLiner*, que es más rápido.

que en notación matricial es

$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -v \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}$$

Puede observarse que el sistema $F'(x', y', z', t')$ se expresa con una *transformación matricial* del sistema $F(x, y, z, t)$. A esto se le llama **transformación de Galileo** en la mecánica newtoniana. Su nombre es en honor de Galileo,[5] quien fue el primero en expresar el principio de que los sistemas en movimiento uniforme entre sí son equivalentes, cuando describió las leyes de la cinemática.

Simplificación de la notación

Al estudiar las transformaciones $\mathbf{R}^n \to \mathbf{R}^m$ a menudo se omiten los paréntesis en $T(\mathbf{u})$ cuando el vector $\mathbf{u}$ se expresa en forma de componentes. Por ejemplo, se escribe $T\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $T(1, 2)$ en lugar de lo más correcto, que es $T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$ y $T((1, 2))$.

Definición de transformaciones con sistemas algebraicos computacionales

Defina de $T(x, y, z) = (x - 2y, y - 3z, z - 4x)$ y calcule $T(1, 2, 3)$:

Maple

```
> T := (x,y,z,) -> [x-2*y,y-3*z,z-4*x];
          T:=(x,y,z)->[x-2y,y-3z,z-4x]
> T(1,2,3);
                        [-3,-7,-1]
```

Mathematica

```
In[1]:=
     T[x_,y_,z_]:={x-2y,y-3z,z-4x}
In[2]:=
     T[1,2,3]
Out[2]=
     {-3, -7, -1}
```

[5] **Galileo Galilei** (1564–1642) nació en Pisa, Italia, y se le considera uno de los físicos más grandes de todos los tiempos. Estudió medicina y matemáticas en la Universidad de Pisa. Enseñó en Pisa y en Padua, y su fama se debe a sus logros fundamentales en física y astronomía. Algo de su trabajo ofendió a la Iglesia, siendo perseguido y encarcelado.

MATLAB

```
% Function file:
function [A] = T(x,y,z)
          A=[x-2*y,y-3*z,z-4*x];
          end
% In session:
T(1,2,3)
ans =
    -3    -7    -1
```

Ejercicios 5.1

Transformaciones generales

En los ejercicios 1 a 5, determine cuáles de las reglas son transformaciones. Si alguna lo es, identifique el dominio y el codominio.

A cada vector 2 (x, y), asigne

1. El vector 2 $(x - y, y)$.
2. El vector 2 $(x - y, y^2)$.
3. El vector 2 $(x - y, \sqrt{y})$.
4. $\begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$
5. $A\begin{bmatrix} x \\ y \end{bmatrix}$, si A es una matriz 3×3.

Entonces

$$T_1: \mathbf{R} \to \mathbf{R}, \quad T_1(x) = x^2$$

En los ejercicios 6 a 8 determine si las transformaciones T_1 y T_2 son iguales.

6. T_2: $\mathbf{R} \to \mathbf{R}$, $T_2(x)$ es la solución de $\sqrt{y} - x = 0$ para y.
7. $T_2 : \mathbf{R} \to \mathbf{R}^+$, $T_2(x) = x^2$ ($\mathbf{R}^+$ es el conjunto de los números reales positivos).
8. $T_2 : \mathbf{R}^+ \to \mathbf{R}$, $T_2(x) = x^2$.

En los ejercicios 9 a 15, para las transformaciones dadas $T : \mathbf{R}^n \to \mathbf{R}^m$, determine lo siguiente:

(a) n y m.

(b) El dominio y el codominio de T.

(c) Todos los vectores del dominio cuya imagen es el vector-m cero.

(d) El contradominio de T.

9. $T(x, y) = (x + y, x - y)$
10. $T(x, y) = (x - y, 0)$
11. $T(x, y, z) = (x + y, x - z)$
12. $T(x, y) = (x - y, x - y, 0)$
13. $T(x, y, z) = (x - z, -x + z, x - z)$
14. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$
15. $T(\mathbf{x}) = \begin{bmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 1 & 0 \end{bmatrix} \mathbf{x}$

Transformaciones matriciales

16. Demuestre el teorema 2.

En los ejercicios 17 a 22 encuentre una base para $R(T) = \mathrm{Col}(A)$.

17. $T : \mathbf{R}^2 \to \mathbf{R}^2, \quad T(\mathbf{x}) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \mathbf{x}$
18. $T : \mathbf{R}^2 \to \mathbf{R}^2, \quad T(\mathbf{x}) = \begin{bmatrix} 1 & 2 \\ -3 & -6 \end{bmatrix} \mathbf{x}$
19. $T : \mathbf{R}^4 \to \mathbf{R}^2, \quad T(\mathbf{x}) = \begin{bmatrix} 1 & 2 & 7 & -5 \\ 0 & 1 & 6 & 6 \end{bmatrix} \mathbf{x}$
20. $T : \mathbf{R}^3 \to \mathbf{R}^2, \quad T(\mathbf{x}) = \begin{bmatrix} 1 & 0 & -8 \\ -2 & 0 & 16 \end{bmatrix} \mathbf{x}$
21. $T : \mathbf{R}^3 \to \mathbf{R}^2, \quad T(\mathbf{x}) = \begin{bmatrix} 1 & 0 & -2 \\ -2 & 0 & 1 \\ 0 & -1 & 4 \end{bmatrix} \mathbf{x}$

22. $T:\mathbf{R}^2 \to \mathbf{R}^4, \quad T(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x}$

Transformaciones matriciales en el plano

En los ejercicios 23 a 30 considere la transformación matricial en el plano $T: \mathbf{R}^2 \to \mathbf{R}^2$, $T(x) = Ax$, con la matriz A respectiva.

(i) Calcule y trace las imágenes de

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

(ii) Identifique a T como una de las transformaciones siguientes:

1. Reflexión
2. Compresión-expansión
3. Deslizamiento
4. Rotación
5. Proyección
6. Ninguna de las anteriores

23. **a.** $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ **b.** $\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$

24. **a.** $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$

25. **a.** $\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}$

26. **a.** $\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$

27. **a.** $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ **b.** $\begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}$

28. **a.** $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ **b.** $\begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$

29. **a.** $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ **b.** $\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$

30. **a.** $\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$ **b.** $\begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix}$

31. Deduzca una fórmula y la matriz para la reflexión respecto a la recta $y = -x$.

32. Obtenga una fórmula y la matriz para la reflexión respecto a la recta $y = 2x$.

33. Deduzca una fórmula y la matriz para la rotación de θ radianes en el sentido de las manecillas del reloj en torno al origen.

34. Trace la imagen del cuadrado con vértices en (0, 0), (1, 0), (1, 1) y (0, 1) bajo el corte o deslizamiento.

$$S_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

35. Demuestre que cualquier transformación matricial $T: \mathbf{R}^2 \to \mathbf{R}^2$ transforma rectas en rectas o en puntos.

Transformaciones matriciales en el espacio

Reflexiones: En el espacio, las principales reflexiones son respecto *al origen, a los planos coordenados, a los ejes coordenados, a las bisectrices de los ejes coordenados* y *a los planos bisectores*.

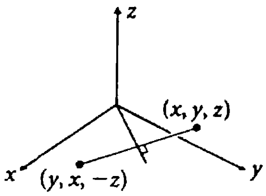


Figura 5.12. Reflexión con respecto a la bisectriz xy.

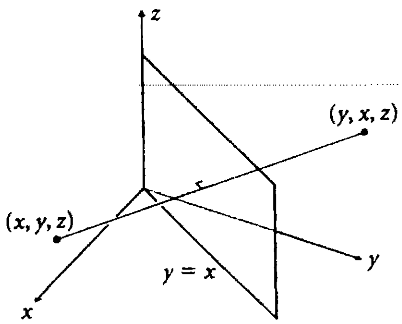


Figura 5.13 Reflexión con respecto al plano $y = x$.

36. En la tabla siguiente anote la respuesta correcta donde hay signo de interrogación.

Reflexiones respecto a:	$T(\mathbf{x})$	Matriz A
?	?	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Plano ?	?	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Eje ?	?	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Recta en primer cuadrante que biseca el plano ? (figura 5.12)	?	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

Reflexiones con respecto al:	$T(\mathbf{x})$	Matriz A
Plano y = ? (figura 5.13)	?	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Rotaciones: La transformación de rotación R_θ^z que hace girar θ radianes a cualquier vector-3 en torno al eje z en la dirección positiva[6] (figura 5.14) se expresa como sigue:

$$R_\theta^z \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

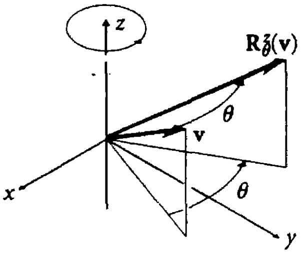


Figura 5.14 Rotación de θ rad con respecto al eje z.

37. Deduzca las fórmulas para las rotaciones correspondientes en torno a los ejes x y y.

Proyecciones: Existen varias proyecciones ortogonales, principalmente en los planos y ejes coordenados.

38. En la tabla siguiente anote la respuesta correcta en los lugares donde aparecen signos de interrogación.

Proyección sobre:	$T(\mathbf{x})$	Matriz A
Plano ? (Figura 5.15)	?	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Eje ? (Figura 5.16)	?	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

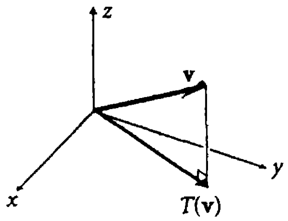


Figura 5.15 Proyección sobre el plano xy.

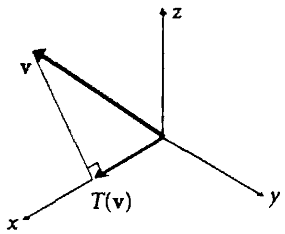


Figura 5.16 Proyección sobre el eje x.

5.2 Transformaciones lineales

Objetivo del estudiante para esta sección
Comprender la definición de una transformación lineal y cómo investigarla.

Las transformaciones lineales son *mapeos* de importancia fundamental en el álgebra lineal y en sus aplicaciones. Son transformaciones entre espacios vectoriales que conservan la suma vectorial y la multiplicación por escalar. En esta sección definiremos una transformación lineal y mostraremos varios ejemplos que el lector debe estudiar con mucho cuidado.

[6] La dirección positiva se determina por la **regla de la mano derecha**. Cuando los dedos de la mano derecha se curvan desde el eje x positivo hacia el eje y positivo, el pulgar apunta hacia el eje z positivo.

Definición y ejemplos

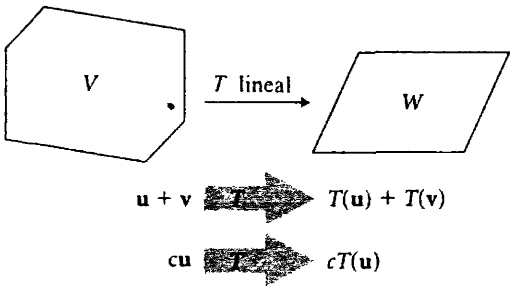
DEFINICIÓN

(Transformaciones lineales)

Sean V, W dos espacios vectoriales. Una **transformación lineal** (o **mapeo lineal**) de V a W es una transformación $T : V \to W$ tal que para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ de V y cualquier escalar c,

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$;
2. $T(c\mathbf{u}) = cT(\mathbf{u})$.

El signo + en $\mathbf{u} + \mathbf{v}$ es una suma en V, mientras que el signo + en $T(\mathbf{u}) + T(\mathbf{v})$ es una suma en W. Del mismo modo, las multiplicaciones escalares $c\mathbf{u}$ y $cT(\mathbf{u})$ se efectúan en V y W, respectivamente. En el caso especial cuando $V = W$, la transformación lineal $T : V \to V$ se llama **operador lineal** de V.



■ EJEMPLO 5 Demuestre que la transformación $T : \mathbf{R}^2 \to \mathbf{R}^2$ definida por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - 3y \\ x + 4y \end{bmatrix}$$

es lineal

SOLUCIÓN Sean $\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ Entonces

$$\begin{aligned}
T(\mathbf{u} + \mathbf{v}) &= T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \\
&= T\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} \\
&= \begin{bmatrix} 2(x_1 + x_2) - 3(y_1 + y_2) \\ (x_1 + x_2) + 4(y_1 + y_2) \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 - 3y_1 \\ x_1 + 4y_1 \end{bmatrix} + \begin{bmatrix} 2x_2 - 3y_2 \\ x_2 + 4y_2 \end{bmatrix} \\
&= T\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + T\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \\
&= T(\mathbf{u}) + T(\mathbf{v})
\end{aligned}$$

Para todo escalar c,

$$T(c\mathbf{u}) = T\begin{bmatrix} cx_1 \\ cy_1 \end{bmatrix}$$

$$= \begin{bmatrix} 2cx_1 - 3cy_1 \\ cx_1 + 4cy_1 \end{bmatrix}$$

$$= c\begin{bmatrix} 2x_1 - 3y_1 \\ x_1 + 4y_1 \end{bmatrix}$$

$$= cT\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$= cT(\mathbf{u})$$

Se satisfacen ambas partes de la definición, de modo que T es lineal.

■ EJEMPLO 6 Demuestre que la transformación $T : \mathbf{R}^3 \to \mathbf{R}^2$ es lineal:

$$T(x, y, z) = (x - z, y + z)$$

SOLUCIÓN Sean $\mathbf{u} = (x_1, y_1, z_1)$ y $\mathbf{v} = (x_2, y_2, z_2)$. Entonces

$$T(\mathbf{u} + \mathbf{v}) = T(x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

$$= ((x_1 + x_2) - (z_1 + z_2), (y_1 + y_2) + (z_1 + z_2))$$

$$= (x_1 - z_1, y_1 + z_1) + (x_2 - z_2, y_2 + z_2)$$

$$= T(\mathbf{u}) + T(\mathbf{v})$$

demuestra la parte 1 de la definición. La parte 2 se deja como práctica.

Los ejemplos más importantes de transformaciones lineales son las transformaciones de matrices. De hecho, como veremos, estos últimos son los *únicos* mapeos lineales de $\mathbf{R}^n$ a $\mathbf{R}^m$.

■ EJEMPLO 7 (Transformaciones matriciales) Demuestre que toda transformación matricial es lineal.

SOLUCIÓN De acuerdo con el teorema 2 de la sección 5.1, ambas propiedades de la definición se cumplen. Practicaremos volviendo a desarrollar la demostración. Si $T : \mathbf{R}^n \to \mathbf{R}^m$, entonces $T(\mathbf{x}) = A\mathbf{x}$ es una transformación matricial, y

$$T(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = T(\mathbf{x}) + T(\mathbf{y})$$

$$T(c\mathbf{x}) = A(c\mathbf{x}) = c(A\mathbf{x}) = cT(\mathbf{x})$$

En consecuencia, T es lineal.

■ EJEMPLO 8 (Transformación lineal geométrica) Demuestre que las reflexiones, cortes o deslizamientos, compresiones-estiramientos, todos con respecto a los ejes coordenados, las rotaciones con respecto al origen y las proyecciones sobre los ejes de coordenadas son transformaciones lineales.

SOLUCIÓN Todos estos mapeos[7] son transformaciones matriciales, basándose en la sección 5.1. En consecuencia, según el ejemplo 7, son lineales. ☐

Propiedades de las transformaciones lineales

La definición de una transformación lineal también puede expresarse como sigue:

TEOREMA 3

> $T : V \to W$ es una transformación lineal si y sólo si para todos los vectores $\mathbf{v}_1$ y $\mathbf{v}_2 \in V$, y todos los escalares c_1 y c_2, se cumple
>
> $$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

DEMOSTRACIÓN Si T es lineal, entonces

$$\begin{aligned} T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) &= T(c_1\mathbf{v}_1) + T(c_2\mathbf{v}_2) \\ &= c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) \end{aligned}$$

de acuerdo con las partes 1 y 2 de la definición.

A la inversa, si T es una transformación tal que $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$ para todos $\mathbf{v}_1, \mathbf{v}_2 \in \mathbf{R}^n$ y todos $c_1, c_2 \in \mathbf{R}$, entonces si igualamos $c_1 = c_2 = 1$ obtenemos la parte 1 de la definición y haciendo $c_2 = 0$ se obtiene la parte 2. ☐

En términos generales, si $\mathbf{v}_1, \ldots, \mathbf{v}_n$ son vectores cualesquiera en V, y $c_1, \ldots, c_n$ son escalares cualesquiera, entonces

$$T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n)$$

Así, las transformaciones lineales mapean una combinación lineal de vectores en la misma combinación lineal de las imágenes de esos vectores.

TEOREMA 4

> $T : V \to W$ es una transformación lineal. Entonces
>
> 1. $T(\mathbf{0}) = \mathbf{0}$;
> 2. $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$.

DEMOSTRACIÓN

1. De acuerdo con la parte 2 de la definición,

$$T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}$$

[7] Las reflexiones con respecto a cualquier recta que pasa por el origen y las proyecciones sobre cualquier recta que pasa por el origen son también transformaciones matriciales (es decir, lineales).

2. Según el teorema 3, haciendo que $c_1 = 1$ y $c_2 = -1$,

$$T(\mathbf{u} - \mathbf{v}) = T(1\mathbf{u} + (-1)\mathbf{v}) = 1T(\mathbf{u}) + (-1)T(\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) \qquad \square$$

■ EJEMPLO 9 ¿Es $f: \mathbf{R}^2 \to \mathbf{R}^2$ una transformación lineal, definida por $f(x, y) = (x, 1)$?

RESPUESTA Si f fuera lineal, entonces $f(0, 0)$ sería $(0, 0)$, con base en la parte 1 del teorema 4. Sin embargo, $f(0, 0) = (0, 1)$, de modo que f es no lineal. $\square$

Más ejemplos

La transformación $\mathbf{0}: V \to W$ que mapea a todos los vectores de V en $\mathbf{0}$ en W se le llama **transformación cero**

$$\mathbf{0}(\mathbf{v}) = \mathbf{0} \qquad \text{para todo } \mathbf{v} \in V$$

■ EJEMPLO 10 Compruebe que la transformación $\mathbf{0}: V \to W$ es lineal.

SOLUCIÓN Si $\mathbf{u}$ y $\mathbf{v}$ son vectores de V y c es un escalar, entonces

$$\mathbf{0}(\mathbf{v} + \mathbf{u}) = \mathbf{0} = \mathbf{0} + \mathbf{0} = \mathbf{0}(\mathbf{v}) + \mathbf{0}(\mathbf{u})$$

y

$$\mathbf{0}(c\mathbf{v}) = \mathbf{0} = c\mathbf{0} = c\mathbf{0}(\mathbf{v}) \qquad \square$$

La transformación $I: V \to V$ que convierte cada vector de V en sí mismo se llama **transformación identidad de** V.

$$I(\mathbf{v}) = \mathbf{v} \qquad \text{para todo } \mathbf{v} \in V$$

■ EJEMPLO 11 Demuestre que la transformación identidad $I: V \to V$ es lineal.

SOLUCIÓN Se deja como ejercicio. $\square$

■ EJEMPLO 12 (Homotecia) Para un escalar fijo c, compruebe que $T: V \to V$ es lineal.

$$T(\mathbf{v}) = c\mathbf{v}$$

SOLUCIÓN Sean $\mathbf{u}, \mathbf{w} \in V$ y $r \in \mathbf{R}$. T es lineal, porque

$$T(\mathbf{u} + \mathbf{w}) = c(\mathbf{u} + \mathbf{w}) = c\mathbf{u} + c\mathbf{w} = T(\mathbf{u}) + T(\mathbf{w})$$

$$T(r\mathbf{u}) = c(r\mathbf{u}) = r(c\mathbf{u}) = rT(\mathbf{u}) \qquad \square$$

A la transformación definida en el ejemplo 12 se le llama **homotecia**. Si $c > 1$, la homotecia es una **dilatación,** y su efecto sobre $\mathbf{v}$ es estirarlo en un factor de c. Si $0 < c < 1$, la homotecia es una **contracción**, y su efecto sobre $\mathbf{v}$ es encogerla en un factor de c (figura 5.17). Si $c < 0$, esta transformación invierte la dirección de $\mathbf{v}$.

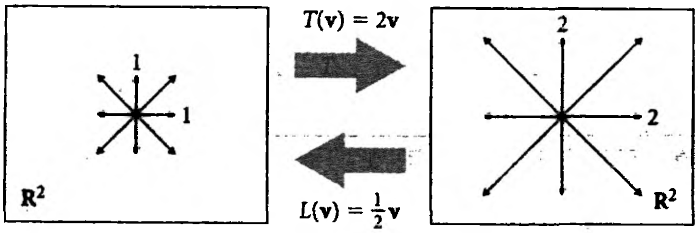


Figura 5.17 Dilatación y contracción por un factor de 2.

■ EJEMPLO 13 (Multiplicación por una matriz fija) Sea A una matriz fija $m \times n$. Demuestre que la transformación $T: M_{nk} \to M_{mk}$ es lineal.

$$T(B) = AB$$

SOLUCIÓN

$$T(B + C) = A(B + C) = AB + AC = T(B) + T(C)$$
$$T(cB) = A(cB) = c(AB) = cT(B)$$

■ EJEMPLO 14 Compruebe que la transformación $T: P_2 \to P_1$ es lineal.

$$T(a + bx + cx^2) = b + 2cx$$

SOLUCIÓN Sean $p_1 = a_1 + b_1x + c_1x^2$ y $p_2 = a_2 + b_2x + c_2x^2$. Entonces

$$\begin{aligned} T(p_1 + p_2) &= T\left((a_1 + a_2) + (b_1 + b_2)x + (c_1 + c_2)x^2\right) \\ &= (b_1 + b_2) + 2(c_1 + c_2)x \\ &= (b_1 + 2c_1x) + (b_2 + 2c_2x) \\ &= T(p_1) + T(p_2) \end{aligned}$$

La verificación de la parte 2 de la definición se deja como ejercicio.

■ EJEMPLO 15 $T: M_{22} \to \mathbf{R}^2$ es lineal. (¿Por qué?)

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} b - a \\ c + d \end{bmatrix}$$

■ EJEMPLO 16 (Producto punto por un vector fijo) Sea $\mathbf{u}$ un vector fijo en $\mathbf{R}^n$. Demuestre que la transformación lineal $T: \mathbf{R}^n \to \mathbf{R}$ es lineal.

$$T(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$$

SOLUCIÓN Sean $\mathbf{v}_1, \mathbf{v}_2 \in \mathbf{R}^n$ y $c_1, c_2 \in \mathbf{R}$. Entonces

$$\begin{aligned} T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) &= \mathbf{u}\cdot(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = \mathbf{u}\cdot(c_1\mathbf{v}_1) + \mathbf{u}\cdot(c_2\mathbf{v}_2) \\ &= c_1(\mathbf{u}\cdot\mathbf{v}_1) + c_2(\mathbf{u}\cdot\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) \end{aligned}$$

Por consiguiente, T es lineal según el teorema 3. ☐

■ EJEMPLO 17 (Producto cruz por vector fijo) Compruebe que la transformación $T: \mathbf{R}^3 \to \mathbf{R}^3$ es lineal.

$$T(\mathbf{v}) = \mathbf{u} \times \mathbf{v}$$

SOLUCIÓN La verificación es semejante a la del ejemplo 16. ☐

■ EJEMPLO 18 (Proyección a la recta que pasa por el origen) Si $\mathbf{u}$ es un vector fijo distinto de cero en $\mathbf{R}^3$, la transformación $T: \mathbf{R}^3 \to \mathbf{R}^3$ definida con la proyección ortogonal de cada $\mathbf{v} \in \mathbf{R}^3$ sobre $\mathbf{u}$ (véase la sección 2.2) es lineal (figura 5.18).

$$T(\mathbf{v}) = \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$$

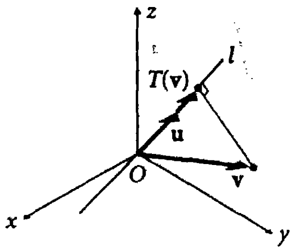


Figura 5.18 Proyección ortogonal sobre la recta l.

SOLUCIÓN Ya que

$$\begin{aligned} T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) &= \frac{(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{c_1\mathbf{v}_1\cdot\mathbf{u} + c_2\mathbf{v}_2\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} \\ &= c_1\frac{\mathbf{v}_1\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} + c_2\frac{\mathbf{v}_2\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) \end{aligned}$$

lo hemos demostrado, según el teorema 3. ☐

■ EJEMPLO 19 (Requiere cálculo) Sea V el espacio vectorial de todas las funciones diferenciables de valor real definidas en $\mathbf{R}$. Compruebe que la transformación $T: V \to V$ definida diferenciando cada $f \in V$

$$T(f) = f'$$

es lineal.

SOLUCIÓN Si $f, g \in V$ y $c \in \mathbf{R}$, de acuerdo con las propiedades básicas de las derivadas se obtiene

$$T(f+g) = (f+g)' = f' + g' = T(f) + T(g)$$
$$T(cf) = (cf)' = cf' = cT(f)$$

De modo que T es lineal. ☐

■ EJEMPLO 20 (Requiere cálculo) Si $C[0, 1]$ es el vector espacial que contiene todos los valores reales continuos de las funciones diferenciables, las cuales están definidas en el intervalo $[0, 1]$; demuestre que la transformación $T: C[0, 1] \to C[0, 1]$, definida por la integración (Riemann).

$$T(f) = \int_0^1 f(x)\,dx$$

es lineal.

SOLUCIÓN Si $f, g \in V$ y $c \in \mathbf{R}$, entonces según las propiedades básicas de la integral se obtiene

$$T(f+g) = \int_0^1 (f(x) + g(x))\,dx = \int_0^1 f(x)\,dx + \int_0^1 g(x)\,dx = T(f) + T(g)$$
$$T(cf) = \int_0^1 cf(x)\,dx = c\int_0^1 f(x)\,dx = cT(f)$$

Por tanto, T es lineal. ☐

■ EJEMPLO 21 ¿Es lineal la transformación $f: \mathbf{R} \to \mathbf{R}$, definida por $f(x) = x^2$?

RESPUESTA No, la parte 1 de la definición no lo es, porque

$$f(x+y) = (x+y)^2 = x^2 + 2xy + y^2$$

y

$$f(x) + f(y) = x^2 + y^2$$

no son iguales si $xy \neq 0$. (Tampoco la parte 2 es lineal. Compruebe por qué.) ☐

■ EJEMPLO 22 Demuestre que no es lineal la transformación $f: \mathbf{R}^2 \to \mathbf{R}$, definida por $f(x, y) = xy$.

SOLUCIÓN Se deja como ejercicio. ☐

Determinación de un mapeo lineal a partir de sus valores en una base

Una de las propiedades más importantes de la transformación lineal es que puede determinarse *únicamente* cuando sus valores están dados en una base. Explicaremos lo anterior con un ejemplo.

■ EJEMPLO 23 Sea $T : \mathbf{R}^2 \to \mathbf{R}^3$ una transformación lineal tal que

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}, \qquad T\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -8 \\ -6 \\ 5 \end{bmatrix}$$

Calcule $T\begin{bmatrix} -9 \\ 6 \end{bmatrix}$ y $T\begin{bmatrix} x \\ y \end{bmatrix}$

SOLUCIÓN Como

$$\begin{bmatrix} -9 \\ 6 \end{bmatrix} = -4\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 5\begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

si agregamos T en ambos lados de la ecuación, obtenemos

$$\begin{aligned} T\begin{bmatrix} -9 \\ 6 \end{bmatrix} &= T\left(-4\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 5\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) = -4T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 5T\begin{bmatrix} -1 \\ 2 \end{bmatrix} \\ &= -4\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} + 5\begin{bmatrix} -8 \\ -6 \\ 5 \end{bmatrix} = \begin{bmatrix} -36 \\ -42 \\ 21 \end{bmatrix} \end{aligned}$$

La segunda igualdad permanece porque T es lineal.

Es facil observar que

$$\begin{bmatrix} x \\ y \end{bmatrix} = \left(\frac{2}{3}x + \frac{1}{3}y\right)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \left(-\frac{1}{3}x + \frac{1}{3}y\right)\begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

Por consiguiente,

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \left(\frac{2}{3}x + \frac{1}{3}y\right)\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} + \left(-\frac{1}{3}x + \frac{1}{3}y\right)\begin{bmatrix} -8 \\ -6 \\ 5 \end{bmatrix} = \begin{bmatrix} 2x - 3y \\ 4x - y \\ -x + 2y \end{bmatrix}$$ $\square$

Notación de la operación matricial

En el ejemplo 23 efectuamos las operaciones matriciales

$$\begin{bmatrix} 1 & -1 & : & x \\ 1 & 2 & : & y \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & : & \frac{2}{3}x + \frac{1}{3}y \\ 0 & 1 & : & \frac{1}{3}y - \frac{1}{3}x \end{bmatrix}$$

y

$$\begin{bmatrix} -1 & -8 \\ 3 & -6 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} \frac{2}{3}x + \frac{1}{3}y \\ \frac{1}{3}y - \frac{1}{3}x \end{bmatrix} = \begin{bmatrix} 2x - 3y \\ 4x - y \\ -x + 2y \end{bmatrix}$$

NOTA Hemos explicado cómo calcular los valores de una transformación lineal dados sus valores sobre una base. Si T no se da sobre toda una base no podemos determinar la imagen de *cada* elemento. Refiriéndonos al ejemplo 23, si sólo se dieran las imágenes de $(1, 1)$ y $(-1, -1)$ no podríamos determinar $T(-9, 6)$ porque $(-9, 6)$ no está en $\text{Gen}\{(-1, 1), (1, -1)\} = \text{Gen}\{(1, 1)\}$.

■ EJEMPLO 24 Sea $T: P_1 \to P_1$ una transformación lineal tal que

$$T(-1 + x) = -7 + 2x, \qquad T(1 + x) = 4 + x$$

Calcular $T(2 + 6x)$.

SOLUCIÓN $\mathcal{B} = \{-1 + x, 1 + x\}$ es una base de P_2. Primero expresaremos $2 + 6x$ como una combinación lineal en $\mathcal{B}$:

$$2 + 6x = 2(-1 + x) + 4(1 + x)$$

A continuación aplicaremos T.

$$T(2 + 6x) = 2T(-1 + x) + 4T(1 + x) = 2(-7 + 2x) + 4(4 + x) = 2 + 8x$$

En otras palabras,

$$\begin{bmatrix} -1 & 1 & : & 2 \\ 1 & 1 & : & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & : & 2 \\ 0 & 1 & : & 4 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -7 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$$

$\square$

En general, se tiene el teorema siguiente.

TEOREMA 5

> Sea $T: V \to W$ es una tranformación lineal y sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ el generador de V. Entonces, el conjunto $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ genera al contradominio de T.

DEMOSTRACIÓN Sea $\mathbf{w} \in R(T)$, entonces existe un $\mathbf{v} \in V$ tal que $T(\mathbf{v}) = \mathbf{w}$. Como $\mathcal{B}$ genera a V, hay escalares $c_1, \ldots, c_n$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. Entonces

$$\mathbf{w} = T(\mathbf{v}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n)$$

De aquí que $\mathbf{w}$ sea una combinación lineal de $T(\mathcal{B})$. $\square$

Valores de transformaciones lineales con sistemas algebraicos de cómputo

Para resolver el ejemplo 24 con un sistema algebraico de cómputo:

Maple

```
> with(linalg):
> col(rref([[-1,1,2],[1,1,6]]),3);
                                    [2,4]

> evalm([[-7,4],[2,1]] &* ");
                                    [2,8]
```

Mathematica

```
In[1]:=
    Map[Last,RowReduce[{{-1,1,2},{1,1,6}}]]
Out[1]=
    {2, 4}
In[2]:=
    {{-7,4},{2,1}} . %
Out[2]=
    {2, 8}
```

MATLAB

```
>> A=rref([-1 1 2; 1 1 6]); A(:,3)
ans =
      2
      4
>> [-7 4; 2 1] * ans
ans =
      2
      8
```

Ejercicios 5.2

En los ejercicios 1 a 12 determine el dominio y el codominio de la transformación T, y determine si T es lineal.

1. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x+y \\ x-2y \end{bmatrix}$
2. $T\begin{bmatrix} x \\ y \end{bmatrix} = x\begin{bmatrix} 3 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ -2 \end{bmatrix}$
3. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-1 \\ x-y \end{bmatrix}$
4. $T\begin{bmatrix} x \\ y \end{bmatrix} = x\begin{bmatrix} 3 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ -2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
5. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix}$
6. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} xy \\ -y \end{bmatrix}$
7. $T(\mathbf{x}) = 3\mathbf{x}, \mathbf{x} \in \mathbf{R}^2$
8. $T(\mathbf{x}) = 3\mathbf{x} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{x} \in \mathbf{R}^2$
9. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \\ 0 \end{bmatrix}$
10. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ y+z \end{bmatrix}$
11. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$
12. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ y+z \\ -7z \end{bmatrix}$

Aplique el teorema 3 para demostrar que las dos transformaciones siguientes son lineales.

13. $T : \mathbf{R}^3 \to \mathbf{R}^2, T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y+z \\ x-y-z \end{bmatrix}$
14. $T : \mathbf{R}^3 \to \mathbf{R}^2, T(\mathbf{x}) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \end{bmatrix}\mathbf{x}$

En los ejercicios 15 a 20 determine si $T: P_1 \to P_1$ es lineal.

15. $T(a + bx) = (3a + b) + (a - 2b)x$

16. $T(a + bx) = (a - b) + (a + b + 1)x$

17. $T(a + bx) = (a - 5b) + abx$

18. $T(p) = 10p$

19. $T(p) = 10p - 2$

20. $T(p) = 10p^2$

21. ¿Es lineal $T: P_2 \to P_1$?

$$T\left(a + bx + cx^2\right) = (a - b) + (b + c)x$$

22. ¿Es lineal $T: M_{22} \to \mathbf{R}^2$?

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$$

23. ¿Es lineal $T: \mathbf{R}^2 \to M_{22}$?

$$T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 2 \end{bmatrix}$$

24. Sea T la transformación lineal tal que

$$T\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -8 \end{bmatrix}, \qquad T\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -9 \\ 4 \end{bmatrix}$$

Determine $T\begin{bmatrix} x \\ y \end{bmatrix}, T\begin{bmatrix} 10 \\ -15 \end{bmatrix}$, y $T\begin{bmatrix} y \\ x \end{bmatrix}$

25. Sea $T: \mathbf{R}^3 \to \mathbf{R}^2$ la transformación lineal tal que

$$T(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$$T(-\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

$$T(\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Encuentre $T(\mathbf{x})$ y $T\begin{bmatrix} -10 \\ 15 \\ -25 \end{bmatrix}$

26. Sea $T: P_1 \to P_1$ la transformación lineal tal que

$$T(-1 + x) = -8 + 5x, \qquad T(2 + 2x) = 4 - 9x$$

Calcule $T(a + bx)$, $T(-15 + 10x)$ y $T(b + ax)$.

27. Sea $T: P_2 \to P_1$ la transformación lineal tal que

$$T\left(1 + x + x^2\right) = -1 + 3x$$

$$T\left(1 + x - x^2\right) = -3 + 2x$$

$$T\left(1 - x + x^2\right) = 1 + 2x$$

Determine $T(a + bx + cx^2)$ y $T(-25 + 15x - 10x^2)$.

28. Sea $T: P \to P$ la transformación lineal que satisface

$$T(x^n) = \frac{1}{n+1}x^{n+1}, \qquad n \geq 0$$

Calcule: $T(x + x^2)$, $T(-1 + x^3)$ y $T((1 + x^2)^2)$.

29. Sea C una matriz invertible $n \times n$. Demuestre que $T: M_{nn} \to M_{nn}$ es lineal.

$$T(X) = C^{-1}XC$$

30. Sea C una matriz invertible $n \times n$. Compruebe que $T: M_{nn} \to M_{nn}$ es lineal.

$$T(X) = C^{-1}XC - X$$

31. Explique por qué una transformación lineal T tal que

$$T\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad T\begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} -6 \\ 2 \end{bmatrix}$$

no puede calcularse en forma única. Determine al menos dos de esas transformaciones lineales.

32. Demuestre que la transformación de identidad es lineal.

33. Cite un ejemplo de una transformación T no lineal que tenga la propiedad $T(\mathbf{0}) = \mathbf{0}$.

34. Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ una transformación lineal. Demuestre que si $\{\mathbf{v}_1, \mathbf{v}_2\} \subseteq \mathbf{R}^2$ es linealmente dependiente, entonces $\{T(\mathbf{v}_1), T(\mathbf{v}_2)\}$ es linealmente dependiente.

35. Dé un ejemplo de una transformación lineal $T: \mathbf{R}^2 \to \mathbf{R}^2$ y linealmente independiente de los vectores 2 $\mathbf{v}_1, \mathbf{v}_2$; de modo que $\{T(\mathbf{v}_1), T(\mathbf{v}_2)\}$ sea linealmente dependiente.

36. Sea $T: V \to W$ una transformación lineal de los espacios vectoriales V y W. Demuestre que si $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq V$ es linealmente dependiente, entonces $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\} \subseteq W$ es linealmente dependiente.

37. Demuestre que cualquier transformación lineal $T: \mathbf{R}^2 \to \mathbf{R}^2$ mapea rectas en rectas o en puntos.

38. Compruebe que

$$T: \mathbf{R}^m \to \mathbf{R}^m, \qquad T(\mathbf{x}) = \mathbf{x} + \mathbf{b}, \qquad \mathbf{b} \neq \mathbf{0}$$

es no lineal. Esa transformación se llama **traslación** de $\mathbf{R}^m$.

39. Determine los tamaños de A y $\mathbf{b}$ y demuestre que

$$T: \mathbf{R}^n \to \mathbf{R}^m, \qquad T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}, \qquad \mathbf{b} \neq \mathbf{0}$$

es no lineal. Esa transformación se llama **transformación afín**.

40. Demuestre que una translación (definida en el ejercicio 38) es un caso especial de transformación afín (definida en el ejercicio 39).

En los ejercicios 41 a 47 demuestre que las transformaciones lineales son transformaciones matriciales y determine sus matrices estándar.

41. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ y \\ x \end{bmatrix}$

42. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ y-4z \end{bmatrix}$

43. $0 : \mathbf{R}^2 \to \mathbf{R}^3$

44. $T : \mathbf{R}^3 \to \mathbf{R}^3,\ T(\mathbf{v}) = -10\mathbf{v}$

45. La reflexión respecto al eje y en el plano.

46. La reflexión respecto a la recta $y = -x$ en el plano.

47. La rotación θ rad en sentido contrario al de las manecillas del reloj en torno al origen en el plano.

48. Sea $\mathcal{B}$ una base de $\mathbf{R}^n$ y sea T la transformación

$$T : \mathbf{R}^n \to \mathbf{R}^n, \quad T(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$$

a. Demuestre que T es lineal.

b. Sean $n = 2$ y $[\mathbf{e}_1]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$, $[\mathbf{e}_2]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$.

Calcule $T\begin{bmatrix} -9 \\ 6 \end{bmatrix}$

49. Determine todas las transformaciones lineales de $\mathbf{R}$ a $\mathbf{R}$.

En el ejemplo 16 demostramos que el producto punto por un vector fijo en $\mathbf{R}^n$ es una transformación lineal. En el siguiente ejercicio comprobaremos que esas transformaciones son las *únicas* lineales de $\mathbf{R}^n$ a $\mathbf{R}$.

50. Sea $T : \mathbf{R}^n \to \mathbf{R}$ lineal. Demuestre que existe un vector n $\mathbf{u}$, tal que

$$T(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \quad \text{para todo } \mathbf{v} \in \mathbf{R}^n$$

(*Sugerencia:* Sea $\mathbf{u} = (T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n))$.)

5.3 Núcleo y contradominio

Objetivos del estudiante para esta sección

1. Aprender a determinar el núcleo y el contradominio de una transformación lineal.
2. Conocer y aplicar el teorema de la dimensión.
3. Comprender los conceptos biunívoco, sobre e isomorfismo.

En esta sección dirigiremos nuestra atención hacia dos subespacios básicos que acompañan a toda transformación lineal, el *núcleo* y el *contradominio*. Además describiremos cómo distinguir entre dos espacios vectoriales distintos y cuándo considerarlos iguales o *isomórficos*.

DEFINICIÓN

(Núcleo)

Sea $T : V \quad W$ una transformación lineal. El **núcleo**, Ker(T)[8] de T contiene todos los vectores en V que se mapean a cero en W.

$$\mathrm{Ker}(T) = \{\mathbf{v} \in V, \quad T(\mathbf{v}) = \mathbf{0} \in W\}$$

Recordemos que el **contradominio** R(T) de T es el conjunto de todas las imágenes de T en W.

$$\mathrm{R}(T) = \{\mathbf{w} \in W, \quad \mathbf{w} = T(\mathbf{v}) \quad \text{para algún } \mathbf{v} \in V\}$$

Observe que tanto Ker(T) como R(T) de una transformación lineal T son conjuntos no vacíos: $T(\mathbf{0}) = \mathbf{0}$ implica que $\mathbf{0} \in V$ está en Ker(T) y que $\mathbf{0} \in W$ está en R(T).

[8] N. del T: La representación "Ker(T)" proviene de la palabra inglesa "*kernel*", que significa "núcleo".

■ EJEMPLO 25 Determine el núcleo y el contradominio de

(a) La transformación lineal cero $\mathbf{0} : V \to W$;
(b) La transformación identidad lineal $I : V \to V$;
(c) La proyección $p : \mathbf{R}^2 \to \mathbf{R}^2$, $p(x, y) = (x, 0)$.

SOLUCIÓN

(a) $\mathbf{0}(\mathbf{v}) = \mathbf{0}$ para todo $\mathbf{v} \in V$, de modo que el núcleo es V. Como $\mathbf{0}$ es la única imagen, el contradominio es $\{\mathbf{0}\}$.

$$\mathrm{Ker}(\mathbf{0}) = V, \qquad \mathrm{R}(\mathbf{0}) = \{\mathbf{0}\}$$

(b) Ya que $I(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v} \in V$, todo vector no cero se mapea en un vector no cero. Así, el núcleo es $\{\mathbf{0}\}$. Como todo $\mathbf{v}$ es su propia imagen, el contradominio es V.

$$\mathrm{Ker}(I) = \{\mathbf{0}\}, \qquad \mathrm{R}(I) = V$$

(c) (x, y) está en $\mathrm{Ker}(p)$ si y sólo si $p(x, y) = (x, 0) = (0, 0)$. Así, $x = 0$ y el núcleo consiste en los puntos $(0, y)$. También (z, w) está en el contradominio si y sólo si hay (x, y) tales que $p(x, y) = (x, 0) = (z, w)$. Por consiguiente, $w = 0$, y el contradominio consiste en los puntos $(x, 0)$ véase la figura 5.19:

$$\mathrm{Ker}(p) = \{(0, y), \quad y \in \mathbf{R}\}, \qquad \mathrm{R}(p) = \{(x, 0), \quad x \in \mathbf{R}\} \qquad \square$$

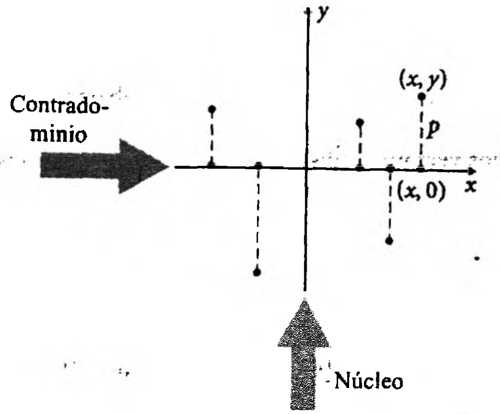


Figura 5.19 El núcleo y el contradominio de la proyección sobre el eje x.

■ EJEMPLO 26 Determine el núcleo de $T : \mathbf{R}^3 \to \mathbf{R}^2$.

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - z \\ y + z \end{bmatrix}$$

SOLUCIÓN Ker(T) es el conjunto de todos los vectores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ tales que $\begin{bmatrix} x-z \\ y+z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ Al resolver el sistema $x - z = 0$, $y + z = 0$, se obtiene $(r, -r, r)$, $r \in \mathbf{R}$. De ahí que,

$$\mathrm{Ker}(T) = \left\{ \begin{bmatrix} r \\ -r \\ r \end{bmatrix}, r \in \mathbf{R} \right\} = \mathrm{Gen} \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\} \qquad \square$$

La propiedad fundamental del núcleo y del contradominio es que ambos son espacios vectoriales.

TEOREMA 6

> Sea $T: V \to W$ una transformación lineal. Entonces
> 1. Ker(T) es un subespacio de V;
> 2. R(T) es un subespacio de W.

DEMOSTRACIÓN

1. Sean $\mathbf{u}, \mathbf{v} \in \mathrm{Ker}(T)$ y sea $c \in \mathbf{R}$. Como Ker(T) es no vacío, bastará demostrar que $\mathbf{u} + \mathbf{v}$, $c\mathbf{u} \in \mathrm{Ker}(T)$. T es lineal, de modo que

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

$$T(c\mathbf{u}) = cT(\mathbf{u}) = c\mathbf{0} = \mathbf{0}$$

Por consiguiente, $\mathbf{u} + \mathbf{v}$, $c\mathbf{u} \in \mathrm{Ker}(T)$. Así, Ker($T$) es un subespacio de V.

2. Sean $\mathbf{u}', \mathbf{v}' \in \mathrm{R}(T)$ y sea $c \in \mathbf{R}$. Entonces hay vectores $\mathbf{u}$ y $\mathbf{v}$ de V tales que $\mathbf{u}' = T(\mathbf{u})$ y $\mathbf{v}' = T(\mathbf{v})$. Como T es lineal,

$$\mathbf{u}' + \mathbf{v}' = T(\mathbf{u}) + T(\mathbf{v}) = T(\mathbf{u} + \mathbf{v})$$

$$c\mathbf{u}' = cT(\mathbf{u}) = T(c\mathbf{u})$$

Ya determinamos vectores $\mathbf{u} + \mathbf{v}$ y $c\mathbf{u}$ que se mapean en $\mathbf{u}' + \mathbf{v}'$ y en $c\mathbf{u}'$, respectivamente. Por tanto, $\mathbf{u}' + \mathbf{v}'$ y $c\mathbf{u}'$ están en R(T). En vista de lo anterior, R(T) es un subespacio de W. $\square$

La dimensión del núcleo se llama **nulidad** de T. La dimensión del contradominio se llama **rango** de T.

■ EJEMPLO 27 Determine la nulidad y el rango de T en el ejemplo 26.

SOLUCIÓN Como $\mathrm{Ker}(T) = \{(r, -r, r), r \in \mathbf{R}\} = \mathrm{Gen}\{(1, -1, 1)\}$, la nulidad es 1. Puesto que $\mathrm{R}(T) = \mathbf{R}^2$, el rango es 2 (figura 5.20). $\square$

■ EJEMPLO 28 Calcule el núcleo y el contradominio de T del ejemplo 16, sección 5.2.

SOLUCIÓN El núcleo contiene todos los vectores $\mathbf{v}$ tales que $\mathbf{u} \cdot \mathbf{v} = 0$, es decir, todos los vectores n *ortogonales* a $\mathbf{u}$. Éste es el hiperplano que pasa por el origen cuya normal es $\mathbf{u}$. Para

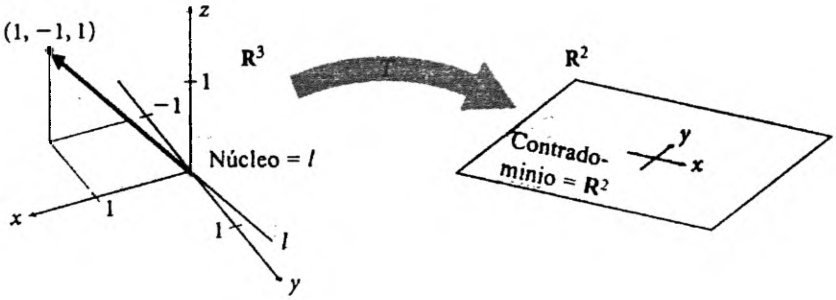


Figura 5.20 El núcleo y el contradominio de $T(x, y, z) = (x - z, y + z)$.

determinar el contradominio se observa que como **u** es distinto de cero,

$$T(\mathbf{u}) = \mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2 > 0.$$

Por consiguiente, el número $\|\mathbf{u}\|^2$ distinto de cero está en el contradominio de T. Así, el contradominio contiene al generador de $\|\mathbf{u}\|^2$, que es **R**. Por tanto, $R(T) = \mathbf{R}$ (figura 5.21). ☐

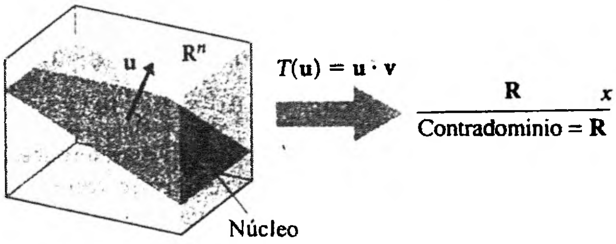


Figura 5.21 Núcleo e imagen del producto punto por un vector fijo.

■ EJEMPLO 29 Determine el núcleo, el contradominio, la nulidad y el rango de T del ejemplo 17, sección 5.2.

SOLUCIÓN Obtendremos *geométricamente* el núcleo y el contradominio, por ser más fácil que un cálculo algebraico. El núcleo está formado por todos los vectores **v** tales que $\mathbf{u} \times \mathbf{v} = \mathbf{0}$, los cuales son paralelos a **u**. Por consiguiente, el núcleo es la recta que pasa por el origen en la dirección de **u**. La nulidad de T es 1.

El contradominio contiene todos los vectores **b** para los que hay un vector **v** tal que $\mathbf{u} \times \mathbf{v} = \mathbf{b}$. Por tanto, **u** y **b** son ortogonales. Así, el contradominio está contenido en el conjunto de vectores ortogonales a **u**. Ahora, si **b** es un vector distinto de cero y ortogonal a **u**. Comprobaremos que se encuentra en el contradominio de T. Los vectores **b** y $(\mathbf{u} \times \mathbf{b}) \times \mathbf{u}$, son paralelos, porque cada uno es perpendicular a **u** y a $\mathbf{u} \times \mathbf{b}$, a la vez. Así, $(\mathbf{u} \times \mathbf{b}) \times \mathbf{u} = c\mathbf{b}$ para algún escalar c distinto de cero. De manera que,

$$\mathbf{u} \times \left(-\frac{1}{c}\mathbf{u} \times \mathbf{b}\right) = \mathbf{b}$$

Así, el vector $-\frac{1}{c}\mathbf{u} \times \mathbf{b}$ se transforma en $\mathbf{b}$; por consiguiente, $\mathbf{b}$ está en el contradominio de T. Con ello hemos demostrado que el contradominio es el plano que pasa por el origen cuya normal es $\mathbf{u}$. El rango de T es 2 (figura 5.22). □

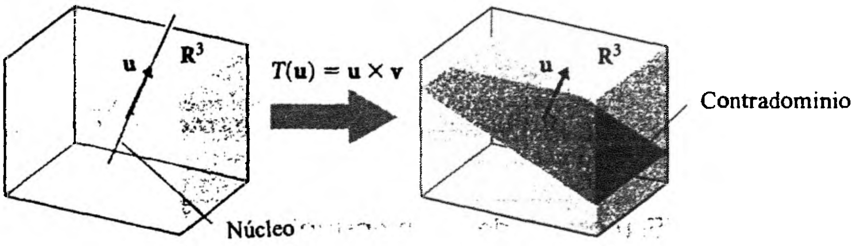


Figura 5.22 Núcleo e imagen del producto cruz por un vector fijo.

En el caso especial en el que la transformación lineal es matricial, se cumple lo siguiente:

TEOREMA 7

Sea $T : \mathbf{R}^n \to \mathbf{R}^m$ es una transformación matricial cuya matriz estándar es A. Entonces

1. $\mathrm{Ker}(T) = V(A)$;
2. $\mathrm{R}(T) = \mathrm{Col}(A)$;
3. $\mathrm{Nulidad}(T) = \mathrm{Nulidad}(A)$;
4. $\mathrm{Rango}(T) = \mathrm{Rango}(A)$.

DEMOSTRACIÓN Puesto que $T(\mathbf{x}) = A\mathbf{x}$, entonces $T(\mathbf{x}) = \mathbf{0}$ si y sólo si $A\mathbf{x} = \mathbf{0}$. De aquí, $\mathrm{Ker}(T) = V(A)$. Igualmente, si $\mathbf{b}$ está en $\mathbf{R}^m$, entonces hay un $\mathbf{x}$ en $\mathbf{R}^n$ tal que $T(\mathbf{x}) = \mathbf{b}$ si y sólo si $A\mathbf{x} = \mathbf{b}$. En consecuencia, $\mathrm{R}(T) = \mathrm{Col}(A)$. Las partes de las nulidades y los rangos se ven a continuación. □

■ EJEMPLO 30 Calcule las bases para el núcleo y el contradominio y determine la nulidad y el rango de

$$T : \mathbf{R}^4 \to \mathbf{R}^3, \qquad T(x, y, z, w) = (x + 3z, y - 2z, w)$$

SOLUCIÓN De acuerdo con el teorema 7, basta expresar a T como transformación matricial y obtener las bases para las columnas y el espacio nulo de su matriz estándar A. A se expresa con

$$A = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

y está en forma escalonada reducida por operaciones de renglón. Por consiguiente, los vectores $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ forman una base para $\mathrm{Col}(A) = \mathrm{R}(T) = \mathbf{R}^3$. Por otra parte, $\{(-3, 2, 1, 0)\}$ es una base para $V(A) = \mathrm{Ker}(T)$. De modo que el rango de T es 3 y la nulidad es 1.

El teorema 7 implica que

$$\mathrm{Ker}(T) = \{\mathbf{0}\} \Leftrightarrow A\mathbf{x} = \mathbf{0} \text{ sólo tiene la solución trivial y}$$

y

$$R(T) = \mathbf{R}^m \Leftrightarrow \text{las columnas de } A \text{ generan a } \mathbf{R}^m \qquad \square$$

El teorema siguiente es fundamental en el álgebra lineal porque generaliza el teorema del rango, y se demostrará en la sección 5.4.

TEOREMA 8

(El teorema de la dimensión)

Si $T: V \to W$ es una transformación lineal de un espacio vectorial V de dimensión finita a un espacio vectorial W, entonces

$$\mathrm{Nulidad}(T) + \mathrm{Rango}(T) = \dim(V)$$

■ EJEMPLO 31 Verifique el teorema de la dimensión para T del ejemplo 14, sección 5.2.

SOLUCIÓN Dejaremos como ejercicio comprobar que

$$\mathrm{Ker}(T) = P_0 = \{a_0, a_0 \in \mathbf{R}\} \subseteq P_2$$

y

$$R(T) = P_1 = \{a_0 + a_1x, a_0, a_1 \in \mathbf{R}\} \subseteq P_2$$

Por consiguiente,

$$\dim(P_2) = 3 = 1 + 2 = \dim(P_0) + \dim(P_1) \qquad \square$$

El teorema 8 es muy útil, porque con frecuencia es mucho más fácil determinar el espacio nulo y la nulidad de una transformación lineal que su contradominio y su rango. como sugiere el trabajo hecho en los ejemplos 28 y 29.

■ EJEMPLO 32 Determine el contradominio de la transformación lineal $T: \mathbf{R}^4 \to P_2$.

$$T(a, b, c, d) = (a - b) + (c + d)x + (2a + b)x^2$$

SOLUCIÓN El espacio nulo de T está generado por $(0, 0, -1, 1)$, que se obtiene de inmediato al resolver el sistema $a - b = 0$, $c + d = 0$, $2a + b = 0$. Por tanto, la nulidad de T es 1. Así, de acuerdo con el teorema de la dimensión,

$$\mathrm{Rango}(T) = \dim(\mathbf{R}^4) - \mathrm{Nulidad}(T) = 4 - 1 = 3$$

De manera que el contradominio es un subespacio tridimensional de P_2, y esto significa que es todo P_2. $\square$

■ EJEMPLO 33 Supongamos que una transformación lineal $T: \mathbf{R}^4 \to \mathbf{R}^3$ tiene su núcleo generado por un vector distinto de cero. ¿Cuál es el contradominio de T?

SOLUCIÓN La nulidad de T es 1. Basándonos en el teorema de la dimensión, tenemos

$$\text{Rango}(T) = 4 - \text{Nulidad}(T) = 4 - 1 = 3.$$

Por consiguiente, el contradominio es un subespacio tridimensional de $\mathbf{R}^3$. Por lo anterior, $R(T) = \mathbf{R}^3$. $\square$

Transformaciones biunívocas, sobre e isomorfismos

Ya sabemos que hay muchos espacios vectoriales en los que se interesan los matemáticos y los científicos. Sin embargo, haciendo a un lado las distintas notaciones de los ejemplos individuales, veremos que muchos de esos espacios son "esencialmente el mismo". En este párrafo analizaremos la noción de que dos espacios vectoriales son el mismo. A esos espacios se les llama isomórficos. Pero comencemos desde lo básico.

La definición de una transformación $T: A \to B$ entre dos conjuntos permite que

1. Dos o más elementos de A tengan la misma imagen;
2. El rango de T esté estrictamente contenido en el codominio B.

Si la condición (1) no se aplica, T es biunívoca.

DEFINICIÓN

(Transformación biunívoca)

Una transformación $T: A \to B$ se llama **biunívoca**, o **uno a uno**, si para cada elemento b del contradominio hay *exactamente un* elemento a cuya imagen es $b = T(a)$. Esto puede enunciarse como sigue:

$$T(a_1) = T(a_2) \Rightarrow a_1 = a_2 \tag{5.2}$$

o bien, en forma equivalente,

$$a_1 \neq a_2 \Rightarrow T(a_1) \neq T(a_2) \tag{5.3}$$

(Véase la figura 5.23.)

Si la afirmación 2 es falsa, T es sobre.

DEFINICIÓN

(Transformación sobre)

Una transformación $T: A \to B$ se llama **sobre** si su contradominio es igual a su codominio, es decir,

$$R(T) = B \tag{5.4}$$

(Véase la figura 5.23.)

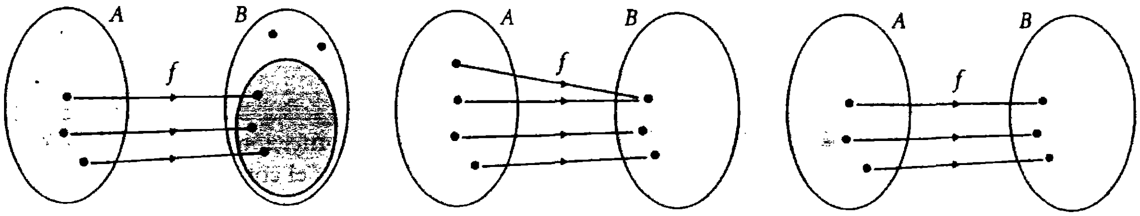


Figura 5.23 Transformación biunívoca y sobre.

■ EJEMPLO 34 ¿Cuáles de las transformaciones son biunívocas? ¿Cuáles son sobre?

(a) $T : \mathbf{R}^2 \to \mathbf{R}^3$, $T(x, y) = (x + y, y, 0)$
(b) $T : \mathbf{R}^3 \to \mathbf{R}^2$, $T(x, y, z) = (x, z)$

SOLUCIÓN

(a) Si $T(x_1, y_1) = T(x_2, y_2)$, entonces $(x_1 + y_1, y_1, 0) = (x_2 + y_2, y_2, 0)$. Por consiguiente, $y_1 = y_2$. Así, $x_1 = x_2$. Por lo anterior, $(x_1, y_1) = (x_2, y_2)$ y T es biunívoca. T no es sobre, porque $(0, 0, 1)$ no es la imagen de algún vector-2.
(b) T no es biunívoca, porque $T(0, 0, 0) = (0, 0) = T(0, 1, 0)$. T es sobre, porque para todo vector-2 (a, b), hay al menos un vector-3 que se transforma en el anterior. Por ejemplo, $T(a, 0, b) = (a, b)$. ☐

El teorema siguiente dice que para demostrar que una transformación lineal es biunívoca, sólo basta demostrar que transforma a 0 y no fijarse en las demás imágenes. Esto reduce la cantidad de trabajo que se invierte para investigar si una transformación *lineal* es biunívoca.

TEOREMA 9

Sea $T : V \to W$ una transformación lineal. Entonces.

$$T \text{ es } \textit{biunívoca} \Leftrightarrow \operatorname{Ker}(T) = \{\mathbf{0}\}$$

DEMOSTRACIÓN Supongamos que T es biunívoca. Si $\mathbf{v}$ está en el núcleo de T, entonces $T(\mathbf{v}) = \mathbf{0}$. Pero sabemos que $T(\mathbf{0}) = \mathbf{0}$. En consecuencia, $T(\mathbf{v}) = T(\mathbf{0})$, lo cual implica que $\mathbf{v} = \mathbf{0}$, porque T es biunívoca. Esto demuestra que todo elemento del núcleo es el vector cero. Por consiguiente, $\operatorname{Ker}(T) = \{\mathbf{0}\}$.

A la inversa, supongamos que $\operatorname{Ker}(T) = \{\mathbf{0}\}$. Demostraremos que es biunívoca. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores de V tales que $T(\mathbf{u}) = T(\mathbf{v})$. Debemos comprobar que $\mathbf{u} = \mathbf{v}$. Como T es lineal.

$$T(\mathbf{u}) = T(\mathbf{v}) \Rightarrow T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{0} \Rightarrow T(\mathbf{u} - \mathbf{v}) = \mathbf{0}$$

Por tanto, $\mathbf{u} - \mathbf{v}$ está en el núcleo de T, que supusimos era $\{\mathbf{0}\}$. Por lo anterior, $\mathbf{u} - \mathbf{v} = \mathbf{0}$, es decir, $\mathbf{u} = \mathbf{v}$. Así, T es biunívoca (figura 5.24). ☐

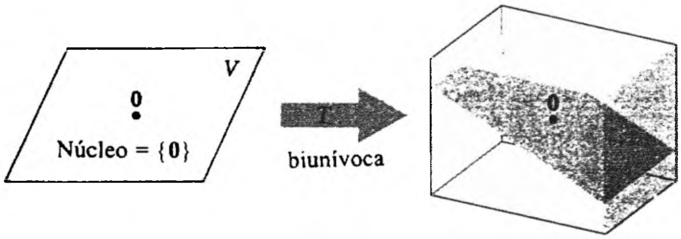


Figura 5.24 El núcleo de una transformación lineal biunívoca es cero.

■ EJEMPLO 35 Demuestre que $T : \mathbf{R}^2 \to \mathbf{R}^2$, $T(x, y) = (x - y, x + 2y)$ es biunívoca y sobre.

SOLUCIÓN Sea $T(x, y) = 0, 0$. Entonces $(x - y, x + 2y) = (0, 0)$ implica que $x = 0$, $y = 0$, es decir $(x, y) = (0, 0)$. Por tanto, según el teorema 9 es biunívoca, y además su nulidad es cero. El rango es 2. Entonces, el contradominio es todo $\mathbf{R}^2$, y la transformación es sobre. □

■ EJEMPLO 36 La transformación del ejemplo 33, ¿es biunívoca? ¿Es sobre?

SOLUCIÓN No es biunívoca, porque el núcleo es distinto de cero. Sí es sobre, porque se determinó que su contradominio se encuentra en $\mathbf{R}^3$. □

■ EJEMPLO 37 Compruebe que $T : P_1 \to P_1$, $T(a + bx) = (a - b) + 2ax$ es biunívoca y sobre.

SOLUCIÓN Se deja como ejercicio demostrar que T es lineal. Sea $p = a + bx$ en el núcleo de T. Entonces, $T(p) = \mathbf{0}$. Así, $(a - b) + 2ax$ es el polinomio cero, lo cual implica que $a - b = 0$ y que $2a = 0$, es decir, $a = b = 0$. Por lo que, $p = \mathbf{0}$ y el núcleo es $\{\mathbf{0}\}$. Entonces T es biunívoca, de acuerdo con el teorema 9, y su nulidad es 0. Por lo anterior el rango es 2, según el teorema de la dimensión. Por tanto, el contradominio es P_1 y la transformación es sobre. □

TEOREMA 10

> Una transformación lineal biunívoca mapea conjuntos linealmente dependientes en conjuntos linealmente independientes. En otras palabras, si $T : V \to W$ es lineal y biunívoca, y si $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un subconjunto linealmente independiente de V, entonces
>
> $$\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$$
>
> es un subconjunto linealmente independiente de W.

DEMOSTRACIÓN Sea

$$c_1 T(\mathbf{v}_1) + \cdots + c_k T(\mathbf{v}_k) = \mathbf{0}$$

entonces

$$T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = \mathbf{0}$$

porque T es lineal. En consecuencia,

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

de acuerdo con el teorema 9. Pero como $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es linealmente independiente, $c_1 = \cdots = c_k = 0$. Entonces, $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ también es linealmente independiente. $\square$

El teorema siguiente permite ahorrar trabajo. Dice que una transformación lineal biunívoca entre espacios de la misma dimensión automáticamente es sobre, y viceversa:

TEOREMA 11

Sea $T: V \to W$ una transformación lineal entre dos espacios vectoriales de dimensiones finitas V y W, siendo $\dim(V) = \dim(W)$. Entonces, T es biunívoca si y sólo si es sobre.

DEMOSTRACIÓN Se deja como ejercicio. (*Sugerencia:* Aplique el teorema 9 y el teorema de la dimensión.) $\square$

DEFINICIÓN

(Isomorfismo)

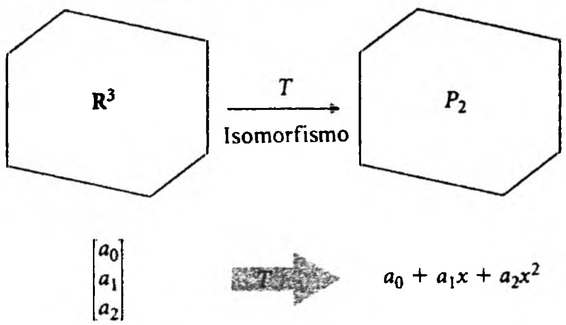
Una transformación lineal entre dos espacios vectoriales que es biunívoca y sobre, al mismo tiempo, se llama **isomorfismo**. Dos espacios vectoriales son **isomorfos** si demuestran esta propiedad entre ellos. Se considera que los espacios isomorfos son iguales, porque sus elementos se corresponden en forma biunívoca y la estructura de las operaciones en el espacio vectorial se conserva a través de la linealidad.

■ EJEMPLO 38 La transformación T del ejemplo 37 es un isomorfismo. $\square$

■ EJEMPLO 39 Demuestre que $\mathbf{R}^n$ y P_{n-1} son isomorfos.

SOLUCIÓN Basta con determinar un isomorfismo entre los dos espacios. Consideraremos $T: \mathbf{R}^n \to P_{n-1}$ definida por

$$T(a_0, \ldots, a_{n-1}) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$$



$\mathbf{R}^3$ y P_2 son isomorfos. $\square$

Dejaremos como ejercicio la demostración de que T es lineal, biunívoca y sobre y, en consecuencia, muestra isomorfismo.

■ EJEMPLO 40 Compruebe que $\mathbf{R}^6$ y M_{23} son isomorfos.

SOLUCIÓN Observamos fácilmente que el mapeo $T: \mathbf{R}^6 \to M_{23}$ definida por

$$T(a_1, \ldots, a_6) = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \end{bmatrix}$$

es lineal, biunívoca y sobre. $\square$

■ EJEMPLO 41 Demuestre que $\mathbf{R}^{mn}$ y M_{mn} son isomorfos.

SOLUCIÓN Se deja como ejercicio. $\square$

■ EJEMPLO 42 Compruebe que M_{32} y M_{23} son isomorfas.

SOLUCIÓN Se ve con facilidad que la transformación $T: M_{32} \to M_{23}$ definida por

$$T\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \\ a_5 & a_6 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \end{bmatrix}$$

es lineal, biunívoca y sobre. $\square$

El criterio más importante para determinar si dos espacios vectoriales son isomorfos es, a la vez, el más sencillo:

TEOREMA 12

Sean V y W espacios vectoriales de dimensión finita. Entonces

$$V \text{ y } W \text{ son isomorfos} \Leftrightarrow \dim(V) = \dim(W)$$

DEMOSTRACIÓN Si V y W son isomorfos, entonces hay un isomorfismo $T: V \to W$. Si $\dim(V) = n$ y $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de V, entonces el conjunto de las imágenes $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ es linealmente independiente, según el teorema 10. Como T es sobre el generador de $T(\mathcal{B})$ está todo en W. Por tanto, $T(\mathcal{B})$ es una base de W porque tiene n elementos. $\dim(W) = n$.

A la inversa, supongamos que V y W tienen la misma dimensión, n. Si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ y $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ son bases de V y W, respectivamente, se establece un isomorfismo $T: V \to W$ en la forma siguiente: si $\mathbf{v}$ está en V, como $\mathcal{B}$ genera a V, hay escalares c_i tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots c_n\mathbf{v}_n$$

Definiremos a T como sigue:

$$T(\mathbf{v}) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$$

La transformación T está bien definida; los coeficientes c_i están determinados en forma única, porque $\mathcal{B}$ es una base. Dejaremos que el lector demuestre que T es lineal. Pero también es biunívoca, porque si $\mathbf{v}$ está en Ker(T), entonces

$$\mathbf{0} = T(\mathbf{v}) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$$

implica que $c_1 = \cdots = c_n = 0$, porque $\mathcal{B}'$ es linealmente independiente. Por consiguiente, $\mathbf{v} = \mathbf{0}$, de acuerdo con el teorema 9. Así, T es sobre, según el teorema 11 y en consecuencia es un isomorfismo.

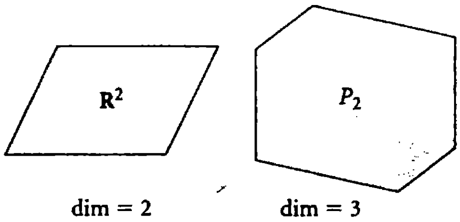
■ EJEMPLO 43 Demuestre que $\mathbf{R}^2$ y $\mathbf{R}^3$ no son isomorfos.

SOLUCIÓN No pueden ser isomorfos, porque no tienen la misma dimensión.

■ EJEMPLO 44 Demuestre que $\mathbf{R}^n$ y P_n no son isomorfos.

SOLUCIÓN Así es, porque

$$n = \dim(\mathbf{R}^n) \neq \dim(P_n) = n + 1$$



$\mathbf{R}^2$ y P_2 *no* son isomorfos.

Núcleo y contradominio con sistemas algebraicos computacionales

Para determinar bases del núcleo y el contradominio de $T(x, y, z) = (x + 2y + 3z, x + 3y + 2z)$:

Maple

```
> with(linalg):
> A:=matrix([[1,2,3],[1,3,2]]):
> nullspace(A);colspace(A);
```

$$\{[-5, 1, 1]\} \qquad \{[0, 1], [1, 0]\}$$

Mathematica

```
In[1]:= A={{1,2,3},{1,3,2}};
In[2]:= NullSpace[A]
        RowReduce[A]
Out[2]= {{-5, 1, 1}}
Out[3]= {{1, 0, 5}, {0, 1, -1}}
```

El contradominio se calcula en forma indirecta

MATLAB

```
>> A=[1 2 3; 1 3 2];
>> null(A), rref(A)
ans =           ans =
  -0.9623
   0.1925       1    0    5
   0.1925       0    1   -1
```

El contradominio se calcula en forma indirecta

Ejercicios 5.3

Núcleo y contradominio

En los ejercicios 1 a 10 determine bases para el núcleo y el contradominio y calcule la nulidad y el rango de T. En cada caso, verifique el teorema de la dimensión.

1. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ y-z \\ -x+z \end{bmatrix}$

2. $T\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x-y \\ y-w \\ z+w \end{bmatrix}$

3. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-2y \\ x-2y \\ 2x-4y \\ 0 \end{bmatrix}$

4. $I : \mathbf{R}^2 \to \mathbf{R}^2$

5. $T : \mathbf{R}^3 \to \mathbf{R}^3$, $T(\mathbf{v}) = -2\mathbf{v}$

6. $T: \mathbf{R}^2 \to \mathbf{R}^2$ es la proyección sobre el eje y.

7. $T : \mathbf{R}^2 \to \mathbf{R}^2$ es la reflexión en la recta $y = -x$.

8. $T : P_2 \to P_2$ de manera que
$$T\left(a + bx + cx^2\right) = (a-b) + (b-c)x + (-a+c)x^2$$

9. $T : P_3 \to P_2$ de modo que
$$T\left(a + bx + cx^2 + dx^3\right) = (a-b) + (b-d)x + (c+d)x^2$$

10. $T : P_1 \to P_3$ de manera que
$$T(a+bx) = (a-2b)x + (a-2b)x^2 + (2a-4b)x^3$$

En los ejercicios 11 a 13 determine una base para el núcleo y el contradominio de $T : P_2 \to P_2$, si T satisface las ecuaciones dadas.

11. $T(1) = x$, $T(x) = x^2$, $T(x^2) = -1$
12. $T(p(x)) = p(1+x)$
13. $T(1) = 0$, $T(x) = 0$, $T(x^2) = 1$

En los ejercicios 14 a 16 determine, para la matriz A, la dimensión del núcleo de la transformación lineal $T : M_{33} \to M_{33}$ definida por
$$T(X) = AX$$

14. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

15. $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

16. $A = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 17 a 19 determine, para n y A dadas, la nulidad y el rango de la transformación lineal $T : M_{22} \to M_{22}$ definida por
$$T(X) = AX - XA$$

17. $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ **18.** $A = E_{12}$ **19.** $A = E_{22}$

20. Con el ejercicio 50 de la sección 5.2, suponga que $\mathbf{u} \neq \mathbf{0}$. Determine el núcleo y el contradominio de T. Si $n = 3$, describa geométricamente al núcleo y al contradominio, y compruebe el teorema de la dimensión.

Transformaciones biunívocas, sobre e isomorfismos

En los ejercicios 21 a 26 aplique el teorema 9 para demostrar que las transformaciones lineales son biunívocas.

21. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ -3x + 4y \end{bmatrix}$

22. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x - y \\ -y + z \\ -3x + z \end{bmatrix}$

23. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - y \\ x - y \\ -x + y \\ x - 2y \end{bmatrix}$

24. $T : P_1 \to P_1$ de manera que

$$T\left(a + bx + cx^2\right) = (2a-b)+(-b+c)x+(-3a+c)x^2$$

25. $T : P_2 \to P_2$ de manera que

$$T(a + bx) = (2a + b) + (-3a + 4b)x$$

26. $T : P_1 \to P_4$ de manera que

$$T(a + bx) = (2a - b) + (a - b)x + (-a + b)x^2 + (a - 2b)x^3$$

En los ejercicios 27 a 32 demuestre que las transformaciones lineales son sobre.

27. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ -x + 2y \end{bmatrix}$

28. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ z \end{bmatrix}$

29. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ -x + y + z \end{bmatrix}$

30. $T(\mathbf{x}) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix}\mathbf{x}$, $\mathbf{x} \in \mathbf{R}^3$

31. $T : P_2 \to P_1$ de modo que

$$T\left(a + bx + cx^2\right) = (a + b) + (a + c)x$$

32. $T : P_2 \to P_2$ de manera que

$$T\left(a + bx + cx^2\right) = c + bx + ax^2$$

En los ejercicios 33 a 37 demuestre que las transformaciones lineales son isomorfismos.

33. $T : P_1 \to P_1$ de modo que

$$T(a + bx) = (a - 2b) + (-2a + b)x$$

34. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - 2y \\ -2x + y \end{bmatrix}$

35. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ -x + y + z \\ -y + z \end{bmatrix}$

36. $T(\mathbf{x}) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & -2 & 6 \end{bmatrix}\mathbf{x}$, $\mathbf{x} \in \mathbf{R}^3$

37. $T : P_2 \to P_2$ de manera que

$$T(a + bx + cx^2) = c + bx + (a - b)x^2$$

¿Cuáles de las transformaciones siguientes son isomorfismos?

38. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ y \end{bmatrix}$ **39.** $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$

40. $I : \mathbf{R}^n \to \mathbf{R}^n, I(\mathbf{x}) = \mathbf{x}$

41. $T : \mathbf{R}^n \to \mathbf{R}^n, T(\mathbf{x}) = -10\mathbf{x}$

42. $T : \mathbf{R}^n \to \mathbf{R}^n, T(\mathbf{x}) = \mathbf{x} + \mathbf{b}, \mathbf{b} \neq \mathbf{0}$

43. $T : \mathbf{R}^n \to \mathbf{R}^m, T(\mathbf{x}) = A\mathbf{x}, m \neq n$

44. Demuestre que todo isomorfismo $T : \mathbf{R}^2 \to \mathbf{R}^2$ transforma líneas rectas en líneas rectas.

45. Sea $\mathcal{B}$ una base de $\mathbf{R}^n$. Demuestre que la transformación de coordenadas es un isomorfismo.

$$T : \mathbf{R}^n \to \mathbf{R}^n, \quad T(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$$

46. Sea A una matriz $n \times n$ de rango n, y sea $\mathbf{b}$ un vector-n. Demuestre que la transformación afín (definida en el ejercicio 39 de la sección 5.2)

$$T : \mathbf{R}^n \to \mathbf{R}^n, \qquad T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}, \qquad \mathbf{b} \neq \mathbf{0}$$

es biunívoca y sobre. ¿Es un isomorfismo?

Algunos resultados teóricos

Si V y W son espacios vectoriales de dimensiones finitas, demuestre las afirmaciones siguientes:

47. Si T es biunívoca, entonces $\dim(V) \leq \dim(W)$.

48. Si T es sobre y un conjunto S genera a V, entonces $T(S)$ genera a W.

49. Si T es sobre, entonces $\dim(V) \geq \dim(W)$.

50. Si T es un isomorfismo y un conjunto $\mathcal{B}$ es una base de V, en ese caso $T(\mathcal{B})$ es una base de W.

5.4 La matriz de una transformación lineal

Objetivos del estudiante para esta sección

1. Aprender a determinar la matriz de una transformación lineal.
2. Evaluar una transformación lineal a partir de su matriz.
3. Calcular la matriz de una transformación lineal con respecto a una base nueva.

En esta sección generalizaremos el concepto de la matriz estándar de una transformación matricial. Demostraremos que *toda* transformación lineal entre espacios vectoriales de dimensiones finitas puede representarse con una transformación matricial. Este resultado útil nos permite evaluar transformaciones lineales mediante la multiplicación de matrices, la cual puede efectuarse en computadora.

TEOREMA 13

(Matriz de una transformación lineal)

Sea $T : V \to W$ una transformación lineal entre dos espacios vectoriales V y W de dimensiones finitas. Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base de V y $\mathcal{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_m\}$ una base de W. La matriz A $m \times n$ cuyas columnas son

$$[T(\mathbf{v}_1)]_{\mathcal{B}'}, \ldots, [T(\mathbf{v}_n)]_{\mathcal{B}'}$$

es la única matriz que satisface

$$[T(\mathbf{v})]_{\mathcal{B}'} = A[\mathbf{v}]_{\mathcal{B}}$$

para todo $\mathbf{v} \in V$.

DEMOSTRACIÓN Como $\mathcal{B}$ genera a V, hay escalares $c_1, \ldots, c_n$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. Así

$$T(\mathbf{v}) = c_1 T(\mathbf{v}_1) + \cdots + c_n T(\mathbf{v}_n)$$

porque T es lineal. Por consiguiente, de acuerdo con el teorema 18 de la sección 4.5,

$$\begin{aligned}[T(\mathbf{v})]_{\mathcal{B}'} &= c_1 [T(\mathbf{v}_1)]_{\mathcal{B}'} + \cdots + c_n [T(\mathbf{v}_n)]_{\mathcal{B}'} \\ &= A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = A[\mathbf{v}]_{\mathcal{B}}\end{aligned}$$

La verificación de que A es la *única* matriz con la propiedad $[T(\mathbf{v})]_{\mathcal{B}'} = A[\mathbf{v}]_{\mathcal{B}}$ para todo $\mathbf{v} \in V$ se deja como ejercicio. ☐

DEFINICIÓN

La matriz A del teorema 13 se llama **matriz de T con respecto a $\mathcal{B}$ y $\mathcal{B}'$**. Si $V = W$ y $\mathcal{B} = \mathcal{B}'$, A se llama **matriz de T con respecto a $\mathcal{B}$** (figura 5.25).

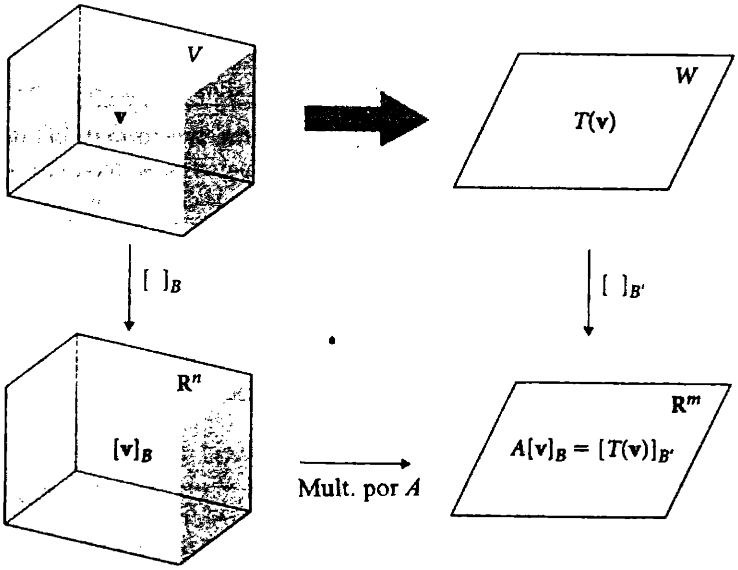


Figura 5.25 Matriz de una transformación lineal.

OBSERVACIONES

1. El teorema 13 es muy útil. Si conocemos A es posible evaluar $T(\mathbf{v})$ calculando $[T(\mathbf{v})]_{\mathcal{B}'}$ como $A[\mathbf{v}]_{\mathcal{B}}$, lo cual es tan sólo una multiplicación de matrices.
2. La matriz de T depende de T, $\mathcal{B}$ y $\mathcal{B}'$. Aun cuando se modifica el orden de los vectores en una de las bases, la matriz de T cambia (véase el ejemplo 48).

El teorema 13 tiene la consecuencia importante de que las únicas transformaciones lineales de $\mathbf{R}^n$ a $\mathbf{R}^m$ son las transformaciones matriciales.

TEOREMA 14

Toda transformación lineal $T : \mathbf{R}^n \to \mathbf{R}^m$ es una transformación matricial.

DEMOSTRACIÓN Sean $\mathcal{B}$ y $\mathcal{B}'$ las bases estándar de $\mathbf{R}^n$ y $\mathbf{R}^m$, respectivamente. Entonces, según el teorema 13, hay una matriz A tal que

$$[T(\mathbf{v})]_{\mathcal{B}'} = A[\mathbf{v}]_{\mathcal{B}}$$

para todo $\mathbf{v} \in \mathbf{R}^n$. Pero como $T(\mathbf{v}) = [T(\mathbf{v})]_{\mathcal{B}'}$ y $A[\mathbf{v}]_{\mathcal{B}} = A\mathbf{v}$ para las bases estándar, tenemos

$$T(\mathbf{v}) = A\mathbf{v}$$

Por consiguiente, T es una transformación matricial cuya matriz estándar es A. ☐

■ EJEMPLO 45 Si V es un espacio vectorial n dimensional, y $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es cualquier base, demuestre que la matriz de la identidad $I : V \to V$ con respecto a $\mathcal{B}$ es I_n.

SOLUCIÓN Como $I(\mathbf{v}_i) = \mathbf{v}_i$ y $[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i$, entonces $[I(\mathbf{v}_i)]_{\mathcal{B}} = \mathbf{e}_i$ para $i = 1, \ldots, n$. Por consiguiente, las columnas de la matriz de I con respecto a $\mathcal{B}$ son $\mathbf{e}_1, \ldots, \mathbf{e}_n$, y en consecuencia esta matriz es I_n. ☐

■ EJEMPLO 46 Sea $T : \mathbf{R}^2 \to \mathbf{R}^3$ la transformación lineal definida por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ x - y \\ x + 4y \end{bmatrix}$$

y sean $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ y $\mathcal{B}' = \{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3'\}$ las bases de $\mathbf{R}^2$ y $\mathbf{R}^3$, donde

$$\mathbf{v}_1 = \mathbf{e}_2, \quad \mathbf{v}_2 = \mathbf{e}_1 \qquad \text{y} \qquad \mathbf{v}_1' = \mathbf{e}_3, \quad \mathbf{v}_2' = \mathbf{e}_2, \quad \mathbf{v}_3' = \mathbf{e}_1$$

respectivamente.

(a) Determine la matriz A de T con respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$.
(b) La matriz estándar de T ¿es igual que A del inciso (a)?
(c) Evalúe $T\begin{bmatrix} -4 \\ 6 \end{bmatrix}$ en forma directa y a partir del inciso (a).

SOLUCIÓN

(a) En este caso,

$$T(\mathbf{e}_2) = T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \quad \text{y} \quad T(\mathbf{e}_1) = T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

A continuación necesitamos $[T(\mathbf{e}_2)]_{\mathcal{B}'}$ y $[T(\mathbf{e}_1)]_{\mathcal{B}'}$. Es fácil verificar que

$$\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}_{\mathcal{B}'} = \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}'} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Por consiguiente,

$$A = \begin{bmatrix} 4 & 1 \\ -1 & 1 \\ 1 & 2 \end{bmatrix}$$

(b) La matriz estándar de T, que también es la matriz de T con respecto a la base estándar, es

$$\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 1 & 4 \end{bmatrix}$$

que no es igual que A.

(c) Al sustituir en la fórmula para T, se obtiene

$$T\begin{bmatrix} -4 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 \\ -10 \\ 20 \end{bmatrix}$$

Por otro lado, para usar A se necesita $\begin{bmatrix} -4 \\ 6 \end{bmatrix}_{\mathcal{B}}$, que es $\begin{bmatrix} 6 \\ -4 \end{bmatrix}$. De manera que,

$$\left[T\begin{bmatrix} -4 \\ 6 \end{bmatrix}\right]_{\mathcal{B}'} = \begin{bmatrix} 4 & 1 \\ -1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -4 \end{bmatrix} = \begin{bmatrix} 20 \\ -10 \\ -2 \end{bmatrix}$$

Así, por la definición de un vector coordenado con respecto a $\mathcal{B}'$,

$$T\begin{bmatrix} -4 \\ 6 \end{bmatrix} = 20\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 10\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -10 \\ 20 \end{bmatrix}$$

■ EJEMPLO 47 Sea $T : P_1 \to P_2$ la transformación lineal definida por $T(a + bx) = ax + bx^2$.

(a) Encuentre la matriz A de T con respecto a las bases $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ y $\mathcal{B}' = \{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3'\}$, siendo

$$\mathbf{v}_1 = x, \quad \mathbf{v}_2 = 1 \qquad \text{y} \qquad \mathbf{v}_1' = 1, \quad \mathbf{v}_2' = x, \quad \mathbf{v}_3' = x^2$$

(b) Evalúe $T(-3 + 4x)$ en forma directa y usando A, $\mathcal{B}$ y $\mathcal{B}'$.

(c) Recupere la fórmula general de T usando A, $\mathcal{B}$ y $\mathcal{B}'$.

SOLUCIÓN

(a) Como $T(x) = x^2$ y $T(1) = x$,

$$[T(x)]_{\mathcal{B}'} = [x^2]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

y

$$[T(1)]_{\mathcal{B}'} = [x]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Por consiguiente,

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(b) La evaluación directa llega a $T(-3 + 4x) = -3x + 4x^2$. Por otro lado, como

$$[-3 + 4x]_{\mathcal{B}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

sucede que

$$[T(-3 + 4x)]_{\mathcal{B}'} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 4 \end{bmatrix}$$

Por tanto,

$$T(-3 + 4x) = 0 \cdot 1 - 3 \cdot x + 4 \cdot x^2$$

y se llega a la misma respuesta.

(c) Puesto que

$$[a + bx]_{\mathcal{B}} = \begin{bmatrix} b \\ a \end{bmatrix}$$

entonces

$$[T(a + bx)]_{\mathcal{B}'} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ a \\ b \end{bmatrix}$$

en consecuencia

$$T(a + bx) = 0 \cdot 1 + a \cdot x + b \cdot x^2$$

■ EJEMPLO 48 Sea $T : P_2 \to P_2$ la transformación lineal definida por $T(a + bx + cx^2) = b + 2cx$.

(a) Obtenga la matriz A de T con respecto a la base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ donde $\mathbf{v}_1 = x^2$, $\mathbf{v}_2 = x$ y $\mathbf{v}_3 = 1$.

(b) Determine la matriz A' de T con respecto a la base de $\mathcal{B}' = \{\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3'\}$, siendo $\mathbf{v}_1' = 1$, $\mathbf{v}_2' = x$ y $\mathbf{v}_3' = x^2$.

(c) Evalúe $T(3x - 4x^2)$: (i) directamente, (ii) a partir de A y (iii) usando de A'.

SOLUCIÓN

(a) Como $T(x^2) = 2x$, $T(x) = 1$ y $T(1) = 0$, entonces

$$[T(x^2)]_{\mathcal{B}} = [2x]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \qquad [T(x)]_{\mathcal{B}} = [1]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

$$[T(1)]_{\mathcal{B}} = [0]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

De ahí que,

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

(b) Para este caso,

$$[T(1)]_{\mathcal{B}'} = [0]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad [T(x)]_{\mathcal{B}'} = [1]_{\mathcal{B}'} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$\left[T(x^2)\right]_{\mathcal{B}'} = [2x]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

Por consiguiente,

$$A' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

(c) (i) Con evaluación directa se obtiene $T(3x - 4x^2) = 3 - 8x$.

(ii) Porque

$$\left[3x - 4x^2\right]_{\mathcal{B}} = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$$

tenemos

$$\left[T(3x - 4x^2)\right]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -8 \\ 3 \end{bmatrix}$$

Por tanto,

$$T(3x - 4x^2) = 0 \cdot x^2 - 8 \cdot x + 3 \cdot 1 = 3 - 8x$$

y se obtiene la misma respuesta.

(iii) De forma semejante, ya que

$$\left[3x - 4x^2\right]_{\mathcal{B}'} = \begin{bmatrix} 0 \\ 3 \\ -4 \end{bmatrix}$$

entonces

$$\left[T(3x - 4x^2)\right]_{\mathcal{B}'} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 3 \\ -8 \\ 0 \end{bmatrix}$$

Por consiguiente

$$T(3x - 4x^2) = 3 \cdot 1 - 8 \cdot x + 0 \cdot x^2 = 3 - 8x$$

Observe que en el ejemplo 48, $A \neq A'$. □

TEOREMA 15

> Sea $T : V \to W$ una transformación lineal entre dos espacios vectoriales de dimensiones finitas, V y W. Sea A la matriz de T con respecto a las bases $\mathcal{B} = \{\mathbf{v}_1, \ldots, V_n\} \subseteq V$ y $\mathcal{B}' = \{\mathbf{v}_1', \ldots, \mathbf{v}_m'\} \subseteq W$. Entonces
>
> 1. **v** está en el núcleo de T si y sólo si $[\mathbf{v}]_{\mathcal{B}}$ está en el espacio nulo de A;
> 2. **w** está en el contradominio de T si y sólo si $[\mathbf{w}]_{\mathcal{B}'}$ se encuentra en el espacio de columnas de A.

DEMOSTRACIÓN

1. La equivalencia es válida, porque

$$\begin{aligned} \mathbf{v} \in \mathrm{Ker}(T) &\Leftrightarrow T(\mathbf{v}) = \mathbf{0} \in W \\ &\Leftrightarrow [T(\mathbf{v})]_{\mathcal{B}'} = \mathbf{0} \in \mathbf{R}^m && \text{Según el teorema 18 de la sección 4.5} \\ &\Leftrightarrow A[\mathbf{v}]_{\mathcal{B}} = \mathbf{0} \in \mathbf{R}^m && \text{Según el teorema 13} \\ &\Leftrightarrow [\mathbf{v}]_{\mathcal{B}} \in \nu(A) \end{aligned}$$

2. Se deja como ejercicio la demostración de esta parte. □

El teorema 15 implica que $\mathrm{Ker}(T) = \{\mathbf{0}\}$ si y sólo si $\mathbf{v}(A) = \{\mathbf{0}\}$ y también que $R(T) = W$ si y sólo si $\mathrm{Col}(A) = \mathbf{R}^m$.

TEOREMA 16

> Siguiendo la notación del teorema 15:
>
> 1. T es biunívoca si y sólo si A tiene n pivotes;
> 2. T es sobre si y sólo si A tiene m pivotes;
> 3. T es un isomorfismo si y sólo si A es invertible.

Cambio de base y la matriz de una transformación lineal

En este párrafo estudiaremos cómo se afecta la matriz de una transformación lineal $T : V \to V$ cuando se modifican las bases en V. En general, una transformación lineal tiene distintas matrices con respecto a diferentes bases (véase el ejemplo 48). A veces hay bases que producen una matriz muy sencilla para T, como una *matriz diagonal*. En este caso, la evaluación de T es muy sencilla. El teorema siguiente nos dice cómo determinar una nueva representación (potencialmente fácil) matricial de T a partir de otra anterior.

TEOREMA 17

> Sea $T : V \to V$ una transformación lineal de un espacio vectorial V de dimensión finita en sí mismo. Sean $\mathcal{B}$ y $\mathcal{B}'$ dos bases de V, y sea P la matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$. Si A es la matriz de T con respecto a $\mathcal{B}$ y si A' es la matriz de T con respecto a $\mathcal{B}'$, entonces
>
> $$A' = P^{-1}AP$$

DEMOSTRACIÓN Como P es la matriz de transición de $\mathcal{B}'$ a $\mathcal{B}$, entonces P^{-1} es la matriz de transición de $\mathcal{B}$ a $\mathcal{B}'$ (según el corolario 21 de la sección 4.5). Por consiguiente, $[\mathbf{w}]_{\mathcal{B}'} = P^{-1}[\mathbf{w}]_{\mathcal{B}}$ para todo $\mathbf{w}$ en V. En particular, $[T(\mathbf{v})]_{\mathcal{B}'} = P^{-1}[T(\mathbf{v})]_{\mathcal{B}}$ para todo $\mathbf{v}$ en V. Así,

$$\begin{aligned}[T(\mathbf{v})]_{\mathcal{B}'} &= P^{-1}[T(\mathbf{v})]_{\mathcal{B}} = P^{-1}(A[\mathbf{v}]_{\mathcal{B}}) \\ &= \left(P^{-1}A\right)[\mathbf{v}]_{\mathcal{B}} = P^{-1}A(P[\mathbf{v}]_{\mathcal{B}'}) = (P^{-1}AP)[\mathbf{v}]_{\mathcal{B}'}\end{aligned}$$

Entonces, la matriz $P^{-1}AP$ satisface $[T(\mathbf{v})]_{\mathcal{B}'} = (P^{-1}AP)[\mathbf{v}]_{\mathcal{B}'}$ para todo $\mathbf{v}$ en V. Por tanto, debe ser la matriz de transición de T con respecto a $\mathcal{B}'$ (figura 5.26). $\square$

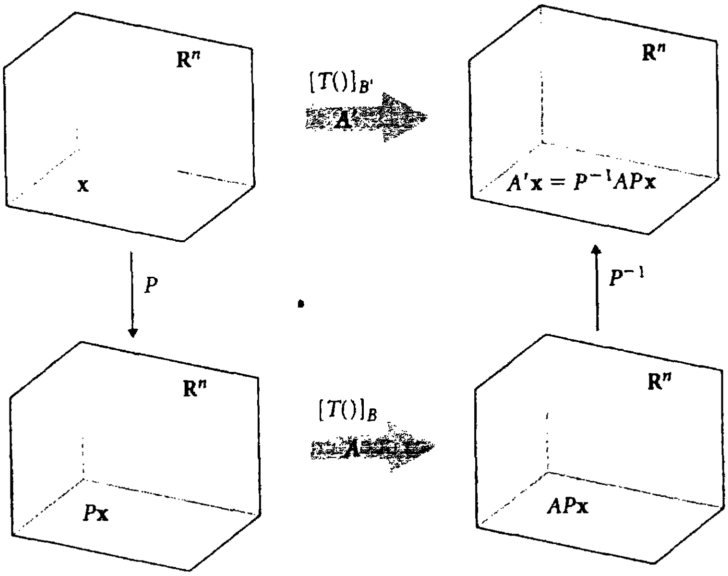


Figura 5.26 El efecto del cambio de base.

Ahora ilustraremos el teorema 17.

■ EJEMPLO 49 Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ la transformación lineal expresada por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -5x + 6y \\ -3x + 4y \end{bmatrix}$$

y sean $\mathcal{B}$ y $\mathcal{B}'$ las bases

$$\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} \qquad \mathcal{B}' = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$$

(a) Determine la matriz A de T con respecto a $\mathcal{B}$. (Es la matriz estándar.)
(b) Obtenga la matriz de transición P de $\mathcal{B}'$ a $\mathcal{B}$.
(c) Aplique el teorema 17 para calcular la matriz de T con respecto a $\mathcal{B}'$.
(d) Calcule la matriz A' de T con respecto a $\mathcal{B}'$ directamente a partir de $\mathcal{B}'$.

SOLUCIÓN

(a) Ya que

$$T\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}-5\\-3\end{bmatrix},\qquad T\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}6\\4\end{bmatrix}$$

la matriz estándar A de T es

$$A=\begin{bmatrix}-5 & 6\\-3 & 4\end{bmatrix}$$

(b) Para encontrar P se necesitan los vectores coordenados de $\mathcal{B}'$ con respecto a $\mathcal{B}$. Tenemos que

$$\begin{bmatrix}1\\1\end{bmatrix}_{\mathcal{B}}=\begin{bmatrix}1\\1\end{bmatrix},\qquad \begin{bmatrix}2\\1\end{bmatrix}_{\mathcal{B}}=\begin{bmatrix}2\\1\end{bmatrix}$$

En consecuencia,

$$P=\begin{bmatrix}1 & 2\\1 & 1\end{bmatrix}$$

(c) Es fácil ver que

$$P^{-1}=\begin{bmatrix}-1 & 2\\1 & -1\end{bmatrix}$$

De acuerdo con el teorema 17,

$$A'=P^{-1}AP=\begin{bmatrix}-1 & 2\\1 & -1\end{bmatrix}\begin{bmatrix}-5 & 6\\-3 & 4\end{bmatrix}\begin{bmatrix}1 & 2\\1 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & -2\end{bmatrix}$$

(d) Calcularemos A' directamente de $\mathcal{B}'$ y de T. Al evaluar T en $\mathcal{B}'$ obtenemos

$$T\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix},\qquad T\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}-4\\-2\end{bmatrix}$$

Como

$$\begin{bmatrix}1\\1\end{bmatrix}_{\mathcal{B}'}=\begin{bmatrix}1\\0\end{bmatrix},\qquad \begin{bmatrix}-4\\-2\end{bmatrix}_{\mathcal{B}'}=\begin{bmatrix}0\\-2\end{bmatrix}$$

tenemos que

$$A'=\begin{bmatrix}1 & 0\\0 & -2\end{bmatrix}$$

que es la misma respuesta que la del inciso (c). ☐

Observamos que, en el ejemplo 49, la matriz A' es una matriz diagonal.

DEFINICIÓN

(Matrices semejantes)

Sean A y B dos matrices (cuadradas) $n \times n$. Se dice que B es **semejante** o **similar** a A, si existe una matriz P invertible tal que $B = P^{-1}AP$.

Mencionaremos los siguientes hechos básicos (cuyas demostraciones se dejan como ejercicios):

1. A es semejante a sí misma.
2. Si B es semejante a A, entonces A es similar a B.
3. Si B es semejante a A y C es similar a B, entonces C es semejante a A.

El teorema 17 se puede reformular diciendo que
Las matrices de una transformación lineal con respecto a dos bases son semejantes.

De hecho, como veremos en los ejercicios, dos matrices semejantes dan lugar a la *misma* transformación lineal con respecto a bases diferentes.

Demostración del teorema de la dimensión

Ahora comprobaremos el teorema de la dimensión (sección 5.3), el cual establece que

Si $T : V \to W$ es una transformación lineal de un espacio vectorial de dimensión finita en un espacio vectorial W, entonces

$$\text{Nulidad}(T) + \text{Rango}(T) = \dim(V)$$

DEMOSTRACIÓN Como V es de dimensión finita, el teorema 5 de la sección 5.2 implica que $R(T)$ es de dimensión finita. En consecuencia, puede considerarse que T es una transformación lineal entre los dos espacios vectoriales de dimensiones finitas V y $R(T)$. Sea A la matriz de T con respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ de V y de $R(T)$, respectivamente. Entonces la cantidad de columnas de A es $\dim(V)$. Los teoremas 18 y 19 de la sección 4.5, junto con el teorema 15, implican que

$$\text{Nulidad}(T) = \text{Nulidad}(A) \quad \text{y que} \quad \text{Rango}(T) = \text{Rango}(A)$$

El teorema, ahora, es consecuencia del teorema del rango (sección 4.6). $\square$

Ejercicios 5.4

Sean $\mathcal{B}$, $\mathcal{B}'$ y $\mathcal{B}''$ las siguientes bases de P_1, P_2 y P_3, respectivamente.

$$\mathcal{B} = \{1 + x, -1 + x\} \subseteq P_1$$
$$\mathcal{B}' = \{-x + x^2, 1 + x, x\} \subseteq P_2$$
$$\mathcal{B}'' = \{-x + x^3, 1 + x^2, x, -1 + x\} \subseteq P_3$$

Sean p, p' y p'' polinomios cualesquiera de P_1, P_2 y P_3, respectivamente.

$$p = a + bx \quad \text{en} \quad P_1$$
$$p' = a + bx + cx^2 \quad \text{en} \quad P_2$$
$$p'' = a + bx + cx^2 + dx^3 \quad \text{en} \quad P_3$$

En los ejercicios 1 a 3, determine la matriz de T con respecto a cada base de las siguientes:

1. $\mathcal{B}$ y $\mathcal{B}'$, si $T(p) = (a - b) + bx + ax^2$.
2. $\mathcal{B}'$ y $\mathcal{B}$, si $T(p') = (a - b) + (b - 4c)x$.
3. $\mathcal{B}'$, si $T(p') = (a - b) + (b - c)x + (-a + 3c)x^2$.
4. Si $T : P_1 \to P_1$, $T(p) = (-a + b) + (2a - 3b)x$ y sea $S = \{2 + x, 1\}$.
   a. Determine la matriz A de T con respecto a $\mathcal{B}$.
   b. Determine la matriz de transición P de S a $\mathcal{B}$.
   c. Aproveche (a) y (b) para determinar la matriz de T con respecto a S.
   d. Determine la matriz A' de T con respecto a S directamente a partir de S.
5. Sea $T : P_2 \to P_2$ la transformación lineal definida por $T(p') = -2c + bx$.

a. Determine la matriz A de T con respecto a la base estándar $\{1, x, x^2\}$.
b. Determine la matriz A' de T con respecto a la base de $\mathcal{B}'$.
c. Evalúe $T(6x - 2x^2)$: (a) directamente, (b) usando A, y (c) empleando A'.

Sea $T : P_1 \to P_2$ una transformación lineal. En los ejercicios 6 y 7, para la T dada:

(a) Determine la matriz A de T con respecto a $\mathcal{B}$ y a $\mathcal{B}'$.
(b) Evalúe $T(5 - 2x)$ directamente y empleando A, $\mathcal{B}$ y $\mathcal{B}'$.

6. $T(p) = b + ax - ax^2$
7. $T(p) = a - ax - bx^2$

Sea $\mathcal{R}$ la siguiente base de M_{22}:

$$\mathcal{R} = \{E_{11} - E_{12}, E_{12} - E_{21}, E_{21} - E_{22}, E_{22} + E_{11}\}$$

8. Obtenga la matriz de $T : M_{22} \to M_{22}$ con respecto a $\mathcal{R}$, donde

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -b & a \\ -d & c \end{bmatrix}$$

9. Determine $T\begin{bmatrix} 1 & -2 \\ -4 & 3 \end{bmatrix}$ si la matriz de la transformación lineal $T : M_{22} \to M_{22}$ con respecto a $\mathcal{R}$ es

$$A = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

10. Demuestre la parte 2 del teorema 15.
11. Demuestre la unicidad de la matriz de transición que afirma el teorema 13.

En los ejercicios 12 a 14, A es la matriz de una transformación lineal $T : P_n \to P_m$. Determine n y m, y una fórmula para $T(q)$, $q \in P_n$ con respecto a cada una de las siguientes:

12. $\mathcal{B}'$, si $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$
13. $\mathcal{B}$ y $\mathcal{B}''$ si $A = \begin{bmatrix} 2 & 0 & 1 & 0 \\ -4 & 1 & 2 & -8 \end{bmatrix}$
14. $\mathcal{B}''$ y $\mathcal{B}$ si $A = \begin{bmatrix} -4 & 2 \\ 0 & 9 \\ 1 & -1 \\ 2 & -3 \end{bmatrix}$
15. Determine la matriz A de $T : M_{22} \to M_{22}$,

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -b & d \\ c & -a \end{bmatrix}$$

con respecto a la base $\mathcal{R}$, demuestre que T es un isomorfismo mediante operaciones de renglón reducido de A. ¿Qué teorema aplicaría?

16. Compruebe que A es semejante a sí misma.
17. Demuestre que si B es semejante a A, entonces A es similar a B.
18. Demuestre que si B es semejante a A, y C es similar a B, entonces C es similar a A.
19. Para dos matrices, A y B $n \times n$ con al menos una de ellas invertible, demuestre que AB es semejante a BA.

5.5 El álgebra de las transformaciones lineales

Objetivos del estudiante para esta sección

1. Conocer las operaciones con las transformaciones lineales y la forma en que se relacionan con las operaciones matriciales.
2. Definir una transformación lineal invertible (isomorfismo) y sus propiedades.

Definiremos las operaciones básicas con las transformaciones lineales, que son suma, multiplicación por escalar e inversión, y las relacionaremos con las operaciones básicas con matrices. También repasaremos y ahondaremos en la investigación del concepto de isomorfismo.

NOTA En toda esta sección, V, W y U son espacios vectoriales de dimensión finita.

Sumas y productos por escalares

Sean $f, g : V \to W$ transformaciones lineales. La **suma** $f + g$ de f y g es la transformación $f + g : V \to W$ definida por

$$(f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v})$$

para todo $\mathbf{v} \in V$. Sea c cualquier escalar. El **múltiplo escalar** cf de f por c es la transformación $cf : V \to W$ definida por

$$(cf)(\mathbf{v}) = cf(\mathbf{v})$$

para todo $\mathbf{v} \in V$.

■ EJEMPLO 50 Evalúe $f + g$ y $5f$ en $a + bx + cx^2$ si

$$f, g : P_2 \to P_1$$

$$f(a + bx + cx^2) = b + cx \qquad \text{y} \qquad g(a + bx + cx^2) = c - ax$$

SOLUCIÓN La suma $f + g$ se expresa como sigue:

$$\begin{aligned}(f + g)(a + bx + cx^2) &= f(a + bx + cx^2) + g(a + bx + cx^2)\\ &= (b + cx) + (c - ax)\\ &= (b + c) + (c - a)x\end{aligned}$$

El múltiplo escalar $5f$ es la transformación con los valores

$$(5f)(a + bx + cx^2) = 5f(a + bx + cx^2) = 5(b + cx) = 5b + 5cx$$ □

TEOREMA 18

> $f + g$ y cf son transformaciones lineales.

DEMOSTRACIÓN Sean $\mathbf{v}_1, \mathbf{v}_2 \in V$ y sean $c_1, c_2 \in \mathbf{R}$. Entonces

$$\begin{aligned}(f + g)(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) &= f(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) + g(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)\\ &= c_1 f(\mathbf{v}_1) + c_2 f(\mathbf{v}_2) + c_1 g(\mathbf{v}_1) + c_2 g(\mathbf{v}_2)\\ &= c_1 (f(\mathbf{v}_1) + g(\mathbf{v}_1)) + c_2 (f(\mathbf{v}_2) + g(\mathbf{v}_2))\\ &= c_1 (f + g)(\mathbf{v}_1) + c_2 (f + g)(\mathbf{v}_2)\end{aligned}$$

En consecuencia, $f + g$ es lineal. La demostración que cf es lineal se deja como ejercicio. □

Observe que pueden formarse combinaciones lineales

$$c_1 f_1 + \cdots + c_n f_n$$

para los escalares $c_1, \ldots, c_n$ y las transformaciones lineales $f_1, \ldots, f_n$.

La suma y la multiplicación por escalar de transformaciones satisfacen propiedades idénticas a las de la suma y la multiplicación por escalar de matrices, enunciadas en el teorema 1 de la sección 3.1. Recuerde que I es la transformación identidad, $I : V \to V$, $I(\mathbf{v}) = \mathbf{v}$ y $\mathbf{0}$ es la transformación $\mathbf{0} : V \to W$, $\mathbf{0}(\mathbf{v}) = \mathbf{0}$.

TEOREMA 19

(Leyes de la suma y multiplicación por escalar)

Sean f, g y h transformaciones lineales tales que pueden llevarse a cabo las operaciones descritas a continuación, cuando c es cualquier escalar. Entonces,

1. $(f+g)+h=f+(g+h)$
2. $f+g=g+f$
3. $f+\mathbf{0}=\mathbf{0}+f=f$
4. $f+(-f)=(-f)+f=\mathbf{0}$
5. $c(f+g)=cf+cg$
6. $(a+b)f=af+bf$
7. $(ab)f=a(bf)=b(af)$
8. $1f=f$
9. $0f=\mathbf{0}$

DEMOSTRACIÓN Se deja como ejercicio. □

Composición de transformaciones lineales

La composición de dos transformaciones $g: A \to B$ y $f: B \to C$ es la nueva transformación $f \circ g: A \to C$ que se obtiene evaluando la segunda transformación f en los valores de la primera, g. Por tanto, $f \circ g(a) = f(g(a))$ para toda $a \in A$. Esta operación es muy importante en todas las matemáticas y sus aplicaciones. Estudiaremos ahora el caso en el que f y g son transformaciones lineales.

DEFINICIÓN

(Composición)

Sean $g: U \to V$ y $f: V \to W$ transformaciones lineales. La **composición de f con g** es la transformación $f \circ g: U \to W$ definida por

$$f \circ g(\mathbf{v}) = f(g(\mathbf{v}))$$

para todo $\mathbf{v} \in U$ (figura 5. 27).

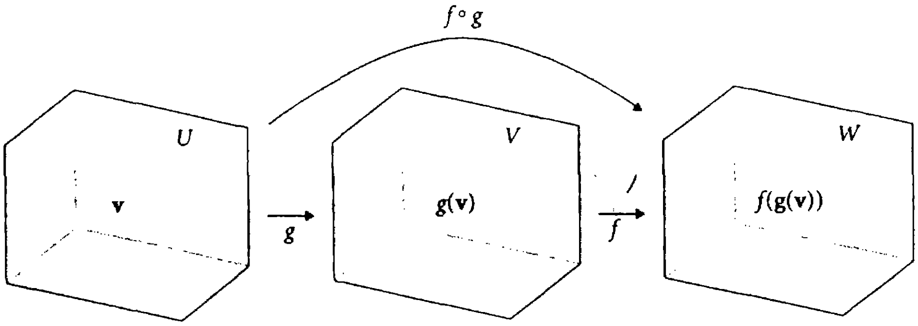


Figura 5.27 Composición de transformaciones.

■ EJEMPLO 51 Sean $g : \mathbf{R}^2 \to \mathbf{R}^3$ y $f: \mathbf{R}^3 \to \mathbf{R}^4$ las transformaciones lineales definidas por

$$g(x,y) = (x + y, x - y, 2x) \quad \text{y} \quad f(x,y,z) = (x - y, x + y, x + z, 2z)$$

Determine (a) $f \circ g(1, 2)$ y (b) $f \circ g(x, y)$.

SOLUCIÓN

(a) La composición es una aplicación de $\mathbf{R}^2$ a $\mathbf{R}^4$, $f \circ g : \mathbf{R}^2 \to \mathbf{R}^4$. Como $g(1, 2) = (3, -1, 2)$, entonces

$$f \circ g(1,2) = f(g(1,2)) = f(3,-1,2) = (4,2,5,4)$$

(b) De igual forma,

$$f \circ g(x,y) = f(g(x,y)) = f(x + y, x - y, 2x) = (2y, 2x, 3x + y, 4x)$$

□

TEOREMA 20

$f \circ g : U \to W$ es una transformación lineal.

DEMOSTRACIÓN Sean $\mathbf{v}_1, \mathbf{v}_2 \in V$ y sean $c_1, c_2 \in \mathbf{R}$. Entonces

$$\begin{aligned} f \circ g(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) &= f(g(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)) \\ &= f(c_1 g(\mathbf{v}_1) + c_2 g(\mathbf{v}_2)) \\ &= c_1 f(g(\mathbf{v}_1)) + c_2 f(g(\mathbf{v}_2)) \\ &= c_1 f \circ g(\mathbf{v}_1) + c_2 f \circ g(\mathbf{v}_2) \end{aligned}$$

Por consiguiente, $f \circ g$ es lineal. □

Las composiciones de transformaciones lineales satisfacen las propiedades siguientes. (Observe la semejanza con el teorema 2 de la sección 3.1, si se reemplaza la multiplicación matricial por la composición.)

TEOREMA 21

(Leyes de la composición)

Sean f, g y h transformaciones lineales en las que pueden llevarse a cabo las siguientes operaciones y sea c cualquier escalar. Entonces se verifica lo siguiente:

1. $(f \circ g) \circ h = f \circ (g \circ h)$
2. $f \circ (g + h) = f \circ g + f \circ h$
3. $(g + h) \circ f = g \circ f + h \circ f$
4. $c(g \circ h) = (cg) \circ h = g \circ (ch)$
5. $I \circ f = f \circ I = f$
6. $\mathbf{0} \circ f = \mathbf{0}, \quad f \circ \mathbf{0} = \mathbf{0}$

Emplearemos los mismos nombres que en el teorema 2 de la sección 3.1, es decir, ley asociativa, ley distributiva izquierda, etcétera.

DEMOSTRACIÓN Comprobaremos las partes 1 y 2, dejando las demás como ejercicios.

1. De acuerdo con la definición de composición,

$$\begin{aligned}(f \circ g) \circ h(\mathbf{v}) &= f \circ g(h(\mathbf{v}))\\ &= f(g(h(\mathbf{v})))\\ &= f(g \circ h(\mathbf{v}))\\ &= f \circ (g \circ h)(\mathbf{v})\end{aligned}$$

Para todo **v**.[9]

2. También para todo **v**,

$$\begin{aligned}f \circ (g + h)(\mathbf{v}) &= f((g + h)(\mathbf{v}))\\ &= f((g(\mathbf{v}) + h(\mathbf{v}))\\ &= f(g(\mathbf{v})) + f(h(\mathbf{v})) && \text{ya que } f \text{ es lineal.}\\ &= f \circ g(\mathbf{v}) + f \circ h(\mathbf{v})\\ &= (f \circ g + f \circ h)(\mathbf{v})\end{aligned}$$

Como en el caso de la multiplicación matricial, la composición no es conmutativa. Así, en general,

$$f \circ g \neq g \circ f$$

Potencias de una transformación lineal

Sea $f: V \to V$ una transformación lineal. La composición $f \circ f$ suele escribirse en la forma f^2. En forma semejante, se escribe f^3 en vez de $(f^2) \circ f$, etc. También se define f^1 como f, y f^0 como I, la transformación de identidad. A esas composiciones se les llama **potencias** de f.

$$f^0 = I, \quad f^1 = f, \quad f^2 = f \circ f, \quad \ldots, \quad f^k = f \circ f \circ \cdots \circ f \qquad (k \text{ “factores”})$$

Transformación lineal y operaciones matriciales

En la sección 5.4 vimos que hay una relación muy estrecha entre matrices y transformaciones lineales, que consiste en que toda transformación lineal $f: V \to W$ puede representarse con una transformación matricial a través de

$$[f(\mathbf{v})]_{\mathcal{B}'} = A\,[\mathbf{v}]_{\mathcal{B}} \qquad \text{para todo } \mathbf{x} \in V \tag{5.5}$$

en donde $\mathcal{B}$ y $\mathcal{B}'$ son bases fijas de V y de W, respectivamente. A es la matriz de f con respecto a $\mathcal{B}$ y a $\mathcal{B}'$. Recuérdese que A es la única matriz que satisface la ecuación (5.5) y se expresa como sigue:

$$A = \left[[f(\mathbf{v}_1)]_{\mathcal{B}'}, \ldots, [f(\mathbf{v}_n)]_{\mathcal{B}'}\right]$$

El teorema siguiente nos explica cómo se corresponden transformaciones lineales con las operaciones matriciales.

[9] Vale la pena hacer notar que no se ha usado la hipótesis de que las transformaciones son lineales en la demostración de 1, porque la asociatividad es válida en general, siempre y cuando estén definidas las transformaciones compuestas.

TEOREMA 22

Sean f y g transformaciones lineales entre espacios vectoriales de dimensiones finitas, con matrices A y B con respecto a bases fijas. Entonces, la matriz de la transformación lineal

1. $f+g$ es $A+B$;
2. $f-g$ es $A-B$;
3. $-f$ es $-A$;
4. cf es cA,
5. $f\circ g$ es AB.

DEMOSTRACIÓN Comprobaremos la parte 1 y dejaremos las demás como ejercicios. Para todo $\mathbf{v}\in V$,

$$\begin{aligned}[(f+g)(\mathbf{v})]_{\mathcal{B}'} &= [f(\mathbf{v})+g(\mathbf{v})]_{\mathcal{B}'}\\ &= [f(\mathbf{v})]_{\mathcal{B}'}+[g(\mathbf{v})]_{\mathcal{B}'}\\ &= A[\mathbf{v}]_{\mathcal{B}}+B[\mathbf{v}]_{\mathcal{B}}\\ &= (A+B)[\mathbf{v}]_{\mathcal{B}}\end{aligned}$$

de modo que $A+B$ es la matriz de $f+g$ con respecto a $\mathcal{B}$ y a $\mathcal{B}'$. $\square$

■ EJEMPLO 52 Compruebe la parte 5 del teorema 22 para las transformaciones del ejemplo 51, empleando las bases estándar.

SOLUCIÓN Se tienen

$$g(1,0)=(1,1,2),\qquad g(0,1)=(1,-1,0)$$

$$f(1,0,0)=(1,1,1,0),\qquad f(0,1,0)=(-1,1,0,0),\qquad f(0,0,1)=(0,0,1,2)$$

En consecuencia, las matrices estándar de f y g son, respectivamente,

$$\begin{bmatrix}1&-1&0\\1&1&0\\1&0&1\\0&0&2\end{bmatrix}\quad \text{y}\quad \begin{bmatrix}1&1\\1&-1\\2&0\end{bmatrix}$$

Por otro lado, la composición $f\circ g$ se expresa por $(f\circ g)(x,y)=(2y,2x,3x+y,4x)$. Por tanto,

$$(f\circ g)(1,0)=(0,2,3,4)\quad \text{y}\quad (f\circ g)(0,1)=(2,0,1,0)$$

Así, la matriz estándar de $f\circ g$ es

$$\begin{bmatrix}0&2\\2&0\\3&1\\4&0\end{bmatrix}=\begin{bmatrix}1&-1&0\\1&1&0\\1&0&1\\0&0&2\end{bmatrix}\begin{bmatrix}1&1\\1&-1\\2&0\end{bmatrix}$$ $\square$

Transformaciones lineales invertibles

DEFINICIÓN

Una transformación lineal $f: V \to V$ es **invertible** si hay una transformación $g: V \to V$ con la propiedad

$$f \circ g = I \quad \text{y} \quad g \circ f = I$$

A la transformación g se le llama **inversa** de f. Si existe una inversa, es única (la demostración es idéntica a la de la unicidad de la inversa de una matriz). Esta inversa única se representa por f^{-1}. De ahí que

$$f \circ f^{-1} = I \quad \text{y} \quad f^{-1} \circ f = I$$

Observe que si f es invertible, entonces

$$f(\mathbf{v}) = \mathbf{w} \Leftrightarrow f^{-1}(\mathbf{w}) = \mathbf{v}$$

En realidad, $f(\mathbf{v}) = \mathbf{w}$ significa que $f^{-1}(f(\mathbf{v})) = f^{-1}(\mathbf{w})$, lo cual implica que $f^{-1} \circ f(\mathbf{v}) = f^{-1}(\mathbf{w})$ o $\mathbf{v} = f^{-1}(\mathbf{w})$. Como estos pasos son reversibles, la equivalencia es el resultado es la equivalencia.

La inversa de una transformación, si existe, invierte el efecto de la transformación (figura 5.28).

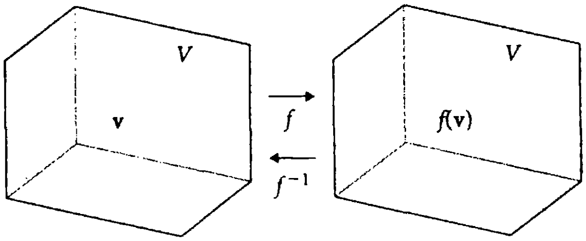


Figura 5.28 Una transformación lineal y su inversa.

El teorema siguiente identifica las transformaciones lineales invertibles y los isomorfismos; es decir, las transformaciones lineales que son biunívocas y sobre, estudiadas en la sección 5.3.

TEOREMA 23

Sea $f: V \to V$ una transformación lineal.

1. f es invertible si y sólo si es un isomorfismo.
2. Si f es invertible, entonces f^{-1} es lineal.

DEMOSTRACIÓN

1. Sea f invertible y sea g su inversa, comprobaremos que f es biunívoca y sobre. Para $\mathbf{v}_1$, $\mathbf{v}_2 \in V$, supongamos que $f(\mathbf{v}_1) = f(\mathbf{v}_2)$. Entonces

$$\begin{aligned} f(\mathbf{v}_1) &= f(\mathbf{v}_2) \\ &\Rightarrow g(f(\mathbf{v}_1)) = g(f(\mathbf{v}_2)) \\ &\Rightarrow g \circ f(\mathbf{v}_1) = g \circ f(\mathbf{v}_2) \\ &\Rightarrow \mathbf{v}_1 = \mathbf{v}_2 \end{aligned}$$

Por consiguiente, f es biunívoca. Sea $\mathbf{w} \in V$ y sea $\mathbf{v} = g(\mathbf{w})$. En tal caso

$$f(\mathbf{v}) = f(g(\mathbf{w})) = f \circ g(\mathbf{w}) = \mathbf{w}$$

Por lo anterior, para cada elemento $\mathbf{w}$ existe un elemento $\mathbf{v}$ que se transforma a $\mathbf{w}$ bajo f. Así, f es sobre.

Por el contrario, supongamos que f es biunívoca y sobre. Definiremos la inversa g de f. Sea $\mathbf{v} \in V$. Existe un $\mathbf{w}$ único tal que $f(\mathbf{w}) = \mathbf{v}$. Definimos $g(\mathbf{v}) = \mathbf{w}$. Por consiguiente, g es bien definida y es la inversa de f. (Compruébelo.)

2. Sea f invertible. Demostraremos que f^{-1} es lineal. Como $\mathbf{v}_1, \mathbf{v}_2 \in V$, entonces hay vectores únicos, $\mathbf{w}_1, \mathbf{w}_2 \in V$ tales que $f(\mathbf{w}_1) = \mathbf{v}_1$ y $f(\mathbf{w}_2) = \mathbf{v}_2$, porque f es biunívoca y sobre, de acuerdo con la parte 1. Por consiguiente, $\mathbf{w}_1 = f^{-1}(\mathbf{v}_1)$ y $\mathbf{w}_2 = f^{-1}(\mathbf{v}_2)$. Así,

$$\begin{aligned} f^{-1}(\mathbf{v}_1 + \mathbf{v}_2) &= f^{-1}(f(\mathbf{w}_1) + f(\mathbf{w}_2)) \\ &= f^{-1}(f(\mathbf{w}_1 + \mathbf{w}_2)) \\ &= f^{-1} \circ f(\mathbf{w}_1 + \mathbf{w}_2) \\ &= I(\mathbf{w}_1 + \mathbf{w}_2) \\ &= \mathbf{w}_1 + \mathbf{w}_2 \\ &= f^{-1}(\mathbf{v}_1) + f^{-1}(\mathbf{v}_2) \end{aligned}$$

Hemos demostrado que $f^{-1}(\mathbf{v}_1 + \mathbf{v}_2) = f^{-1}(\mathbf{v}_1) + f^{-1}(\mathbf{v}_2)$. Dejaremos como ejercicio la verificación que para cada $\mathbf{v} \in V$ y cada escalar c,

$$f^{-1}(c\mathbf{v}) = c f^{-1}(\mathbf{v})$$

Por lo anterior, f^{-1} es una transformación lineal. □

TEOREMA 24

> Sea $f: V \to V$ una transformación lineal cuya matriz es A con respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$ de V. Entonces
>
> 1. f es invertible si y sólo si A es invertible.
> 2. Si f es invertible, entonces A^{-1} es la matriz de f^{-1} con respecto a $\mathcal{B}'$ y $\mathcal{B}$.

DEMOSTRACIÓN Se deja como ejercicio. □

■ EJEMPLO 53 Demuestre que la transformación $T: \mathbf{R}^3 \to \mathbf{R}^3$ es invertible y determine su inversa:

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y+z \\ z \end{bmatrix}$$

SOLUCIÓN La matriz estándar A de T es invertible. Con más precisión,

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

En consecuencia, según el teorema 24, T es invertible y la matriz estándar de T^{-1} es A^{-1}. En consecuencia,

$$T^{-1}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y+z \\ y-z \\ z \end{bmatrix}$$ □

Ejercicios 5.5

En los ejercicios 1 a 3 evalúe $f+g$ y $-4f$ en (x, y, z) y en $(-1, 2, 0)$.

1. $f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y+z \\ x+y \end{bmatrix}$

   $g\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x+y \\ 2x-y-z \end{bmatrix}$

2. $f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ x+2y \\ z \end{bmatrix}$

   $g\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y-z \\ 2x-2y \\ x-z \end{bmatrix}$

3. $f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x-2y \\ x-z \end{bmatrix}$

   $g\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y+z \\ x-3y \end{bmatrix}$

4. Evalúe $f+g$ y $-3f$ en $a+bx$ y en $-6+7x$, si

$$f(a+bx) = b - ax, \quad g(a+bx) = (3a+b) - bx$$

En los ejercicios 5 a 7 compruebe los enunciados 1 y 4 del teorema 22 para las f y g dadas. Use $c = -2$.

5. f y g como en el ejercicio 1.
6. f y g como en el ejercicio 2.
7. f y g como en el ejercicio 3.

En los ejercicios 8 a 10 determine $f \circ g(x, y)$ y $f \circ g(-1, -3)$.

8. $f\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ x+y \end{bmatrix}$

   $g\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -5x+y \\ x+3y \end{bmatrix}$

9. $f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y+z \\ x+2z \end{bmatrix}$

   $g\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x-4y \\ x-y \end{bmatrix}$

10. $f\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -x-2y+z \\ x-w \end{bmatrix}$

    $g\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x-3y \\ x-y \\ x \end{bmatrix}$

En los ejercicios 11 a 13 verifique el enunciado 5 del teorema 22 para las f y g dadas.

11. f y g como en el ejercicio 8.
12. f y g como en el ejercicio 9.
13. f y g como en el ejercicio 10.

14. Con los datos del ejercicio 4, determine $f \circ g$ y $g \circ f$.

15. Explique por qué $f \circ g$ es indefinida. ¿Está definida $g \circ f$?

a. $f(x,y) = (x - y, x + y, y)$
$g(x,y) = (-x + y, x, y)$

b. $f(x,y,z) = (x + z, x + y)$
$g(x,y) = (-x + y, x + 2y)$

c. $f(x,y) = (x - y, x + y)$
$g(x,y) = (-x + y, x + y, x)$

16. Determine dos transformaciones lineales, f y g, tales que $f \circ g \neq g \circ f$.

17. Sean $f\colon \mathbf{R}^2 \to \mathbf{R}^2$, $f(x, y) = (x - y, x + y)$. Calcule $f^3(x, y)$ y $f^3(1, -1)$.

18. Demuestre que f y g son inversas entre sí.

$$f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x + y \\ x - z \\ x + y - z \end{bmatrix}$$

$$g\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x - y + z \\ -y + z \\ -x - 2y + z \end{bmatrix}$$

19. Compruebe que f es invertible. No calcule la inversa.

$$f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x + y + z \\ x - 2z \\ 2x - y \end{bmatrix}$$

20. Demuestre que f es invertible. No calcule la inversa.

$$f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + 3y + 2z \\ 3x + 2y + z \\ 3x + 3y + z \end{bmatrix}$$

En los ejercicios 21 a 25 demuestre que la transformación es invertible y calcule su inversa.

21. $f(x,y,z) = (-2x - z, -y - 2z, -2z)$

22. $f(x,y,z) = (x + 2y - z, x - 2y - z, x + 6y + z)$

23. $f(x,y,z) = (x + y - z, x + 2y - z, 3x + 4y + 3z)$

24. $$f\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x + z \\ x + \frac{2}{3}y - \frac{4}{3}z \\ -y + z \end{bmatrix}$$

25. $f\colon \mathbf{R}^4 \to \mathbf{R}^4$, expresada por

$$f(\mathbf{x}) = \begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}\mathbf{x}$$

26. Termine la demostración del teorema 18.

27. Demuestre el teorema 19.

28. Compruebe los enunciados 3 a 6 del teorema 21.

29. Demuestre los enunciados 2 a 5 del teorema 22.

30. Compruebe el teorema 24.

Inversas derecha e izquierda

En este párrafo presentaremos las inversas derecha e izquierda de transformaciones lineales, como lo hicimos para las matrices en los ejercicios de la sección 3.3.

Sea $f\colon V \to W$ una transformación lineal entre espacios vectoriales de dimensiones finitas. Se dice que la transformación lineal g es **inversa derecha** de f si $f \circ g = I$. En forma parecida, h es una **inversa izquierda** de f si $h \circ f = I$. Por ejemplo, si $p(x, y, z) = (x, y)$ y $q(x, y) = q(x, y, 0)$, entonces p es una inversa izquierda de q, y q a su vez es una inversa derecha de p, porque

$$p \circ q(x,y) = (x,y) \tag{5.6}$$

Sea A la matriz de f con respecto a las bases fijas de V y W. Recordemos las definiciones de inversas izquierda y derecha de matrices, en los ejercicios de la sección 3.3.

31. Demuestre que las afirmaciones siguientes son equivalentes:

a. f tiene una inversa derecha.
b. f es sobre.
c. A tiene una inversa derecha.

32. Compruebe que las afirmaciones siguientes son equivalentes:

a. f tiene una inversa izquierda.
b. f es biunívoca.
c. A tiene una inversa izquierda.

33. Confirme que si f tiene a la vez una inversa derecha g y una inversa izquierda h, entonces es válido lo siguiente.

a. $g = h$.
b. f es biunívoca y sobre.
c. f es un isomorfismo.
d. A tiene una inversa derecha y una izquierda que coinciden.
e. A es invertible.

Potencias negativas de transformaciones invertibles

Así como hicimos con las matrices invertibles, podemos determinar las potencias negativas de transformaciones invertibles. Si f es invertible y n es un entero positivo, se define

$$f^{-n} = (f^{-1})^n = \underbrace{f^{-1} \circ f^{-1} \circ \cdots \circ f^{-1}}_{n \text{ factores}}$$

34. Calcule $f^{-3}(x, y)$ y $f^{-3}(-1, 2)$ si

$$f\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ x + y \end{bmatrix}$$

35. Compare $f^{-2} \circ f^{-1}(x, y)$ y $f^{-3}(x, y)$ si

$$f\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - 2y \\ 2x + y \end{bmatrix}$$

5.6 Aplicaciones

Objetivo del estudiante para esta sección
Adquirir una idea de las aplicaciones relacionadas con las transformaciones lineales.

En esta sección describiremos algunas aplicaciones importantes del material estudiado en este capítulo, subrayando las relaciones de las transformaciones lineales con las gráficas de computadora y con los fractales.

Transformaciones afines y gráficas de computadora

DEFINICIÓN

Sea A una matriz $m \times n$. Una **transformación afín** $T : \mathbf{R}^n \to \mathbf{R}^m$ tiene la forma

$$T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

para algún vector m fijo $\mathbf{b}$. Esta transformación es *no lineal* si $\mathbf{b} \neq \mathbf{0}$. Por lo anterior, $T(\mathbf{0}) \neq \mathbf{0}$. En el caso especial en que $m = n$ y A sea la matriz I de identidad, $n \times n$, entonces

$$T(\mathbf{x}) = I\mathbf{x} + \mathbf{b} = \mathbf{x} + \mathbf{b}$$

A esa T se le llama **traslación por b**.

Una traslación por un vector $\mathbf{b} \neq \mathbf{0}$ desplaza a una figura sumando $\mathbf{b}$ a todos sus puntos. Una transformación afín es una transformación lineal seguida de una traslación. La figura

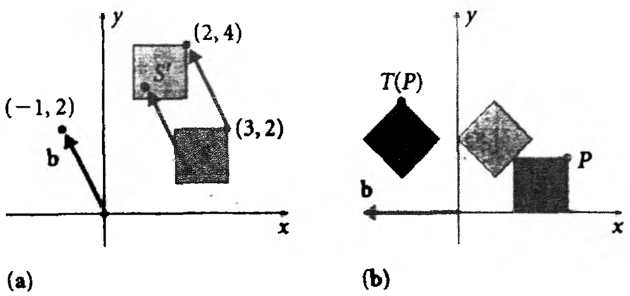


Figura 5.29 (a) Traslación, (b) transformación afín: rotación y después traslación.

5.29(a) muestra la imagen S' del cuadrado S después de la traslación por (2, −1). La figura 5.29(b) muestra la imagen S' del cuadrado S bajo la transformación afín

$$T(\mathbf{x}) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \mathbf{x} + \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

T consiste en una rotación de 45° seguida de una traslación por (−2, 0).

Las transformaciones afines con $n = m = 2$ y $n = m = 3$ son muy útiles para las gráficas en computadora.

■ EJEMPLO 54 Obtenga la transformación afín T que convirtió la imagen izquierda de la figura 5.30 en la de la derecha, puesto que se usaron los puntos siguientes:

$$(1,0), (0.7, 0.7), (0,1), (-0.7, 0.7), (-1,0), (-0.7, -0.7), (0,-1), (0.7, -0.7), (1,0)$$

cuyas respectivas imágenes fueron

$$(2,-1), (2.05, -0.3), (1.5, 0), (0.65, -0.3), (0,-1), (-0.05, -1.7), (0.5, -2),$$
$$(1.35, -1.7), (2, -1)$$

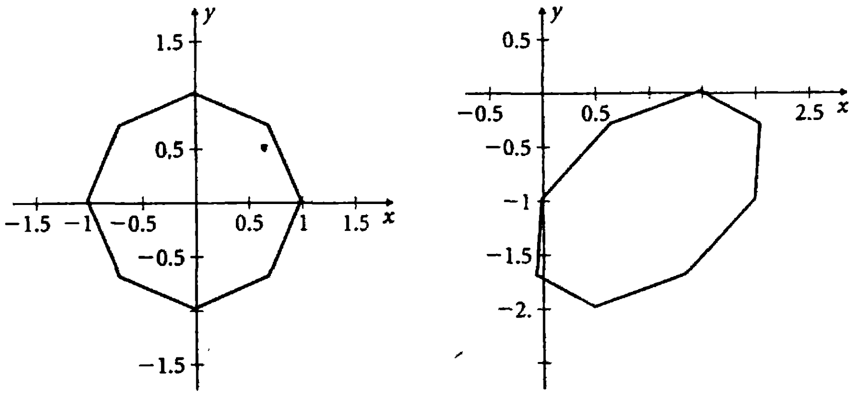


Figura 5.30 Traslación con deslizamiento de un polígono.

SOLUCIÓN Sean $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ con $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ y $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Entonces

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} ax_1 + bx_2 + b_1 \\ cx_1 + dx_2 + b_2 \end{bmatrix}$$

Ya que

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a + b_1 \\ c + b_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b + b_1 \\ d + b_2 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0 \end{bmatrix}$$

$$T\begin{bmatrix}-1\\0\end{bmatrix} = \begin{bmatrix}-a+b_1\\-c+b_2\end{bmatrix} = \begin{bmatrix}0\\-1\end{bmatrix}$$

llegamos al sistema

$$\begin{aligned} a + b_1 &= 2 \\ c + b_2 &= -1 \\ b + b_1 &= 1.5 \\ d + b_2 &= 0 \\ -a + b_1 &= 0 \\ -c + b_2 &= -1 \end{aligned}$$

cuya solución es $a = 1$, $b = 0.5$, $c = 0$, $d = 1$ y $b_1 = 1$, $b_2 = -1$. De modo que,

$$T(\mathbf{x}) = A\mathbf{x} + \mathbf{b} = \begin{bmatrix}1 & 0.5\\0 & 1\end{bmatrix}\mathbf{x} + \begin{bmatrix}1\\-1\end{bmatrix}$$

siempre y cuando esta ecuación sea válida para el resto de los puntos que se usaron en la gráfica. (Sí lo es, compruébelo.)

Por lo anterior, T es el deslizamiento en 0.5 a lo largo del eje x seguido de la traslación por $(1, -1)$. $\square$

Si observamos la figura 5.31, la transformación afín T se describe con la rotación de 45° en dirección positiva en torno al eje z seguida de una traslación por $(1, 1, 1)$, esto se aplicó al tetraedro de la izquierda y produjo el tetraedro de la derecha.

$$T(\mathbf{x}) = \begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{x} + \begin{bmatrix}1\\1\\1\end{bmatrix}$$

En este caso se calcularon las imágenes de una lista bastante larga de puntos, para producir la imagen transformada.

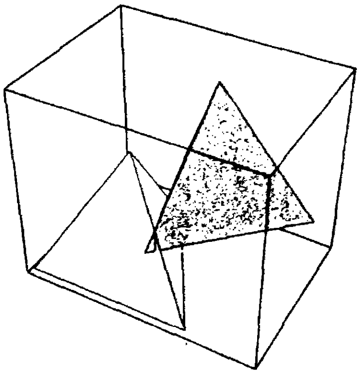


Figura 5.31 Tetraedro girado y trasladado.

Transformaciones afines y fractales

En los últimos años ha surgido un área nueva en las matemáticas, llamada *geometría fractal*. Aunque esta geometría tiene sus raíces en varias contribuciones importantes como las de Cantor, Sierpinski, von Koch, Peano y otros matemáticos del siglo XIX, no fue sino hasta finales la década de los sesenta que llegó a ser un "campo nuevo". Esto se debió al trabajo precursor de Benoit Mandelbrot de la IBM Corporation, y a la disponibilidad de computadoras rápidas. La palabra *fractal*, introducida por Mandelbrot, se usa para describir figuras con "infinitas repeticiones de la misma forma" (figuras 5.32 y 5.33).[10] A continuación describiremos dos fractales, el triángulo de Sierpinski y un fractal que se ve como un abeto. (Es un análogo del helecho de M. Barnsley.)

M. Barnsley observó que pueden obtenerse muchos objetos "fractaloides" graficando *iteraciones* de ciertas transformaciones afines.

El triángulo de Sierpinski

Sean f_1, f_2 y f_3 las tres transformaciones afines de $\mathbf{R}^2$ a $\mathbf{R}^2$ expresadas por

$$f_1(\mathbf{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \mathbf{x}$$

$$f_2(\mathbf{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$$

$$f_3(\mathbf{x}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}$$

El triángulo de Sierpinski puede generarse como sigue: comenzamos con un triángulo, por ejemplo, aquel cuyos vértices están en (0, 0), (1, 0), (0, 1) y elegimos un punto dentro de él

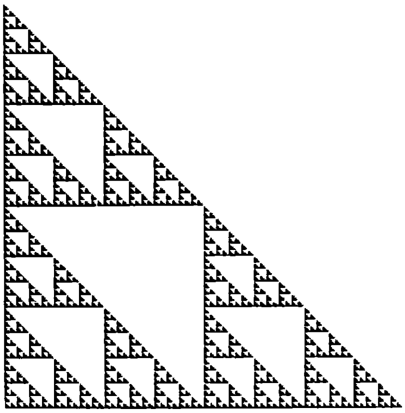


Figura 5.32 Un triángulo de Sierpinski.

[10] Para conocer más información acerca de los fractales, recomendamos *Fractals Everywhere*, de M. Barnsley, 2a. ed. (San Diego, CA: Academic Press, 1993); y *An Introduction to Chaotic Dynamical Systems*, de R. L. Devaney, 2a. ed. (Reading, MA: Addison-Wesley, 1989).

y lo graficamos, digamos, en el punto $(\frac{1}{2}, \frac{1}{2})$. A continuación seleccionamos al azar una de las transformaciones f_1, f_2 o f_3, digamos f_i, y se calcula y grafica $f_i(\frac{1}{2}, \frac{1}{2})$. Partiremos de este nuevo punto y repetiremos el proceso tanto como lo deseemos. La imagen que resulta es un "objeto fractal" que parece un triángulo con agujeros triangulares (si se grafican bastantes puntos). Véase la figura 5.32.

Un abeto parecido al helecho de Barnsley

Sean f_1, f_2, f_3 y f_4 las cuatro transformaciones afines de $\mathbf{R}^2$ a $\mathbf{R}^2$ representadas por

$$f_1(\mathbf{x}) = \begin{bmatrix} 0.75 & 0.03 \\ -0.07 & 0.7 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 10 \\ 150 \end{bmatrix}$$

$$f_2(\mathbf{x}) = \begin{bmatrix} -0.15 & 0.51 \\ 0.5 & 0.15 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 10 \\ 40 \end{bmatrix}$$

$$f_3(\mathbf{x}) = \begin{bmatrix} 0.2 & -0.25 \\ 0.21 & 0.4 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 30 \\ 150 \end{bmatrix}$$

$$f_4(\mathbf{x}) = \begin{bmatrix} 0.02 & -0.05 \\ 0.03 & 0.2 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 10 \\ 1 \end{bmatrix}$$

El fractal parecido a un abeto de la figura 5.33 puede generarse como sigue: se elige y grafica cualquier punto, por ejemplo (5, 5). A continuación se selecciona aleatoriamente una de las funciones f_1, f_2, f_3 o f_4, digamos f_i, y se calcula y grafica $f_i(5, 5)$. Haciendo que éste sea un nuevo punto de partida, se repite el proceso. La imagen que resulta parece una parte de un abeto.

Figura 5.33 Abeto de fractales.

A continuación describiremos el procedimiento que generó ambos fractales. *Este procedimiento produce una imagen fractal para* algunos *conjuntos de transformaciones afines.*[11]

[11] No todas las iteraciones de cualquier conjunto de transformaciones afines producen imágenes fractaloides.

Algoritmo

(Generador de imagen fractal)

1. Comenzar con un conjunto *adecuado* de transformaciones afines $S = \{f_1, f_2, \ldots, f_n\}$ y un punto inicial (x_k, y_k).
2. Elegir al azar una transformación afín de S, por ejemplo f_i.
3. Calcular y graficar el punto $f_i(x_k, y_k)$. Igualar $(x_k, y_k) = f_i(x_k, y_k)$.
4. Ir al paso 2 y repetir tanto como se quiera.

Ejercicios 5.6

En los ejercicios 1 a 4 determine las imágenes del vector cero y las de los vectores ~~de base~~ estándar para las transformaciones afines dadas.

1. $T(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

2. $T(\mathbf{x}) = \begin{bmatrix} 4 & -3 \\ 2 & 5 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

3. $T(\mathbf{x}) = \begin{bmatrix} -1 & 2 & 0 \\ 1 & 2 & -1 \\ 0 & -4 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

4. $T(\mathbf{x}) = \begin{bmatrix} 1 & 2 & -1 & -3 \\ 0 & -4 & 1 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

En los ejercicios 5 a 8 escriba las transformaciones afines dadas en la forma $T(\mathbf{x}) = A(\mathbf{x}) + \mathbf{b}$.

5. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ -x + y - 1 \end{bmatrix}$

6. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2y - 1 \\ 2x + 7y - 1 \end{bmatrix}$

7. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x + 3y + 1 \\ x - z \\ x - 5y + z - 1 \end{bmatrix}$

8. $T\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x - z - 9w + 1 \\ -3z - 6w \end{bmatrix}$

En los ejercicios 9 a 13 determine A y $\mathbf{b}$ para la transformación afín $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ si

9. $T : \mathbf{R}^2 \to \mathbf{R}^2, T(\mathbf{e}_1) = \begin{bmatrix} -1 \\ 3 \end{bmatrix},$

$T(\mathbf{e}_2) = \begin{bmatrix} 4 \\ -7 \end{bmatrix}, T(\mathbf{0}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

10. $T : \mathbf{R}^2 \to \mathbf{R}^2, T(\mathbf{e}_1) = \begin{bmatrix} 4 \\ -2 \end{bmatrix},$

$T(\mathbf{e}_2) = \begin{bmatrix} -5 \\ 9 \end{bmatrix}, T(\mathbf{0}) = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$

11. $T : \mathbf{R}^2 \to \mathbf{R}^2, T(\mathbf{e}_1) = \begin{bmatrix} -2 \\ 5 \end{bmatrix},$

$T(\mathbf{e}_2) = \begin{bmatrix} -5 \\ 6 \end{bmatrix}, T(\mathbf{e}_1 + \mathbf{e}_2) = \begin{bmatrix} -4 \\ 9 \end{bmatrix}$

12. $T : \mathbf{R}^3 \to \mathbf{R}^3$ y

$$T(\mathbf{e}_1) = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

$$T(\mathbf{e}_3) = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}, \quad T(\mathbf{0}) = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

13. $T : \mathbf{R}^3 \to \mathbf{R}^2$ y

$$T(\mathbf{e}_1) = \begin{bmatrix} -2 \\ -5 \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

$$T(\mathbf{e}_3) = \begin{bmatrix} 4 \\ -7 \end{bmatrix}, \quad T(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = \begin{bmatrix} 3 \\ -6 \end{bmatrix}$$

14. Demuestre que si T es una transformación afín con $T(\mathbf{0}) = \mathbf{b}$, entonces $L(\mathbf{x}) = T(\mathbf{x}) - \mathbf{b}$ es una transformación lineal.

15. Sea $T : \mathbf{R}^n \to \mathbf{R}^m$, $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ cualquier transformación afín. Compruebe que T queda determinado en forma única por los valores

$$T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n) \quad \text{y} \quad T(\mathbf{0})$$

(*Sugerencia:* Observe que $T(\mathbf{0}) = \mathbf{b}$ y que $L(\mathbf{x}) = T(\mathbf{x}) - \mathbf{b}$ es una transformación matricial lineal.)

16. Demuestre que toda transformación afín $T: \mathbf{R}^2 \to \mathbf{R}^2$, $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ convierte líneas rectas en líneas rectas o en puntos.

17. Sea $T: \mathbf{R}^n \to \mathbf{R}^m$, $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ una transformación afín. ¿Cuál es la relación entre el conjunto $\{\mathbf{x} \in \mathbf{R}^n, T(\mathbf{x}) = \mathbf{0}\}$ y el conjunto de soluciones del sistema $[A : -\mathbf{b}]$?

18. Trace la imagen de la recta $x - y = -1$ bajo

$$T(\mathbf{x}) = \begin{bmatrix} 1 & -3 \\ 0 & 5 \end{bmatrix}\mathbf{x} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

5.7 Miniproyectos

1 ■ Algunas transformaciones afines especiales

En este proyecto el lector se explicará las bases de algunas transformaciones afines especiales: las similitudes. Éstas se usan mucho en gráficas por computadora, sistemas dinámicos y en fractales.

Una **transformación de similaridad** o **similitud** $T: \mathbf{R}^2 \to \mathbf{R}^2$ es una transformación afín especial que tiene una de las formas siguientes:

$$T(\mathbf{x}) = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}\mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$T(\mathbf{x}) = \begin{bmatrix} r\cos\theta & r\sin\theta \\ r\sin\theta & -r\cos\theta \end{bmatrix}\mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

para algún escalar $r \neq 0$, algún ángulo θ, $0 \leq \theta < 2\pi$, y algunos escalares b_1 y b_2. Las similitudes son rotaciones escaladas seguidas de traslaciones, o rotaciones reflejadas y escaladas seguidas de traslaciones. Por esa causa conservan los ángulos, como veremos después.

Problema A

Demuestre que las transformaciones siguientes son similitudes:

1. Toda rotación en torno al origen.
2. Las reflexiones en torno a los ejes, la diagonal o el origen.
3. Cualquier homotecia de $\mathbf{R}^2$.

Problema B

1. ¿Son similitudes todos los deslizamientos? Compruébelo con los deslizamientos cuyas matrices estándar son

$$\begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0.5 \\ -0.5 & 0 \end{bmatrix}$$

2. ¿Son similitudes las proyecciones sobre los ejes?
3. ¿Son similitudes las traslaciones?

Problema C

1. Deduzca una fórmula para la similitud T que transforma el triángulo (0, 0), (1, 0), (0, 1) en el triángulo (1, 1), (−1, 1), (1, −1).
2. Sea S_1 la imagen, bajo T, del rectángulo S cuyos vértices están en (0, 0), (2, 0), (2, 1) y (0, 1), y sea S_2 la imagen de S_1. Calcule las áreas (S), (S_1) y (S_2) y compare las relaciones (S_2) : (S_1) y (S_1) : (S).
3. Deduzca la fórmula de la similitud R que hace girar 45° a cualquier punto en torno al origen, después lo escala en un factor de 2 y por último lo traslada por (1, 1).
4. Determine la imagen L_1, bajo R, del triángulo L cuyos vértices están en (0, 0), (1, 0) y (0, 1), y determine la imagen L_2 de L_1. Calcule las relaciones de áreas $(L_2) : (L_1)$ y $(L_1) : (L)$. ¿Qué ha observado?

2 ■ Otro fractal

Por lo general, las imágenes de fractales no pueden graficarse sin el auxilio de una computadora. En este proyecto estudiaremos un fractal que, hasta cierto punto, puede visualizarse al graficarlo a mano.

Se tienen los rectángulos con los vértices indicados:

$$\begin{aligned} A:&\quad (1,-1),(1,1),(-1,1),(-1,-1)\\ B:&\quad (1,0),(1,1),(-1,1),(-1,0)\\ C:&\quad (2,0),(2,1),(-2,1),(-2,0)\end{aligned}$$

También se consideran las siguientes transformaciones afines:

$$R(\mathbf{x}) = \begin{bmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}\mathbf{x}$$

$$T(\mathbf{x}) = \begin{bmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \end{bmatrix}$$

Sean $A_1^R = R(A)$ (la imagen del rectángulo A bajo R), $A_2^R = R(A_1^R)$, $A_3^R = R(A_2^R)$, $A_4^R = R(A_3^R)$. En forma parecida, sean A_1^T, A_2^T, A_3^T y A_4^T las imágenes correspondientes bajo T. También se tienen las imágenes consecutivas de B, que son B_1^R, B_2^R, B_3^R y B_4^R bajo R, y B_1^T, B_2^T, B_3^T y B_4^T bajo T. Las imágenes de C se definen de la misma forma.

El problema A tiene por objeto mostrar al lector los efectos de R y T sobre A, B y C y sobre sus imágenes iteradas.

Problema A

1. Trace A, A_1^R, A_2^R, A_3^R y A_4^R en una gráfica y A, A_1^T, A_2^T, A_3^T y A_4^T en otra.
2. Trace B, B_1^R, B_2^R, B_3^R y B_4^R en una gráfica y B, B_1^T, B_2^T, B_3^T y B_4^T en otra.
3. Trace C, C_1^R, C_2^R, C_3^R y C_4^R en una gráfica y C, C_1^T, C_2^T, C_3^T y C_4^T en otra.

El problema B tiene por objeto mostrar la imagen fractal generada al aplicar R y T al origen e iterar.

Problema B

Sea $P(0, 0)$. Determine las dos imágenes de P, P_1 y P_2, bajo R y T. A continuación determine las imágenes P_3, P_4 de P_1 bajo R y T, y las imágenes P_5, P_6 bajo T. Continúe con este proceso hasta donde desee. Después grafique *todos* los puntos que determinó. Se necesitan de 5 a 6 iteraciones para ver la formación de un objeto fractal.

El problema C le indicará cómo se afecta una imagen fractal si se comienza en un punto distinto.

Problema C

Resuelva el problema B comenzando con el punto $Q(0.5, 0.5)$.

5.8 Ejercicios en computadora

Esta sesión de computadora le ayudará a dominar los comandos de sus programas que se relacionan con los temas de este capítulo. Además, servirá para repasar algunos de los conceptos básicos.

Sean

$$\begin{aligned}
T_1 &: \mathbf{R}^3 \to \mathbf{R}^3, & T_1(x,y,z) &= (2x - y + z, x + y, 2y - 3z)\\
T_2 &: \mathbf{R}^2 \to \mathbf{R}^3, & T_2(x,y) &= (3x - 4y, x + 3y, -y)\\
T_3 &: \mathbf{R}^3 \to \mathbf{R}^2, & T_3(x,y,z) &= (2x - y + z, x - z)\\
T_4 &: \mathbf{R}^3 \to \mathbf{R}^3, & T_4(x,y,z) &= (-y + z, x + 1, 2y)
\end{aligned}$$

1. Defina T_1, T_2, T_3 y T_4.
2. Calcule $T_1(1, 1, 1)$, $T_2(1, 1)$, $T_3(1, 1, 1)$, $T_4(1, 1, 1)$.
3. Demuestre que T_1, T_2 y T_3 son lineales.
4. Compruebe que T_4 es no lineal.
5. Determine las matrices estándar de T_1, T_2, T_3.

Una transformación lineal T es tal que

$$T(1,2,3,4) = (1,0,-1,1)$$

$$T(1,3,5,7) = (0,1,0,-1)$$

$$T(3,3,4,4) = (1,1,1,-1)$$

$$T(4,4,4,5) = (1,1,-1,1)$$

6. Obtenga la matriz estándar de T.
7. Calcule $T(2, 2, -2, -2)$.

Sean

$$T_1(x,y,z,w) = (x + 2y + 3z + 4w, 2x + 3y + 4z + 5w, 3x + 4y + 5z + 6w)$$

$$T_2(x,y,z) = (x + 2y + 3z + 4w, 2x + 2y + 3z + 4w, 3x + 3y + 3z + 4w)$$

$$T_3(x,y,z,w,t) = (x + 2y + 3z + 4w, 2x + 2y + 3z + 4w, 3x + 3y + 3z + 4w,$$
$$4x + 4y + 4z + 4t)$$

8. ¿Cuál(es) de $\mathbf{v}_1 = (-1, 0, 3, -2)$, $\mathbf{v}_2 = (18, -31, 8, 5)$ y $\mathbf{v}_3 = (1, -1, 8, 7)$ está(n) en $\mathrm{Ker}(T_1)$?
9. ¿Cuál(es) de $\mathbf{w}_1 = (2, 7, 12)$, $\mathbf{w}_2 = (42, 59, 76)$, $\mathbf{w}_3 = (42, 59, 77)$ está(n) en $\mathrm{R}(T_1)$?

Para cada transformación T_1, T_2 y T_3:

10. Determine la matriz estándar.
11. Calcule una base para el núcleo. ¿Cuál es la nulidad?
12. Obtenga una base para el contradominio. ¿Cuál es el rango?
13. Verifique el teorema de la dimensión.
14. ¿Cuál(es) de la(s) transformación(es) T_1, T_2 y T_3 es biunívoca, sobre, un isomorfismo o nada de lo anterior?
15. ¿Cierto o falso?

    a. $\mathrm{R}(T_1) = \mathbf{R}^3$
    b. $\mathrm{R}(T_2) = \mathbf{R}^3$
    c. $\mathrm{R}(T_3) = \mathbf{R}^4$

16. Defina T y evalúe $T(x + 1)$. Demuestre que T es lineal.

$$T : P_1 \to P_2, \qquad T(ax + b) = (3a - 4b)x^2 + (a + 3b)x - b$$

17. Calcule $L(2x^3 + 2x^2 - 2x - 2)$, si L es lineal de tal forma que

$$L(x^3 + 2x^2 + 3x + 4) = x^3 - x + 1$$
$$L(x^3 + 3x^2 + 5x + 7) = x^2 - 1$$
$$L(3x^3 + 3x^2 + 4x + 4) = x^3 + x^2 + x - 1$$
$$L(4x^3 + 4x^2 + 4x + 5) = x^3 + x^2 - x + 1$$

18. Encuentre una base para el espacio nulo de F, donde

$$\begin{aligned} F(ax^3 + bx^2 + cx + d) = {} & (a + 2b + 3c + 4d)x^2 \\ & + (2a + 3b + 4c + 5d)x \\ & + (3a + 4b + 5c + 6d) \end{aligned}$$

19. Determine la matriz M de T con respecto a $\mathcal{B} = \{x - 1, x + 1\}$ y $\mathcal{B}' = \{x^2 - 1, x + 1, x - 1\}$.

$$T : P_1 \to P_2, \qquad T(ax + b) = (3a - 4b)x^2 + (a + 3b)x - b$$

Soluciones seleccionadas con Maple

```
with(linalg):    # Se carga el paquete para el conjunto completo.
# Ejercicios 1 a 5.
# En lugar de T1:= (x,y,z) -> vector([2*x-y+z, x+y, 2*y-3*z]); etc.,
# también se puede usar lo siguiente, más sencillo:
T1 := (x, y, z) -> [2*x-y+z, x+y, 2*y-3*z];
T2 := (x, y, z) -> [3*x-4*y, x+3*y, -y];
T3 := (x, y, z) -> [2*x-y+z, x-z];
T4 := (x, y, z) -> [-y+z, x+1, 2*y];
```

```
# También T1 := proc(x,y,z) [2*x-y+z, x+y, 2*y-3*z] end, etc,.
T1 (1, 1, 1);            # También T2(1,1); T3(1,1,1); T4(1,1,1);
# A continuación se necesita el comando equal, cargado con el paquete
# linalg. equal prueba vectores y matrices para ver si hay igualdad.
# Primero comprobar la parte 1 de la definición:
equal(T1(x1+x2, y1+y2, z1+z2), T1(x1, y1, z1)+T1(x2, y2, z2));
equal(T2(x1+x2, y1+y2), T2(x1, y1)+T2(x2, y2));
equal(T3(x1+x2, y1+y2, z1+z2), T3(x1, y1, z1)+T3(x2, y2, z2));
equal(T4(x1+x2, y1+y2, z1+z2), T4(x1, y1, z1)+T4(x2, y2, z2));
# En la segunda parte tener cuidado. Un producto como c*[x-y,x+y]
# no se simplifica automáticamente a [c*x-c*y, c*x+c*y].
# Primero usar evalm(c*[x-y,x+y]); para pasar el escalar al interior:
# [c*(x-y),c*(x+y)]. Ahora expand([c*(x-y), c*(x+y)]); no se aventaja.
# Se necesita pasar al vector o enlistar y desarrollar
# cada componente por separado con map(expand, [c*(x-y),c*(x+y)]);.
equal(T1(c*x, c*y, c*z), map(expand, evalm(c*T1(x, y, z))));
equal(T2(c*x, c*y), map(expand, evalm(c*T2(x, y))));
equal(T3(c*x, c*y, c*z), map(expand, evalm(c*T3(x, y, z))));
equal(T4(c*x, c*y, c*z), map(expand, evalm(c*T4(x, y, z))));
MT1 := augment (T1(1, 0, 0), T1(0, 1, 0), T1(0, 0, 1));     # Evaluación en los
MT2 := augment (T2(1, 0), T2(0, 1));                        # vectores de base para obtener
MT3 := augment (T3(1, 0, 0), T3(0, 1, 0), T3(0, 0, 1));     # la matriz estándar.

# Ejercicios 6 y 7.

M := matrix([[1, 1, 3, 4], [2, 3, 3, 4], [3, 5, 4, 4], [4, 7, 4, 5]]);      # Los vectores del dominio.
Id := matrix([[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]])      # I_4.
N := matrix([[1, 0, 1, 1], [0, 1, 1, 1], [-1, 0, 1, -1], [1, -1, -1, 1]])   # Los valores.
M1 := rref(augment(M, Id));                                                 # Reducción de [M:I_4].
M2 := delcols(M1, 1..4);
STM := evalm(N & M2);                                                       # La matriz estándar.
evalm(STM &* vector([2, 2, -2, -2]));                                       # T(2,2,-2,-2).

# Ejercicios 8 a 15.

T1 := (x, y, z, w)-> [x+2*y+3*z+4*w, 2*x+3*y+4*z+5*w, 3*x+4*y+5*z+6*w];
T2 := (x, y, z, w) -> [x+2*y+3*z+4*w, 2*x+2*y+3*z+4*w, 3*x+3*y+3*z+4*w];
T3 := (x, y, z, w) -> [x+2*y+3*z+4*w, 2*x+2*y+3*z+4*w, 3*x+3*y+3*z+4*w, 4*x+4*y+4*z+4*w];
T1(-1, 0, 3, -2);     # v1 en el núcleo.
T1(18, -31, 8, 5);    # v2 en el núcleo.
T1(1, -1, 8, 7);      # v3 no está en el núcleo.
M1 := augment (T1(1, 0, 0, 0), T1(0, 1, 0, 0), T1(0, 0, 1, 0), T1(0, 0, 0, 1));   # Matriz estándar.
Mw1 := rref(augment (M1, vector([2, 7, 12])));     # En contradominio (T1). Última columna es pivote.
Mw2 := rref(augment (M1, vector([42, 59, 76])));   # En contradominio (T1). Última columna es pivote.
Mw2 := rref(augment (M1, vector([42, 59, 77])));   # No en contradominio (T1). Última columna no es
                                                     pivote.
evalm(M1);    # La matriz estándar de T1 ya se determinó. Las demás son:
M2 := augment (T2(1, 0, 0, 0), T2(0, 1, 0, 0), T2(0, 0, 1, 0), T2(0, 0, 0, 1));
M3 := augment (T3(1, 0, 0, 0), T3(0, 1, 0, 0), T3(0, 0, 1, 0), T3(0, 0, 0, 1));
              # O bien, kernel(M2) etcétera.
              # k1 tiene dos vectores, así que nulidad 2.
```

```
k1 := nullspace(M1);            # Núcleo no cero, T1 no es biunívoca.
                                # Por tanto, no es isomorfismo.
# También, k1:=kernel(T1);
                                # k2 tiene un vector, por tanto, nulidad 2.
k2 := nullspace(M2);            # Núcleo no cero, T2 no es biunívoca.
                                # Por tanto, no es isomorfismo.
# También, k2:=kernel(T2);
K3 := nullspace(M3);            # k3 no tiene vectores, por tanto nulidad 0. T3 es biunívoca.
# También, k3:=kernel(T3);
                                # Las 2 primeras columnas forman una base para el contradominio.
r1 := rref(M1);                 # El rango es 2. 2 + 2 = cantidad de columnas.
                                # El tercer renglón no es pivote, así que no es sobre.
# También, r1:=range(T1);
                                # Las 3 primeras columnas forman una base para el contradominio.
r2 := rref(M2)                  # El rango es 3. 3 + 1 = cantidad de columnas.
                                # Cada renglón tiene un pivote, así que es sobre.
# También, r2:=range(T2);
                                # Todas las columnas forman una base para el contradominio.
r3 := rref(M3);                 # El rango es 4. 4 + 0 = cantidad de columnas.
                                # T3 es biunívoca y sobre, por tanto es un isomorfismo.
También, r3:=range(T3);
# Falso. T1 no es sobre.
# Cierto. T2 es sobre.
# Cierto. T3 es sobre.

# Ejercicio 16.

T := (a, b) -> [3*a-4*b, a+3*b, -b];     # Usaremos [a,b,c] para indicar ax^2+bx+c.
T(1, 1);                                 # T(x+1).
equal(a1+a2, b1+b2), T(a1, b1)+T(a2, b2)); # linalg[equal] prueba si hay igualdad de matrices.
# Para simplificar c*[a-b,a+b] a [c*a-c*b,c*a+c*b] se usa evalm(c*[a-b,a+b]);
# para pasar el escalar al interior a continuación "map(expand())" para desarrollar cada componente.
equal (T(c*a, c*b), map(expand, evalm(c*T(a, b))));

# Ejercicio 17.

# Primero se forma una matriz con los coeficientes de los polinomios dados.
M := matrix([[1, 1, 3, 4], [2, 3, 3, 4], [3, 5, 4, 4], [4, 7, 4, 5]]);
# después una matriz con los coeficientes de sus valores.
N := matrix([[1, 0, 1,1], [0, 1, 1, 1], [-1, 0, 1, -1), [1, -1, -1, 1]]);
M1 := rref(augment(M, vector([2, 2, -2, -2])));
# La última columna de M1 tiene como elementos los coeficientes de 2x^3+2x^2-2x-2
delcols(M1, 1..4);              # en función de los polinomios dados.
evalm(N &* ("));                # Multiplicación por N para evaluar T(2,2,-2,-2).

# Ejercicio 18.

# Igualar F=0. Se necesita resolver a+2b+3c+4d=0,2a+3b+4c+5d=0,3a+4b+5c+6d=0.
M := matrix(3, 4, [1, 2, 3, 4,  2, 3, 4, 5,  3, 4, 5, 6]);    # Se calcula una base para
NM := nullspace(M);             # el espacio nulo de la matriz de coeficientes.
evalm(matrix([NM[1], NM[2]] &*vector ([x^3, x^2, x, 1])); # En forma de polinomio.
```

```
# Ejercicio 19.
T := (a, b) -> matrix (3, 1, [3*a-4*b, a+3*b, -b]);     # La transformación.
b2 := matrix([[1, 0, 0], [0, 1, 1], [-1, 1, -1]]);     # Los coeficientes de B'.
aug := augment(b2, T(1, -1), T(1, 1));     # Los coeficientes de T(x-1),T(x+1)
rref(aug);               # en términos de B' se calculan con rref(aug);.
delcols(", 1..3);        # Las 2 últimas columnas forman la matriz de T.
```

Soluciones seleccionadas con Mathematica

```
<<LinearAlgebra 'MatrixManipulation';     (* Se carga de una vez todo el paquete. *)
(* Ejercicios 1 a 5. *)
T1[x_, y_, z_] := {2 x-y+z, x+y, 2 y-3 z}     (* Definiciones. El subrayado _ *)
T2[x_, y_] := {3 x-4 y, x+3 y, -y}            (* declara una posición para una *)
T3[x_, y_, z_] := {2 x-y+z, x-z}              (* variable. x_ es la declaración de *)
T4[x_, y_, z_] := {-y+z, x+1, 2 y}            (* una variable llamada x.  *)
(* Dos puntos igual := es el operador de asignación demorada. Se usó de tal forma *)
(* que la función se evalúa cuando se llama, y no durante la definición. *)
T1[1, 1, 1]     (* También T2[1,1] T3[1,1,1]  T4[1,1,1]   *)
(* Para probar la igualdad se usa SameQ o su sinónimo ===. Por ejemplo, *)
(* tanto SameQ[a, b] como a===b investigan la igualdad a = b. *)
(* Antes de probar la igualdad hay que desarrollar los dos lados. *)
SameQ[expand [T1[x1+x2, y1+y2, z1+z2]], Expand [T1[x1, y1, z1]+T1[x2, y2, z2]]]
SameQ[Expand [c T1[x, y, z]], Expand [T1[c x, c y, c z]]]
SameQ(Expand [T2[x1+x2, y1+y2]], Expand [T2[x1, y1]+T2[x2, y2]]]
SameQ[Expand [c T2[x, y]], Expand [T2[c x, c y]]]
SameQ[Expand [T3[x1+x2, y1+y2, z1+z2]], Expand [T3[x1, y1, z1]+T3[x2, y2, z2]]]
SameQ[Expand [c T3[x, y, z]], Expand [T3[c x, c y, c z]]]
SameQ[Expand [T4[x1+x2, y1+y2, z1+z2]], Expand [T4[x1, y1, z1]+T4[x2, y2, z2]]]
SameQ[Expand [c T4[x, y, z]], Expand [T4[c x, c y, c z]]]
Transpose[{T1[1, 0, 0], T1[0, 1, 0], T1[0, 0, 1]}]     (* Evaluación en los vectores *)
Transpose[{T2[1, 0], T2[0, 1]}]                         (* de la base para obtener *)
Transpose[{T3[1, 0, 0], T3[0, 1, 0], T3[0, 0, 1]}]     (* la matriz estándar.  *)
(* Ejercicios 6 y 7.  *)
M = {{1, 1, 3, 4}, {2, 3, 3, 4}, {3, 5, 4, 4}, {4, 7, 4, 5}}     (* Los vectores del dominio.  *)
n = {{1, 0, 1, 1}, {0, 1, 1, 1}, {-1, 0, 1, -1}, {1, -1, -1, 1}}  (* Los valores.   *)
M1 = RowReduce [AppendRows[M, IdentityMatrix[4]]]        (* Reducción de [M:I_4]. *)
M2 = Take Columns[M1, {5, 8}]
STM = n . M2                                             (* La matriz estándar.  *)
STM . {{2}, {2}, {-2}, {-2}}                             (* T(2,2,-2,-2) .  *)
(* Ejercicios 8 a 15.  *)
T1[x_, y_, z_, w_] := {x+2y+3z+4w, 2x+3y+4z+5w, 3x+4y+5z+6w}
T2[x_, y_, z_, w_] := {x+2y+3z+4w, 2x+2y+3z+4w, 3x+3y+3z+4w}
T3[x_, y_, z_, w_] := {x+2y+3z+4w, 2x+2y+3z+4w, 3x+3y+3z+4w, 4x+4y+4z+4w}
T1[-1, 0, 3, -2]     (* v1 está en el núcleo.  *)
T1[18, -31, 8, 5]    (* v2 está en el núcleo.  *)
T1[1, -1, 8, 7]      (* v3 no está en el núcleo.  *)
(* Matriz estándar de T1:  *)
```

```
M1 = Transpose [{T1[1, 0, 0, 0], T1[0, 1, 0, 0], T1[0, 0, 1, 0], T1[0, 0, 0, 1]}]
Mw1 = RowReduce [AppendRows [M1, {{2}, {7}, {12}}]]
                                (* En contradominio(T1). Última columna es pivote.  *)
Mw2 = RowReduce [AppendRows [M1, {{42}, {59}, {76}}]]
                                (* En contradominio(T1). Última columna es pivote.  *)
Mw3 = RowReduce [AppendRows [M1, {{42}, {59}, {77}}]]
                  (* No en contradominio(T1). Última columna no es pivote.  *)
M1                (* Ya se determinó la matriz estándar de T1. Las demás son: *)
M2 = Transpose[{T2[1, 0, 0, 0], T2[0, 1, 0, 0], T2[0, 0, 1, 0], T2[0, 0, 0, 1]}]
M3 = Transpose[{T3[1, 0, 0, 0], T3[0, 1, 0, 0], T3[0, 0, 1, 0], T3[0, 0, 0, 1]}]
                        (* k1 tiene dos vectores, así nulidad 2.  *)
k1 = NullSpace[M1]      (* Núcleo no cero, T1 no es biunívoca.  *)
                        (* Por tanto, no es isomorfismo.  *)

                        (* k2 tiene un vector así nulidad 2.  *)
k2 = NullSpace[M2]      (* Núcleo no cero, T2 no es biunívoca.  *)
                        (* Por tanto, no es un isomorfismo.  *)

k3 = NullSpace[M3]      (* k3 no tiene vectores, entonces nulidad 0. T3 es biunívoca.  *)

                        (* Las 2 primeras columnas forman una base para el contradominio.  *)
r1 = RowReduce[M1]      (* El rango es 2. 2 + 2 = cantidad de columnas.  *)
                        (* Cada renglón tiene un pivote, así que es sobre.  *)

                        (* El tercer renglón no tiene pivote, así que no es sobre  *)
r2 = RowReduce[M2]      (* Las 3 primeras columnas forman una base para el contradominio  *)
(* El rango es 3. 3 + 1 = cantidad de columnas.  *)

                        (* Todas las columnas forman una base para el contradominio.  *)
r3 = RowReduce[M3]      (* El rango es 4. 4 + 0 = cantidad de columnas.  *)
                        (* T3 es biunívoca y sobre, por tanto, un isomorfismo.  *)
(* Falso. T1 no es sobre.  *)
(* Cierto. T2 es sobre.  *)
(* Cierto. T3 es sobre.  *)

(* Ejercicio 16.  *)

t[a_, b_] := {3 a-4 b, a+3 b, -b}   (* Usamos {a,b,c} para indicar ax^2+bx+c.  *)
T[1, 1]                             (* T(1,1) .  *)
(* Antes de comprobar la igualdad hay que desarrollar los dos lados.  *)
SameQ[Expand [T[a1+a2, b1+b2]], Expand[T[a1, b1]+T[a2, b2]]]
SameQ[Expand [c T[a, b]], Expand [T[c a, c b]]]

(* Ejercicio 17.  *)

(* Primero se forma una matriz con los coeficientes de los polinomios dados.  *)
M = {{1, 1, 3, 4}, {2, 3, 3, 4}, {3, 5, 4, 4}, {4, 7, 4, 5}}
n = {{1, 0, 1, 1}, {0, 1, 1, 1}, {-1, 0, 1, -1}, {1, -1, -1}}
(* a continuación una matriz con los coeficientes de sus valores.  *)
M1 = RowReduce[AppendRows[M, {{2}, {2}, {-2}, {-2}}]]
```

```
(* La última columna de M1 tiene como elementos los coeficientes de 2x^3+2x^2-2x-2  *)
TakeColumns[M1, {5}]  (* en términos de los polinomios dados.   *)
n . %                 (* Multiplicación por N para evaluar T(2,2,-2,-2). *)

(* Ejercicio 18. *)

(* Igualar F=0. Se necesita resolver a+2b+3c+4d=0,2a+3b+4c+5d=0,3a+4b+5c+6d=0.  *)
M = {{1, 2, 3, 4}, {2, 3, 4, 5}, {3, 4, 5, 6}}   (* Se calcula una base para el   *)
MN = NullSpace[M]              (* espacio nulo de la matriz de los coeficientes.  *)
MN . {{x^3}, {x^2}, {x}, {1}}  (* La respuesta en forma de polinomio.  *)

(* Ejercicio 19.  *)

T(a_, b_] := {3 a-4 b, a+3 b, -b}
b2={{1, 0, 0}, {0, 1, 1}, {-1, 1, -1}}     (* Los coeficientes de B' y los de *)
aug = AppendRows[b2, Transpose[{T[1, -1], T[1, 1,]}]]    (* T(x-1),T(x+1) *)
RowReduce[aug]              (* en términos de B' se calculan con rref(aug). *)
TakeColumns[%, {4, 5}]      (* Las 2 últimas columnas forman la matriz de T.  *)
```

Soluciones seleccionadas con MATLAB

```
% Ejercicios 1 a 5.

function [A] = T1(x, y, z)             % Definir la función T1 editando y guardando
    A=[2*x-y+z; x+y; 2*y-3*z];         % un archivo llamado T1.m en el directorio
    end                                % actual de trabajo y teclear los 3 renglones a la izquierda.
function [A] = T2(x, y)                % Repetir con T2.
    A = [3*x-4*y; x+3*y; -y];
    end
function [A] = T3(x, y, z)              % T3.
    A = [2*x-y+z; x-z];
    end
function [A] = T4(x, y, z)              % T4.
    A = [-y+z; x+1; 2*y];
    end                                % A continuación, en sesión de MATLAB, escribir:
T1(1, 1, 1)                             % También, T2(1,1)  T3(1,1,1)  T4(1,1, 1).
T4(0, 0, 0)                            % Se obtiene un valor distinto de cero, entonces T4 es no lineal.
T4(0, 0, 0) == [0; 0; 0]               % También se puede investigar la igualdad.
[T1(1, 0, 0) T1(0, 1, 0) T1(0, 0, 1)]  % Evaluación en los vectores de base
[T2(1, 0) T3(0, 1)]                    % para obtener la matriz estándar.
[T3(1, 0, 0) T3(0, 1, 0) T3(0, 0, 1)]

% Ejercicios 6 y 7.

M = [1 1 3 4; 2 3 3 4; 3 5 4 4; 4 7 4 5]    % Los vectores del dominio.
N = [1, 0, 1, 1; 0, 1, 1, 1; -1, 0, 1, -1; 1, -1; 1, -1, -1, 1]    % Los valores.
M1 = rref([M eye(4)])                       % Reducción de [M:I_4].
M2 = M1(:, 5:8)
STM = N * M2                                % La matriz estándar
STM * [2; 2; -2; -2]                        % T(2,2,-2,-2]
```

```
% Ejercicios 8 a 15.

% Como siempre, definir las funciones editando y guardando los archivos llamados T1.m, T2.m
% y T3.m en la sesión actual. (El código sigue.) Estando en sesión MATLAB
% evaluar cada función cuando sea necesario.
function [A] = T1(x, y, z, w]
      A = [x+2*y+3*z+4*w; 2*x+3*y+4*z+5*w; 3*x+4*y+5*z+6*w]; end
function [A] = T2(x, y, z, w)
      A = [x+2*y+3*z+4*w; 2*x+2*y+3*z+4*w; 3*x+3*y*3*z+4*w]; end
function [A] = T3(x, y, z, w)
      A = [x+2*y+3*z+4*w; 2*x+2*y+3*z+4*w; 3*x+3*y+3*z+4*w; 4*x+4*y+4*z+4*w]; end
T1(-1, 0, 3, -2)        % v1 en el núcleo.
T1(18, -31, 8, 5)       % v2 en el núcleo.
T1(1, -1, 8, 7)         % v3 no está en el núcleo.
M1=[T1(1, 0, 0, 0) T1(0, 1, 0, 0) T1(0, 0, 1, 0) T1(0, 0, 0, 1)]     % Matriz estándar.
Mw1=rref([M1 [2; 7; 12]])    % En contradominio(T1). Última columna es pivote.
Mw2=rref([M1 [42; 59; 76]])  % En contradominio(T1). Última columna es pivote.
Mw3=rref([M1 [42; 59; 77]])  % No en contradominio(T1). Última columna no es pivote.
M1                   % Ya se determinó matriz estándar de T1. Las demás son:
M2=[T2(1, 0, 0, 0) T2(0, 1, 0, 0) T2(0, 0, 1, 0) T2(0, 0, 0, 1)
M3=[T3(1, 0, 0, 0) T3(0, 1, 0, 0) T3(0, 0, 1, 0) T3(0, 0, 0, 1)]
                     % k1 tiene dos vectores, luego nulidad 2.
k1=null(M1)          % Núcleo no cero, T1 no es biunívoca.
                     % Por tanto, no es un isomorfismo.
                     % k2 tiene un vector, así que nulidad 2?
k2=null(M2)          % Núcleo no cero, T2 no es biunívoca
                     % En consecuencia no es un isomorfismo
k3=null(M3)          % k3 no tiene vectores, entonces nulidad 0. T3 es biunívoca.
                     % Las 2 primeras columnas forman una base para el contradominio.
r1=rref(M1)          % El rango es 2. 2 + 2 = cantidad de columnas.
                     % El tercer renglón no tiene pivote, luego no es sobre.
                     % Las 3 primeras columnas forman una base para el contradominio.
r2=rref(M2)          % El rango es 3. 3 + 1 = cantidad de columnas.
                     % Cada renglón tiene un pivote; por tanto, sobre.
                     % Todas las columnas forman una base para el contradominio.
r3=rref(M3)          % El rango es 4. 4 + 0 = cantidad de columnas.
                     % T3 es biunívoca y sobre, por tanto, un isomorfismo.
% Falso. T1 no es sobre.
% Cierto. T2 es sobre.
% Cierto. T3 es sobre.

% Ejercicio 16.

function [A] = T(a, b)      % En un archivo de función se teclea el código:
      A = [3*a-4*b; a+3*b; -b];
      end
T(1, 1)

% Ejercicio 17.

% Primero se forma una matriz con los coeficientes de los polinomios dados,
M = [1 1 3 4; 2 3 3 4; 3 5 4 4; 4 7 4 5]
```

```
% después una matriz con los coeficientes de sus valores.
N = [1 0 1 1; 0 1 1 1; -1 0 1 -1; 1 -1 -1 1]
% La última columna de M1 tiene por elementos los coeficientes de 2x^3+2x^2-2x-2
M1 = rref([M[2; 2; -2; -2]]      % en términos de los polinomios dados.
M1(: , 5), N * ans               % Multiplicación por N para evaluar T(2,2,-2,-2).

% Ejercicio 18.

% Igualar F=0. Se necesita resolver a+2b+3c+4d=0,2a+3b+4c+5d=0, 3a+4b+5c+6d=0.
M = [1 2 3 4; 2 3 4 5; 3 4 5 6]      % Se determina una base para
null(M)              % el espacio nulo de la matriz de coeficientes.
% ADVERTENCIA: La respuesta es muy distinta de  lo correcto N/norm(N), donde N=[1 2;-2 -3;1 0;0 1].
% En forma polinomial esto es {0.4082x^3-0.8165x²+0.4082x,0.5345x^3-0.8018x^2+0.2673}

% Ejercicio 19.

function [A] = T(a, b)           % En un archivo de función teclear la transformación.
    A = [3*a-4*b; a+3*b; -b];
b2=[1, 0, 0; 0, 1, 1; -1, 1, -1]   % Los coeficientes de 𝓑'.
aug = [b2, T (1, -1), T (1, 1) ]   % Los coeficientes de T(x-1),T(x+1).
rref(aug)                          % Los coeficientes de 𝓑' se calculan con rref(aug).
and(:, 4:5)                        % Las últimas 2 columnas forman la matriz de T.
```

6

Determinantes

El álgebra es generosa; con frecuencia da más de lo que se le pide.
Jean Le Rond D'Alembert (1717-1783)

Introducción

Los determinantes es uno de los temas más útiles del álgebra lineal, con muchas aplicaciones en ingeniería, física, economía, matemáticas y otras ciencias. En la geometría ofrecen una forma natural de escritura de fórmulas muy elegantes que calculan áreas y volúmenes, y también ecuaciones de objetos geométricos como rectas, círculos, planos, esferas, etcétera.

Dirichlet nos dice que los determinantes fueron introducidos por Leibniz, en una carta a L'Hôpital fechada el 28 de abril de 1693. También hay pruebas de que Seki Takakazu, matemático japonés, ya los usaba en 1683. Los principales contribuyentes en esta área han sido Laplace, Cauchy, Jacobi, Bezout, Sylvester y Cayley.

El problema de las vacas y los campos de Newton

En 1707, Sir Isaac Newton[1] propuso el siguiente problema. Supongamos que

a_1 vacas consumen b_1 campos en c_1 días
a_2 vacas consumen b_2 campos en c_2 días
a_3 vacas consumen b_3 campos en c_3 días

Considerando que todos los campos producen la misma cantidad de pasto, que el crecimiento diario de los campos permanece constante y que las vacas comen la misma cantidad cada día, ¿qué relación existe entre los 9 números $a_1, \ldots, c_3$?

[1] **Sir Isaac Newton** (1642–1727) nació en la Navidad del año que Galileo murió, en Woolsthorpe-by-Colsterworth, Inglaterra. Revolucionó tanto las matemáticas como la física al inventar el cálculo y la teoría de la gravitación en mecánica. Su trabajo sobre la gravedad explica los movimientos planetarios descritos por Kepler. Acerca de sus asombrosos descubrimientos, dijo: "Si he visto más allá que otros es porque me he apoyado en hombros de gigantes." Murió el 20 de marzo de 1727, y fue enterrado en la Abadía de Westminster. Junto con Arquímedes y Gauss, Newton se yergue entre los tres matemáticos más grandes de todos los tiempos.

Al principio, lo sorprendente es que pueda haber alguna relación. Sin embargo, como explicaremos en la sección 6.5, esos números satisfacen la siguiente condición. El determinante de la matriz

$$\begin{bmatrix} a_1c_1 & b_1c_1 & b_1 \\ a_2c_2 & b_2c_2 & b_2 \\ a_3c_3 & b_3c_3 & b_3 \end{bmatrix}$$

es cero.

Aunque los determinantes aparecieron en las publicaciones a fines del siglo XV (mucho antes que las matrices),[2] el primer trabajo que los estudió en forma sistemática fue escrito por Vandermonde[3] en 1772.

> CONVENCIÓN
>
> A menos que se diga otra cosa, en este capítulo todas las matrices son cuadradas.

6.1 Determinantes y desarrollo en cofactores

> Objetivo del estudiante para esta sección
>
> Calcular las determinantes aplicando el desarrollo por cofactores.

Sea $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. El determinante de A es el número

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

■ EJEMPLO 1

$$\det\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2, \quad \det\begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} = 2 \cdot (-2) - (-1) \cdot 4 = 0$$

Sea

$$B = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

[2] Véase una historia breve sobre el tema en *Lessons Introductory to the Higher Modern Algebra*, de George Salmon, D. D., 5a. ed. (Chelsea Publishing Company, 1885), pp. 338-339.

[3] **A. T. Vandermonde** (1735–1796), francés, autor de algunas de las primeras contribuciones a la teoría de los determinantes. También trabajó en geometría; Gauss conoció sus trabajos. (No debe confundirse con C. A. Vandermonde, matemático contemporáneo que trabajó en la deducción de una fórmula para las raíces de las ecuaciones polinomiales.)

El determinante de B puede escribirse en términos de 2×2.

$$\det(B) = a_{11}\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

o en la forma explícita siguiente:

$$\det(B) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

También hay un artificio para memorizar esta fórmula, llamado **esquema de Sarrus**. Se agregan las dos primeras columnas a la derecha de B y se forman los productos de los elementos que atraviesan las flechas. A los productos de las flechas que van de la izquierda superior a la derecha inferior se les asigna el signo más y a los otros el signo menos. A continuación se suman todos los productos con signo.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

$$\det(B) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

ADVERTENCIA El esquema de Sarrus *no se aplica* a los determinantes 4×4, 5×5, y a las de más alto orden.

Sean

$$C = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & -2 \\ 0 & 2 & -1 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 2 & 0 & 1 \\ -1 & 1 & 2 & 0 \\ -2 & 1 & 0 & -2 \\ 1 & 0 & 2 & -1 \end{bmatrix}$$

■ EJEMPLO 2

$$\det(C) = 1\det\begin{bmatrix} 0 & -2 \\ 2 & -1 \end{bmatrix} - 2\det\begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix} + 0\det\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = 1\cdot 4 - 2\cdot(-1) + 0\cdot 2 = 6$$

Del mismo modo pueden definirse los determinantes 4×4:

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = a_{11}\det\begin{bmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{bmatrix} - a_{12}\det\begin{bmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{bmatrix}$$

$$+ a_{13}\det\begin{bmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{bmatrix} - a_{14}\det\begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

■ EJEMPLO 3

$$\det(D) = 1\det\begin{bmatrix}1 & 2 & 0\\ 1 & 0 & -2\\ 0 & 2 & -1\end{bmatrix} - 2\det\begin{bmatrix}-1 & 2 & 0\\ -2 & 0 & -2\\ 1 & 2 & -1\end{bmatrix} + 0\det\begin{bmatrix}-1 & 1 & 0\\ -2 & 1 & -2\\ 1 & 0 & -1\end{bmatrix}$$

$$- 1\det\begin{bmatrix}-1 & 1 & 2\\ -2 & 1 & 0\\ 1 & 0 & 2\end{bmatrix} = 1\cdot 6 - 2\cdot(-12) + 0\cdot(-3) - 1\cdot 0 = 30$$

Podemos continuar en la misma forma y definir los determinantes $n \times n$ en términos de los de $(n-1)\times(n-1)$, los cuales se llaman *menores*. El **menor** (i, j) de una matriz A se representa por M_{ij} y es el determinante que se obtiene eliminando el i-ésimo renglón y la j-ésima columna de A.

Hemos presentado lo que se llama *desarrollo de un determinante por cofactores respecto al primer renglón*. Cada elemento de este renglón se multiplicó por su menor correspondiente. Cada uno de esos productos se multiplicó por ± 1, dependiendo de la posición del elemento. Los productos con signo se sumaron. De hecho, no hay nada especial acerca de la elección del primer renglón en el cálculo del determinante. Pudimos haber usado *cualquier otro renglón o columna*. Veamos cómo.

Sea A es una matriz cuadrada. Primero asignamos un signo a cada elemento de A siguiendo una distribución de tablero de ajedrez, con signos más y menos:

$$\begin{bmatrix}+ & - & + & \cdots\\ - & + & - & \cdots\\ + & - & + & \cdots\\ \vdots & & \vdots & \ddots\end{bmatrix}$$

En seguida se elige cualquier renglón o columna y se multiplica cada elemento con su signo que da la tabla, por el menor correspondiente. Por último, se suman todos esos productos. Observe que el signo de la (i, j)-ésima posición en esa distribución de tablero de ajedrez se obtiene con el factor $(-1)^{i+j}$.

■ EJEMPLO 4 El $\det(C)$ desarrollado respecto al tercer renglón es

$$\det(C) = 0\det\begin{bmatrix}2 & 0\\ 0 & -2\end{bmatrix} - 2\det\begin{bmatrix}1 & 0\\ 1 & -2\end{bmatrix} - 1\det\begin{bmatrix}1 & 2\\ 1 & 0\end{bmatrix} = 6$$

■ EJEMPLO 5 El $\det(D)$ que se amplió respecto a la segunda columna es

$$\det(D) = -2\det\begin{bmatrix}-1 & 2 & 0\\ -2 & 0 & -2\\ 1 & 2 & -1\end{bmatrix} + 1\det\begin{bmatrix}1 & 0 & 1\\ -2 & 0 & -2\\ 1 & 2 & -1\end{bmatrix} - 1\det\begin{bmatrix}1 & 0 & 1\\ -1 & 2 & 0\\ 1 & 2 & -1\end{bmatrix}$$

$$+ 0\det\begin{bmatrix}1 & 0 & 1\\ -1 & 2 & 0\\ -2 & 0 & -2\end{bmatrix} = -2\cdot(-12) + 1\cdot 0 - 1\cdot(-6) + 0\cdot 10 = 30$$

NOTA *Por lo común se trata de desarrollar un determinante respecto al renglón o a la columna que tenga más ceros.* Con esto se evita el cálculo de algunos de los menores.

En forma más general, sea

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

El (i, j)-ésimo **cofactor**, C_{ij}, de A es el (i, j)-ésimo menor con signo.

$$C_{ij} = (-1)^{i+j} M_{ij}$$

Desarrollo por cofactores respecto al i-ésimo renglón

El determinante de A puede desarrollarse con respecto al i-ésimo renglón en términos de cofactores como sigue:

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

Desarrollo mediante cofactores respecto a la j-ésima columna

El determinante de A puede desarrollarse con respecto a la j-ésima columna en términos de cofactores como sigue:

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

Este método de calcular determinantes por medio de cofactores se llama **desarrollo por cofactores, o desarrollo de Laplace**, y se atribuye a Vandermonde y a Laplace.[4]

■ EJEMPLO 6 Sea

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 4 & 3 & -2 \\ -5 & 0 & 3 \end{bmatrix}$$

Entonces

$$M_{11} = \det \begin{bmatrix} 3 & -2 \\ 0 & 3 \end{bmatrix} = 9 \qquad C_{11} = (-1)^{1+1} M_{11} = 9$$

$$M_{12} = \det \begin{bmatrix} 4 & -2 \\ -5 & 3 \end{bmatrix} = 2 \qquad C_{12} = (-1)^{1+2} M_{12} = -1 \cdot 2 = -2$$

$$M_{13} = \det \begin{bmatrix} 4 & 3 \\ -5 & 0 \end{bmatrix} = 15 \qquad C_{13} = (-1)^{1+3} M_{13} = 15$$

[4] **(Marqués de) Pierre Simon Laplace** (1749–1827) nació en Beumont–en–Auge, Normandía, Francia. Escribió su primer trabajo publicado, sobre cálculo de diferencias finitas, a los 16 años. Es autor de contribuciones importantes al cálculo, la mecánica celeste y la teoría de la probabilidad. Desempeñó por poco tiempo el cargo de ministro del interior bajo el imperio de Napoleón, llegó a ser presidente del senado y después se le otorgó el título de conde.

$$M_{21} = \det\begin{bmatrix} 2 & 2 \\ 0 & 3 \end{bmatrix} = 6 \qquad C_{21} = (-1)^{2+1}M_{21} = -1\cdot 6 = -6$$

$$M_{22} = \det\begin{bmatrix} -1 & 2 \\ -5 & 3 \end{bmatrix} = 7 \qquad C_{22} = (-1)^{2+2}M_{22} = 7$$

$$M_{23} = \det\begin{bmatrix} -1 & 2 \\ -5 & 0 \end{bmatrix} = 10 \qquad C_{23} = (-1)^{2+3}M_{23} = -1\cdot 10 = -10$$

$$M_{31} = \det\begin{bmatrix} 2 & 2 \\ 3 & -2 \end{bmatrix} = -10 \qquad C_{31} = (-1)^{3+1}M_{31} = -10$$

$$M_{32} = \det\begin{bmatrix} -1 & 2 \\ 4 & -2 \end{bmatrix} = -6 \qquad C_{32} = (-1)^{3+2}M_{32} = (-1)(-6) = 6$$

$$M_{33} = \det\begin{bmatrix} -1 & 2 \\ 4 & 3 \end{bmatrix} = -11 \qquad C_{33} = (-1)^{3+3}M_{33} = -11$$

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = (-1)9 + 2(-2) + 2\cdot 15 = 17$$
$$\det A = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} = 4(-6) + 3\cdot 7 + (-2)(-10) = 17$$
$$\det A = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} = (-5)(-10) + 0\cdot 6 + 3(-11) = 17$$
$$\det A = a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} = (-1)9 + 4(-6) + (-5)(-10) = 17$$
$$\det A = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} = 2(-2) + 3\cdot 7 + 0\cdot 6 = 17$$
$$\det A = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33} = 2\cdot 15 + (-2)(-10) + 3(-11) = 17$$

NOTA Si multiplicamos los elementos de un renglón (o columna) por los cofactores correspondientes de *otro* renglón (o columna) el resultado siempre es *cero*. (Véase la demostración del teorema 10, sección 6.3.) Por ejemplo,

$$a_{11}C_{21} + a_{12}C_{22} + a_{13}C_{23} = (-1)(-6) + 2\cdot 7 + 2(-10) = 0$$
$$a_{11}C_{12} + a_{21}C_{22} + a_{31}C_{32} = (-1)(-2) + 4\cdot 7 + (-5)6 = 0$$

OBSERVACIÓN El desarrollo por cofactores implica que el determinante de cualquier matriz triangular superior o inferior es el producto de los elementos de su diagonal principal.

Por ejemplo, el desarrollo repetido respecto a la primera columna da como resultado

$$\det\begin{bmatrix} 4 & 5 & 6 \\ 0 & 7 & 8 \\ 0 & 0 & 9 \end{bmatrix} = 4\cdot\det\begin{bmatrix} 7 & 8 \\ 0 & 9 \end{bmatrix} = 4\cdot 7\cdot 9 = 252$$

La notación | |

Con frecuencia se escribe $|A|$ en lugar de $\det(A)$, que no debe confundirse con el valor absoluto. Por ejemplo,

$$\det\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} = \begin{vmatrix}1 & 2\\ 3 & 4\end{vmatrix} = -2$$

A veces conviene hablar de determinantes de 1×1. El determinante de una matriz $A = [a]$ 1×1 es tan sólo a, el único elemento. Por ejemplo, $\det[-2] = -2$, y $\det[3] = 3$.

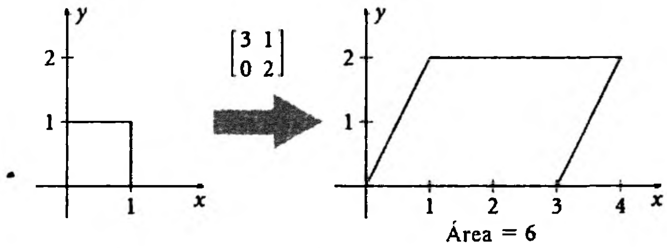
La sorprendente geometría del determinante

En esta sección investigaremos en forma breve las propiedades más sorprendentes del determinante. En general, si se aplica una transformación lineal a una región en el plano, se modifica el área de su imagen. Entonces surge la pregunta siguiente: dada la matriz de transformación, ¿cómo podemos predecir el área de la imagen? A continuación demostraremos que el área del cuadrado unitario se escala con un factor igual al valor absoluto del determinante de la matriz de la transformación.

Por ejemplo, veamos el efecto de

$$T(\mathbf{x}) = \begin{bmatrix}3 & 1\\ 0 & 2\end{bmatrix}\mathbf{x}$$

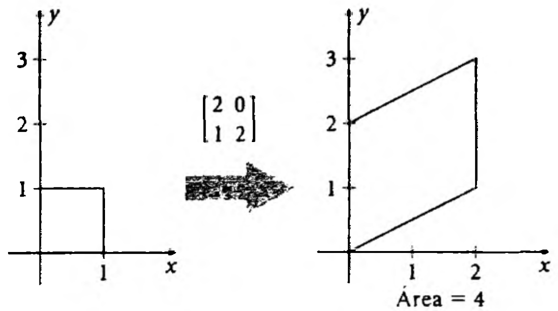
sobre el cuadrado unitario. La imagen es el paralelogramo con vértices en (0, 0), (3, 0), (1, 2) y (4, 2). El área de la imagen es 6, y sucede que es igual al determinante de la matriz.



De modo semejante, veamos el efecto de

$$T(\mathbf{x}) = \begin{bmatrix}2 & 0\\ 1 & 2\end{bmatrix}\mathbf{x}$$

sobre el cuadrado unitario. La imagen es el paralelogramo con vértices en (0, 0), (2, 1), (0, 2) y (2, 3). De nuevo, el área de la imagen, que es 4, es igual al determinante de la matriz.



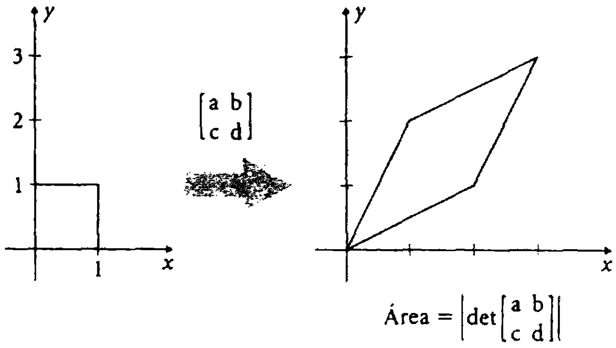
En general, al aplicar

$$T(\mathbf{x}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mathbf{x}$$

al cuadrado unitario, las imágenes de (0, 0), (1, 0), (0, 1), (1, 1) son (0, 0), (a, c), (b, d), $(a + b, c + d)$, respectivamente. Definen un paralelogramo si (a, c) no es proporcional a (b, d), es decir, si la matriz es invertible (lo cual también significa que el determinante $ad - bc \neq 0$), ¿por qué? En este caso podemos calcular con rapidez el área del paralelogramo usando productos cruz. Recordamos, del capítulo 2, sección 2.6, que el área del paralelogramo con dos vectores-3 dados, **u** y **v**, como lados adyacentes, es igual a la longitud de $\mathbf{u} \times \mathbf{v}$. En este caso, sean $\mathbf{u} = (a, c, 0)$ y $\mathbf{v} = (b, d, 0)$, al convertir los vectores-2 a vectores-3 podemos usar el producto cruz. Así, el área de la imagen de T es

$$\|\mathbf{u} \times \mathbf{v}\| = |ad - bd|$$

que es el valor absoluto del determinante de T.



Determinantes con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> det(matrix([[1,7,-8], [5,2,-3], [1,-3,-2]]));
                              172
```

Mathematica

```
In[1]:=
     Det[{{1,7,-8},{5,2,-3},{1,-3,-2}}]
Out[1]=
     172
```

MATLAB

```
>> det([1 7 -8; 5 2 -3; 1 -3 -2])
ans =
   172
```

Ejercicios 6.1

En los ejercicios 1 a 11 calcule los determinantes.

1. a. $\begin{vmatrix} 3 & 6 \\ 4 & 7 \end{vmatrix}$ b. $\begin{vmatrix} 3 & 4 \\ 6 & 7 \end{vmatrix}$

2. a. $\begin{vmatrix} 0 & -1 \\ -1 & 100 \end{vmatrix}$ b. $\begin{vmatrix} \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \end{vmatrix}$

3. a. $\begin{vmatrix} 1 & 0 & 2 \\ 0 & 5 & 0 \\ 3 & 0 & 4 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 5 \end{vmatrix}$

4. a. $\begin{vmatrix} 0 & 0 & 3 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{vmatrix}$ b. $\begin{vmatrix} 0 & 7 & -2 \\ 1 & 3 & 8 \\ 0 & 1 & 1 \end{vmatrix}$

5. a. $\begin{vmatrix} 11 & 1 & 0 \\ -2 & 7 & 0 \\ 3 & 0 & 2 \end{vmatrix}$ b. $\begin{vmatrix} -3 & 5 & 1 \\ 1 & 0 & -2 \\ 0 & 3 & 0 \end{vmatrix}$

6. a. $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{vmatrix}$

7. a. $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 7 & 0 & 3 & 0 \\ 9 & -1 & 4 & 0 \end{vmatrix}$

8. a. $\begin{vmatrix} 1 & 2 & -3 & -4 \\ 0 & 2 & 5 & 6 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 0 & 4 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & -1 \\ 3 & 0 & 1 & 1 \\ 4 & 0 & 1 & 0 \end{vmatrix}$

9. $\begin{vmatrix} 1 & -1 & 2 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 4 & 6 & 7 & 8 & 0 \\ -4 & 7 & 0 & 1 & 1 \end{vmatrix}$

10. $\begin{vmatrix} 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & -1 \\ 1 & 3 & 0 & 1 & 1 \\ 1 & 4 & 0 & 1 & 0 \end{vmatrix}$

11. $\begin{vmatrix} a & a & a \\ a & a & a \\ b & c & d \end{vmatrix}$

12. Calcule el determinante de la matriz correspondiente al problema de las vacas y los campos, de Newton.

$$\begin{vmatrix} a_1c_1 & b_1c_1 & b_1 \\ a_2c_2 & b_2c_2 & b_2 \\ a_3c_3 & b_3c_3 & b_3 \end{vmatrix}$$

En los ejercicios 13 y 14 calcule el determinante de la matriz de coeficientes del sistema.

13.
$$\begin{aligned} 5x - 2y &= 1 \\ -x + y + 2z &= -3 \\ -7y + 2z &= -2 \end{aligned}$$

14.
$$\begin{aligned} y - z &= 0 \\ x \quad - z &= 1 \\ -x + y \quad &= -1 \end{aligned}$$

15. Sean

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$$

a. Demuestre que $\det(A) = \det(A^T)$.

b. Compruebe que $\det(AB) = \det(A)\det(B)$.

c. Demuestre que $\det(A^{-1}) = 1/\det(A)$.

16. Escriba todos los menores M_{ij} y todos los cofactores C_{ij} de A.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

17. a. Determine todos los menores y todos los cofactores de A.

$$A = \begin{bmatrix} 1 & -2 & 2 \\ 3 & 5 & -4 \\ 7 & 0 & -6 \end{bmatrix}$$

b. Calcule $\det(A)$ desarrollando por cofactores respecto a

(b_1) el primer renglón
(b_2) el segundo renglón
(b_3) el tercer renglón
(b_4) la primera columna
(b_5) la segunda columna
(b_6) la tercera columna

Sean

$$A = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

18. Compruebe que $\det(AB) = \det(A)\det(B)$.
19. Compruebe que $\det(A^{-1}) = 1/\det(A)$.
20. Verifique la igualdad

$$\left|\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} e & f \\ g & h \end{bmatrix}\right| = \begin{vmatrix} a & b \\ c & d \end{vmatrix}\begin{vmatrix} e & f \\ g & h \end{vmatrix}$$

21. Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ con $ad - bc \neq 0$. Demuestre que

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Determinantes de matrices elementales

En los ejercicios 22 a 24 calcule los determinantes de las matrices elementales.

22. $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ r & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & r & 1 \end{bmatrix}$

23. $E_3 = \begin{bmatrix} r & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & r \end{bmatrix}$

24. $E_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $E_6 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

25. Con base en sus cálculos de los ejercicios 22 a 24, haga una conjetura acerca de los determinantes de los tres tipos de matrices elementales: los obtenidos a partir de I por (a) eliminación, (b) escalamiento y (c) intercambio.

Sea

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

y sean

	Obtenidas de A por
A_1	$R_2 + rR_1 \to R_2$
A_2	$R_3 + rR_2 \to R_3$
A_3	$rR_1 \to R_1$
A_4	$rR_3 \to R_3$
A_5	$R_2 \leftrightarrow R_3$
A_6	$R_1 \leftrightarrow R_3$

En los ejercicios 26 a 28 demuestre las identidades.

26. a. $\det(E_1A) = \det(E_1)\det(A) = \det(A_1)$
b. $\det(E_2A) = \det(E_2)\det(A) = \det(A_2)$

27. a. $\det(E_3A) = \det(E_3)\det(A) = \det(A_3)$
b. $\det(E_4A) = \det(E_4)\det(A) = \det(A_4)$

28. a. $\det(E_5A) = \det(E_5)\det(A) = \det(A_5)$
b. $\det(E_6A) = \det(E_6)\det(A) = \det(A_5)$

Ecuaciones con determinantes

En los ejercicios 29 a 31 calcule todos los valores de λ (reales o complejos) tales que los determinantes sean cero.

29. a. $\begin{vmatrix} 1-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix}$
b. $\begin{vmatrix} 1-\lambda & 5 \\ 2 & 10-\lambda \end{vmatrix}$

30. a. $\begin{vmatrix} \lambda & 1 \\ 4 & \lambda \end{vmatrix}$
b. $\begin{vmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix}$

31. $\begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda-2 & 1 \\ 0 & 2 & \lambda-1 \end{vmatrix}$

32. $\begin{vmatrix} 2-\lambda & 0 & 0 \\ 1 & 3-\lambda & 0 \\ 0 & 1 & 1-\lambda \end{vmatrix}$

33. Calcule el valor de x.

$$\begin{vmatrix} x & x \\ 1 & 2 \end{vmatrix} = \begin{vmatrix} x-1 & 1 \\ -1 & x-1 \end{vmatrix}$$

34. Despeje el valor de x.

$$\begin{vmatrix} x & 0 & 2 \\ 0 & x-3 & -1 \\ 2 & 0 & x-6 \end{vmatrix} = \begin{vmatrix} x-1 & 1 & 0 \\ -2 & x-7 & 0 \\ 3 & 0 & x-2 \end{vmatrix}$$

35. Determine a, b tales que

$$\begin{vmatrix} a & b \\ 1 & 2 \end{vmatrix} = 0 = \begin{vmatrix} a & 4 \\ 1 & a \end{vmatrix}$$

Volúmenes y determinantes

Sea R un cubo unitario en el espacio. Éste es un cubo de lado 1 en el primer octante con lados adyacentes a los planos coordinados.

36. Sea T la transformación lineal cuya matriz es

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

Calcule el volumen de la imagen de R bajo T y relaciónelo con el determinante de A.

37. Sea T la transformación lineal general cuya matriz es

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Compruebe que el volumen de la imagen de R bajo T es igual a $|\det(A)|$. (*Sugerencia:* Si T es invertible, use la fórmula de la sección 2.6 para el volumen del paralelepípedo determinado por los tres vectores.)

6.2 Propiedades de los determinantes

Objetivos del estudiante para esta sección

1. Comprender y aplicar las propiedades básicas de los determinantes.
2. Simplificar los determinantes con una reducción correcta de renglones o de columnas.

El desarrollo de los determinantes por cofactores es bastante tedioso, a menos que la matriz sea pequeña ó que tenga muchos ceros. El mejor método es la eliminación de Gauss. Pero antes de continuar necesitamos estudiar los efectos de las operaciones elementales de renglón sobre los determinantes.

Operaciones elementales y determinantes

Las demostraciones de los casos especiales del teorema siguiente se dejan como ejercicios. Una demostración completa se describe en el conjunto de ejercicios de la sección 6.4.

TEOREMA 1 (Propiedades básicas)

Sea A una matriz $n \times n$. (Para ilustrar, tomaremos $n = 3$.)

1. A y su transpuesta tienen el mismo determinante: $\det(A) = \det(A^T)$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

2. Si B se obtiene de A multiplicando uno de sus renglones (o columnas) por una constante distinta de cero, entonces $\det(B) = k\det(A)$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ kb_1 & kb_2 & kb_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = k\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \qquad \begin{vmatrix} a_1 & a_2 & ka_3 \\ b_1 & b_2 & kb_3 \\ c_1 & c_2 & kc_3 \end{vmatrix} = k\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

3. Si B se obtiene de A intercambiando dos renglones (o columnas) cualesquiera, en ese caso $\det(B) = -\det(A)$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = -\begin{vmatrix} b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \qquad \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = -\begin{vmatrix} a_3 & a_2 & a_1 \\ b_3 & b_2 & b_1 \\ c_3 & c_2 & c_1 \end{vmatrix}$$

4. Si B se obtiene de A sumando un múltiplo de un renglón (o columna) a otro renglón (o columna), entonces $\det(B) = \det(A)$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ ka_1 + b_1 & ka_2 + b_2 & ka_3 + b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & a_2 & ka_2 + a_3 \\ b_1 & b_2 & kb_2 + b_3 \\ c_1 & c_2 & kc_2 + c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Vemos que la eliminación $R_i + kR_j \to R_i$ no cambia el determinante; en el escalamiento $kR_i \to R_i$ se escala el determinante por k; y en el intercambio, $R_i \leftrightarrow R_j$, se cambia el signo del determinante.

■ EJEMPLO 7

$$\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 1 \\ 3 & 2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = -\begin{vmatrix} 3 & 1 \\ 4 & 2 \end{vmatrix}$$

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 0 & -2 \end{vmatrix} \quad \text{por} \quad \boxed{-3R_1 + R_2 \to R_2}$$

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} \quad \text{por} \quad \boxed{-C_1 + C_2 \to C_2}$$

□

ERROR FRECUENTE A veces se aplica mal la propiedad 4 del teorema 1. Un renglón (o columna) se reemplaza por un múltiplo de otro sumado al primero, y *no un múltiplo de él.* Si el renglón (o la columna) original se escala, también se escala el determinante. Por ejemplo,

$$\begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ 0 & \frac{5}{3} \end{vmatrix} = 5 \quad \text{por} \quad \boxed{-\tfrac{1}{3}R_1 + R_2 \to R_2}$$

mientras que $\boxed{R_1 - 3R_2 \to R_2}$ implica que $\begin{vmatrix} 3 & 1 \\ 0 & -5 \end{vmatrix} = -15$

Pueden emplearse las propiedades que menciona el teorema 1 para escribir un determinante en forma triangular, por eliminación de Gauss, y a continuación se calcula el producto de los elementos diagonales.

■ EJEMPLO 8

$$\begin{vmatrix} 2 & 4 & 6 & -2 & 16 \\ 0 & 0 & 4 & 2 & -1 \\ 0 & -5 & 5 & 3 & 7 \\ 0 & 0 & 0 & 1 & 6 \\ 1 & 2 & 3 & -2 & -9 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 & -1 & 8 \\ 0 & 0 & 4 & 2 & -1 \\ 0 & -5 & 5 & 3 & 7 \\ 0 & 0 & 0 & 1 & 6 \\ 1 & 2 & 3 & -2 & -9 \end{vmatrix} \quad \text{por} \quad \boxed{\tfrac{1}{2}R_1 \to R_1}$$

$$= 2\begin{vmatrix} 1 & 2 & 3 & -1 & 8 \\ 0 & 0 & 4 & 2 & -1 \\ 0 & -5 & 5 & 3 & 7 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & -1 & -17 \end{vmatrix} \quad \text{por} \quad \boxed{-R_1 + R_5 \to R_5}$$

$$= -2\begin{vmatrix} 1 & 2 & 3 & -1 & 8 \\ 0 & -5 & 5 & 3 & 7 \\ 0 & 0 & 4 & 2 & -1 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & -1 & -17 \end{vmatrix} \quad \text{por} \quad \boxed{R_2 \leftrightarrow R_3}$$

$$= -2\begin{vmatrix} 1 & 2 & 3 & -1 & 8 \\ 0 & -5 & 5 & 3 & 7 \\ 0 & 0 & 4 & 2 & -1 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & -11 \end{vmatrix} \quad \text{por} \quad \boxed{R_4 + R_5 \to R_5}$$

$$= -2 \cdot (-5) \cdot 4 \cdot 1(-11) = -440$$

El método del ejemplo 8 también da como resultado una fórmula para el determinante: primero se observa que siempre es posible reducir cualquier matriz a su forma escalonada *sin usar operación alguna de escalamiento*. (En el ejemplo 8 eso significaría no hacer $\frac{1}{2}R_1 \to R_1$ y no sacar el factor adicional 2.)

Sea A una matriz $n \times n$ reducida mediante operaciones de renglón, sin escalamiento, a la matriz triangular superior B. Las únicas operaciones que alteran el $\det(A)$ (y sólo por un signo) son los intercambios. Por consiguiente,

$$\det(A) = (-1)^k \det(B)$$

donde k es la cantidad de intercambios en el proceso de reducción. Si A es invertible, entonces B tiene n pivotes, por ejemplo $p_1, \ldots, p_n$, todos en la diagonal principal, porque $A \sim B \sim I$. Así, $\det(A) = (-1)^k \det(B) = (-1)^k p_1 p_2 \cdots p_n$. Si A es no invertible, B tiene al menos un renglón de ceros, así que $\det(A) = \det(B) = 0$. Con ello hemos demostrado el teorema 2.

TEOREMA 2

> En la notación anterior,
>
> $$\det(A) = \begin{cases} (-1)^k p_1 p_2 \cdots p_n & \text{si } A \text{ es invertible} \\ 0 & \text{si } A \text{ es no invertible} \end{cases}$$

Acerca del ejemplo 8, si reducimos sin escalar, los pivotes de la matriz reducida son $-2, -5, 4, 1, -11$. Su producto es el determinante. Ya que los pivotes siempre son distintos de cero, el teorema 2 implica al siguiente teorema básico.

TEOREMA 3

> Una matriz A $n \times n$ es invertible si y sólo si $\det(A) \neq 0$.

■ **EJEMPLO 9** A, B y C no son invertibles, porque sus determinantes son iguales a 0.

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}, \qquad B = \begin{bmatrix} 2a & a \\ 2 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

□

■ **EJEMPLO 10** ¿Son linealmente independientes los renglones que contiene D? ¿Y las columnas?

$$D = \begin{bmatrix} 2 & -4 & 5 \\ 8 & 0 & -3 \\ 5 & -2 & 1 \end{bmatrix}$$

SOLUCIÓN No, porque $\det(D) = 0$. Por consiguiente, los renglones y las columnas de D son linealmente dependientes, de acuerdo con el teorema 3 y el teorema 15 de la sección 3.3. □

■ **EJEMPLO 11** ¿Son linealmente independientes los vectores siguientes?

$$\begin{bmatrix} 2 \\ 8 \\ 5 \end{bmatrix}, \qquad \begin{bmatrix} -4 \\ -7 \\ -2 \end{bmatrix}, \qquad \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$$

RESPUESTA Sí, porque $\begin{vmatrix} 2 & -4 & 5 \\ 8 & -7 & -3 \\ 5 & -2 & 1 \end{vmatrix} = 161 \neq 0.$ □

Como consecuencia del teorema 3 y del teorema 15, de la sección 3.3, tenemos el siguiente y muy útil teorema.

TEOREMA 4

> El sistema homogéneo cuadrado $A\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales si y sólo si $\det(A) = 0$.

■ EJEMPLO 12 ¿Tiene soluciones no triviales el sistema siguiente?

$$\begin{aligned} 2x + 3y + z &= 0 \\ x - y + 2z &= 0 \\ x + 4y - z &= 0 \end{aligned}$$

RESPUESTA Sí, porque de acuerdo con el teorema 4

$$\begin{vmatrix} 2 & 3 & 1 \\ 1 & -1 & 2 \\ 1 & 4 & -1 \end{vmatrix} = 0$$

Ahora llegamos a algunas conclusiones sencillas, pero importantes, del teorema 1 y de los desarrollos por cofactores.

TEOREMA 5

1. Si A tiene un renglón (o columna) de ceros, entonces $\det(A) = 0$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ 0 & 0 & 0 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0, \qquad \begin{vmatrix} a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \\ c_1 & c_2 & 0 \end{vmatrix} = 0$$

2. Si A tiene dos renglones (o columnas) que son iguales, entonces $\det(A) = 0$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0, \qquad \begin{vmatrix} a_1 & a_2 & a_1 \\ b_1 & b_2 & b_1 \\ c_1 & c_2 & c_1 \end{vmatrix} = 0$$

3. Si A tiene dos renglones (o columnas) que son múltiplos entre sí, entonces $\det(A) = 0$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ ka_1 & ka_2 & ka_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0, \qquad \begin{vmatrix} a_1 & a_2 & ka_1 \\ b_1 & b_2 & kb_1 \\ c_1 & c_2 & kc_1 \end{vmatrix} = 0$$

4. Si un renglón (o columna) de A es igual a la suma de múltiplos de otros dos renglones (o columnas), entonces $\det(A) = 0$.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ ka_1 + lc_1 & ka_2 + lc_2 & ka_3 + lc_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0, \qquad \begin{vmatrix} a_1 & a_2 & ka_1 + la_2 \\ b_1 & b_2 & kb_1 + lb_2 \\ c_1 & c_2 & kc_1 + lc_2 \end{vmatrix} = 0$$

DEMOSTRACIÓN La propiedad 1 es fácil (mediante el desarrollo por cofactores respecto al renglón o columna de ceros). Basta demostrar la propiedad 4, porque la 2 y la 3 son consecuencia de esta última si hacemos que $k = 1$, $l = 0$ y $l = 0$, respectivamente. Para ilustrarlo, emplearemos un determinante 3×3.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ ka_1 + lc_1 & ka_2 + lc_2 & ka_3 + lc_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ lc_1 & lc_2 & lc_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \quad \text{de acuerdo con el enunciado 4 del teorema 1.}$$

$$= l\begin{vmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \quad \text{según el enunciado 2, teorema 1.}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \\ 0 & 0 & 0 \end{vmatrix} \quad \text{según el enunciado 4, teorema 1.}$$

El último determinante es cero, de acuerdo con la parte 1. □

■ EJEMPLO 13

$$\begin{vmatrix} 10 & -11 & 3 \\ 20 & 50 & 40 \\ 0 & 0 & 0 \end{vmatrix} = 0 \qquad \begin{vmatrix} 1 & 2 & 4 & 4 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 4 & 4 \\ 1 & -1 & -4 & -4 \end{vmatrix} = 0$$

$$\begin{vmatrix} 10 & 15 & -5 \\ 12 & 21 & -6 \\ -120 & 26 & 60 \end{vmatrix} = \begin{vmatrix} (-2)\cdot(-5) & 15 & -5 \\ (-2)\cdot(-6) & 21 & -6 \\ (-2)\cdot 60 & 26 & 60 \end{vmatrix} = 0$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 1 & -1 & -3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 3-2\cdot 1 & 3-2\cdot 2 & 3-2\cdot 3 \end{vmatrix} = 0$$

□

Operaciones matriciales y determinantes

Veamos cómo se afectan los determinantes con las operaciones matriciales $A + B$, kA y AB.

Desafortunadamente, no hay una fórmula fácil para el determinante de la suma de dos matrices, $\det(A + B)$. En general,

$$\det(A + B) \neq \det(A) + \det(B)$$

■ EJEMPLO 14

$$\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = -1 + 1 = 0, \quad \text{pero} \quad \det\left(\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\right) = \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} = 1$$

□

De las igualdades de la derecha llegamos al teorema 6.

TEOREMA 6

(Suma de renglones)

Si todo elemento en cualquier renglón (o columna) de un determinante es igual a la suma de otros dos, entonces el determinante es igual a la suma de otros dos.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 + d_1 & c_2 + d_2 & c_3 + d_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ d_1 & d_2 & d_3 \end{vmatrix}$$

DEMOSTRACIÓN Se deja como ejercicio. (*Sugerencia:* Si el i-ésimo renglón es $[c_{i1} + d_{i1}, \ldots, c_{in} + d_{in}]$, entonces aplique el desarrollo por cofactores con respecto a él.) ☐

El determinante del producto por un escalar, $\det(kA)$, puede calcularse usando el teorema siguiente.

TEOREMA 7

(Determinante del producto por un escalar)

Si A es cualquier matriz $n \times n$ y k es cualquier escalar, entonces

$$\det(kA) = k^n \det(A)$$

DEMOSTRACIÓN Aplicando repetidamente la propiedad 2 del teorema 1, factorizamos cada renglón con una k a la vez. Si $A = [\mathbf{a}_1\mathbf{a}_2 \cdots \mathbf{a}_n]$, entonces

$$\begin{aligned}\det(kA) &= \det[k\mathbf{a}_1 k\mathbf{a}_2 \cdots k\mathbf{a}_n] = k\det[\mathbf{a}_1 k\mathbf{a}_2 \cdots k\mathbf{a}_n] \\ &= k^2 \det[\mathbf{a}_1\mathbf{a}_2 \cdots k\mathbf{a}_n] = \cdots = k^n \det[\mathbf{a}_1\mathbf{a}_2 \cdots \mathbf{a}_n] = k^n \det(A)\end{aligned}$$

☐

A continuación pasaremos al $\det(AB)$. Esta vez contamos con una bella y útil fórmula, conocida como **teorema de Cauchy**.

TEOREMA 8

(Determinante de un producto)

El determinante de un producto de matrices es el producto de los determinantes de los factores.

$$\det(AB) = \det(A)\det(B)$$

$$\det(A_1A_2 \cdots A_m) = \det(A_1)\det(A_2)\cdots\det(A_m)$$

La demostración, debida a Cauchy,[5] se describe en los ejercicios.

Sean

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 2 & 0 \\ -5 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

■ EJEMPLO 15 Compruebe que $\det(AB) = \det(A)\det(B)$.

SOLUCIÓN

$$\det(A) = -3 \qquad \det(B) = 10 \qquad \det(A)\det(B) = -30$$

$$AB = \begin{bmatrix} -5 & 0 & 0 \\ -5 & 2 & 0 \\ 0 & 2 & 3 \end{bmatrix} \qquad \det(AB) = -30$$

☐

[5] Este hecho también fue descubierto por Gauss para las matrices 2×2 y 3×3.

El teorema 8 tiene una implicación importante.

TEOREMA 9

> Si A es invertible, entonces
> $$\det(A^{-1}) = \frac{1}{\det(A)}$$

DEMOSTRACIÓN A es invertible, de modo que $AA^{-1} = I$. Así, $\det(AA^{-1}) = \det(A)\det(A^{-1}) = \det(I) = 1$, de acuerdo con el teorema 8. Por consiguiente, $\det(A) \neq 0$. La división de $\det(A)\det(A^{-1}) = 1$ entre $\det(A)$ da como resultado la fórmula anterior. ☐

■ EJEMPLO 16 Demuestre que A, B son invertibles. Verifique la igualdad $\det(B^{-1}) = 1/\det(B)$.

SOLUCIÓN $\det(A) = -3$, $\det(B) = 10$. Los determinantes son distintos de cero; por consiguiente, de acuerdo con el teorema 9, las matrices son invertibles.

$$B^{-1} = \begin{bmatrix} 0 & -\frac{1}{5} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \det(B^{-1}) = \frac{1}{10} = \frac{1}{\det(B)}$$ ☐

Ejercicios 6.2

En los ejercicios 1 a 7 evalúe los determinantes sólo por inspección.

1. a. $\begin{vmatrix} 5 & 1 & 14 \\ 0 & 10 & 2 \\ 0 & 0 & -1 \end{vmatrix}$ b. $\begin{vmatrix} 1 & -1 & 1 \\ 2 & -1 & 2 \\ 4 & -2 & 4 \end{vmatrix}$

2. a. $\begin{vmatrix} 1 & 1 & 2 \\ 2 & 2 & 3 \\ 4 & 4 & 8 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{vmatrix}$

3. a. $\begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix}$

4. a. $\begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$ b. $\begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{vmatrix}$

5. a. $\begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 \\ 3 & 6 & 3 & 0 & 0 \\ 4 & 6 & 6 & 4 & 0 \\ 5 & 5 & 5 & 5 & -1 \end{vmatrix}$

7. $\begin{vmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 & 2 \\ 0 & 6 & 3 & 0 & 3 \\ 0 & 6 & 6 & 4 & 5 \\ -1 & 5 & 5 & 5 & 5 \end{vmatrix}$

Sea

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 3$$

En los ejercicios 8 a 15 explique las identidades. No haga cálculos.

8. $\begin{vmatrix} a & b & c \\ d & e & f \\ 2g & 2h & 2i \end{vmatrix} = 6$

9. $\begin{vmatrix} g & h & i \\ d & e & f \\ a & b & c \end{vmatrix} = -3$

10. $\begin{vmatrix} a-4c & b & c \\ d-4f & e & f \\ g-4i & h & i \end{vmatrix} = 3$

11. $\begin{vmatrix} 2b-4c & b & c \\ 2e-4f & e & f \\ 2h-4i & h & i \end{vmatrix} = 0$

12. $\begin{vmatrix} c & b & -a \\ f & e & -d \\ i & h & -g \end{vmatrix} = 3$

13. $\begin{vmatrix} a & b & c \\ d-a & e-b & f-c \\ 3g & 3h & 3i \end{vmatrix} = 9$

14. $\begin{vmatrix} -1 & a & e & i \\ 0 & a & b & c \\ 0 & d & e & f \\ 0 & g & h & i \end{vmatrix} = -3$

15. $\begin{vmatrix} 2a & 2b & 2c \\ 2d & 2e & 2f \\ 2g & 2h & 2i \end{vmatrix} = 24$

21. $\begin{vmatrix} 2 & -4 & 2 & 8 \\ -2 & 3 & 0 & -7 \\ 0 & 1 & 5 & -1 \\ 1 & 0 & 1 & 0 \end{vmatrix}$

22. $\begin{vmatrix} 1 & -1 & 2 & 0 & 0 \\ 0 & 1 & 2 & -2 & 7 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 4 & 0 & 3 \\ 0 & 3 & 0 & 0 & 1 \end{vmatrix}$

23. $\begin{vmatrix} 1 & 0 & -1 & 2 & 1 \\ 2 & 2 & 4 & 2 & 8 \\ 0 & 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{vmatrix}$

24. $\begin{vmatrix} 2 & 1 & 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & -2 & 3 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 4 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{vmatrix}$

En los ejercicios 16 y 17 explique, sin hacer cálculos, por qué las sustituciones $x = 0, 2$, hacen que los determinantes sean iguales a cero.

16. $\begin{vmatrix} x & x & 2x \\ 0 & 1 & 0 \\ 2 & 2 & 4 \end{vmatrix}$

17. $\begin{vmatrix} x & x & 2 \\ 0 & x^2 & 0 \\ 8 & 8 & x^3 \end{vmatrix}$

18. Explique, sin hacer cálculos, por qué son iguales los determinantes siguientes.

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & -b & c \\ -d & e & -f \\ g & -h & i \end{vmatrix}$$

En los ejercicios 19 a 24 evalúe los determinantes con reducciones de renglón.

19. a. $\begin{vmatrix} 2 & 2 & 4 \\ -1 & -1 & -1 \\ 0 & 2 & 0 \end{vmatrix}$ b. $\begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix}$

20. a. $\begin{vmatrix} 2 & -2 & 4 \\ 0 & 1 & 7 \\ -2 & 9 & 0 \end{vmatrix}$ b. $\begin{vmatrix} 1 & -2 & 5 \\ -2 & 6 & -4 \\ 3 & -5 & 0 \end{vmatrix}$

25. Demuestre las propiedades de los determinantes siguientes.

a. $\begin{vmatrix} a_1 & a_2 & ka_3 \\ b_1 & b_2 & kb_3 \\ c_1 & c_2 & kc_3 \end{vmatrix} = k \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$

b. $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = - \begin{vmatrix} c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{vmatrix}$

26. Compruebe las propiedades de los determinantes siguientes.

a. $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

b. $\begin{vmatrix} a_1+kb_1 & a_2+kb_2 & a_3+kb_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$

27. Demuestre el teorema 6.

En los ejercicios 28 a 31 aplique el teorema 3 para determinar qué matrices son invertibles.

28. $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix}$ $B = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$

29. $C = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ $D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

30. $E = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 0 \\ 2 & 3 & -4 \end{bmatrix}$ $F = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

31. $G = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix}$ $H = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

En los ejercicios 32 y 33 encuentre todos los valores de k tales que las matrices no sean invertibles.

32. $\begin{vmatrix} k & k-1 & 1 \\ 0 & k+1 & 4 \\ k & 0 & k \end{vmatrix}$

33. $\begin{vmatrix} k & k^2 & 0 \\ 0 & k^3 & 4 \\ -k & 0 & k \end{vmatrix}$

Sean A y B matrices $n \times n$.

34. Demuestre la identidad

$$\det(AB) = \det(BA)$$

(Es cierta aun en el caso $AB \neq BA$.)

35. Si B es invertible, demuestre que

$$\det(B^{-1}AB) = \det(A)$$

36. Sea A una matriz 3×3 cuyo $\det(A) = -2$. Calcule cada uno de los valores siguientes.

a. $\det(A^3)$ b. $\det(A^{-1})$
c. $\det(A^{-3})$ d. $\det(A^T)$
e. $\det(AA^T)$ f. $\det(-7A)$

37. Sean A y B matrices 4×4 tales que el $\det(A) = -2$ y el $\det(B) = 7$. Calcule cada uno de los valores siguientes.

a. $\det(AB)$ b. $\det(A^3B)$
c. $\det(ABA^{-1})$ d. $\det(BAB^{-1})$
e. $\det(AB^T)$ f. $\det(3AB)$

38. Demuestre que el cuadrado de cualquier determinante, $\det(A)^2$, puede expresarse como el determinante de una matriz simétrica.[a] (*Sugerencia:* Considere la matriz *simétrica* AA^T).

En los ejercicios 39 a 42 use reducción por operaciones de renglón para demostrar las identidades.

39. $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)(c-b)(c-a)$

40. $\begin{vmatrix} 1 & a & b \\ 1 & a^2 & b^2 \\ 1 & a^3 & b^3 \end{vmatrix} = ab(b-a)(a-1)(b-1)$

41. $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix} = (b-a)(c-b)(c-a)(a+b+c)$

42. $\begin{vmatrix} 1 & a & b \\ 1 & a^2 & b^2 \\ 1 & a^4 & b^4 \end{vmatrix} = ab(b-a)(a-1)(b-1)(a+b+1)$

Una demostración del teorema de Cauchy

Los ejercicios siguientes describen una demostración de la fórmula básica $\det(AB) = \det(A)\det(B)$ aplicando el teorema 1.

43. Sea E una matriz elemental. Use el teorema 1 para comprobar que

	Si E se obtiene de I por
$\det(E) = 1$	$R_i + cR_j \rightarrow R_i$
$\det(E) = c$	$cR_i \rightarrow R_i$
$\det(E) = -1$	$R_i \leftrightarrow R_j$

44. Si A y E son matrices $n \times n$, y E es elemental. Emplee el ejercicio 43 para demostrar que

$$\det(EA) = \det(E)\det(A)$$

45. Si A y B son matrices $n \times n$, y A no es invertible. Compruebe que

$$\det(AB) = \det(A)\det(B)$$

(*Sugerencia:* Aplique el teorema 16, de la sección 3.3 y el teorema 3.)

46. Si A y B son matrices $n \times n$, y A es invertible. Demuestre que

$$\det(AB) = \det(A)\det(B)$$

(*Sugerencia:* Escriba A como el producto de matrices elementales. A continuación use el resultado del ejercicio 44.)

[a] (a) Recuerde que B es simétrica si $B^T = B$. (b) La afirmación de este ejercicio se debe a Lagrange.

6.3 La adjunta; la regla de Cramer

Objetivos del estudiante para esta sección

1. Calcular la matriz inversa empleando a la matriz adjunta.
2. Resolver sistemas lineales $n \times n$ empleando la regla de Cramer.

Adjunta e inversa

DEFINICIÓN

Sea A una matriz $n \times n$. La matriz cuyo (i, j)-ésimo elemento es el cofactor C_{ij} de A es la **matriz de cofactores** de A. Su transpuesta es la **adjunta de** A, y se representa por $\operatorname{Adj}(A)$.

$$\mathbf{Adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

En la sección 6.1 vimos que los cofactores de

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 4 & 3 & -2 \\ -5 & 0 & 3 \end{bmatrix}$$

son

$$\begin{array}{lll} C_{11} = 9 & C_{12} = -2 & C_{13} = 15 \\ C_{21} = -6 & C_{22} = 7 & C_{23} = -10 \\ C_{31} = -10 & C_{32} = 6 & C_{33} = -11 \end{array}$$

■ EJEMPLO 17 La matriz de cofactores de A es

$$\begin{bmatrix} 9 & -2 & 15 \\ -6 & 7 & -10 \\ -10 & 6 & -11 \end{bmatrix}$$

La adjunta de A es

$$\operatorname{Adj}(A) = \begin{bmatrix} 9 & -6 & -10 \\ -2 & 7 & 6 \\ 15 & -10 & -11 \end{bmatrix}$$

□

En el capítulo 3 describimos un algoritmo con operaciones de renglón —pero *no presentamos la fórmula*— para calcular A^{-1} por reducción de $[A : I]$. Ahora, con los determinantes a nuestro alcance, podemos presentar una *fórmula explícita* de A^{-1}.

TEOREMA 10

Sea A una matriz $n \times n$. Entonces

$$A\,\mathrm{Adj}(A) = \det(A)I_n = \mathrm{Adj}(A)A$$

DEMOSTRACIÓN Examinemos el producto $\mathrm{Adj}(A)A$.

$$\mathrm{Adj}(A)A = \begin{bmatrix} C_{11} & \cdots & C_{j1} & \cdots & C_{n1} \\ \vdots & & \vdots & & \vdots \\ C_{1i} & \cdots & C_{ji} & & C_{ni} \\ \vdots & & \vdots & & \vdots \\ C_{1n} & \cdots & C_{jn} & \cdots & C_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}$$

El (i, j)-ésimo elemento es

$$C_{1i}a_{1j} + C_{2i}a_{2j} + \cdots + C_{ni}a_{nj}$$

Esta suma puede considerarse como el desarrollo del determinante por cofactores respecto a la j-ésima columna de la matriz A', que se obtiene de A reemplazando la j-ésima columna por la i-ésima. Si $i = j$, la suma es el $\det(A)$, porque $A = A'$. Si $i \neq j$, la suma es 0, porque A tiene una columna repetida, de acuerdo con el teorema 5. Por consiguiente,

$$\mathrm{Adj}(A)A = \begin{bmatrix} \det(A) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \det(A) \end{bmatrix} = \det(A)I_n$$

Con un argumento semejante se demuestra que $A\,\mathrm{Adj}(A) = \det(A)I_n$. ☐

TEOREMA 11

(Inversa a través de la adjunta)

Sea A una matriz invertible. Entonces

$$A^{-1} = \frac{1}{\det(A)}\mathrm{Adj}(A)$$

DEMOSTRACIÓN Si multiplicamos $A\,\mathrm{Adj}(a) = \det(A)I_n$ del teorema 10, por la izquierda de A^{-1}, se obtiene

$$\mathrm{Adj}(A) = A^{-1}\det(A)$$

Como A es invertible, $\det(A) \neq 0$ según el teorema 3. Por consiguiente,

$$\frac{1}{\det(A)}\mathrm{Adj}(A) = A^{-1}$$ ☐

■ EJEMPLO 18 Sea A la matriz del ejemplo 17. Calcule A^{-1} aplicando el teorema 11.

SOLUCIÓN $\det(A) = 17$. Por consiguiente,

$$A^{-1} = \frac{1}{\det(A)}\text{Adj}(A) = \frac{1}{17}\begin{bmatrix} 9 & -6 & -10 \\ -2 & 7 & 6 \\ 15 & -10 & -11 \end{bmatrix} = \begin{bmatrix} \frac{9}{17} & -\frac{6}{17} & -\frac{10}{17} \\ -\frac{2}{17} & \frac{7}{17} & \frac{6}{17} \\ \frac{15}{17} & -\frac{10}{17} & -\frac{11}{17} \end{bmatrix}$$

■ EJEMPLO 19 Sea

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Entonces

$$\det(A) = ad - bc \quad \text{y} \quad \begin{matrix} C_{11} = d & C_{12} = -c \\ C_{21} = -b & C_{22} = a \end{matrix}$$

Por lo tanto,

$\text{Adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$; por tanto, $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ así, de acuerdo con el teorema 11.

Así hemos obtenido la fórmula familiar de la inversa para una matriz general 2×2.

NOTA Para calcular Adj(A) es preciso calcular n^2 determinantes de tamaño $(n-1) \times (n-1)$. Si $n = 10$, necesitaríamos 100 determinantes 9×9. Debido a esta intensidad de cómputo casi no se usa el teorema 11 para el cálculo de A^{-1}. La reducción de $[A : I]$ *sigue siendo el método a elegir*. Sin embargo, el teorema puede ser muy útil para demostrar propiedades teóricas de las inversas. También podemos emplearlo en el tema siguiente, para comprobar una fórmula de la solución de un sistema lineal cuadrado.

Regla de Cramer

Sea $A\mathbf{x} = \mathbf{b}$ un sistema cuadrado, donde

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Sea A_i la matriz obtenida de A reemplazando la i-ésima columna con $\mathbf{b}$.

$$A_i = \begin{bmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{bmatrix}$$

La regla de Cramer es una fórmula explícita para resolver cualquier sistema cuadrado consistente.

TEOREMA 12

(Regla de Cramer)

Si $\det(A) \neq 0$, el sistema $A\mathbf{x} = \mathbf{b}$ presenta una solución única, que se expresa como sigue:

$$x_1 = \frac{\det(A_1)}{\det(A)}, x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, x_n = \frac{\det(A_n)}{\det(A)}$$

DEMOSTRACIÓN Como el $\det(A) \neq 0$, de acuerdo con el teorema 3, A es invertible. Por consiguiente, $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$. Pero A^{-1} se puede calcular con el teorema 11,

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det(A)}\operatorname{Adj}(A)\mathbf{b} = \frac{1}{\det(A)}\begin{bmatrix} C_{11}b_1 + C_{21}b_2 + \cdots + C_{n1}b_n \\ \vdots \\ C_{1n}b_1 + C_{2n}b_2 + \cdots + C_{nn}b_n \end{bmatrix}$$

Así, el i-ésimo componente x_i de $\mathbf{x}$ es igual al i-ésimo componente del lado derecho.

$$x_i = \frac{1}{\det(A)}(C_{1i}b_1 + C_{2i}b_2 + \cdots + C_{ni}b_n)$$

Como A y A_i difieren sólo en la i-ésima columna, los cofactores de esa columna son iguales. En consecuencia, $C_{1i}b_1 + C_{2i}b_2 + \cdots + C_{ni}b_n$ es $\det(A_i)$, por el desarrollo en cofactores respecto a su i-ésima columna. Por tanto,

$$x_i = \frac{\det(A_i)}{\det(A)}$$

para $i = 1, \ldots, n$, como se afirmó. □

■ EJEMPLO 20 Aplique la regla de Cramer para determinar la solución del sistema

$$\begin{aligned} x + y - z &= 2 \\ x - y + z &= 3 \\ -x + y + z &= 4 \end{aligned}$$

SOLUCIÓN Primero se calcula el determinante de la matriz de coeficientes

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

y también los determinantes de las matrices

$$A_1 = \begin{bmatrix} 2 & 1 & -1 \\ 3 & -1 & 1 \\ 4 & 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 3 & 1 \\ -1 & 4 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 3 \\ -1 & 1 & 4 \end{bmatrix}$$

Obteniendo

$$\det(A) = -4 \qquad \det(A_1) = -10 \qquad \det(A_2) = -12 \qquad \det(A_3) = -14$$

con lo cual se llega a

$$x = \frac{\det(A_1)}{\det(A)} = \frac{5}{2}, \qquad y = \frac{\det(A_2)}{\det(A)} = 3, \qquad z = \frac{\det(A_3)}{\det(A)} = \frac{7}{2} \qquad \square$$

■ EJEMPLO 21 Suponga que $ad - bc \neq 0$. Utilice la regla de Cramer para determinar la solución del sistema lineal general

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

SOLUCIÓN

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad A_1 = \begin{bmatrix} e & b \\ f & d \end{bmatrix}, \qquad A_2 = \begin{bmatrix} a & e \\ c & f \end{bmatrix}$$

$$\det(A) = ad - bc \qquad \det(A_1) = de - bf \qquad \det(A_2) = af - ce$$

Como $\det(A) = ad - bc \neq 0$, podemos aplicar la regla de Cramer para obtener

$$x = \frac{de - bf}{ad - bc}, \qquad y = \frac{af - ce}{ad - bc} \qquad \square$$

Ejercicios 6.3

En los ejercicios 1 a 6 use el teorema 11 para calcular las inversas de las matrices dadas.

1. a. $\begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}$ b. $\begin{bmatrix} -1 & 2 \\ -3 & 4 \end{bmatrix}$

2. a. $\begin{bmatrix} -3 & 4 \\ 8 & -11 \end{bmatrix}$ b. $\begin{bmatrix} 2.2 & 1.0 \\ 3.2 & 1.5 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix}$

7. Sea

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 2 & 1 & 1 \end{bmatrix}$$

   a. Determine A^{-1} reduciendo con operaciones de renglón.
   b. Calcule A^{-1} aplicando el teorema 11.
   c. ¿Cuál método es más eficiente?

8. Sea A una matriz cuadrada con elementos *enteros*. Demuestre lo siguiente.
   a. $\det(A)$ es un entero.
   b. $\mathrm{Adj}(A)$ tiene elementos enteros.
   c. Si el $\det(A)$ es divisor exacto de todos los elementos de la $\mathrm{Adj}(A)$, entonces A^{-1} tiene elementos enteros.
   d. Si el $\det(A) = \pm 1$, entonces A^{-1} tiene elementos enteros.

En los ejercicios 9 a 11 aplique la regla de Cramer para resolver los sistemas.

9. a. $\begin{aligned} x + y &= 1 \\ x - y &= 1 \end{aligned}$ b. $\begin{aligned} x + 2z &= 1 \\ 3x + 4z &= 1 \end{aligned}$

10. $x + y + z = 1$
$x - y + z = 1$
$x + y - z = 1$

11. $x + 2y + 3z = 1$
$y + 4z = 1$
$z = 1$

12. Aplique la regla de Cramer para despejar x y y del sistema siguiente.

$$(\cos\theta)x - (\operatorname{sen}\theta)y = \cos\theta - 2\operatorname{sen}\theta$$
$$(\operatorname{sen}\theta)x + (\cos\theta)y = \operatorname{sen}\theta + \cos\theta$$

En los ejercicios 13 y 14 use la regla de Cramer para despejar z.

13. $3x - 3y - 2z = 3$
$-x - 4y + 2z = 2$
$5x + 4y + z = 1$

14. $x - y - z - w = 0$
$-x - y + z + w = 2$
$x + y - z + w = 1$
$x + y + z + w = 1$

15. Sea A cualquier matriz $n \times n$. Demuestre que

$$\det(\operatorname{Adj}(A)) = (\det(A))^{n-1}$$

(*Sugerencia:* Use la ecuación $A \operatorname{Adj}(A) = \det(A)I_n$ de la demostración del teorema 11.)

16. Sea A 4×4 con $\det(A) = 3$. Calcule al $\det(\operatorname{Adj}(A))$.

17. Compruebe que una matriz A $n \times n$ es invertible si y sólo si $\operatorname{Adj}(A)$ es invertible. (*Sugerencia:* Use el resultado del ejercicio 15.)

6.4 Determinantes con permutaciones

Objetivos del estudiante para esta sección

1. Conocer qué es una permutación y aprender a determinar su signo.
2. Comprender la relación entre permutaciones y determinantes.
3. Usar el desarrollo completo para calcular un determinante.

En la sección 6.1 aprendimos que

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Éstos son ejemplos de lo que se llama **desarrollo completo**; dicha fórmula expresa el determinante en términos de los *elementos* de la matriz.[7] Para una matriz A $n \times n$, esta fórmula se obtiene como sigue:

1. Se forman todos los productos, donde cada uno consiste en n elementos de A que provienen de *distintos* renglones y columnas.
2. Se asigna un signo a cada producto. Se suman todos los productos con signo.

Todos los términos de los dos determinantes que acabamos de ver se apegan al paso 1. Por ejemplo, $-a_{13}a_{22}a_{31}$ en el determinante 3×3 es el producto cuyo signo se formó tomando a_{13}, a_{22} y a_{31} del primer renglón y la tercera columna, del segundo renglón y la segunda co-

[7] Contrasta con el desarrollo por cofactores, que calcula un determinante en términos de *determinantes* de menor tamaño.

lumna y del tercer renglón y la primera columna, respectivamente. Para calcular todos esos productos, llamados **productos elementales**, se recurre al concepto de la permutación.

DEFINICIÓN

> Una **permutación** del conjunto de enteros $\{1, 2, \ldots, n\}$ es un rearreglo de estos números.

■ EJEMPLO 22 Las permutaciones de $\{1, 2\}$, $\{1, 2, 3\}$ y $\{1, 2, 3, 4\}$ son

$\{1, 2\}$:

$$(1,2) \qquad (2,1)$$

$\{1, 2, 3\}$:

$$\begin{array}{ccc} (1,2,3) & (2,1,3) & (3,1,2) \\ (1,3,2) & (2,3,1) & (3,2,1) \end{array}$$

$\{1, 2, 3, 4\}$:

$$\begin{array}{cccccc} (1,2,3,4) & (1,4,2,3) & (2,3,1,4) & (3,1,2,4) & (3,4,1,2) & (4,2,1,3) \\ (1,2,4,3) & (1,4,3,2) & (2,3,4,1) & (3,1,4,2) & (3,4,2,1) & (4,2,3,1) \\ (1,3,2,4) & (2,1,3,4) & (2,4,1,3) & (3,2,1,4) & (4,1,2,3) & (4,3,1,2) \\ (1,3,4,2) & (2,1,4,3) & (2,4,3,1) & (3,2,4,1) & (4,1,3,2) & (4,3,2,1) \end{array}$$

□

Son $6 = 1 \cdot 2 \cdot 3 = 3!$ permutaciones de $\{1, 2, 3\}$ y $24 = 4!$ permutaciones de $\{1, 2, 3, 4\}$. En general, la cantidad de permutaciones de $\{1, 2, \ldots, n\}$ es $n!$ Esto puede visualizarse como sigue: para llenar la primera posición hay n alternativas, puesto que podemos usar cada uno de los números. Para la segunda posición existen $n - 1$ alternativas, porque un número ya se utilizó en la primera. Así, para llenar las dos primeras posiciones, hay $n \cdot (n - 1)$ alternativas. Se continúa en la misma forma con el resto de las posiciones para obtener un total de $n \cdot (n - 1) \cdots 2 \cdot 1 = n!$ alternativas. Observe que $n!$ crece con mucha rapidez en función de n.

$$\begin{array}{ll} 0! = 1 & 6! = 720 \\ 1! = 1 & 7! = 5040 \\ 2! = 2 & 8! = 40\,320 \\ 3! = 6 & 9! = 362\,880 \\ 4! = 24 & 10! = 3\,628\,800 \\ 5! = 120 & 11! = 39\,916\,800 \end{array}$$

Por lo anterior, es difícil escribir todas las permutaciones, aun para n pequeño.

DEFINICIÓN

> Sea $p = (j_1, j_2, \ldots, j_n)$ cualquier permutación de $\{1, 2, \ldots, n\}$. Se dice que p tiene una **inversión** (j_i, j_k) si un entero mayor, j_i, preceda a uno menor, j_k. También, p es **par** o **impar** si tiene una cantidad total par o impar de inversiones. El **signo** de p. indicado por $\text{sign}(p)$ es 1 si p es par y -1 si es impar. Se considera que $(1, 2, \ldots, n)$ es par con signo 1.

Por ejemplo, $(1, 3, 2, 4)$ tiene una inversión: $(3, 2)$. De modo que es impar con signo -1. La permutación $(4, 2, 1, 3)$ tiene cuatro inversiones: $(4, 2)$, $(4, 1)$, $(4, 3)$ y $(2, 1)$; por tanto, es par con signo 1.

■ EJEMPLO 23

Permutación	Inversiones	Par/impar	Signo
(1, 2, 3)	Ninguna	Par	1
(1, 3, 2)	(3, 2)	Impar	−1
(2, 1, 3)	(2, 1)	Impar	−1
(2, 3, 1)	(2, 1), (3, 1)	Par	1
(3, 1, 2)	(3, 1), (3, 2)	Par	1
(3, 2, 1)	(3, 2), (3, 1), (2, 1)	Impar	−1

NOTA Si r es la cantidad de inversiones de una permutación p, entonces $\operatorname{sign}(p) = (-1)^r$.

Veamos ahora cómo intervienen las permutaciones en el cálculo de los determinantes.[8] Para el determinante 2×2 y obtener todos los productos elementales formamos

$$a_{1\_}a_{2\_}$$

con cada número de renglón representado en los primeros subíndices. Los espacios en blanco en los segundos subíndices deben llenarse con todas las alternativas de números de columna. En este caso son ya sea 1, 2, o bien 2, 1. Son justamente las permutaciones de $\{1, 2\}$. Al continuar con este proceso se asegura que no haya dos elementos que provengan del mismo renglón o columna. Además, el signo de cada producto elemental es el de la permutación de los índices de columna. El signo de $a_{11}a_{12}$ es 1, porque (1, 2) es par, mientras que el signo de $a_{12}a_{21}$ es −1, porque (2, 1) es impar.

De manera semejante, en el determinante 3×3 todos los productos elementales se obtienen partiendo de

$$a_{1\_}\,a_{2\_}\,a_{3\_}$$

y los espacios se llenan con las permutaciones de $\{1, 2, 3\}$. Por ejemplo, $a_{13}a_{21}a_{32}$ corresponde a la permutación (3, 1, 2). El signo de cada término es el signo de la permutación correspondiente. El signo de $a_{13}a_{21}a_{32}$ es 1, porque (3, 1, 2) es par. El signo de $a_{12}a_{21}a_{33}$ es −1, puesto que (2, 1, 3) es impar. En general, podemos **definir** un determinante $n \times n$ usando a las permutaciones de la manera siguiente:

DEFINICIÓN

Si A es una matriz $n \times n$ con elementos a_{ij}, el **desarrollo completo** del determinante de A es

$$\det(A) = \sum \operatorname{sign}(j_1, j_2, \ldots, j_n)a_{1j_1}a_{2j_2}\cdots a_{nj_n} \tag{6.1}$$

donde la suma es sobre todas las permutaciones $(j_1, j_2, \ldots, j_n)$ de $(1, 2, \ldots, n)$ y el signo de $(j_1, j_2, \ldots, j_n)$.

OBSERVACIONES Esta definición nos permite demostrar *todas las propiedades de los determinantes, incluyendo el desarrollo por cofactores y el teorema 1* (véase el conjunto de ejercicios). También es muy útil para demostrar las propiedades teóricas de los determinantes. Sin

[8] Este método de estudiar los determinantes se debe a Bezout y a Laplace.

embargo, cuando éstos se calculan no es muy práctico aplicar la ecuación (6.1). Por ejemplo, para un determinante de 11×11 se necesitarían 39 916 800 términos, cada uno formado por 11 factores. Para ello se requiere un total de 439 084 800 operaciones, siempre y cuando se cuente con un buen método para escribir todos los productos elementales posibles. Con una rapidez de 10 000 operaciones por minuto se necesitarían 30 años de cálculos. Ya nos hemos encontrado con un método mucho más práctico: el desarrollo por cofactores precedido por la (aún presente) eliminación de Gauss. El uso de la ecuación (6.1) sólo es práctico en matrices con pocos productos elementales distintos de cero.

■ EJEMPLO 24 Calcule el determinante

$$\begin{vmatrix} 3 & 0 & 4 \\ 0 & 5 & 0 \\ 6 & 0 & 7 \end{vmatrix}$$

SOLUCIÓN Como cada producto elemental no debe tener factores de la misma columna o el mismo renglón, los únicos términos distintos de cero son

$$3 \cdot 5 \cdot 7 \text{ y } 4 \cdot 5 \cdot 6$$

que corresponden a

$$a_{11}a_{22}a_{33} \text{ y } a_{13}a_{22}a_{31}$$

y en consecuencia a las permutaciones

$$(1, 2, 3) \quad \text{y} \quad (3, 2, 1)$$

respectivamente. La primera permutación es par y la segunda impar. En consecuencia, los signos son 1 y -1, respectivamente, y el determinante es

$$3 \cdot 5 \cdot 7 - 4 \cdot 5 \cdot 6 = 105 - 120 = -15$$

Concluiremos esta sección describiendo una forma interesante para determinar el signo de una permutación. Esto se hace empleando diagramas de cruce. (Sin embargo, en la sección de ejercicios se presenta otra forma, el método de *matrices de permutación*.)

Diagramas de Cruce (Opcional)

DEFINICIÓN

El **diagrama de cruce** de una permutación $(j_1, j_2, \ldots, j_n)$ está formado por dos columnas de vértices, cada columna con los elementos $1, 2, \ldots, n$, en el cual se ha trazado un segmento de recta (o de curva) desde cada nodo, por ejemplo i, a la i-ésima coordenada j_i de la permutación. El **número de cruce** de este diagrama su cantidad de cruces. Dicho número es impar si la permutación es impar, y par si la permutación es par. El caso de 0 cruces corresponde a $(1, 2, \ldots, n)$, que es una permutación par.

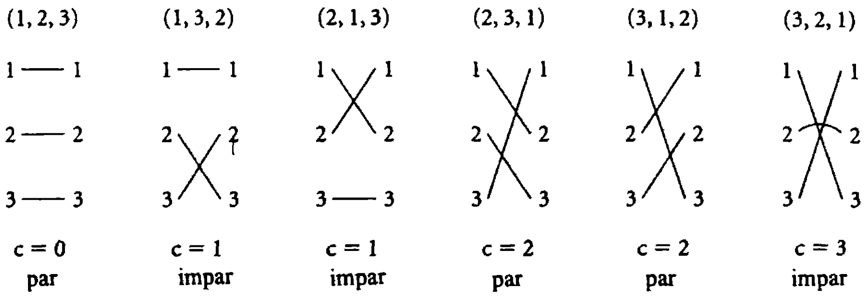


Figura 6.1 Diagramas de cruzamiento de las permutaciones de $\{1, 2, 3\}$.

Consideremos la permutación (3, 1, 2). Escribimos dos columnas, ambas formadas por 1, 2, 3. Como $j_1 = 3$, $j_2 = 1$, $j_3 = 2$, trazamos una línea que conecte el 1 de la primera columna con el 3 de la segunda; otra que conecte el 2 de la primera con el 1 de la segunda y una tercera línea que una al 3 de la primera columna con el 2 de la segunda (figura. 6.1). La cantidad de cruces en el diagrama es 2, así que el número de cruce c es 2, que es número par. Por consiguiente (3, 1, 2) es una permutación par. En la figura. 6.1 se muestran los diagramas de cruce de todas las permutaciones de $\{1, 2, 3\}$.

■ EJEMPLO 25 Aplique la ecuación (6.1) y los diagramas de cruce para calcular el determinante

$$\begin{vmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 3 & 0 & 4 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 6 & 0 & 7 & 0 \\ 8 & 0 & 0 & 0 & 9 \end{vmatrix}$$

SOLUCIÓN Un producto elemental distinto de cero puede tener un factor de 1 o 2 del primer renglón. Si comienza con 1, el último factor es 9 y no 8, porque 1 y 8 están en la misma columna. Igualmente, si un producto comienza con 2, el último factor es 8. Por lo anterior, sólo tenemos productos de la forma

$$1 \cdot \_ \cdot \_ \cdot \_ \cdot 9 \quad \text{y} \quad 2 \cdot \_ \cdot \_ \cdot \_ \cdot 8$$

El resto de los factores proviene de la submatriz

$$\begin{bmatrix} 3 & 0 & 4 \\ 0 & 5 & 0 \\ 6 & 0 & 7 \end{bmatrix}$$

que se usó en el ejemplo 24. Los productos posibles en este caso son $3 \cdot 5 \cdot 7$ y $4 \cdot 5 \cdot 6$. Por consiguiente, se obtiene un total de cuatro productos distintos de cero, que son $1 \cdot 3 \cdot 5 \cdot 7 \cdot 9$, $1 \cdot 4 \cdot 5 \cdot 6 \cdot 9$, $2 \cdot 3 \cdot 5 \cdot 7 \cdot 8$ y $2 \cdot 4 \cdot 5 \cdot 6 \cdot 8$. Los signos de las permutaciones correspondientes se determinan con los diagramas de cruce de la figura. 6.2.

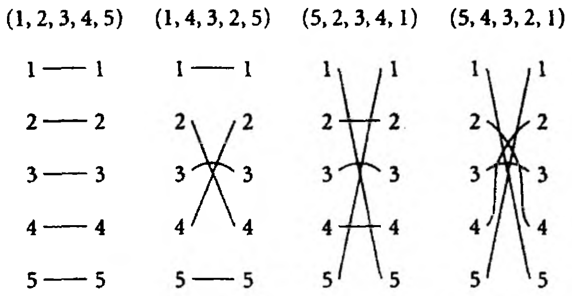


Figura 6.2 Diagramas de cruzamiento de algunas permutaciones de {1, 2, 3, 4, 5}.

Llegamos a lo siguiente:

Producto de elementos	Permutación	Par/impar	Signo	Valor
$1 \cdot 3 \cdot 5 \cdot 7 \cdot 9 = 945$	(1, 2, 3, 4, 5)	Par	+	945
$1 \cdot 4 \cdot 5 \cdot 6 \cdot 9 = 1080$	(1, 4, 3, 2, 5)	Impar	–	–1080
$2 \cdot 3 \cdot 5 \cdot 7 \cdot 8 = 1680$	(5, 2, 3, 4, 1)	Impar	–	–1680
$2 \cdot 4 \cdot 5 \cdot 6 \cdot 8 = 1920$	(5, 4, 3, 2, 1)	Par	+	1920

El determinante es $945 - 1080 - 1680 + 1920 = 105$. □

Permutaciones con sistemas algebraicos computacionales

Maple

```
# Permutations of {1,2,3}
> with(combinat) :
> permute(3);
              [[1,2,3],[1,3,2],[2,1,3],[2,3,1],[3,1,2],[3,2,1]]
```

Mathematica

```
(* Permutations of {1,2,3}*)
In[1]:=
      Permutations[{1,2,3}]
Out[1]= {{1, 2, 3}, {1, 3, 2}, {2, 1, 3}, {2, 3, 1},
        {3, 1, 2}, {3, 2, 1}}
```

MATLAB

```
% A random permutation
>> randperm(3)
ans =
     2     1     3
```

Ejercicios 6.4

En los ejercicios 1 a 3 determine el signo y clasifique las permutaciones en par e impar.

1. (2, 1, 3, 4), (1, 4, 2, 3), (1, 5, 2, 4, 3), (1, 4, 3, 5, 2)
2. (2, 3, 1, 4), (3, 1, 4, 2), (2, 3, 4, 1, 5), (5, 3, 1, 4, 2)
3. (3, 1, 4, 2), (4, 2, 1, 3), (3, 4, 2, 1, 5), (4, 2, 5, 1, 3)

En los ejercicios 4 a 9 calcule los determinantes de las matrices empleando desarrollo completo.

4. a. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ b. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & 4 & 0 \end{bmatrix}$

5. a. $\begin{bmatrix} 0 & 2 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ b. $\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 4 & 0 & 0 \end{bmatrix}$

6. a. $\begin{bmatrix} 0 & 0 & 2 \\ 3 & 0 & 0 \\ 0 & 4 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 0 & 0 & 2 \\ 0 & 3 & 0 \\ 4 & 0 & 0 \end{bmatrix}$

7. a. $\begin{bmatrix} 2 & 0 & 5 \\ 0 & 3 & 0 \\ 6 & 0 & 4 \end{bmatrix}$ b. $\begin{bmatrix} -2 & 0 & 5 \\ 0 & 3 & 0 \\ -6 & 0 & -4 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 2 & 6 & 0 \\ 0 & 7 & 3 & 0 \\ 8 & 0 & 0 & 4 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 5 & 0 & 0 \end{bmatrix}$

10. Determine los signos de las siguientes permutaciones usando diagramas de cruce.

$$(3, 1, 2, 4) \qquad (4, 2, 3, 1)$$
$$(1, 4, 2, 3, 5) \qquad (5, 1, 2, 4, 3)$$

Matrices de permutación

Una *matriz de permutación* es una matriz cuadrada formada por unos y ceros de tal forma que hay exactamente un uno en cada renglón y en cada columna (véase también la sección 3.4). Las matrices A, B, C y D son matrices de permutación.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Una matriz de permutación da lugar a una, y sólo una permutación, como sigue: para cada renglón de la matriz escriba el número de columna del elemento que tiene 1. Todos esos números forman los elementos de la permutación correspondiente. Por ejemplo, las permutaciones que corresponden a A, B, C y D son (1, 2, 3), (1, 3, 2), (1, 3, 2, 4) y (2, 1, 4, 3, 5), respectivamente.

11. Encuentre todas las matrices de permutación de 3×3. Para cada una de ellas escriba su permutación correspondiente.

12. Escriba las matrices de permutación que corresponden a las siguientes permutaciones:

$$(4, 1, 2, 3) \qquad (3, 2, 4, 1)$$
$$(4, 2, 3, 1, 5) \qquad (3, 5, 4, 2, 1)$$

13. Demuestre que el signo de una permutación es igual al determinante de la matriz de permutación correspondiente.

14. Determine los signos de las permutaciones siguientes, calculando el determinante de la matriz de permutación correspondiente.

$$p = (1, 4, 2, 3) \qquad q = (3, 2, 1, 4)$$
$$s = (2, 1, 3, 4, 5) \qquad t = (1, 5, 4, 2, 3)$$

En los dos apartados siguientes describiremos las demostraciones del teorema 1 y del desarrollo por cofactores usando permutaciones. Para estas partes, A y B son dos matrices $n \times n$ cuyos elementos respectivos son a_{ij} y b_{ij}.

Demostración del Teorema 1

15. Si B se obtiene de A multiplicando uno de sus renglones por una constante k distinta de cero, demuestre que el $\det(B) = k\det(A)$. (*Sugerencia:* Si $kR_i \to R_i$ produce B a partir de A, entonces $\det(B) = \Sigma \pm a_{1j_1} \ldots (ka_{ij_i}) \ldots a_{nj_n}$. Saque a k como factor común.)
16. Demuestre que si se intercambian dos elementos *consecutivos* cualesquiera en una permutación, la cantidad de inversiones aumenta o disminuye en 1.
17. Use el resultado del ejercicio 16 para demostrar que si se intercambian dos elementos *cualesquiera* en una permutación, la cantidad de inversiones cambia en un entero *impar*.
18. Con el resultado del ejercicio 17 demuestre que si se intercambian dos elementos *cualesquiera* en una permutación, las permutaciones nueva y anterior tienen signos opuestos.
19. Si B se obtiene de A al intercambiar dos renglones cualesquiera, demuestre que $\det(B) = -\det(A)$. (*Sugerencia:* Si $R_i \leftrightarrow R_l$ $(i < l)$ produce B partiendo de A, entonces $\det(B) = \Sigma \pm a_{1j_1} \ldots a_{lj_i} \ldots a_{ij_l} \ldots a_{nj_n}$. Use el resultado del ejercicio 18.)
20. Si dos renglones de A son iguales, compruebe que $\det(A) = 0$. (*Sugerencia:* Intercambie los dos renglones similares y aplique el resultado del ejercicio 19.)
21. Si B se obtiene de A sumando un múltiplo de un renglón a otro renglón, entonces $\det(B) = \det(A)$. (*Sugerencia:* Si $R_i + kR_l \to R_i$ produce B partiendo de A, en ese caso $\det(B) = \Sigma \pm a_{1j_1} \ldots (a_{ij_i} + ka_{lj_i}) \ldots a_{nj_n}$. Sepárelo en dos sumas. La segunda suma es cero, según el resultado del ejercicio 20, después de sacar a k como factor común.)
22. La matriz A y su transpuesta tienen el mismo determinante, $\det(A) = \det(A^T)$. (*Sugerencia:* $\det(A^T) = \Sigma \pm a_{j_1 1} \ldots a_{j_n n}$. Reordene $a_{j_1 1} \ldots a_{j_n n}$ en la forma $a_{1l_1} \ldots a_{nl_n}$ y compare los signos de $(j_1, \ldots, j_n)$ y de $(l_1, \ldots, l_n)$.
23. Demuestre el teorema 1.

Demostración del desarrollo por cofactores.

Recuerde que los términos de la suma

$$\det(A) = \sum \pm a_{1j_1} \ldots a_{ij_i} \ldots a_{nj_n}$$

contienen sólo un elemento de cada renglón y de cada columna. Por consiguiente, el factor a_{i1} del i-ésimo renglón se presenta exactamente en $(n-1)!$ términos, mientras que a_{i2} podemos encontrarlo en $(n-1)!$ términos, diferentes de los primeros, y por último, a_{in} se presenta en $(n-1)!$ términos, distintos de los anteriores. Como la suma de todos esos términos es igual al $\det(A)$, puede escribirse

$$\det(A) = a_{i1}D_{i1} + a_{i2}D_{i2} + \cdots + a_{in}D_{in}$$

donde D_{ij} es la suma, en el $\det(A)$, que queda después de sacar a_{ij} como factor común de todos los términos que lo contienen. En los dos ejercicios siguientes demostraremos que $D_{ij} = C_{ij}$, el (i, j) cofactor de A, y también el desarrollo de $\det(A)$ por cofactores respecto al i-ésimo renglón.

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \tag{6.2}$$

NOTACIÓN Representaremos con $A(i, j)$ a la matriz que se obtiene de A omitiendo el i-ésimo renglón y la j-ésima columna.

24. Demuestre que $D_{11} = C_{11}$. (*Sugerencia:* $D_{11} = \Sigma \pm a_{2j_2} \ldots a_{nj_n}$, donde la suma esta sobre todas las permutaciones de la forma $(j_2, \ldots, j_n)$, porque $j_1 = 1$. Pero éste es el determinante de $A(1, 1)$.)
25. Demuestre que $D_{ij} = C_{ij}$. (*Sugerencia:* Considere que A' es la matriz obtenida de A mediante $i - 1$ intercambios sucesivos de renglones adyacentes y de $j - 1$ intercambios sucesivos de columnas adyacentes, con los que se lleva a a_{ij} a la posición superior izquierda, manteniendo el orden relativo de los demás elementos. Entonces, el $\det(A) = (-1)^{i+j}\det(A')$, a través de varias aplicaciones de la propiedad 3 del teorema 1. Observe que $a_{ij} = a'_{11}$ y que el $\det(A(i, j)) = \det(A'(1, 1))$. Ahora aplique el resultado del ejercicio 24.
26. Desarrolle el $\det(A)$ por cofactores respecto al i-ésimo renglón. (Es la fórmula (6.2).)
27. Deduzca la fórmula para el desarrollo de $\det(A)$ por cofactores respecto a la j-ésima columna.

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

(*Sugerencia:* Use los resultados de los ejercicios 22 y 26.)

6.5 Aplicaciones

Objetivo del estudiante para esta sección
Conocer la extensa aplicación de los determinantes.

Existen numerosas aplicaciones de los determinantes en ingeniería, física, matemáticas y otras ciencias. En esta sección describiremos unas cuantas. El fundamento de todas ellas es el teorema 4.

Geometría: ecuaciones de objetos geométricos

En geometría analítica, los determinantes desempeñan un papel básico en el cálculo de áreas y volúmenes, y en la formulación de ecuaciones de objetos geométricos como rectas, círculos, elipses, parábolas, planos y esferas. En esta sección nos concentraremos en las ecuaciones de algunos objetos geométricos.

Recta que pasa por dos puntos

Podemos formular la ecuación de la recta que pasa por dos puntos distintos, cuyas coordenadas son (x_1, y_1), (x_2, y_2).

Sea $ax + by + c = 0$ la ecuación de la recta. Como ambos puntos están en ella, entonces (x_1, y_1) y (x_2, y_2) deben satisfacer esta ecuación; por tanto,

$$\begin{aligned} ax + by + c &= 0 \\ ax_1 + by_1 + c &= 0 \\ ax_2 + by_2 + c &= 0 \end{aligned}$$

Es un sistema homogéneo cuyas incógnitas son a, b y c. De acuerdo con el teorema 4, tiene una solución no trivial si y sólo si

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0 \tag{6.3}$$

Al desarrollar el determinante, esta relación es la ecuación de la recta.

■ EJEMPLO 26 Deduzca la ecuación de la recta que pasa por los puntos (1, 2), (−2, 0).

SOLUCIÓN Al sustituir los dos puntos en la ecuación (6.3) se obtiene

$$\begin{vmatrix} x & y & 1 \\ 1 & 2 & 1 \\ -2 & 0 & 1 \end{vmatrix} = 0$$

Al desarrollar el determinante se obtiene $2x - 3y + 4 = 0$. □

Tres puntos en la misma recta

Podemos demostrar que una condición necesaria y suficiente para que tres puntos cuyas coordenadas son (x_1, y_1), (x_2, y_2) y (x_3, y_3) estén en la misma recta es

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0 \tag{6.4}$$

De acuerdo con la ecuación (6.3), la recta definida por (x_2, y_2) y (x_3, y_3) se expresa mediante

$$\begin{vmatrix} x & y & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0 \tag{6.5}$$

Si el punto cuyas coordenadas son (x_1, y_1) también está en la recta, entonces al sustituir $x = x_1$ y $y = y_1$ se obtiene la ecuación (6.4).

De manera recíproca, la ecuación (6.4) implica que la ecuación de la recta (6.5) que pasa por (x_2, y_2) y (x_3, y_3), se satisface si hacemos que $x = x_1$ y $y = y_1$. Por consiguiente, el punto (x_1, y_1) está en la recta definida por (x_2, y_2) y (x_3, y_3).

Círculo que pasa por tres puntos

Como otra aplicación, deduciremos la ecuación del círculo que pasa por los puntos no colineales (x_1, y_1), (x_2, y_2) y (x_3, y_3).

Sea $(x - a)^2 + (y - b)^2 = r^2$ la ecuación del círculo de radio r con centro en (a, b). Al desarrollar esta igualdad se llega a la forma $A(x^2 + y^2) + Bx + Cy + D = 0$. (Observe que $A = 1$.) Como los puntos (x_1, y_1), (x_2, y_2), (x_3, y_3) satisfacen la ecuación, entonces

$$\begin{aligned} A(x^2 + y^2) + Bx + Cy + D &= 0 \\ A(x_1^2 + y_1^2) + Bx_1 + Cy_1 + D &= 0 \\ A(x_2^2 + y_2^2) + Bx_2 + Cy_2 + D &= 0 \\ A(x_3^2 + y_3^2) + Bx_3 + Cy_3 + D &= 0 \end{aligned}$$

Puede considerarse que éste es un sistema homogéneo de 4 ecuaciones con 4 incógnitas, A, B, C y D (una de las cuales ya se conoce, $A = 1$). De acuerdo con el teorema 4, una solución no trivial debe forzar a que el determinante de los coeficientes sea cero. En consecuencia

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0 \tag{6.6}$$

Al desarrollarlo se obtiene la ecuación del círculo.

■ EJEMPLO 27 Defina la ecuación del círculo que pasa por (1, 4), (3, 2) y (−1, 2).

SOLUCIÓN Al sustituir los tres puntos en la ecuación (6.6) se obtiene

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ 17 & 1 & 4 & 1 \\ 13 & 3 & 2 & 1 \\ 5 & -1 & 2 & 1 \end{vmatrix} = 0$$

Esta ecuación se reduce a

$$-8x^2 - 8y^2 + 32y + 16x - 8 = 0$$

Al dividir entre −8 y completar los dos cuadrados se tiene una ecuación de un círculo de radio 2 centrado en (1, 2).

$$(x - 1)^2 + (y - 2)^2 = 4$$

Plano que pasa por tres puntos

Con el uso de determinantes puede deducirse la ecuación del plano que pasa por los puntos no colineales (x_1, y_1, z_1), (x_2, y_2, z_2) y (x_3, y_3, z_3).

Sea $ax + by + cz + d = 0$ la ecuación del plano. Al sustituir los tres puntos se obtienen

$$\begin{aligned} ax + by + cz + d &= 0 \\ ax_1 + by_1 + cz_1 + d &= 0 \\ ax_2 + by_2 + cz_2 + d &= 0 \\ ax_3 + by_3 + cz_3 + d &= 0 \end{aligned}$$

Una vez más llegamos a un sistema homogéneo de cuatro ecuaciones con las cuatro incógnitas a, b, c y d. De acuerdo con el teorema 4, una solución no trivial debe hacer que el determinante de los coeficientes sea cero. Por consiguiente,

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0 \tag{6.7}$$

Esta relación se reduce a la ecuación del plano.

■ EJEMPLO 28 Deduzca la ecuación del plano que pasa por los puntos (1, 1, 7), (3, 2, 6) y (−2, −2, 4).

SOLUCIÓN Al sustituir los tres puntos en la ecuación (6.7) se obtiene

$$\begin{vmatrix} x & y & z & 1 \\ 1 & 1 & 7 & 1 \\ 3 & 2 & 6 & 1 \\ -2 & -2 & 4 & 1 \end{vmatrix} = 0$$

Esta ecuación se reduce a

$$-6x + 9y + (-3z) + 18 = 0$$

es decir, a

$$2x - 3y + z - 6 = 0$$

Parábola que pasa por tres puntos

Por último, deduciremos la ecuación de una parábola, con la forma $ay + bx^2 + cx + d = 0$, que pasa por los tres puntos no colineales (x_1, y_1), (x_2, y_2), (x_3, y_3).

Si sustituimos las coordenadas de los tres puntos se obtiene

$$\begin{aligned} ay \;+ bx^2 + cx \;+ d &= 0 \\ ay_1 + bx_1^2 + cx_1 + d &= 0 \\ ay_2 + bx_2^2 + cx_2 + d &= 0 \\ ay_3 + bx_3^2 + cx_3 + d &= 0 \end{aligned}$$

Igual que antes llegamos a un sistema homogéneo de cuatro ecuaciones con cuatro incógnitas, a, b, c y d. Según el teorema 4, una solución no trivial implica que el determinante de los coeficientes es cero. Por tanto,

$$\begin{vmatrix} y & x^2 & x & 1 \\ y_1 & x_1^2 & x_1 & 1 \\ y_2 & x_2^2 & x_2 & 1 \\ y_3 & x_3^2 & x_3 & 1 \end{vmatrix} = 0 \tag{6.8}$$

que se reduce a la ecuación de la parábola.

■ EJEMPLO 29 Determine la ecuación de la parábola de la forma $ay + bx^2 + cx + d = 0$, que pasa por los puntos cuyas coordenadas son $(1, 0)$, $(2, 1)$ y $(0, 3)$.

SOLUCIÓN Al sustituir los tres puntos en la ecuación (6.8) se obtiene

$$\begin{vmatrix} y & x^2 & x & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 4 & 2 & 1 \\ 3 & 0 & 0 & 1 \end{vmatrix} = 0$$

Por consiguiente,

$$-2y + 4x^2 - 10x + 6 = 0 \quad \Rightarrow \quad y = 2x^2 - 5x + 3$$

Álgebra: teoría de la eliminación y resultantes

En esta teoría se trata de eliminar cierta cantidad de incógnitas de un sistema de ecuaciones polinomiales con una o más variables, para determinar las soluciones comunes de esas ecuaciones.[9]

La teoría de la eliminación floreció a fines del siglo XIX y principios del XX, con Sylvester, Cayley, Dixon y Macaulay; entre los precursores se incluyen Euler y Bezout. Aunque es una de las teorías más elegantes y útiles de las matemáticas, pasó de moda durante casi todo el siglo XX.[10] En la actualidad hay un interés cada vez más renovado, debido parcialmente a pro-

---

[9] El caso especial en el que los polinomios son lineales pertenece, como hemos visto, al dominio del álgebra lineal.

[10] Véase el trabajo de S. S. Abyankar, de 1976, titulado "Historical Ramblings in Algebraic Geometry and Related Algebra," *American Mathematical Monthly*, 83 (6): 409–448.

gramas simbólicos como Maple y Mathematica, y a sus requerimientos de resolver en forma eficiente las ecuaciones polinomiales. Todavía más importante es que, por la existencia de esos paquetes, la exploración y el mayor desarrollo de esta teoría es más factible que nunca.

Supongamos que hay un sistema de dos ecuaciones polinomiales[11] en una variable, x. Si nos interesaran las soluciones comunes, un método burdo consistiría en resolver por separado cada ecuación (por lo general una tarea difícil) para después determinar las raíces comunes de los conjuntos solución. Si sólo nos interesa la *existencia* de soluciones, este método es impráctico. Lo mismo ocurre en caso de resolver un sistema de dos ecuaciones polinomiales con dos variables, x y y.

Sea

$$\begin{aligned} a_1x^2 + b_1x + c_1 &= 0 \\ a_2x^2 + b_2x + c_2 &= 0 \end{aligned}$$

un sistema de dos polinomios cuadráticos, p_1 y p_2, en x. Ahora determinaremos una condición necesaria y suficiente para la existencia de una solución común. Si p_1 y p_2 tienen una solución común, entonces también deben compartir un factor lineal que lo sea, por ejemplo Q. Sean $q_1 = p_1/Q$ y $q_2 = p_2/Q$ los dos cocientes lineales, y sea $q_1 = A_1x + B_1$, y $q_2 = -A_2x - B_2$ (en seguida veremos que tiene sentido la peculiar elección de los signos). En ese caso, $Q = p_1/q_1 = p_2/q_2$; por consiguiente, $p_1q_2 = p_2q_1$. En forma explícita:

$$(a_1x^2 + b_1x + c_1)(-A_2x - B_2) = (a_2x^2 + b_2x + c_2)(A_1x + B_1)$$

Al desarrollar y agrupar los términos con potencias iguales de x se obtiene

$$\begin{aligned} &(a_1A_2 + a_2A_1)x^3 + (b_2A_1 + b_1A_2 + a_1B_2 + a_2B_1)x^2 \\ &\qquad + (c_1A_2 + c_2A_1 + b_2B_1 + b_1B_2)x + c_1B_2 + c_2B_1 = 0 \end{aligned}$$

Como esta ecuación polinomial es válida para toda x, los coeficientes de x^3, x^2, x^1 y x^0 deben ser cero. Por tanto,

$$\begin{aligned} a_1A_2 + a_2A_1 &= 0 \\ b_2A_1 + b_1A_2 + a_1B_2 + a_2B_1 &= 0 \\ c_1A_2 + c_2A_1 + b_2B_1 + b_1B_2 &= 0 \\ c_1B_2 + c_2B_1 &= 0 \end{aligned}$$

Éste es un sistema homogéneo en A_2, B_2, A_1 y B_1. De acuerdo con el teorema 4, tiene soluciones no triviales si y sólo si el determinante de los coeficientes es cero. Así,

$$\begin{vmatrix} a_1 & 0 & a_2 & 0 \\ b_1 & a_1 & b_2 & a_2 \\ c_1 & b_1 & c_2 & b_2 \\ 0 & c_1 & 0 & c_2 \end{vmatrix} = 0$$

[11] En esta sección escribiremos los polinomios en orden descendente de potencias de x, para que las fórmulas finales coincidan con las que aparecen en otros libros.

La mayoría de los autores usan el determinante de la transpuesta de esta matriz. En vista de ello, una condición necesaria y suficiente para la existencia de una raíz común es que

$$\begin{vmatrix} a_1 & b_1 & c_1 & 0 \\ 0 & a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 & 0 \\ 0 & a_2 & b_2 & c_2 \end{vmatrix} = 0$$

A este determinante se le llama **(determinante) resultante de Sylvester** [12] de p_1 y p_2. Su tamaño es 4 y consiste en los coeficientes de los dos polinomios rellenados de ceros. En general, el resultante de Sylvester de dos polinomios de grados m y n tiene tamaño $m + n$. Por ejemplo,

$$a_1x^2 + b_1x + c_1 = 0$$
$$a_2x^3 + b_2x^2 + c_2x + d_2 = 0$$

Entonces el resultante de Sylvester de este sistema tiene el tamaño 2 + 3 = 5, y se expresa como sigue:

$$\begin{vmatrix} a_1 & b_1 & c_1 & 0 & 0 \\ 0 & a_1 & b_1 & c_1 & 0 \\ 0 & 0 & a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 & d_2 & 0 \\ 0 & a_2 & b_2 & c_2 & d_2 \end{vmatrix}$$

Este determinante es cero si y sólo si el sistema tiene una solución común.

TEOREMA 13

(Anulación del resultante de Sylvester)

Sean f y g dos polinomios en x. El sistema $f = 0$, $g = 0$ tiene una solución si y sólo si el resultante de Sylvester es cero, suponiendo que uno de los coeficientes de la mayor potencia de x no sea cero.

■ EJEMPLO 30 Sin resolver las ecuaciones, muestre que el sistema siguiente tiene una solución

$$x^2 - 5x + 6 = 0$$
$$x^2 + 2x - 8 = 0$$

SOLUCIÓN El resultante de Sylvester del sistema es

$$\begin{vmatrix} 1 & -5 & 6 & 0 \\ 0 & 1 & -5 & 6 \\ 1 & 2 & -8 & 0 \\ 0 & 1 & 2 & -8 \end{vmatrix} = 0$$

Por consiguiente, el sistema tiene una raíz común, de acuerdo con el teorema 13. □

[12] Originalmente debido a Euler, la formulación actual se debe a Sylvester.

NOTA El ejemplo 30 sólo sirvió para ilustrar el método del resultante de Sylvester, y no como una demostración de su poder. Ciertamente podríamos resolver cada una de las ecuaciones cuadráticas para ver que la raíz común es 2. Sin embargo, lo que hace potente al método es que es muy general, y que casi no impone restricciones a los coeficientes de los polinomios. Los coeficientes mismos pueden ser polinomios en otra variable. Esta observación puede aprovecharse para resolver sistemas polinomiales multivariados.

■ EJEMPLO 31 Resuelva el sistema

$$x^2 + y^2 - 1 = 0$$
$$x^2 - 2x + y^2 - 2y + 1 = 0$$

SOLUCIÓN Consideremos que éste es un sistema en y con coeficientes polinomiales en x.

$$y^2 + (x^2 - 1) = 0$$
$$y^2 - 2y + (x^2 - 2x + 1) = 0$$

De acuerdo con el teorema 13, si hay una solución, entonces el resultante de Sylvester es cero.

$$\begin{vmatrix} 1 & 0 & x^2 - 1 & 0 \\ 0 & 1 & 0 & x^2 - 1 \\ 1 & -2 & x^2 - 2x + 1 & 0 \\ 0 & 1 & -2 & x^2 - 2x + 1 \end{vmatrix} = 0$$

Al desarrollar el determinante y simplificarlo se obtiene

$$8x^2 - 8x = 0$$

Por tanto, $x = 0$ y 1. Si $x = 0$, al sustituir en la primera ecuación se implica que $y = -1$, 1. Sin embargo, al sustituirla en la segunda deducimos que y sólo puede ser 1. De forma semejante, si $x = 1$, la sustitución en el sistema implica que $y = 0$. Llegamos a la conclusión que hay dos soluciones comunes: (1, 0) y (0, 1).

Observe que el sistema puede reformularse como $x^2 + y^2 = 1$ y $(x - 1)^2 + (y - 1)^2 = 1$. De este modo se tienen las ecuaciones de los círculos de radio 1 centrados en (0, 0) y en (1, 1), respectivamente. Su intersección es geométricamente obvia, y son los puntos (1, 0) y (0, 1) de la figura 6.3. □

Ingeniería eléctrica y teoría de las gráficas: árboles generadores

En 1847, en una publicación famosa, G. R. Kirchhoff estableció las bases para el estudio de los circuitos eléctricos en lo que desde entonces se llaman **leyes de Kirchhoff**. Varias de las muchas fórmulas que se presentaron ahí dependen sólo de la geometría del circuito, y no de los resistores, inductores ni fuentes de voltaje presentes. Para estudiar las propiedades geométricas, Kirchhoff reemplazó al circuito eléctrico con la gráfica inherente (figura. 6.4).

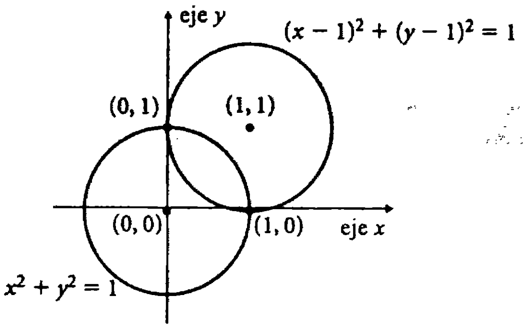


Figura 6.3 Círculos que se intersectan.

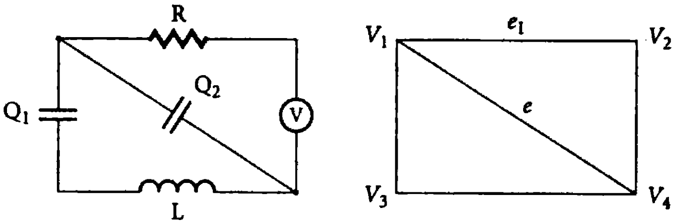


Figura 6.4 Circuito eléctrico y su correspondiente gráfica G.

Sea G una gráfica[13] con n nodos o vértices identificados por $v_1, v_2, \ldots, v_n$. Supongamos que G no tiene ramas ni lazos múltiples. A esa gráfica se le llama **simple**. Un **circuito** de G en un nodo v_j es un conjunto de ramas distintas que conectan a los nodos $v_{j1}, v_{j2}, \ldots, v_{jk}$ tales que $v_{j1} = v_j$ y $v_{j_k} = v_j$, y el resto de los nodos son distintos. En otras palabras, una trayectoria cerrada sin autocruzamientos y con ramas que sólo se recorren una vez. Por ejemplo, las ramas e_1, e_2 y e forman un circuito en el nodo V_1 de la figura. 6.4.

En una de las fórmulas de Kirchhoff se pide la cantidad de árboles generadores en una gráfica. Un **árbol generador** de una gráfica G contiene *todos* los nodos de G junto con algunas ramas de tal manera que

1. El árbol generador no tenga circuitos;
2. Dos vértices cualesquiera puedan conectarse con ramas que pertenezcan al árbol generador.

La **matriz de árbol** de la gráfica G es, entonces, la matriz T de $n \times n$ cuyos elementos t_{ij} se definen como sigue:

$$t_{ij} := \begin{cases} \text{si } i = j, & t_{ii} = \text{cantidad de ramas incidentes en el vertice } v_i \\ \text{si } i \neq j, & t_{ij} = \begin{cases} -1 & \text{si } v_i \text{ y } v_j \text{ son adyacentes} \\ 0 & \text{si } v_i \text{ y } v_j \text{ no son adyacentes} \end{cases} \end{cases}$$

NOTA Una matriz de árbol A es *simétrica*, es decir, $A = A^T$.

[13] Recuérdese que en el capítulo 3 se describieron algunas de las bases de la teoría de las gráficas.

■ EJEMPLO 32 La matriz de árbol de la gráfica en la figura. 6.4 es

$$T = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

El teorema siguiente, deducido por Kirchhoff, cuenta la cantidad de árboles generadores de una gráfica. Se conoce como el *teorema de la matriz de árbol.*

TEOREMA 14

(Teorema de la matriz de árbol)

Todos los cofactores de la matriz de árbol T de la gráfica G son iguales y su valor común es la cantidad de árboles generadores de G.

■ EJEMPLO 33 Veamos la gráfica de la figura. 6.4. En ésta, la cantidad de árboles generadores se obtiene calculando sólo uno de los cofactores iguales.

$$C_{11} = (-1)^{1+1} M_{11} = \begin{vmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 3 \end{vmatrix} = 8$$

Por consiguiente, la cantidad de árboles generadores es 8. En este caso pueden trazarse con facilidad todos ellos para contarlos (figura 6.5).

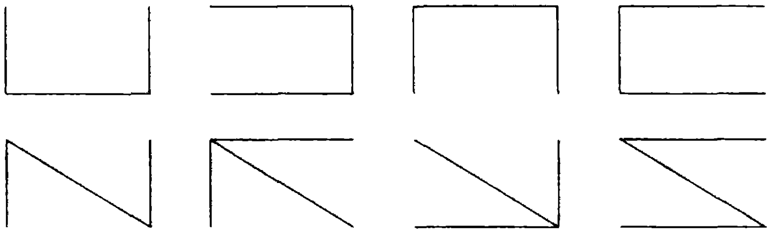


Figura 6.5 Árboles generadores de G.

El poder del teorema 14 puede apreciarse aún más cuando es necesario contar los árboles generadores de una gráfica como la de la figura. 6.6. La matriz del árbol es

$$T = \begin{vmatrix} 2 & -1 & 0 & 0 & 0 & -1 \\ -1 & 4 & -1 & -1 & -1 & 0 \\ 0 & -1 & 4 & -1 & -1 & -1 \\ 0 & -1 & -1 & 3 & -1 & 0 \\ 0 & -1 & -1 & -1 & 4 & -1 \\ -1 & 0 & -1 & 0 & -1 & 3 \end{vmatrix}$$

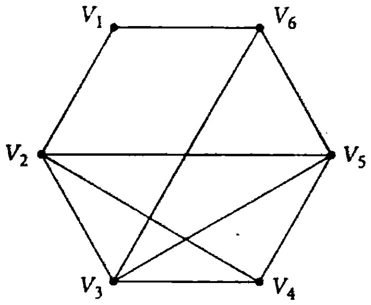


Figura 6.6

Al calcular el cofactor C_{11} se obtiene $C_{11} = (-1)^{1+1}M_{11} = 115$. Así, hay 115 árboles generadores. En este caso, trazarlos y contarlos es bastante tedioso.

El problema de las vacas y los campos de Newton

Considere que

1. a_1 vacas consumen b_1 campos en c_1 días;
2. a_2 vacas consumen b_2 campos en c_2 días;
3. a_3 vacas consumen b_3 campos en c_3 días.

Suponiendo que todos los campos tienen la misma cantidad de pasto, que el crecimiento diario de los campos permanece constante y que las vacas comen la misma cantidad cada día, ¿qué relación existe entre las 9 magnitudes $a_1, \ldots, c_3$?

Sea x la cantidad inicial de pasto en cada campo, sea y el crecimiento diario y sea z el consumo diario por vaca. Según las hipótesis, x, y y z permanecen constantes.

Veamos la afirmación 1. En c_1 días, a_1 vacas comen a_1c_1z cantidad de pasto, la cual es igual a la cantidad inicial de pasto, b_1x, más la que ha crecido en c_1 días, b_1c_1y. En consecuencia $b_1x + b_1c_1y = a_1c_1z$. De igual forma, las afirmaciones 2 y 3 conducen a $b_2x + b_2c_2y = a_2c_2z$ y $b_3x + b_3c_3y = a_3c_3z$, respectivamente. Así llegamos al sistema homogéneo

$$\begin{aligned} b_1x + b_1c_1y - a_1c_1z &= 0 \\ b_2x + b_2c_2y - a_2c_2z &= 0 \\ b_3x + b_3c_3y - a_3c_3z &= 0 \end{aligned}$$

que tiene a x, y y z como incógnitas. De acuerdo con el teorema 4, el sistema tiene soluciones no triviales si y sólo si

$$\begin{vmatrix} a_1c_1 & b_1c_1 & b_1 \\ a_2c_2 & b_2c_2 & b_2 \\ a_3c_3 & b_3c_3 & b_3 \end{vmatrix} = 0$$

Esta es la condición necesaria requerida que relaciona los nueve números $a_1, \ldots, c_3$ aparentemente no relacionados.

Resultantes con sistemas algebraicos computacionales

Maple

```
> resultant(a*x^2+b*x+c,2*a*x+b,x);
```

$$4a^2c - b^2a$$

Mathematica

```
In[1]:=
     Resultant[a x^2+b x+c, 2a x+b, x]
Out[1]=
         2      2
     -(a b ) + 4 a  c
```

Ejercicios 6.5

En los ejercicios 1 a 4 deduzca la ecuación de la recta que pasa por P y Q.

1. $P(-1, 2)$ y $Q(1, 1)$
2. $P(2, 1)$ y $Q(1, -1)$
3. $P(0, 1)$ y $Q(6, -8)$
4. $P(-3, -5)$ y $Q(4, 7)$

En los ejercicios 5 y 6 determine si los puntos P, Q y R están en la misma recta.

5. $P(-2, 0)$, $Q(0, 1)$ y $R(2, 1)$
6. $P(-1, 2)$, $Q(0, 0)$ y $R(2, -1)$

En los ejercicios 7 a 10 obtenga la ecuación, el centro y el radio del círculo que pasa por P, Q y R.

7. $P(0, 0)$, $Q(-1, -1)$ y $R(0, -2)$
8. $P(2, 2)$, $Q(4, 0)$ y $R(6, 2)$
9. $P(5, 5)$, $Q(1, -1)$ y $R(0, 0)$
10. $P(7, 7)$, $Q(1, 1)$ y $R(-3, 2)$

En los ejercicios 11 a 13 deduzca la ecuación del tipo $y = ax^2 + bx + c$ para la parábola que pasa por P, Q y R.

11. $P(0, 4)$, $Q(1, 3)$ y $R(-1, 9)$
12. $P(2, 0)$, $Q(0, 1)$ y $R(1, 0)$
13. $P(2, 2)$, $Q(3, 1)$ y $R(1, 7)$.

En los ejercicios 14 a 16 encuentre la ecuación del plano que pasa por P, Q y R.

14. $P(1, 1, 1)$, $Q(0, -1, 1)$, y $R(4, 3, -1)$
15. $P(-1, 1, 1)$, $Q(-1, 4, 3)$, y $R(4, 0, 2)$
16. $P(5, 4, 3)$, $Q(-1, 2, 2)$, y $R(4, 4, 4)$

17. Aplique el resultante de Sylvester para resolver el sistema

$$x^2 + y^2 - 1 = 0$$
$$x^2 + 2x + y^2 - 2y + 1 = 0$$

18. Use el resultante de Sylvester para resolver el sistema

$$x^2 + y^2 - 1 = 0$$
$$x^2 - 2x + y^2 - 1 = 0$$

19. Deduzca la ecuación de la esfera que pasa por los puntos $(1, 2, 7)$, $(5, 2, 3)$, $(1, 6, 3)$, $(1, 2, -1)$. (*Sugerencia:* Recuerde que la ecuación de una esfera tiene la forma $(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2$, o bien $A(x^2 + y^2 + z^2) + Bx + Cy + Dz + E = 0$, donde $A = 1$. Forme un sistema homogéneo con A, B, C, D y E como incógnitas.)

En los ejercicios 20 a 22 trace las gráficas correspondientes a los circuitos eléctricos dados.

20. Para la figura 6.7

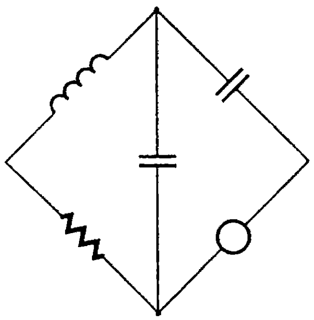


Figura 6.7

21. Para la figura 6.8

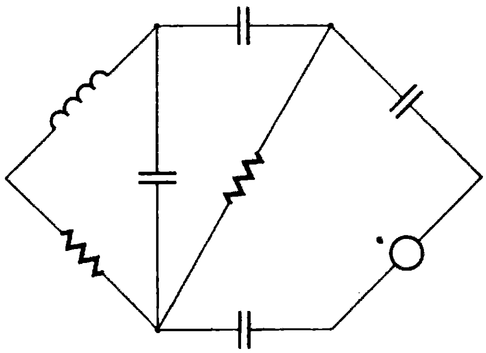


Figura 6.8

22. Para la figura 6.9

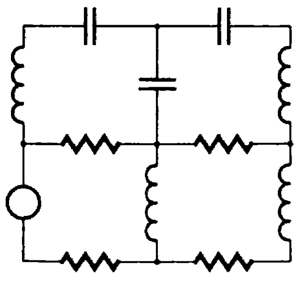


Figura 6.9

23. **a.** Trace todos los árboles generadores de la gráfica G_1 en la figura 6.10, y cuéntelos.

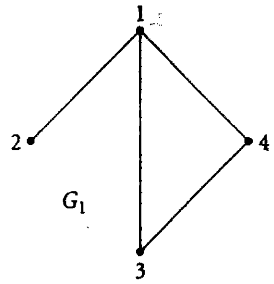


Figura 6.10

b. Escriba la matriz de árbol para la gráfica, y use el teorema 14 para calcular la cantidad de árboles generadores.

24. Escriba la matriz de la gráfica G_2 de la figura 6.11 y use el teorema 14 para calcular la cantidad de árboles generadores.

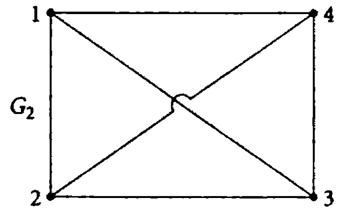


Figura 6.11

6.6 Miniproyectos

1 ■ Determinantes de matrices de bloques

Sean

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 1 & 4 \\ 2 & 0 & -2 \\ 3 & 1 & 1 \end{bmatrix}$$

y sean

$$D = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{bmatrix}, \qquad E = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 3 & 1 & 1 \end{bmatrix}$$

Observe que D es la matriz de bloques formada colocando A y B a lo largo de la diagonal de los bloques, y rellenando con ceros el resto de los elementos. E se formó de la misma manera a partir de A y C.

Problema A

Deduzca una relación entre

1. $\det(A)$, $\det(B)$ y $\det(D)$;
2. $\det(A)$, $\det(C)$ y $\det(E)$.

En general, sean $A_1, \ldots, A_m$ m matrices cuadradas, no necesariamente del mismo tamaño. Sea A la matriz de bloques en diagonal que tiene a $A_1, A_2, \ldots, A_m$ en la diagonal principal.

$$A = \begin{bmatrix} A_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & A_2 & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & A_m \end{bmatrix}$$

Problema B

Demuestre que el determinante de la matriz de bloques A es el producto de los determinantes de las matrices en la diagonal.

$$\det(A) = \begin{vmatrix} A_1 & \mathbf{0} & & \mathbf{0} \\ \mathbf{0} & A_2 & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & A_m \end{vmatrix} = \det(A_1)\det(A_2)\cdots\det(A_m)$$

(*Sugerencia:* Primero observe que, con una reducción cuidadosa con operaciones de renglón, el determinante de cualquier matriz cuadrada es igual al determinante de una matriz triangular superior. Después vea que este último es igual al producto de los elementos de la diagonal principal.)

Si A, B, C y D son matrices $n \times n$, forme la matriz de bloques $\mathbf{M}$ $2n \times 2n$:

$$\mathbf{M} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

Podemos calcular el determinante de $\mathbf{M}$ basándonos en la proposición siguiente, que enunciaremos sin demostrar.

Proposición 15

Si A, B, C y D conmutan entre sí, es decir, si

[redacted] $AC = CA$ [redacted]

entonces

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - BC)$$

Problema C

1. Compruebe las hipótesis y la conclusión de la proposición 15 con

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & 1 \\ 0 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

2. Determine en qué matrices E, F, G y H 2×2 no aplica la proposición 15.

2 ■ Determinantes de Vandermonde

Sea

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 9 & 25 \end{bmatrix}$$

Observe que los elementos en cada columna son potencias: $2^0 = 1$, $2^1 = 1$, $2^2 = 4$ en la primera columna, $3^0 = 1$, $3^1 = 3$, $3^2 = 9$ para la segunda columna, y $5^0 = 1$, $5^1 = 5$, $5^2 = 25$ para la tercera columna. Una matriz con esta propiedad se llama *matriz de Vandermonde*.

DEFINICIÓN

Una matriz A_n, $n \times n$, es una **matriz de Vandermonde** si hay números $x_1, x_2, \ldots, x_n$ tales que

$$A_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}$$

Hay una fórmula sencilla para calcular el determinante de una matriz de Vandermonde.

Proposición 16

(Determinante de Vandermonde)

El determinante V_n de la matriz de Vandermonde A_n se expresa como sigue:

$$V_n = \det(A_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

En otras palabras,

$$\begin{aligned} V_n = {} & (x_n - x_{n-1})(x_n - x_{n-2}) \cdots (x_n - x_2)(x_n - x_1) \times \\ & (x_{n-1} - x_{n-2})(x_{n-1} - x_{n-3}) \cdots (x_{n-1} - x_2)(x_{n-1} - x_1) \times \\ & \cdots \\ & (x_4 - x_3)(x_4 - x_2)(x_4 - x_1) \times \\ & (x_3 - x_2)(x_3 - x_1) \times \\ & (x_2 - x_1) \end{aligned}$$

Problema A

1. Compruebe la proposición 16 para las matrices de Vandermonde siguientes:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 9 & 25 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 4 \end{bmatrix}$$

2. Use la proposición 16 para calcular el determinante de las matrices de Vandermonde siguientes:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 10 & 11 & 12 \\ 100 & 121 & 144 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{bmatrix}$$

3. Aplique la proposición 16 para calcular el determinante de las matrices siguientes:

$$A = \begin{bmatrix} 1 & 5 & 25 \\ 1 & 9 & 81 \\ 1 & 12 & 144 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 27 \\ 1 & 5 & 25 & 125 \\ 1 & 7 & 49 & 343 \end{bmatrix}$$

Problema B

Encuentre una condición necesaria y suficiente para que el determinante de la matriz de Vandermonde sea igual a cero.

Problema C

Demuestre la proposición 16.

(*Sugerencias:*

1. Efectúe las operaciones siguientes:

$$R_n - x_n R_{n-1} \rightarrow R_n$$

$$R_{n-1} - x_n R_{n-2} \to R_{n-1}$$
$$R_{n-2} - x_n R_{n-3} \to R_{n-2}$$
$$\vdots$$
$$R_2 - x_n R_1 \to R_2$$

para obtener una matriz cuya i-ésima columna sea

$$\begin{bmatrix} 1 \\ x_i - x_n \\ x_i(x_i - x_n) \\ x_i^2(x_i - x_n) \\ \vdots \\ x_i^{n-2}(x_i - x_n) \end{bmatrix} \quad \text{si} \quad i < n \quad \text{y} \quad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{si } i = n$$

2. Desarrolle el determinante que se obtiene respecto a la n-ésima columna para deducir el cofactor $(1, n)$: $C_{1,n} = (-1)^n M_{1,n}$.
3. Aplique el teorema 1 para factorizar el producto

$$(x_1 - x_n)(x_2 - x_n) \cdots (x_{n-1} - x_n)$$

como factor común del menor $M_{1,n}$. El determinante que queda es justamente V_{n-1}. Continúe con este proceso.)

3 ■ Resultante de Bezout

Este proyecto es una introducción rápida a la cada vez más importante resultante de Bezout,[14] dada a conocer en 1774, de un sistema de dos polinomios en una variable. Describiremos el *enunciado de Cayley* del método de Bezout.

Sean $f(x)$ y $g(x)$ dos polinomios. Se desea definir una condición necesaria para que el sistema $f(x) = 0, g(x) = 0$ tenga una solución común. Sea a una segunda variable independiente de x. Representaremos los polinomios $f(x)$ y $g(x)$ mediante $f(a)$ y $g(a)$ cuando x se reemplaza por a. Examinemos el determinante 2×2:

$$\Delta(x, a) = \begin{vmatrix} f(x) & g(x) \\ f(a) & g(a) \end{vmatrix} = f(x)g(a) - g(x)f(a)$$

Observe que Δ es cero para cualquier solución común, x, del sistema $f(x) = 0$, $g(x) = 0$. Es más, $\Delta = 0$ si $x = a$. Por consiguiente, $x - a$ es divisor exacto de $\Delta(x, a)$. Por lo anterior, el cociente

$$\delta(x, a) = \frac{\Delta(x, a)}{x - a} = \frac{f(x)g(a) - g(x)f(a)}{x - a}$$

es cero para cualquier solución del sistema original. Asimismo, $\delta(x, a)$ es un polinomio en a y en x. Para cualquier cero común del sistema, por ejemplo $x = x_0$, $\delta(x_0, a)$ es cero para toda a; de esta manera, los coeficientes de las potencias de a en $\delta(x, a)$ son iguales a cero. Con esta igualdad se obtiene un sistema homogéneo en x, cuyas soluciones son no triviales si el deter-

[14] En especial la versión multivariada debida a Dixon (1908).

minante de la matriz de coeficientes es cero. A este último determinante se le llama **resultante de Bezout** del sistema. Si el sistema original tiene un cero común, el resultante de Bezout es cero.

Por ejemplo, si

$$f(x) = x^2 - 5x + 6$$
$$g(x) = x^2 + 2x - 8$$

Entonces

$$\delta(x,a) = \frac{1}{x-a}\begin{vmatrix} x^2 - 5x + 6 & x^2 + 2x - 8 \\ a^2 - 5a + 6 & a^2 + 2a - 8 \end{vmatrix}$$

$$= \frac{1}{x-a}\left[(x^2 - 5x + 6)(a^2 + 2a - 8) - (x^2 + 2x - 8)(a^2 - 5a + 6)\right]$$

que se simplifica a

$$\delta(x,a) = 7ax - 14x - 14a + 28$$

Para cualquier cero común del sistema, $\delta(x, a) = 0$ para toda a. De ahí que los coeficientes de todas las potencias de a deben ser cero. Como el coeficiente de a^0 es $28 - 14x$ y el coeficiente de a^1 es $-14 + 7x$, entonces

$$28 - 14x = 0$$
$$-14 + 7x = 0$$

Aquí determinante de la matriz de coeficientes es el resultante de Bezout, y como éste es cero, el sistema tiene una solución común.

$$\begin{vmatrix} 28 & -14 \\ -14 & 7 \end{vmatrix} = 0$$

NOTA Si f y g, los dos polinomios, son del mismo grado, los resultantes de Bezout y de Sylvester son idénticos. Con frecuencia se prefiere el resultante de Bezout, porque el tamaño del determinante Δ (= máx(grado(f), grado(g))) es mucho menor que el del determinante de Sylvester (= grado(f) + grado(g)).

Problema A

Con el método de Bezout podemos eliminar una variable de un sistema de dos ecuaciones polinomiales con dos variables. Por ejemplo,

$$x^2 + y^2 - 1 = 0$$
$$x^2 - 2x + y^2 - 2y + 1 = 0$$

como un sistema en y con coeficientes polinomiales en x:

$$y^2 + (x^2 - 1) = 0$$
$$y^2 - 2y + (x^2 - 2x + 1) = 0$$

1. Use el método de Bezout para eliminar y.
2. Iguale el resultante de Bezout y despeje x.
3. Sustituya en el sistema original y calcule todos los ceros comunes.

Problema B

Repita el proceso del problema A para el sistema siguiente, y llegue a la conclusión que no hay soluciones comunes:

$$x^2 + y^2 - 1 = 0$$
$$x^2 - 6x + y^2 - 2y + 6 = 0$$

4 ■ Productos cruz en $\mathbf{R}^n$

En este proyecto describiremos una generalización interesante del producto cruz[15] en $\mathbf{R}^n$.

Sean $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ los vectores comunes de la base $\mathbf{R}^n$ y suponga que $n \geq 3$. Fije tres vectores $n - 3$ cualesquiera, $\mathbf{a}_1, \ldots, \mathbf{a}_{n-3}$ en $\mathbf{R}^n$ y defina un producto $\times$ en $\mathbf{R}^n$ por medio del determinante

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \\ & \mathbf{a}_1 & \\ & \vdots & \\ & \mathbf{a}_{n-3} & \\ & \mathbf{u} & \\ & \mathbf{v} & \end{vmatrix}, \qquad \mathbf{u}, \mathbf{v} \in \mathbf{R}^n$$

■ **EJEMPLO 34** Sean $n = 4$, $\mathbf{a}_1 = (a_1, a_2, a_3, a_4)$, $\mathbf{u} = (u_1, u_2, u_3, u_4)$ y $\mathbf{v} = (v_1, v_2, v_3, v_4)$. Entonces, en $\mathbf{R}^4$:

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ a_1 & a_2 & a_3 & a_4 \\ u_1 & u_2 & u_3 & u_4 \\ v_1 & v_2 & v_3 & v_4 \end{vmatrix}$$

$$\begin{aligned} &= (a_2u_3v_4 - a_2u_4v_3 - u_2a_3v_4 + u_2a_4v_3 + v_2a_3u_4 - v_2a_4u_3)\mathbf{e}_1 \\ &\quad + (-v_4u_3a_1 + v_3u_4a_1 + v_4a_3u_1 - v_3a_4u_1 - u_4a_3v_1 + u_3a_4v_1)\mathbf{e}_2 \\ &\quad + (v_4a_1u_2 - u_4a_1v_2 - v_4u_1a_2 + a_4u_1v_2 + u_4v_1a_2 - a_4v_1u_2)\mathbf{e}_3 \\ &\quad + (-v_3a_1u_2 + u_3a_1v_2 + v_3u_1a_2 - a_3u_1v_2 - u_3v_1a_2 + a_3v_1u_2)\mathbf{e}_4 \end{aligned}$$

■ **EJEMPLO 35** Si $\mathbf{a}_1 = (1, 1, 1, 1)$, tenemos que

$$(2, 5, 10, 17) \times (3, 1, -1, -3) = (4, -12, 12, -4)$$

Problema A

1. Si $\mathbf{a}_1 = (1, -1, 1, 1)$, determine $(2, 5, 2, -3) \times (2, 3, -1, -3)$ en $\mathbf{R}^4$.
2. Si $\mathbf{a}_1 = (1, 1, 1, 1, 1)$ y $\mathbf{a}_2 = (1, -1, 1, 0, 1)$, encuentre $(2, 0, 2, -3, 1) \times (1, 0, 3, -1, -3)$ en $\mathbf{R}^5$.

[15] Para conocer más detalles sobre el tema, véase el artículo de Dittmer en la revista *American Mathematical Monthly* (noviembre de 1944) y las referencias que allí se citan.

Problema B

Fije vectores $\mathbf{a}_1, \ldots, \mathbf{a}_{n-3}$ en $\mathbf{R}^n$. Con las propiedades de los determinantes[16] demuestre que para cualesquiera vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ en $\mathbf{R}^n$, son válidas las siguientes:

1. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$
2. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$
3. $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$
4. $k(\mathbf{u} \times \mathbf{v}) = (k\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (k\mathbf{v})$
5. $\mathbf{0} \times \mathbf{u} = \mathbf{u} \times \mathbf{0} = \mathbf{0}$
6. $\mathbf{u} \times \mathbf{u} = \mathbf{0}$

Problema C

El producto cruz generalizado es no asociativo, es decir, para cualquier elección de $\mathbf{a}_i$ distinto de cero, hay vectores n, $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ tales que

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) \neq (\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$$

Demuestre esta afirmación para $n = 4$.

6.7 Ejercicios en computadora

En esta sección el lector aprenderá cómo calcular determinantes con los programas que usa, al efectuar las partes computacionales de los ejercicios. También ayudará a repasar algo del material básico de este capítulo. Observe que un ejercicio identificado con [S] requiere manipulación simbólica.[17]

Sean

$$A = \begin{bmatrix} 6 & 7 & 1 \\ 6 & -7 & 2 \\ 6 & 7 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \\ 13 & 15 & 17 \end{bmatrix}, \qquad D = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

1. Calcule los determinantes de A, B y C. ¿Cuáles de esas matrices son invertibles?
2. Encuentre $\det(B)$ con desarrollo por cofactores respecto a la segunda columna.
3. Calcule (a) $\det(A^5) - (\det(A))^5$, (b) $\det(A) - \det(A^T)$, (c) $\det(AB) - \det(A)\det(B)$. En cada caso, explique por qué la respuesta debe ser cero.
4. Encuentre (a) $\det(5B) - 125\det(B)$, (b) $\det(B^{-1}) - 1/\det(B)$, (c) $\det(B^{-2}) - 1/\det(B)^2$. En cada caso, explique por qué la respuesta debe ser cero.
5. [S] Calcule y compare $\det(D)$ y $\det(D^T)$. Repítalo con el $\det(D^TD)$ y $\det(D)^2$. Explique los resultados de las comparaciones.

---

[16] Aun cuando el primer renglón es renglón de vectores en el producto cruz.

[17] Omita esos ejercicios si no dispone de manipulación simbólica.

Sea M_n la sucesión de matrices:

$$M_2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad M_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix},$$

$$M_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 4 \end{bmatrix}, \ldots$$

6. Defina M_n como una función de n. Use esta función para calcular $\det(M_2)$, $\det(M_3)$, $\det(M_4) \cdots$ ¿Ve usted un patrón para el $\det(M_n)$?
7. Calcule $\det(S^TS)$ para varias matrices S 3×2. Encuentre una relación entre S^TS cuando es invertible y la dependencia o independencia lineal de las columnas de S.
8. Resuelva el sistema siguiente (a) por inversión de matrices y (b) con la regla de Cramer. (c) Calcule la adjunta de la matriz de coeficientes.

$$\begin{aligned} -85x - 55y - 37z &= -306 \\ -35x + 97y + 50z &= 309 \\ 79x + 56y + 49z &= 338 \end{aligned}$$

9. Sea $A\mathbf{x} = \mathbf{b}$ un sistema cuadrado. Escriba el programa para dos funciones, CramerDisplay y CramerSolve, cada uno con tres argumentos, A, $\mathbf{b}$ e i. CramerDisplay debe desplegar la matriz A_i obtenida de A reemplazando su i-ésima columna con $\mathbf{b}$. CramerSolve debe calcular las x_i con la regla de Cramer. Pruebe el programa mostrando A_1, A_2 y A_3 y resolviendo el sistema anterior.
10. Si dispone del comando adecuado, determine *todas* las permutaciones de $\{1, 2, 3, 4\}$. Compruebe que el número calculado de permutaciones sea el correcto.
11. Si dispone del comando adecuado, determine los signos de las permutaciones: $\{1, 4, 2, 3\}$ y $\{4, 2, 3, 1\}$.
12. Si dispone del comando adecuado, determine una permutación aleatoria de $\{1, 2, 3, 4\}$.
13. [S] Utilice determinantes para definir una función $f(x_1, y_1, x_2, y_2, x_3, y_3)$ que determine la ecuación, con x y y como variables, de un círculo que pase por los tres puntos (x_1, y_1), (x_2, y_2) y (x_3, y_3).
14. [S] Use f para determinar la ecuación del círculo C_1 que pasa por $(-1, 1)$, $(1, 1)$, $(2, 4)$.
15. [S] Determine el o los puntos con abscisa -2 en el círculo C_1 anterior. También, demuestre que C_1 no tiene puntos cuya abscisa sea -3.
16. Con determinantes, defina una función $g(x_1, y_1, x_2, y_2, x_3, y_3)$ que investigue si uno de los cuatro puntos cuyas coordenadas son (x, y), (x_1, y_1), (x_2, y_2) y (x_3, y_3) están en el mismo círculo.
17. Use g para comprobar si los puntos $A(-2, 2)$, $B(-2, 4)$ y $C(-1, 2)$ están en el círculo C_1.
18. Sean

$$\begin{aligned} p_1 &= x^4 - 3x^3 - 63x^2 - 85x + 150 \\ p_2 &= x^4 + 22x^3 - 103x^2 - 2740x + 5700 \\ p_3 &= x^5 - 4x^3 + 8x^2 - 12x - 1 \end{aligned}$$

Calcule los resultantes de pares de polinomios (p_1, p_2), (p_1, p_3) y (p_2, p_3). ¿Cuáles de ellos tienen una solución común?

19. Calcule la cantidad de árboles generadores de la gráfica formada por un hexágono regular y todas sus diagonales.

Determinantes y permutaciones en Maple, Mathematica y MATLAB

Sea A cualquier matriz cuadrada y n cualquier entero positivo. Entonces:

función	Maple	Mathematica	MATLAB
determinante de A	det(A);	Det[A]	det(A)
permutación de $\{1,..., n\}$	permute(n);	Permutations[n]	
signo de la permutación		Signature[p]	
permutación aleatoria	randperm(n);		randperm(n)
resultante	resultant	Resultant	
matriz de Sylvester	sylvester		

En Maple, permute(n) muestra todas las permutaciones de $\{1, 2, \ldots, n\}$. Primero, necesita cargar el paquete combinat tecleando con(combinat);. Entonces, randperm(n) despliega una permutación aleatoria de $\{1, 2, \ldots, n\}$ Los comandos resultant y sylvester están en el paquete linalg.

En Mathematica, Permutations[n] muestra todas las permutaciones de $\{1, 2, \ldots, n\}$, y Signature[{2,5,4,1,3}] despliega el signo de la permutación $\{2, 5, 4, 1, 3\}$.

En MATLAB, randperm(n) muestra una permutación aleatoria de $\{1, 2, \ldots, n\}$.

Soluciones seleccionadas con Maple

```
with(linalg):
A := matrix([[6, 7, 1], [6, -7, 2], [6, 7, 3]]);                # DATOS
B := matrix([1/3, 1/4, 1/5], [1/4, 1/4, 1/5], [1/5, 1/5, 1/5]]);
C := matrix([[1, 3, 5], [7, 9, 11], [13, 15, 17]]);
D1 := matrix([[a, b, c], [d, e, f], [g, h, i]]);    # D se usa para diferenciar.
# Ejercicios 1 a 4.
det(A); det(B); det(C);    # C es la única no invertible, porque det(C)=0.
-(1/4) *det(minor(B, 1, 2)+(1/4)*det(minor(B, 2, 2)) - (1/5)*det(minor(B, 3, 2));
det(A^5)-det(A)^5; det(A)-(transpose(A)); det(A &* B) -det(A) * det(B);
det(5*B)-125*det(B); det(inverse(B))-1/det(B); det(B^(-2)) -1/det(B)^2;
# Ejercicio 5.
det(D1); det(transpose(D1));
det(transpose(D1) &* D1); det(D1)^2;
""-expand(");                      # Desarrollar la diferencia para obtener 0.
# Ejercicio 6
m:=proc(n) matrix(n,n, (i, j)->if i=j the i else 1 fi) end:      # M_n.
det(m(2)); det(m(3)); det(m(4));                                 # Etc. .
# Patrón: det (M_n) = (n-1)!
# Ejercicio 7 - Comentario.
# (S^T)S es invertible sólo si las columnas de S son linealmente independientes.
```

```
# Ejercicio 8.
A :=matrix([[-85, -55, -37], [-35, 97, 50], [79, 56,49]]);
b :=vector([-306, 309, 338]);
sol :=evalm(inverse(A)&*b);     # (a)
A1 :=delcols(A,1..1);           # Elimina columna 1 de A
A1 :=augment(b,A1);             # después agrega b como primera columna
A2 :=delcols(A, 1..2);          # la última columna de A
A2 :=augment(b, A2);            # seguida de b
A22 :=delcols((a, 2..3);        # la primera columna de A
A2 :=augment(A22, A2);          # unida con las columnas de A22
A3 :=delcols(A, 3..3);          # elimina la columna 3 de A
A3 :=augment(A3, b);            # y después agrega b como última columna
x :=det(A1)/det(A);             # (b)
y :=det(A2)/det(A);             # (b)
z :=det(A3)/det(A)              # (b)
adj :=adjoint(A);               # (c)

# Ejercicio 9.
CramerDisplay := proc (A, b, i) local AA, j; AA:= copy(A);
                   for j from 1 to rowdim(A) do
                   AA[j, i] :=b[j] od:            # Reemplazo de la iésima
                   evalm(AA)                      # columna con b.
                 end:
CramerSolve := proc (A, b, i) det(CramerDisplay(A, b, i))/det(A) end:
CramerDisplay(A,b,1); CramerDisplay(A, b, 2); CramerDisplay(A, b, 3);
Cramersolve(A, b, 1); CramerSolve(A, b, 2); CramerSolve(A, b, 3);

# Ejercicios 10 y 12.
with(combinat);       # Cargar el paquete combinat.
permute(4);           # Las permutaciones de {1,2,3,4}.
nops(");              # La cantidad de permutaciones calculadas.
4!;                   # La respuesta esperada.
randperm(4);          # Una permutación aleatoria de {1,2,3,4}.

# Ejercicio 13.
circleeqn := proc(x1, y1, x2, y2, x3, y3)
             det(matrix(4, 4, [x^2+y^2, x, y, 1,
                               x1^2+y1^2,x1,y1,1,
                               x2^2+y2^2,x2,y2,1,
                               x3^2+y3^2,x3,y3,1]))
       end:

# Ejercicio 14.
ce :=circleeqn(-1, 1, 1, 1, 2, 4);    # La ecuación del círculo.

# Ejercicio 17.
subs(x=-2, y=2,ce);                   # Sustitución de los puntos
subs(x=-2, y=4,ce);                   # en la ecuación del círculo.
subs(8x=-1, y=4,ce);                  # El último punto no está en el círculo.

# Ejercicio 18 - parcial.
p1:=x^4-3*x^3-63*x^2-85*x+150;
```

```
p2 :=x^4+22*x^3-103*x^2-2740*x+5700;
resultant(p1, p2, x);          # También se necesita declarar la variable x.
```

Nota En Maple se incluye un comando circle que calcula, entre otras cosas, la ecuación de un círculo. Este comando es parte del paquete geometry, que se debe cargar primero.

Soluciones seleccionadas con Mathematica

```
A = {{6, 7, 1}, {6, -7, 2}, {6, 7, 3}}                     (* DATOS *)
B = {{1/3, 1/4, 1/5}, {1/4, 1/4, 1/5} {1/5, 1/5, 1/5}}
C1 = {{1, 3, 5}, {7, 9, 11}, {13, 15, 17}}  (* C se usa en constantes de ecuaciones diferenciales. *)
D1 = {{a, b, c}, {d, e, f}, {g, h, i}}        (* D se usa para diferenciación.*)

(* Ejercicios 1 a 4. *)

Det[A]
Det[B]
Det[C1]                     (* C es la única no invertible, porque Det[C]=0. *)
(* Cuidarse de la numeración inversa en el comando Minors. *)
-(1/4) Minors[B, 2] [[3, 2]]+(1/4) Minors[B, 2] [[2, 2]]-(1/5) Minors[B, 2] [[1, 2]]
Det[MatrixPower[A, 5]]-Det[A]^5
Det[A]-Det[Transpose[A]]
Det[A.B]-Det[A] Det[B]
Det[5B]-125Det[B]
Det[Inverse[B]]-1/Det[B]
Det[MatrixPower[B, -2]]-1/Det[B]^2

(* Ejercicio 5. *)

Det[D1]
Det[Transpose[D1]]
Det[Transpose[D1].D1]
Expand[% - Det[D1]^2]          (* Desarrollar la diferencia para obtener 0.  *)

(* Ejercicio 6.  *)

m[n_] := Table[If[i==j, i, 1], {i, 1, n}, {j, 1, n}]              (* M_n.   *)
{Det[m[2]], Det[m[3]], Det[m[4]], Det[m[5]]}                      (* Etc. .  *)
(* Patrón: det (M_n) = (n-1)!  *)

(* Ejercicio 7 - Comentario.  *)

(* (S^T)S es invertible sólo si las columnas de S son linealmente independientes. *)

(* Ejercicio 8.  *)

<<LinearAlgebra`MatrixManipulation`    (* Carga el paquete de manipulación de matrices. *)
A={{-85, -55, -37}, {-35, 97, 50}, {79, 56, 49}}
b={{-306}, {309}, {338}}
sol=Inverse[A] . b               (* (a) *)
A1=TakeColumns[A, {2, 3}]        (* Elimina la columna 1 de A    *)
A1=AppendRows[b, A1]             (* después agrega b como primera columna    *)
A2=TakeColumns[A, {3}]           (* toma la columna 3 de A     *)
A2AppendRows[b, A2]              (* agrega b como primera columna.     *)
A22=TakeColumns[A, {1}]          (* toma la columna 1 de A    *)
```

```
A2=AppendRows[A22, A2]        (* une las columnas de A22 y A2  *)
A3=TakeColumns[A, {1, 2}]     (* elimina la columna 3 de A  *)
A3=AppendRows[A3, b]          (* a continuación agrega b como última columna.  *)
x=Det[A1]/Det[A]              (*  (b)  *)
y=Det[A2]/Det[A]              (*  (b)  *)
z=Det[A3]/Det[A]              (*  (b)  *)
adj=Det[A] Inverse[A]         (*  (c)  *)
(* Ejercicio 9. *)
CramerDisplay[A_, b_, i_] := Module[{AA=A, j},
            For[j=1, j<=Length[A], j++,   (* Reemplazo de la iésima  *)
              AA[[j, i]]=b[[j, 1]]]; AA]  (* columna con b.  *)
CramerSolve[A_, b_, i_] := Det[CramerDisplay[A, b, i]]/Det[A]]
CramerDisplay[A, b, 1]     (* También CramerDisplay[A, b, 2] y CramerDisplay[A, b, 3] *)
CramerSolve[A, b, 1]       (* También CramerSolve[A, b, 2] y CramerSolve[A, b, 3]  *)
(* Ejercicios 10 y 11.  *)
Permutations[{1, 2, 3, 4}]         (* Las permutaciones de {1,2,3,4}.  *)
Length[%]                          (* La cantidad calculada de permutaciones.  *)
4!                                 (* La respuesta esperada.  *)
Signature[{1, 4, 2, 3}]            (* Signo.  *)
Signature[{4, 2, 3, 1}]
(* Ejercicio 13.  *)
Clear[x, y, z]                     (* Borrar los valores del ejercicio 8.  *)
circleeqn[x1_, y1_, x2_, y2_, x3_, y3_] := Det[{{x^2+y^2, x, y, 1}
                                   {x1^2+y1^2, x1, y1, 1}
                                   {x2^2+y2^2, x2, y2, 1}
                                   {x3^2+y3^2, x3, y3, 1}}]
(* Ejercicio 14.  *)
ce=circleeqn[-1, 1, 1, 1, 2, 4]    (* La ecuación del círculo.  *)
(* Ejercicio 17.  *)
ce /. {x->-2, y->2}                (* Sustitución de los puntos en la  *)
ce /. {x->-2, y->4}                (* ecuación del círculo.  *)
ce /. {x->-1, y->2}                (* El último punto no está en el círculo.  *)
(* Ejercicio 18. - Parcial.  *)
p1=x^4-3x^3-63x^2-85x+150
p2=x^4+22x^3-103x^2-2740x+5700
Resultant[p1, p2, x]   (* También se necesita declarar la variable x.  *)
```

Soluciones seleccionadas con MATLAB

Nota Todo renglón de comando indicado con (ST) requiere los métodos simbólicos *symbolic toolbox*.

```
A = [6 7 1; 6 -7 2; 6 7 3]                    % DATOS
B = [1/3 1/4 1/5; 1/4 1/4 1/5; 1/5 1/5 1/5]
C = [1 3 5; 7 9 11; 13 15 17]
D = sym('[a b c; d e f; g h i]')              % (ST)
```

```
% Ejercicios 1 a 4.

det(A), det(B), det(C)       % C es la única no invertible, porque det(C) = 0.
-(1/4)*det(B([2 3], [1 3]))+(1/4)*det(B(1 3], [1 3]))-(1/5)*det(B([1 2], [1 3]))
det(A^5)-det(A)^5
det(A)-det(A.')
det(A*B)-det(A)*det(B)
det(5*(-125*det(B)
det(inv(B))-1/det(B)
det(B^(-2))-1/det(B)^2

% Ejercicio 5.

determ(D)                                  % (ST)
determ(transpose(D))                       % (ST)
d1=determ(symmul(transpose(D), D))         % (ST)
d2=expand(sympow(determ(D), 2))            % (ST) Primero elevar al cuadrado y después desarrollar
                                             simbólicamente.
symsub(d1, d2)                             % (ST) La diferencia es cero.

% Ejercicio 6.

% Crear un archivo original llamado m.m que tenga los renglones siguientes:
function a = m(n)
for i=1:n;
  for j=1:n,
    if i==j
      a(i, j)=i;
      else a(i, j)=1;
    end
  end
end
% Después teclear
det(m(2)), det(m(3)), det(m(4)), det(m5))     % Etcétera..
% Patrón: det(M - n) = (n - 1)!

% Ejercicio 7 - Comentario.

% (S^T)S es invertible sólo si las columnas de S son linealmente independientes.

% Ejercicio 8.

A=[-85 -55 -37; -35 97 50; 79 56 49]
b=[-306; 309; 338]
sol=A\b                  % (a)
A1=[b A(:, 2:3)]         % columna b y columnas 2 y 3 de A
A2=[A(:, 1) b A(:, 3)]   % columna 1 de A columna b columna 3 de A
A3=[A(:, 1:2) b]         % columnas 2 y 3 de A y columna b
x=det(A1)/det(A)         % (b)
y=det(A2)/det(A)         % (b)
z=det(A3)/det(A)         % (b)
adj=det(A)*inv(A)        % (c)

% Ejercicio 9.

% En un archivo llamado CramerD.m, teclee y guarde el programa:
```

```
function [B] = CramerD (A, b, i)
                        B = [A(:, 1:i-1) b A(:, i+1:length(A))];
                        end
% En un archivo llamado CramerS.m, teclee y guarde el programa:
function [B] = CramerS (A, b, i)
                        B = det([A(:, 1:i-1) b A(:, i+1:length(A))])/det(A);
                        end
% Después, en sesión de MATLAB, teclee:
CramerD(A, b, 1), CramerD(A, b, 2), CramerD(A, b, 3)
CramerS(A, b, 1), CramerS(A, b, 2), CramerS(A, b, 3)

% Ejercicio 12.

randperm(4)                         % Una permutación aleatoria de {1,2,3,4}.

% Ejercicios 16 y 17.

% Crear un archivo original llamado «circ.m» que contenga:
function [a] = circ(x, y, x1, y1, x2, y2, x3, y3)   % x,y se usan como
             a = det([x^2+y^2 x y 1;                % argumentos de la función
                 x1^2+y1^2 x1 y1 1;                 % porque MATLAB no
                 x2^2+y2^2 x2 y2 1;                 % los acepta como
                 x3^2+y3^2 x3 y3 1]);               % variables simbólicas
% después regrese a su sesión de MATLAB y teclee
circ(-2, 2, -1, 1, 1, 1, 2, 4)     % evaluación de la función
circ(-2, 4, -1, 1, 1, 1, 2, 4)     % en los cuatro puntos
circ(-1, 2, -1, 1, 1, 1, 2, 4)     % (-1,2) no está en el círculo, porque
                                   % la respuesta no es cero.
% Ejercicio 18. Calcular el det. de la matriz de Sylvester tecleada manualmente.
det([1 -3 - 63 -85 150 0 0 0; 0 1 -3 -63 -85 150 0 0; 0 0 1 -3 -63 -85 150 0; ...
0 0 0 1 -3 -63 -85 150; 1 22 -103 -2740 5700 0 0 0; 0 1 22 -103 -2740 5700 0 0;...
0 0 1 22 -103 -2740 5700 0; 0 0 0 1 22 -103 -2740 5700])     %%%% Etcétera..
```

Eigenvalores y eigenvectores[1]

Ninguna investigación humana se puede calificar como realmente científica si no puede ser demostrada matemáticamente.
Leonardo Da Vinci (1452-1519)

Introducción

Los eigenvalores y eigenvectores pertenecen a los temas de mayor utilidad del álgebra lineal. Se usan en varias áreas de las matemáticas, física, mecánica, ingeniería eléctrica y nuclear, hidrodinámica y aerodinámica, etc. De hecho, es raro encontrar un área de la ciencia aplicada donde nunca se hayan usado.

Puede parecer muy extraño, pero los eigenvalores de las matrices aparecieron publicados antes que las matrices. Esto se debe al hecho insólito de que, parafraseando a Cayley, la teoría de las matrices estaba bien desarrollada (a través de la teoría de los determinantes) antes de que siquiera se definieran las matrices. Según Morris Kline,[2] los eigenvalores se originaron en el contexto de formas cuadráticas y en la mecánica celeste (el movimiento de los planetas), conociéndose como *raíces características de la ecuación escalar*. Desde aproximadamente 1740, Euler usaba de manera implícita los eigenvalores para describir geométricamente las formas cuadráticas en tres variables. Estas funciones (que estudiaremos en el capítulo 8) son de la forma

$$q(x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

En la década de 1760, Lagrange estudió un sistema de seis ecuaciones diferenciales del movimiento de los planetas (sólo se conocían seis) y de ahí dedujo una ecuación polinomial de sexto grado, cuyas raíces eran los eigenvalores de una matriz 6×6 (en la sección 7.1 se explican las definiciones). En 1820, Cauchy se dio cuenta de la importancia de los eigenvalores para determinar los "ejes principales" de una forma cuadrática con n variables. También aplicó sus descubrimientos a la teoría del movimiento planetario. Fue Cauchy quien, en 1840, usó por primera vez los términos **valores característicos** y **ecuación característica** para indicar los eigenvalores y la ecuación polinomial básica que satisfacen.

---

[1] N. del T.: Sin embargo, este término también tiene amplio uso como "valores propios" y "vectores propios".

[2] En *Mathematical Thought from Ancient to Modern Times* (Fair Lawn, N. J.: Oxford University Press, 1972).

CONVENCIÓN

A menos que se indique otra cosa, todas las matrices de este capítulo serán cuadradas.

7.1 Eigenvalores y eigenvectores

Objetivo del estudiante para esta sección

Definir y aprender a calcular los eigenvalores y los eigenvectores

Si A es una matriz $n \times n$, $A\mathbf{v}$ no se relaciona por lo general con el vector n $\mathbf{v}$. Un caso muy interesante se da cuando $A\mathbf{v}$ resulta ser proporcional (paralelo) a $\mathbf{v}$. Así, geométricamente, $\mathbf{v}$ y $A\mathbf{v}$ están en la misma recta que pasa por el origen. En ese caso, se dice que $\mathbf{v}$ es un *eigenvector*, o *vector propio* de A, y que la constante de proporcionalidad es un *eigenvalor* o *valor propio* de A.

DEFINICIÓN

Sea A una matriz $n \times n$. Un vector $\mathbf{v}$ distinto de cero es un **eigenvector** de A si para cierto escalar λ,

$$A\mathbf{v} = \lambda\mathbf{v} \tag{7.1}$$

El escalar λ (que puede ser cero) se llama **eigenvalor** de A *correspondiente a* (o *asociado con*) el eigenvector $\mathbf{v}$. Los eigenvalores también se conocen como **valores característicos**, o **valores propios** (*eigen* en alemán quiere decir *propio*), y también **raíces latentes**.

■ EJEMPLO 1 Sea

$$A = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}$$

Demuestre que $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ son eigenvectores de A. ¿Cuáles son los eigenvalores correspondientes?

SOLUCIÓN Ya que

$$\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(2, 1) es un eigenvector cuyo eigenvalor correspondiente es $\lambda = 3$. También,

$$\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix} = -2\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Por consiguiente (1, −2) es un eigenvector cuyo eigenvalor correspondiente es $\lambda = -2$.

Observe que cualquier múltiplo escalar de un eigenvector **v** también es un eigenvector, porque si $\mathbf{w} = r\mathbf{v}$, entonces

$$\begin{aligned} A\mathbf{w} &= A(r\mathbf{v}) = rA\mathbf{v} \\ &= r(\lambda\mathbf{v}) = \lambda(r\mathbf{v}) \\ &= \lambda\mathbf{w} \end{aligned}$$

Además, **v** y **w** tienen el mismo eigenvalor.

Así, en el ejemplo 1 vemos que los vectores distintos de cero en las líneas l_1 y l_2 determinados por (2, 1) y (1, − 2) son eigenvectores de A. La transformación lineal $A\mathbf{x}$ **estira** los vectores de l_1 en un factor $\lambda = 3$. Los vectores a lo largo de l_2 se reflejan respecto al origen y después se estiran en un factor de 2 (figura. 7.1).

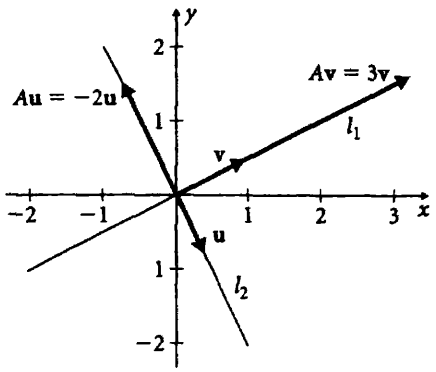


Figura 7.1

■ EJEMPLO 2 Determine geométricamente los eigenvalores y eigenvectores de A en cada caso.

(a) $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (*Sugerencia:* Recuerde que $A\mathbf{x}$ es la reflexión de **x** respecto a la recta $y = x$.)

(b) A es la matriz estándar de la rotación de 30° en $\mathbf{R}^3$ en torno al eje z, en dirección positiva.

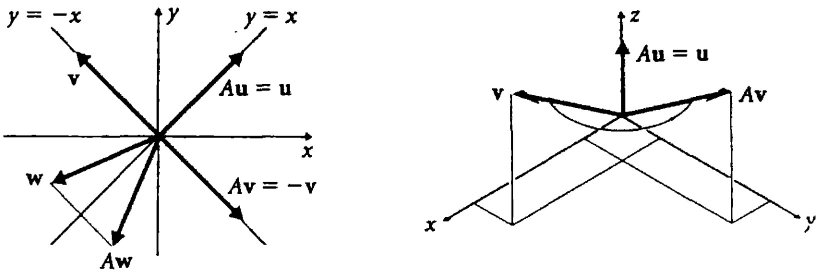
SOLUCIÓN

(a) Los únicos vectores que permanecen en la misma recta después de la rotación son todos aquellos que se encuentran en las rectas $y = x$ y $y = -x$. Sin considerar al origen, éstos son los únicos eigenvectores. Para **v** a lo largo de $y = x$ se cumple $A\mathbf{v} = 1\mathbf{v}$, así que **v** es un eigenvector con su correspondiente eigenvalor igual a 1. Cuando **v** se localiza en la recta $y = -x$, tenemos que $A\mathbf{v} = -1\mathbf{v}$, en consecuencia **v** es un eigenvector cuyo eigenvalor correspondiente es 1 (figura. 7.2(a)).

(b) Los únicos vectores que permanecen en la misma recta después de la rotación son todos los que se encuentran en el eje z (figura 7.2(b)). Éstos, sin considerar al origen, son los únicos eigenvectores. El eigenvalor correspondiente es 1.

OBSERVACIONES

1. El ejemplo 2 nos indica que los eigenvalores y los eigenvectores se *relacionan muy estrechamente con las transformaciones lineales*. En el inciso (b) ni siquiera tuvimos que mostrar la matriz para poder calcular los eigenvectores y eigenvalores. Fue suficiente una interpretación geométrica de la transformación matricial correspondiente.
2. Hay limitaciones serias cuando se trata de determinar geométricamente los eigenvalores y eigenvectores. Si el tamaño de A es mayor que 3, nos abandona nuestra intuición geométrica. Además, hay muchas transformaciones tridimensionales que son demasiado complicadas para poder explicarse geométricamente.



(a) Reflexión respecto a la recta $y = x$. (b) Rotación en torno al eje z.

Figura 7.2 Visualización de eigenvectores a través de la transformación lineal.

Cálculo de los eigenvalores y los eigenvectores

Ahora describiremos cómo calcular los eigenvalores y eigenvectores en general. Ya que

$$\begin{aligned} A\mathbf{v} = \lambda\mathbf{v} &\Rightarrow A\mathbf{v} = \lambda I\mathbf{v} \\ &\Rightarrow A\mathbf{v} - \lambda I\mathbf{v} = \mathbf{0} \\ &\Rightarrow (A - \lambda I)\mathbf{v} = \mathbf{0} \end{aligned}$$

vemos que $\mathbf{v}$ es un eigenvector si y sólo si es una solución no trivial del sistema homogéneo $(A - \lambda I)\mathbf{v} = \mathbf{0}$. En este caso, $\mathbf{v}$ es un vector distinto de cero llamado *espacio nulo* de $A - \lambda I$. El sistema tendrá una solución no trivial si y sólo si el determinante de la matriz de coeficientes es cero. Así, λ es un eigenvalor de A si y sólo si el $\det(A - \lambda I) = 0$. Hemos demostrado el teorema siguiente:

TEOREMA 1

Sea A una matriz cuadrada.

1. Un escalar λ es un eigenvalor de A si y sólo si

$$\det(A - \lambda I) = 0 \tag{7.2}$$

2. Un vector $\mathbf{v}$ es un eigenvector de A correspondiente a un eigenvalor λ si y sólo si $\mathbf{v}$ es una solución no trivial del sistema

$$(A - \lambda I)\mathbf{v} = \mathbf{0} \tag{7.3}$$

A la ecuación (7.2) se le llama **ecuación característica** de A, y *es una de las más importantes en todas las matemáticas.* También, el $\det(A - \lambda I)$ es un polinomio de grado n en λ, y se llama **polinomio característico** de A. La matriz $A - \lambda I$ se llama **matriz característica** de A. Si un eigenvalor λ es una raíz de multiplicidad k del polinomio característico, se dice que λ tiene **multiplicidad algebraica** igual a k. El espacio nulo de $A - \lambda I$, representado por E_λ, se llama **eigenespacio** de A correspondiente al eigenvalor λ.

$$E_\lambda = \nu(A \quad \lambda I)$$

Así, E_λ es un *subespacio* de $\mathbf{R}^n$, y está formado por eigenvectores de A y por el vector cero. La dimensión de E_λ se llama **multiplicidad geométrica** de λ. Es un hecho básico que *la multiplicidad geométrica nunca es mayor que la algebraica* (véase el ejercicio 38).

Según el teorema 1 y las observaciones precedentes, llegamos a lo siguiente.

Algoritmo

(Cálculo de eigenvalores, eigenvectores y bases de eigenespacios)

DATO: Matriz A de $n \times n$.

1. Calcular el polinomio característico $\det(A - \lambda I)$.
2. Determinar los eigenvalores de A despejando λ del $\det(A - \lambda I) = 0$.
3. Para cada eigenvalor λ_i, resolver completamente el sistema homogéneo $(A - \lambda_i)\mathbf{v} = \mathbf{0}$, reduciendo la matriz aumentada

$$[A - \lambda_i I : \mathbf{0}]$$

   Las soluciones no triviales son los eigenvectores de A que corresponden a λ_i.
4. Escribir la solución general de $(A - \lambda_i I)\mathbf{v} = \mathbf{0}$ en el paso 3 en forma de combinación lineal de vectores cuyos coeficientes son las variables libres. Esos vectores forman una base para E_{λ_i}.

RESULTADO: Eigenvalores $\lambda_1, \ldots, \lambda_n$ de A; una base de eigenvectores para cada E_{λ_i}.

En los ejemplos 3 a 7 deduciremos los eigenvalores y eigenvectores, y calcularemos las bases para cada eigenespacio de la matriz dada A.

■ EJEMPLO 3

$$A = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}$$

SOLUCIÓN La ecuación característica es

$$\det(A - \lambda I) = \begin{vmatrix} 5-\lambda & 1 \\ 1 & 5-\lambda \end{vmatrix} = \lambda^2 - 10\lambda + 24 = 0$$

Por consiguiente, los eigenvalores son

$$\lambda_1 = 4, \qquad \lambda_2 = 6$$

Con el fin de determinar los eigenvectores correspondientes haremos lo siguiente. Para $\lambda_1 = 4$, resolveremos el sistema con la matriz aumentada $[A - \lambda_1 I : \mathbf{0}]$.

$$[A - 4I : \mathbf{0}] = \begin{bmatrix} 1 & 1 & : & 0 \\ 1 & 1 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & : & 0 \\ 0 & 0 & : & 0 \end{bmatrix}$$

para obtener $(-r, r)$ para cualquier $r \in \mathbf{R}$. Así, todos los eigenvectores que corresponden a $\lambda_1 = 4$ tienen la forma $(-r, r)$, $r \neq 0$, y el eigenespacio E_4 es

$$E_4 = \left\{ \begin{bmatrix} -r \\ r \end{bmatrix}, r \in \mathbf{R} \right\} = \text{Gen} \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

El eigenvector $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ define la base $\{\mathbf{v}_1\}$ para E_4.

Para $\lambda_2 = 6$, se resuelve el sistema cuya matriz aumentada es $[A - \lambda_2 I : \mathbf{0}]$.

$$[A - 6I : \mathbf{0}] = \begin{bmatrix} -1 & 1 & : & 0 \\ 1 & -1 & : & 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 1 & : & 0 \\ 0 & 0 & : & 0 \end{bmatrix}$$

para obtener (r, r), para cualquier $r \in$ R. Por tanto, los eigenvectores correspondientes a $\lambda_2 = 6$ tienen la forma (r, r), $r \neq 0$, y el eigenespacio E_6 es

$$E_6 = \left\{ \begin{bmatrix} r \\ r \end{bmatrix}, r \in \mathbf{R} \right\} = \text{Gen} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

El eigenvector $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ define la base $\{\mathbf{v}_2\}$ para E_6.

Vemos que los eigenespacios son las rectas $y = -x$ y $y = x$. El efecto de la transformación lineal $A\mathbf{x}$ en esas dos rectas es el **estiramiento** en un factor de 4 a lo largo de $y = -x$, y de 6 a lo largo de $y = x$. □

■ EJEMPLO 4

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

SOLUCIÓN La ecuación característica es

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = -\lambda(1-\lambda)^2 = 0$$

Por consiguiente, los eigenvalores son

$$\lambda_1 = 0, \qquad \lambda_2 = \lambda_3 = 1$$

Para $\lambda_1 = 0$,

$$[A - 0I : \mathbf{0}] = \begin{bmatrix} 0 & 0 & 1 & : & 0 \\ 0 & 1 & 0 & : & 0 \\ 0 & 0 & 1 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & : & 0 \\ 0 & 0 & 1 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

La solución general es $(r, 0, 0)$ para $r \in \mathbf{R}$. En consecuencia,

$$E_0 = \left\{ \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix}, r \in \mathbf{R} \right\} = \text{Gen} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

y el eigenvector $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ define a la base $\{\mathbf{v}_1\}$ de E_0.

Cuando $\lambda_2 = \lambda_3 = 1$ (con multiplicidad algebraica 2), tenemos

$$[A - 1I : \mathbf{0}] = \begin{bmatrix} -1 & 0 & 1 & : & 0 \\ 0 & 0 & 0 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

cuya solución general es (r, s, r) para $r \in \mathbf{R}$. Como $(r, s, r) = r(1, 0, 1) + s(0, 1, 0)$,

$$E_1 = \left\{ \begin{bmatrix} r \\ s \\ r \end{bmatrix}, r \in \mathbf{R} \right\} = \text{Gen} \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Los eigenvectores generadores $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ también son linealmente independientes.

Por tanto, $\{\mathbf{v}_2, \mathbf{v}_3\}$ es una base para E_1. La multiplicidad geométrica de $\lambda = 1$ es 2.

■ EJEMPLO 5

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -4 & 2 \\ 0 & 0 & -2 \end{bmatrix}$$

SOLUCIÓN Para esta matriz,

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & -1 & 0 \\ 0 & -4-\lambda & 2 \\ 0 & 0 & -2-\lambda \end{vmatrix} = -(\lambda - 1)(\lambda + 2)(\lambda + 4) = 0$$

Por tanto, los eigenvalores son

$$\lambda_1 = 1, \qquad \lambda_2 = -2, \qquad \lambda_3 = -4$$

Al reducir $[A - 1I : \mathbf{0}]$, $[A - (-2)I : \mathbf{0}]$ y $[A - (-4)I : \mathbf{0}]$ con operaciones de renglón llegamos a

$$E_1 = \text{Gen} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \qquad E_{-2} = \text{Gen} \left\{ \begin{bmatrix} \frac{1}{3} \\ 1 \\ 1 \end{bmatrix} \right\}, \qquad E_{-4} = \text{Gen} \left\{ \begin{bmatrix} \frac{1}{5} \\ 1 \\ 0 \end{bmatrix} \right\}$$

Los eigenvectores generadores definen bases para los eigenespacios correspondientes.

■ EJEMPLO 6

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ -1 & 1 & -2 \end{bmatrix}$$

SOLUCIÓN En este caso,

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 0 & 3 \\ 1 & -1-\lambda & 2 \\ -1 & 1 & -2-\lambda \end{vmatrix} = -\lambda^3 - 2\lambda^2 = -\lambda^2(2+\lambda) = 0$$

Por consiguiente, los eigenvalores son

$$\lambda_1 = \lambda_2 = 0, \quad \lambda_3 = 2$$

Por reducción de $[A - 0I : \mathbf{0}]$ y $[A - (-2)I : \mathbf{0}]$ con operaciones de renglón se obtiene

$$E_0 = \text{Gen}\left\{\begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}\right\}, \qquad E_{-2} = \text{Gen}\left\{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\right\}$$

y los eigenvectores generadores definen bases para los eigenespacios correspondientes. Observe que aun cuando la multiplicidad algebraica de $\lambda = 0$ es 2, la multiplicidad geométrica sólo es 1.

■ EJEMPLO 7 (Eigenvalores complejos)

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

SOLUCIÓN No hay eigenvalores reales, porque

$$\begin{vmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 + 1 = 0 \Rightarrow \lambda = \pm i$$

Si se aceptan soluciones complejas, pueden calcularse los eigenvectores siguientes, que ahora son complejos

$$\begin{bmatrix} -i & -1 & : & 0 \\ 1 & -i & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -i & : & 0 \\ 0 & 0 & : & 0 \end{bmatrix}$$

$$\begin{bmatrix} i & -1 & : & 0 \\ 1 & i & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & i & : & 0 \\ 0 & 0 & : & 0 \end{bmatrix}$$

para obtener los eigenvectores básicos $(i, 1)$ y $(-i, 1)$ para $\lambda = i$ y $\lambda = -i$, respectivamente. También podemos escribir

$$E_i = \text{Gen}\left\{\begin{bmatrix} i \\ 1 \end{bmatrix}\right\}, \qquad E_{-i} = \text{Gen}\left\{\begin{bmatrix} -i \\ 1 \end{bmatrix}\right\}$$

en la que los escalares empleados en los generadores son números *complejos*.

Aunque las raíces de la ecuación característica pueden ser números complejos. nos interesan más las raíces reales, y en consecuencia los eigenvalores *reales*.

Un artificio (raro)

Por lo general es difícil o imposible resolver con exactitud la ecuación característica. Sin embargo, si el polinomio característico tiene coeficientes enteros, a veces se puede usar un truco que describiremos en el siguiente ejemplo.

■ EJEMPLO 8 Determine los eigenvalores de una matriz cuyo polinomio característico es

$$p(\lambda) = -\lambda^3 - 2\lambda^2 + 3\lambda + 6$$

SOLUCIÓN $p(\lambda)$ tiene coeficientes *enteros*. Si se confirma que tiene una raíz entera, esa raíz debe ser un divisor del término constante (es un teorema del álgebra). Los divisores de 6 son $\pm 1, \pm 2, \pm 3, \pm 6$. Si sustituimos $\lambda = -2$ en p, se obtiene cero. De aquí podemos deducir que -2 es una raíz, y que $\lambda + 2$ divide exactamente a p. Aplicando la división larga, vemos que

$$p(\lambda) = -(\lambda + 2)(\lambda^2 - 3)$$

Así,

$$\lambda_1 = -2, \lambda_2 = \sqrt{3}, \lambda_3 = -\sqrt{3}$$

son los eigenvalores. □

Algo acerca de los eigenvalores de matrices invertibles

Como una matriz cuadrada A es invertible si y sólo si

$$\det(A) \neq 0 \Leftrightarrow \det(A - 0I) \neq 0$$

se llega al teorema siguiente.

TEOREMA 2 Una matriz cuadrada A es invertible si y sólo si 0 **no es** un eigenvalor de A.

Eigenvalores de matrices triangulares

Si $A = [a_{ij}]$ es una matriz triangular, también lo es $A - \lambda I$. De manera que en este caso particular,

$$\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda)$$

TEOREMA 3 Los eigenvalores de una matriz triangular son sus elementos diagonales.

Cálculo rápido de $A^k\mathbf{x}$

Ahora describiremos una aplicación muy interesante que se usa en distintos problemas de matemáticas, física e ingeniería (principalmente a través de ecuaciones diferenciales y sistemas dinámicos).

Considere que necesitamos calcular la matriz producto $A^k\mathbf{x}$ para un vector n x y una matriz A $n \times n$. Esto puede ser tedioso, en especial para k o n grande. Ahora **supongamos** que puede escribirse en forma de una combinación lineal de eigenvectores $\mathbf{v}_1, \ldots, \mathbf{v}_m$ de A, por ejemplo

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m$$

Si $\lambda_1, \ldots, \lambda_m$ son los eigenvalores correspondientes, entonces

$$\begin{aligned} A\mathbf{x} &= A(c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m) \\ &= c_1A\mathbf{v}_1 + \cdots + c_mA\mathbf{v}_m \\ &= c_1\lambda_1\mathbf{v}_1 + \cdots + c_m\lambda_m\mathbf{v}_m \end{aligned}$$

Repetimos este proceso empleando $A\mathbf{x}$ en lugar de $\mathbf{x}$ e iteramos $k-1$ veces para obtener

$$A^k\mathbf{x} = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_m\lambda_m^k\mathbf{v}_m \tag{7.4}$$

Observamos que una vez conocidas las c_i, las λ_i y las $\mathbf{v}_i$, se facilita mucho el cálculo de $A^k\mathbf{x}$, porque el lado derecho de la ecuación (7.4) no implica multiplicaciones de matrices.

■ EJEMPLO 9 Calcule

$$\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}^8 \begin{bmatrix} -6 \\ -8 \end{bmatrix}$$

SOLUCIÓN (1, −2) y (2, 1) son los eigenvectores de la matriz, y los eigenvalores son −2 y 3. Entonces

$$\begin{bmatrix} -6 \\ -8 \end{bmatrix} = 2\begin{bmatrix} 1 \\ -2 \end{bmatrix} - 4\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Por tanto, de acuerdo con la ecuación (7.4), el producto anterior es igual a

$$2 \cdot (-2)^8 \begin{bmatrix} 1 \\ -2 \end{bmatrix} - 4 \cdot 3^8 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -51\,976 \\ -27\,268 \end{bmatrix}$$

□

PUNTO DÉBIL DEL MÉTODO Este método sólo se aplica si $\mathbf{x}$ puede expresarse en forma de una combinación lineal de eigenvectores de A. Esta hipótesis puede no cumplirse para algunas A y $\mathbf{x}$.

Eigenvalores de transformaciones lineales

También es posible definir los eigenvalores y eigenvectores para las transformaciones lineales. Si V es un espacio vectorial y $T: V \to V$ es una transformación lineal, entonces un vector $\mathbf{v}$ distinto de cero es un **eigenvector** de T si

$$T(\mathbf{v}) = \lambda\mathbf{v}$$

para algunos escalares λ (posiblemente cero). Como antes, λ se llama **eigenvalor** de T y se dice que el eigenvector $\mathbf{v}$ pertenece a (o corresponde a, o está asociado con) λ.

■ EJEMPLO 10 Si V es cualquier espacio vectorial y si r es un escalar fijo, determine los eigenvalores y eigenvectores de la homotecia

$$T: V \to V, \qquad T(\mathbf{v}) = r\mathbf{v}$$

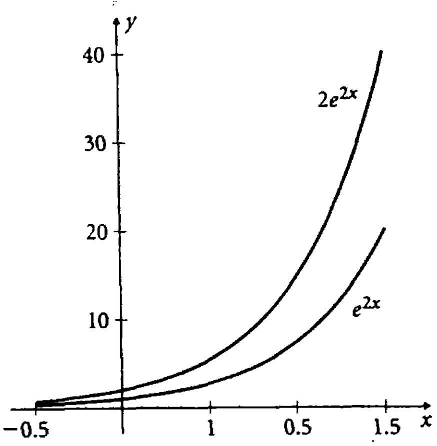
SOLUCIÓN Como $T(\mathbf{v}) = r\mathbf{v}$, todo vector distinto de cero es un eigenvector, cuyo eigenvalor correspondiente es r.

■ EJEMPLO 11 (Se requiere cálculo) Sea $V = C^1[a, b]$ un espacio vectorial de todas las funciones de valor real de x, definidas en $[a, b]$ y diferenciables. El operador diferenciación $\frac{d}{dx}: V \to V$ es una transformación lineal. Si r es un escalar fijo, la función e^{rx} está en V. Demuestre que e^{rx} es un eigenvector de $\frac{d}{dx}$. Calcule el eigenvalor correspondiente.

SOLUCIÓN Puesto que

$$\frac{d}{dx}(e^{rx}) = re^{rx}$$

en ese caso e^{rx} es un eigenvector de $\frac{d}{dx}$ y r es el eigenvalor correspondiente.



e^{2x} es un eigenvector de $\frac{d}{dx}$ con eigenvalor igual a 2.

■ EJEMPLO 12 Determine los eigenvalores y eigenvectores de

$$T: \mathbf{R}^2 \to \mathbf{R}^2, T(x, y) = (x + y, x + y)$$

SOLUCIÓN Se trata de buscar números λ y vectores (x, y) distintos de cero tales que

$$\begin{bmatrix} x + y \\ x + y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$$

o que

$$\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \lambda I\right)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Así, el problema se reduce a encontrar los eigenvectores y eigenvalores de $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, que es la matriz estándar de T. Vemos ahora que 0 y 2 son los únicos eigenvalores, y que $(-1, 1)$ y $(1, 1)$ son los eigenvectores correspondientes. $\square$

El teorema siguiente nos ayudará a generalizar.

TEOREMA 4

Sea V un espacio vectorial de dimensiones finitas. Sea $T: V \to V$ una transformación lineal cuya matriz es A, con respecto a una base $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V. Entonces

$$T(\mathbf{v}) = \lambda\mathbf{v} \Leftrightarrow A[\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}} \tag{7.5}$$

Por consiguiente,

1. λ es un eigenvalor de T si y sólo si es un eigenvalor de A;
2. $\mathbf{v}$ es un eigenvector de T si y sólo si $[\mathbf{v}]_{\mathcal{B}}$ es un eigenvector de A.

DEMOSTRACIÓN De acuerdo con el teorema 13 de la sección 5.4,

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}}$$

También sabemos que $[\lambda\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}}$, según el teorema 18 de la sección 4.5. Por tanto,

$$\begin{aligned} T(\mathbf{v}) = \lambda\mathbf{v} &\Leftrightarrow [T(\mathbf{v})]_{\mathcal{B}} = [\lambda\mathbf{v}]_{\mathcal{B}} \\ &\Leftrightarrow A[\mathbf{v}]_{\mathcal{B}} = [\lambda\mathbf{v}]_{\mathcal{B}} \\ &\Leftrightarrow A[\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}} \end{aligned}$$

Esto demuestra la ecuación (7.5), cuya consecuencia son las afirmaciones 1 y 2.

Eigenvalores y eigenvectores con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> eigenvects([[3,3,3],[2,2,2],[1,1,1]]);
```

$$[6, 1, \{[3, 2, 1]\}], [0, 2, \{[-1, 0, 1], [-1, 1, 0]\}]$$

Mathematica

```
In[1]:= Eigensystem[{{3,3,3},{2,2,2},{1,1,1}}]

Out[1]= {{0, 0, 6},
        {{-1, 0, 1}, {-1, 1, 0}, {3, 2, 1}}}
```

MATLAB

```
>> [V,D]=eig([3 3 3; 2 2 2; 1 1 1])
V =
    0.8018    0.8018   -0.7193
    0.5345   -0.5345    0.0250
    0.2673   -0.2673    0.6943
D =
    6.0000         0         0
         0    0.0000         0
         0         0         0
```

Ejercicios 7.1

Sean

$$A = \begin{bmatrix} 3 & -2 \\ -3 & 2 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

1. Demuestre que **u** es un eigenvector de A. ¿Cuál es el eigenvalor correspondiente?

2. Compruebe que cualquier múltiplo distinto de cero, $b\mathbf{u}$, de **u** por un escalar, también es un eigenvector de A.

3. Demuestre que **v** es un eigenvector de A. ¿Cuál es el eigenvalor correspondiente?

4. ¿Por qué $c\mathbf{v}$ es un eigenvector de A para todo escalar $c \neq 0$?

5. ¿Es $\mathbf{u} + \mathbf{v}$ un eigenvector de A?

6. ¿Cuál es el error en el razonamiento siguiente? Como $b\mathbf{u}$ y $c\mathbf{v}$ son eigenvectores de A, también lo es $b\mathbf{u} + c\mathbf{v}$. De ahí que todo vector $2 \neq \mathbf{0}$ es un eigenvector de A.

7. Si A se reduce a su forma escalonada mediante operaciones de renglón, ¿siguen siendo eigenvectores **u** y **v**?

En los ejercicios 8 y 9 determine cuáles de los vectores dados son eigenvectores de la matriz $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$. Si un vector es eigenvector, calcule el eigenvalor correspondiente.

8. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

9. $\begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix}$

10. ¿Cuál(es) de $\mathbf{e}_1$ y $\mathbf{e}_2$ es (son) eigenvector(es) de A?

$$A = \begin{bmatrix} 7 & 0 & 0 \\ 0 & -2 & 4 \\ 0 & 6 & 0 \end{bmatrix}$$

11. Demuestre que $-6, 4, 7$ son los eigenvalores de A en el ejercicio 10.

En los ejercicios 12 a 15 determine *geométricamente* los eigenvalores y eigenvectores de cada matriz.

12. $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (reflexión con respecto al eje x).

13. $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (proyección sobre el eje x).

14. $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ (estiramiento en un factor de 3, alejándose del origen).

15. $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (rotación de 90° respecto al origen).

Para las matrices de los ejercicios 16 a 23, determine lo siguiente.

(a) El polinomio característico.
(b) Los eigenvalores.
(c) Bases de los eigenvectores para todos los eigenespacios.
(d) La multiplicidad algebraica y geométrica de cada eigenvalor.

16. a. $\begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix}$ b. $\begin{bmatrix} 3 & 6 \\ 9 & 0 \end{bmatrix}$ c. $\begin{bmatrix} 1 & 11 \\ 11 & 1 \end{bmatrix}$

17. a. $\begin{bmatrix} -2 & 4 \\ 6 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 0 & -9 \\ 1 & -6 \end{bmatrix}$ c. $\begin{bmatrix} 0 & 7 \\ 5 & 2 \end{bmatrix}$

18. a. $\begin{bmatrix} -2 & 17 \\ 17 & -2 \end{bmatrix}$ b. $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$

19. a. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

20. a. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 4 & 0 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

21. a. $\begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

22. a. $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$ b. $\begin{bmatrix} 0 & -1 & 0 \\ -4 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

23. a. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ -3 & 0 & 3 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ -4 & 0 & 3 \end{bmatrix}$

24. Encuentre una matriz 3×3 con tres eigenvalores distintos.

25. Determine una matriz 3×3 que sólo tenga dos eigenvalores distintos.

En los ejercicios 26 y 27 encuentre, sin hacer cálculos, los eigenvalores de las matrices.

26. a. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ b. $\begin{bmatrix} -3 & 1 \\ 0 & 13 \end{bmatrix}$

!7. a. $\begin{bmatrix} 1 & 0 \\ 7 & -2 \end{bmatrix}$ b. $\begin{bmatrix} 0 & 8 \\ 2 & 0 \end{bmatrix}$

:n los ejercicios 28 y 29 obtenga, sin hacer cálculos, un igenvector y el eigenvalor correspondiente para la matriz dada.

8. $\begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$

9. $\begin{bmatrix} a & b & c \\ a & b & c \\ a & b & c \end{bmatrix}$

). ¿Por qué una matriz $n \times n$ no puede tener más de n eigenvalores distintos?

l. Compruebe que cualquier matriz cuadrada A y su transpuesta A^T tienen el mismo polinomio característico. Llegue a la conclusión de que tienen los mismos eigenvalores.

:. Demuestre que una matriz A $n \times n$ tiene un eigenvalor si y sólo si $\nu(A) \neq \{\mathbf{0}\}$. En este caso, demuestre que $\nu(A) = E_0$.

33. Sea $\mathbf{v}$ un eigenvector de la matriz A cuyo eigenvalor correspondiente es igual a 2. Determine una solución del sistema $A\mathbf{x} = \mathbf{v}$.

Sea $\mathbf{v}$ un eigenvector de A, con su correspondiente eigenvalor λ.

34. **(Potencia)** Demuestre que $\mathbf{v}$ también es un eigenvector de A^k con el correspondiente eigenvalor λ^k.

35. **(Inversa)** Si A es invertible, demuestre que $\mathbf{v}$ también es un eigenvector de A^{-1}, con el eigenvalor correspondiente λ^{-1}.

36. **(Desplazamiento del origen)** Si $\mathbf{v}$ es un eigenvector de la matriz A, con el eigenvalor λ, y si c es cualquier escalar compruebe que $\mathbf{v}$ es un eigenvector de $A - cI$ con el eigenvalor correspondiente $\lambda - c$.

37. **(Matrices semejantes)** Sean A y B matrices $n \times n$ semejantes entre sí, de tal manera que hay una matriz P invertible tal que $P^{-1}AP = B$. Demuestre lo siguiente.

   a. A y B tienen el mismo polinomio característico.

   b. A y b tienen los mismos eigenvalores.

   c. Si $\mathbf{v}$ es un eigenvector de B con eigenvalor λ, entonces $P\mathbf{v}$ es un eigenvector de A con eigenvalor λ.

   d. Si $\mathbf{u}$ es un eigenvector de A con eigenvalor λ, entonces $P^{-1}\mathbf{u}$ es un eigenvector de B con eigenvalor λ.

38. Sea r un eigenvalor de una matriz $n \times n$. Demuestre que la multiplicidad geométrica de r es menor que o igual a la multiplicidad algebraica de r. (*Sugerencia:* Amplíe una base de eigenvectores de r a una base de $\mathbf{R}^n$. Si A' es la matriz de $T(\mathbf{x}) = A\mathbf{x}$ relativa a esa base, entonces $A = P^{-1}A'P$ para una matriz invertible P. Aplique los resultados del ejercicio 37.)

39. Sea A una matriz cuadrada, y $\mathbf{v} \in \nu(A)$, tal que $\mathbf{v} \neq \mathbf{0}$. Compruebe que $\mathbf{v}$ es un eigenvector de A. ¿Cuál es su eigenvalor?

40. Determine las matrices A y B tales que los eigenvalores de $A + B$ no sean las sumas de los eigenvalores de A y B.

41. Encuentre las matrices A y B tales que los eigenvalores de AB no sean los productos de los eigenvalores de A y B.

42. Demuestre que los vectores n de base estándar $\mathbf{e}_1, \ldots, \mathbf{e}_n$ son eigenvectores de cualquier matriz diagonal A $n \times n$. ¿Cuáles son los eigenvalores correspondientes?

43. Suponga que una matriz A $n \times n$ tiene a todo vector n distinto de cero como eigenvector. Compruebe que A es una matriz escalar.

44. Demuestre que si A es **nilpotente** (es decir, que si $A^k = \mathbf{0}$ para un entero positivo k), entonces su único eigenvalor es 0. (*Su-

gerencia: Considere a v distinto de cero tal que $A\mathbf{v} = \lambda\mathbf{v}$. Si $\lambda \neq 0$, entonces $A\mathbf{v} \neq \mathbf{0}$. Por consiguiente, $A^2\mathbf{v} = \lambda A\mathbf{v} \neq \mathbf{0}$. Continúe.)

45. Compruebe que si A es nilpotente (refiérase al ejercicio 44), la multiplicidad geométrica de 0 es igual a la nulidad de A.

La **traza**, tr(A), de una matriz cuadrada $A = [a_{ij}]$ es la suma de sus elementos diagonales:

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

46. Sean $\lambda_1, \ldots, \lambda_n$ todos los eigenvalores (repetidos si son múltiples) de una matriz A $n \times n$. Demuestre que

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 \cdots + \lambda_n$$

y que

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$$

(*Sugerencia:* El $\det(A - \lambda I) = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$. Suponga $\lambda = 0$ para demostrar la segunda afirmación. Para la primera, observe que el coeficiente de λ^{n-1} se puede determinar a partir del producto $(a_{11} - \lambda)\cdots(a_{nn} - \lambda)$.)

Matriz asociada del polinomio

Sea $p(x)$ el polinomio

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

La matriz de $n \times n$ siguiente se llama **matriz asociada de** p:

$$C(p) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}$$

47. Determine la matriz asociada $C(p)$ de $p(x) = x^2 + 2x - 15$ y a continuación encuentre el polinomio característico de $C(p)$.

48. Compruebe que la matriz asociada $C(p)$ de $p(x) = x^2 + ax + b$ tiene el polinomio característico $\lambda^2 + a\lambda + b$.

49. Encuentre una matriz no diagonal cuyos eigenvalores son $4, -5$. (*Sugerencia*: Utilice el ejercicio 48.)

50. Compruebe que la matriz asociada $C(p)$ de $p(x) = x^3 + ax^2 + bx + c$ tiene el polinomio característico $-(\lambda^3 + a\lambda^2 + b\lambda + c)$. Demuestre también que para cualquier eigenvalor λ de $C(p)$, el vector $(1, \lambda, \lambda^2)$ es un eigenvector de $C(p)$.

51. Determine una matriz no triangular cuyos eigenvalores sean $4, -5$ y -2. (*Sugerencia:* Aplique los resultados del ejercicio 50.)

52. Encuentre una matriz cuyos eigenvectores sean $(1, 2, 4)$, $(1, 3, 9)$ y $(1, 4, 16)$. (*Sugerencia:* Considere la matriz asociada de $(x - 2)(x - 3)(x - 4)$.)

53. Demuestre (por inducción) que $C(p)$, la matriz asociada de $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ tiene el polinomio característico $(-1)^n p(\lambda)$.

Eigenvalores y transformaciones lineales

54. Determine los eigenvalores y eigenvectores de la proyección p de $\mathbf{R}^3$ sobre el plano xy.

55. Calcule los eigenvalores y eigenvectores de $T : P_2 \to P_2$, $T(a + bx) = b + ax$.

56. (**Se requiere cálculo**) Determine los eigenvalores y eigenvectores de la diferenciación $\frac{d}{dx} : P_2 \to P_2$.

7.2 Diagonalización

> **Objetivos del estudiante para esta sección**
>
> 1. Conocer cuáles matrices pueden diagonalizarse y cómo hacerlo.
> 2. Calcular A^n con eficiencia, si A puede diagonalizarse.
> 3. Aprender a diagonalizar una transformación lineal.

Las ideas y los métodos de esta sección son muy útiles en las ecuaciones diferenciales, en sistemas dinámicos, en procesos de Markov, en el estudio de curvas y superficies, en la teoría de las gráficas y en muchos otros campos.

Quienes trabajan con aritmética de matrices prefieren las matrices diagonales, porque es más fácil hacer cálculos con ellas. Esto es más notorio en la multiplicación matricial. Por ejemplo, una matriz diagonal D no "mezcla" los componentes de $\mathbf{x}$ en el producto $D\mathbf{x}$.

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2a \\ 3b \end{bmatrix}$$

y no mezcla los renglones de A en un producto DA (o columnas en AD),

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} 2a & 2b & 2c \\ 3d & 3e & 3f \end{bmatrix}$$

Además es muy fácil calcular las potencias D^k:

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}$$

Estudiaremos matrices que pueden transformarse en matrices diagonales, y trataremos de aprovechar la ventaja de la fácil aritmética. Desarrollaremos criterios para identificar esas matrices e investigar sus propiedades básicas. Todo esto a través de los eigenvalores y eigenvectores.

Diagonalización

Si una matriz $n \times n$ A es semejante a una matriz *diagonal* D se llama **diagonalizable**. También se dice que A **se puede diagonalizar**. Esto significa que existe una matriz $n \times n$ P tal que $P^{-1}AP$ es una matriz diagonal D.

$$P^{-1}AP = D$$

El proceso para determinar matrices P y D como las que se citan se llama **diagonalización**, y se dice que P y D **diagonalizan** a A.

■ EJEMPLO 13 Demuestre que

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

se diagonaliza con $P = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ y $D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

SOLUCIÓN Esto es cierto porque

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Vale la pena hacer notar que $P' = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ y $D' = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ también diagonalizan a A. (¿Por qué?) □

■ EJEMPLO 14 Compruebe que

$$B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

no es diagonalizable.

SOLUCIÓN Si B pudiera diagonalizarse con

$$P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad ad - cb \neq 0 \quad \text{y} \quad D = \begin{bmatrix} e & 0 \\ 0 & f \end{bmatrix}$$

entonces $PD = BP$. Así,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & 0 \\ 0 & f \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

es decir

$$\begin{bmatrix} ae & bf \\ ce & df \end{bmatrix} = \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix}$$

Por consiguiente, $ce = 0$. Si $c \neq 0$, entonces $e = 0$. Así, $ae = a0 = 0 = c$. Por tanto, c debe ser cero. Igualmente, $d = 0$. Pero si $ad - cb = 0$, entonces P sería no invertible. Llegamos a la conclusión que A no puede diagonalizarse.

En los ejemplos 13 y 14 vimos que

1. *No* todas las matrices cuadradas se pueden diagonalizar;
2. Las matrices P y D que diagonalizan a una matriz A *no son únicas*.

Diagonalización de una matriz cuadrada

Veamos ahora cuándo es diagonalizable una matriz cuadrada A, y cómo determinar las matrices P y D que la diagonalizan.

Primero, vale la pena observar que si D es una matriz diagonal con elementos diagonales $\lambda_1, \ldots, \lambda_n$, entonces, para $i = 1, \ldots, n$,

$$D\mathbf{e}_i = \lambda_i \mathbf{e}_i$$

En consecuencia, los vectores de base estándar $\mathbf{e}_1, \ldots, \mathbf{e}_n$ son eigenvectores de D. En particular, estos eigenvectores son *linealmente independientes*. Para generalizar aún más se cuenta con el teorema 5.

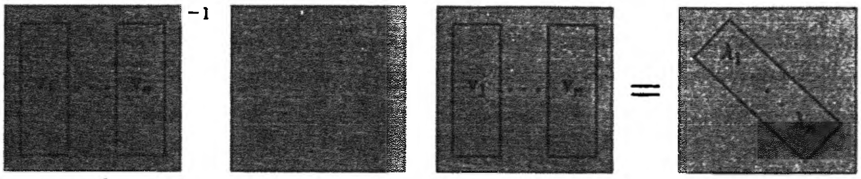
TEOREMA 5

(Criterio para la diagonalización)

Sea A una matriz de $n \times n$.

1. A es diagonalizable si y sólo si tiene n eigenvectores linealmente independientes.
2. Si A es diagonalizable con $P^{-1}AP = D$, entonces las columnas de P son eigenvectores de A y los elementos diagonales de D son los eigenvalores correspondientes.
3. Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ son eigenvectores linealmente independientes de A con los eigenvalores correspondientes $\lambda_1, \ldots, \lambda_n$, entonces A se puede diagonalizar con

$$P = [\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_n] \quad \text{y} \quad D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$



P^{-1} A P D

DEMOSTRACIÓN Sea P cualquier matriz cuyas columnas sean cualesquiera vectores-n, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y sea D cualquier matriz diagonal cuyos elementos diagonales son $\lambda_1, \ldots, \lambda_n$. Entonces

$$AP = A[\mathbf{v}_1\mathbf{v}_2 \cdots \mathbf{v}_n] = [A\mathbf{v}_1 A\mathbf{v}_2 \cdots A\mathbf{v}_n] \tag{7.6}$$

y

$$[\lambda_1\mathbf{v}_1\lambda_2\mathbf{v}_2 \cdots \lambda_n\mathbf{v}_n] = [\mathbf{v}_1\mathbf{v}_2 \cdots \mathbf{v}_n]\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = PD \tag{7.7}$$

Si A es diagonalizable con $P^{-1}AP = D$, entonces $AP = PD$. Por consiguiente, $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, $i = 1, \ldots, n$, de acuerdo con las ecuaciones (7.6) y (7.7). Así, las λ_i son eigenvalores y las $\mathbf{v}_i$ son los eigenvectores correspondientes. Esto demuestra la propiedad 2, y la implicación directa es la propiedad 1.

Supongamos que A tiene n eigenvectores linealmente independientes, digamos $\mathbf{v}_1, \ldots, \mathbf{v}_n$ (las columnas de P). Si $\lambda_1, \ldots, \lambda_n$ son los eigenvalores correspondientes, entonces $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, $i = 1, \ldots n$. Si D es diagonal y sus elementos diagonales son $\lambda_1, \ldots, \lambda_n$, en ese caso $AP = PD$, de acuerdo con las ecuaciones (7.6) y (7.7). Como P es cuadrada con columnas linealmente independientes, es invertible. Por tanto, $P^{-1}AP = D$ y A es diagonalizable. Esto demuestra la propiedad 3, y la implicación opuesta en la propiedad 1. □

TEOREMA 6 Una matriz A $n \times n$ es diagonalizable si y sólo si $\mathbf{R}^n$ tiene una base de eigenvectores de A.

DEMOSTRACIÓN Esta afirmación es cierta porque n vectores n linealmente independientes forman una base de $\mathbf{R}^n$. □

■ EJEMPLO 15 Haga de nuevo el ejemplo 14 aplicando el teorema 5.

SOLUCIÓN Podemos comprobar con facilidad que el único eigenvalor de B es 0, y $E_0 = \text{Gen}\{(1, 0)\}$. B no tiene dos eigenvectores linealmente independientes y, según la propiedad 1 del teorema 5, no es diagonalizable. (Observe lo fácil que fue esta vez llegar a la misma conclusión.) □

En los ejemplos 16 a 18 determinaremos si la matriz dada es diagonalizable. Si lo es, calcularemos P y D para diagonalizarla.

■ EJEMPLO 16

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

SOLUCIÓN En el ejemplo 4, de la sección 7.1 vimos que

$$\lambda_1 = 0, \quad \lambda_2 = \lambda_3 = 1$$

y

$$E_0 = \text{Gen}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}, \qquad E_1 = \text{Gen}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$$

A tiene tres eigenvectores linealmente independientes, de modo que es diagonalizable y podemos decir que

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ ☐

■ EJEMPLO 17

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -4 & 2 \\ 0 & 0 & -2 \end{bmatrix}$$

SOLUCIÓN En el ejemplo 5 de la sección 7.1 encontramos que

$$\lambda_1 = 1, \qquad \lambda_2 = -2, \qquad \lambda_3 = -4$$

y que

$$E_1 = \text{Gen}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}, \qquad E_{-2} = \text{Gen}\left\{\begin{bmatrix} \frac{1}{3} \\ 1 \\ 1 \end{bmatrix}\right\}, \qquad E_{-4} = \text{Gen}\left\{\begin{bmatrix} \frac{1}{5} \\ 1 \\ 0 \end{bmatrix}\right\}$$

La matriz A tiene tres eigenvectores linealmente independientes, por tanto es diagonalizable y

$$P = \begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{5} \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$ ☐

■ EJEMPLO 18

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 1 & -1 & 2 \\ -1 & 1 & -2 \end{bmatrix}$$

SOLUCIÓN En el ejemplo 6, de la sección 7.1, llegamos a

$$\lambda_1 = \lambda_2 = 0, \qquad \lambda_3 = -2$$

y

$$E_0 = \text{Gen}\left\{\begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}\right\}, \qquad E_{-2} = \text{Gen}\left\{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\right\}$$

Esta vez A tiene cuando mucho $2(<3)$ eigenvectores linealmente independientes, de modo que *no es* diagonalizable. □

TEOREMA 7

Sean $\lambda_1, \ldots, \lambda_n$ eigenvalores distintos cualesquiera de una matriz A $n \times n$.

1. Cualesquiera eigenvectores $\mathbf{v}_1, \ldots, \mathbf{v}_l$ correspondientes son linealmente independientes.
2. Si $\mathcal{B}_1, \ldots, \mathcal{B}_l$ son bases para los eigenespacios correspondientes, entonces $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_l$ es linealmente independiente.
3. Sea l la cantidad de *todos* los eigenvalores distintos de A. Entonces A es diagonalizable si y sólo si $\mathcal{B}$ (definido como en la propiedad 2) tiene exactamente n elementos.

DEMOSTRACIÓN

1. Si los $\mathbf{v}$ no son linealmente independientes, $\mathbf{v}_k$ es el primero que se puede escribir en forma de combinación lineal de los anteriores. Por consiguiente,

$$\mathbf{v}_k = a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1} \tag{7.8}$$

para escalares $a_1, \ldots, a_{k-1}$ no todos cero (ya que $\mathbf{v} \neq \mathbf{0}$ como un eigenvector) y linealmente independiente $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$. Multiplicamos por la izquierda por A para obtener

$$\begin{aligned} A\mathbf{v}_k &= A(a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1}) \\ &= a_1A\mathbf{v}_1 + \cdots + a_{k-1}A\mathbf{v}_{k-1} \end{aligned}$$

Por tanto,

$$\lambda_k\mathbf{v}_k = a_1\lambda_1\mathbf{v}_1 + \cdots + a_{k-1}\lambda_{k-1}\mathbf{v}_{k-1} \tag{7.9}$$

Si ahora multiplicamos la ecuación (7.8) por $-\lambda_k$ y la sumamos a la (7.9), tenemos

$$a_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_{k-1} = \mathbf{0}$$

En consecuencia, $a_1(\lambda_1 - \lambda_k) = 0, \ldots + a_{k-1}(\lambda_{k-1} - \lambda_k) = 0$, porque $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ son linealmente independientes. Una de las a, digamos a_i, es distinta de cero, así que $\lambda_i - \lambda_k = 0$, es decir, $\lambda_i = \lambda_k$, lo cual contradice la hipótesis de que las λ son distintas. Llegamos a la conclusión de que ninguno de los eigenvectores puede escribirse como combinación lineal de los anteriores, de modo que son linealmente independientes.

2. Para ahorrar escritura sólo consideraremos dos eigenvalores distintos, λ_1 y λ_2, y dos bases, $\mathcal{B}_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ y $\mathcal{B}_2 = \{\mathbf{w}_1, \ldots, \mathbf{w}_q\}$, para E_{λ_1} y E_{λ_2}. Para el caso general, la idea es la misma. Demostraremos que $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ es linealmente independiente. Sea

$$c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p + d_1\mathbf{w}_1 + \cdots + d_q\mathbf{w}_q = 0$$

Entonces $\mathbf{u} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ puede ser un eigenvector de λ_1, o cero. En el caso de $\mathbf{w} = d_1\mathbf{w}_1 + \cdots + d_q\mathbf{w}_q$ también puede obtenerse eigenvector de λ_2 o cero. Además, $\mathbf{u} + \mathbf{v} = \mathbf{0}$. Pero si $\mathbf{u}$ y $\mathbf{w}$ fueran ambos eigenvectores, deberían ser linealmente independientes, según la propiedad 1. Esto contradice que su suma es cero. Llegamos a la conclusión que

$$\mathbf{u} = \mathbf{w} = \mathbf{0}$$

Por consiguiente, $c_1 = \cdots = c_p = 0$ y $d_1 = \cdots = d_q = 0$, porque $\mathcal{B}_1$ y $\mathcal{B}_2$ son linealmente independientes; y por esa causa $\mathcal{B}$ también lo es.

3. De acuerdo con la propiedad 2, $\mathcal{B}$ tiene n vectores linealmente independientes. Luego, A es diagonalizable. En forma recíproca, si A es diagonalizable, entonces tiene n eigenvectores linealmente independientes. Si de esos eigenvectores hay exactamente n_i que corresponden al eigenvalor λ_i, entonces $\mathcal{B}_i$ tiene al menos n_i elementos, porque los eigenvectores son linealmente independientes. Nuestra conclusión es que $\mathcal{B}$ tiene cuando menos o exactamente n elementos.

Ahora podremos deducir algunos corolarios interesantes del teorema 7.

TEOREMA 8

Toda matriz A $n \times n$ con n eigenvalores distintos es diagonalizable.

DEMOSTRACIÓN Según el teorema 5, basta demostrar que los eigenvectores correspondientes son linealmente independientes. Pero esto lo garantiza la propiedad 1 del teorema 7.

PRECAUCIÓN *Una matriz diagonalizable no necesita tener eigenvalores distintos*, como vimos en el ejemplo 16.

TEOREMA 9

La matriz A es diagonalizable si y sólo si para cada eigenvalor λ, las multiplicidades geométrica y algebraica son iguales.

DEMOSTRACIÓN Se deja como ejercicio.

El teorema 7 nos permite usar el proceso de diagonalización siguiente que ya ilustraron los ejemplos 16 a 18. Esta vez no necesitaremos demostrar que $\mathcal{B}$ es linealmente independiente, porque el teorema lo garantiza.

Algoritmo

(Proceso de diagonalización)

DATOS: Matriz A $n \times n$.

1. Determinar las bases $\mathcal{B}_1, \ldots, \mathcal{B}_l$ para todos los eigenespacios de A. Formar la unión $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_l$.
2. Si $\mathcal{B}$ tiene menos de n elementos, detenerse: A *no es* diagonalizable.
3. Si $\mathcal{B}$ tiene n elementos, entonces A es diagonalizable.
4. A se puede diagonalizar con P, cuyas columnas son los elementos de $\mathcal{B}$; y D cuyos elementos diagonales son los eigenvalores correspondientes.

RESULTADOS: P y D diagonalizan A.

Potencias de matrices diagonalizables

Como ya sabemos, el cálculo de las potencias A^k puede ser bastante tedioso. Sin embargo, si A es diagonalizable y hemos calculado P y D, entonces $A = PDP^{-1}$, así que

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD^2P^{-1}$$

Iteramos para llegar a $A^k = PD^kP^{-1}$. Como el cálculo de D^k equivale a elevar sólo los elementos diagonales de D a la k-ésima potencia, vemos que A^k es fácil de obtener *dadas* P, P^{-1} y D.

Si sucede que A no es invertible, entonces 0 no es eigenvalor de A, de acuerdo con el teorema 2. Por consiguiente, D^{-1} existe y

$$A^{-1} = (PDP^{-1})^{-1} = PD^{-1}P^{-1}$$

De nuevo, podemos iterar para llegar a $A^{-k} = PD^{-k}P^{-1}$.

TEOREMA 10

Si A es diagonalizada por P y D, entonces, para $k = 0, 1, 2, \ldots$,

$$A^k = PD^kP^{-1} \tag{7.10}$$

Además, si A es invertible, entonces esta ecuación también es válida para $k = -1, -2, -3, \ldots$

■ **EJEMPLO 19** Encuentre una fórmula para A^k, $k = 0, 1, 2, \ldots$, donde

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}$$

SOLUCIÓN Los eigenvalores de A son 0, 2 y 4, y los eigenvectores básicos correspondientes $(-1, 0, 1)$, $(0, 1, 0)$ y $(1, 0, 3)$ son linealmente independientes. Por consiguiente,

$$A^k = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}^k \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 4^k \end{bmatrix} \begin{bmatrix} -\frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 1 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} \end{bmatrix}$$

Por tanto,

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}^k = \begin{bmatrix} 4^{k-1} & 0 & 4^{k-1} \\ 0 & 2^k & 0 \\ 3 \cdot 4^{k-1} & 0 & 3 \cdot 4^{k-1} \end{bmatrix}$$

Un cambio importante de variables

Ahora describiremos una idea que es la base de la mayor parte de las aplicaciones de la diagonalización. Sea A una matriz diagonalizada por P y Q. Con mucha frecuencia, una ecuación matricial

$$f(A, \mathbf{x}) = \mathbf{0}$$

se puede simplificar bastante si $\mathbf{x}$ se reemplaza por un nuevo vector y tal que

$$\mathbf{x} = P\mathbf{y}, \text{ es decir } \mathbf{y} = P^{-1}\mathbf{x} \tag{7.11}$$

y se reemplaza A con PDP^{-1} para obtener a una ecuación de la forma

$$g(D, \mathbf{y}) = \mathbf{0}$$

que involucra a la matriz diagonal D y al nuevo vector $\mathbf{y}$.

El cambio de variables ecuación (7.11) es *muy importante*. Entre otras cosas, se aplica a las ecuaciones diferenciales y a los sistemas dinámicos.

El ejemplo siguiente nos enseña *cómo* usar la ecuación (7.11), aunque esta vez el método no proporciona ventajas computacionales.

■ EJEMPLO 20 Determine (x, y) en el sistema, diagonalizando la matriz de los coeficientes.

$$\begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

SOLUCIÓN Sea A la matriz de los coeficientes, y sea $\mathbf{a} = (a, b)$. Entonces $D = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$, $P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$, y $P^{-1} = \frac{1}{2}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$.

Ahora consideremos el nuevo vector variable $\mathbf{y} = (x', y')$ definido por $\mathbf{y} = P\mathbf{x}$. En ese caso,

$$\begin{aligned} A\mathbf{x} = \mathbf{a} &\Leftrightarrow PA\mathbf{x} = P\mathbf{a} \\ &\Leftrightarrow PAP^{-1}\mathbf{y} = P\mathbf{a} \\ &\Leftrightarrow D\mathbf{y} = P\mathbf{a} \end{aligned}$$

La última ecuación es el sistema "diagonal"

$$\begin{bmatrix} 4x' \\ 6y' \end{bmatrix} = \begin{bmatrix} -a + b \\ a + b \end{bmatrix}$$

del cual es fácil despejar $\mathbf{y}$. Obtenemos $\mathbf{y} = (-\frac{1}{4}a + \frac{1}{4}b, \frac{1}{6}a + \frac{1}{6}b)$. Por consiguiente,

$$\mathbf{x} = P^{-1}\mathbf{y} = \frac{1}{24}\begin{bmatrix} 5a - b \\ -a + 5b \end{bmatrix}$$

□

Diagonalización de transformaciones lineales

La diagonalización se relaciona estrechamente con las transformaciones lineales. Sea A una matriz $n \times n$ diagonalizable, diagonalizada por P cuyas columnas son $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y por D, cuyos elementos diagonales son $\lambda_1, \ldots, \lambda_n$. Consideremos la transformación matricial T definida por A,

$$T : \mathbf{R}^n \to \mathbf{R}^n, \qquad T(\mathbf{x}) = A\mathbf{x}$$

De acuerdo con los teoremas 5 y 6 sabemos que las $\mathbf{v}_i$ son eigenvectores y que las λ_i son eigenvalores de A, y que además el conjunto $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ forma una base para $\mathbf{R}^n$. Según el teorema 17, de la sección 5.4, la matriz de T con respecto a esta nueva base $\mathcal{B}$ es

$$P^{-1}AP = D$$

puesto que P es la matriz de transición de $\mathcal{B}$ a la base estándar, de acuerdo con el teorema 20 de la sección 4.5. Así, *la matriz de T con respecto a una base de eigenvectores de A es diagonal.*

■ EJEMPLO 21 Si

$$A = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

Determine una base $\mathcal{B}$ de $\mathbf{R}^2$ tal que la matriz D de $T(\mathbf{x}) = A\mathbf{x}$ con respecto a su diagonal. También encuentre la matriz de transición P de $\mathcal{B}$ a la base estándar.

SOLUCIÓN Los eigenvalores de la matriz son 0.5 y 1.5, con eigenvectores $(-1, 1)$ y $(1, 1)$. En consecuencia

$$D = \begin{bmatrix} 0.5 & 0 \\ 0 & 1.5 \end{bmatrix} \quad \text{y} \quad P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

Si $T: V \to V$ es una transformación lineal de un espacio vectorial V de *dimensiones finitas*, es preferible tener una base $\mathcal{B}$ de V tal que la matriz D de T con respecto a $\mathcal{B}$ sea diagonal, porque en ese caso es más fácil calcular T. Si existe esa base, se dice que T es **diagonalizable**, y que $\mathcal{B}$ **diagonaliza a** T. Sucede que los vectores de $\mathcal{B}$ son eigenvectores de T, lo cual nos conduce al teorema siguiente (cuya demostración se deja como ejercicio).

TEOREMA 11

Sea $T: V \to V$ una transformación lineal de un espacio vectorial V de dimensiones finitas. Entonces

1. T es diagonalizable si y sólo si V tiene una base de eigenvectores de T;
2. T es diagonalizable si y sólo si la matriz de T con respecto a cualquier base de V es diagonalizable;
3. Si T es diagonalizada por $\mathcal{B}$, entonces los vectores de $\mathcal{B}$ son eigenvectores de T.

■ EJEMPLO 22 Demuestre que la transformación lineal $T: P_2 \to P_2$, $T(a + bx) = b + ax$ es diagonalizable. Determine una base $\mathcal{B}$ de P_2 que diagonalice a T. Evalúe T usando $\mathcal{B}$.

SOLUCIÓN Ya que

$$T(1) = 0 \cdot 1 + 1 \cdot x, \qquad T(x) = 1 \cdot 1 + 0 \cdot x$$

la matriz de T con respecto a la base estándar $\{1, x\}$ es $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Según el teorema 11, ¿por qué T y A son diagonalizables? Es posible obtener los eigenvectores de T y A partir de los de A (de acuerdo con el teorema 4, de la sección 7.1) y son $1 + x$ y $-1 + x$. Son linealmente independientes en P_2; en consecuencia, $\mathcal{B} = \{1 + x, -1 + x\}$ diagonaliza a T. Los eigenvalores correspondientes son 1 y -1. Podemos usar $\mathcal{B}$ para evaluar a T en la forma siguiente:

$$\begin{aligned} T(a + bx) &= T\left(\frac{a+b}{2}(1 + x) + \frac{b-a}{2}(-1 + x)\right) \\ &= \frac{a+b}{2}T(1 + x) + \frac{b-a}{2}T(-1 + x) \\ &= \frac{a+b}{2} \cdot 1 \cdot (1 + x) + \frac{b-a}{2} \cdot (-1) \cdot (-1 + x) \\ &= b + ax \end{aligned}$$

■ **EJEMPLO 23** (Se requiere cálculo) Compruebe que la diferenciación $\frac{d}{dx} : P_2 \to P_2$ no es diagonalizable.

SOLUCIÓN Como

$$\frac{d}{dx}(1) = 0 = 0 \cdot 1 + 0 \cdot x, \qquad \frac{d}{dx}(x) = 1 = 1 \cdot 1 + 0 \cdot x$$

la matriz de $\frac{d}{dx}$ con respecto a la base estándar $\{1, x\}$ es

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

la cual no es diagonalizable, por lo que $\frac{d}{dx}$ tampoco lo es, según el teorema 11. Por consiguiente, P_2 no tiene una base de eigenvectores $\frac{d}{dx}$.

Ejercicios 7.2

En los ejercicios 1 a 5 diagonalice la matriz si A ésta es diagonalizable; es decir, si es posible, encontrar P invertible y a D diagonal tales que $P^{-1}AP = D$.

1. $A = \begin{bmatrix} -2 & 5 \\ 5 & -2 \end{bmatrix}$

2. $A = \begin{bmatrix} -2 & 0 \\ 5 & -2 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -2 & 5 \\ 0 & 5 & -2 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 5 \\ 2 & 0 & -2 \end{bmatrix}$

5. $A = \begin{bmatrix} -4 & 0 & 0 \\ 0 & 2 & 1 \\ 2 & -1 & 0 \end{bmatrix}$.

En los ejercicios 6 a 9 compruebe que S es un conjunto linealmente independiente de eigenvectores de A. Diagonalice A usando S.

6. $S = \{(-2, 2), (5, 5)\}$ y $A = \begin{bmatrix} -3 & 6 \\ 6 & -3 \end{bmatrix}$

7. $S = \{(10, 0), (6, 5)\}$ y $A = \begin{bmatrix} -1 & 6 \\ 0 & 4 \end{bmatrix}$

8. $S = \{(-2, 0, 0), (0, -3, 3), (0, 2, 3)\}$ y

$$A = \begin{bmatrix} 7 & 0 & 0 \\ 0 & -2 & 4 \\ 0 & 6 & 0 \end{bmatrix}$$

9. $S = \{(0, 2, 0), (1, 0, 2), (1, 0, -2)\}$ y

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 4 & 0 & 0 \end{bmatrix}$$

En los ejercicios 10 a 14, S es un conjunto linealmente independiente de eigenvectores de A, y E es el conjunto de los eigenvalores correspondientes. Determine A.

10. $S = \{(-1,1),(1,1)\}, E = \{-10,12\}$

11. $S = \{(1,0,0),(1,1,0),(0,0,1)\}, E = \{1,2,3\}$

12. $S = \{(0,1,0),(1,0,1),(-1,0,1)\}, E = \{1,1,-1\}$

13. $S = \{(1,1,1),(-3,0,1),(-2,1,0)\}, E = \{6,0,0\}$

14. $S = \{(-1,1,0),(1,1,0),(0,0,1)\}, E = \{-2,2,3\}$

15. Suponga que una matriz A de 3×3 tiene los eigenvalores $3, 0, -7$. ¿A es diagonalizable? ¿Por qué sí o por qué no?

16. Suponga que una matriz A 3×3 es triangular superior, y que sus elementos diagonales son $2, 1, -5$. Demuestre que A es diagonalizable. ¿Cuál es D?

17. Determine una base de $\mathbf{R}^3$ formada por los eigenvectores de

$$\begin{bmatrix} 1 & 2 & 2 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

18. Encuentre una base de $\mathbf{R}^3$ formada por los eigenvectores de

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & 0 & 0 \\ 1 & 2 & 2 \end{bmatrix}$$

19. Demuestre que A es diagonalizable.

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$$

20. Compruebe que A no es diagonalizable.

$$A = \begin{bmatrix} 2 & 3 & -5 \\ 2 & 3 & -5 \\ 2 & 3 & -5 \end{bmatrix}$$

21. Demuestre que A no es diagonalizable.

$$A = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

22. Demuestre que A no es diagonalizable.

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

23. Demuestre que A es diagonalizable si y sólo si $a + b + c \neq 0$.

$$A = \begin{bmatrix} a & b & c \\ a & b & c \\ a & b & c \end{bmatrix}$$

Supondremos que al menos uno de a, b y c es distinto de cero, de modo que A es distinta de cero.

24. ¿Para qué valores de a es diagonalizable la matriz siguiente? Sólo tenga en cuenta la diagonalización con matrices cuyos elementos sean *reales*.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & a \end{bmatrix}$$

25. ¿Para cuáles valores de a es diagonalizable la matriz siguiente?

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & a & 0 \end{bmatrix}$$

26. Demuestre que la matriz siguiente es diagonalizable para todos los valores reales de a.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ a & 0 & 0 \end{bmatrix}$$

27. Aplique la diagonalización para determinar A^6 y A^9 donde

$$A = \begin{bmatrix} 0 & 8 \\ 2 & 0 \end{bmatrix}$$

28. Deduzca una fórmula para determinar A^n, si A es como en el ejercicio 27.

29. Use la diagonalización para calcular

$$\begin{bmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}^7$$

30. Emplee la diagonalización para calcular

$$\begin{bmatrix} -2 & 0 & 3 \\ 0 & 1 & 0 \\ 3 & 0 & -2 \end{bmatrix}^{-n}$$

31. Demuestre que para $k = 0, 1, 2, \ldots$,

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}^{k+1} = \begin{bmatrix} 2^k & 2^k \\ 2^k & 2^k \end{bmatrix}$$

32. Verifique la identidad

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} r_1^{k+1} - r_2^{k+1} \\ r_1^k - r_2^k \end{bmatrix} \qquad (7.12)$$

en la que $r_1 = (1 + \sqrt{5})/2$, $r_2 = (1 - \sqrt{5})/2$, y k es un entero positivo.

33. Compruebe el teorema 9.

Diagonalización de transformaciones lineales

34. (**Se requiere cálculo**) Demuestre que la diferenciación $\frac{d}{dx}$: $P_3 \to P_3$ no es una transformación lineal diagonalizable.

En los ejercicios 35 a 41 determine si la transformación mencionada es diagonalizable. Si lo es, calcule una base que la diagonalice.

35. $T: P_3 \to P_3$, $T(p(x)) = p(0)$.

36. $T: P_2 \to P_2$, $T(p(x)) = p(x + 1)$.

37. Reflexión de $\mathbf{R}^2$ en el eje x.

38. Reflexión de $\mathbf{R}^2$ respecto a la recta $y = -x$.

39. Proyección de $\mathbf{R}^2$ sobre el eje x.

40. Reflexión de $\mathbf{R}^3$ en el plano xy.

41. Proyección de $\mathbf{R}^3$ sobre el plano xy.

42. Demuestre el teorema 11.

7.3 Aproximaciones de eigenvalores y eigenvectores

Objetivo del estudiante para esta sección
Ver cómo se calculan los eigenvalores y los eigenvectores en la práctica.

Hasta ahora hemos calculado los eigenvalores resolviendo la ecuación característica y después determinando los eigenvectores correspondientes. Cuando las matrices son grandes, éste *no es* un método práctico aun cuando se acepten soluciones aproximadas. El problema principal estriba en que el cálculo del determinante de una matriz característica grande es *muy costoso*. También puede ser que la ecuación característica resultante, de grado alto, sea difícil de resolver.

En la práctica se usan métodos mucho más eficientes que trabajar hacia atrás: primero aproximan un eigenvector y después el eigenvalor correspondiente.

En esta sección analizaremos dos métodos iterativos: el *método de la potencia* y el *método de la potencia inversa*. En el capítulo 8 analizaremos otro método eficiente basado en la descomposición QR de una matriz.

El método de potencias

Este método aproxima un eigenvector de una matriz A calculando potencias de productos

$$A^k\mathbf{x}, \quad k = 1, 2, \ldots$$

comenzando con cualquier vector n $\mathbf{x}$. *Con frecuencia* sucede que a medida que k crece, $A^k\mathbf{x}$ se hace paralelo a un eigenvector de A. Para ilustrar esto veamos el ejemplo 24.

■ EJEMPLO 24 Sean

$$A = \begin{bmatrix} 8 & 7 \\ 1 & 2 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Primero se encuentran los eigenvalores y los eigenvectores de A. A continuación se determina la dirección de $A^k\mathbf{x}$ cuando k es grande, y se relaciona con la dirección de los eigenvectores.

SOLUCIÓN A tiene los eigenvalores 1 y 9 y sus eigenvectores correspondientes son $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Ya que

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{13}{8}\begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{3}{8}\begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

entonces, de acuerdo con la ecuación (7.4) de la sección 7.1,

$$A^k\mathbf{x} = \frac{13}{8}\cdot 1^k\begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{3}{8}\cdot 9^k\begin{bmatrix} 7 \\ 1 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} -13 + 21\cdot 9^k \\ 13 + 3\cdot 9^k \end{bmatrix}$$

Para encontrar la dirección de $A^k\mathbf{x}$, para k grandes, calcularemos la relación de sus componentes cuando $k \to \infty$:

$$\lim_{k\to\infty} \frac{-13 + 21\cdot 9^k}{13 + 3\cdot 9^k} = \lim_{k\to\infty} \frac{-13/9^k + 21}{13/9^k + 3} = \frac{21}{3} = 7$$

porque $\lim_{k\to\infty}(1/9)^k = 0$. Así, $A^k\mathbf{x}$ tiende a hacerse paralelo a $\begin{bmatrix} 7 \\ 1 \end{bmatrix}$, el eigenvector cuyo eigenvalor tiene el máximo valor absoluto. Por ejemplo, cuando $k = 6$,

$$A^6\mathbf{x} = \begin{bmatrix} 1\,395\,031 \\ 199\,292 \end{bmatrix} \text{ con la relación de componentes } \cong 6.9999$$

La solución en este ejemplo sugiere que la dirección de $A^k\mathbf{x}$ tiende a la dirección del eigenvector cuyo valor absoluto es máximo. □

Sea A una matriz de $n \times n$, cuyos eigenvalores sean $\lambda_1, \ldots, \lambda_n$. Un eigenvalor, digamos λ_1, se llama **eigenvalor dominante** si

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$$

Por ejemplo, si A tiene los eigenvalores 1, 4 y –5, entonces –5 es el eigenvalor dominante. Si los eigenvalores de A son 1, 4 y – 4, *no* hay un eigenvalor dominante, porque dos eigenvalores tienen el valor absoluto máximo $|4| = |-4|$.

Observe que un eigenvalor dominante no puede ser cero. También, debe ser *real* si la matriz es real, porque los eigenvalores complejos en este tipo de matrices se presentan en pares conjugados y, como ya sabemos, los números conjugados tienen los mismos valores absolutos.

El teorema siguiente es la espina dorsal del método de las potencias, y resalta la importancia de las matrices con un eigenvalor dominante.

TEOREMA 12

Sea A una matriz diagonalizable $n \times n$ con eigenvectores básicos $\mathbf{v}_1, \ldots, \mathbf{v}_n$, con los eigenvalores correspondientes $\lambda_1, \ldots, \lambda_n$, y con un eigenvalor dominante, digamos λ_1. Sea $\mathbf{x}$ un vector resultante de la combinación lineal de los $\mathbf{v}$:

$$\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$$

de modo que $c_1 \neq 0$. Entonces, cuando k crece, un múltiplo escalar de $A^k\mathbf{x}$ tiende a un múltiplo de $\mathbf{v}_1$ por un escalar. En particular, la dirección de $A^k\mathbf{x}$ se aproxima a la de $\mathbf{v}_1$.

DEMOSTRACIÓN De acuerdo con la ecuación (7.4) de la sección 7.1,

$$A^k\mathbf{x} = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n$$

Como $\lambda_1 \neq 0$, esta ecuación implica que

$$\frac{1}{\lambda_1^k}A^k\mathbf{x} = c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n \tag{7.13}$$

Aprovechando que $r^k \to 0$ cuando $k \to \infty$ si $|r| < 1$,

$$\frac{1}{\lambda_1^k}A^k\mathbf{x} \to c_1\mathbf{v}_1 \qquad \text{cuando } k \to \infty$$

porque $|\lambda_2/\lambda_1|, \ldots, |\lambda_n/\lambda_1| < 1$. Por consiguiente, $A^k\mathbf{x}$, escalado por $1/\lambda_1^k$ se hace paralelo a $\mathbf{v}_1$ cuando k es grande, siempre que $c_1 \neq 0$.

Así, para obtener la aproximación de un eigenvector del eigenvalor dominante, comenzando con un vector inicial $\mathbf{x}$, se calcula $A^k\mathbf{x}$ con la siguiente iteración. Si $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ se definen como sigue:

$$\mathbf{x} = \mathbf{x}_0, \qquad \mathbf{x}_1 = A\mathbf{x}_0, \qquad \mathbf{x}_2 = A^2\mathbf{x}_0 = A\mathbf{x}_1, \ldots,$$

Así, $\mathbf{x}_k = A^k\mathbf{x}$ y

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \qquad k = 0, 1, \ldots \tag{7.14}$$

Para aproximar el eigenvalor λ_1 correspondiente, se observa que si $\mathbf{x}_k$ es un eigenvector (aproximado), entonces $\mathbf{x}_{k+1} = A\mathbf{x}_k \cong \lambda_1\mathbf{x}_k$. Por tanto, λ_1 puede calcularse determinando el cociente de las dos componentes correspondientes de $\mathbf{x}_{k+1}$ y $\mathbf{x}_k$. $\square$

Utilizaremos el teorema 12 para deducir un eigenvector de una matriz, volviendo a resolver el ejemplo 24.

■ EJEMPLO 25 Aproxime el eigenvalor dominante de

$$A = \begin{bmatrix} 8 & 7 \\ 1 & 2 \end{bmatrix}$$

y un eigenvector correspondiente aplicando la iteración (7.14) con $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

SOLUCIÓN Necesitamos despejar $\mathbf{x}_k$ para varios valores de k. La tabla siguiente fue compilada con Maple. Presenta los valores de k, $\mathbf{x}_k$, el cociente d_k entre los componentes de cada $\mathbf{x}_k$ y el cociente l_k entre los primeros componentes de $\mathbf{x}_k$ y $\mathbf{x}_{k-1}$.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 22 \\ 5 \end{bmatrix}$	$\begin{bmatrix} 211 \\ 32 \end{bmatrix}$	$\begin{bmatrix} 1912 \\ 275 \end{bmatrix}$	$\begin{bmatrix} 17\,221 \\ 2462 \end{bmatrix}$	$\begin{bmatrix} 155\,002 \\ 22\,145 \end{bmatrix}$
d_k	0.5	4.4	6.5938	6.9527	6.9947	6.9994
l_k		22.0	9.5909	9.0616	9.0068	9.0008

Es evidente que la relación d_k de los componentes de $\mathbf{x}_k$ tiende a 7. Así, tenemos que un eigenvector aproximado es un múltiplo de (7, 1) por un escalar. Como los valores de l_k tienden a 9, también lo hace el eigenvalor correspondiente. De este modo hemos determinado el eigenvalor de mayor valor absoluto, y un eigenvector correspondiente. En la figura. 7.3 mostramos sólo $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$ y la recta l que pasa por $\mathbf{0}$ y (7, 1). Las demás $\mathbf{x}$ están demasiado lejos y sus líneas demasiado cerca de l para poder mostrarlas.

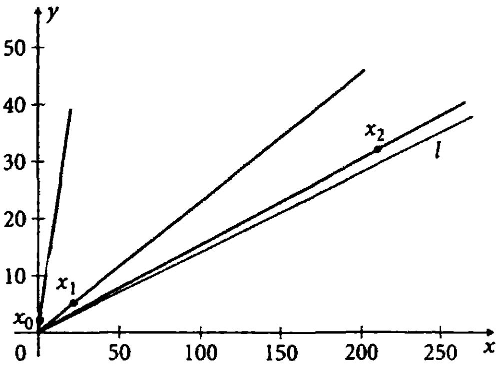


Figura 7.3 El método de potencias.

En los cálculos del ejemplo 25 surge un problema. Los componentes de $\mathbf{x}_k$ crecen con mucha rapidez, casi fuera de control, mientras que todo lo que necesitamos es (7, 1) o algún múltiplo escalar de él. Como nos interesa principalmente la dirección de (7, 1), podemos escalar cada $\mathbf{x}_k$ para mantener pequeños los números. Una forma es hacer que $\mathbf{x}_k$ sea vector unitario, multiplicándolo por $1/\|\mathbf{x}_k\|$. Un método más fácil es escalar $\mathbf{x}_k$ de manera que su mayor elemento sea 1. En el ejemplo 25 comenzamos con $\mathbf{x}_0$ y lo escalamos:

$$\mathbf{y}_0 = \frac{1}{2}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1.0 \end{bmatrix}$$

A continuación determinamos

$$\mathbf{x}_1 = \begin{bmatrix} 8 & 7 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0.5 \\ 1.0 \end{bmatrix} = \begin{bmatrix} 11.0 \\ 2.5 \end{bmatrix}$$

y lo escalamos:

$$\mathbf{y}_1 = \frac{1}{11.0}\begin{bmatrix} 11.0 \\ 2.5 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 0.22727 \end{bmatrix}$$

y así sucesivamente. En este entorno podemos aproximar el eigenvalor λ_1 como sigue: supongamos que llegamos a un $\mathbf{y}_k$ escalado. A continuación calculamos $\mathbf{x}_{k+1} = A\mathbf{y}_k$; antes de escalarlo guardamos el componente que corresponde al componente de $\mathbf{y}_k$ que tiene 1. Este componente es una aproximación de λ_1. (¿Por qué?) En la tabla siguiente presentamos los $\mathbf{x}_k$, $\mathbf{y}_k$ y los componentes l_k que aproximan a λ_1 para las A y $\mathbf{x}_0$ del ejemplo 25.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix}1\\2\end{bmatrix}$	$\begin{bmatrix}11.0\\2.5\end{bmatrix}$	$\begin{bmatrix}9.5909\\1.4545\end{bmatrix}$	$\begin{bmatrix}9.0616\\1.3033\end{bmatrix}$	$\begin{bmatrix}9.0068\\1.2877\end{bmatrix}$	$\begin{bmatrix}9.0008\\1.2859\end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix}0.5\\1.0\end{bmatrix}$	$\begin{bmatrix}1.0\\0.22727\end{bmatrix}$	$\begin{bmatrix}1.0\\0.15165\end{bmatrix}$	$\begin{bmatrix}1.0\\0.14383\end{bmatrix}$	$\begin{bmatrix}1.0\\0.14297\end{bmatrix}$	$\begin{bmatrix}1.0\\0.14287\end{bmatrix}$
l_k	2	11.0	9.5909	9.0616	9.0068	9.0008

Esta vez vemos que los números son más fáciles de manejar. (En la figura. 7.4 se muestran sólo $\mathbf{y}_0$, $\mathbf{y}_1$ y $\mathbf{y}_7$. Los restantes $\mathbf{y}$ están demasiado cerca de $\mathbf{y}_7$, que también está próximo a la recta l que pasa por $\mathbf{0}$ y $(7, 1)$. Repasemos los pasos de este método de potencias modificado.)

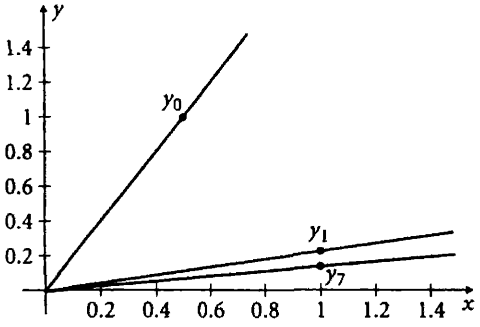


Figura 7.4 El método de potencias (modificado).

Algoritmo 1

(Método de potencias, elemento máximo 1)

Sea A una matriz diagonalizable $n \times n$, con un eigenvalor dominante; sea $\mathbf{x}_0$ cualquier vector y k la cantidad deseada de iteraciones.

DATOS: A, $\mathbf{x}_0$ y k.

1. Sea l_0 el componente de $\mathbf{x}_0$ de máximo valor absoluto.
2. Igualar $\mathbf{y}_0 = (1/l_0)\mathbf{x}_0$.
3. Par $i = 1, \ldots, k$:

   (a) Calcular $A\mathbf{y}_{i-1}$. Sea $\mathbf{x}_i = A\mathbf{y}_{i-1}$;
   (b) Sea l_i el componente de $\mathbf{x}_i$ de máximo valor absoluto;
   (c) Sea $\mathbf{y}_i = (1/l_i)\mathbf{x}_i$.

RESULTADOS: $\mathbf{y}_k$ y l_k.
Para la mayor parte de $\mathbf{x}_0$: l_k se aproxima al eigenvalor dominante de A, y $\mathbf{y}_k$ al correspondiente eigenvector.

OBSERVACIONES

1. Dijimos que el método de las potencias se aplica a casi todo $\mathbf{x}_0$, porque no conocemos previamente si el coeficiente c_1 de $\mathbf{v}_1$ en términos del eigenvector es $\neq 0$, como lo requiere el teorema 12. Sin embargo, el método parece trabajar para cualquier vector inicial, porque es probable que los errores de redondeo de la computadora reemplacen los ceros por números pequeños de punto flotante.
2. El método de las potencias es un **método autocorregido**. Si hemos calculado mal $\mathbf{x}_i$ en cualquier punto, podemos proseguir, como si estuviéramos comenzando en este vector por primera vez.
3. La rapidez o lentitud de convergencia de la iteración depende de la relación $|\lambda_2/\lambda_1|$, siendo λ_2 el eigenvalor con el segundo valor absoluto mayor en la ecuación (7.13). Si $|\lambda_2/\lambda_1|$ se aproxima a 0, la convergencia es muy rápida. Éste fue el caso en el ejemplo 25, donde $|\lambda_2/\lambda_1| = \frac{1}{9} \cong 0.11111$. Si $|\lambda_2/\lambda_1|$ se acerca a 1, la convergencia es lenta.
4. El método de las potencias funciona aunque tengamos un **eigenvalor dominante repetido**.

$$\lambda_1 = \cdots = \lambda_r \quad \text{y} \quad |\lambda_1| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_n|$$

   Por ejemplo, supongamos que A tiene los eigenvalores 4, 1, –5 y –5 (de modo que $\lambda_1 = \lambda_2 = -5$, $\lambda_3 = 4$ y $\lambda_4 = 1$). Esto se comprueba con facilidad, porque la ecuación (7.13) del teorema 12 sería entonces

$$(1/\lambda_1^k)A^k\mathbf{x} = \sum_{i=1}^{r} c_i\mathbf{v}_i + \sum_{i=r+1}^{n} c_i(\lambda_i/\lambda_1)^k\mathbf{v}_i$$

   de modo que $(1/\lambda_1^k)A^k\mathbf{x}$ tiende a $\Sigma_{i=1}^{r} c_i\mathbf{v}_i$, que es un eigenvector de λ_1.
5. El método de las potencias funciona aun cuando A no sea diagonalizable.[3] Por lo general, en este caso la convergencia es lenta. Las hipótesis cuidadosas y las demostraciones son bastante complicadas.

Cocientes de Rayleigh (o método de Rayleigh-Ritz)

Hay una variante común del método de las potencias, donde se normaliza dividiendo entre la *norma* del vector; en ese caso se usa el **cociente de Rayleigh** $r(x)$,

$$r(x) = \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

para aproximar los eigenvalores. Esto funciona porque si $\mathbf{x}$ es un eigenvector, entonces $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T\lambda\mathbf{x} = \lambda(\mathbf{x}^T\mathbf{x})$. Para ilustrarlo utilizaremos los datos del ejemplo 25, sea

$$\mathbf{y}_0 = \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|} = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}0.44721\\0.89443\end{bmatrix}$$

A continuación calculamos

$$\mathbf{x}_1 = A\mathbf{y}_0 = \begin{bmatrix}9.8387\\2.2361\end{bmatrix}$$

[3] Los analistas numéricos llaman *defectuosa* a una matriz no diagonalizable.

Después se calcula el cociente de Rayleigh

$$r(\mathbf{y}_0) = \frac{\mathbf{y}_0^T A \mathbf{y}_0}{\mathbf{y}_0^T \mathbf{y}_0} = \mathbf{y}_0 \cdot \mathbf{x}_1$$

porque $\mathbf{y}_0^T \mathbf{y}_0 = \mathbf{y}_0 \cdot \mathbf{y}_0 = \|\mathbf{y}_0\|^2 = 1$. Se continúa en la misma forma:

$$\mathbf{y}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}, \qquad \mathbf{x}_2 = A\mathbf{y}_1, \qquad r(\mathbf{y}_1) = \mathbf{y}_1 \cdot \mathbf{x}_2, \cdots$$

Los cocientes de Rayleigh

$$r(\mathbf{y}_0), \qquad r(\mathbf{y}_1), \qquad r(\mathbf{y}_2), \ldots,$$

tienden al eigenvalor dominante. En nuestro caso,

k	0	1	2	3	4	5
$\mathbf{y}_k$	$\begin{bmatrix} 0.44721 \\ 0.89443 \end{bmatrix}$	$\begin{bmatrix} 0.97509 \\ 0.22162 \end{bmatrix}$	$\begin{bmatrix} 0.9887 \\ 0.14994 \end{bmatrix}$	$\begin{bmatrix} 0.98981 \\ 0.14236 \end{bmatrix}$	$\begin{bmatrix} 0.98994 \\ 0.14152 \end{bmatrix}$	$\begin{bmatrix} 0.98995 \\ 0.14143 \end{bmatrix}$
$\mathbf{x}_{k+1}$	$\begin{bmatrix} 9.8387 \\ 2.2361 \end{bmatrix}$	$\begin{bmatrix} 9.3521 \\ 1.4183 \end{bmatrix}$	$\begin{bmatrix} 8.9592 \\ 1.2886 \end{bmatrix}$	$\begin{bmatrix} 8.915 \\ 1.2745 \end{bmatrix}$	$\begin{bmatrix} 8.9102 \\ 1.273 \end{bmatrix}$	$\begin{bmatrix} 8.9096 \\ 1.2728 \end{bmatrix}$
$r(y_k)$	6.4	9.4335	9.0512	9.0056	9.0007	9.0001

Algoritmo 2

(Cocientes de Rayleigh, o método de Rayleigh-Ritz)

Sea A una matriz diagonalizable $n \times n$, con un eigenvalor dominante; sea $\mathbf{x}_0$ cualquier vector, y k la cantidad deseada de iteraciones.

DATOS: A, $\mathbf{x}_0$, k.
Para $i = 0, \ldots, k-1$,

1. Sea $\mathbf{y}_i = (1/\|\mathbf{x}_i\|)\mathbf{x}_i$;
2. Sea $\mathbf{x}_{i+1} = A\mathbf{y}_i$;
3. Sea $r_i = \mathbf{y}_i \cdot \mathbf{x}_{i+1}$.

RESULTADOS: $\mathbf{y}_{k-1}$ y r_{k-1}.
Para la mayor parte de $\mathbf{x}_0$, r_k se aproxima al eigenvalor dominante de A, y $\mathbf{y}_k$ al eigenvector correspondiente.

Para matrices **simétricas**, este método es muy eficiente y requiere menos iteraciones[4] para lograr la misma exactitud. Como ejemplo veamos a la matriz $\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}$ cuyos eigenvalores son 3 y −2. El método de potencias es lento, porque $|(-2)/3|$ se acerca a 1. Si usamos este método para obtener $\lambda = 3.0$ con 5 cifras decimales se necesitan 33 iteraciones, pero sólo 18 con los cocientes de Rayleigh, si comenzamos en $\mathbf{x}_0 = (1, 2)$. (Para este cálculo empleamos Mathematica.)

[4] Requiere casi la mitad de las iteraciones.

Desplazamientos del origen

Ahora que sabemos cómo calcular el eigenvalor dominante, podemos aplicar un artificio sencillo para ***determinar el eigenvalor más alejado del dominante*** (si es que lo hay).

Como vimos en el ejercicio 36 de la sección 7.1, si λ es un eigenvalor de A con el correspondiente eigenvector $\mathbf{v}$, entonces, para cualquier escalar c, la matriz $A - cI$ tiene el eigenvalor $\lambda - c$ y el eigenvector correspondiente $\mathbf{v}$, porque

$$(A - cI)\mathbf{v} = A\mathbf{v} - c\mathbf{v} = (\lambda - c)\mathbf{v}$$

De modo que si $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los eigenvalores de A, entonces $0, \lambda_2 - \lambda_1, \ldots, \lambda_n - \lambda_1$ son los eigenvalores de $B = A - \lambda_1 I$.

Podemos combinar esta observación con el método de la potencia, para obtener el eigenvalor λ_r más alejado del eigenvalor dominante, digamos λ_1. Primero se calcula λ_1. Ahora, $\lambda_r - \lambda_1$ es el eigenvalor dominante de B, de manera que podemos recurrir a cualquier método de potencia para determinarlo. Por último, sumamos λ_1 para obtener λ_r.

Como ilustración, calcularemos el menor de los eigenvalores de A en el ejemplo 25, teniendo como dato el eigenvalor dominante ya obtenido, que es $\lambda_1 = 9$. Se forma la matriz

$$B = A - 9I = \begin{bmatrix} 8-9 & 7 \\ 1 & 2-9 \end{bmatrix} = \begin{bmatrix} -1 & 7 \\ 1 & -7 \end{bmatrix}$$

A continuación se calcula el eigenvalor dominante de B usando, por ejemplo, el método de potencias con cocientes de Rayleigh para obtener una convergencia muy rápida. Iniciaremos en $\mathbf{x}_0 = (1, 2)$:

k	0	1	2
y_k	$\begin{bmatrix} 0.44721 \\ 0.89443 \end{bmatrix}$	$\begin{bmatrix} 0.7071 \\ -0.7071 \end{bmatrix}$	$\begin{bmatrix} -0.70711 \\ 0.70711 \end{bmatrix}$
$r(y_k)$	-2.6	-7.9998	-8.0001

Por consiguiente, el eigenvalor dominante de B es -8. Entonces el eigenvalor de A más alejado de $\lambda_1 = 9$ es $9 - 8 = 1$. El eigenvector correspondiente es $\begin{bmatrix} -0.70711 \\ 0.70711 \end{bmatrix}$, es decir $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Método de las potencias inversas

A continuación demostraremos cómo usar el método de las potencias para determinar *el eigenvalor más cercano al origen (si es que lo hay).*

El método de las potencias inversas se basa en la observación de que si λ es un eigenvalor de A con el eigenvector correspondiente $\mathbf{v}$, entonces λ^{-1} es un eigenvalor de A^{-1} con el mismo eigenvector (ejercicio 35 de la sección 7.1). Así, para determinar el eigenvalor de A más cercano al origen, necesitamos calcular sólo el eigenvalor dominante de A^{-1}. El cálculo de A^{-1} es costoso, pero lo podemos evitar sólo con resolver el sistema

$$A\mathbf{x}_{k+1} = \mathbf{x}_k$$

para determinar a $\mathbf{x}_{k+1}$. A continuación presentamos una versión de este método, que emplea los cocientes de Rayleigh.

Algoritmo 3

(Método de las potencias inversas)

DATOS: A, $\mathbf{x}_0$, k.
Para $i = 0, \ldots, k-1$,

1. Sea $\mathbf{y}_i = (1/\|\mathbf{x}_i\|)\mathbf{x}_i$;
2. Despejar $\mathbf{z}$ del sistema $A\mathbf{z} = \mathbf{y}_i$;
3. Sea $\mathbf{x}_{i+1} = \mathbf{z}$;
4. Sea $r_i = \mathbf{y}_i \cdot \mathbf{x}_{i+1}$.

RESULTADOS: $\mathbf{y}_{k-1}$ y r_{k-1}.
Para la mayor parte de $\mathbf{x}_0$, r_k^{-1} se aproxima al eigenvalor de A más cercano al origen, $\mathbf{y}_k$ es el eigenvector correspondiente.

Para ilustrarlo, supongamos que se desea conocer el eigenvalor de A, del ejemplo 25, que sea más cercano al origen. Entonces, para $\mathbf{x}_0 = (1, 2)$,

$$\mathbf{y}_0 = \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|} = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}0.44721\\0.89443\end{bmatrix}$$

A continuación se reduce el sistema con operaciones de renglón:

$$\begin{bmatrix}8 & 7 & 0.44721\\1 & 2 & 0.89443\end{bmatrix} \sim \begin{bmatrix}1 & 0 & -0.59629\\0 & 1 & 0.74536\end{bmatrix}$$

e igualamos

$$\mathbf{x}_1 = \begin{bmatrix}-0.59629\\0.74536\end{bmatrix} \quad \text{y} \quad r_0 = \mathbf{x}_1 \cdot \mathbf{y}_0 = 0.40001$$

Continuamos del mismo modo para obtener

$$\mathbf{y}_1 = \begin{bmatrix}-0.62469\\0.78087\end{bmatrix}, \qquad \mathbf{y}_2 = \begin{bmatrix}-0.69893\\0.71519\end{bmatrix}$$

y

$$r_1 = 1.0623, \qquad r_2 = 1.0075$$

Si nos detenemos ahora, el eigenvalor más cercano al origen es aproximadamente, $1/r_2 = 1/1.0075 = 0.99256$, que ya se acercó bastante a 1, el eigenvalor verdadero. El eigenvector correspondiente es $\mathbf{y}_2$, paralelo a $\begin{bmatrix}-1\\1.0233\end{bmatrix}$. Está suficientemente cerca del eigenvector verdadero, que es $\begin{bmatrix}-1\\1\end{bmatrix}$.

El algoritmo 3 se usa con más eficiencia con una factorización LU de A. En este caso el paso (2) se sustituye por:

(2a) Resolver $L\mathbf{y} = \mathbf{y}_i$.
(2b) Resolver $L\mathbf{z} = \mathbf{y}$.

Método de la potencia inversa desplazada

Por último, combinaremos la iteración inversa con un desplazamiento del origen para calcular el *eigenvalor más cercano a determinado número*. Por ejemplo, si un número μ se aproxima más al eigenvalor λ que a cualquier otro eigenvalor, entonces $1/(\lambda - \mu)$ es un eigenvalor dominante de $(A-\mu I)^{-1}$, que podemos calcular con el método de las potencias. De la misma forma que en el caso del método de la potencia inversa, es más eficiente resolver el sistema

$$(A - \mu I)\mathbf{x}_{k+1} = \mathbf{x}_k$$

que usar la matriz inversa. Llegamos entonces a lo siguiente.

Algoritmo 4

(Método de la potencia inversa desplazada)

DATOS: A, $\mathbf{x}_0$, k. Aproximación inicial del eigenvalor: μ.
Para $i = 0, \ldots, k-1$.

1. Sea $\mathbf{y}_i = 1/(\|\mathbf{x}_i\|)\mathbf{x}_i$;
2. Despejar $\mathbf{z}$ del sistema $(A - \mu I)\mathbf{z} = \mathbf{y}_i$;
3. Sea $\mathbf{x}_{i+1} = \mathbf{z}$;
4. Sea $r_i = \mathbf{y}_i \cdot \mathbf{x}_{i+1}$.

RESULTADOS: $\mathbf{y}_{k-1}$ y r_{k-1}.
Para la mayor parte de $\mathbf{x}_0$, $\mu + r_k^{-1}$ se aproxima al eigenvalor de A más cercano a μ; $\mathbf{y}_k$ es el eigenvector correspondiente.

La iteración funciona admirablemente si el número dado es un valor inicial de un eigenvalor que se desee calcular. Cuanto mejor se haga esta aproximación, la convergencia de la iteración será más rápida.

En el ejemplo 25, supongamos que el valor inicial es $\mu = 2$, para uno de los eigenvalores de A. Entonces, para $\mathbf{x}_0 = (1, 2)$,

$$\mathbf{y}_0 = \frac{\mathbf{x}_0}{\|\mathbf{x}_0\|} = \begin{bmatrix} 0.44721 \\ 0.89443 \end{bmatrix}$$

Ahora reducimos el sistema $[A - 2I : \mathbf{y}_0]$ con operaciones de renglón,

$$\begin{bmatrix} 6 & 7 & 0.44721 \\ 1 & 0 & 0.89443 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0.89443 \\ 0 & 1 & -0.70277 \end{bmatrix}$$

e igualamos

$$\mathbf{x}_1 = \begin{bmatrix} 0.89443 \\ -0.70277 \end{bmatrix} \quad \text{y} \quad r_0 = \mathbf{x}_1 \cdot \mathbf{y}_0 = -0.22858$$

Continuamos en la misma forma para llegar a

$$\mathbf{y}_1 = \begin{bmatrix} 0.78631 \\ -0.61782 \end{bmatrix}, \qquad \mathbf{y}_2 = \begin{bmatrix} -0.69347 \\ 0.72049 \end{bmatrix}$$

y

$$r_1 = -0.88237, \qquad r_2 = -1.016$$

Si nos detenemos en este punto, el eigenvalor más cercano a 2 es $2 + 1/r_2 = 2 + 1/(-1.016) = 1.0157$, que ya se acerca bastante a 1, el eigenvalor verdadero.

De nuevo, el algoritmo 4 se usa casi siempre con una factorización LU de A.

Aplicación a raíces de polinomios

Las aproximaciones numéricas de los eigenvalores pueden ser tan eficientes que, en lugar de calcular los eigenvalores mediante la ecuación característica, se determinan las raíces de esta última o de *cualquier* ecuación polinomial, aproximando los eigenvalores de la **matriz asociada**. En los ejercicios 47 a 53 de la sección 7.1 definimos la matriz asociada, $C(p)$, de un polinomio mónico[5] $p(x)$, y vimos que sus eigenvalores son las raíces de $p(x)$.

■ EJEMPLO 26 (Raíces como eigenvalores) Aproxime una raíz de $p(x) = x^3 - 15x^2 + 59x - 45$ usando el valor inicial $x = 10$.

SOLUCIÓN Aplicaremos el método de la potencia inversa para determinar el eigenvalor de $C(p)$ más cercano a 10

$$C(p) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 45 & -59 & 15 \end{bmatrix}$$

Comenzando con un vector inicial unitario, por ejemplo $\mathbf{y}_0 = \left(\frac{2}{3}, \frac{2}{3}, \frac{1}{3}\right)$, reducimos $[C(p) - 10I : \mathbf{y}_0]$ con lo que obtenemos $\mathbf{x}_1$. A continuación calculamos $\mathbf{y}_0 \cdot \mathbf{x}_1$ para obtener $r(y_0)$, y así sucesivamente. Después de 4 iteraciones llegamos a

$r(y_0)$	$r(y_1)$	$r(y_2)$	$r(y_3)$
0.17778	−1.2099	−1.0303	−1.0051

Por tanto, una raíz aproximada de $p(x)$ es $10 + (-1.0051)^{-1} = 9.0051$, que ya se acerca bastante a la raíz verdadera $x = 9$. En forma semejante podemos obtener las otras dos raíces: 1 y 5.

NOTA Con frecuencia, para calcular las raíces de polinomios se elige el método de eigenvalores aproximados de la matriz asociada, en los paquetes de programas informáticos numéricos. Por ejemplo, el comando `roots` de MATLAB se basa en este método.

Eigenvalores aproximados con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> evalf(eigenvals(matrix(
        [[1,1,1],[1,1,0],[1,0,0]])));
```

$$[-.8019377352, .5549681326, 2.246979605]$$

---

[5] N. del T.: Un polinomio mónico es aquel cuyo coeficiente principal es +1.

Mathematica

```
In[1]:=
     N[Eigenvalues[{{1,1,1},{1,1,0},{1,0,0}}]]
Out[1]=
                                 -16                         -17
     {2.24698+1.11022 10    I,-0.801938-5.55112 10    I,
                                    -16
        0.554958-1.11022 10    I
```

MATLAB

```
>> eig([1 1 1; 1 1 0; 1 0 0])
ans =
    0.5550
   -0.8019
    2.2478
```

Ejercicios 7.3 (Se recomienda utilizar calculadora)

En los ejercicios 1 y 2 use la información acerca de la matriz A 2×2 y el vector 2 $\mathbf{x}_0$ para lo siguiente:

(a) Estime un eigenvalor de A.
(b) Calcule un eigenvector con elemento máximo igual a 1.

1. $A^4\mathbf{x}_0 = \begin{bmatrix} 937 \\ 938 \end{bmatrix}, A^5\mathbf{x}_0 = \begin{bmatrix} 4687 \\ 4688 \end{bmatrix}$

2. $A^5\mathbf{x}_0 = \begin{bmatrix} 1561 \\ -1564 \end{bmatrix}, A^6\mathbf{x}_0 = \begin{bmatrix} -7811 \\ 7814 \end{bmatrix}$

Sean $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$. En los ejercicios 3 a 5 determine un eigenvector y un eigenvalor de cierta matriz A (no especificada) 2×2 si para cierto vector 2 $\mathbf{x}$ (no especificado):

3. $A^6\mathbf{x} = \mathbf{u}$ y $A^7\mathbf{x} = \mathbf{v}$

4. $A^6\mathbf{x} = \mathbf{u}$ y $A^8\mathbf{x} = \mathbf{v}$

5. $A^{-6}\mathbf{x} = \mathbf{u}$ y $A^{-7}\mathbf{x} = \mathbf{v}$

En los ejercicios 6 a 19 aplique el método de las potencias ([illegible] goritmo 1) con $k = 4$ para hacer lo siguiente:

(a) Aproxime el eigenvalor dominante, y el e[illegible] corres- p[illegible]ente de la matriz dada.

(b) Compare el eigenvalor aproximado con el verdadero.

6. $\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$

7. $\begin{bmatrix} -3 & 2 \\ 2 & -3 \end{bmatrix}$

8. $\begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$

9. $\begin{bmatrix} -4 & [illegible] \\ 2 & 4 \end{bmatrix}$

10. $\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$

11. $\begin{bmatrix} -5 & 4 \\ 4 & -5 \end{bmatrix}$

12. $\begin{bmatrix} -8 & 7 \\ 1 & -2 \end{bmatrix}$

13. $\begin{bmatrix} -6 & 5 \\ 2 & -3 \end{bmatrix}$

14. $\begin{bmatrix} -7 & 6 \\ 1 & -2 \end{bmatrix}$

15. $\begin{bmatrix} 7 & 6 \\ 4 & 5 \end{bmatrix}$

16. $\begin{bmatrix} 10 & 9 \\ 2 & 3 \end{bmatrix}$

17. $\begin{bmatrix} -10 & 9 \\ 2 & -3 \end{bmatrix}$

18. $\begin{bmatrix} -7 & 6 \\ 4 & -5 \end{bmatrix}$

19. $\begin{bmatrix} -7 & 6 \\ 7 & -8 \end{bmatrix}$

En los ejerci[illegible] 24 use el método de Raileigh-Ritz (algo[illegible]o 2) con k [illegible] [illegible]aproximar el eigenvalor dominante de la m[illegible]iz.

20 [illegible]a matriz [illegible] ejercic[illegible]
21. La matriz d[illegible]l ejercic[illegible]
22. La matriz del ejercic[illegible]
23. La matriz del ejercici[illegible]
24. La matriz del ejercicio [illegible]

En los ejercicios 25 a 34 use el método de la potencia inversa (algoritmo 3) con $k = 4$ para ap[illegible]ximar el eigenvalor más cercano al origen, y compare su re[illegible]a con el eigenvalor verdadero para cada matriz.

25. La matriz del ejercicio 6
26. La matriz del ejercicio 7
27. La matriz del ejercicio 8
28. La matriz del ejercicio 9
29. La matriz del ejercicio 10
30. La matriz del ejercicio 11
31. La matriz del ejercicio 12
32. La matriz del ejercicio 13
33. La matriz del ejercicio 14
34. La matriz del ejercicio 15

En los ejercicios 35 a 39 aplique el método de la potencia inversa desplazada en la matriz asociada para aproximar la raíz más cercana a r de cada polinomio $p(x)$.

35. $p(x) = x^2 - 5x + 4, r = 5$
36. $p(x) = x^2 - 8x + 12, r = 7$
37. $p(x) = x^2 - 7x + 6, r = 5$
38. $p(x) = x^2 - 3x - 4, r = 5$
39. $p(x) = x^2 - 8x - 9, r = 10$

7.4 Aplicaciones a sistemas dinámicos

Objetivo del estudiante para esta sección
Aprender una de las aplicaciones más importantes de los eigenvalores.

Sistemas dinámicos discretos

Los eigenvalores y los eigenvectores se usan para resolver muchos problemas de física, matemáticas e ingeniería. Aquí los usaremos para estudiar con mayor amplitud los sistemas dinámicos discretos,[6] que fueron presentados en la sección 2.8.

Un **sistema dinámico**, o **ecuación de diferencias**, es una ecuación donde interviene una cantidad vectorial dependiente del tiempo, $\mathbf{x}(t)$. En un sistema dinámico *discreto* la variable tiempo es un entero, y $x(t)$ se representa con $\mathbf{x}_k$. Un **sistema dinámico homogéneo y discreto de primer orden** es una ecuación vectorial con la siguiente forma:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \tag{7.15}$$

en la que A es una matriz cuadrada fija, cuyo tamaño es igual al del vector $\mathbf{x}_k$. Sólo manejaremos A con elementos *reales* que *no* dependan de k.

[6] De nuevo recomendaremos *Discrete Dynamical Systems, Theory and Applications*, de James T. Sandefur (Oxford: Clarendon Press, 1990).

La ecuación (7.15) expresa el siguiente valor de $\mathbf{x}$ en función del actual; $\mathbf{x}_k$ puede calcularse por medio de aplicaciones repetidas de esa ecuación,

$$\mathbf{x}_k = A\mathbf{x}_{k-1} = A^2\mathbf{x}_{k-2} = \cdots$$

Por consiguiente,

$$\mathbf{x}_k = A^k\mathbf{x}_0 \tag{7.16}$$

La ecuación (7.16) se llama **solución** del sistema dinámico, y expresa $\mathbf{x}_k$ en función de un vector inicial $\mathbf{x}_0$.

El cálculo de $\mathbf{x}_k$ con la ecuación (7.16) tiene un inconveniente principal: el cálculo de A^k puede ser tedioso. Además, con frecuencia nos interesa el *comportamiento a largo plazo* del sistema. Esto es, deseamos conocer el vector límite

$$\lim_{k\to\infty} \mathbf{x}_k = \lim_{k\to\infty} A^k\mathbf{x}_0$$

si es que existe. Pero el cálculo de $A^k\mathbf{x}_0$ cuando k es grande es una tarea seria.

Supongamos que el vector $\mathbf{x}_0$ puede escribirse como una combinación lineal de los eigenvectores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ de A, por ejemplo

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Sean $\lambda_1, \ldots, \lambda_n$ los eigenvalores correspondientes. Entonces, de acuerdo con la ecuación (7.4), de la sección 7.1,

$$A^k\mathbf{x}_0 = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_k\lambda_n^k\mathbf{v}_n$$

En consecuencia, la solución del sistema puede simplificarse a

$$\mathbf{x}_k = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_k\lambda_n^k\mathbf{v}_n \tag{7.17}$$

La ecuación (7.17) no implica *potencias de matrices*, y su lado derecho es fácil de calcular, siempre que se conozcan las c_i, los $\mathbf{v}_i$ y las λ_i.

En el caso especial en el que la matriz A es *diagonalizable*, este método se aplica a *cualquier* vector n inicial $\mathbf{x}_0$. Porque entonces $\mathbf{R}^n$ tiene una base de eigenvectores de A y, por tanto, cualquier vector n puede escribirse como una combinación lineal de ellos. Hemos demostrado el teorema siguiente.

TEOREMA 13

> Sea A una matriz diagonalizable $n \times n$ con eigenvectores linealmente independientes $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y sus eigenvalores correspondientes $\lambda_1, \ldots, \lambda_n$. La solución del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ se expresa como sigue:
>
> $$\mathbf{x}_k = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n$$
>
> donde los coeficientes $c_1, \ldots, c_n$ son tales que
>
> $$\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Comportamiento de sistemas dinámicos a largo plazo

¿Qué sucede si k crece en un sistema dinámico? El teorema 12 de la sección 7.3 nos dice que si A es diagonalizable y tiene un eigenvalor dominante λ_1, con su correspondiente eigenvector $\mathbf{v}_1$, y si

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n, \quad \text{tal que } c_1 \neq 0,$$

a medida que k crece, la dirección de $\mathbf{x}_k$ tiende a la de $\mathbf{v}_1$.

Como ilustración, tenemos el caso con $A = \begin{bmatrix} -2 & -2 \\ -2 & 1 \end{bmatrix}$ (cuyos eigenvalores son −3, 2) y $\mathbf{x}_0 = \begin{bmatrix} 5 \\ 15 \end{bmatrix}$. Sabemos que la dirección de $A^k\mathbf{x}_0 = \mathbf{x}_{k+1}$ tiende hacia a la de $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, el eigenvector que corresponde a $\lambda = -3$.

Por ejemplo, para $k = 11$,

$$\mathbf{x}_{11} = \begin{bmatrix} -1\,423\,320 \\ -696\,300 \end{bmatrix} \quad \text{cuya relación de componentes es} \cong 2.0441.$$

En la figura. 7.5 se ven los puntos $\mathbf{x}_0, \ldots, \mathbf{x}_k$ para $k = 2, 3, 6$. Los puntos consecutivos se unen con un segmento de recta.

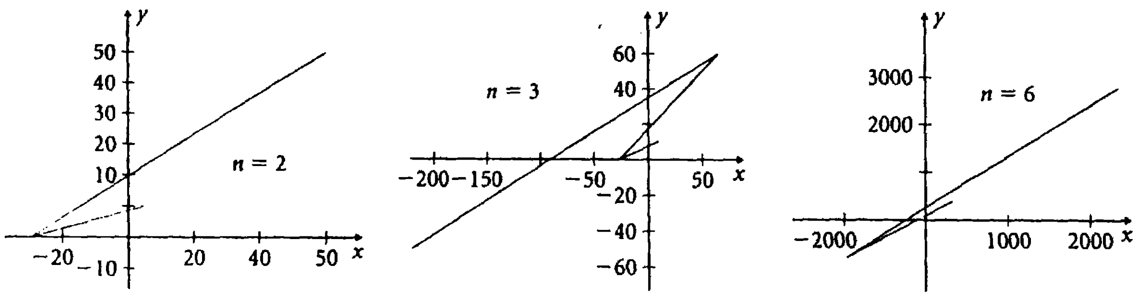


Figura 7.5

Atractores, repulsores y puntos silla

En los siguientes ejemplos describiremos el comportamiento a largo plazo, y graficaremos algunas de las soluciones del sistema dinámico para las matrices dadas A. Todas las gráficas incluyen $(\pm 1, \pm 1)$, $(\pm 1, \pm 2)$ y $(\pm 2, \pm 1)$ como puntos de partida. Los puntos consecutivos $\mathbf{x}_k, \mathbf{x}_{k+1}$ se unen con segmentos de recta. Las poligonales que resultan son las **trayectorias** de las soluciones.

■ EJEMPLO 27 Estudie las soluciones de los sistemas dinámicos definidos por las siguientes matrices.

(a) $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} .2 & 0 \\ 0 & .3 \end{bmatrix}$

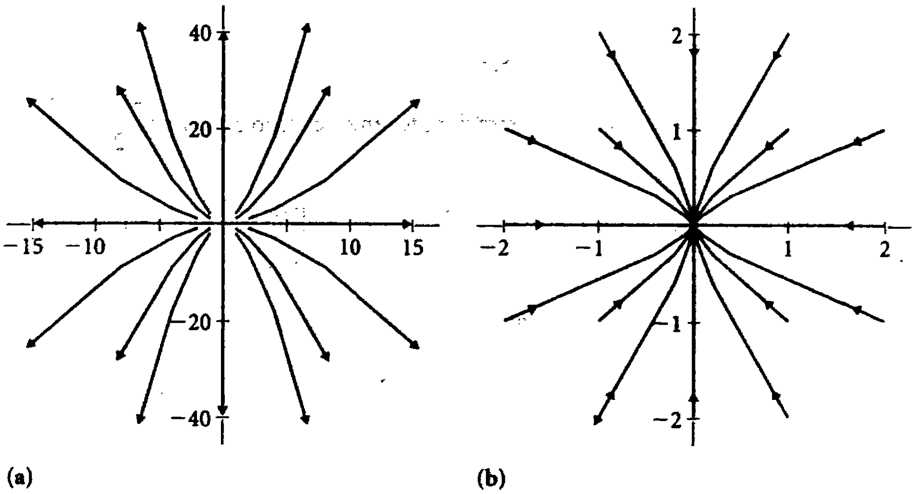


Figura 7.6

SOLUCIÓN

(a) Los eigenvalores son 3 y 2, y los eigenvectores correspondientes son (0, 1) y (1, 0). Por tanto, si

$$\mathbf{x}_0 = c_1(0,1) + c_2(1,0) = (c_2, c_1)$$

entonces

$$\mathbf{x}_k = c_1 3^k(0,1) + c_2 2^k(1,0) = \left(2^k c_2, 3^k c_1\right) \tag{7.18}$$

Es por eso que los componentes de $\mathbf{x}_k$ van hasta $\pm\infty$, dependiendo de los signos de c_1 y c_2. De acuerdo con el teorema 12, la dirección de $\mathbf{x}_k$ se aproxima a la de (0, 1) si $c_1 \neq 0$. A la larga, todas las trayectorias se hacen paralelas al eje y y *se alejan* del origen, exceptuando los puntos del eje x (donde $c_1 = 0$); los cuales se mueven hacia $\pm\infty$, dependiendo de si $\mathbf{x}_0$ es mayor o menor que 0. (¿Por qué?) En la figura. 7.6(a) se ven las trayectorias hasta $k = 2$.

(b) La segunda matriz tiene los mismos eigenvectores, y los eigenvalores son 0.3 y 0.2. Entonces,

$$\mathbf{x}_k = (0.2^k c_2, 0.3^k c_1) \quad \text{siendo} \quad \mathbf{x}_0 = (c_2, c_1)$$

como vimos anteriormente. Por tanto, ambos componentes de $\mathbf{x}_k$ tienden a cero. Según el teorema 12, la dirección de $\mathbf{x}_k$ se acerca a la de (0, 1) cuando k es grande. En consecuencia, las trayectorias se hacen paralelas al eje y y se mueven hacia el origen, excepto los puntos del eje x (donde $c_1 = 0$), los cuales permanecen en dicho eje y se desplazan hacia el origen. (¿Por qué?) (Figura 7.6(b), $k = 2$.) □

■ EJEMPLO 28 Estudie las soluciones de los sistemas dinámicos definidos por las siguientes matrices:

(a) $\begin{bmatrix} 2.5 & 0.5 \\ 0.5 & 2.5 \end{bmatrix}$

(b) $\begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{bmatrix}$

SOLUCIÓN

(a) Los eigenvalores de A son 3 y 2, con los eigenvectores correspondientes (1, 1) y (−1, 1). De modo que para

$$\mathbf{x}_0 = c_1(1,1) + c_2(-1,1) = (c_1 - c_2, c_1 + c_2)$$

se tiene

$$\mathbf{x}_k = c_1 3^k(1,1) + c_2 2^k(-1,1) = \left(c_1 3^k - c_2 2^k, c_1 3^k + c_2 2^k\right) \tag{7.19}$$

Cuando k se hace grande, $\mathbf{x}_k$ se dirige hacia (∞, ∞) si $c_1 > 0$ y hacia $(-\infty, -\infty)$ si $c_1 < 0$. Así, cuando $c_1 \neq 0$, todas las trayectorias terminan en el *primero* o en el *tercer* cuadrante. La dirección de $\mathbf{x}_k$ a largo plazo es paralela a (1, 1), de acuerdo con el teorema 12, con excepción de los puntos $\mathbf{x}_0$ con $c_1 = 0$. Son los puntos de la recta l que pasa por el origen y por (−1, 1). Permanecen en l, y su distancia al origen aumenta con k, porque

$$\mathbf{x}_k = \left(-c_2 2^k, c_2 2^k\right) = 2^k c_2(-1,1)$$

para $c_1 = 0$. (Figura 7.7(a), $k = 2$.)

(b) Para la segunda matriz tenemos los mismos eigenvectores y los eigenvalores son 0.6 y 0.4. De nuevo, las trayectorias tienden a hacerse paralelas a (1, 1), pero se mueven hacia el origen, a excepción de los puntos de la recta l (donde $c_1 = 0$), los cuales permanecen sobre l y también se mueven hacia el origen. (Figura 7.7(b), $k = 4$.) □

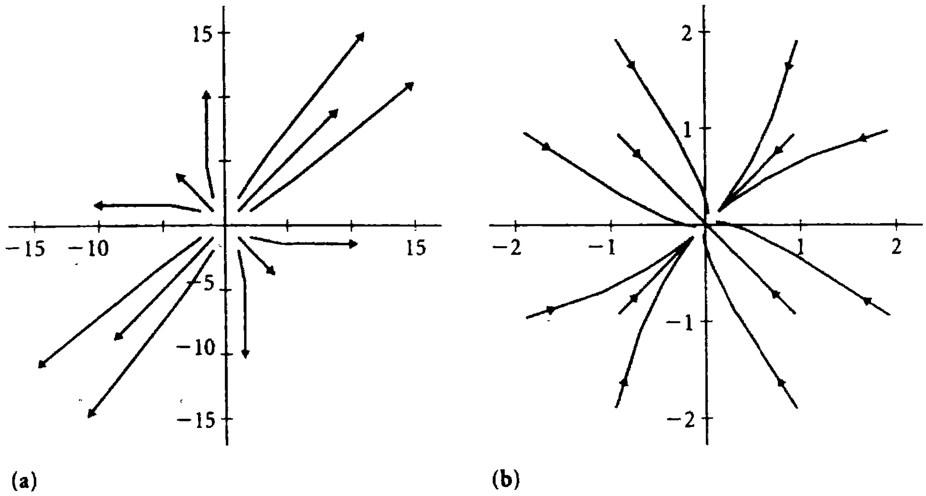


Figura 7.7

OBSERVACIONES

1. En los ejemplos 27 y 28 vimos que si todos los eigenvalores tienen valor absoluto menor que 1, todas las trayectorias tienden al origen. Se dice que el origen es un **atractor**. Esto se cumple en general, porque cada término de $\mathbf{x}_k = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n$ iría a cero si todas las $|\lambda_i| < 1$. Por otro lado, si todos los eigenvalores tienen valor absoluto mayor que 1, entonces todas las trayectorias se alejan del origen. En ese caso se dice que el origen es un **repulsor**. Por último, si algunas trayectorias se acercan al origen y otras se alejan de él, se dice que el origen es un **punto silla**.
2. Las trayectorias se hacen paralelas al eigenvector que tiene el eigenvalor con mayor valor absoluto, a excepción de los puntos que permanecen sobre la recta del otro eigenvector. Lo mismo es cierto (con zigzagueo) si uno o ambos eigenvalores son negativos. (Pruebe algunas trayectorias para $\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}$ y $\begin{bmatrix} -1 & -2 \\ 1 & -4 \end{bmatrix}$.)
3. Las gráficas del ejemplo 28 se parecen a las del ejemplo 27, donde las matrices eran diagonales. Ahora, los **eigenespacios** desempeñan los papeles de los ejes. Las gráficas para las matrices no diagonales pueden obtenerse a partir de las diagonales (o viceversa) mediante el **cambio de variables** que se describió en la sección 7.1.

De hecho, si A es cualquier matriz diagonalizable $n \times n$ y P es la matriz cuyas columnas son los elementos de una base $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de eigenvectores de A, definiremos las nuevas variables $\mathbf{y}_k$ de manera que

$$\mathbf{x}_k = P\mathbf{y}_k \tag{7.20}$$

Así, $\mathbf{y}_k = [\mathbf{x}_k]_{\mathcal{B}}$. En este caso, el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ puede escribirse como

$$P\mathbf{y}_{k+1} = AP\mathbf{y}_k \Rightarrow \mathbf{y}_{k+1} = P^{-1}AP\mathbf{y}_k$$

Por consiguiente,

$$\mathbf{y}_{k+1} = D\mathbf{y}_k \tag{7.21}$$

en la que D es, como de costumbre, la matriz diagonal cuyos elementos son los eigenvalores de A. El cambio de variables de la ecuación (7.20) transformó a $\mathbf{x}_{k+1} = A\mathbf{x}_k$ en un sistema dinámico **diagonal**, o **desacoplado,** ecuación (7.21). La ventaja de trabajar con sistemas desacoplados es que el i-ésimo componente de $\mathbf{y}_{k+1}$ sólo depende del i-ésimo componente de $\mathbf{y}_k$.

Para ejemplificar, si $A = \begin{bmatrix} 2.5 & 0.5 \\ 0.5 & 2.5 \end{bmatrix}$, $P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, y $\mathbf{x}_k$ es como en la ecuación (7.19), entonces

$$\mathbf{y}_k = P^{-1}\mathbf{x}_k = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} c_1 3^k - c_2 2^k \\ c_1 3^k + c_2 2^k \end{bmatrix} = \begin{bmatrix} c_2 2^k \\ c_1 3^k \end{bmatrix}$$

que es la fórmula (7.18) del ejemplo 27.

Veamos ahora el caso en el que un eigenvalor tiene un valor absoluto menor que 1 y otro mayor que 1.

■ EJEMPLO 29 Estudie las soluciones de la ecuación (7.15) para

$$A = \begin{bmatrix} 0.5 & 0 \\ 0 & 1.5 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

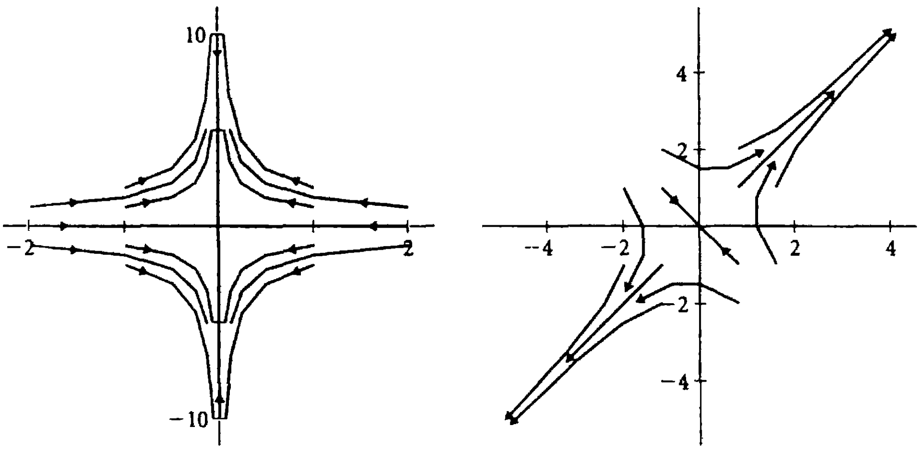
SOLUCIÓN Los eigenvalores, para ambas matrices, son 0.5 y 1.5. Los eigenvectores correspondientes $\mathbf{v}_1$ y $\mathbf{v}_2$ para la primera matriz son (0, 1) y (1, 0), y (1, 1) y (− 1, 1) para la segunda. En consecuencia,

$$\mathbf{x}_k = c_1(1.5)^k(0,1) + c_2(0.5)^k(1,0), \quad \text{siendo} \quad \mathbf{x}_0 = c_1(0,1) + c_2(1,0)$$

para la primera matriz y

$$\mathbf{x}_k = c_1(1.5)^k(1,1) + c_2(0.5)^k(-1,1), \quad \text{siendo} \quad \mathbf{x}_0 = c_1(1,1) + c_2(-1,1)$$

para la segunda. En cada caso, cuando $k \to \infty$, $(0.5)^k \to 0$ y $(1.5)^k \to \infty$, de modo que si c_1 no es cero, todas las trayectorias se hacen paralelas a $\mathbf{v}_1$. Si c_1 es cero, hay vectores a lo largo de la dirección de $\mathbf{v}_2$, y sus trayectorias se dirigen a cero. La figura 7.8(a) muestra las trayectorias para el primer sistema ($k = 4$), y la figura 7.8(b) para el segundo ($k = 3$). Esta vez el origen es un *punto silla*. ☐



(a) (b)

Figura 7.8

Eigenvalor repetido

Si nuestra matriz 2 × 2 sólo tiene un eigenvalor λ con dos eigenvectores linealmente independientes $\mathbf{v}_1$ y $\mathbf{v}_2$, entonces para $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ tenemos

$$\begin{aligned} \mathbf{x}_k &= c_1\lambda^k\mathbf{v}_1 + c_2\lambda^k\mathbf{v}_2 \\ &= \lambda^k(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) \\ &= \lambda^k\mathbf{x}_0 \end{aligned}$$

Por consiguiente, $\mathbf{x}_k$ y $\mathbf{x}_0$ están en la misma recta. Las partes (a) y (b) de la figura 7.9 lo muestran para las matrices $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ y $\begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}$. Para la primera matriz, el origen es un repulsor y para la segunda es un atractor.

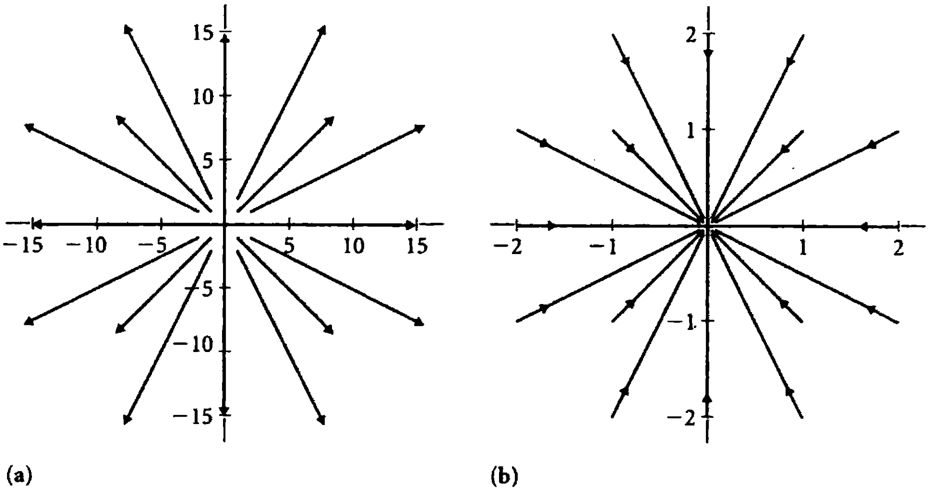


Figura 7.9

Eigenvalores complejos

Si A tiene eigenvalores complejos, las trayectorias suelen describir una espiral en torno al origen, acercándose o alejándose de él, lo cual depende de si la magnitud de los eigenvalores es mayor o menor que 1. También pueden describir círculos a su alrededor.

■ EJEMPLO 30 Estudie las soluciones de la ecuación (7.15) para

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

SOLUCIÓN

(a) Los eigenvalores de la primera matriz son $1 + i$, $1 - i$, cuyos eigenvectores correspondientes son $\mathbf{v}_1 = (i, 1)$ y $\mathbf{v}_2 = (-i, 1)$. Por consiguiente, si $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, entonces $\mathbf{x}_k$ tiene la forma

$$\mathbf{x}_k = c_1(1+i)^k \begin{bmatrix} i \\ 1 \end{bmatrix} + c_2(1-i)^k \begin{bmatrix} -i \\ 1 \end{bmatrix} \tag{7.22}$$

Observe que como los componentes de $\mathbf{x}_k$ son reales, el lado derecho de la ecuación (7.22) debe ser un número real. Veamos la trayectoria que comienza en $(1, 1)$. Puesto que

$$(1, 1) = c_1 \begin{bmatrix} i \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

implica que $c_1 = \frac{1}{2} - \frac{1}{2}i$ y $c_2 = \frac{1}{2} + \frac{1}{2}i$, entonces

$$\mathbf{x}_k = \left(\frac{1}{2} - \frac{1}{2}i\right)(1+i)^k \begin{bmatrix} i \\ 1 \end{bmatrix} + \left(\frac{1}{2} + \frac{1}{2}i\right)(1-i)^k \begin{bmatrix} -i \\ 1 \end{bmatrix} \tag{7.23}$$

Por tanto, para $k = 1, 2, 3, 4, \ldots$, obtenemos

$$\begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ -4 \end{bmatrix}, \ldots$$

Esos vectores tienen una magnitud cada vez mayor, y en consecuencia, las trayectorias describen una espiral que se aleja del origen, figura 7.10(a). De hecho, este comportamiento en espiral puede predecirse a partir de la ecuación (7.23), pero omitiremos los detalles.

(b) Para la segunda matriz, con un cálculo semejante se obtiene

$$\begin{aligned} \mathbf{x}_k = {} & c_1\left(\tfrac{1}{2} + \tfrac{1}{2}i\sqrt{3}\right)^k \begin{bmatrix} \frac{1}{2} - \frac{1}{2}i\sqrt{3} \\ 1 \end{bmatrix} \\ & + c_2\left(\tfrac{1}{2} - \tfrac{1}{2}i\sqrt{3}\right)^k \begin{bmatrix} \frac{1}{2} + \frac{1}{2}i\sqrt{3} \\ 1 \end{bmatrix} \end{aligned} \tag{7.24}$$

y para $\mathbf{x}_0 = (1, 1)$ es fácil demostrar que $c_1 = \frac{1}{2} + \frac{1}{6}i\sqrt{3}$ y $c_2 = \frac{1}{2} - \frac{1}{6}i\sqrt{3}$. En este caso, para $k = 0, \ldots, 6$, se obtiene

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ldots$$

Note que $\mathbf{x}_6$ es igual a $\mathbf{x}_0$. Por consiguiente, $\mathbf{x}_7$ es igual a $\mathbf{x}_1$, y así sucesivamente. Esta vez tenemos un **ciclo**-6, es decir, los vectores se repiten cada 6 unidades de tiempo, figura 7.10(b).

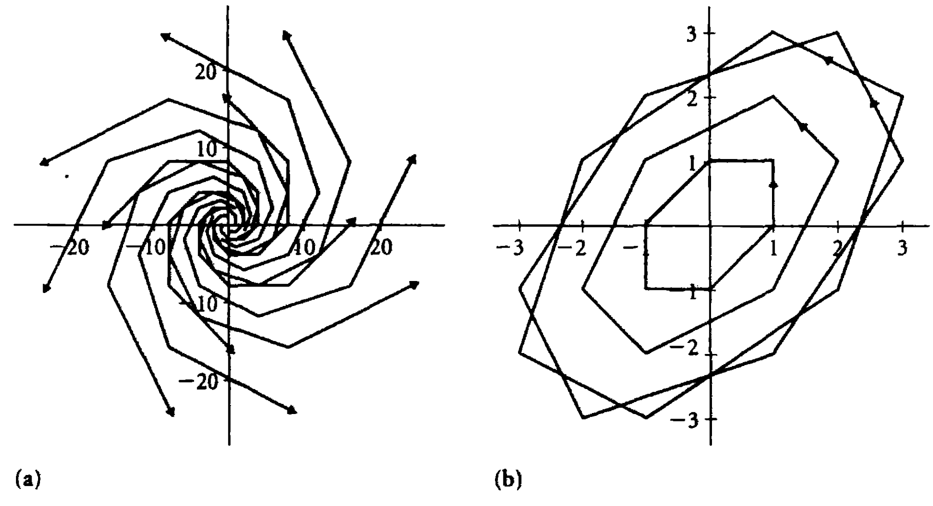


Figura 7.10

Así, para $k = 0, 1, 2, \ldots$

$$\mathbf{x}_{k+6} = \mathbf{x}_k$$

También puede escribirse

$$\mathbf{x}_k = \mathbf{x}_r$$

donde r es el residuo de la división de k entre 6. Por ejemplo,

$$\mathbf{x}_{44} = \mathbf{x}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Este comportamiento cíclico se debe a que los eigenvalores son raíces sextas de 1, es decir, $\left(\frac{1}{2} \pm \frac{1}{2} i\sqrt{3}\right)^6 = 1$. Así, en la ecuación (7.24) vemos que los valores de $\mathbf{x}_k$ se repiten cada vez que k aumenta en 6. Esto es válido para cualquier elección de c_1 y c_2.

Observe que no hay nada especial acerca de los ciclos-6. También existen casos con ciclos 2, 3, 4, . . .

Un problema de crecimiento de población

Aplicaremos nuestros métodos al problema de población de insectos de la sección 2.8.

Una población de insectos se divide en tres grupos de edad: A, B y C. El grupo A está formado por insectos de 0 a 1 semana de edad, en el grupo B están los de 1 a 2 semanas, y en el C se encuentran los de 2 a 3 semanas. Supongamos que los grupos tienen A_k, B_k y C_k insectos al final de la k-ésima semana. Deseamos estudiar cómo varían A, B y C a través del tiempo, dadas las dos condiciones siguientes:

1. (Tasa de supervivencia) Sólo el 10% del grupo A sobrevive una semana. En consecuencia,

$$B_{k+1} = \frac{1}{10}A_k \tag{7.25}$$

y sólo el 40% del grupo B sobrevive una semana. Así,

$$C_{k+1} = \frac{2}{5}B_k \tag{7.26}$$

2. (Tasa de nacimientos) Cada insecto del grupo A tiene $\frac{1}{2}$ de descendientes, cada uno del grupo B tiene 4 descendientes y cada uno de C tiene 5 descendientes. En la semana $k + 1$, los insectos del grupo A son descendientes de los insectos en la semana k. Por consiguiente,

$$A_{k+1} = \frac{2}{5}A_k + 4B_k + 5C_k \tag{7.27}$$

■ PROBLEMA Si inicialmente hay 1 000 insectos en cada grupo de edad, ¿cuál es la distribución de la población a largo plazo?

SOLUCIÓN Replantearemos las ecuaciones (7.25), (7.26) y (7.27) como el sistema dinámico

$$\mathbf{x}_{k+1} = A\mathbf{x}_k$$

siendo

$$A = \begin{bmatrix} \frac{2}{5} & 4 & 5 \\ \frac{1}{10} & 0 & 0 \\ 0 & \frac{2}{5} & 0 \end{bmatrix} \quad \text{y} \quad \mathbf{x}_k = \begin{bmatrix} A_k \\ B_k \\ C_k \end{bmatrix}, \qquad k = 0, 1, 2, \ldots,$$

En consecuencia,

$$\mathbf{x}_{k+1} = A^k \mathbf{x}_0$$

El cálculo de los eigenvalores de A da como resultado $\lambda = 1$, $\bar{r}/10$, donde $r = -3 - i\sqrt{11}$ y $\bar{r} = -3 + i\sqrt{11}$ (el conjugado de r). Los eigenvectores correspondientes son (50, 5, 2), $(r^2, r, 4)$ y $(\bar{r}^2, \bar{r}, 4)$. Por tanto, si

$$\mathbf{x}_0 = c_1 \begin{bmatrix} 50 \\ 5 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} r^2 \\ r \\ 4 \end{bmatrix} + c_3 \begin{bmatrix} \bar{r}^2 \\ \bar{r} \\ 4 \end{bmatrix} \tag{7.28}$$

entonces

$$\mathbf{x}_k = c_1 1^k \begin{bmatrix} 50 \\ 5 \\ 2 \end{bmatrix} + c_2 \left(\frac{r}{10}\right)^k \begin{bmatrix} r^2 \\ r \\ 4 \end{bmatrix} + c_3 \left(\frac{\bar{r}}{10}\right)^k \begin{bmatrix} \bar{r}^2 \\ \bar{r} \\ 4 \end{bmatrix} \tag{7.29}$$

Observe que $|r/10| = |\bar{r}/10| = 1/\sqrt{5} < 1$. De manera que los números positivos $|r/10|^k$ y $|\bar{r}/10|^k$ tienden a 0 cuando $k \to \infty$. Así, los números complejos

$$\left(\frac{r}{10}\right)^k \to 0, \qquad \left(\frac{\bar{r}}{10}\right)^k \to 0 \quad \text{cuando } k \to \infty$$

Entonces, cuando k es grande, la ecuación (7.29) se reduce a

$$\mathbf{x}_k \cong c_1 \begin{bmatrix} 50 \\ 5 \\ 2 \end{bmatrix} \tag{7.30}$$

Y para cualquier vector inicial dado $\mathbf{x}_0 = (A_0, B_0, C_0)$, basta calcular c_1 de la ecuación (7.28) y sustituirlo en la (7.30) para determinar $\mathbf{x}_k$ (cuando k es grande). Podemos despejar c_1, por ejemplo, con la regla de Cramer:

$$c_1 = \begin{vmatrix} A_0 & r^2 & \bar{r}^2 \\ B_0 & r & \bar{r} \\ C_0 & 4 & 4 \end{vmatrix} \begin{vmatrix} 50 & r^2 & \bar{r}^2 \\ 5 & r & \bar{r} \\ 2 & 4 & 4 \end{vmatrix}^{-1}$$

para obtener

$$c_1 = \frac{1}{90}(A_0 + 6B_0 + 5C_0)$$

después de una simplificación (bastante larga). Por consiguiente, para $\mathbf{x}_0 = (1\,000, 1\,000, 1\,000)$ resulta

$$\mathbf{x}_k \cong \frac{1}{90} \cdot 12 \cdot 1000 \cdot \begin{bmatrix} 50 \\ 5 \\ 2 \end{bmatrix} \cong \begin{bmatrix} 6666.66 \\ 666.66 \\ 266.66 \end{bmatrix}$$

En consecuencia, bajo las tasas dadas de supervivencia y natalidad de los insectos en los grupos A, B y C de edades, las poblaciones tienden a 6666.66, 666.66 y 266.66, respectivamente. En la figura 7.11 vemos que la trayectoria describe una espiral hacia el punto cuyas coordenadas son los números anteriores. □

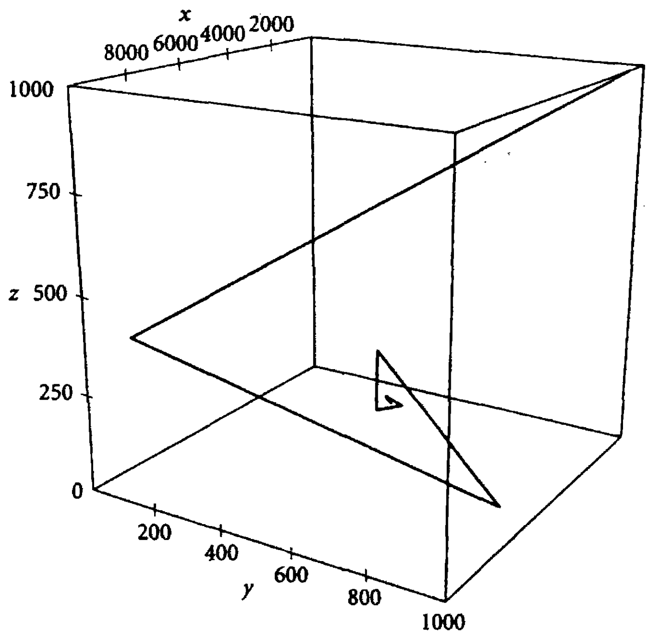


Figura 7.11

Ejercicios 7.4

En los ejercicios 1 a 9 suponga que la matriz A tiene los eigenvectores $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ con los eigenvalores correspondientes λ_1 y λ_2, que se dan. Considere el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, cuando el vector inicial es $x_0 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$.

(a) Deduzca una fórmula para $\mathbf{x}_k$.

(b) Calcule $A\mathbf{x}_0$ y $A^2\mathbf{x}_0$.

(c) Indique si el origen es un atractor, repulsor o ninguno de ellos.

1. $\lambda_1 = 1, \lambda_2 = 5$

2. $\lambda_1 = 2, \lambda_2 = 10$

3. $\lambda_1 = -7, \lambda_2 = -1$

4. $\lambda_1 = 1, \lambda_2 = 9$

5. $\lambda_1 = \frac{1}{2}, \lambda_2 = \frac{5}{2}$

6. $\lambda_1 = 2, \lambda_2 = 14$

7. $\lambda_1 = -13, \lambda_2 = -1$

8. $\lambda_1 = -\frac{1}{6}, \lambda_2 = \frac{5}{6}$

9. $\lambda_1 = -\frac{1}{10}, \lambda_2 = \frac{1}{2}$

En los ejercicios 10 a 15 considere el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, con la matriz dada A y con $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Determine $\mathbf{x}_5$ usando (a) $A^5\mathbf{x}_0$, y (b) eigenvalores.

10. $A = \begin{bmatrix} 7 & 6 \\ 4 & 5 \end{bmatrix}$

11. $A = \begin{bmatrix} -3 & 2 \\ 2 & -3 \end{bmatrix}$

12. $A = \begin{bmatrix} -6 & 5 \\ 2 & -3 \end{bmatrix}$

13. $A = \begin{bmatrix} \frac{3}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{3}{10} \end{bmatrix}$

14. $A = \begin{bmatrix} \frac{3}{2} & 1 \\ 1 & \frac{3}{2} \end{bmatrix}$

15. $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$

En los ejercicios 16 a 18 use eigenvalores y eigenvectores para calcular $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ para el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ con la matriz A dada y $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Trace la trayectoria que pasa por $\mathbf{x}_0$, . . ., $\mathbf{x}_3$ y determine si el origen es un atractor, repulsor o ninguno de ellos

16. $A = \begin{bmatrix} 2 & -2 \\ 2 & 2 \end{bmatrix}$

17. $A = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

18. $A = \begin{bmatrix} 0 & \frac{1}{2} \\ -2 & 1 \end{bmatrix}$

19. Demuestre que todas las soluciones del sistema dinámico

$$\mathbf{x}_{k+1} = \begin{bmatrix} 0 & 2 \\ -\frac{1}{2} & 1 \end{bmatrix} \mathbf{x}_k \text{ son ciclos-6.}$$

20. Compruebe que todas las soluciones del sistema dinámico

$$\mathbf{x}_{k+1} = \begin{bmatrix} \sqrt{3} & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}_k \text{ son ciclos-12.}$$

21. La k-ésima generación de una población animal consiste en A_k hembras y B_k machos. Suponga que la generación siguiente depende de la actual, de acuerdo con

$$A_{k+1} = 0.8A_k + 0.7B_k$$
$$B_{k+1} = 0.2A_k + 0.3B_k$$

Escriba este sistema dinámico en notación matricial. Si al principio había 100 hembras y 300 machos, ¿cuál es la población aproximada (a) inmediatamente después de la tercera generación? (b) ¿A la larga? ¿Cuál sexo predominará al final?

22. Repita el ejercicio 21 con las siguientes dependencias:

$$A_{k+1} = 0.7B_k$$
$$B_{k+1} = A_k + 0.3B_k$$

7.5 Aplicaciones a las cadenas de Markov

Objetivo del estudiante para esta sección

Aplicar los eigenvalores en el estudio del comportamiento de las cadenas de Markov a largo plazo.

Una de las aplicaciones más interesantes de los eigenvalores es el cálculo de etapas avanzadas de las cadenas de Markov que estudiamos en la sección 3.5. Recuerde que en una cadena de Markov, el estado siguiente de un sistema sólo depende de su estado actual. Para dar un ejemplo en esa sección regresaremos al estudio de los fumadores y los no fumadores.

Supongamos que la probabilidad de que un fumador continúe fumando un año después es 65%, mientras que la probabilidad de que un no fumador continúe sin fumar es 85%. Esta información se tabuló usando la matriz estocástica de las probabilidades de transición:

$$A = \begin{bmatrix} 0.65 & 0.15 \\ 0.35 & 0.85 \end{bmatrix}$$

Por ejemplo, el elemento 0.35 indica que un fumador tiene una probabilidad de 35% de dejar de fumar un año después, y 0.15 quiere decir que un no fumador tiene 15% de probabilidades de volverse fumador.

■ EJEMPLO 31 ¿Cuáles son los porcentajes de fumadores y no fumadores, a la larga, si inicialmente $100p\%$ son fumadores y $100q\%$ son no fumadores?

SOLUCIÓN Primero notaremos que $p + q = 1$. Recuerde que en la sección 3.5 se mencionó que los nuevos porcentajes, en k años, pueden calcularse como sigue:

$$\begin{bmatrix} 0.65 & 0.15 \\ 0.35 & 0.85 \end{bmatrix}^k \begin{bmatrix} p \\ q \end{bmatrix}$$

Por consiguiente, necesitamos el valor de este vector cuando k tiende a ∞. La diagonalización de A produce

$$A^k = \begin{bmatrix} 3 & -1 \\ 7 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}^k \begin{bmatrix} 3 & -1 \\ 7 & 1 \end{bmatrix}^{-1}$$

Como $\left(\frac{1}{2}\right)^k$ tiende a 0 cuando $k \to \infty$, entonces

$$\lim_{k\to\infty} A^k = \begin{bmatrix} 3 & -1 \\ 7 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 7 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0.3 & 0.3 \\ 0.7 & 0.7 \end{bmatrix}$$

Por tanto,

$$\lim_{k\to\infty} A^k \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} 0.3 & 0.3 \\ 0.7 & 0.7 \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} 0.3p + 0.3q \\ 0.7p + 0.7q \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}$$

ya que $p + q = 1$. Entonces, a largo plazo, los fumadores serán 30% en comparación con 70% de no fumadores. Esto es cierto para cualquier vector porcentaje inicial (p, q) con $p + q = 1$. $\square$

Un vector cuyos componentes son todos positivos y suman 1 se llama vector de **probabilidad**. Por ejemplo,

$$\begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}$$

son vectores probabilidad. Sin embargo, $\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0.8 \\ 0.1 \end{bmatrix}$ no lo son, porque $-1 < 0$ y también porque $0.1 + 0.8 \neq 1$. En el ejemplo 31 demostramos que para cualquier vector $\mathbf{v}$ de probabilidad, el límite de $A^k\mathbf{v}$ es $(0.3, 0.7)$ cuando $k \to \infty$.

Como segundo ejemplo, recordaremos los juegos entre el Ejército y la Marina de la sección 3.5. La probabilidad de que la Marina gane el siguiente año es 70%; la probabilidad de que la Marina sea el campeón un año y el Ejército lo sea el siguiente es 30%. Esto puede expresarse mediante la matriz doblemente estocástica de las probabilidades de transición:

$$B = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

■ EJEMPLO 32 (Juegos del Ejército contra la Marina) Dado que la Marina ganó el juego de este año, ¿cuál es la probabilidad de que gane después? ¿Y si la Marina perdió este año?

SOLUCIÓN La diagonalización de B da como resultado

$$B^k = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{2}{5} \end{bmatrix}^k \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1}$$

Ya que $\left(\frac{2}{5}\right)^k$ tiende a 0 cuando $k \to \infty$,

$$\lim_{k\to\infty} B^k = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

Por consiguiente, para toda p y q tales que $p + q = 1$,

$$\lim_{k\to\infty} B^k \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}$$

$$= \begin{bmatrix} 0.5p + 0.5q \\ 0.5p + 0.5q \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

Así, a largo plazo, la Marina ganará aproximadamente 50% de los juegos, si gana ($p = 1$, $q = 0$) o pierde ($p = 0$, $q = 1$) este año. □

Límites de las matrices estocásticas

Acabamos de ver cómo emplear la diagonalización para determinar límites de potencias de matrices estocásticas, pero ¿se ve claro que siempre existen esos límites?

Por ejemplo, veamos la matriz estocástica $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Entonces

$$B^2 = I, \qquad B^3 = B, \qquad B^4 = I, \qquad B^5 = B, \ldots$$

y es claro que no existe $\lim_{k\to\infty} B^k$, aun cuando B sea diagonalizable.

PREGUNTA ¿Cuándo tenemos la garantía que existe ese límite?

La clave para contestar esta pregunta reside en la definición siguiente: una matriz estocástica A es **regular** si cierta potencia A^k (k entero positivo) consiste en elementos estrictamente positivos.

$A = \begin{bmatrix} 0.5 & 1 \\ 0.5 & 0 \end{bmatrix}$ es regular, porque

$$A^2 = \begin{bmatrix} 0.75 & 0.5 \\ 0.25 & 0.5 \end{bmatrix}$$

sólo tiene elementos positivos. Por otro lado, B no es regular, porque todas sus potencias tienen *algunos* elementos cero.

■ EJEMPLO 33 Demuestre que A es regular.

$$A = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0 & 1 \\ 0 & 0.5 & 0 \end{bmatrix}$$

SOLUCIÓN Es fácil ver que A^4 es la primera potencia sólo con elementos positivos.

El siguiente teorema contesta la pregunta. Su demostración se puede encontrar en el libro *Finite Markov Chains*, de J. G. Kemeny y J. L. Snell (Nueva York: Springer-Verlag, 1976).

TEOREMA 14

Sea A una matriz estocástica regular $n \times n$. Entonces, cuando $k \to \infty$, A^k tiende a una matriz L, $n \times n$, de la forma

$$L = \begin{bmatrix} \mathbf{v} & \mathbf{v} & \cdots & \mathbf{v} \end{bmatrix}$$

donde $\mathbf{v}$ es un vector-n de probabilidad, con todos sus elementos mayores que 0.

Así, para cualquier matriz estocástica regular existe el límite L de potencias. Sin embargo, el cálculo de L empleando límites es bastante ineficiente. Un método mucho más eficiente es una consecuencia de nuestro resultado siguiente.

TEOREMA 15

Si A es una matiz estocástica regular, y L y $\mathbf{v}$ son como en el teorema 14, entonces

1. Para cualquier vector $\mathbf{x}_0$ de probabilidad inicial, $A^k\mathbf{x}_0$ tiende a $\mathbf{v}$ cuando $k \to \infty$, esto es,

$$\lim_{k\to\infty} (A^k\mathbf{x}_0) = \mathbf{v}$$

2. $\mathbf{v}$ es el único vector de probabilidad que satisface

$$A\mathbf{v} = \mathbf{v}$$

Así, $\mathbf{v}$ es un eigenvector de A con el eigenvalor $\lambda = 1$.

DEMOSTRACIÓN

1. Sea $\mathbf{x}_0 = (x_1, \ldots, x_n)$. Según el teorema 14,

$$\begin{aligned}\lim_{k\to\infty} (A^k\mathbf{x}_0) &= \left(\lim_{k\to\infty} A^k\right)\mathbf{x}_0 = L\mathbf{x}_0 \\ &= x_1\mathbf{v} + \cdots + x_n\mathbf{v} \\ &= (x_1 + \cdots + x_n)\mathbf{v} = \mathbf{v}\end{aligned}$$

porque $x_1 + \cdots + x_n = 1$.

2. Se tiene que

$$\mathbf{v} = \lim_{k\to\infty}(A^k\mathbf{x}_0) = \lim_{k\to\infty}(A^{k+1}\mathbf{x}_0) = A\lim_{k\to\infty}(A^k\mathbf{x}_0) = A\mathbf{v}$$

Se deja como ejercicio la demostración de la unicidad de $\mathbf{v}$.

Un vector $\mathbf{v}$ distinto de cero que satisface $A\mathbf{v} = \mathbf{v}$ se llama **vector de estado estable** (o **de equilibrio**) de A. Veamos ahora cómo calcular $\mathbf{v}$ sin límites. Como $\mathbf{v}$ es un eigenvector de A con eigenvalores 1, tan sólo resolvemos el sistema

$$(A - I)\mathbf{x} = \mathbf{0}$$

y elegimos la solución cuyos elementos sumen 1.

■ EJEMPLO 34 Determine v y L para

$$A = \begin{bmatrix} 0.5 & 1 \\ 0.5 & 0 \end{bmatrix}$$

SOLUCIÓN Tenemos que

$$[A - I : \mathbf{0}] = \begin{bmatrix} -0.5 & 1 & 0 \\ 0.5 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Así, la solución es $(2r, r)$, $r \in \mathbf{R}$. Se desea que

$$2r + r = 1$$

Por consiguiente, $r = \frac{1}{3}$, y así

$$\mathbf{v} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \quad \text{y} \quad L = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

(También podemos aplicar la diagonalización para obtener a L como el límite de A^k.)

NOTA La demostración de la propiedad 1 del teorema 15 muestra que si A es regular, entonces para *cualquier* vector inicial $\mathbf{x}_0$ (no necesariamente un vector de probabilidad), cuando $k \to \infty$:

$$A^k\mathbf{x}_0 - r\mathbf{v}$$

donde $r = x_1 + \cdots + x_n$. Así, para cualquier vector inicial, el sistema dinámico $\mathbf{x}_k = A^k\mathbf{x}_0$ tiene un límite, $r\mathbf{v}$, que es un vector de estado estable de A y que puede calcularse con facilidad.

Ejercicios 7.5

1. ¿Cuál(es) de los siguientes es (son) vector(es) de probabilidad?

$$\begin{bmatrix} \frac{1}{3} \\ \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

2. Demuestre que las matrices estocásticas siguientes son regulares.

a. $\begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix}$ **b.** $\begin{bmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{bmatrix}$

3. Compruebe que las matrices siguientes no son regulares.

a. $\begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 1 \end{bmatrix}$ **b.** $\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$

4. ¿Cuál(es) de las matrices siguientes es (son) regular(es)?

a. $\begin{bmatrix} 0.2 & 0.5 & 0 \\ 0.2 & 0.5 & 1 \\ 0.6 & 0 & 0 \end{bmatrix}$ **b.** $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

c. $\begin{bmatrix} 0 & 0.5 & 1 \\ 0 & 0.5 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ **d.** $\begin{bmatrix} 0 & 0.5 & 1 \\ 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0 \end{bmatrix}$

5. Determine los vectores de estado estable de las matrices del ejercicio 2.

6. Deduzca los vectores de estado estable de las matrices siguientes:

a. $\begin{bmatrix} 0.6 & 0.5 \\ 0.4 & 0.5 \end{bmatrix}$ **b.** $\begin{bmatrix} 0.7 & 0.5 \\ 0.3 & 0.5 \end{bmatrix}$

7. Encuentre los vectores de estado estable de las matrices siguientes:

a. $\begin{bmatrix} 0.2 & 0.5 & 0.5 \\ 0.2 & 0.5 & 0 \\ 0.6 & 0 & 0.5 \end{bmatrix}$ **b.** $\begin{bmatrix} 0.4 & 0.2 & 0.2 \\ 0.4 & 0.4 & 0.3 \\ 0.2 & 0.4 & 0.5 \end{bmatrix}$

8. Demuestre que $\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$ tiene un vector de estado estable aunque no sea regular.

9. Verifique la unicidad del vector de estado estable que afirma el teorema 15.

10. **(Demografía)** En cierta ciudad un residente tiene 40% de probabilidades de permanecer en ella después de 1 año, y

60% de mudarse al área suburbana que la rodea. Un residente suburbano tiene 20% de probabilidades de mudarse a una ciudad después de un año. Formule una matriz estocástica de transición de probabilidades para este caso. ¿Cuál es la distribución, a largo plazo, de una población que vive en esta ciudad y en las áreas suburbanas que la rodean?

11. (**Economía**) En la actualidad hay tres planes de inversión, A, B y C, disponibles para los empleados de una empresa. Un empleado sólo puede usar un plan a la vez, y puede cambiar de uno a otro sólo al final de cada año. La probabilidad de que alguien en el plan A continúe con él es 20%; de que elija el plan B es 50% y de que elija el plan C es 50%. La matriz M de probabilidades de transición para los empleados que participan es la siguiente.

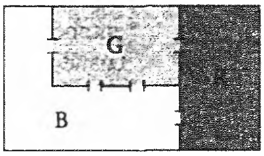
$$\begin{array}{cc} & \text{Este año} \\ & \begin{array}{ccc} \text{A} & \text{B} & \text{C} \end{array} \\ \text{Año próximo} \begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \end{array} & \begin{bmatrix} 0.2 & 0.5 & 0.5 \\ 0.2 & 0.5 & 0 \\ 0.6 & 0 & 0.5 \end{bmatrix} \end{array}$$

Demuestre que M es regular y determine su equilibrio. ¿Cuáles son los planes más popular y el menos popular a largo plazo?

12. (**Psicología**) Un psicólogo coloca 40 ratas en una caja con tres compartimientos coloreados: uno azul (B), uno verde (G) y uno rojo (R). Cada uno tiene puertas que conducen a los otros, como se ve en la figura de abajo. Las ratas se mueven constantemente hacia una puerta, de modo que la probabilidad de que permanezcan en un compartimento es 0. Una rata en B tiene $\frac{3}{4}$ de probabilidades de ir a G y $\frac{1}{4}$ de ir a R, según la distribución de las puertas. De igual forma, una rata en R tiene $\frac{1}{2}$ de probabilidades de ir a G y $\frac{1}{2}$ de ir a B. De este modo, la matriz de probabilidades de transición A tiene la forma

$$A = \begin{bmatrix} 0 & * & \frac{1}{2} \\ \frac{3}{4} & 0 & \frac{1}{2} \\ \frac{1}{4} & * & 0 \end{bmatrix}$$

Reemplace los asteriscos en A con las probabilidades correctas. Demuestre que A es regular y calcule su vector de estado estable. A largo plazo, ¿cuál es la distribución de las ratas? ¿Cuál es la probabilidad a largo plazo de que una rata esté en G?



7.6 Miniproyectos

1 ■ El teorema de Cayley-Hamilton

Si A es una matriz cuadrada de elementos escalares y $p(x)$ es un polinomio como

$$p(x) = a_0 + a_1x + \cdots + a_kx^k$$

entonces la matriz $p(A)$ está representada por

$$p(A) = a_0I + a_1A + \cdots + a_kA^k$$

Por ejemplo, si $A = \begin{bmatrix} 2 & -2 \\ 1 & 4 \end{bmatrix}$ y $p(x) = 1 - 3x + x^2$, tenemos que

$$p(A) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 3\begin{bmatrix} 2 & -2 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 2 & -2 \\ 1 & 4 \end{bmatrix}^2 = \begin{bmatrix} -3 & -6 \\ 3 & 3 \end{bmatrix}$$

En este proyecto usted comprobará el siguiente e importante teorema.

TEOREMA 16

(Cayley-Hamilton)

Toda matriz cuadrada satisface su ecuación característica. Así, si $p(x)$ es el polinomio característico de A, entonces

$$p(A) = \mathbf{0}$$

Para citar un ejemplo, es fácil observar que $p(x) = 10 - 6x + x^2$ es el polinomio característico de la matriz A. En ese caso

$$p(A) = 10\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 6\begin{bmatrix} 2 & -2 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 2 & -2 \\ 1 & 4 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Problema A

Compruebe el teorema de Cayley-Hamilton para las matrices siguientes.

1. $\begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}$, 2. $\begin{bmatrix} -5 & 6 \\ 8 & 1 \end{bmatrix}$, 3. $\begin{bmatrix} -1 & -1 & 0 \\ 1 & -\frac{3}{2} & \frac{3}{2} \\ 0 & 1 & -1 \end{bmatrix}$

A continuación deberá demostrar el teorema de Cayley-Hamilton para el caso especial en el que A es diagonalizable. Sólo siga las instrucciones.

Problema B

1. $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ Genera a $\mathbf{R}^n$ y sea B una matriz de $n \times n$ tal que

$$B\mathbf{v}_1 = \mathbf{0}, \ldots, B\mathbf{v}_k = \mathbf{0}$$

   Demuestre que B es la matriz cero.
2. Si λ *es* un eigenvalor de una matriz cuadrada A y su eigenvector correspondiente es $\mathbf{v}$, demuestre que para cualquier entero positivo k

$$A^k\mathbf{v} = \lambda^k\mathbf{v}$$

3. Sea A una matriz diagonalizable con polinomio característico $p(x)$. Compruebe para A el teorema de Cayley-Hamilton para dicha matriz como sigue: demuestre que para cualquier eigenvector $\mathbf{v}$ de A, el vector $p(A)\mathbf{v}$ es cero (aplicando la parte 2). A continuación use la parte 1 para llegar a la conclusión de que $p(A) = \mathbf{0}$.

A continuación lo conduciremos a que compruebe el teorema de Cayley-Hamilton para cualquier matriz cuadrada.

Si B es una matriz $n \times n$ cuyos elementos son polinomios en x, hay matrices únicas B_0, $B_1, \ldots, B_k$, con elementos escalares tales que

$$B = B_0 + B_1x + \cdots + B_kx^k$$

Por ejemplo,

$$\begin{bmatrix} 1 + x - 3x^2 & -1 + x \\ 2 + 5x & -6x + x^2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 5 & -6 \end{bmatrix}x + \begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}x^2$$

Gran parte de la aritmética para las matrices ordinarias se extiende a las matrices con elementos polinomiales. En particular, la fórmula siguiente generaliza el teorema 10, de la sección 6.3, a matrices B con elementos polinomiales.

$$\mathrm{Adj}(B)B = \det(B)I_n$$

En los siguientes renglones indicaremos cómo demostrar el teorema de Cayley-Hamilton para cualquier matriz A 2×2. Si el polinomio característico de A es

$$p(\lambda) = a + b\lambda + \lambda^2$$

entonces considere la matriz

$$B = A - \lambda I$$

Como el grado máximo en λ de los elementos de $\mathrm{Adj}(B)$ es 1, hay matrices únicas B_0 y B_1, con elementos escalares tales que

$$\mathrm{Adj}(B) = B_0 + B_1\lambda$$

En consecuencia,

$$\mathrm{Adj}(B)B = (B_0 + B_1\lambda)(A - \lambda I) = B_0A + (-B_0 + B_1A)\lambda - B_1\lambda^2$$

Por otro lado,

$$\mathrm{Adj}(B)B = \det(B)I = p(\lambda)I = aI + bI\lambda + I\lambda^2$$

Entonces, por la unicidad,

$$aI = B_0A, \qquad -B_0 + B_1A = bI, \qquad -B_1 = I$$

Así,

$$p(A) = aI + bA + A^2 = B_0A + (-B_0 + B_1A)A + A^2 = \mathbf{0}$$

Problema C

Siga los pasos anteriores para demostrar el teorema de Cayley-Hamilton para cualquier matriz cuadrada.

2 ■ Los números de Fibonacci (parte II)

En este proyecto regresamos a los números de Fibonacci $f_0, f_1, f_2, \ldots$ de la sección 3.6, con el fin de estudiar su comportamiento a largo plazo usando eigenvalores. Esos números surgen al contar los pares macho y hembra de conejos que se reproducen cada mes, y crean otro par macho-hembra. Recuerde que

$$f_0 = 1, \qquad f_1 = 1, \qquad f_k = f_{k-1} + f_{k-2}, \qquad k \geq 2$$

Para simplificar la notación agregaremos un número adicional $f_{-1} = 0$ a esta lista. Así, los primeros términos son

$$f_{-1} = 0, \quad f_0 = 1, \quad f_1 = 1, \quad f_2 = 2, \quad f_3 = 3, \quad f_4 = 5, \quad f_5 = 8, \ldots$$

Sean $\mathbf{x}_k = \begin{bmatrix} f_k \\ f_{k-1} \end{bmatrix}$, $k = 0, 1, \ldots,$ y $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

Problema

1. Demuestre que $\mathbf{x}_{k+1} = A\mathbf{x}_k$.
2. Compruebe (o aplique, si ya está demostrada) la identidad del ejercicio 32 de la sección 7.2,

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} r_1^{k+1} - r_2^{k+1} \\ r_1^k - r_2^k \end{bmatrix} \tag{7.31}$$

en donde r_1 y r_2 son los eigenvalores de A.

3. Concluya que para $k = -1, 0, 1, 2, \ldots,$

$$f_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^{k+1} - \left(\frac{1 - \sqrt{5}}{2} \right)^{k+1} \right) \tag{7.32}$$

4. Sin desarrollar, deduzca el entero

$$\frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^{8} - \left(\frac{1 - \sqrt{5}}{2} \right)^{8} \right)$$

5. Al observar que $r_2^k \to 0$ cuando $k \to \infty$ (porque $|r_2| < 1$), llegue a la conclusión que

$$\lim_{k \to \infty} \frac{f_k}{f_{k-1}} = \frac{1 + \sqrt{5}}{2} \cong 1.618$$

El número $r_1 = (1 + \sqrt{5})/2 \cong 1.618$ se llama **relación dorada**, o **promedio dorado (de Platón)**. Fue conocido por los griegos debido a sus propiedades interesantes. Los artistas saben que la relación de los lados en los rectángulos más agradable es $r_1 : 1$.

3 ■ Probabilidades de transición (parte II)

Regresamos al proyecto 3 de la sección 3.6, para contestar algunas otras preguntas.

Problema A

Un grupo de personas compra automóviles cada 4 años, que son de 3 marcas, A, B y C. Las probabilidades de transición de cambiar de una marca a otra se expresan con la matriz

$$R = \begin{bmatrix} 0.5 & 0.4 & 0.6 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.1 \end{bmatrix}$$

1. Emplee eigenvalores para calcular $\lim_{n \to \infty} R^n$.
2. Finalmente, ¿alguna de las marcas dominará el mercado, sin importar cuáles sean las ventas iniciales?

Problema B

La siguiente matriz estocástica de probabilidades de transición

$$T = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix}$$

expresa el flujo de clientes de y hacia los mercados A y B después de una compra. Recuerde que un **equilibrio de mercado** es un vector de las acciones (a, b) que permanece igual de una compra a la siguiente.

1. Demuestre que un equilibrio de mercado es un eigenvector de la matriz de probabilidades de transición. ¿Cuál es el eigenvalor correspondiente?
2. Compruebe que T tiene un equilibrio de mercado.
3. Calcule $\lim_{n\to\infty} T^n$.
4. Al final, ¿alguno de los mercados dominará al otro?

7.7 Ejercicios en computadora

En esta sección practicará con los comandos de su programa para calcular eigenvalores y eigenvectores. Al usarlos investigará otros temas. El ejercicio identificado con [S] requiere manipulación simbólica.[7]

Sean

$$A = \begin{bmatrix} 3 & 3 & 4 & 4 \\ 3 & 3 & 4 & 4 \\ 5 & 5 & 6 & 6 \\ 5 & 5 & 6 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} a & a & 1 \\ a & a & a \\ 1 & a & a \end{bmatrix}$$

$$C = \begin{bmatrix} 0.2 & 0.3 & 0.8 \\ 0.2 & 0.3 & 0.1 \\ 0.6 & 0.4 & 0.1 \end{bmatrix} \qquad R = \begin{bmatrix} 0.2 & 0 & 0.8 \\ 0 & 0 & 0.2 \\ 0.8 & 1 & 0 \end{bmatrix}$$

1. Sin hacer cálculos, determine un eigenvalor de A. En seguida use su programa para calcular numéricamente todos los eigenvalores y los eigenvectores básicos y, si es posible, exactamente. Confirme su respuesta demostrando que los eigenvalores calculados satisfacen la ecuación característica y que los eigenvectores básicos calculados sean realmente eigenvectores de A.
2. Diagonalice A determinando D y P. A continuación confirme su respuesta demostrando que $A = PDP^{-1}$.
3. [S] Calcule todos los valores de a para los que B tenga a cero como un eigenvalor.
4. Determine las raíces del polinomio $p(x) = x^5 - 15x^3 + 36x + 74$ en forma directa, y calculando los eigenvalores de la matriz compañera.
5. Deduzca una aproximación con cuatro cifras decimales para el límite de la matriz estocástica C, $\lim_{n\to\infty} C^n$: (a) directamente, calculando C^n para un valor grande de n, y (b) empleando eigenvalores.
6. Demuestre que R es una matriz regular. ¿Es cierto que si una matriz S $n \times n$ es regular, S^n sólo debe tener elementos distintos de cero? Examinando las matrices regulares y sus potencias, conjeture el entero positivo mínimo k tal que S^k sólo esté formado por elementos distintos de cero.
7. Compruebe el teorema de Cayley-Hamilton para A (analizado en el miniproyecto 1).

[7] Omita esos ejercicios si no dispone de manipulación simbólica.

8. Sea A_n la matriz $n \times n$ con elementos 1. Deduzca una fórmula para sus eigenvalores y eigenvectores básicos.

$$A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \ldots$$

9. Sea B_n una matriz de $n \times n$ con elementos diagonales n, y los elementos restantes 1. Deduzca una fórmula para sus eigenvalores y eigenvectores básicos.

$$B_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}, \ldots$$

El ejercicio siguiente está modelado de acuerdo con un ejemplo conocido en dinámica de población, atribuido a H. Bernadelli, P. H. Leslie y E. G. Leslie.

10. Una especie de escarabajos vive 3 años. Sean A, B y C las hembras de 0 a 1 año, 1 a 2 años y 2 a 3 años de edad, respectivamente. Ninguna hembra del grupo A se reproduce. Cada hembra del grupo B produce 8 hembras, y cada una del grupo C produce 24 hembras. Suponga que sólo sobreviven $\frac{1}{4}$ del grupo A para pasar al grupo B, y que sólo $\frac{1}{6}$ del grupo B sobrevive y pasa al grupo C. Si A_k, B_k y C_k son las cantidades de hembras en A, B y C después de k años, determine una matriz M tal que $M(A_k, B_k, C_k)$ sea $(A_{k+1}, B_{k+1}, C_{k+1})$. Si $A_0 = 100$, $B_0 = 40$ y $C_0 = 20$, use eigenvalores y eigenvectores para determinar si la especie se extinguirá.

Soluciones seleccionadas con Maple

Comandos charpoly, charmat, companion, eigenvals, Eigenvals, eigenvects.

```
# Ejercicio 1

# La matriz no es invertible (renglones repetidos) así que 0 es un eigenvalor.
A := matrix([[3, 3, 4, 4], [3, 3, 4, 4] , [5, 5, 6 ,6], [5, 5, 6, 6]]);
evas:= eigenvals(A);                    # Eigenvalores exactamente
evalf(");                               # y numéricamente.
eves := eigenvects (A);                 # Eigenvectores exactamente
evalf (");                              # y numéricamente.
# También se puede usar la versión inerte:
# Eigenvals (A); evalf ("); para la aproximación.
p:=charpoly (A,x);                      # Polinomio característico de A.
subs(x=evas[4], p);                     # Sustituyendo, por ejemplo, el 4o
simplify(")                             # eigenvalor y simplificando para obtener 0.
e := evas [3];                          # Elegir un eigenvalor y
v := eves [1] [3] [1];                  # el eigenvector correspondiente.
evalm(A &* v - e*v);                    # Calcular Av-ev y simplificar para
map(simplify, ");                       # obtener el vector cero.

# Ejercicio 3

B := matrix([[a, a, 1], [a, a, a], [1, a, a]]);
eigenvals(B);                           # Eigenvalores simbólicos.
solve("[1]); solve(" "[2]); solve(" " "[3]);   # Igualarlos a 0 y resolver. Así a=0,1.
```

```
# Ejercicio 4
p:=x^5-15*x^3+36*x+74;
solve(p);                      # Las raíces no pueden calcularse exactamente.
allvalues(");                  # Pero con allvalues podrán aproximarse.
eigenvals(companion p, x))     # Los eigenvalores de la TRANSPUESTA de la
allvalues(");                  # matriz asociada se aproximan. Igual respuesta.
# Ejercicio 5
C := matrix ([[.2, .3, .8] , [.2, .3, .1] , [.6, .4, .1]]);
evalm(C&^80);      # Produce columnas idénticas. Límite a la exactitud presentada.
v:=eigenvects (C);
DD:=diag(v[1] [1] , v[2] [1], v[3] [1]);      # DD, porque D ya lo usa Maple.
P:=concat (v[1] [3] [1], v [2] [3] [1], v[3] [3] [1]);      # P.
evalm(P&*DD&^80&*inverse (P));      # Igual respuesta con 6 cifras decimales.
# Ejercicio 6 - Parcial
R := matrix ([[.2, 0, .8] , [0, 0, .2] , [.8, 1, 0]]);
evalm (R&^2); evalm (R&^3); evalm(R&^4);      # R^4 tiene todos los elementos distintos de cero,
                                              # así que R es regular.
# Ejercicio 7
subs (x=A, charpoly(A, x));      # Sustituir la matriz en el polinomio característico.
evalm (");                       # Evaluar para obtener una matriz cero.
# Ejercicio 8 - Parcial
An := proc (n) local i, j; matrix (n, n, (i, j) -> 1) end:      # Definir A_n.
eigenvects (An(2)); eigenvects (An (3));      # Etcétera.
# Ejercicio 9 - Parcial
Bn: =proc (n) subs (nn=n, matrix(n, n, (i, j) -> if <>j then 1 else nn fi)) end:
eigenvects (Bn (2)) ; eigenvects (Bn (3));      # Etcétera.
# NOTA: proc(n) matrix (n,n, (i, j)->si i<>j then 1 else n fi))end: falla
# pase la n correcta en la diagonal debido a las reglas de Maple.
# Ejercicio 10 - Parcial
M := matrix ([[0, 8, 24] , [1/4, 0, 0][0, 1/6,0]] ;      # La matriz correcta.
```

Soluciones seleccionadas con Mathematica

Comandos Eigenvalues, Eigenvectors, Eigensystem.

```
(* Ejercicio 1 *)
(* La matriz es no invertible (renglones repetidos) así que 0 es un eigenvalor. *)
A = {{3, 3, 4, 4} , {3, 3, 4, 4} , {5, 5, 6, 6} ,{5, 5, 6, 6}}
evas = Eigenvalues [A]          (* Eigenvalores exactamente *)
N [%]                           (* y numéricamente. *)
Eigenvectors [A]                (* Eigenvectores exactamente. *)
eves = Simplify [%]             (* Se necesita simplificar la respuesta. *)
N [%]                           (* Eigenvectores numéricamente. *)
Eigensystem [A] // Simply       (* O bien, eigenvalores y eigenvectores juntos. *)
p=Det [A-x*IdentityMatrix[4]]   (* Polinomio característico. *)
p / . x->evas [[4]]             (* Sustituyendo, por ejemplo, el 4o *)
Simplify [%]                    (* eigenvalor y simplifique para obtener 0. *)
```

```
e = evas [[3]]                          (* Elija un eigenvalor y el  *)
v = eves [[3]]                          (* eigenvector correspondiente.  *)
A . v - e v                             (* Calcule Av-ev y simplifique para  *)
Simplify [%]                            (* obtener el vector cero.  *)

(* Ejercicio 3  *)

B = {{a, a, 1}, {a, a, a}, {1, a, a}}
ee = Eigenvalues [B] // Simplify        (* Eigenvalores simbólicos.  *)
Solve [ee[[1]]==0, a]                   (* Iguale a 0 y resuelva  *)
Solve [ee[[2]]==0, a]                   (* para obtener a=0,1.  *)
Solve [ee[[3]]==0, a]

(* Ejercicio 4  *)

p1 =x^5-15x^3+36x+74
Solve [p1==0, x]                                  (* Las raíces no pueden calcularse  *)
N [%]                                             (* de manera exacta, sino aproximada.  *)
Eigenvalues [N[{{0, 1, 0, 0, 0,} , {0, 0, 1, 0, 0},   (* Los eigenvalores numéricos  *)
{0, 0, 0, 1, 0,}, {0, 0, 0, 0, 1}{-74, -36, 0, 15, 0}}]]   (* de la matriz compañera.  *)

(* Ejercicio 5  *)

CC = {{.2, .3, .8} , {.2, .3,.1}{.6, .4, .1}}      (* Mathematica usa C.  *)
MatrixPower [CC, 80]     (* Produce columnas idénticas. Límite a la exactitud mostrada.  *)
v=Eigensystem [CC]                                   (* A continuación formar D y P.  *)
DD=DiagonalMatrix [{v [[1, 1]], v[[1, 2]], v[[1, 3]]}]  (* De se usa para diferenciación.  *)
P=Transpose [v[[2]]]                                 (* P.  *)
P . MatrixPower [DD, 80] . Inverse [P]     (* La misma respuesta con >15 cifras decimales.  *)

(* Ejercicio 6 - Parcial  *)

R = {{.2, 0, .8}, {0, 0, .2}, {.8, 1, 0}}          (* Todos los elementos de R^4 son distintos de 0
For [i=1, i <=4,i++, Print [MatrixPower[R, i]]]    de modo que R es regular. *)

(* Ejercicio 7  *)

p / . {Power -> MatrixPower, r-> A}     (* Sustituir A en el carácter poly y  *)
(* convierta Power a MatrixPower para obtener la matriz cero, como se espera.  *)

(* Ejercicio 8 - Parcial  *)

An [n_] := Table [1, {i, 1, n}, {j, 1, n}]
Eigensystem [An[2]]                 (* A continuación Eigensystem[An[[3]]], etcétera.  *)

(* Ejercicio 9 - Parcial  *)

Bn [n_] := Table [If[i==j, n, 1], {i, 1, n} , {j, 1, n}]
Eigensystem [Bn[2]]                 (* A continuación Eigensystem[Bn[[3]]], etcétera.  *)

(* Ejercicio 10 - Parcial  *)

M = {{0, 8, 24} , {1/4, 0, 0}, {0, 1/6, 0}}       (* La matriz correcta. Etcétera.  *)
```

Soluciones seleccionadas con MATLAB

Comandos balance, eig, poly, polyeig, polyval, polivalm, qz, y desde las herramientas simbólicas: Charpoly, eigensys, sym.

```
% Ejercicio 1
% La matriz es no invertible (renglones repetidos) de modo que un eigenvalor es 0.
```

```
A= [3 3 4 4; 3 3 4 4; 5 5 6 6; 5 5 6 6]; eig (A)      % Eigenvalores numéricos. También puede
[eves, evas] =eig(A)          % usar el formato [,]. eves es la matriz cuyas columnas son
% los eigenvectores, y evas es la matriz diagonal con eigenvalores en la diagonal.
eigensys(A)                   % (ST) Los eigenvalores simbólicamente.
[AA, BB]= eigensys (A)        % (ST) Los eigenvectores y eigenvalores simbólicamente.
p =poly (A)                   % El polinomio característico de A.
roots(p)                      % Todos los eigenvalores se obtienen como raíces de p.
polyval (p, evas (1, 1))      % También evaluar p en un eigenvalor. De 0 a 11 cifras decimales.
e = evas (1, 1), v = eves (:, 1)     % Elegir un eigenvalor y el eigenvector correspondiente.
A * v - e * v                        % La respuesta tiene de 0 a 13 cifras decimales.
```

```
% Ejercicio 3
B = sym('[a a 1; a a a; 1 a a] '), eigensys (B)      % (ST) Matriz y eigenvalores simbólicos.
solve (' a-1')                                 % (ST) copiar y resolver cada uno de los
solve (' a+1/2-1/2*(8*a^2+1)^(1/2)')           % (ST) eigenvalores simbólicos (aparecen
solve (' a+1/2+1/2*(8*a^2+1)^(1/2)')           % (ST) como cadenas) para obtener a = 0,1.
```

```
% Ejercicio 4
p1 = [1 0 -15 0 36 74] , roots(p1)    % ¡roots determina los eigenvalores de la asociada!
eig ([0 1 0 0 0; 0 0 1 0 0; 0 0 0 1 0; 0 0 0 0 1; -74 -36 0 15 0])       % Lo mismo.
```

```
% Ejercicio 5
C = [.2 .3 .8; .2 .3 .1; .6 .4 .1];      % Produce columnas idénticas.
[P, D] = eig(C)                          % P y D.
P * D^80 * P^(-1)                        % Igual respuesta mediante la diagonalización.
```

```
% Ejercicio 6 - Parcial
R = [.2 0 .8; 0 0 .2; .8 1 0]
R^2,R^3, R^4            % R^4 tiene todos los elementos distintos de cero, de modo que R es regular.
```

```
% Ejercicio 7
polyvalm (poly(A), A)      % Sustituir A en el polinomio característico poly(A) y
% evaluar en sentido matricial para obtener una matriz cercana a cero; polyval no funciona en este caso.
```

```
% Ejercicio 8 - Parcial
[P2, D2] = eig(ones(2)), [P3, D3] = eig(ones(3))      % An es sólo ones(n). Sin embargo, el cálculo
[P2, D2] = eigensys (ones (2))      % (ST) exacto ayuda a determinar una pauta con más facilidad.
```

```
% Ejercicio 9 - Parcial
% Para definir Bn crear un archivo de guión con nombre Bn.m que tenga los siguientes renglones:
function a = B (n)
for i=1 : n,
  for j=1 : n,
    if i==j
      a(i, j) = n;
        else a(i, j) =1;
      end
  end
end                                                % Ahora teclear
[P2, D2] = eig(Bn(2)), [P3, D3] = eig (Bn (3))     % Etcétera. Sin embargo, el
[P2, D2] = eigensys (Bn(2))      % (ST) cálculo exacto ayuda a determinar con patrón más fácil.
```

```
% Ejercicio 10 - Parcial
M = [0 8 24; 1/4 0 0; 0 1/6 0]         % La matriz correcta. Etcétera.
```

8

Productos punto e interno

El conocimiento que busca la geometría es el conocimiento de lo eterno.
Platón (Aprox. 429-347 a. C.), *La República*, VII, 52.

Introducción

En este capítulo estudiaremos muchas propiedades útiles del producto punto de $\mathbf{R}^n$. El material es de interés tanto para los lectores con inclinación teórica como para quienes buscan aplicaciones. También estudiaremos los espacios vectoriales que contienen un "producto interno", y que son generalización del producto punto para vectores abstractos. Los productos internos tienen una gama muy amplia de usos, desde en el análisis teórico hasta en el procesamiento aplicado de señales.

Mínimos cuadrados

Una de las aplicaciones más interesantes en este capítulo es el método de *mínimos cuadrados*. Con frecuencia, al tratar de comprender datos experimentales, deseamos determinar una recta o una curva que "encaje" o "se ajuste" más (o describa mejor) esos datos. Por ejemplo, imaginemos que un profesor de álgebra lineal mantiene las estadísticas (que se muestran a continuación) del porcentaje de calificaciones B otorgadas durante un periodo de seis semestres.

Semestre	1	2	3	4	5	6
Porcentaje de calificaciones B	0.20	0.25	0.20	0.30	0.45	0.40

Si el profesor quisiera trazar una recta que se acerque a los puntos en la tabla tendrá muchas opciones. Sin embargo, hay una que se ajusta mejor a esos datos, bajo cierto criterio. En la sección 8 veremos que esa recta es $y = 0.13333 + 0.05x$ (figura 8.1).

El método de los mínimos cuadrados fue inventado por Karl Friedrich Gauss, y lo usó para resolver un problema de astronomía. En 1801 el asteroide *Ceres* se había observado mucho más brillante durante más de un mes antes de desaparecer cuando se acercó al Sol. Con base en las observaciones disponibles, los astrónomos deseaban aproximar la órbita de Ceres para poder observarlo de nuevo cuando se alejara del Sol. Gauss empleó los mínimos cuadrados e impactó a la comunidad científica al predecir la hora y el lugar correctos (unos 10 meses después) para localizar al asteroide.

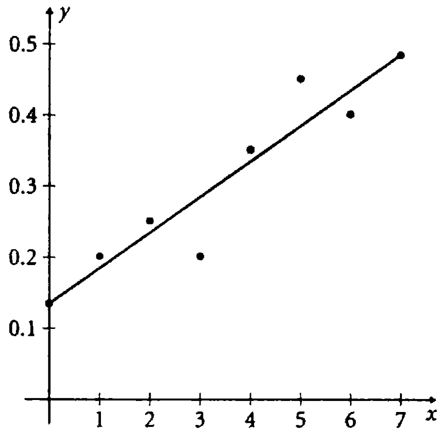


Figura 8.1 Recta que se ajusta mejor a los datos.

8.1 Conjuntos ortogonales y matrices

Objetivos del estudiante para esta sección

1. Definir y conocer las definiciones y las propiedades básicas de los conjuntos ortogonales y ortonormales.
2. Conocer la definición y las propiedades básicas de las matrices ortogonales.

En el capítulo 2 definimos a dos vectores como ortogonales si su producto punto es cero. En esta sección estudiaremos conjuntos completos de vectores cuya ortogonalidad se da en pares. Esos conjuntos se llaman ortogonales y comparten muchas propiedades interesantes que los hacen muy útiles en los cálculos.

Primero demostraremos una identidad que usaremos varias veces en este capítulo. Para cualquier matriz A $m \times n$, y un vector-n $\mathbf{u}$ y un vector-m $\mathbf{v}$, se cumple

$$(A\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (A^T\mathbf{v}) \tag{8.1}$$

DEMOSTRACIÓN

$$(A\mathbf{u}) \cdot \mathbf{v} = (A\mathbf{u})^T\mathbf{v} = (\mathbf{u}^TA^T)\mathbf{v} = \mathbf{u}^T(A^T\mathbf{v}) = \mathbf{u} \cdot (A^T\mathbf{v})$$

Conjuntos ortogonales

Se dice que un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de vectores-n es **ortogonal** si dos vectores distintos cualesquiera en él son ortogonales. Esto quiere decir que

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \text{si } i \neq j$$

Si $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \subseteq \mathbf{R}^3$ es ortogonal, entonces todos los pares posibles de vectores distintos, $\{\mathbf{v}_1, \mathbf{v}_2\}$, $\{\mathbf{v}_1, \mathbf{v}_3\}$ y $\{\mathbf{v}_2, \mathbf{v}_3\}$ deben ser ortogonales. En consecuencia, S forma un marco de coordenadas perpendiculares (figura 8.2)

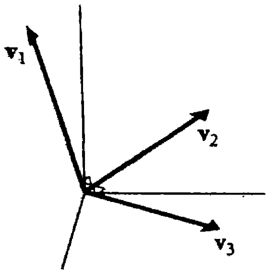


Figura 8.2 Vectores ortogonales.

■ EJEMPLO 1 Demuestre que $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \subseteq \mathbf{R}^4$ es ortogonal, siendo

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

SOLUCIÓN Esta afirmación es cierta, porque

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_2 &= 2 \cdot 0 + 2 \cdot 2 + 4 \cdot (-1) + 0 \cdot 1 &&= 0 \\ \mathbf{v}_1 \cdot \mathbf{v}_3 &= 2 \cdot (-2) + 2 \cdot 0 + 4 \cdot 1 + 0 \cdot 1 &&= 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_3 &= 0 \cdot (-2) + 2 \cdot 0 + (-1) \cdot 1 + 1 \cdot 1 &&= 0 \end{aligned}$$

Una de las propiedades más importantes de los conjuntos ortogonales se resume en el teorema siguiente. No debemos olvidar que para calcular los coeficientes de un vector $\mathbf{u}$ en forma de combinación lineal en $\mathbf{v}_1, \ldots, \mathbf{v}_k$ puede ser necesaria una tediosa reducción por operaciones de renglón. Sin embargo, si $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es ortogonal, hay una fórmula *fácil* para determinar esos coeficientes. $\square$

TEOREMA 1

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto ortogonal de vectores distintos de cero. Si $\mathbf{u}$ está en Gen(S) y

$$\mathbf{u} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \tag{8.2}$$

entonces

$$c_i = \frac{\mathbf{u} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}, \qquad i = 1, \ldots, k \tag{8.3}$$

DEMOSTRACIÓN Para una $i = 1, \ldots, k$ fija, se forma el producto punto de cada lado de la ecuación (8.2) con $\mathbf{v}_i$:

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v}_i &= (c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i \\ &= c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) \\ &= c_i(\mathbf{v}_i \cdot \mathbf{v}_i) \end{aligned}$$

porque $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ cuando $i \neq j$, por su ortogonalidad. Por consiguiente, $c_i = (\mathbf{u} \cdot \mathbf{v}_i)/(\mathbf{v}_i \cdot \mathbf{v}_i)$, como se afirmó. Observe que $\mathbf{v}_i \cdot \mathbf{v}_i = \|\mathbf{v}_i\|^2 \neq 0$, puesto que $\mathbf{v}_i \neq \mathbf{0}$. □

Los escalares $c_i = (\mathbf{u} \cdot \mathbf{v}_i)/(\mathbf{v}_i \cdot \mathbf{v}_i)$ pueden definirse para *cualquier* vector-n $\mathbf{u}$ (no tan sólo uno en Gen(S)), y con frecuencia se les llama **coeficientes de Fourier** de $\mathbf{u}$ con respecto a S.

Una consecuencia *importante* del teorema 1 es el siguiente.

TEOREMA 2

> Cualquier conjunto ortogonal $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ de vectores-n distintos de cero es linealmente independiente.

DEMOSTRACIÓN Igualemos a $\mathbf{0}$ una combinación lineal:

$$c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$$

De acuerdo con el teorema 1, con $\mathbf{u} = \mathbf{0}$,

$$c_i = \frac{\mathbf{0} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} = 0, \qquad i = 1, \ldots, k$$

Por consiguiente, S es linealmente independiente. □

Vemos que un conjunto ortogonal de vectores distintos de cero es una **base** para su generador, y que los coeficientes c_i de la ecuación (8.2) están determinados **en forma única** por la ecuación (8.3).

Si una base de un subespacio V de $\mathbf{R}^n$ es un conjunto ortogonal, se llama **base ortogonal**. Estas bases ortogonales son muy útiles, porque las coordenadas de los vectores pueden calcularse con facilidad con la ecuación (8.3).

■ EJEMPLO 2 Compruebe que el conjunto $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortogonal de $\mathbf{R}^3$. Escriba $\mathbf{u}$ en forma de combinación lineal de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} \frac{5}{7} \\ \frac{4}{7} \\ \frac{1}{7} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 12 \\ -6 \\ 6 \end{bmatrix}$$

SOLUCIÓN Es fácil ver que $\mathcal{B}$ es ortogonal. En consecuencia, $\mathcal{B}$ es un conjunto independiente de tres vectores-3, según el teorema 2. De manera que es una base ortogonal de $\mathbf{R}^3$. Sea $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Entonces

$$\begin{array}{lll} \mathbf{u} \cdot \mathbf{v}_1 = 42 & \mathbf{u} \cdot \mathbf{v}_2 = -24 & \mathbf{u} \cdot \mathbf{v}_3 = 6 \\ \mathbf{v}_1 \cdot \mathbf{v}_1 = 14 & \mathbf{v}_2 \cdot \mathbf{v}_2 = 12 & \mathbf{v}_3 \cdot \mathbf{v}_3 = 6/7 \end{array}$$

Por tanto, de acuerdo con la ecuación (8.3),

$$\begin{aligned} \mathbf{u} &= \frac{\mathbf{u} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{u} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 + \frac{\mathbf{u} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3}\mathbf{v}_3 \\ &= \frac{42}{14}\mathbf{v}_1 + \frac{-24}{12}\mathbf{v}_2 + \frac{6}{6/7}\mathbf{v}_3 \\ &= 3\mathbf{v}_1 - 2\mathbf{v}_2 + 7\mathbf{v}_3 \end{aligned}$$

Este cálculo es mucho más fácil que la reducción por operaciones de renglón de la matriz

$$[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \mathbf{u}]$$

TEOREMA 3

Sea $\mathcal{B}$ una base ortogonal de un subespacio V de $\mathbf{R}^n$. Si un vector $\mathbf{u}$ de V es ortogonal a cada vector de $\mathcal{B}$, entonces $\mathbf{u} = \mathbf{0}$.

DEMOSTRACIÓN Se deja como ejercicio. (*Sugerencia:* Aplique el teorema 1.)

Una propiedad interesante de los conjuntos ortogonales se relaciona con las matrices y sus transpuestas.

TEOREMA 4

Si las columnas de una matriz A $m \times n$ forman un conjunto ortogonal, entonces A^TA es una matriz diagonal $n \times n$. Con más precisión, si $A = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, entonces

$$A^TA = \begin{bmatrix} \|\mathbf{v}_1\|^2 & 0 & \cdots & 0 \\ 0 & \|\mathbf{v}_2\|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \|\mathbf{v}_n\|^2 \end{bmatrix}$$

En forma contraria, si la ecuación (8.4) es válida, en ese caso las columnas de A forman un conjunto ortogonal.

DEMOSTRACIÓN Sea c_{ij} el (i, j)-ésimo elemento de A^TA, y $\mathbf{r}_i$ el i-ésimo renglón de A^T. Tenemos que $\mathbf{r}_i = \mathbf{v}_i$ como vectores-n. Por definición de la multiplicación matricial,

$$c_{ij} = \mathbf{r}_i \cdot \mathbf{v}_j = \mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} \|\mathbf{v}_i\|^2 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

porque $\mathbf{v}_i \cdot \mathbf{v}_i = \|\mathbf{v}_i\|^2$ y $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, por la ortogonalidad. Se deja como ejercicio la demostración de la inversa.

Para ilustrarlo, si las columnas de A son los vectores del ejemplo 1:

$$A^TA = \begin{bmatrix} 2 & 2 & 4 & 0 \\ 0 & 2 & -1 & 1 \\ -2 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & -2 \\ 2 & 2 & 0 \\ 4 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 24 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

Es claro que los elementos diagonales son $\|\mathbf{v}_1\|^2$, $\|\mathbf{v}_2\|^2$ y $\|\mathbf{v}_3\|^2$, respectivamente.

Conjuntos ortonormales

Se dice que un conjunto de vectores es **ortonormal** si es ortogonal y está formado por vectores *unitarios* (figura 8.3). Así $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es ortonormal si

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0, \qquad i \neq j \qquad \text{and} \qquad \|\mathbf{v}_i\| = 1, \qquad i = 1, \ldots, k$$

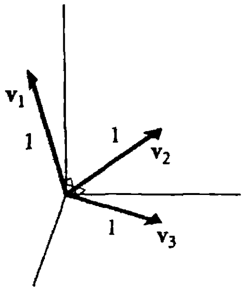


Figura 8.3 Vectores ortonormales.

Como $\|\mathbf{v}_i\| = 1 \Leftrightarrow \|\mathbf{v}_i\|^2 = \mathbf{v}_i \cdot \mathbf{v}_i = 1$, entonces

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \text{ ortonormal } \Leftrightarrow \mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases} \tag{8.5}$$

■ EJEMPLO 3 La base estándar de $\mathbf{R}^n$ es ortonormal. □

■ EJEMPLO 4 El conjunto $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ es ortonormal, donde

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Ya que $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ y $\|\mathbf{v}_1\| = 1$, y $\|\mathbf{v}_2\| = 1$. □

Cualquier conjunto ortogonal de vectores distintos de cero puede normalizarse para obtener un conjunto ortonormal:

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \text{ ortogonal} \Rightarrow \left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \ldots, \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|} \right\} \text{ortonormal}$$

Esto se debe a que cada vector $\mathbf{v}_i/\|\mathbf{v}_i\|$ es unitario y, para $i \neq j$,

$$\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|} \cdot \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|} = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\|\,\|\mathbf{v}_j\|} = 0$$

Al normalizar $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ del ejemplo 2 se obtiene el conjunto ortonormal $S' = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$,

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ -\frac{2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} \frac{5}{\sqrt{42}} \\ \frac{4}{\sqrt{42}} \\ \frac{1}{\sqrt{42}} \end{bmatrix} \tag{8.6}$$

Un conjunto ortonormal que es una base de un subespacio V de $\mathbf{R}^n$ se llama **base ortonormal** de V. Por ejemplo, S' es una base ortonormal de $\mathbf{R}^3$. Al expresar el teorema 1 para bases ortonormales toma la siguiente forma especial:

TEOREMA 5

Si $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es una base ortonormal del subespacio V de $\mathbf{R}^n$, entonces cada vector-n $\mathbf{u}$ de V puede escribirse en la forma única

$$\mathbf{u} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + \cdots + (\mathbf{u} \cdot \mathbf{v}_k)\mathbf{v}_k \tag{8.7}$$

De este modo es fácil calcular los componentes de un vector con respecto a una base ortonormal.

Siguiendo estas directrices también contamos con la muy útil desigualdad de Bessel, que aparece en muchas aplicaciones. En ésta se afirma que el cuadrado de la longitud de un vector es, cuando menos, la suma de los cuadrados de sus coeficientes de Fourier.

TEOREMA 6

(Desigualdad de Bessel)

Si $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un subconjunto ortonormal de $\mathbf{R}^n$ (no necesariamente una base) y $\mathbf{u}$ es cualquier vector-n. Entonces

$$\|\mathbf{u}\|^2 \geq (\mathbf{u} \cdot \mathbf{v}_1)^2 + \cdots + (\mathbf{u} \cdot \mathbf{v}_k)^2 \tag{8.8}$$

DEMOSTRACIÓN Vea el ejercicio 27. □

El teorema 4 también tiene un caso especial importante. La matriz $A = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ tiene columnas ortonormales si y sólo si cada elemento diagonal $\|\mathbf{v}_i\|^2$ de A^TA es 1. Esto equivale a que $A^TA = I$.

TEOREMA 7

Las columnas de una matriz A $m \times n$ forman un conjunto ortonormal (por consiguiente, $m \geq n$) si y sólo si

$$A^TA = I_n$$

Matrices ortogonales

Una matriz A es **ortogonal** si

1. es cuadrada, y
2. tiene columnas *ortonormales*.

■ EJEMPLO 5 I es ortogonal. También lo es la matriz cuyas columnas son $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$ de la ecuación (8.6). □

Quizá el mejor nombre de una matriz ortogonal sería *ortonormal*, pero este término no se usa. Observe que una matriz *no cuadrada* con columnas ortonormales *no* se llama ortogonal.

Nuestra primera observación acerca de las matrices ortogonales es que son **invertibles**, porque son cuadradas y sus columnas son linealmente independientes. De hecho, cuando $m = n$, el teorema 7 implica el resultado importante que sigue:

TEOREMA 8

Una matriz cuadrada A es ortogonal si y sólo si

$$A^T A = I \quad \text{es decir} \quad A^{-1} = A^T$$

Por consiguiente, la inversa de una matriz ortogonal es su transpuesta. En este caso no hay que hacer inversiones complicadas.

■ EJEMPLO 6 Demuestre que A y B son ortogonales y calcule sus inversas.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

SOLUCIÓN Es claro que las columnas de A son ortonormales, de modo que A es ortogonal. En consecuencia,

$$A^{-1} = A^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Las columnas de B son ortonormales, porque $\begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix} \cdot \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix} = 0$ y

$$\begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix} = \cos^2\theta + \operatorname{sen}^2\theta = 1 = \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix} \cdot \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix}$$

Por tanto,

$$B^{-1} = B^T = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

□

Observe que A es una matriz de *permutación*. En general, toda matriz de permutación es ortogonal.

■ EJEMPLO 7 Calcule la inversa de $A = [\mathbf{u}_1\, \mathbf{u}_2\, \mathbf{u}_3]$ de las ecuaciones (8.6).

SOLUCIÓN Como A es ortogonal,

$$A^{-1} = A^T = \begin{bmatrix} \frac{1}{\sqrt{14}} & -\frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{5}{\sqrt{42}} & \frac{4}{\sqrt{42}} & \frac{1}{\sqrt{42}} \end{bmatrix}$$

□

A la transformación matricial $A\mathbf{x}$ definida por una matriz ortogonal A también se le llama ortogonal. Las transformaciones matriciales ortogonales preservan los productos punto, y por ende también las longitudes y ángulos. A la inversa, si una transformación matricial preserva los productos punto, entonces su matriz (estándar) es ortogonal.

TEOREMA 9

> Sea A una matriz $n \times n$. Las siguientes afirmaciones son equivalentes:
>
> 1. A es ortogonal.
> 2. $A\mathbf{u} \cdot A\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ para todos los vectores-n $\mathbf{u}$ y $\mathbf{v}$ (preservación de los productos punto).
> 3. $\|A\mathbf{v}\| = \|\mathbf{v}\|$ para todo vector-n $\mathbf{v}$ (preservación de las normas).

DEMOSTRACIÓN

(1) $\Rightarrow$ (2) Si A es ortogonal, entonces $A^TA = I$. Así, según la ecuación (8.1),

$$A\mathbf{u} \cdot A\mathbf{v} = \mathbf{u} \cdot (A^T A\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$$

(2) $\Rightarrow$ (1) Supongamos que $A\mathbf{u} \cdot A\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$. En particular, $A\mathbf{e}_i \cdot A\mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{e}_j$. Pero la base estándar es ortonormal. De modo que,

$$A\mathbf{e}_i \cdot A\mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

que demuestra que A es ortogonal, de acuerdo con la ecuación (8.5), porque $A\mathbf{e}_i$ es la i-ésima columna de A.
(2) $\Leftrightarrow$ (3) La demostración de esta equivalencia se deja como ejercicio. $\square$

El teorema 9 tiene dos implicaciones interesantes.

TEOREMA 10

> 1. Si A y B son ortogonales $n \times n$, también lo es AB.
> 2. Si A es ortogonal, también lo es A^{-1}.

DEMOSTRACIÓN

1. De acuerdo con el teorema 9, basta con demostrar que AB preserva las normas:

$$\|AB\mathbf{v}\| = \|A(B\mathbf{v})\| = \|B\mathbf{v}\| = \|\mathbf{v}\|$$

La demostración de la afirmación 2 se deja como ejercicio. $\square$

NOTA La inversa, y en consecuencia la transpuesta, de una matriz A ortogonal es ortogonal, llegamos a la conclusión de que *los renglones de una matriz ortogonal también son ortonormales*.

La segunda implicación del teorema 9 es la siguiente:

TEOREMA 11

> Si λ es un eigenvalor de la matriz ortogonal A, entonces $\lambda = 1$.

DEMOSTRACIÓN Si $\mathbf{v}$ es un eigenvector de A, de acuerdo con la parte 3 del teorema 9, tenemos que

$$\|\mathbf{v}\| = \|A\mathbf{v}\| = \|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$$

De ahí que $|\lambda| = 1$, porque $\|\mathbf{v}\| \neq 0$.

El teorema 11 es válido aun para eigenvalores complejos de A. Por ejemplo, los eigenvalores de $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ son $1, -\frac{1}{2} + \frac{1}{2}i\sqrt{3}$ y $-\frac{1}{2} - \frac{1}{2}i\sqrt{3}$, y el valor absoluto de los tres es 1. (Compruébelo.) ☐

Ejercicios 8.1

En los ejercicios 1 a 4 demuestre que el conjunto de los vectores-n proporcionados es ortogonal. ¿Cuál(es) de ellos forma(n) una base ortogonal para $\mathbf{R}^n$?

1. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$

2. $\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 4 \end{bmatrix}$

3. $\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

4. $\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$

5. Cite un ejemplo de un conjunto de vectores $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ tal que los pares $\mathbf{v}_1, \mathbf{v}_2$ y $\mathbf{v}_2, \mathbf{v}_3$ sean ortogonales, pero que S no lo sea.

En los ejercicios 6 a 9 demuestre que cada conjunto de vectores forma una base ortogonal para $\mathbf{R}^3$. Emplee el teorema 1 para expresar $\mathbf{u} = (1, 1, 1)$ como una combinación lineal de esos vectores.

6. $\begin{bmatrix} 6 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \\ 20 \end{bmatrix}$

7. $\begin{bmatrix} 6 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

8. $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$

9. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$

En los ejercicios 10 a 14 determine si el conjunto ortogonal dado de vectores es ortonormal. Si no lo es, normalice los vectores para que lo sea.

10. $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

11. $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{2}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

12. $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$

13. $\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 15 y 16 demuestre que $\mathcal{B}$ es una base ortonormal para $\mathbf{R}^n$ (para n adecuada). Use el teorema 5 para escribir $\mathbf{e}_1 \in \mathbf{R}^n$ como combinación lineal en $\mathcal{B}$.

15. $\mathcal{B} = \left\{ \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} \right\}$

16. $\mathcal{B} = \left\{ \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} \right\}$

En los ejercicios 17 a 20 determine si la matriz dada es ortogonal. Si lo es, calcule su inversa.

17. $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

18. $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

19. $\begin{bmatrix} \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{21}} \\ -\frac{2}{\sqrt{14}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{21}} \\ \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{6}} & \frac{4}{\sqrt{21}} \end{bmatrix}$

20. $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$

21. Suponga que las columnas de una matriz A $m \times n$ forman un conjunto ortonormal. ¿Por qué $m \geq n$? Si A sólo tuviera columnas ortogonales, ¿eso implicaría que $m \geq n$?

22. Demuestre que los renglones de una matriz ortogonal $n \times n$ forman una base para $\mathbf{R}^n$.

23. Demuestre el teorema 3.

24. Termine la demostración del teorema 4.

25. Finalice la demostración del teorema 9.

26. Concluya la demostración del teorema 10.

27. Demuestre la desigualdad de Bessel planteada en el teorema 6. (*Sugerencia:* Sea $\mathbf{v} = \Sigma_{i=1}^{k}(\mathbf{u} \cdot \mathbf{v}_i)\,\mathbf{v}_i$ y sea $\mathbf{r} = \mathbf{u} - \mathbf{v}$. Compruebe que $\mathbf{r} \cdot \mathbf{v} = \mathbf{0}$. A continuación use el teorema de Pitágoras para llegar a la conclusión que $\|\mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{r}\|^2$.)

8.2 Proyecciones ortogonales: proceso de Gram-Schmidt

Objetivos del estudiante para esta sección

1. Comprender el concepto y las propiedades de una proyección ortogonal.
2. Aprender a determinar una base ortonormal a partir de una base dada de un subespacio de $\mathbf{R}^n$.

En esta sección estudiaremos *proyecciones* y *complementos* ortogonales. Ellos generalizan las nociones básicas del capítulo 2, como la proyección de un vector sobre otro, la proyección de un vector en un plano, el vector normal a un plano y a una línea, o un plano perpendicular a un vector. A continuación aprovecharemos nuestras nuevas herramientas para construir una base ortogonal, partiendo de una base ordinaria de cualquier subespacio de $\mathbf{R}^n$.

CONVENCIÓN

Todas las líneas y planos en esta sección pasan por el origen.

Complementos ortogonales

Sabemos que un vector normal a un plano es ortogonal a cualquier vector en ese plano. En general, si $\mathbf{u}$, vector-n, es ortogonal a todo vector de un subespacio V de $\mathbf{R}^n$, se dice que $\mathbf{u}$ es **ortogonal** a V. En la práctica, para comprobar que $\mathbf{u}$ es ortogonal a V no se calcula una cantidad infinita de productos punto, sino sólo los productos punto de $\mathbf{u}$ cuyos elementos son una *base* de V (o un conjunto generador finito).

TEOREMA 12

El vector-n $\mathbf{u}$ es ortogonal a $V = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq \mathbf{R}^n$ si y sólo si

$$\mathbf{u} \cdot \mathbf{v}_i = 0, \qquad i = 1, \ldots, k \tag{8.9}$$

DEMOSTRACIÓN Si $\mathbf{u}$ es ortogonal a V, entonces es válida la ecuación (8.9). A la inversa, si se supone la ecuación (8.9) y que $\mathbf{v}$ es cualquier elemento de V, entonces hay escalares c_i tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k$$

Al formar el producto punto con $\mathbf{u}$ se obtiene

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= \mathbf{u} \cdot (c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k)\\ &= c_1(\mathbf{u} \cdot \mathbf{v}_1) + \cdots + c_k(\mathbf{u} \cdot \mathbf{v}_k)\\ &= c_1 0 + \cdots + c_k 0 = 0\end{aligned}$$

Así, $\mathbf{u}$ y $\mathbf{v}$ son ortogonales, y en consecuencia $\mathbf{u}$ también lo es a V, como se afirmó. □

DEFINICIÓN

El conjunto de todos los vectores-n ortogonales a V se llama **complemento ortogonal** de V y se representa por $V^\perp$ (se lee "V perpendicular").

■ EJEMPLO 8 En $\mathbf{R}^3$, el complemento ortogonal de un plano que pasa por $\mathbf{0}$ es la recta que pasa por el origen y es perpendicular al plano. También, el complemento ortogonal de una recta que pasa por $\mathbf{0}$ es el plano que pasa por el origen y es perpendicular a la recta (figura 8.4). □

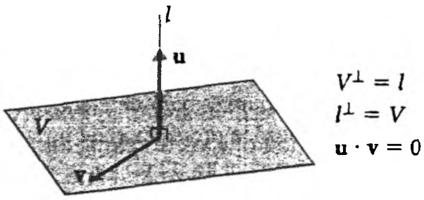


Figura 8.4 Complemento ortogonal de una recta y de un plano.

TEOREMA 13

> Sea V un subespacio de $\mathbf{R}^n$. Entonces
>
> 1. $V^{\perp}$ es un subespacio de $\mathbf{R}^n$;
> 2. $(V^{\perp})^{\perp} = V$.

DEMOSTRACIÓN

1. Si $\mathbf{u}_1$ y $\mathbf{u}_2$ están en $V^{\perp}$ y si $\mathbf{v}$ es cualquier vector de V, entonces $\mathbf{u}_1 \cdot \mathbf{v} = 0$ y $\mathbf{u}_2 \cdot \mathbf{v} = 0$. Para cualesquiera escalares c_1 y c_2,

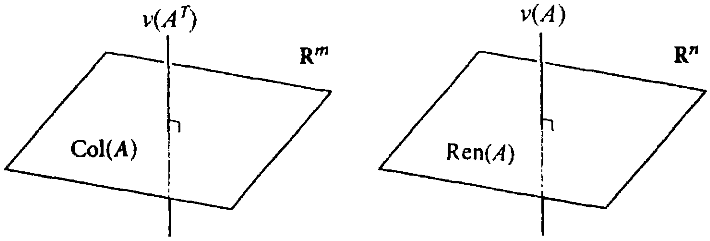
$$(c_1\mathbf{u}_1 + c_2\mathbf{u}_2) \cdot \mathbf{v} = c_1\mathbf{u}_1 \cdot \mathbf{v} + c_2\mathbf{u}_2 \cdot \mathbf{v} = c_1 0 + c_2 0 = 0$$

Por consiguiente, $c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ es ortogonal a cualquier vector de V y por tanto está en $V^{\perp}$. De manera que $V^{\perp}$ es un subespacio de $\mathbf{R}^n$.

2. Esta demostración se deja como ejercicio. $\square$

La noción de complemento ortogonal nos permite expresar una relación importante entre el espacio de columnas de una matriz y el espacio nulo de su transpuesta.

TEOREMA 14

> Sea A cualquier matriz $m \times n$.[1] Entonces
>
> $$(\mathrm{Col}(A))^{\perp} = \nu(A^T)$$
>
> Al pasar de las columnas a los renglones, también
>
> $$(\underbrace{\mathrm{Ren}(A)}_{EF})^{\perp} = \nu(A)$$
>
> 
>

DEMOSTRACIÓN Sea $A = [\mathbf{v}_1 \cdots \mathbf{v}_n]$. Entonces $\mathbf{u}$ está en $(\mathrm{Col}(A))^{\perp}$ si y sólo si $\mathbf{u}$ es ortogonal a todo vector de Col(A). En forma equivalente, $\mathbf{u}$ es ortogonal a las columnas de A (que generan a Col(A)). Así

$$\begin{aligned}
\mathbf{u} \in (\mathrm{Col}(A))^{\perp} &\Leftrightarrow \mathbf{u} \cdot \mathbf{v}_1 = 0, \ldots, \mathbf{u} \cdot \mathbf{v}_n = 0 \\
&\Leftrightarrow \mathbf{v}_1^T\mathbf{u} = 0, \ldots, \mathbf{v}_n^T\mathbf{u} = 0 \\
&\Leftrightarrow A^T\mathbf{u} = 0 \\
&\Leftrightarrow \mathbf{u} \in \nu(A^T)
\end{aligned}$$

$\square$

[1] Algunos autores llaman *teorema fundamental del álgebra lineal* a este teorema.

■ EJEMPLO 9 Compruebe el teorema 14 con

$$A = \begin{bmatrix} 1 & 2 \\ -2 & 0 \\ 1 & 4 \end{bmatrix}$$

SOLUCIÓN Según el teorema 12, basta demostrar que cada vector de alguna base de Col(A) es ortogonal a todo vector de una base de $v(A^T)$. Como las columnas de A son linealmente independientes, forman una base para Col(A). Reduciendo $[A^T : \mathbf{0}]$ es fácil demostrar que $\{(4, 1, -2)\}$ es una base para $v(A^T)$. Ahora comprobaremos que

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix} = 0 = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}$$

como afirma el teorema 14. $\square$

Proyecciones ortogonales

En la sección 2.2 vimos que, para los vectores dados en un plano o en el espacio, $\mathbf{u}$ y $\mathbf{v}$ ($\neq \mathbf{0}$), es posible expresar $\mathbf{u}$ como una suma de dos vectores *ortogonales*, $\mathbf{u}_{pr}$ y $\mathbf{u}_c$,

$$\mathbf{u} = \mathbf{u}_{pr} + \mathbf{u}_c$$

tales que $\mathbf{u}_c$ esté a lo largo de la recta definida por $\mathbf{v}$. En este caso, $\mathbf{u}_{pr}$ es la **proyección ortogonal** de $\mathbf{u}$ sobre, o a lo largo, de $\mathbf{v}$, y $\mathbf{u}_c$ es el **componente de u ortogonal** a $\mathbf{v}$. Esto se aplica en física, cuando se trata de descomponer un vector fuerza en componentes más manejables.

Además, vimos que $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ pueden calcularse muy bien en función de productos punto como sigue:

$$\mathbf{u}_{pr} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} \tag{8.10}$$

y

$$\mathbf{u}_c = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} \tag{8.11}$$

(véase la figura 8.5). El hecho de que $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ sean perpendiculares es consecuencia de

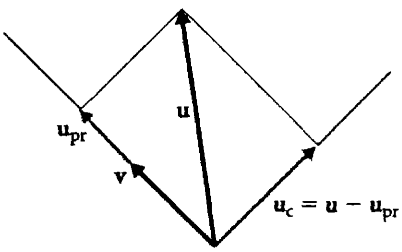


Figura 8.5 La proyección ortogonal $\mathbf{u}_{pr}$ de $\mathbf{u}$ sobre $\mathbf{v}$.

$$\mathbf{u}_c \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{v} - \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} \cdot \mathbf{v}$$
$$= \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{v} = 0$$

Observe que Gen$\{\mathbf{v}, \mathbf{u}\}$= Gen$\{\mathbf{v}, \mathbf{u}_c\}$, de acuerdo con la ecuación (8.11).

También vale la pena notar que $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ permanecen inalterados si reemplazamos $\mathbf{v}$ por $c\mathbf{v}$, $c \neq 0$, porque

$$\frac{\mathbf{u} \cdot c\mathbf{v}}{c\mathbf{v} \cdot c\mathbf{v}} c\mathbf{v} = \frac{c(\mathbf{u} \cdot \mathbf{v})}{c^2(\mathbf{v} \cdot \mathbf{v})} c\mathbf{v} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v}$$

Así, $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ *sólo* dependen de Gen$\{\mathbf{v}\}$ y no de $\mathbf{v}$ mismo.

Es evidente que la longitud de $\mathbf{u}_c$ es la distancia más corta de la punta de $\mathbf{u}$ a la recta l = Gen$\{\mathbf{v}\}$.

■ EJEMPLO 10 Calcule la distancia más corta d, de $\mathbf{u} = (1, -2, 3)$ a la recta $\mathbf{p} = (1, 1, 1)t$, $t \in \mathbf{R}$.

SOLUCIÓN Si $\mathbf{v} = (1, 1, 1)$, entonces $\mathbf{u} \cdot \mathbf{v} = 2$ y $\mathbf{v} \cdot \mathbf{v} = 3$. Por consiguiente,

$$\mathbf{u}_{pr} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} = \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)$$

Así (véase figura 8.6),

$$d = \|\mathbf{u}_c\| = \|\mathbf{u} - \mathbf{u}_{pr}\| = \left\|\left(\frac{1}{3}, -\frac{8}{3}, \frac{7}{3}\right)\right\| = \frac{\sqrt{114}}{3} \qquad \square$$

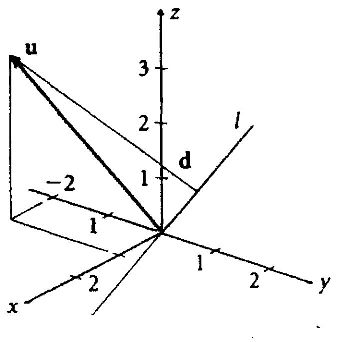


Figura 8.6 La longitud del componente ortogonal a l es la distancia más corta a l.

En lugar de proyectar vectores-2 o vectores-3 sobre una recta, veamos cómo podemos hacerlo con un vector-n sobre un subespacio V de $\mathbf{R}^n$ de tal manera que se preserven las proyecciones del plano o del espacio. Para ampliar las ecuaciones (8.10) y (8.11) con facilidad, supondremos que V tiene una base ortogonal. Esto no es una restricción porque, como veremos después, siempre existe una base ortogonal.

DEFINICIÓN

Sea $\mathbf{u}$ un vector-n y sea V un subespacio de $\mathbf{R}^n$ con una base ortogonal $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Entonces, la **proyección ortogonal** de $\mathbf{u}$ sobre V es el vector

$$\mathbf{u}_{pr} = \frac{\mathbf{u}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 + \cdots + \frac{\mathbf{u}\cdot\mathbf{v}_k}{\mathbf{v}_k\cdot\mathbf{v}_k}\mathbf{v}_k \tag{8.12}$$

La diferencia $\mathbf{u}_c = \mathbf{u} - \mathbf{u}_{pr}$ se llama la **componente de u ortogonal a** V.

$$\mathbf{u}_c = \mathbf{u} - \frac{\mathbf{u}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \cdots - \frac{\mathbf{u}\cdot\mathbf{v}_k}{\mathbf{v}_k\cdot\mathbf{v}_k}\mathbf{v}_k \tag{8.13}$$

$$\mathbf{u} = \mathbf{u}_{pr} + \mathbf{u}_c \tag{8.14}$$

Vemos que $\mathbf{u}_c$ es ortogonal a todos los vectores de $\mathcal{B}$.

$$\mathbf{u}_c\cdot\mathbf{v}_1 = 0, \ldots, \mathbf{u}_c\cdot\mathbf{v}_k = 0 \tag{8.15}$$

Tenemos

$$\begin{aligned}\mathbf{u}_c\cdot\mathbf{v}_i &= \mathbf{u}\cdot\mathbf{v}_i - \frac{\mathbf{u}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1\cdot\mathbf{v}_i - \cdots - \frac{\mathbf{u}\cdot\mathbf{v}_k}{\mathbf{v}_k\cdot\mathbf{v}_k}\mathbf{v}_k\cdot\mathbf{v}_i \\ &= \mathbf{u}\cdot\mathbf{v}_i - \frac{\mathbf{u}\cdot\mathbf{v}_i}{\mathbf{v}_i\cdot\mathbf{v}_i}\mathbf{v}_i\cdot\mathbf{v}_i \\ &= \mathbf{u}\cdot\mathbf{v}_i - \mathbf{u}\cdot\mathbf{v}_i = 0\end{aligned}$$

porque $\mathbf{v}_j\cdot\mathbf{v}_i = 0$, cuando $j \neq i$. Según el teorema 12, observamos que $\mathbf{u}_c$ es ortogonal a V. En consecuencia,

$$\mathbf{u}_{pr} \in V \quad \text{y} \quad \mathbf{u}_c \in V^{\perp} \tag{8.16}$$

La ecuación (8.13) implica que

$$\text{Gen}\{\mathbf{v}_1,\ldots,\mathbf{v}_k, \mathbf{u}\} = \text{Gen}\{\mathbf{v}_1,\ldots,\mathbf{v}_k, \mathbf{u}_c\} \tag{8.17}$$

Desde el punto de vista geométrico, si $n = 3$, notamos que $\mathbf{u}_{pr}$ es la suma vectorial de las proyecciones de $\mathbf{u}$ sobre $\mathbf{v}_1$ y $\mathbf{v}_2$, y está en el plano $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$. También, $\mathbf{u}_c$ es la normal a este plano tal que $\mathbf{u} = \mathbf{u}_{pr} + \mathbf{u}_c$ (figura 8.7).

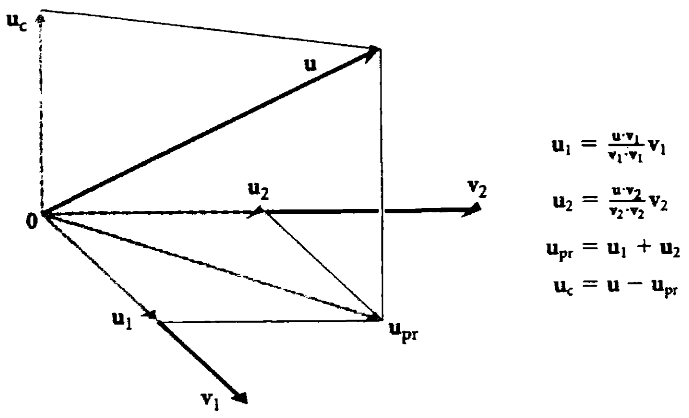


Figura 8.7 La proyección ortogonal, $\mathbf{u}_{pr}$, de $\mathbf{u}$ sobre $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$.

El vector $\mathbf{u}_{pr}$ tiene una propiedad de lo más interesante. Es el punto en V más cercano a $\mathbf{u}$, es decir, *la mejor aproximación de los vectores de V hacia* $\mathbf{u}$ (figura 8.8).

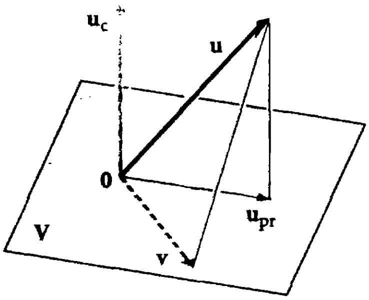


Figura 8.8 La proyección ortogonal $\mathbf{u}_{pr}$ es el único vector de V que es más cercano a **u.**

TEOREMA 15

(La mejor aproximación)

Con la notación precedente,

$$\|\mathbf{u}_c\| = \|\mathbf{u} - \mathbf{u}_{pr}\| < \|\mathbf{u} - \mathbf{v}\|$$

para cualquier vector $\mathbf{v}$ de V distinto de $\mathbf{u}_{pr}$.

DEMOSTRACIÓN Los vectores $\mathbf{u}_{pr} - \mathbf{v}$ y $\mathbf{u} - \mathbf{u}_{pr}$ son ortogonales, porque el primero está en V y el segundo en $V^{\perp}$. Por tanto, de acuerdo con el teorema de Pitágoras para vectores n (sección 2.2),

$$\begin{aligned}\|\mathbf{u} - \mathbf{u}_{pr}\|^2 + \|\mathbf{u}_{pr} - \mathbf{v}\|^2 &= \|(\mathbf{u} - \mathbf{u}_{pr}) + (\mathbf{u}_{pr} - \mathbf{v})\|^2 \\ &= \|\mathbf{u} - \mathbf{v}\|^2\end{aligned}$$

En consecuencia, $\|\mathbf{u} - \mathbf{u}_{pr}\| < \|\mathbf{u} - \mathbf{v}\|$, porque $\mathbf{u}_{pr} - \mathbf{v} \neq \mathbf{0}$. $\square$

■ EJEMPLO 11 Determine el vector, en el plano generado por los vectores ortogonales $\mathbf{v}_1 = (-1, 4, 1)$ y $\mathbf{v}_2 = (5, 1, 1)$ que se aproxime mejor a $\mathbf{u} = (1, -1, 2)$.

SOLUCIÓN El vector que necesitamos es $\mathbf{u}_{pr}$. Puesto que

$$\mathbf{u}\cdot\mathbf{v}_1 = -3 \qquad \mathbf{u}\cdot\mathbf{v}_2 = 6$$
$$\mathbf{v}_1\cdot\mathbf{v}_1 = 18 \qquad \mathbf{v}_2\cdot\mathbf{v}_2 = 27$$

entonces

$$\mathbf{u}_{pr} = \frac{\mathbf{u}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{u}\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2$$

$$= \frac{-3}{18}(-1,4,1) + \frac{6}{27}(5,1,1)$$

$$= \left(\frac{23}{18}, -\frac{4}{9}, \frac{1}{18}\right) \qquad \square$$

El teorema de la mejor aproximación implica que la proyección ortogonal $\mathbf{u}_{pr}$ es única, y que *no depende de la base ortogonal* de V que se usó para calcularla. Otra base ortogonal $\mathcal{B}'$ produciría el mismo $\mathbf{u}_{pr}$ y el mismo $\mathbf{u}_c$. De hecho, poco más adelante veremos que $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ sólo dependen de $\mathbf{u}$ y de V.

El proceso de Gram-Schmidt

En este párrafo describiremos un método *muy importante*, llamado **proceso de Gram-Schmidt**[2], que nos permite "ortogonalizar" cualquier base $\mathcal{B}$ de cualquier subespacio V de $\mathbf{R}^n$; es decir, transformar a ésta en una nueva base de V que tenga vectores ortogonales.

Sea V cualquier subespacio de $\mathbf{R}^n$ y $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ cualquier base de V. Se desea reemplazar en forma gradual a los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$ por los vectores $\mathbf{u}_1, \ldots, \mathbf{u}_k$ que sean ortogonales y que sigan formando una base de V. Primero, reemplazamos el conjunto $\{\mathbf{v}_1, \mathbf{v}_2\}$ por un conjunto ortogonal $\{\mathbf{u}_1, \mathbf{u}_2\}$ tal que $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\} = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$. Tan sólo hacemos que $\mathbf{u}_1$ sea $\mathbf{v}_1$, y que $\mathbf{u}_2$ sea el componente de $\mathbf{v}_2$ ortogonal a $\mathbf{v}_1$. Según la ecuación (8.15), $\{\mathbf{u}_1, \mathbf{u}_2\}$ es ortogonal. Según la ecuación (8.13), $\{\mathbf{v}_1, \mathbf{v}_2\}$ y $\{\mathbf{u}_1, \mathbf{u}_2\}$ tienen el mismo generador. También

$$\mathbf{u}_1 = \mathbf{v}_1$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1$$

Continuamos de la misma manera, para ortogonalizar el conjunto $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3\}$. Sustituimos $\mathbf{v}_3$ por $\mathbf{u}_3$, el componente de $\mathbf{v}_2$ ortogonal a $\text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$. Entonces $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es ortogonal y genera a $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ Además,

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2$$

(figura 8.9). Por inducción continuamos hasta que todo $\mathcal{B}$ queda reemplazado con $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, que es ortogonal y genera al generador de $\mathcal{B}$, cuya totalidad se encuentra en V. Es el proceso

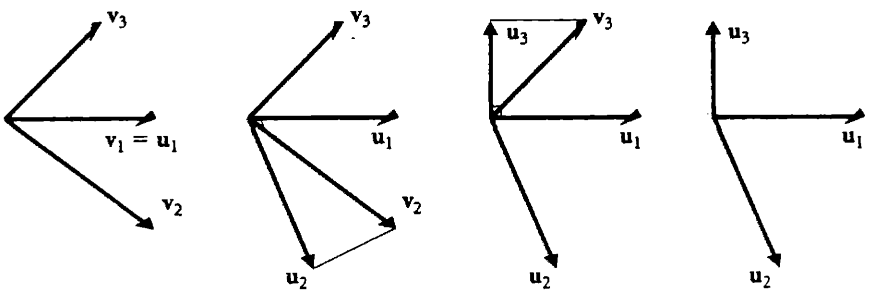


Figura 8.9 Etapas del proceso de Gram-Schmidt.

[2] En honor del matemático y actuario danés **Jörgen Pedersen Gram** (1850-1916) y **Erhard Schmidt**, matemático alemán (1876-1959).

completo. Si deseamos obtener una base ortonormal, tan sólo normalizamos a $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. Resumiendo:

TEOREMA 16

(Proceso de Gram-Schmidt)

Todo subespacio V de $\mathbf{R}^n$ tiene al menos una base ortogonal y una base ortonormal. Si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es cualquier base de V, entonces $\mathcal{B}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es una base ortogonal, donde

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \\
&\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \frac{\mathbf{v}_k \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_k \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \cdots - \frac{\mathbf{v}_k \cdot \mathbf{u}_{k-1}}{\mathbf{u}_{k-1} \cdot \mathbf{u}_{k-1}}\mathbf{u}_{k-1}
\end{aligned}$$

y

$$\mathrm{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_i\} = \mathrm{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_i\}, \qquad i = 1, \ldots, k$$

Una base ortonormal $\mathcal{B}''$ se obtiene normalizando $\mathcal{B}'$:

$$\mathcal{B}'' = \left\{\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \ldots, \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}\right\}$$

■ EJEMPLO 12 Determine una base ortogonal y una ortonormal de $\mathbf{R}^3$ aplicando el proceso de Gram-Schmidt a la base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, en la cual

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix}$$

SOLUCIÓN Sea $\mathbf{u}_1 = \mathbf{v}_1$. Como $\mathbf{v}_2 \cdot \mathbf{u}_1 = -6$ y $\mathbf{u}_1 \cdot \mathbf{u}_1 = 3$, en ese caso

$$\begin{aligned}
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\
&= \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix} - \frac{-6}{3}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}
\end{aligned}$$

Ya que $\mathbf{v}_3 \cdot \mathbf{u}_1 = -5$, $\mathbf{v}_3 \cdot \mathbf{u}_2 = -2$ y $\mathbf{u}_2 \cdot \mathbf{u}_2 = 2$, entonces

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2$$

$$= \begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix} - \frac{-5}{3}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \frac{-2}{2}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{8}{3} \\ \frac{4}{3} \\ -\frac{4}{3} \end{bmatrix}$$

Así, la base ortogonal es $\mathcal{B}' = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, donde

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} \frac{8}{3} \\ \frac{4}{3} \\ -\frac{4}{3} \end{bmatrix}$$

Por último, normalizamos para obtener una base ortonormal $\mathcal{B}''$:

$$\mathcal{B}'' = \left\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{bmatrix} \right\} \qquad \square$$

El teorema del proceso de Gram-Schmidt asegura que existen bases ortogonales. Esto, combinado con el teorema 3 de la sección 8.1, da lugar a la importante observación de que si $\mathbf{v}$ está en V *y también* en $V^\perp$, entonces $\mathbf{v} = \mathbf{0}$.

TEOREMA 17

Sea V cualquier subespacio de $\mathbf{R}^n$. Entonces

$$V \cap V^\perp = \{\mathbf{0}\}$$

La definición de los vectores $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ supuso inicialmente la existencia de una base ortogonal, pero ahora que lo comprobamos llegamos a la conclusión de que $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ *siempre* están definidas, para todo $\mathbf{u}$ y todo V. El teorema de la mejor aproximación implica que $\mathbf{u}_{pr}$ y $\mathbf{u}_c$ no dependen de la base ortogonal particular que los definió, sino sólo de $\mathbf{u}$ y de V.

TEOREMA 18

(Descomposición ortogonal)

Si $\mathbf{u}$ es cualquier vector-n y V es cualquier subespacio de $\mathbf{R}^n$, entonces $\mathbf{u}$ siempre tendrá una proyección ortógonal $\mathbf{u}_{pr}$ sobre V y un componente $\mathbf{u}_c$ ortogonal a V.

$$\mathbf{u} = \mathbf{u}_{pr} + \mathbf{u}_c, \quad \text{siendo} \quad \mathbf{u}_{pr} \in V, \mathbf{u}_c \in V^\perp \tag{8.18}$$

$\mathbf{u}_{pr}$ y $\mathbf{u}_c$ pueden determinarse con las ecuaciones (8.12) y (8.13), si se conoce una base ortogonal. Además, la descomposición representada por (8.18) es única. En otras palabras, si

$$\mathbf{u} = \mathbf{v} + \mathbf{v}^\perp \quad \text{siendo} \quad \mathbf{v} \in V, \quad \mathbf{v}^\perp \in V^\perp$$

entonces

$$\mathbf{v} = \mathbf{u}_{pr} \quad \text{y} \quad \mathbf{v}^\perp = \mathbf{u}_c$$

DEMOSTRACIÓN Lo único que necesitamos comprobar es la unicidad. Tenemos que

$$\mathbf{u} = \mathbf{u}_{pr} + \mathbf{u}_c = \mathbf{v} + \mathbf{v}^{\perp} \Rightarrow \mathbf{u}_{pr} - \mathbf{v} = \mathbf{v}^{\perp} - \mathbf{u}_c$$

Este vector común es cero, de acuerdo con el teorema 17, porque $\mathbf{u}_{pr} - \mathbf{v} \in V$ y $\mathbf{v}^{\perp} - \mathbf{u}_c \in V^{\perp}$. Por consiguiente, $\mathbf{v} = \mathbf{u}_{pr}$ y $\mathbf{v}^{\perp} = \mathbf{u}_c$, tal como se afirmó.

La descomposición única de $\mathbf{u}$ representada por la ecuación (8.18), con uno de los sumandos en V y otro en $V^{\perp}$, se llama **descomposición ortogonal** de $\mathbf{u}$ con respecto a V.

Observe que, en el caso especial en que $\mathbf{u}$ ya está en V, entonces $\mathbf{u}_{pr} = \mathbf{u}$, y $\mathbf{u}_c = \mathbf{0}$. Esto es consecuencia de $\mathbf{u} = \mathbf{u} + \mathbf{0}$ y de la unicidad de la descomposición, porque $\mathbf{u} \in V$ y $\mathbf{0} \in V^{\perp}$.

$$\text{Si } \mathbf{u} \in V \quad \Rightarrow \quad \mathbf{u}_{pr} = \mathbf{u} \quad \text{y} \quad \mathbf{u}_c = \mathbf{0}$$

■ EJEMPLO 13 Determine la descomposición ortogonal de $\mathbf{u} = (1, 1, 1)$ con respecto a $V = \mathrm{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_1$ y $\mathbf{v}_2$ los del ejemplo 12.

SOLUCIÓN En el ejemplo 12 ortogonalizamos $\{\mathbf{v}_1, \mathbf{v}_2\}$ y obtuvimos $\{\mathbf{u}_1, \mathbf{u}_2\}$, donde $\mathbf{u}_1 = (1, -1, 1)$ y $\mathbf{u}_2 = (0, 1, 1)$. Por consiguiente,

$$\mathbf{u}_{pr} = \frac{\mathbf{u} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{u} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{4}{3} \end{bmatrix}$$

$$\mathbf{u}_c = \mathbf{u} - \mathbf{u}_{pr} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}$$

Así,

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{4}{3} \end{bmatrix} + \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}$$

tenemos que $(\frac{1}{3}, \frac{2}{3}, \frac{4}{3}) \in V$ y $(\frac{2}{3}, \frac{1}{3}, -\frac{1}{3}) \in V^{\perp}$. (Para comprobarlo: $\mathbf{u}_{pr} = \frac{7}{3}\mathbf{v}_1 + \mathbf{v}_2$ y $\mathbf{u}_c \cdot \mathbf{v}_1 = 0 = \mathbf{u}_c \cdot \mathbf{v}_2$.) □

NOTA El proceso de Gram-Schmidt no se adapta bien a los cálculos numéricos grandes. Casi siempre hay una pérdida de ortogonalidad en las $\mathbf{u}_i$ calculadas. En la práctica se usan variantes del método QR que se describió en la sección 8.3, y una versión del proceso de Gram-Schmidt, llamada **Gram-Schmidt modificado**, que tiene mucho mejores propiedades numéricas.

Proceso de Gram-Schmidt con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> v1:=vector([1,-1,1]):v2:=vector([-2,3,-1]):
  v3:=vector([1,2,-4]):
> GramSchmidt({v1,v2,v3});
```

$$\left\{[1,-1,1],[0,1,1],\left[\frac{8}{3},\frac{4}{3},\frac{-4}{3}\right]\right\}$$

Mathematica

```
In[1]:=
   <<LinearAlgebra'Orthogonalization'
   GramSchmidt[{{1,-1,1},{-2,3,-1},{1,2,-4}}]
Out[2]=
          1            1            1
   {{---------, -(---------), ---------},
      Sqrt[3]      Sqrt[3]      Sqrt[3]

             1           1
      {0, ---------, ---------},
           Sqrt[2]     Sqrt[2]

            2        1            1
      {Sqrt[-], ---------, -(---------)}}
            3     Sqrt[6]      Sqrt[6]
```

MATLAB

```
>> orth([1 -2 1; -1 3 2; 1 -1 -4])
ans=
   0.1133   -0.7713    0.6262
  -0.6577    0.4142    0.6292
   0.7447    0.4832    0.4684
```

Ejercicios 8.2

Proyecciones ortogonales

1. Determine la proyección ortogonal de $\mathbf{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ sobre la recta l que pasa por $\mathbf{p} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y por el origen.

2. Con los datos del ejercicio 1, calcule la distancia mínima de $\mathbf{u}$ a l.

3. Determine la proyección de $\mathbf{u} = (3, -1, 2)$ sobre la recta $l = \{(1, 1, -3)t, t \in \mathbf{R}\}$.

4. Con los datos del ejercicio 3, calcule la distancia mínima de $\mathbf{u}$ a l.

En los ejercicios 5 y 6 determine la descomposición ortogonal de **u** con respecto a V.

5. $\mathbf{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, V = \text{Gen}\left\{\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right\}$

6. $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, V = \text{Gen}\left\{\begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}\right\}$

En los ejercicios 7 y 8 exprese **u** como una suma de dos vectores ortogonales, uno en V y el otro en $V^{\perp}$.

7. $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, V = \text{Gen}\left\{\begin{bmatrix} -2 \\ 2 \end{bmatrix}\right\}$

8. $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, V = \text{Gen}\left\{\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}\right\}$

En los ejercicios 9 y 10 determine el vector en el plano V que se aproxima mejor a **u**.

9. $V = \text{Gen}\left\{\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}\right\}, \mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

10. $V = \text{Gen}\left\{\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}\right\}, \mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

11. Finalice la demostración del teorema 13.

12. Compruebe el teorema 14 para $A = \begin{bmatrix} 4 & 2 \\ -2 & 2 \\ 1 & 2 \end{bmatrix}$.

Proceso de Gram-Schmidt

En los ejercicios 13 y 14 calcule una base ortogonal y una ortonormal de $\mathbf{R}^3$, aplicando el proceso de Gram-Schmidt a la base $\mathcal{B}$.

13. $\mathcal{B} = \left\{\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right\}$

14. $\mathcal{B} = \left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \\ -5 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right\}$

En los ejercicios 15 y 16 aplique el proceso de Gram-Schmidt para determinar una base ortogonal para V.

15. $V = \text{Gen}\left\{\begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right\}$

16. $V = \text{Gen}\left\{\begin{bmatrix} 3 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -2 \\ 2 \end{bmatrix}\right\}$

En los ejercicios 17 y 18 determine la descomposición ortogonal de **u** con respecto a V. En cada caso se necesitará aplicar el proceso de Gram-Schmidt para determinar una base ortogonal de V.

17. $\mathbf{u} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, V = \text{Gen}\left\{\begin{bmatrix} 5 \\ 1 \\ -4 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right\}$

18. $\mathbf{u} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 2 \end{bmatrix}, V = \text{Gen}\left\{\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}\right\}$

Distancia de un vector a un subespacio

Sea **u** un vector-n y V un subespacio de $\mathbf{R}^n$. Si $\mathbf{u}_{pr}$ es la proyección ortogonal de **u** sobre V, entonces, de acuerdo con el teorema de la mejor aproximación, $\mathbf{u}_c = \mathbf{u} - \mathbf{u}_{pr}$ tiene la longitud mínima de todos los vectores $\mathbf{u} - \mathbf{v}$, cuando **v** está en V. A la longitud $\|\mathbf{u}_c\|$ se le llama **distancia** de **u** a V. Observe que si la distancia de **u** a V es cero, entonces **u** está en V.

19. Calcule la distancia de (4, –4, 4) a

$$\text{Gen}\{(1, 1, -2), (2, 0, 1)\}$$

20. Calcule la distancia de (2, 0, 4) a

$$\text{Gen}\{(1, 1, -2), (2, 3, 1)\}$$

(Necesitará una base ortogonal.)

21. Calcule la distancia de (1, 2, 1, 2) a

$$\text{Gen}\{(1, 1, -2, 2), (2, 0, 1, 0)\}$$

8.3 La factorización QR

Objetivos del estudiante para esta sección

1. Aprender a determinar la factorización QR de las matrices que la tienen.
2. Conocer cómo aplicar el método QR para aproximar eigenvalores.

En esta sección estudiaremos un tema muy interesante que tiene muchas aplicaciones: la factorización QR de una matriz. Constituye un método importante para aproximar numéricamente eigenvalores y eigenvectores.

La factorización QR

La ortogonalización de las columnas de una matriz A conduce a cierta factorización de A que es muy útil en los cálculos numéricos, en especial para aproximar eigenvalores y eigenvectores como veremos poco después.

TEOREMA 19

(Factorización QR)

Si A es una matriz $m \times n$ con columnas linealmente independientes (en consecuencia, $m \geq n$), entonces A puede factorizarse en la forma

$$A = QR$$

en la que Q es una matriz con columnas ortonormales y R es una matriz triangular superior invertible.

DEMOSTRACIÓN Sean $\mathbf{v}_1, \ldots, \mathbf{v}_n$ las columnas de A y $\mathbf{u}_1, \ldots, \mathbf{u}_n$ los vectores obtenidos al ortonormalizarlas, en tal forma que $\mathrm{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_i\} = \mathrm{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_i\}$, $i = 1, \ldots, n$. Por ejemplo, el proceso de Gram-Schmidt seguido por la normalización garantiza estas condiciones (sección 8.2). Sea

$$Q = [\mathbf{u}_1\, \mathbf{u}_2 \cdots \mathbf{u}_n]$$

Cada $\mathbf{v}_i$ es una combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_i$ y, en consecuencia, una combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_n$ de la forma

$$\mathbf{v}_i = r_{1i}\mathbf{u}_1 + \cdots + r_{ni}\mathbf{u}_n = Q \begin{bmatrix} r_{1i} \\ \vdots \\ r_{ni} \end{bmatrix}, \qquad i = 1, \ldots, n \tag{8.19}$$

siendo

$$r_{i+1,i} = \cdots = r_{ni} = 0, \qquad i = 1, \ldots, n \tag{8.20}$$

Por tanto,

$$A = [\mathbf{v}_1 \cdots \mathbf{v}_n] = \left[Q \begin{bmatrix} r_{11} \\ \vdots \\ 0 \end{bmatrix} \cdots Q \begin{bmatrix} r_{1n} \\ \vdots \\ r_{nn} \end{bmatrix} \right] = QR$$

donde R es la matriz cuyo (i, j)-ésimo elemento es r_{ij}, $i, j = 1, \ldots, n$. Q y R son las matrices que buscamos. Q tiene columnas ortonormales y R es triangular superior, de acuerdo con la ecuación (8.20). También, R es *invertible*, porque el sistema homogéneo $R\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial. Si no lo fuera, el sistema $QR\mathbf{x} = \mathbf{0}$, o $A\mathbf{x} = \mathbf{0}$, tendría una solución no trivial y las columnas de A serían linealmente *dependientes*. $\square$

NOTAS

1. Es fácil, pero no necesario, llegar a fórmulas para Q y R con base en las ecuaciones del proceso de Gram-Schmidt. Lo que se hace en la práctica es ortonormalizar las columnas de A para obtener Q. A continuación se determina R con

$$R = Q^T A$$

porque

$$Q^T A = Q^T(QR) = (Q^T Q)R = IR = R$$

y $Q^T Q = I$, según el teorema 7 de la sección 8.1.
2. La matriz R puede arreglarse de tal forma que sus elementos diagonales sean siempre **estrictamente positivos**. Si $r_{ii} < 0$ en la ecuación (8.19), se reemplaza $\mathbf{u}_i$ con $-\mathbf{u}_i$. Al hacerlo, Q es única, porque al ortonormalizar las $\mathbf{u}_i$ son únicas incluyendo al signo.
3. Las columnas de Q forman una **base ortonormal** para Col(A). Además,

$$\text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_i\} = \text{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_i\}$$

para $i = 1, \ldots, n$.
4. En el caso especial cuando A es cuadrada, Q es una matriz **ortogonal**.

■ EJEMPLO 14 Determine la factorización QR de A.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

SOLUCIÓN Notaremos primero que las columnas $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ de A son linealmente independientes; en consecuencia, existe una factorización QR. A continuación necesitamos "ortonormalizar" $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Con el proceso de Gram-Schmidt llegamos a

$$\mathbf{u}_1 = \mathbf{v}_1 = (1,1,1,1), \qquad \mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = \left(\frac{1}{2}, -\frac{3}{2}, \frac{1}{2}, \frac{1}{2}\right)$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \left(-\frac{2}{3}, 0, \frac{1}{3}, \frac{1}{3}\right)$$

y formamos la matriz

$$Q = \left[\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \; \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \; \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}\right]$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{6} & -\frac{\sqrt{6}}{3} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{6} \\ \frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{6} \end{bmatrix}$$

Como $R = Q^T A$, entonces

$$R = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} \\ -\frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 & 1 \\ 0 & \sqrt{3} & \frac{\sqrt{3}}{3} \\ 0 & 0 & \frac{\sqrt{6}}{3} \end{bmatrix}$$

$\square$

El método QR para eigenvalores

La descomposición QR puede usarse para aproximar los eigenvalores de una matriz cuadrada.[3] Al algoritmo resultante se le llama **método QR** y constituye una herramienta muy importante para aproximaciones numéricas de los eigenvalores que estudiamos en el capítulo 7. Comparado con los demás, el método QR determina *todos* los eigenvalores de una matriz. También se usa para resolver sistemas lineales. Además, en la ortonormalización se emplean sus variantes para reemplazar el proceso de Gram-Schmidt, que es inestable.

Primero describiremos el método. Se comienza con una matriz invertible A, $n \times n$. Se calcula la factorización QR de $A = QR$, y a continuación se forma la matriz $A_1 = RQ$. Observaremos que A y A_1 son semejantes, porque

$$Q^{-1}AQ = Q^{-1}(QR)Q = RQ = A_1$$

Por tanto, tienen los mismos eigenvalores. Continuamos con la factorización QR, R_1Q_1 de A_1, y se forma la matriz $A_2 = Q_1R_1$ que tiene los mismos eigenvalores que A. Iteramos para obte-

[3] El método QR, en su forma actual, fue introducido por J. G. F. Francis en 1961, y en forma independiente por V. N. Kublanovskaya. Sin embargo, H. Rutishauser (1958) es el autor de la idea, quien usó la factorización LU de una matriz para calcular eigenvalores, y llamó a las iteraciones *transformaciones LR*.

ner una sucesión de matrices

$$A, A_1, A_2, A_3, \ldots$$

Resulta que si A tiene n eigenvalores de distintas magnitudes, entonces esta sucesión tiende a una matriz triangular superior $\hat{R}$ semejante a A. Por consiguiente, los elementos diagonales de $\hat{R}$ son *todos* los eigenvalores de A.

El algoritmo y teorema siguiente, cuya demostración omitiremos, es el núcleo del método QR que acabamos de describir.

Algoritmo

(El método QR)

DATOS: Para una matriz A invertible $n \times n$, cuyos eigenvalores son $\lambda_1, \ldots, \lambda_n$, tal que

$$|\lambda_1| < |\lambda_2| < \cdots < |\lambda_n|$$

1. Igualar $A_0 = A$.
2. Para $i = 1, 2, \ldots, k-1$
   (a) Determinar la descomposición QR de A_i, por ejemplo $A_i = Q_iR_i$.
   (b) Igualar $A_{i+1} = R_iQ_i$.

RESULTADOS: A_k, que se aproxima a una matriz triangular $\hat{R}$ cuyos elementos diagonales son todos los eigenvalores de A.

Advertencias numéricas con el método de Gram-Schmidt y con el método QR

Como hicimos notar en la sección 8.2, al aplicar el proceso de Gram-Schmidt en cálculos numéricos, no se obtiene exactamente una base ortonormal. Esto se debe al error acumulado de redondeo, que puede ser inmenso para este proceso. El resultado es que en la descomposición QR, la matriz R no es exactamente triangular después del cálculo numérico. Los elementos que deberían ser cero con frecuencia son números muy pequeños. En el ejemplo que sigue calculamos con exactitud todas las factorizaciones; a continuación cada valor se aproximó y redondeó a cuatro decimales. Esto sólo fue para demostrar que las matrices R_i son realmente triangulares superiores, y que las Q_i son (casi) ortogonales. En la práctica se pasa directamente a las aproximaciones.

■ EJEMPLO 15 Determine los eigenvalores de

$$A = \begin{bmatrix} 9 & 8 \\ 1 & 2 \end{bmatrix}$$

aplicando el método QR.

SOLUCIÓN Primero obtenemos la factorización QR de $A = A_0$.

$$Q_0 = \begin{bmatrix} \frac{9}{\sqrt{82}} & -\frac{1}{\sqrt{82}} \\ \frac{1}{\sqrt{82}} & \frac{9}{\sqrt{82}} \end{bmatrix} \cong \begin{bmatrix} 0.9938 & -0.1104 \\ 0.1104 & 0.9938 \end{bmatrix}$$

$$R_0 = \begin{bmatrix} \sqrt{82} & \frac{37}{41}\sqrt{82} \\ 0 & \frac{5}{41}\sqrt{82} \end{bmatrix} \cong \begin{bmatrix} 9.0553 & 8.1719 \\ 0 & 1.1043 \end{bmatrix}$$

En consecuencia,

$$A_1 = R_0Q_0 = \begin{bmatrix} 9.9024 & 7.1219 \\ 0.1219 & 1.0975 \end{bmatrix}$$

A continuación repetimos factorizando A_1,

$$Q_1 = \begin{bmatrix} 0.9999 & -0.0123 \\ 0.0123 & 0.9999 \end{bmatrix}, \qquad R_1 = \begin{bmatrix} 9.9031 & 7.1349 \\ 0 & 1.0097 \end{bmatrix}$$

$$A_2 = R_1Q_1 = \begin{bmatrix} 9.9903 & 7.0124 \\ 0.0124 & 1.0096 \end{bmatrix}$$

y factorizando A_2,

$$Q_2 = \begin{bmatrix} 0.9999 & -0.0012 \\ 0.0012 & 0.9999 \end{bmatrix} \qquad R_2 = \begin{bmatrix} 9.9903 & 7.0136 \\ 0 & 1.0009 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 9.9990 & 7.0012 \\ 0.0012 & 1.0009 \end{bmatrix}$$

Observamos que Q_2 es casi ortogonal y R_2 es triangular superior cuyos elementos en la diagonal son casi cercanos a 10 y a 1, que son los verdaderos eigenvalores de A. $\square$

Factorización QR con sistemas algebraicos computacionales

Maple

```
> with(linalg):
> R:=QRdecomp(matrix([[1,-2],[-1,3]]),Q='q');
```

$$R = \begin{bmatrix} \sqrt{2} & -\frac{5}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} \end{bmatrix}$$

```
> evalm(q);
```

$$\begin{bmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix}$$

Mathematica

```
In[1]:=
   QRDecomposition[{{1.,-2},{-1,3}}]
Out[1]=
{{{-0.707107, 0.707107},
   {-0.707107, -0.707107}},
  {{-1.41421, 3.53553}, {0, -0.707107}}}
```

MATLAB

```
>> [Q,R]=qr([1 -2; -1 3])
 Q =
   -0.7071    0.7071
    0.7071    0.7071
 R =
   -1.4142    3.5355
         0    0.7071
```

Ejercicios 8.3

En los ejercicios 1 a 7 determine una factorización QR de A.

1. $A = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$

3. $A = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & -2 \\ 1 & 4 \\ 1 & 0 \\ -1 & 2 \end{bmatrix}$

7. $A = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 2 \\ -1 & 4 \end{bmatrix}$

En los ejercicios 8 a 10 calcule la matriz triangular R tal que $A = QR$, sabiendo que Q se obtuvo ortonormalizando las columnas de A.

8. $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ -1 & 0 \end{bmatrix}, Q = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

9. $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & -1 \\ -1 & -2 \end{bmatrix}, Q = \begin{bmatrix} \frac{1}{2} & -\frac{1}{\sqrt{6}} \\ \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{2} & -\frac{\sqrt{6}}{3} \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & 1 \end{bmatrix}, Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$

11. Obtenga una factorización QR para una matriz ortogonal A.

12. Demuestre que A es invertible si y sólo si $A = QR$ para alguna matriz ortogonal Q y alguna matriz triangular superior R cuyos elementos en la diagonal principal sean distintos de cero.

13. Si A tiene columnas linealmente independientes, y $A = QR$ es una factorización QR, demuestre que

$$A\mathbf{x} = \mathbf{b} \text{ es consistente } \Leftrightarrow$$
$$Q\mathbf{y} = \mathbf{b} \text{ es consistente}$$

Llegue a la conclusión que A y Q tienen el mismo espacio de columnas.

14. Si A es una matriz con columnas distintas de cero, que forman un conjunto ortogonal. Si $A = QR$ es una factorización QR, demuestre que R es diagonal.

En los ejercicios 15 a 17 determine A_1, A_2 y A_3 del método QR. Use A_3 para estimar los eigenvalores de A. ¿Cuál es el error en cada caso?

15. $A = \begin{bmatrix} 8 & 7 \\ 1 & 2 \end{bmatrix}$

16. $A = \begin{bmatrix} 10 & 8 \\ 1 & 3 \end{bmatrix}$

17. $A = \begin{bmatrix} 12 & 11 \\ 2 & 3 \end{bmatrix}$

8.4 Mínimos cuadrados

Objetivo del estudiante para esta sección

Aprenda a resolver problemas básicos de mínimos cuadrados.

En esta sección estudiaremos un tema de mucho interés en las aplicaciones: el método de los mínimos cuadrados, que citamos en la introducción.

Un problema de mínimos cuadrados

Al tratar de resolver un problema es frecuente obtener puntos experimentales y buscar una función cuya gráfica pase por esos puntos. Por lo general, la naturaleza del problema determina el tipo de función que se necesita. Por ejemplo, si un automóvil se mueve con rapidez constante y medimos la distancia $s(t)$ recorrida cada minuto, cabe esperar que la gráfica de $s(t)$ en función de t sea una recta. Un polinomio de grado superior, o una función exponencial, serían inadecuados en este caso.

Supongamos que nuestro problema sugiere una recta, y que tenemos los puntos $(1, 2)$, $(2, 4)$ y $(3, 3)$. En este caso, $y = b + mx$ es la ecuación de esa recta. Queremos determinar la pendiente m y la ordenada al origen b. Como la recta debe pasar por los tres puntos tenemos,

$$2 = b + m \cdot 1, \qquad 4 = b + m \cdot 2, \qquad 3 = b + m \cdot 3$$

Desafortunadamente, el sistema lineal resultante en las incógnitas m y b,

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$$

resulta ser *inconsistente*, lo cual se aprecia con facilidad, de manera que nuestro problema no se puede resolver con exactitud.[4] Lo menos que podemos hacer es tratar de determinar la recta que se "ajuste" mejor a esos puntos.

El mejor ajuste es algo que puede tener distintos significados, dependiendo de cuáles aspectos de la solución se necesita resaltar. En este caso, supongamos que lo deseable es que la mejor recta sea tal que si los errores en la dirección y son δ_1, δ_2 y δ_3,

$$\delta_1 = 2 - b - m \cdot 1, \qquad \delta_2 = 4 - b - m \cdot 2, \qquad \delta_3 = 3 - b - m \cdot 3$$

entonces el número

$$\delta_1^2 + \delta_2^2 + \delta_3^2$$

es mínimo (figura 8.10). Una solución para m y b que minimice esta suma de los cuadrados de los errores se llama **solución de mínimos cuadrados**.

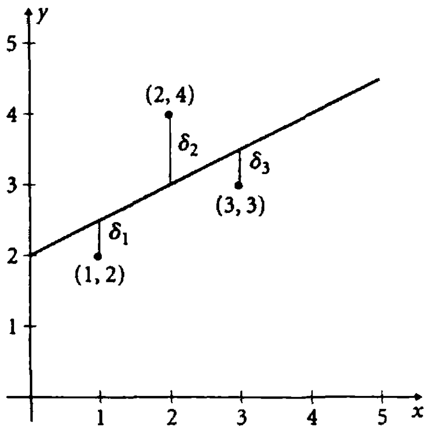


Figura 8.10 Mínimos cuadrados: minimizar a $\delta_1^2 + \delta_2^2 + \delta_3^2$.

Podemos expresar todo lo anterior en notación vectorial. Si Δ es el vector error

$$\Delta = (\delta_1, \delta_2, \delta_3)$$

se desea minimizar $\delta_1^2 + \delta_2^2 + \delta_3^2 = \|\Delta\|^2$ o, lo que es igual, minimizar $\|\Delta\|$.

■ EJEMPLO 16 Determine cuál de las rectas produce el menor error de mínimos cuadrados para los puntos (1, 2), (2, 4) y (3, 3).

(a) $y = 2x$
(b) $y = 3$
(c) $y = 0.5x + 2$

[4] Observe que la función cuadrática $-\frac{3}{2}x^2 + \frac{13}{2}x - 3$ pasa por esos puntos, pero no es la que necesitamos.

SOLUCIÓN Contamos con los datos siguientes.

	$y = 2x$ $(m = 2,\ b = 0)$	$y = 3$ $(m = 0,\ b = 3)$	$y = 0.5x + 2$ $(m = 0.5,\ b = 2)$
δ_1	$2 - 0 - 1 \cdot 2 = 0$	$2 - 3 - 1 \cdot 0 = -1$	$2 - 2 - 1 \cdot 0.5 = -0.5$
δ_2	$4 - 0 - 2 \cdot 2 = 0$	$4 - 3 - 2 \cdot 0 = 1$	$4 - 2 - 2 \cdot 0.5 = 1$
δ_2	$3 - 0 - 3 \cdot 2 = -3$	$3 - 3 - 3 \cdot 0 = 0$	$3 - 2 - 3 \cdot 0.5 = -0.5$
$\|\Delta\|^2$	$0^2 + 0^2 + (-3)^2 = 9$	$(-1)^2 + 1^2 + 0^2 = 2$	$(-0.5)^2 + (-0.5)^2 + 1 = 1.5$

Notamos que la recta $y = 0.5x + 2$ produce el menor error de mínimos cuadrados con cualquiera de las tres rectas. De hecho, como veremos, esta recta produce el menor error en comparación con el de *cualquier* otra recta.

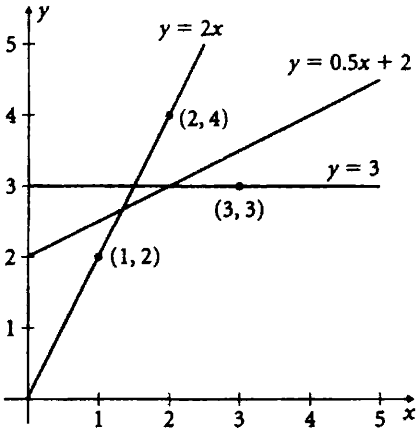


Figura 8.11

Solución del problema de mínimos cuadrados

Veamos ahora cómo determinar la solución, por mínimos cuadrados, para los puntos de la figura 8.11, y en general. Supongamos que se tiene un sistema lineal inconsistente

$$A\mathbf{x} = \mathbf{b} \tag{8.21}$$

en el que A es una matriz $m \times n$. Como para cualquier vector-n $\mathbf{x}$, el producto $A\mathbf{x}$ nunca es igual a $\mathbf{b}$, el error resultante Δ,

$$\Delta = \mathbf{b} - A\mathbf{x}$$

es un vector-m no cero para todos los vectores-n $\mathbf{x}$. Resolver el problema de mínimos cuadrados para la ecuación (8.21) equivale a determinar un vector-n $\tilde{\mathbf{x}}$, tal que la longitud de $\Delta = \mathbf{b} - A\tilde{\mathbf{x}}$ sea mínima. Entonces $\tilde{\mathbf{x}}$ sería nuestra *solución por* (o *de*) *mínimos cuadrados*.

Problema de mínimos cuadrados: Determine $\tilde{\mathbf{x}}$ tal que $\|\mathbf{b} - A\tilde{\mathbf{x}}\|$ sea mínima.

A medida que varía $\mathbf{x}$, $A\mathbf{x}$ genera Col(A). Así, $\|\mathbf{b} - A\tilde{\mathbf{x}}\|$ es mínima sólo si $A\tilde{\mathbf{x}}$ es la proyección ortogonal $\mathbf{b}_{pr}$ de $\mathbf{b}$ sobre la Col(A), de acuerdo con el teorema de la mejor aproximación (sección 8.2). Por tanto,

$$\|\Delta\| = \text{mín} \Leftrightarrow A\mathbf{x} = \mathbf{b}_{pr} \Leftrightarrow \mathbf{b} - A\tilde{\mathbf{x}} = \mathbf{b}_c$$

Llegamos a la conclusión de que **siempre existe** una solución $\tilde{\mathbf{x}}$ de mínimos cuadrados para $A\mathbf{x} = \mathbf{b}$, porque $\mathbf{b}_{pr}$ siempre existe.

TEOREMA 20

> Para cualquier matriz A $m \times n$ y cualquier vector-m $\mathbf{b}$, hay una solución $\tilde{\mathbf{x}}$ de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. Además, si $\mathbf{b}_{pr}$ es la proyección ortogonal de $\mathbf{b}$ sobre la Col(A), entonces
>
> $$A\tilde{\mathbf{x}} = \mathbf{b}_{pr} \qquad (8.22)$$

Como $\mathbf{b}_c = \mathbf{b} - A\tilde{\mathbf{x}}$ es ortogonal a Col(A), entonces $A\mathbf{x}$ y $\mathbf{b} - A\tilde{\mathbf{x}}$ debe ser ortogonal para cualquier vector-n $\mathbf{x}$. De manera que

$$\begin{aligned}
&(\mathbf{b} - A\tilde{\mathbf{x}}) \cdot A\mathbf{x} &&= 0 \\
\Leftrightarrow\ &A^T(\mathbf{b} - A\tilde{\mathbf{x}}) \cdot \mathbf{x} &&= 0 \\
\Leftrightarrow\ &A^T(\mathbf{b} - A\tilde{\mathbf{x}}) &&= \mathbf{0} \qquad \text{Según el teorema 3 de la sección 8.1.} \\
\Leftrightarrow\ &A^T\mathbf{b} - A^TA\tilde{\mathbf{x}} &&= \mathbf{0} \\
\Leftrightarrow\ &A^TA\tilde{\mathbf{x}} &&= A^T\mathbf{b}
\end{aligned}$$

Así, $\tilde{\mathbf{x}}$ es una solución de mínimos cuadrados si y sólo si satisface al sistema $A^TA\tilde{\mathbf{x}} = A^T\mathbf{b}$. A este sistema se le llama **ecuaciones normales para** $\tilde{\mathbf{x}}$. Ahora ya sabemos cómo determinar $\tilde{\mathbf{x}}$. Además, podemos usar $\|\Delta\| = \|\mathbf{b} - A\tilde{\mathbf{x}}\|$ para calcular el **error de mínimos cuadrados** incurrido.

TEOREMA 21

> (Soluciones por cuadrados mínimos)
>
> Si A es una matriz $m \times n$, siempre hay $\tilde{\mathbf{x}}$ soluciones por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. Además,
>
> 1. $\tilde{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ si y sólo si $\tilde{\mathbf{x}}$ es una solución de las ecuaciones normales
>
> $$A^TA\tilde{\mathbf{x}} = A^T\mathbf{b} \qquad (8.23)$$
>
>   El error de mínimos cuadrados, $\|\Delta\|$ se define por
>
> $$\|\Delta\| = \|\mathbf{b} - A\tilde{\mathbf{x}}\|$$
>
> 2. A tendrá columnas linealmente independientes si y sólo si A^TA invertible. En este caso, la solución por mínimos cuadrados es única y puede calcularse con
>
> $$\tilde{\mathbf{x}} = \left(A^TA\right)^{-1}A^T\mathbf{b}$$

DEMOSTRACIÓN Sólo precisamos demostrar el último enunciado del teorema. Primero comprobaremos que A y A^TA tienen el mismo espacio nulo. En realidad, si $\mathbf{v} \in \nu(A)$, entonces

$A\mathbf{v} = \mathbf{0}$, de modo que $A^TA\mathbf{v} = \mathbf{0}$, y esto implica que $\mathbf{v} \in \nu(A^TA)$. Por consiguiente, $\nu(A) \subseteq \nu(A^TA)$. Por otro lado,

$$\begin{aligned}
&\mathbf{v} \in \nu(A^TA) \\
\Rightarrow\ & A^TA\mathbf{v} &&= \mathbf{0} \\
\Rightarrow\ & A^TA\mathbf{v} \cdot \mathbf{v} &&= \mathbf{0} \cdot \mathbf{v} = 0 \\
\Rightarrow\ & A\mathbf{v} \cdot A\mathbf{v} &&= 0 \\
\Rightarrow\ & \|A\mathbf{v}\|^2 &&= 0 \\
\Rightarrow\ & A\mathbf{v} &&= \mathbf{0} \\
\Rightarrow\ & \mathbf{v} \in \nu(A)
\end{aligned}$$

De ahí que $\nu(A^TA) \subseteq \nu(A)$, y los dos espacios nulos son iguales. Entonces, Nulidad(A) = Nulidad(A^TA). Pero de acuerdo con el teorema de la dimensión,

$$\text{Rango}(A) + \text{Nulidad}(A) = n = \text{Rango}(A^TA) + \text{Nulidad}(A^TA)$$

Así, Rango(A) = Rango(A^TA). Si A tiene columnas linealmente independientes, entonces Rango$(A) = n$. Por lo anterior, Rango$(A^TA) = n$. Por consiguiente, la matriz A^TA de $n \times n$ tiene columnas linealmente independientes. Luego, A^TA es invertible. Por el contrario, si A^TA es invertible, entonces Rango$(A^TA) = n$; por consiguiente, Rango$(A) = n$. De modo que A tiene columnas linealmente independientes. Si A^TA es invertible, $A^TA\tilde{\mathbf{x}} = A^T\mathbf{b}$ implica que la solución única es $\tilde{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$. $\square$

■ EJEMPLO 17 Resuelva el siguiente problema de mínimos cuadrados y calcule el error de mínimos cuadrados para el sistema $A\mathbf{x} = \mathbf{b}$,

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$$

Use la solución para determinar la recta que produzca el menor error de mínimos cuadrados para los puntos (1, 2), (2, 4) y (3, 3) mencionados en la introducción de esta sección.

SOLUCIÓN De acuerdo con el teorema 21, basta resolver las ecuaciones normales $A^TA\tilde{\mathbf{x}} = A^T\mathbf{b}$:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$$

es decir

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} 9 \\ 19 \end{bmatrix}$$

La solución de este sistema es la solución por mínimos cuadrados

$$\tilde{\mathbf{x}} = \begin{bmatrix} 2 \\ \frac{1}{2} \end{bmatrix}$$

cuyo error de mínimos cuadrados es (figura 8.12)

$$\begin{aligned}\|\Delta\| &= \|\mathbf{b} - A\tilde{\mathbf{x}}\| \\ &= \left\| \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ \frac{1}{2} \end{bmatrix} \right\| \\ &= \left\| \begin{bmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix} \right\| = \frac{\sqrt{6}}{2}\end{aligned}$$

Como $\tilde{\mathbf{x}} = (2, 0.5)$, la pendiente de la recta de mínimos cuadrados es 0.5, y su ordenada al origen es 2, de modo que su ecuación es $y = 0.5x + 2$. (Es exactamente la línea que se aprecia en la figura 8.10.) □

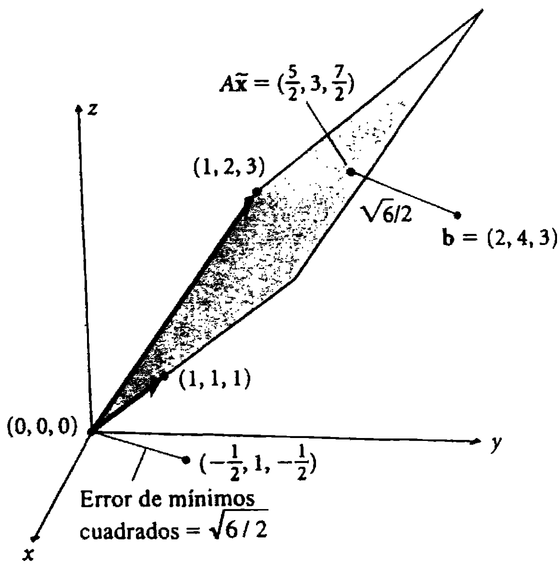


Figura 8.12 Error de mínimos cuadrados.

■ EJEMPLO 18 Determine la recta de mínimos cuadrados para las calificaciones que ha reunido el profesor de álgebra lineal, mencionadas al principio de este capítulo. Además, calcule el porcentaje esperado de B después del décimo semestre.

Semestre	1	2	3	4	5	6
Porcentaje de calificaciones B	0.20	0.25	0.20	0.35	0.45	0.40

SOLUCIÓN Sean

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 0.20 \\ 0.25 \\ 0.20 \\ 0.35 \\ 0.45 \\ 0.40 \end{bmatrix}$$

Por tanto,

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 21 \\ 21 & 91 \end{bmatrix}$$

y

$$A^T \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0.20 \\ 0.25 \\ 0.20 \\ 0.35 \\ 0.45 \\ 0.40 \end{bmatrix} = \begin{bmatrix} 1.85 \\ 7.35 \end{bmatrix}$$

Así, las ecuaciones normales son

$$\begin{bmatrix} 6 & 21 \\ 21 & 91 \end{bmatrix} \begin{bmatrix} \tilde{b} \\ \tilde{m} \end{bmatrix} = \begin{bmatrix} 1.85 \\ 7.35 \end{bmatrix}$$

Pero

$$\begin{bmatrix} 6 & 21 & 1.85 \\ 21 & 91 & 7.35 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0.13333 \\ 0 & 1 & 0.05 \end{bmatrix}$$

Por consiguiente, $\tilde{m} = 0.05$ y $\tilde{b} = 0.13333$. De manera que la recta es

$$y = 0.13333 + 0.05x$$

Para $x = 10$ se obtiene $y = 0.13333 + 0.05 \cdot 10 = 0.63333$. Esto significa que más o menos esperaríamos 63.3% de calificaciones B al terminar el décimo semestre, si continúa esta tendencia de calificaciones. $\square$

Si A no tiene columnas linealmente independientes, habrá varias soluciones de mínimos cuadrados.

■ EJEMPLO 19 Deduzca todas las soluciones por mínimos cuadrados del sistema

$$x - y = 1$$

$$x - y = 5$$

SOLUCIÓN $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$. Por consiguiente,

$$A^T A\widetilde{\mathbf{x}} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix} = A^T\mathbf{b} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

La reducción de la matriz aumentada

$$\begin{bmatrix} 2 & -2 & 6 \\ -2 & 2 & -6 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

da como resultado $\widetilde{x} = 3 + r,\ \widetilde{y} = r,\ r \in \mathbf{R}$. □

Mínimos cuadrados con factorización QR

Las soluciones de mínimos cuadrados que hemos descrito están sujetas a un problema frecuente. La matriz A^TA de las ecuaciones normales está **mal condicionada**; por lo general, esto significa que un error numérico pequeño en una reducción por operaciones de renglón causa un error grande en la solución.

Por lo común, la eliminación de Gauss de A^TA no produce buenas soluciones aproximadas cuando el tamaño $n \geq 5$. Como respuesta a este grave problema puede usarse la factorización QR de A. La idea fundamental de este método es que debido a que las matrices ortogonales preservan las longitudes, también deben hacerlo con la longitud del vector error.

Si A tiene columnas linealmente independientes, y $A = QR$ es una factorización QR como se describió en la sección 8.3, entonces para que $\widetilde{\mathbf{x}}$ sea una solución de $A\mathbf{x} = \mathbf{b}$ por mínimos cuadrados,

$$\begin{aligned} & A^TA\widetilde{\mathbf{x}} = A^T\mathbf{b} \\ \Leftrightarrow\ & (QR)^T(QR)\widetilde{\mathbf{x}} = (QR)^T\mathbf{b} \\ \Leftrightarrow\ & R^TQ^TQR\widetilde{\mathbf{x}} = R^TQ^T\mathbf{b} \\ \Leftrightarrow\ & R^TR\widetilde{\mathbf{x}} = R^TQ^T\mathbf{b} && \text{por ser } Q^TQ = I \\ \Leftrightarrow\ & R\widetilde{\mathbf{x}} = Q^T\mathbf{b} && \text{por ser invertible } R^T \end{aligned}$$

Observe que $R\widetilde{\mathbf{x}} = Q^T\mathbf{b}$ equivale a $\widetilde{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$. Sin embargo, en lugar de invertir R, es más fácil usar la sustitución en reversa para el sistema $R\widetilde{\mathbf{x}} = Q^T\mathbf{b}$. Hemos demostrado el siguiente teorema:

TEOREMA 22

> Si A es una matriz $m \times n$ con columnas linealmente independientes, y si $A = QR$ es una factorización QR, en ese caso la única solución $\widetilde{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ por mínimos cuadrados se expresa teóricamente con
>
> $$\widetilde{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$$
>
> y suele calcularse resolviendo el sistema
>
> $$R\widetilde{\mathbf{x}} = Q^T\mathbf{b}$$

■ EJEMPLO 20 Determine la solución de $A\mathbf{x} = \mathbf{b}$, por mínimos cuadrados, usando la factorización QR.

$$A = \begin{bmatrix} 2 & 2 & 6 \\ 1 & 4 & -3 \\ 2 & -4 & 9 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

SOLUCIÓN Al aplicar el método de la sección 8.3 tenemos

$$A = QR = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & 0 & 9 \\ 0 & 6 & -6 \\ 0 & 0 & -3 \end{bmatrix}$$

Pero

$$Q^T\mathbf{b} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}$$

Por consiguiente, de acuerdo con el teorema 22, $\tilde{\mathbf{x}}$ puede calcularse resolviendo $R\tilde{\mathbf{x}} = Q^T\mathbf{b}$:

$$\begin{bmatrix} 3 & 0 & 9 \\ 0 & 6 & -6 \\ 0 & 0 & -3 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}$$

Sustituyendo en reversa obtenemos con facilidad $\tilde{\mathbf{x}} = \begin{bmatrix} 1 \\ -\frac{1}{2} \\ 0 \end{bmatrix}$ ▭

Ejercicios 8.4

En los ejercicios 1 a 3 use las ecuaciones normales para determinar la solución de mínimos cuadrados para $A\mathbf{x} = \mathbf{b}$.

1. $A = \begin{bmatrix} 1 & -2 \\ -1 & 1 \\ 1 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$

2. $A = \begin{bmatrix} -1 & 1 \\ 2 & 2 \\ -1 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ -4 \\ 1 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & 1 \\ -1 & 1 \\ 2 & 2 \\ -1 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$

4. Resuelva el sistema $\begin{bmatrix} 1 & -2 \\ 1 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$. También determine su solución por mínimos cuadrados. ¿Por qué obtuvo la misma respuesta?

5. Escriba su propio sistema 3×3 cuya solución coincida con su solución por mínimos cuadrados.

6. Obtenga su propio sistema inconsistente 3×2 y calcule su solución por mínimos cuadrados.

En los ejercicios 7 a 9 determine la línea de mínimos cuadrados para los puntos dados. En cada caso trace los puntos y la recta.

7. $(-1, -2), (0, 1), (1, 5)$
8. $(1, 2), (2, 3), (4, 4)$
9. $(-1, -1), (1, 2), (3, 2), (4, 4)$
10. Determine todas las soluciones del sistema por mínimos cuadrados.

$$x - y = 0$$
$$x - y = 1$$
$$x - y = 2$$

11. Deduzca todas las soluciones para $A\mathbf{x} = \mathbf{b}$ por mínimos cuadrados, siendo

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -2 & 0 & 2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

En los ejercicios 12 a 14 use la factorización QR mencionada y obtenga la solución para $A\mathbf{x} = \mathbf{b}$, por mínimos cuadrados.

12. $A = \begin{bmatrix} -2 & 0 \\ 1 & -1 \\ 2 & -1 \end{bmatrix}, Q = \begin{bmatrix} -\frac{2}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix},$

$R = \begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix}$, y $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$

13. $A = \begin{bmatrix} 0 & 3 \\ 0 & 4 \\ 5 & 10 \end{bmatrix}, Q = \begin{bmatrix} 0 & \frac{3}{5} \\ 0 & \frac{4}{5} \\ 1 & 0 \end{bmatrix},$

$R = \begin{bmatrix} 5 & 10 \\ 0 & 5 \end{bmatrix}$, y $\mathbf{b} = \begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}, Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix},$

$R = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, y $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix}$

Mínimos cuadrados cuando A tiene columnas ortogonales

Si la matriz A tiene columnas ortogonales distintas de cero, es fácil determinar la solución por mínimos cuadrados para $A\mathbf{x} = \mathbf{b}$.

15. Si A es una matriz $m \times n$ con columnas ortogonales $\mathbf{a}_i$ **distintas de cero** y si $\tilde{\mathbf{x}}$ es la solución por mínimos cuadrados del sistema $A\mathbf{x} = \mathbf{b}$, demuestre que

$$\tilde{\mathbf{x}} = \begin{bmatrix} \dfrac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1} \\ \dfrac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2} \\ \vdots \\ \dfrac{\mathbf{b}\cdot\mathbf{a}_n}{\mathbf{a}_n\cdot\mathbf{a}_n} \end{bmatrix} \tag{8.24}$$

(*Sugerencia:* A^TA tiene una forma muy especial bajo las hipótesis para A.)

16. ¿Cómo se simplifica la ecuación (8.24) cuando A tiene columnas ortonormales?

17. Aplique el ejercicio 15 para calcular $\tilde{\mathbf{x}}$ si

$$A = \begin{bmatrix} -2 & -2 \\ 1 & -2 \\ 2 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix}$$

18. Use el ejercicio 15 para determinar $\tilde{\mathbf{x}}$ si

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix}$$

19. Con el ejercicio 15 determine $\tilde{\mathbf{x}}$ si

$$A = \begin{bmatrix} -\frac{2}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix}$$

8.5 Ortogonalización de matrices simétricas

Objetivo del estudiante para esta sección

Aprender a diagonalizar ortogonalmente una matriz simétrica.

En esta sección estudiaremos algunas propiedades muy interesantes de las matrices simétricas. Recuerde que A es **simétrica** si cualquier elemento fuera de la diagonal principal tiene una imagen especular respecto a la diagonal. En otras palabras,

$$A^T = A$$

En particular, A es cuadrada.

Por ejemplo, $\begin{bmatrix} 1 & 3 \\ 3 & -2 \end{bmatrix}$ y $\begin{bmatrix} -1 & -1 & 1 \\ -1 & 2 & 4 \\ 1 & 4 & 2 \end{bmatrix}$ son simétricas, pero $\begin{bmatrix} 1 & -3 \\ 3 & -2 \end{bmatrix}$ no lo es.

En la sección 7.2 vimos que si una matriz A $n \times n$ es diagonalizable (es decir, semejante a una matriz diagonal), podemos definir una matriz P invertible cuyas columnas sean n eigenvectores linealmente independientes de A, y una matriz diagonal D cuyos elementos diagonales sean los eigenvalores correspondientes, tales que

$$P^{-1}AP = D$$

En este caso nos interesan aquellas matrices diagonalizables para las que P pueda ser *ortogonal*.

DEFINICIÓN

Se dice que A es **ortogonalmente diagonalizable** si se puede diagonalizar por una matriz Q invertible y una matriz diagonal D, de manera que Q sea ortogonal. Así, $Q^{-1}AQ = D$, o lo que es igual,

$$Q^TAQ = D$$

porque $Q^{-1} = Q^T$.

En general se dice que dos matrices A y B, de tamaño $n \times n$, son **ortogonalmente semejantes** si hay una matriz Q ortogonal tal que $Q^{-1}AQ = B$, es decir

$$Q^TAQ = B$$

Así, una matriz ortogonalmente diagonalizable es ortogonalmente semejante a una matriz diagonal.

TEOREMA 23

Una matriz ortogonalmente diagonalizable es simétrica.

DEMOSTRACIÓN Si A es ortogonalmente diagonalizable, entonces $Q^{-1}AQ = D$ para una matriz ortogonal Q. De modo que

$$A = QDQ^{-1} = QDQ^T$$

que implica que

$$A^T = (QDQ^T)^T = Q^{TT}D^TQ^T = QDQ^T = A$$

porque $D^T = D$ (por ser diagonal D). Por consiguiente, A es simétrica. $\square$

Lo sorprendente (y asombroso) es que el inverso de este teorema es cierto. Es decir, *cualquier matriz simétrica es ortogonalmente diagonalizable*. No es fácil demostrar esta afirmación.

Una propiedad interesante de las matrices simétricas es la siguiente.

TEOREMA 24

Una matriz simétrica real sólo tiene eigenvalores reales.

DEMOSTRACIÓN (OPCIONAL) Para la demostración se requieren algunas propiedades sencillas de los conjugados de los números complejos (que se describen en el apéndice A).

Preparación: El **(complejo) conjugado** de $a + bi$ es el número complejo $a - bi$ y se representa por $\overline{a+bi}$. Por ejemplo,

$$\overline{1+i} = 1 - i, \qquad \overline{-i} = i, \qquad \overline{1} = 1$$

Como la conjugación sólo cambia el signo de la parte imaginaria, vemos que un número es real si y sólo si es igual a su conjugado.

$$z \text{ es real} \iff \bar{z} = z \tag{8.25}$$

La conjugación compleja satisface las siguientes propiedades básicas:

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \qquad \overline{z_1 - z_2} = \overline{z_1} - \overline{z_2}, \qquad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2} \tag{8.26}$$

La matriz compleja conjugada $\overline{A}$ de A es aquella cuyos elementos son los complejos conjugados de los elementos de A. Es fácil aplicar las propiedades anteriores para demostrar que para matrices A y B compatibles,

$$\overline{A+B} = \overline{A} + \overline{B}, \qquad \overline{A-B} = \overline{A} - \overline{B}, \qquad \overline{AB} = \overline{A}\,\overline{B} \tag{8.27}$$

Demostración principal: Si A es una matriz simétrica, y si $\mathbf{v}$ es un eigenvector con eigenvalor correspondiente λ, entonces $A\mathbf{v} = \lambda\mathbf{v}$. También, $\overline{A} = A$, porque A tiene elementos reales, según las ecuaciones (8.26). Por consiguiente, de acuerdo con las ecuaciones (8.27),

$$\begin{aligned}
\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} &\Rightarrow \overline{A}\,\overline{\mathbf{v}} = \overline{\lambda}\,\overline{\mathbf{v}} \\
&\Rightarrow A\overline{\mathbf{v}} = \overline{\lambda}\,\overline{\mathbf{v}} \\
&\Rightarrow (A\overline{\mathbf{v}})^T = (\overline{\lambda}\,\overline{\mathbf{v}})^T \\
&\Rightarrow \overline{\mathbf{v}}^T A^T = \overline{\lambda}\,\overline{\mathbf{v}}^T \\
&\Rightarrow \overline{\mathbf{v}}^T A = \overline{\lambda}\,\overline{\mathbf{v}}^T \\
&\Rightarrow (\overline{\mathbf{v}}^T A)\mathbf{v} = (\overline{\lambda}\,\overline{\mathbf{v}}^T)\mathbf{v} \\
&\Rightarrow \overline{\mathbf{v}}^T (A\mathbf{v}) = \overline{\lambda}(\overline{\mathbf{v}}^T\mathbf{v}) \\
&\Rightarrow \overline{\mathbf{v}}^T (\lambda\mathbf{v}) = \overline{\lambda}(\overline{\mathbf{v}}^T\mathbf{v}) \\
&\Rightarrow (\lambda - \overline{\lambda})(\overline{\mathbf{v}}^T\mathbf{v}) = 0
\end{aligned}$$

Pero $\overline{\mathbf{v}}^T\mathbf{v} \neq 0$, porque si $\mathbf{v} = (a_1 + ib_1, \ldots, a_n + ib_n)$, entonces $\overline{\mathbf{v}} = (a_1 - ib_1, \ldots, a_n - ib_n)$, y así

$$\overline{\mathbf{v}}^T\mathbf{v} = (a_1^2 + b_1^2) + \cdots + (a_n^2 + b_n^2) \neq 0$$

porque $\mathbf{v} \neq 0$. En consecuencia, $\lambda - \overline{\lambda} = 0$ o $\lambda = \overline{\lambda}$. Así, el eigenvalor λ es real, de acuerdo con las ecuaciones (8.26). ☐

TEOREMA 25

Dos eigenvectores cualesquiera de una matriz simétrica A que correspondan a dos eigenvalores distintos son ortogonales.

DEMOSTRACIÓN Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ dos eigenvectores de A cuyos eigenvalores correspondientes sean λ_1 y λ_2, tales que $\lambda_1 \neq \lambda_2$. Como $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$, $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ y $A^T = A$, entonces

$$\begin{aligned}\lambda_1\mathbf{v}_1 \cdot \mathbf{v}_2 &= (A\mathbf{v}_1) \cdot \mathbf{v}_2 \\ &= \mathbf{v}_1 \cdot (A^T\mathbf{v}_2) \\ &= \mathbf{v}_1 \cdot A\mathbf{v}_2 \\ &= \mathbf{v}_1 \cdot \lambda_2\mathbf{v}_2 \\ &= \lambda_2\mathbf{v}_1 \cdot \mathbf{v}_2\end{aligned}$$

Por tanto,

$$(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$$

Pero $\lambda_1 - \lambda_2 \neq 0$. Así, $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, como se afirmaba. $\square$

■ EJEMPLO 21 Compruebe los teoremas 24 y 25 para A y B.

$$A = \begin{bmatrix} -1 & -1 & 1 \\ -1 & 2 & 4 \\ 1 & 4 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 3 & 3 \\ 3 & 0 & 3 \\ 3 & 3 & 0 \end{bmatrix}$$

SOLUCIÓN

1. Es fácil ver que los eigenvalores de A son -3, 0, 6, y que los eigenvectores correspondientes son

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Observe que $\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_3 = \mathbf{v}_2 \cdot \mathbf{v}_3 = 0$. En consecuencia todos los eigenvalores son reales y los eigenvectores son ortogonales, que es lo afirmado por los teoremas 24 y 25.

2. B tiene el polinomio característico $(x-6)(x+3)^2$ y, por tanto, los eigenvalores reales -3 y 6. También

$$E_{-3} = \text{Gen}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}, \qquad E_6 = \text{Gen}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$$

Para los eigenvectores $(-1, 1, 0) \cdot (1, 1, 1) = 0$, y $(-1, 0, 1) \cdot (1, 1, 1) = 0$. Sin embargo, note que $(-1, 1, 0) \cdot (-1, 0, 1) = 1 \neq 0$. $\square$

OBSERVACIÓN Aunque los eigenvectores que corresponden a *distintos* eigenvalores son ortogonales, los que pertenecen al *mismo* eigenvalor *no tienen que ser ortogonales*, como acabamos de ver.

■ EJEMPLO 22 Diagonalice ortogonalmente las matrices A y B del ejemplo 21.

SOLUCIÓN

1. En el ejemplo 21 determinamos los eigenvectores básicos de A que ya eran ortogonales. Si los normalizamos seguirán siendo ortogonales. Así

$$Q = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

es ortogonal y sus columnas son eigenvectores de A. Entonces Q diagonaliza ortogonalmente a A. No es difícil comprobar que

$$Q^T A Q = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

2. Para B, como los eigenvectores $(-1, 1, 0)$ y $(-1, 0, 1)$ no eran ortogonales, podemos aplicar el proceso de Gram-Schmidt para ortogonalizarlos. Con facilidad obtenemos $(-1, 1, 0)$ y $(-\frac{1}{2}, -\frac{1}{2}, 1)$. Es muy importante notar que como el proceso de Gram-Schmidt no altera el generador de los vectores originales, el nuevo vector $(-\frac{1}{2}, -\frac{1}{2}, 1)$ sigue estando en E_{-3}. (Compruébelo.) En consecuencia, es un eigenvector de A que corresponde a -3, y debe ser ortogonal a $(1, 1, 1)$. Como $(-1, 1, 0)$, $(-\frac{1}{2}, -\frac{1}{2}, 1)$ y $(1, 1, 1)$ ya son mutuamente ortogonales, todo lo que se necesita es normalizarlos para obtener una matriz ortogonal Q que diagonalice a B:

$$Q = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

De nuevo, no es difícil comprobar que

$$Q^T B Q = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

□

NOTA En el ejemplo 22 vimos cómo diagonalizar ortogonalmente una matriz simétrica en la práctica. Si es necesario, se aplica el proceso de Gram-Schmidt para ortogonalizar los eigenvectores que corresponden al mismo eigenvalor. A continuación se normalizan los eigenvectores ortogonales y se usan como columnas de Q. Lo que todavía no hacemos es verificar que cualquier matriz simétrica es diagonalizable. Para comprobar este hecho básico aplicaremos un teorema clásico debido a Schur (1909). Su demostración es bastante difícil, y se encuentra al final de la sección, dedicada al lector interesado.

TEOREMA 26

(Descomposición de Schur)

Toda matriz A cuadrada real que sólo contiene eigenvalores reales es ortogonalmente semejante a una matriz triangular superior T. Por consiguiente, existe una matriz ortogonal Q y una matriz triangular superior T tales que $Q^TAQ = T$. En forma equivalente,

$$A = QTQ^T$$

Finalmente hemos llegado a un resultado principal relacionado con matrices simétricas: el teorema espectral.

TEOREMA 27

(El teorema espectral)

Una matriz cuadrada es simétrica real si y sólo si es ortogonalmente diagonalizable.

DEMOSTRACIÓN Ya sabemos que si A es ortogonalmente diagonalizable, debe ser simétrica, según el teorema 23. Por el contrario, si A es simétrica tiene eigenvalores reales, según el teorema 24. Por tanto, de acuerdo con el teorema de descomposición de Schur, hay una matriz ortogonal Q y una matriz triangular D tales que

$$Q^TAQ = D$$

Como $A^T = A$,

$$D^T = (Q^TAQ)^T = Q^TA^TQ^{TT} = Q^TAQ = D$$

Vemos que D es simétrica y triangular superior, y por tanto es diagonal. Por lo anterior, A es ortogonalmente diagonalizable. $\square$

Ahora que sabemos que las matrices simétricas son ortogonalmente diagonalizables, podemos describir el procedimiento aplicado en el ejemplo 22, para determinar Q y D.

Algoritmo

(Diagonalización de una matriz simétrica)

DATOS: Matriz simétrica A $n \times n$.

1. Calcular todos los eigenvalores de A. Sean $\lambda_1, \ldots, \lambda_n$ todos los eigenvalores *distintos*. (Son todos reales y algunos posiblemente sean múltiples.)
2. Determinar una base $\mathcal{B}_i$ de eigenvectores para cada eigenespacio E_{λ_i}, $i = 1, \ldots, k$. (La unión $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ es una base de eigenvectores de A, porque A es simétrica y, en consecuencia, diagonalizable.)

3. Aplicar, si es necesario, el proceso de Gram-Schmidt a cada $\mathcal{B}_i$ para obtener conjuntos ortogonales $\mathcal{B}_i'$. (Así, cada $\mathcal{B}_i'$ es linealmente independiente, en forma automática. Como Gen($\mathcal{B}_i$) = Gen($\mathcal{B}_i'$), cada $\mathcal{B}_i'$ forma una base ortogonal para E_{λ_i}.)
4. Sean $\mathbf{u}_1, \ldots, \mathbf{u}_n$ los vectores de $\mathcal{B}_1', \ldots, \mathcal{B}_k'$. Forman una base ortogonal de eigenvectores de A. (Como los vectores de la misma $\mathcal{B}_i'$ son ortogonales y los eigenvectores de distintas $\mathcal{B}_i'$ son eigenvectores que corresponden a eigenvalores distintos, éstos también son ortogonales).
5. Sean $\mathbf{v}_1, \ldots, \mathbf{v}_n$ las normalizaciones de los $\mathbf{u}_i$. Forman una base ortonormal de eigenvectores de A.
6. Si $Q = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, Q es ortogonal.
7. D es la matriz diagonal cuyos elementos diagonales son los eigenvalores correspondientes, en el mismo orden, y con elementos repetidos para los eigenvalores múltiples.
8. Q y D diagonalizan ortogonalmente a A.

RESULTADOS: Q es ortogonal y D diagonal, tales que $Q^TAQ = D$.

Demostración y ejemplo del teorema de descomposición de Schur

Sea $\mathbf{v}_1$ un eigenvector unitario de A, y λ_1 su eigenvalor (real) correspondiente. Siguiendo el proceso de Gram-Schmidt, hay una matriz ortogonal $Q_1 = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ cuya primera columna es $\mathbf{v}_1$. Entonces

$$Q_1^TAQ_1 = \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} [A\mathbf{v}_1 \cdots A\mathbf{v}_n] = \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} [\lambda_1\mathbf{v}_1 \cdots A\mathbf{v}_n] = \begin{bmatrix} \lambda_1 & \vdots & * \\ \ldots & \vdots & \ldots \\ \mathbf{0} & \vdots & A_1 \end{bmatrix}$$

porque $\lambda_1\mathbf{v}_1^T\mathbf{v}_1 = \lambda\mathbf{v}_1 \cdot \mathbf{v}_1 = \lambda_1$ (porque $\mathbf{v}_1$ es unitario) y $\lambda_1\mathbf{v}_i^T\mathbf{v}_1 = \lambda_1\mathbf{v}_1 \cdot \mathbf{v}_i = \lambda_1 0 = 0$, cuando $i \neq 1$, por la ortogonalidad. Ahora, si λ_2 es un eigenvalor de A_1, también lo es de A y en consecuencia λ_2 es real. Aplicamos el mismo procedimiento a la matriz A_1 de $(n-1) \times (n-1)$ para determinar la matriz Q_2 ortogonal tal que

$$Q_2^TA_1Q_2 = \begin{bmatrix} \lambda_2 & \vdots & * \\ \ldots & \vdots & \ldots \\ \mathbf{0} & \vdots & A_2 \end{bmatrix}$$

La matriz $Q_2' = \begin{bmatrix} 1 & \vdots & \mathbf{0} \\ \ldots & \vdots & \ldots \\ \mathbf{0} & \vdots & Q_2 \end{bmatrix}$ tiene la propiedad siguiente:

$$Q_2'^T(Q_1^T A Q_1)Q_2' = \begin{bmatrix} 1 & \vdots & \mathbf{0} \\ \cdots & \vdots & \cdots \\ \mathbf{0} & \vdots & Q_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 & \vdots & * \\ \cdots & \vdots & \cdots \\ \mathbf{0} & \vdots & A_1 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \cdots & \cdots \\ \mathbf{0} & Q_2 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & * \\ \cdots & \\ \mathbf{0} & Q_2^T A_1 Q_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ \cdots & & \cdots \\ \mathbf{0} & \mathbf{0} & A_2 \end{bmatrix}$$

Continuando en la misma forma, después de $n-1$ pasos, obtenemos una matriz ortogonal Q.

$$Q = Q_1 \begin{bmatrix} 1 & \mathbf{0} \\ \cdots & \cdots \\ \mathbf{0} & Q_2 \end{bmatrix} \begin{bmatrix} I_2 & \mathbf{0} \\ \cdots & \cdots \\ \mathbf{0} & Q_3 \end{bmatrix} \cdots \begin{bmatrix} I_{n-2} & \mathbf{0} \\ \cdots & \cdots \\ \mathbf{0} & Q_{n-1} \end{bmatrix}$$

tal que $Q^T AQ$ es triangular superior:

$$Q^T AQ = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & & * \\ \vdots & & & \\ 0 & 0 & & \lambda_n \end{bmatrix}$$

con lo que se completa la demostración del teorema de la descomposición de Schur.

■ EJEMPLO 23 Determine la descomposición de Schur de

$$A = \begin{bmatrix} 9 & 8 \\ 1 & 2 \end{bmatrix}$$

SOLUCIÓN Los eigenvalores 1 y 10 de A son reales y 1 tiene el eigenvector unitario $\mathbf{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. De acuerdo con la demostración del teorema, se necesita una matriz ortogonal $Q_1 = [\mathbf{v}_1\ \mathbf{v}_2]$. Podemos tomar $Q_1 = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. (Compruébelo.) Entonces $T = Q_1^T AQ_1 = \begin{bmatrix} 1 & -7 \\ 0 & 10 \end{bmatrix}$. Así

$$\begin{bmatrix} 9 & 8 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & -7 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$\square$

Ejercicios 8.5

En los ejercicios 1 a 4 determine cuáles matrices son simétricas.

1. $A = \begin{bmatrix} 0 & 2 \\ -2 & 3 \end{bmatrix}$
2. $A = \begin{bmatrix} 0 & -2 \\ -2 & 3 \end{bmatrix}$
3. $A = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 0 & 1 \\ 2 & -1 & 5 \end{bmatrix}$
4. $A = \begin{bmatrix} 1 & 3 & 0 \\ 3 & 0 & 1 \\ 0 & 1 & 5 \end{bmatrix}$
5. Deduzca una matriz A 2×2 que sea simétrica y ortogonal a la vez. Calcule A^2.
6. Demuestre que si una matriz A es simétrica y ortogonal al mismo tiempo, entonces $A^2 = I$. Llegue a la conclusión de que $A = A^{-1}$.
7. Sin hacer cálculos, determine la inversa de $\begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}$.

En los ejercicios 8 a 15, diagonalice ortogonalmente las matrices.

8. $A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$
9. $A = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$
10. $A = \begin{bmatrix} -4 & 2 \\ 2 & -4 \end{bmatrix}$
11. $A = \begin{bmatrix} 6 & -4 \\ -4 & 6 \end{bmatrix}$
12. $A = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 2 \end{bmatrix}$
13. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 2 \end{bmatrix}$
14. $A = \begin{bmatrix} 5 & -4 & 0 \\ -4 & 3 & 4 \\ 0 & 4 & 1 \end{bmatrix}$
15. $A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & -2 & 2 \end{bmatrix}$
16. Suponga que A y B son $n \times n$ y ortogonalmente diagonalizables, con elementos reales, y sea c cualquier escalar real. Aplique el teorema espectral para demostrar que las siguientes también son ortogonalmente diagonalizables.

    a. $A + B$ **b.** $A - B$
    c. cA **d.** A^2
17. Demuestre que si A es real y simétrica, entonces A^2 tiene eigenvalores no negativos. (*Sugerencia:* Aplique el teorema espectral y escriba Q^TA^2Q en la forma $(Q^TAQ)(Q^TAQ)$.)
18. Verifique que las multiplicidades geométrica y algebraica de cada eigenvalor de una matriz simétrica son iguales.

En los ejercicios 19 a 22 determine la descomposición de Schur para A.

19. $A = \begin{bmatrix} 9 & 8 \\ 2 & 3 \end{bmatrix}$
20. $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$
21. $A = \begin{bmatrix} 10 & 9 \\ 2 & 3 \end{bmatrix}$
22. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & 1 & 2 \end{bmatrix}$

Descomposición espectral

23. Suponga que $A = QDQ^T$ es una diagonalización ortogonal de A. Si $Q = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ y D tiene los elementos diagonales $\lambda_1, \ldots, \lambda_n$, compruebe que

$$A = \lambda_1\mathbf{v}_1\mathbf{v}_1^T + \cdots + \lambda_n\mathbf{v}_n\mathbf{v}_n^T$$

    A esto se le llama **descomposición espectral** de A.
24. Determine la descomposición espectral de $A = \begin{bmatrix} 12 & 3 \\ 3 & 4 \end{bmatrix}$
25. Determine la descomposición espectral de $A = \begin{bmatrix} 13 & 4 \\ 4 & 7 \end{bmatrix}$

8.6 Formas cuadráticas y secciones cónicas

Objetivos del estudiante para esta sección
1. Definir y diagonalizar una forma cuadrática.
2. Identificar una sección cónica que no esté en la posición canónica.

En esta sección aplicaremos la diagonalización de matrices simétricas (que describimos en la sección 8.5) para estudiar *expresiones cuadráticas* tales como

$$ax^2 + by^2 + cxy + dx + ey + f$$

o como

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + kz + l$$

En realidad, sólo nos interesan los términos cuadráticos (o *principales*) de esas sumas:

$$ax^2 + by^2 + cxy \tag{8.28}$$

y

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz \tag{8.29}$$

A las expresiones semejantes a (8.28) y (8.29) se les llama **formas cuadráticas**, y se usan en una gran variedad de problemas de matemáticas, física, mecánica, economía, estadística, robótica y procesamiento de imágenes, y en muchas aplicaciones industriales.

Las formas cuadráticas pueden escribirse como productos matriciales $\mathbf{x}^T A\mathbf{x}$. Por ejemplo,

$$3x^2 + 7y^2 - 2xy = [\,x \quad y\,]\begin{bmatrix} 3 & -1 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

o hasta

$$3x^2 + 7y^2 - 2xy = [\,x \quad y\,]\begin{bmatrix} 3 & -2 \\ 0 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Se prefiere la primera descomposición frente a la segunda, porque la primera matriz cuadrada es simétrica. De hecho, al escribir una forma cuadrática como $\mathbf{x}^T A\mathbf{x}$ siempre podremos suponer que A es simétrica. A es la matriz **asociada** de la forma. Así,

$$ax^2 + by^2 + cxy = [\,x \quad y\,]\begin{bmatrix} a & \frac{1}{2}c \\ \frac{1}{2}c & b \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

y

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz = [\,x \quad y \quad z\,]\begin{bmatrix} a & \frac{1}{2}d & \frac{1}{2}e \\ \frac{1}{2}d & b & \frac{1}{2}f \\ \frac{1}{2}e & \frac{1}{2}f & c \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

La razón principal por la que se desea que A sea simétrica es que este tipo de matrices pueden *diagonalizarse ortogonalmente*. En general, es posible definir a las formas cuadráticas con n variables usando sólo matrices simétricas.

DEFINICIÓN

Una **forma cuadrática** (con n variables) es una función $q : \mathbf{R}^n \to \mathbf{R}$ de la forma

$$q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

para una matriz A *simétrica* $n \times n$ y cualquier vector n $\mathbf{x}$. Se dice que A es la matriz **asociada** con q.

■ EJEMPLO 24 Demuestre que el producto punto en $\mathbf{R}^n$ define a una forma cuadrática q como

$$q(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} \quad \text{es decir} \quad q(\mathbf{x}) = \|\mathbf{x}\|^2$$

¿Cuál es la matriz asociada?

SOLUCIÓN Puesto que

$$q(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = \mathbf{x}^T\mathbf{x} = \mathbf{x}^T I\mathbf{x}$$

vemos que q es una forma cuadrática cuya matriz asociada es I. □

Las formas cuadráticas con dos variables, $q(x, y) = ax^2 + by^2 + cxy$, se relacionan estrechamente con las secciones **cónicas**, que se estudian en el álgebra y la geometría analítica. De hecho, si $c = 0$, el término **mixto** o **cruzado** cxy es cero y la ecuación $q(x, y) = ax^2 + by^2 = 1$ representa en general una *elipse* o una *hipérbola*.[5]

■ EJEMPLO 25 Describa la sección cónica que define cada una de las ecuaciones siguientes:

(a) $x^2 + 2y^2 = 1$
(b) $2x^2 + y^2 = 1$
(c) $x^2 - 2y^2 = 1$
(d) $-x^2 + 2y^2 = 1$

SOLUCIÓN Las ecuaciones (a) y (b) representan elipses, y las (c) y (d) son hipérbolas (figura 8.13). □

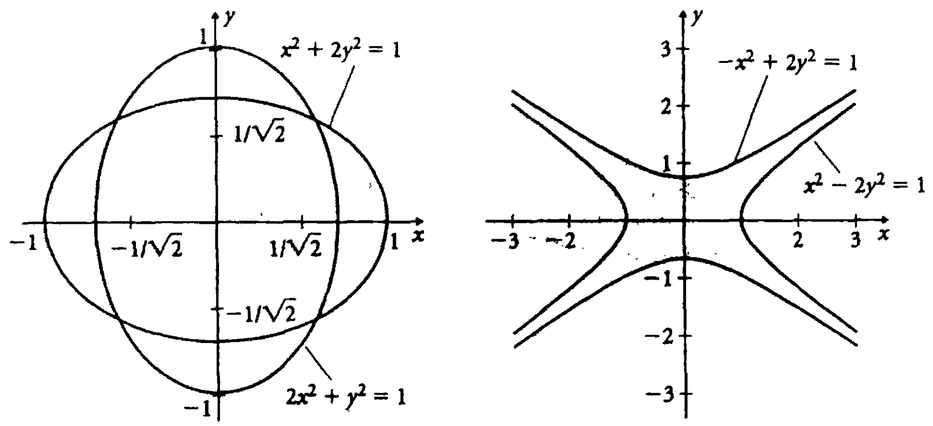


Figura 8.13

[5] Las parábolas (como $y = 2x^2$ o $y^2 = 3x$) también son secciones cónicas, pero como contienen términos no cuadráticos no las incluiremos en nuestra explicación.

Una forma cuadrática con dos variables sin términos mixtos casi siempre representa una elipse o una hipérbola en **posición canónica** (figuras 8.13 y 8.14). Esto quiere decir que los ejes principales de esas cónicas son los ejes x y y. En este caso, la matriz de la forma cuadrática es **diagonal**.

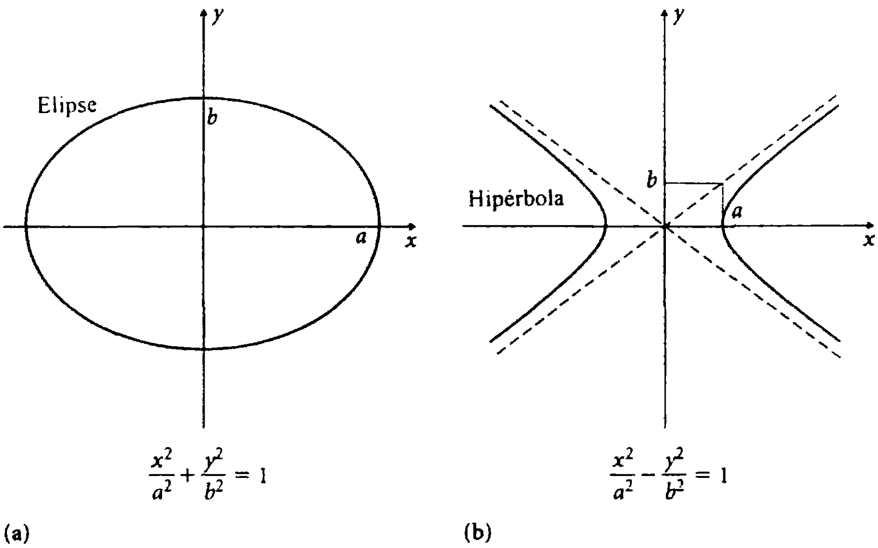


Figura 8.14

¿Qué sucede si nuestra forma cuadrática $q(x, y) = ax^2 + by^2 + cxy$ *tiene* un término mixto, es decir, si $c \neq 0$? Aplicamos un cambio de variables de modo que, con respecto a las nuevas variables x' y y', no haya términos mixtos: $q(x', y') = a'x'^2 + b'y'^2$. Entonces puede identificarse a q, en el nuevo sistema de coordenadas, como una sección cónica. Éste será el tema siguiente.

Diagonalización de formas cuadráticas

Sea $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ cualquier forma cuadrática con n variables. Como A es simétrica, podemos diagonalizarla ortogonalmente, por ejemplo, por medio de Q ortogonal y D diagonal. Recuerde que en la sección 8.5 dijimos que Q tiene n eigenvectores ortonormales de A como columnas, y que D es diagonal, con sus elementos diagonales iguales a los eigenvalores correspondientes. Si aplicamos el cambio de variables familiar del capítulo 7,

$$\mathbf{x} = Q\mathbf{y}, \quad \text{o sea} \quad \mathbf{y} = Q^{-1}\mathbf{x} = Q^T\mathbf{x} \tag{8.30}$$

llegamos a

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^T A\mathbf{x} \\ &= (Q\mathbf{y})^T AQ\mathbf{y} \\ &= \mathbf{y}^T Q^T AQ\mathbf{y} \\ &= \mathbf{y}^T D\mathbf{y} \end{aligned}$$

Así, q tiene los mismos valores de una forma cuadrática con las nuevas variables **y**, cuya matriz es D. En las nuevas variables no hay términos mixtos, porque D es diagonal. A este proceso se le llama **diagonalización** de q. Como los elementos diagonales en D son los eigenvalores de A, hemos demostrado el teorema 28.

TEOREMA 28

(Ejes principales)

Si A es una matriz simétrica $n \times n$, diagonalizada ortogonalmente por Q y D; entonces, el cambio de variables $\mathbf{x} = Q\mathbf{y}$ transforma la forma cuadrática $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ en la forma $\mathbf{y}^T D\mathbf{y}$, que no tiene términos mixtos. De hecho, si $\lambda_1, \ldots, \lambda_n$ son eigenvalores de A, y si $\mathbf{y} = (y_1, \ldots, y_n)$, entonces

$$q(\mathbf{x}) = q(\mathbf{y}) = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

■ EJEMPLO 26 Escribir $q(x, y) = x^2 - 2xy + y^2$ sin términos mixtos. Calcule $q(1, -1)$ empleando las variables anteriores y nuevas.

SOLUCIÓN La matriz de q es $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ cuyos eigenvectores básicos son $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, y sus eigenvalores correspondientes son 0 y 2. Como los eigenvectores ya son ortogonales, sólo necesitamos normalizarlos para obtener Q. Entonces

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \qquad \text{y} \qquad D = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

Consideremos ahora las nuevas variables $\mathbf{y} = (x', y')$ tales que $\mathbf{x} = (x, y) = Q\mathbf{y}$.

$$\begin{aligned} \mathbf{y} = \begin{bmatrix} x' \\ y' \end{bmatrix} &= Q^T \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} \frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} \\ \frac{x}{\sqrt{2}} - \frac{y}{\sqrt{2}} \end{bmatrix} \end{aligned}$$

y

$$q(\mathbf{y}) = \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = 2y'^2$$

Por consiguiente, $q(\mathbf{y}) = 2y'^2$ sin términos mixtos. Para evaluar q en $(1, -1)$ en las variables anteriores, tenemos

$$q(1, -1) = 1^2 - 2 \cdot 1 \cdot (-1) + 1^2 = 4$$

En las nuevas variables primero determinamos

$$\mathbf{y} = \begin{bmatrix} x' \\ y' \end{bmatrix} = Q^T \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\dfrac{2}{\sqrt{2}} \end{bmatrix}$$

y a continuación

$$\begin{bmatrix} 0 & -\dfrac{2}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ -\dfrac{2}{\sqrt{2}} \end{bmatrix} = 4$$

Nos interesa la diagonalización *ortogonal* de las formas cuadráticas, porque entonces el cambio de variables se hace con una **transformación ortogonal** $Q\mathbf{x}$ (quiere decir que Q es ortogonal). Así, se conservan las longitudes o normas, y los ángulos de cualquier vector transformado. Por tanto, las formas de las curvas, superficies, cuerpos, etc., también se preservan en las nuevas coordenadas.

Aplicaciones de las formas cuadráticas a la geometría

Cuando $n = 2, 3$, puede usarse el teorema de los ejes principales para identificar secciones cónicas, o también superficies cuádricas con términos mixtos.

Secciones cónicas: elipses e hipérbolas

■ EJEMPLO 27 Use la diagonalización para identificar a las secciones cónicas $q_1(x, y) = 1$ y $q_2(x, y) = 1$, donde

$$q_1(x,y) = 2x^2 + 2y^2 - 2xy$$
$$q_2(x,y) = -x^2 - y^2 + 6xy$$

SOLUCIÓN Sean $A_1 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ y $A_2 = \begin{bmatrix} -1 & 3 \\ 3 & -1 \end{bmatrix}$ las matrices correspondientes. A_1 y A_2 pueden diagonalizarse fácilmente con Q_1, D_1 y Q_2, D_2, siendo

$$Q_1 = \begin{bmatrix} \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix} = Q_2$$

y

$$D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \qquad D_2 = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$$

Por consiguiente,

$$q_1(x', y') = [x' \quad y'] \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = x'^2 + 3y'^2$$

y

$$q_2(x', y') = [x' \quad y'] \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = 2x'^2 - 4y'^2$$

Así, en las nuevas variables tenemos una elipse $x'^2 + 3y'^2 = 1$ y una hipérbola $2x'^2 - 4y'^2 = 1$. Para trazar esas curvas sólo necesitamos conocer qué vectores transforman a (1, 0) y a (0, 1) en x' y y'. Como

$$Q_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = Q_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}$$

$$Q_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = Q_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} \end{bmatrix}$$

vemos que en el nuevo sistema (1, 0) y (0, 1) son $(1/\sqrt{2}, 1/\sqrt{2})$, y $(1/\sqrt{2}, -1/\sqrt{2})$ en el anterior. En consecuencia, la elipse y la parábola están giradas 45° respecto a la posición canónica (figura 8.15). □

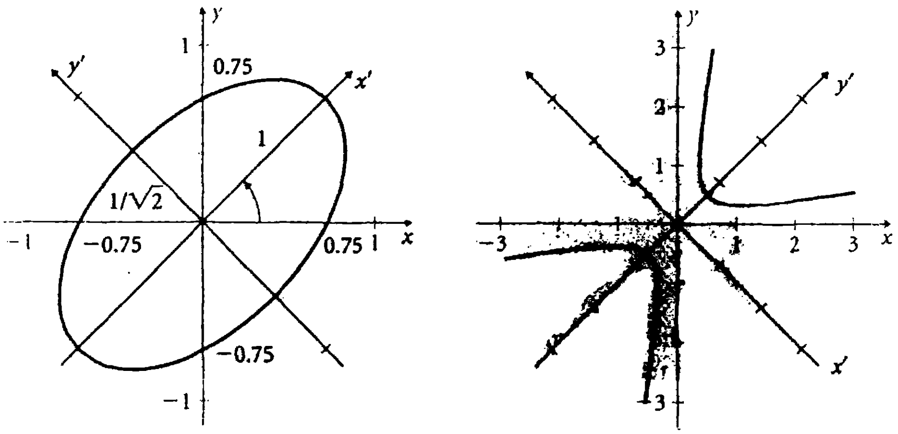


Figura 8.15 Cónicas en posición no canónica.

■ EJEMPLO 28 Use la diagonalización para identificar las secciones cónicas $q_1(x, y) = 1$ y $q_2(x, y) = 1$, donde

$$q_1(x,y) = 2x^2 + 2y^2 + 2xy$$
$$q_2(x,y) = 2x^2 + 2y^2 - 4xy$$

SOLUCIÓN Es fácil observar que q_1 y q_2 se diagonalizan, respectivamente, con

$$Q_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

y con

$$Q_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \qquad D_2 = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$$

Por consiguiente,

$$q_1(x', y') = [\, x' \quad y' \,] \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = x'^2 + 3y'^2$$

y

$$q_2(x', y') = \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = 4y'^2$$

Encontramos que $q_1(x', y') = x'^2 + 3y'^2 = 1$ es una elipse en el sistema $x'y'$, igual que en el ejemplo 27. Sin embargo, esta vez $Q_1(1, 0) = (-1/\sqrt{2}, 1/\sqrt{2})$ y $Q_1(0, 1) = (1/\sqrt{2}, 1/\sqrt{2})$, de modo que el eje de x' positivas es la semirrecta que forma un ángulo de 135°, y el eje de y' positivas es la semirrecta que forma un ángulo de 45° con el eje positivo de las x anteriores. De hecho, como

$$Q_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

la transformación definida por Q_1 es una rotación de −45° (la segunda matriz) seguida de una reflexión respecto a la primera matriz $y = x$ (figura 8.16).

Como $q_2(x', y') = 4y'^2 = 1$, resulta $y' = \pm\frac{1}{2}$, así que esta vez no obtenemos una elipse ni una hipérbola, sino dos rectas en el sistema $x'y'$ (figura 8.16). Éste es un ejemplo de una forma cuadrática **degenerada**. □

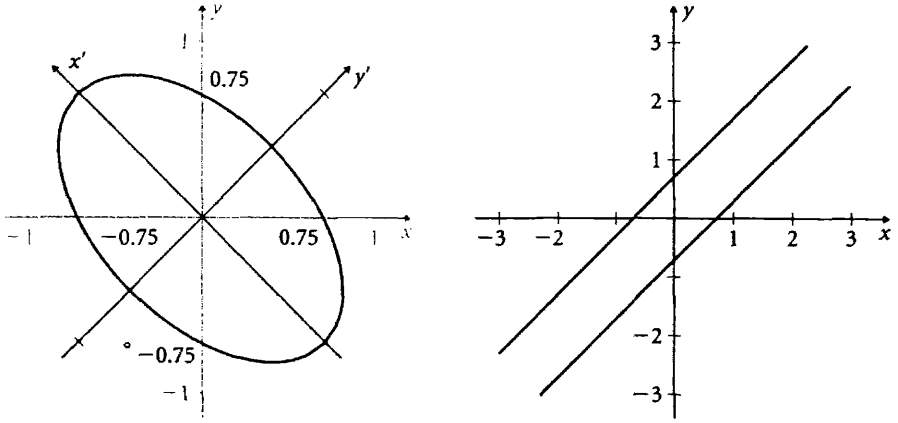


Figura 8.16 (a) Rotación seguida por reflexión; (b) forma degenerada: dos rectas paralelas.

Superficies cuádricas: elipsoides

Podemos aplicar estos métodos para identificar superficies cuádricas. Por ejemplo, una ecuación de la forma

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

para $a, b, c > 0$ es un elipsoide en posición canónica. Los cortes transversales de esa superficie con los planos coordenados son elipses. La figura 8.17(a) muestra el elipsoide $3x^2 + 6y^2 + 9z^2 = 1$.

■ EJEMPLO 29 Identifique la superficie cuádrica $5x^2 + 6y^2 + 7z^2 + 4xy + 4yz = 1$.

SOLUCIÓN La matriz de $q(x, y, z) = 5x^2 + 6y^2 + 7z^2 + 4xy + 4yz$ es

$$A = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$$

cuyos eigenvalores son 2, 6, 9, y los eigenvectores correspondientes son $(2, -2, 1)$, $(2, 1, -2)$ y $(1.\ 2, 2)$. Como todos son ortogonales, los normalizamos para tener

$$Q = \frac{1}{3}\begin{bmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

Usaremos el cambio de variables $\mathbf{y} = Q^T\mathbf{x}$ para obtener

$$q(x, y, z) = [\,x' \quad y' \quad z'\,] \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = 3x'^2 + 6y'^2 + 9z'^2$$

Por consiguiente, $q(x, y, z) = 1$ toma la forma $3x'^2 + 6y'^2 + 9z'^2 = 1$ en el nuevo sistema. Es claro que la gráfica es un elipsoide en las coordenadas $x'y'z'$. La figura 8.17 (b) muestra que este elipsoide tiene algún giro e inclinación, en comparación con el de la misma ecuación en posición canónica. Dejaremos que el estudiante ubique y trace los ejes nuevos.

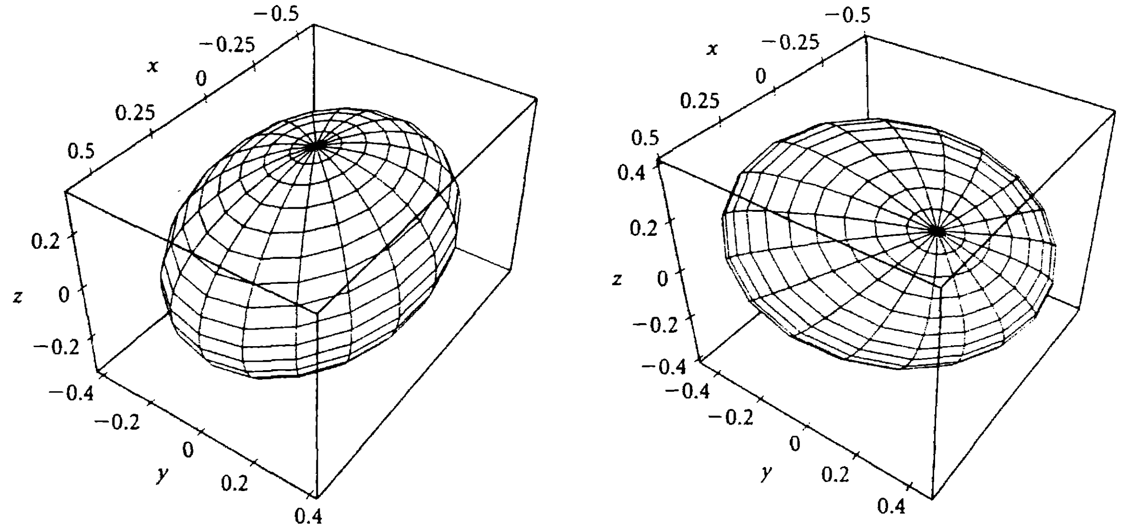


Figura 8.17 Elipsoides en posición (a) canónica y (b) no canónica.

Formas cuadráticas definidas positiva y negativa

En esta sección clasificaremos las formas cuadráticas según sus valores posibles. Como los valores de cualquier forma cuadrática $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ son números reales, pueden ser mayores que, menores que o iguales a 0. De este modo podemos clasificar a q ya sea que sus valores sean *siempre* positivos en vectores distintos de cero, o *siempre* negativos. (Tenemos que excluir a $\mathbf{0}$, porque $q(\mathbf{0}) = 0$.)

DEFINICIÓN

Sea $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ la forma cuadrada con A simétrica.

1. Si $q(\mathbf{x}) > 0$ para todo $\mathbf{x} \neq \mathbf{0}$, q se llama **definida positiva**.
2. Si $q(\mathbf{x}) < 0$ para todo $\mathbf{x} \neq \mathbf{0}$, q se llama **definida negativa**.
3. Si $q(\mathbf{x})$ toma valores positivos y negativos, q se llama **indefinida**.

También usaremos la misma terminología para la matriz simétrica asociada A. Por ejemplo, A es **definida positiva** si $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ es una forma cuadrática definida positiva.

Además de las formas de tipo básico, están las formas cuadráticas y matrices simétricas **semidefinidas, positivas** y **negativas**, ya sea que $q(\mathbf{x}) \geq 0$ o $q(\mathbf{x}) \leq 0$ para todo $\mathbf{x} \neq \mathbf{0}$.

El teorema de los ejes principales puede aplicarse con facilidad para identificar el tipo de una forma cuadrática examinando los signos de los eigenvalores de su matriz.

TEOREMA 29

Una forma cuadrática $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ con A simétrica es

1. Definida positiva si y sólo si todos los eigenvalores de A son > 0.
2. Definida negativa si y sólo si todos los eigenvalores de A son < 0.
3. Indefinida si y sólo si A tiene eigenvalores positivos y negativos.

DEMOSTRACIÓN Se deja como ejercicio. □

■ EJEMPLO 30 (Relatividad) Demuestre que la forma cuadrática q, que se usa en la teoría de la relatividad para definir la distancia en el espacio tiempo, es indefinida.

$$q(\mathbf{x}) = [x \quad y \quad z \quad t] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = x^2 + y^2 + z^2 - t^2$$

SOLUCIÓN Según el teorema 29, q es indefinida porque su matriz tiene eigenvalores positivos y negativos a la vez. □

NOTA La aparición de los signos en la fórmula de una forma cuadrática a veces puede ser decepcionante. Por ejemplo, podríamos sentirnos inclinados a decir que la forma $q(\mathbf{x}) = x^2 + y^2 + 10xy$ es definida positiva, pero en realidad es indefinida porque los eigenvalores de $\begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$ son -4 y 6. De hecho, $q(1, -1) = -8 < 0$.

Ejercicios 8.6

En esta sección todas las matrices no especificadas tienen elementos reales. En los ejercicios 1 a 4 evalúe la forma cuadrática $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ para A y $\mathbf{x}$ dados.

1. $A = \begin{bmatrix} -2 & 2 \\ 2 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$

2. $A = \begin{bmatrix} 4 & 7 \\ 7 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 0 & 1 \\ 2 & 1 & 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 0 & 1 \\ 2 & 1 & 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

En los ejercicios 5 a 12 determine la matriz simétrica A de la forma cuadrática.

5. $q(x,y) = 3x^2 - 6xy + 3y^2$

6. $q(x,y) = -x^2 + 10xy - y^2$

7. $q(x,y) = -4x^2 + 2xy - 4y^2$

8. $q(x,y) = 6x^2 - 2xy + 6y^2$

9. $q(x,y,z) = 2x^2 + 2xz + 2z^2$

10. $q(x,y,z) = x^2 + 2y^2 + 8yz + 2z^2$

11. $q(x,y,z) = 5x^2 - 8xy + 3y^2 + 12yz + z^2$

12. $q(x,y,z,w) = x^2 + 2xy + y^2 + 2z^2 + 4zw + 2w^2$

En los ejercicios 13 a 19 diagonalice ortogonalmente la forma cuadrática. Use un cambio de variables para replantear la fórmula sin términos mixtos.

13. $q(x,y) = 3x^2 - 2xy + 3y^2$

14. $q(x,y) = -x^2 + 4xy - y^2$

15. $q(x,y) = -4x^2 + 4xy - 4y^2$

16. $q(x,y) = 6x^2 - 8xy + 6y^2$

17. $q(x,y,z) = 2x^2 - 2xz + 2z^2$

18. $q(x,y,z) = x^2 + 2y^2 - 4yz + 2z^2$

19. $q(x,y,z) = 5x^2 - 8xy + 3y^2 + 8yz + z^2$

20. Identifique la sección cónica $3x^2 - 2xy + 3y^2 = 1$.

21. Defina la sección cónica $5x^2 - 8xy + 5y^2 = 1$.

22. Identifique la superficie cuadrática $6x^2 + 8xy + 4xz + 10y^2 + 12yz + 11z^2 = 1$.

23. Compruebe el teorema 29.

24. Demuestre que la forma cuadrática $q(x, y) = ax^2 + bxy + cy^2$ es positiva definida si y sólo si $a > 0$ y $b^2 - 4ac < 0$.

25. Verifique que si A es simétrica, entonces la forma cuadrática $q(\mathbf{x}) = \mathbf{x}^TA^2\mathbf{x}$ es semidefinida positiva. (*Sugerencia:* Use el teorema espectral de la sección 8.5 y escriba Q^TA^2Q en la forma $(Q^TAQ)(Q^TAQ)$.)

26. Sea $A = P^TP$, en la que P es una matriz invertible. Demuestre que A es definida positiva. (*Sugerencia:* A es simétrica y semidefinida, porque $\mathbf{x}^TA\mathbf{x} = \mathbf{x}^TP^TP\mathbf{x} = (P\mathbf{x})^TP\mathbf{x} = \|P\mathbf{x}\|^2 \geq 0$. A continuación demuestre que si $\mathbf{x} \neq \mathbf{0}$, entonces $\mathbf{x}^TA\mathbf{x} > 0$.)

27. Si A es una matriz definida positiva (y por tanto simétrica), demuestre que existe una matriz invertible P tal que $A = P^TP$. (*Sugerencia:* Use el teorema espectral de la sección 8.5, y escriba A en la forma QDQ^T, donde Q es ortogonal y D es diagonal. A continuación determine una matriz B tal que $D = B^TB$. ¿Qué debe ser P?)

28. Compruebe que A es positiva definida si y sólo si existe una matriz invertible P tal que $A = P^TP$.

Completar el cuadrado

La fórmula familiar para completar el cuadrado

$$A^2 + bA = \left(A^2 + 2\left(\frac{b}{2}\right)A + \frac{b^2}{4}\right) - \frac{b^2}{4}$$
$$= \left(A + \frac{b}{2}\right)^2 - \frac{b^2}{4}$$

puede usarse para convertir una forma cuadrática con dos variables en una sin términos mixtos.

29. Sea $q(x, y) = ax^2 + bxy + cy^2$. Si $a \neq 0$, complete el cuadrado para escribir q en la forma $aX^2 + By^2$ para una B constante y una nueva variable X que depende de x y de y.

30. Aplique la fórmula del ejercicio 29 para escribir $q(x, y) = 3x^2 - 2xy + 3y^2$ sin términos mixtos.

31. Para el ejercicio 29, si $a = 0$ y $c \neq 0$, ¿es posible completar el cuadrado y escribir q sin términos mixtos?

32. Para el ejercicio 29, si $a = 0$ y $c = 0$, ¿se puede escribir q sin términos mixtos?

8.7 La descomposición en valores singulares (SVD)

Objetivos del estudiante para esta sección

1. Aprender a calcular la descomposición de una matriz en valores singulares.
2. Valorar la importancia teórica y práctica de este método.

Hemos visto que las descomposiciones de matrices en factores con propiedades especiales tienen mucha utilidad. Entre los ejemplos se incluyen la LU, la QR, la diagonalización, la diagonalización ortogonal, la descomposición de Schur, etc. Una factorización tiene interés especial si algunos de los factores son matrices ortogonales. La razón es que las transformaciones ortogonales preservan normas y ángulos; en particular, preservan las longitudes de los vectores de error que son inevitables en los cálculos numéricos.

En esta sección estudiaremos una de las factorizaciones más importantes que se aplica a cualquier matriz A $m \times n$, la **descomposición en valores singulares** (SVD, por sus siglas en inglés). Este método tiene interés teórico y aplicado a la vez, y es bastante antigua.[6] De hecho, es tan útil que merece mucho más atención y crédito. Entre sus muchas aplicaciones está la estimación más fiable del rango de una matriz.

Nuestro objetivo es factorizar cualquier matriz A $m \times n$ en la forma

$$A = U\Sigma V^T$$

donde U es $m \times m$, V es $n \times n$ y ambas son ortogonales. También, Σ es una matriz $m \times n$ que contiene un bloque diagonal superior izquierdo con elementos positivos de magnitud decreciente, y los elementos restantes son 0. Así,

$$\Sigma = \begin{bmatrix} D & \vdots & \mathbf{0} \\ \cdots & \cdots & \cdots \\ \mathbf{0} & & \mathbf{0} \end{bmatrix}, \qquad \text{donde } D = \begin{bmatrix} \sigma_1 & & 0 \\ \vdots & & \vdots \\ 0 & & \sigma_r \end{bmatrix} \tag{8.31}$$

y

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \qquad r \leq m, n$$

A continuación presentamos algunos ejemplos de Σ con $r = 2$:

$$\begin{bmatrix} 6 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 9 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las D correspondientes son

$$\begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 9 & 0 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix}, \quad \begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix}$$

Valores singulares: determinación de V, Σ y U

Primero definiremos V, después determinaremos las σ_i a lo largo de la diagonal de D para formar Σ. Consideremos la matriz simétrica A^TA, $n \times n$. Según el teorema espectral, A^TA es ortogonalmente diagonalizable y tiene eigenvalores reales, digamos $\lambda_1, \ldots, \lambda_n$. Si $\mathbf{v}_1, \ldots, \mathbf{v}_n$ son los eigenvectores correspondientes que forman una base ortonormal de $\mathbf{R}^n$, V es simplemente

$$V = [\mathbf{v}_1\ \mathbf{v}_2 \cdots \mathbf{v}_n]$$

A continuación, observaremos que todos los eigenvalores son no negativos (y así A^TA es semidefinida positiva). Como $(A^TA)\mathbf{v}_i = \lambda_i\mathbf{v}_i$ y $\|\mathbf{v}_i\| = 1$, entonces

[6] Se menciona en "Sulle Funzioni Bilineari", de E. Beltrami, *Giornale di Mathematische 11* (1873), pp. 98-106.

$$\begin{aligned}
0 &\le \|A\mathbf{v}_i\|^2 \\
&= (A\mathbf{v}_i)^T A\mathbf{v}_i \\
&= \mathbf{v}_i^T A^T A\mathbf{v}_i \\
&= \mathbf{v}_i^T \lambda_i \mathbf{v}_i \\
&= \lambda_i \|\mathbf{v}_i\|^2 \\
&= \lambda_i
\end{aligned}$$

Por tanto, $\lambda_i \ge 0$ para $i = 1, \ldots, n$. Si es necesario reordenamos las λ de mayor a menor y calculamos sus raíces cuadradas, a las que llamaremos σ:

$$\sigma_1 = \sqrt{\lambda_1} \ge \cdots \ge \sigma_n = \sqrt{\lambda_n} \ge 0$$

Así,

$$\sigma_i = \|A\mathbf{v}_i\|, \qquad i = 1, \ldots, n \tag{8.32}$$

Los números $\sigma_1, \ldots, \sigma_n$ se llaman **valores singulares** de A y contienen información importante acerca de A. Si r es el entero positivo tal que

$$\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0 \qquad \text{y} \qquad \sigma_{r+1} = \cdots = \sigma_n = 0$$

Así, $\sigma_1, \ldots, \sigma_r$ son valores singulares de A distintos de cero y ordenados por magnitud. Éstos son los elementos diagonales de D en Σ.

■ EJEMPLO 31 Determine V y Σ en cada caso.

(a) $A = \begin{bmatrix} 2 & 4 \\ 1 & -4 \\ -2 & 2 \end{bmatrix}$

(b) $A = \begin{bmatrix} -2 & 1 & 2 \\ 6 & 6 & 3 \end{bmatrix}$

(c) $A = \begin{bmatrix} 0 & 6 & 6 \\ -6 & -3 & 0 \\ 6 & 0 & -3 \end{bmatrix}$

SOLUCIÓN Necesitamos A^TA, sus eigenvalores y los eigenvectores básicos correspondientes.

	A^TA	Eigenvalores	Eigenvectores
(a)	$\begin{bmatrix} 9 & 0 \\ 0 & 36 \end{bmatrix}$	$36, 9$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
(b)	$\begin{bmatrix} 40 & 34 & 14 \\ 34 & 37 & 20 \\ 14 & 20 & 13 \end{bmatrix}$	$81, 9, 0$	$\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$

	A^TA	Eigenvalores	Eigenvectores
(c)	$\begin{bmatrix} 72 & 18 & -18 \\ 18 & 45 & 36 \\ -18 & 36 & 45 \end{bmatrix}$	$81, 81, 0$	$\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$

(a) Los valores singulares de A son $\sigma_1 = \sqrt{36} = 6$, y $\sigma_2 = \sqrt{9} = 3$. Como los eigenvectores son ortonormales,

$$V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{y} \quad \Sigma = \begin{bmatrix} 6 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$$

(b) Los valores singulares de A son $\sigma_1 = 9$, $\sigma_2 = 3$ y $\sigma_3 = 0$. Los eigenvectores son ortogonales y necesitan normalización. Entonces

$$V = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \quad \text{y} \quad \Sigma = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

(c) Los valores singulares de A son $\sigma_1 = 9$, $\sigma_2 = 9$ y $\sigma_3 = 0$. Ahora necesitamos ortonormalizar los eigenvectores. Si hacemos ortogonales los dos primeros que pertenecen a E_{81}, obtenemos $(-2, 0, 1)$, $(\frac{2}{5}, 1, \frac{4}{5})$ con un proceso de Gram-Schmidt en una etapa. Así, al normalizar el conjunto ortogonal $\{(-2, 0, 1), (\frac{2}{5}, 1, \frac{4}{5}), (1, -2, 2)\}$ se tiene

$$V = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} & \frac{1}{3} \\ 0 & \frac{5}{3\sqrt{5}} & -\frac{2}{3} \\ \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & \frac{2}{3} \end{bmatrix} \quad \text{y} \quad \Sigma = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \square$$

NOTA Σ es única porque está determinada por los valores singulares ordenados. Sin embargo, el cálculo de V implica alternativas, así que V no es única. En (c) del ejemplo 31, en lugar de los eigenvectores $\mathbf{v}_1 = (-2, 0, 1)$, $\mathbf{v}_2 = (2, 1, 0)$, podríamos haber usado las combinaciones lineales $\mathbf{v}_1 + 2\mathbf{v}_2 = (2, 2, 1)$ y $2\mathbf{v}_1 + \mathbf{v}_2 = (-2, 1, 2)$. Las cuales son ortogonales, y después de la normalización se obtiene una V distinta:

$$V = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

Por último, llegamos a la definición de U. La haremos en dos etapas:

1. Primero, se forma

$$\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i \qquad \text{para } i = 1, \ldots, r \tag{8.33}$$

Estos vectores son *ortonormales*:

$$\begin{aligned}\mathbf{u}_i \cdot \mathbf{u}_j &= \frac{1}{\sigma_i\sigma_j}(A\mathbf{v}_i \cdot A\mathbf{v}_j)\\ &= \frac{1}{\sigma_i\sigma_j}(A^TA\mathbf{v}_i)\cdot\mathbf{v}_j\\ &= \frac{\lambda_i}{\sigma_i\sigma_j}\mathbf{v}_i\cdot\mathbf{v}_j\\ &= \begin{cases}0 & i\neq j\\ 1 & i=j\end{cases}\end{aligned} \tag{8.34}$$

Para $i \neq j$, los $\mathbf{v}_i$ son ortogonales, de modo que $\mathbf{v}_i \cdot \mathbf{v}_j = 0$. También, para $i = j$ tenemos $\lambda_i/\sigma_i^2 = 1$, por la definición de los valores singulares, y $\mathbf{v}_i \cdot \mathbf{v}_i = 1$, porque $\mathbf{v}_i$ es unitario.

2. A continuación extendemos el conjunto $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ hasta una base ortonormal $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ de $\mathbf{R}^m$. Esto sólo es necesario si $r < m$. Definimos

$$U = [\mathbf{u}_1\ \mathbf{u}_2 \cdots \mathbf{u}_m]$$

■ EJEMPLO 32 Determine una descomposición de A en valores singulares para el ejemplo 31(b).

SOLUCIÓN Como $\sigma_1 = 9$, $\sigma_2 = 3$, $\mathbf{v}_1 = (\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$, y $\mathbf{v}_2 = (-\frac{2}{3}, \frac{1}{3}, \frac{2}{3})$, entonces

$$\mathbf{u}_1 = \frac{1}{9}\begin{bmatrix}-2 & 1 & 2\\ 6 & 6 & 3\end{bmatrix}\begin{bmatrix}\frac{2}{3}\\ \frac{2}{3}\\ \frac{1}{3}\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{3}\begin{bmatrix}-2 & 1 & 2\\ 6 & 6 & 3\end{bmatrix}\begin{bmatrix}-\frac{2}{3}\\ \frac{1}{3}\\ \frac{2}{3}\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$$

Debido a que $m = r = 2$, $\{\mathbf{u}_1, \mathbf{u}_2\}$ no necesita extensión. Por tanto,

$$U = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}.$$

Una descomposición en valores singulares de $A = U\Sigma V^T$ es

$$\begin{bmatrix}-2 & 1 & 2\\ 6 & 6 & 3\end{bmatrix} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}9 & 0 & 0\\ 0 & 3 & 0\end{bmatrix}\begin{bmatrix}\frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3}\end{bmatrix}^T$$

□

A continuación se describe una forma de extender un conjunto ortonormal $S = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ a una base ortonormal $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$.

1. Forme $S' = \{\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{e}_1, \dots, \mathbf{e}_m\}$ y encuentre las columnas pivote de la matriz cuyas columnas son esos vectores.
2. Deduzca el subconjunto S'' de S' que contiene en las columnas pivote. S'' es una base de $\mathbf{R}^m$.
3. Aplique el proceso de Gram-Schmidt a S'' y normalice los vectores resultantes para obtener $\mathcal{B}$.

■ EJEMPLO 33 (Extensión a la base ortonormal) Determine una descomposición en valores singulares de A en el ejemplo 31(a).

SOLUCIÓN Para determinar U tenemos

$$\mathbf{u}_1 = \frac{1}{6}\begin{bmatrix} 2 & 4 \\ 1 & -4 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{3}\begin{bmatrix} 2 & 4 \\ 1 & -4 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

Ahora necesitamos extender $\{\mathbf{u}_1, \mathbf{u}_2\}$ a una base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ de $\mathbf{R}^3$. Ya que

$$\begin{bmatrix} \frac{2}{3} & \frac{2}{3} & 1 & 0 & 0 \\ -\frac{2}{3} & \frac{1}{3} & 0 & 1 & 0 \\ \frac{1}{3} & -\frac{2}{3} & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -2 & -1 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 2 & 2 \end{bmatrix}$$

las tres primeras columnas son pivote, de modo que $\{\mathbf{u}_1, \mathbf{u}_2, (1, 0, 0)\}$ forma una base de $\mathbf{R}^3$. En este caso, el proceso de Gram-Schmidt y la normalización dan como resultado $\mathbf{u}_3 = (\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$. Así,

$$U = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

Entonces, una descomposición de A en valores singulares de $A = U\Sigma V^T$ es

$$\begin{bmatrix} 2 & 4 \\ 1 & -4 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 6 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^T$$

Dejaremos que el lector compruebe la descomposición siguiente en valores singulares de A en el ejemplo 31(c), y también que determine otra basada en la primera V.

$$\begin{bmatrix} 0 & 6 & 6 \\ -6 & -3 & 0 \\ 6 & 0 & -3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}^T$$

Casi hemos demostrado el teorema básico de esta sección.

TEOREMA 30

Cuando A es cualquier matriz $m \times n$, y $\sigma_1, \ldots, \sigma_r$ son todos sus valores singulares distintos de cero, hay matrices ortogonales U ($m \times m$) y V ($n \times n$) y una matriz Σ $m \times n$ de la forma de la ecuación (8.31) tal que

$$A = U\Sigma V^T$$

DEMOSTRACIÓN U, V y Σ (de los tamaños indicados) se definieron explícitamente. Además, U y V son ortogonales por su construcción. Para demostrar que $A = U\Sigma V^T$, sólo hay que comprobar la igualdad $AV = U\Sigma$ porque $V^T = V^{-1}$. De acuerdo con la ecuación (8.33),

$$\sigma_i \mathbf{u}_i = A\mathbf{v}_i \quad \text{para} \quad i = 1, \ldots, r$$

y según la ecuación (8.32), $\|A\mathbf{v}_i\| = \sigma_i = 0$, para $i = r + 1, \ldots, n$. Entonces

$$A\mathbf{v}_i = 0 \quad \text{para } i = r + 1, \ldots, n \tag{8.35}$$

Por consiguiente,

$$\begin{aligned} AV &= [A\mathbf{v}_1 \cdots A\mathbf{v}_n] \\ &= [A\mathbf{v}_1 \cdots A\mathbf{v}_r\ \mathbf{0} \cdots \mathbf{0}] \\ &= [\sigma_1\mathbf{u}_1 \cdots \sigma_r\mathbf{u}_r\ \mathbf{0} \cdots \mathbf{0}] \\ &= [\mathbf{u}_1 \cdots \mathbf{u}_m] \begin{bmatrix} \sigma_1 & & 0 & & \\ & & & & \mathbf{0} \\ 0 & & \sigma_r & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & \mathbf{0} & & \mathbf{0} \end{bmatrix} \\ &= U\Sigma \end{aligned}$$

Esto formaliza la demostración del teorema. ☐

U, Σ, V y r (la cantidad de valores singulares distintos de cero) proporcionan información importante acerca de A.

TEOREMA 31

Sean V, Σ y U los valores singulares de las matrices de descomposición de A, $m \times n$. Sean $\sigma_1, \ldots, \sigma_r$ todos los valores singulares de A distintos de cero.

1. El rango de A es r.
2. $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ es una base ortonormal para Col(A).
3. $\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}$ es una base ortonormal para $\nu(A^T)$.
4. $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ es una base ortonormal para Ren(A).
5. $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ es una base ortonormal para $\nu(A)$.

DEMOSTRACIÓN

1 y 2. Sea $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$. Entonces $\mathcal{B}$ es ortonormal y por tanto linealmente independiente, de acuerdo con la ecuación (8.34), y es un subconjunto de la Col(A) según la ecuación (8.33). Como $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de $\mathbf{R}^n$, el conjunto $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$ genera a Col(A). Por consiguiente, $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ envuelve a Col(A), según la ecuación (8.35). Así, la dimensión de Col(A) es $\leq r$, y entonces es exactamente igual a r, porque $\mathcal{B}$ es un subconjunto linealmente independiente con r elementos. Por lo que $\mathcal{B}$ es una base ortonormal de Col(A) y Rango (A) = r.

3. De acuerdo con la propiedad 2, $\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}$ es una base ortonormal para el complemento ortogonal de Col(A). Ahora, la afirmación es consecuencia de $(\text{Col}(A))^\perp = \nu(A^T)$ del teorema 14, de la sección 8.2.

4. $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ es un subconjunto ortonormal de $\nu(A)$, según la ecuación (8.35). Pero de acuerdo con el teorema del rango, la nulidad de A es n – Rango $(A) = n - r$. De ahí que la dimensión de $\nu(A)$ sea $n - r$, y que $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ sea una base ortonormal.

5. De acuerdo con el punto 4, $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ sea una base ortonormal para el complemento ortogonal de $\nu(A)$. Pero $(\nu(A))^{\perp} = (\text{Ren}(A)^{\perp})^{\perp} = \text{Ren}(A)$, según el teorema 14 de la sección 8.2. La afirmación es consecuencia de lo anterior. ☐

Nota sobre el cálculo numérico del rango

Una de las aplicaciones más importantes de la descomposición en valores singulares (SVD) es el cálculo del rango de una matriz, mediante el teorema 31. La reducción numérica de matrices grandes produce, con frecuencia, un rango equivocado, debido a la acumulación de errores de redondeo. Los elementos que debieran ser cero podrían estar remplazados por números pequeños. Esto se propaga, se repite y se amplifica durante la reducción. De modo que la eliminación de Gauss puede ser insegura para el cálculo del rango. Por otro lado, cuando una matriz se factoriza por descomposición en valores singulares, puede demostrarse que la mayor parte de los errores de redondeo se presentan en el cálculo de Σ. Así, por lo general se descartan valores muy pequeños de σ_i, como si fueran 0, y se cuentan los σ_i restantes para obtener el rango.

Seudoinversa

Si A es cualquier matriz $m \times n$, puede aplicarse la descomposición de A en valores singulares para definir una matriz A^+ $n \times m$ tal que $A^+ = A^{-1}$, para el caso especial en el que A sea invertible (por tanto, $m = n$). La matriz A^+ tiene varias propiedades interesantes, y proporciona una solución óptima en el problema de mínimos cuadrados que estudiamos en la sección 8.4.

DEFINICIÓN

Sea $A = U\Sigma V^T$ una descomposición en valores singulares de una matriz A, $m \times n$. La **seudoinversa**, o **inversa de Moore-Penrose**, de A, es la matriz A^+ $n \times m$ definida por

$$A^+ = V\Sigma^+ U^T \qquad \textbf{(8.36)}$$

en la que Σ^+ es la matriz $n \times m$

$$\Sigma^+ = \begin{bmatrix} D^{-1} & \vdots & \mathbf{0} \\ \cdots & \cdots & \cdots \\ \mathbf{0} & \vdots & \mathbf{0} \end{bmatrix}$$

D es, como antes, la diagonal $r \times r$ cuyos elementos diagonales son los valores singulares positivos $\sigma_1 \geq \cdots \geq \sigma_r > 0$ de A.

■ EJEMPLO 34 Calcule la seudoinversa de

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -6 \end{bmatrix}$$

SOLUCIÓN Es fácil determinar una descomposición de A en valores singulares. Por ejemplo,

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}^T$$

En consecuencia, $\Sigma^+ = \begin{bmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}$, y

$$A^+ = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^T$$

$$= \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & -\frac{1}{6} \end{bmatrix}$$

Observe que si A es $n \times n$ e invertible, entonces $n = r$ y $\Sigma = D$. Por lo que, Σ es $n \times n$ e invertible. Además, $\Sigma\Sigma^+ = I_n$. Por consiguiente,

$$AA^+ = AV\Sigma^+U^T = U\Sigma V^TV\Sigma^+U^T = U\Sigma\Sigma^+U^T = UU^T = I$$

Esto sólo es válido si A es invertible. En este caso, $A^+ = A^{-1}$.

Penrose demostró que A^+ es la *única* matriz B que satisface las **condiciones de Moore-Penrose**:

1. $ABA = A$
2. $BAB = B$
3. $(AB)^T = AB$
4. $(BA)^T = BA$

Al comprobar esas condiciones para el par (A, A^+) del ejemplo 34 podemos aprender mucho. La verificación para cualquier par (A, A^+) se describe en los ejercicios. Aunque no lo demostraremos, *usaremos* la parte de la unicidad de la afirmación de Penrose. Si podemos demostrar que el par (A, B) satisface las condiciones, entonces B es la única seudoinversa de A. Así, $B = A^+$. $\square$

Una aplicación muy importante de la seudoinversa A^+ se da en la solución del problema de los mínimos cuadrados, nuestro tema siguiente.

Descomposición en valores singulares y mínimos cuadrados

Recordemos, de la sección 8.4, que una solución por mínimos cuadrados de un sistema $A\mathbf{x} = \mathbf{b}$ posiblemente inconsistente, es un vector $\tilde{\mathbf{x}}$ que minimiza la longitud del vector de error $\Delta = \mathbf{b} - A\tilde{\mathbf{x}}$,

$$\|\Delta\| = \|\mathbf{b} - A\tilde{\mathbf{x}}\| = \text{mín}$$

El vector $\tilde{\mathbf{x}}$ no es necesariamente único. Si A es $m \times n$ y su rango es $r < n$, su nulidad es mayor que o igual a 1. En este caso, cualquier vector de la forma $\tilde{\mathbf{x}} + \mathbf{z}$ con $\mathbf{z} \neq 0$ en $\nu(A)$ también será una solución por cuadrados mínimos, porque

$$\mathbf{b} - A(\tilde{\mathbf{x}} + \mathbf{z}) = \mathbf{b} - A\tilde{\mathbf{x}} - A\mathbf{z} = \mathbf{b} - A\tilde{\mathbf{x}}$$

Sin embargo, si pedimos que $\tilde{\mathbf{x}}$ también tenga una longitud mínima, entonces esa condición es única, y puede calcularse usando la inversa de Moore-Penrose de A.

TEOREMA 32

El problema de mínimos cuadrados $A\mathbf{x} = \mathbf{b}$ tiene una solución única $\tilde{\mathbf{x}}$ por mínimos cuadrados, de longitud mínima, expresada por

$$\tilde{\mathbf{x}} = A^{+}\mathbf{b}$$

DEMOSTRACIÓN Sea $\mathbf{x}$ un vector n y sea $\mathbf{y} = (y_1, \ldots, y_n)$ igual a $V^T\mathbf{x}$. La matriz U^T es ortogonal, porque U lo es. Por tanto, $\|U^T\mathbf{z}\| = \|\mathbf{z}\|$ para cualquier vector m $\mathbf{z}$. Entonces

$$\|\mathbf{b} - A\mathbf{x}\| = \|\mathbf{b} - U\Sigma V^T\mathbf{x}\| = \|U^T\mathbf{b} - \Sigma V^T\mathbf{x}\| = S_1 + S_2$$

en la que

$$S_1 = (\mathbf{u}_1^T\mathbf{b} - \sigma_1 y_1)^2 + \cdots + (\mathbf{u}_r^T\mathbf{b} - \sigma_r y_r)^2$$
$$S_2 = (\mathbf{u}_{r+1}^T\mathbf{b})^2 + \cdots + (\mathbf{u}_m^T\mathbf{b})^2$$

porque Σ sólo tiene r elementos distintos de cero, ubicados en el bloque superior izquierdo $r \times r$.

Como la suma S_2 es fija, $\|\mathbf{b} - A\mathbf{x}\|$ se minimiza si la suma S_1 es mínima. De hecho, podríamos elegir $\mathbf{x} = (x_1, \ldots, x_n)$ tal que

$$\mathbf{u}_i^T\mathbf{b} = \sigma_i y_i, \qquad i = 1, \ldots, r$$

entonces S_1 seria 0. Por lo que sólo necesitamos un $\mathbf{x}$ de la forma

$$\mathbf{x} = \left(V\frac{\mathbf{u}_1^T\mathbf{b}}{\sigma_1}, \ldots, V\frac{\mathbf{u}_r^T\mathbf{b}}{\sigma_r}, *, \ldots, *\right)$$

Cualquier vector $\mathbf{x}$ como éste sería una solución por mínimos cuadrados, porque minimiza a $\|\mathbf{b} - A\mathbf{x}\|$. Para obtener la magnitud mínima de $\mathbf{x}$, igualaremos a 0 las últimas $n - r$ coordenadas. Por tanto,

$$\tilde{\mathbf{x}} = \left(V\frac{\mathbf{u}_1^T\mathbf{b}}{\sigma_1}, \ldots, V\frac{\mathbf{u}_r^T\mathbf{b}}{\sigma_r}, 0, \ldots, 0\right)$$

es la única solución por mínimos cuadrados con longitud mínima. Además podemos reformular $\tilde{\mathbf{x}}$ como sigue:

$$\tilde{\mathbf{x}} = V\Sigma^{+}U^T\mathbf{b} \Rightarrow \tilde{\mathbf{x}} = A^{+}\mathbf{b}$$

■ EJEMPLO 35 Determine la solución por mínimos cuadrados con longitud mínima de

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -6 \end{bmatrix}\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

SOLUCIÓN Si A es la matriz de coeficientes tenemos, de acuerdo con el teorema 32 y el ejemplo 34,

$$\tilde{\mathbf{x}} = A^+ \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & -\frac{1}{6} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{3} \end{bmatrix}$$

La descomposición polar de una matriz cuadrada

Una consecuencia interesante y útil de la descomposición en valores singulares para una matriz **cuadrada** A es la **descomposición polar** de A.

TEOREMA 33

(Descomposición polar)

Toda matriz cuadrada A puede factorizarse en la forma

$$A = PQ \tag{8.37}$$

en donde P es semidefinida positiva y Q es ortogonal.

DEMOSTRACIÓN Si A es $n \times n$, también lo son U, Σ y V en una descomposición de valores singulares de A. Por consiguiente,

$$A = U\Sigma V^T = U\Sigma(U^TU)V^T = (U\Sigma U^T)UV^T$$

y se iguala $P = U\Sigma U^T$ y $Q = UV^T$. Se deja como ejercicio la comprobación de que P es semidefinida positiva y Q ortogonal.

La descomposición polar es similar a expresar un número complejo z en la forma polar $z = re^{i\theta}$, donde $r \geq 0$ es la magnitud de z, y θ es su argumento, con $|e^{i\theta}| = 1$. P desempeña el papel de r y Q el de $e^{i\theta}$.

■ EJEMPLO 36 Calcule la descomposición polar de

$$A = \begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix}$$

SOLUCIÓN De la descomposición de A en valores singulares,

$$\begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^T$$

definimos

$$P = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$$

y

$$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^T = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Descomposición en valores singulares con sistemas algebraicos computacionales

Maple

```
> evalf(Svd(matrix([[1,2],[3,4]]),U,V));
```

$$[5.464985704, .3659661906]$$

```
> evalf(evalm(U));evalf(evalm(V));
```

$$\begin{bmatrix} -.4045535848 & -.9145142957 \\ -.9145142957 & .4045535848 \end{bmatrix}$$

$$\begin{bmatrix} -.5760484368 & .8174155605 \\ -.8174155605 & -.5760484368 \end{bmatrix}$$

Mathematica

```
In[1]:=
     {U,S,V}=SingularValues[{{1.,2},{3,4}}]
Out[1]=
     {{{-0.404554, -0.914514},
        {-0.914514, 0.404554}},
       {5.46499, 0.365966},
       {{-0.576048, -0.817416},
        {0.817416, -0.576048}}}
```

MATLAB

```
>> [U,S,V]=svd([1 2; 3 4])
U =
    0.4046    0.9145
    0.9145   -0.4046
S =
    5.4650         0
         0    0.3660
V =
    0.5760   -0.8174
    0.8174    0.5760
```

Ejercicios 8.7

En los ejercicios 1 a 3 determine los valores singulares de la matriz correspondiente.

1. $\begin{bmatrix} 0 & 0 \\ 0 & -2 \\ 3 & 0 \end{bmatrix}$

2. $\begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

En los ejercicios 4 a 11 determine una descomposición en valores singulares para la matriz correspondiente.

4. $\begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 0 & 5 \end{bmatrix}$

5. $\begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

6. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}$

7. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

9. $\begin{bmatrix} 2 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 8 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 6 & -4 \\ -2 & 6 & 2 \\ 2 & 3 & 4 \end{bmatrix}$

11. $\begin{bmatrix} 2 & 6 & -4 \\ -4 & 6 & 2 \\ 4 & 3 & 4 \end{bmatrix}$

En los ejercicios 12 y 13 encuentre una descomposición en valores singulares trabajando con la transpuesta de la matriz.

12. $\begin{bmatrix} 2 & -4 & 4 \\ 6 & 6 & 3 \\ -4 & 2 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 2 & 1 & -2 \\ 0 & 0 & 0 \\ 6 & -6 & 3 \end{bmatrix}$

En los ejercicios 14 y 15 calcule las seudoinversas y verifique las propiedades de Moore-Penrose.

14. a. $\begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 0 & 5 \end{bmatrix}^+$ b. $\begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}^+$

15. a. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}^+$ b. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{bmatrix}^+$

En los ejercicios 16 y 17 calcule y compare A^+ y A^{-1}.

16. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}$

17. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{bmatrix}$

18. Si A es cualquier matriz, demuestre que el par (A, A^+) satisface las condiciones de Moore-Penrose. (*Sugerencia:* Compruebe primero las condiciones para (Σ, Σ^+).)

En los ejercicios 19 y 20 demuestre las identidades por medio de las condiciones de Moore-Penrose.

19. $\begin{bmatrix} -3 & 0 \\ 0 & 0 \\ 0 & 4 \end{bmatrix}^+ = \begin{bmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$

20. $\begin{bmatrix} -2 & 6 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}^+ = \begin{bmatrix} -\frac{2}{9} & \frac{1}{9} & \frac{2}{9} \\ \frac{2}{27} & \frac{2}{27} & \frac{1}{27} \end{bmatrix}$

21. Demuestre que $A^{++} = A$. (*Sugerencia:* Compruebe las condiciones de Moore-Penrose para (A^+, A).)

22. Compruebe que $(A^T)^+ = (A^+)^T$. Llegue a la conclusión de que si A es simétrica, también lo es A^+. (*Sugerencia:* Verifique las condiciones de Moore-Penrose para $(A^T, (A^+)^T)$.)

En los ejercicios 23 a 25 resuelva el problema de mínimos cuadrados para $A\mathbf{x} = \mathbf{b}$ usando A^+.

23. $A = \begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 0 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

24. $A = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

25. $A = \begin{bmatrix} -2 & 6 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

En los ejercicios 26 y 27 calcule la descomposición polar de cada una de las matrices.

26. a. $\begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ b. $\begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}$

27. a. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 6 & -4 \\ -2 & 6 & 2 \\ 2 & 3 & 4 \end{bmatrix}$

8.8 Productos internos

Objetivos del estudiante para esta sección

1. Conocer la definición y las propiedades básicas de los productos interiores.
2. Encontrar más propiedades de los productos internos, análogas a las de los productos punto.

En esta sección describiremos una generalización muy útil del producto punto para espacios vectoriales. Aunque el producto punto en $\mathbf{R}^n$ y sus propiedades están en el núcleo de muchos resultados teóricos y aplicados, a veces puede ser muy restrictivo, sobre todo cuando se trabaja directamente con vectores-n. Esto se evidencia al manejar espacios vectoriales de polinomios o de funciones, porque esos conjuntos tienen sus propias notaciones naturales.

Introduciremos un "producto punto" llamado *producto interno* que funciona con vectores generales. Para definirlo usaremos las propiedades básicas del producto punto descritas por el teorema 4, de la sección 2.2. Como las propiedades restantes del producto punto fueron demostradas por medio de este teorema, las del producto interno tienen demostraciones *idénticas* a las análogas para el producto punto. Así, el material de esta sección parecerá bastante familiar.

DEFINICIÓN

Un **producto interno** en un espacio vectorial (real) V es una función que asocia a cada par de vectores $\mathbf{u}$ y $\mathbf{v}$, un número real $\langle \mathbf{u}, \mathbf{v} \rangle$ que satisface las propiedades o **axiomas** siguientes.

Para todo vector $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ cualesquiera, y cualquier escalar c,

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ **(Axioma de simetría)**
2. $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$ **(Axioma de aditividad)**
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ **(Axioma de homogeneidad)**
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$ **(Axioma de positividad)**

Todo espacio vectorial real con un producto interno se llama **espacio (de) producto interno**.

Para comprobar que un espacio vectorial es un espacio producto interno, debemos tener primero una función que asocie un número a cada par de vectores. A continuación debemos verificar los cuatro axiomas para esa función. Por lo general, la igualdad $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Rightarrow \mathbf{u} = \mathbf{0}$ del axioma de positividad es la más difícil de comprobar.

Los axiomas de un producto interno implican las propiedades básicas adicionales siguientes:

TEOREMA 34

Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ cualesquiera vectores en un espacio producto interno, y c cualquier escalar.

1. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
2. $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
3. $\langle \mathbf{u} - \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{w}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{w} \rangle$
5. $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

DEMOSTRACIÓN Comprobaremos la primera propiedad y dejaremos las restantes como ejercicios.

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle &= \langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle && \text{por la simetría} \\
&= \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle && \text{por la aditividad} \\
&= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle && \text{por la simetría}
\end{aligned}$$

$\square$

■ EJEMPLO 37 Si $\mathbf{u} = (u_1, \ldots, u_n)$ y $\mathbf{v} = (v_1 \ldots, v_n)$, son cualesquiera vectores n, demuestre que el producto punto de $\mathbf{R}^n$,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = u_1 v_1 + \cdots + u_n v_n$$

es un producto interno.

SOLUCIÓN Todos los axiomas son válidos, de acuerdo con el teorema 4 de la sección 2.2.

■ EJEMPLO 38 Si $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$ son cualesquiera vectores 2, compruebe que

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1 v_1 + 4u_2 v_2$$

define a un producto interno en $\mathbf{R}^2$.

SOLUCIÓN

Simetría: La simetría es válida, porque

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1 v_1 + 4u_2 v_2 = 3v_1 u_1 + 4v_2 u_2 = \langle \mathbf{v}, \mathbf{u} \rangle$$

Aditividad: Si $\mathbf{w} = (w_1, w_2)$, entonces

$$\begin{aligned} \langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle &= 3(u_1 + w_1)v_1 + 4(u_2 + w_2)v_2 \\ &= (3u_1 v_1 + 4u_2 v_2) + (3w_1 v_1 + 4w_2 v_2) \\ &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle \end{aligned}$$

Homogeneidad: Para cualquier escalar c,

$$\begin{aligned} \langle c\mathbf{u}, \mathbf{v} \rangle &= 3(cu_1)v_1 + 4(cu_2)v_2 \\ &= c(3u_1 v_1 + 4u_2 v_2) \\ &= c\langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

Positividad:

$$\langle \mathbf{u}, \mathbf{u} \rangle = 3u_1 u_1 + 4u_2 u_2 = 3u_1^2 + 4u_2^2 \geq 0$$

que también implica que

$$\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Leftrightarrow (u_1 = 0 \quad \text{y} \quad u_2 = 0) \Leftrightarrow \mathbf{u} = \mathbf{0}$$

Por consiguiente, todos los axiomas son válidos y $\langle , \rangle$ define a un producto interno.

Acabamos de encontrar un producto interno de $\mathbf{R}^2$ que no es el producto punto. Así, un espacio vectorial puede tener *varios productos internos distintos*.

El ejemplo 38 es un caso especial del ejemplo 39.

■ EJEMPLO 39 (Producto punto ponderado) Si $w_1, \ldots, w_n$ son cualesquiera números *positivos*, y si $\mathbf{u} = (u_1, \ldots, u_n)$ y $\mathbf{v} = (v_1, \ldots, v_n)$ son cualesquiera vectores n. Demuestre que

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + \cdots + w_n u_n v_n \tag{8.38}$$

define un producto interno en $\mathbf{R}^n$.

SOLUCIÓN Todos los axiomas se comprueban justamente como en el ejemplo 38.

El producto interno del ejemplo 39 se llama **producto punto ponderado** de $\mathbf{R}^n$ **con pesos** $w_1, \ldots, w_n$. Es importante que todos los pesos (o *factores de ponderación*) $w_1, \ldots, w_n$ sean positivos. Si no es así, el axioma de positividad no se aplicará. La fórmula (8.38) también puede escribirse en notación matricial como sigue:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T W \mathbf{v}, \qquad \text{donde } W = \begin{bmatrix} w_1 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix}$$

■ EJEMPLO 40 Sean

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$$

matrices 2×2 con elementos reales (es decir, los elementos de M_{22}). Es fácil demostrar que

$$\langle A, B \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3 + a_4 b_4$$

define a un producto interno en M_{22}.

Por ejemplo, si $A = \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & -2 \\ -2 & 1 \end{bmatrix}$, entonces

$$\langle A, B \rangle = 1 \cdot 0 + 2 \cdot (-2) + (-3) \cdot (-2) + 4 \cdot 1 = 6$$

■ EJEMPLO 41 Sean $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ y $q(x) = b_0 + b_1 x + \cdots + b_n x^n$ polinomios en P_n. Es fácil comprobar que

$$\langle p, q \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n$$

define a un producto interno en P_n.

Por ejemplo, si $p(x) = 1 - x^2$ y $q(x) = -2x + x^2$, entonces

$$\langle p, q \rangle = 1 \cdot 0 + 0 \cdot (-2) + (-1) \cdot 1 = -1$$

■ EJEMPLO 42 Sean $r_0, r_1, \ldots, r_n$ $n + 1$ números reales distintos. y $p(x)$ y $q(x)$ cualesquiera polinomios en P_n. Demuestre que

$$\langle p, q \rangle = p(r_0)q(r_0) + \cdots + p(r_n)q(r_n)$$

define a un producto interno en P_n.

SOLUCIÓN Los axiomas 1 a 3 se comprueban con facilidad. Para el axioma de positividad

$$\langle p, p\rangle = p(r_0)^2 + \cdots + p(r_n)^2 \geq 0$$

y

$$\langle p, p\rangle = 0 \Leftrightarrow p(r_0) = 0, \ldots, p(r_n) = 0 \Leftrightarrow p = 0$$

porque el polinomio p tiene un grado máximo igual a n, de modo que si tiene más de n raíces debe ser el polinomio cero.

Por ejemplo, sean $r_0 = -2$, $r_1 = 0$, $r_2 = 1$, $p(x) = 1 - x^2$, y $q(x) = -2x + x^2$. Ya que

$$p(-2) = -3, \quad q(-2) = 8, \quad p(0) = 1, \quad q(0) = 0, \quad p(1) = 0, \quad q(1) = -1$$

entonces

$$\langle p, q\rangle = p(-2)q(-2) + p(0)q(0) + p(1)q(1) = -24$$

Nuestro ejemplo siguiente generaliza el producto punto ponderado, y es una fuente importante de productos internos.

■ EJEMPLO 43 Si A es cualquier matriz de $n \times n$ **definida positiva**, y por tanto simétrica, demuestre que para **u** y **v** cualesquiera vectores n,

$$\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^T A\mathbf{v}$$

define un producto interno de $\mathbf{R}^n$.

SOLUCIÓN Necesitamos verificar los cuatro axiomas.

Simetría: Como $A = A^T$, tenemos que

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v}\rangle &= \mathbf{u}^T A\mathbf{v} = \mathbf{u} \cdot A\mathbf{v} = A^T\mathbf{u} \cdot \mathbf{v} \\ &= A\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot A\mathbf{u} = \mathbf{v}^T A\mathbf{u} = \langle \mathbf{v}, \mathbf{u}\rangle\end{aligned}$$

Aditividad:

$$\begin{aligned}\langle \mathbf{u} + \mathbf{w}, \mathbf{v}\rangle &= (\mathbf{u} + \mathbf{w})^T A\mathbf{v} \\ &= \mathbf{u}^T A\mathbf{v} + \mathbf{w}^T A\mathbf{v} = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle\end{aligned}$$

Homogeneidad:

$$\langle c\mathbf{u}, \mathbf{v}\rangle = (c\mathbf{u})^T A\mathbf{v} = c\mathbf{u}^T A\mathbf{v} = c\langle \mathbf{u}, \mathbf{v}\rangle$$

Positividad: Como A es definida positiva, también lo es la forma cuadrática $q(\mathbf{u}) = \mathbf{u}^T A\mathbf{u}$. Por consiguiente,

$$\langle \mathbf{u}, \mathbf{u}\rangle = \mathbf{u}^T A\mathbf{u} > 0 \quad \text{para todo } \mathbf{u} \neq \mathbf{0}$$

Con esto se comprueba el último axioma.

■ EJEMPLO 44 Demuestre que

$$\langle \mathbf{u}, \mathbf{v} \rangle = 6u_1v_1 - 2u_2v_1 - 2u_1v_2 + 3u_2v_2$$

define a un producto interno en $\mathbf{R}^2$.

SOLUCIÓN La notación $\langle \mathbf{u}, \mathbf{v} \rangle$ puede escribirse en la forma $\mathbf{u}^T A \mathbf{v}$:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Como $\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix}$ tiene los eigenvalores positivos (2 y 7), es definida positiva de acuerdo con el teorema 29 de la sección 8.6. Por lo que $\langle \mathbf{u}, \mathbf{v} \rangle$ se define como producto interno de $\mathbf{R}^2$, de acuerdo con el ejemplo 43.

Observe que si A no es definida positiva, podría ser que $\mathbf{u}^T A \mathbf{v}$ no definiera un producto interno. Por ejemplo,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 2u_1v_1 - 2u_2v_1 - 2u_1v_2 + 2u_2v_2$$

no es un producto interno, porque

$$\langle (1, 1), (1, 1) \rangle = 0$$

$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$ tiene los eigenvalores 0 y 4, de modo que no es definida positiva; es semidefinida positiva. $\square$

■ EJEMPLO 45 (Se requiere cálculo) Si $f(x)$ y $g(x)$ están en $C[a, b]$, el espacio vectorial de las funciones continuas de valor real definidas en $[a, b]$; entonces

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx$$

define un producto interno en $C[a, b]$.

SOLUCIÓN

Simetría: Tenemos que

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx = \int_a^b g(x)f(x)\, dx = \langle g, f \rangle$$

Aditividad:

$$\begin{aligned} \langle f + h, g \rangle &= \int_a^b (f(x) + h(x))g(x)\, dx \\ &= \int_a^b (f(x)g(x) + h(x)g(x))\, dx \\ &= \int_a^b f(x)g(x) + \int_a^b h(x)g(x)\, dx \\ &= \langle f, g \rangle + \langle h, g \rangle \end{aligned}$$

Homogeneidad:

$$\langle cf, g\rangle = \int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx = c\langle f, g\rangle$$

Positividad: Para toda función $f(x)$ de $C[a, b]$, $f(x)^2 \geq 0$. Por tanto,

$$\langle f, f\rangle = \int_a^b f(x)^2\,dx \geq 0$$

Cuando $g(x) = f(x)^2$, g es continua y no negativa, de modo que según un teorema del cálculo,

$$\int_a^b g(x)\,dx = 0 \Leftrightarrow g = \mathbf{0}$$

($\mathbf{0}$ es la función cero, es decir, $\mathbf{0}(x) = 0$ para toda $x \in [a, b]$.) Así,

$$\langle f, f\rangle = \int_a^b f(x)^2\,dx = 0 \Leftrightarrow f = \mathbf{0}$$

Con esto finaliza la demostración del axioma de positividad. □

Longitud y ortogonalidad

En un espacio de producto interno pueden definirse longitudes, distancias y vectores ortogonales mediante fórmulas idénticas a las del producto punto.

DEFINICIÓN

Sea V un espacio de producto interno. Dos vectores **u** y **v** se llaman **ortogonales** si su producto interno es cero.

$$\mathbf{u} \text{ y } \mathbf{v} \text{ son ortogonales si } \langle \mathbf{u}, \mathbf{v}\rangle = 0$$

La **norma** (o **longitud**, o **magnitud**) de **v** es el número no negativo $\|\mathbf{v}\|$ definido por:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle} \tag{8.39}$$

Se define la raíz cuadrada positiva, porque $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$, según el axioma de positividad. De igual forma,

$$\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle \tag{8.40}$$

También se define la **distancia** entre dos vectores, **u** y **v**, por medio de

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| \tag{8.41}$$

Observe que

$$d(\mathbf{0}, \mathbf{v}) = d(\mathbf{v}, \mathbf{0}) = \|\mathbf{v}\|$$

Un vector cuya norma es 1 se llama vector **unitario**. El conjunto S de todos los vectores unitarios de V se denomina **círculo unitario** o **esfera unitaria**.

$$S = \{\mathbf{v}, \mathbf{v} \in V \quad \text{y} \quad \|\mathbf{v}\| = 1\} \tag{8.42}$$

> La esfera unitaria está formada por todos los vectores de V a la distancia 1 del origen. Es la forma como se definen el círculo unitario en $\mathbf{R}^2$ y la esfera unitaria en $\mathbf{R}^3$, con respecto a la norma ordinaria (producto punto), justificando así esos nombres. Sin embargo, es preciso notar que un círculo unitario en $\mathbf{R}^2$ puede **no** tener la gráfica de un círculo en el sistema de coordenadas cartesianas.

■ EJEMPLO 46 Para el producto interno $\langle \mathbf{u}, \mathbf{v}\rangle = 3u_1v_1 + 4u_2v_2$ del ejemplo 38, haga lo siguiente:

(a) calcule $\|(-2, 1)\|$.
(b) deduzca $d(\mathbf{e}_1, \mathbf{e}_2)$.
(c) demueste que (4, 3) y (1, −1) son ortogonales.
(d) Describe, y trace una gráfica del círculo unitario.

SOLUCIÓN

(a) Desarrollando,

$$\|(-2,1)\|^2 = \langle(-2,1),(-2,1)\rangle = 3(-2)(-2) + 4 \cdot 1 \cdot 1 = 16$$

Tenemos que,

$$\|(-2,1)\| = \sqrt{16} = 4$$

(b) Como

$$\|(1,0) - (0,1)\|^2 = \|(1,-1)\|^2 = 3 \cdot 1 \cdot 1 + 4 \cdot (-1) \cdot (-1) = 7$$

entonces

$$d(\mathbf{e}_1, \mathbf{e}_2) = \sqrt{7}$$

(c) Los vectores (4, 3) y (1, −1) son ortogonales con respecto a este producto interno (pero no con respecto al producto punto) porque

$$\langle(4,3),(1,-1)\rangle = 3 \cdot 4 \cdot 1 + 4 \cdot 3 \cdot (-1) = 0$$

(d) En vista de que $\|(v_1, v_2)\| = 1$ equivale a $\|(v_1, v_2\|^2 = 1$,

$$\mathbf{S} = \{(v_1, v_2), \quad \text{tal que} \quad 3v_1^2 + 4v_2^2 = 1\} \subseteq \mathbf{R}^2$$

Así, la esfera unitaria (círculo) con respecto a este producto interno se parece a una **elipse** en el sistema de coordenadas que contiene el producto punto, los ángulos y las distancias ordinarios (figura 8.18). □

Desigualdades e igualdades básicas

Los axiomas pueden combinarse con la identidad del siguiente teorema, con el fin de generalizar las identidades familiares de la sección 2.2.

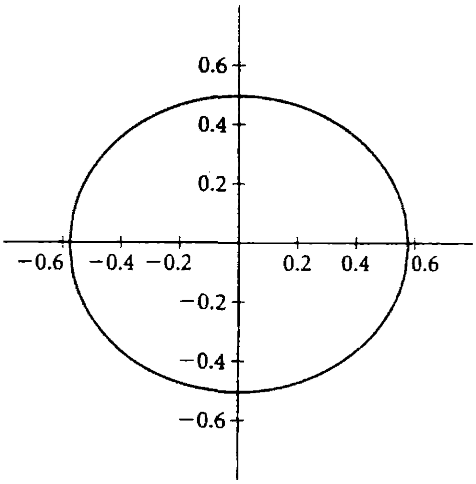


Figura 8.18 El círculo unitario para el producto interno $3u_1v_1 + 4u_2v_2$.

TEOREMA 35

Sea V un espacio de producto interno. Para cualesquiera vectores $\mathbf{u}$ y $\mathbf{v}$ de V,

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle \tag{8.43}$$

DEMOSTRACIÓN Tenemos lo siguiente. (Explique cuáles axiomas se usan.)

$$\begin{aligned}
\|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle \\
&= \langle \mathbf{u}, \mathbf{u} + \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle \\
&= \langle \mathbf{u}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\
&= \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle
\end{aligned}$$

□

Sustituyendo $\mathbf{v}$ con $-\mathbf{v}$ en la ecuación (8.43) se obtiene

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle \tag{8.44}$$

A continuación presentaremos una generalización de la ley del paralelogramo (véase el ejercicio 22 de la sección 2.2).

TEOREMA 36

(Ley del paralelogramo)

Sea V un espacio de producto interno. Para cualesquiera vectores $\mathbf{u}$ y $\mathbf{v}$ de V, tenemos

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

DEMOSTRACIÓN Sume las ecuaciones (8.43) y (8.44). □

La identidad del teorema siguiente expresa el producto interno en función de la norma (véase el ejercicio 23 de la sección 2.2).

TEOREMA 37

> **(Identidad de polarización)**
>
> Sea V un espacio de producto interno. Para cualesquiera vectores $\mathbf{u}$ y $\mathbf{v}$ de V se cumple
>
> $$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$$

DEMOSTRACIÓN Se resta la ecuación (8.44) de la (8.43) y se despeja $\langle \mathbf{u}, \mathbf{v} \rangle$. $\square$

También existe una generalización del teorema de Pitágoras, de la sección 2.2.

TEOREMA 38

> **(Teorema de Pitágoras)**
>
> Sea V un espacio de producto interno. Los vectores $\mathbf{u}$ y $\mathbf{v}$ de V son ortogonales si y sólo si
>
> $$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

DEMOSTRACIÓN Se deja como ejercicio. $\square$

La desigualdad de Cauchy-Schwarz-Bunyakovsky (desigualdad de CSB)

Una de las consecuencias más útiles de los axiomas es una generalización de la desigualdad de Cauchy-Schwarz-Bunyakovsky, que se llama desigualdad de CSB.

TEOREMA 39

> **(Desigualdad de Cauchy-Schwarz-Bunyakovsky)**
>
> $$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$$
>
> Además, la igualdad es válida si y sólo si $\mathbf{u}$ y $\mathbf{v}$ son múltiplos por escalar entre sí.

DEMOSTRACIÓN De acuerdo con el teorema 34,

$$0 \leq \langle x\mathbf{u} + \mathbf{v}, x\mathbf{u} + \mathbf{v} \rangle = x^2 \langle \mathbf{u}, \mathbf{u} \rangle + 2x\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \tag{8.45}$$

para todos los escalares x. Éste es un polinomio cuadrático $p(x) = ax^2 + bx + c$ con $a = \langle \mathbf{u}, \mathbf{u} \rangle$, $b = 2\langle \mathbf{u}, \mathbf{v} \rangle$ y $c = \langle \mathbf{v}, \mathbf{v} \rangle$. Como $a \geq 0$ y $p(x) \geq 0$ para toda x, la gráfica de $p(x)$ es una parábola en el semiplano superior que se abre hacia arriba. En consecuencia, la parábola está arriba del eje x, en cuyo caso $p(x)$ tiene dos raíces complejas. o bien es tangente al eje x, y en ese caso posee una raíz real repetida. Por consiguiente, $b^2 - 2ac \leq 0$. Así,

$$\left(2\langle \mathbf{u}, \mathbf{v} \rangle\right)^2 - 4\langle \mathbf{u}, \mathbf{u} \rangle\langle \mathbf{v}, \mathbf{v} \rangle \leq 0 \qquad \text{o sea} \qquad 4\langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\|\mathbf{u}\|^2\|\mathbf{v}\|^2 \leq 0$$

que implica la desigualdad CSB.

La igualdad es válida si y sólo si $b^2 - 2ac = 0$, o bien si y sólo si $p(x)$ tiene una raíz real doble, digamos r. Por tanto, según la ecuación (8.45) con $x = r$,

$$\begin{aligned} & \langle r\mathbf{u} + \mathbf{v}, r\mathbf{u} + \mathbf{v} \rangle = 0 \\ \Leftrightarrow \quad & \|r\mathbf{u} + \mathbf{v}\| = 0 \\ \Leftrightarrow \quad & r\mathbf{u} + \mathbf{v} = \mathbf{0} \\ \Leftrightarrow \quad & \mathbf{v} = -r\mathbf{u} \end{aligned}$$

De manera que $\mathbf{v}$ es el producto de $\mathbf{u}$ por un escalar. Con esto se demuestra la última afirmación del teorema. □

■ EJEMPLO 47 Compruebe la desigualdad CSB para el ejemplo 42, con $r_0 = -2$, $r_1 = 0$, $r_2 = 1$, $p(x) = 1 - x^2$ y $q(x) = -2x + x^2$.

SOLUCIÓN Para los datos anteriores,

$$\begin{aligned} \langle p, q \rangle &= p(-2)q(-2) + p(0)q(0) + p(1)q(1) = -24 \\ \langle p, p \rangle &= p(-2)^2 + p(0)^2 + p(1)^2 = 10 \\ \langle q, q \rangle &= q(-2)^2 + q(0)^2 + q(1)^2 = 65 \end{aligned}$$

Por lo que,

$$\|p\| = \sqrt{10} \qquad \|q\| = \sqrt{65}$$

y

$$|\langle p, q \rangle| = |-24| = 24 \leq \sqrt{10} \cdot \sqrt{65} \cong 25.495$$

comprueba la desigualdad de Cauchy-Schwarz-Bunyakovsky. □

Sobre la historia de la desigualdad CSB

Se acredita a Cauchy[7] la desigualdad para vectores y a Schwarz[8] para los productos internos integrales, como en el ejemplo 45. Sin embargo, fue Bunyakovsky[9] quien demostró y publicó la desigualdad de Schwarz en una monografía, 25 años antes que Schwarz.

---

[7] **Augustin Louis Cauchy** (1789-1857) nació en París, Francia, y murió en una villa cercana a esa misma ciudad. Es autor de trabajos importantes sobre ecuaciones diferenciales, series infinitas, determinantes, probabilidad, grupos de permutación y de física matemática. En 1814 publicó una memoria que se convirtió en el fundamento de la teoría de las funciones complejas. Su trabajo se conoce por su rigor. Publicó 789 artículos y ocupó puestos en la Facultad de Ciencias, en el Colegio de Francia y en la Escuela Politécnica, todos en París. Hay muchos términos y teoremas que tienen su apellido. Fue un realista fiel y vivió en Suiza, Turín y Praga, después de rehusarse a jurar la alianza. Regresó a París en 1838 y recuperó su puesto en la Academia de Ciencias. En 1848 recuperó su lugar en la Sorbona, que mantuvo hasta su muerte.

[8] **Karl Herman Amandus Schwarz** (1843-1921) nació en Hermsdorf, Polonia (hoy Alemania), y murió en Berlín. Estudió química en Berlín, pero se cambió a matemáticas, obteniendo el doctorado. Desempeñó puestos académicos en Halle, Zurich y Göttingen. Reemplazó a Weierstrass en Berlín y enseñó allí hasta 1917. Trabajó en cálculo de variaciones y en superficies mínimas. Su memoria en ocasión del 70 aniversario de Weierstrass contiene, entre otros temas importantes, la desigualdad para integrales que hoy se conoce como desigualdad de Schwarz.

Una aplicación muy útil del teorema 35 y de la desigualdad CSB es la *desigualdad del triángulo*.

TEOREMA 40

(La desigualdad del triángulo)

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

DEMOSTRACIÓN

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle && \text{Según el teorema 35} \\ &\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| && \text{Según la desigualdad CSB} \\ &= \left(\|\mathbf{u}\| + \|\mathbf{v}\|\right)^2 \end{aligned}$$

Por consiguiente, $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$. $\square$

El proceso de Gram-Schmidt

El proceso de Gram-Schmidt puede ampliarse fácilmente a los productos internos en general, estableciendo con ello la existencia de bases ortogonales para espacios de productos internos con dimensiones finitas. Las fórmulas son las mismas que antes, pero se reemplaza el producto punto con un producto interno general.

■ EJEMPLO 48 (Gram-Schmidt generalizado) Determine una base ortogonal de P_2, comenzando con $1, x, x^2$ y usando el producto interno del ejemplo 42, con $r_0 = 0$, $r_1 = 1$ y $r_2 = 2$.

SOLUCIÓN Sea $p_1 = 1$. Puesto que

$$\langle 1, 1\rangle = 1^2 + 1^2 + 1^2 = 3, \qquad \langle x, 1\rangle = 0 \cdot 1 + 1 \cdot 1 + 2 \cdot 1 = 3$$

se define a

$$p_2 = x - \frac{\langle x, 1\rangle}{\langle 1, 1\rangle} 1 = x - 1$$

De igual manera,

$$\langle x - 1, x - 1\rangle = 2, \qquad \langle x^2, 1\rangle = 5, \qquad \langle x^2, x - 1\rangle = 4$$

Tenemos que

$$\begin{aligned} p_3 &= x^2 - \frac{\langle x^2, 1\rangle}{\langle 1, 1\rangle} 1 - \frac{\langle x^2, x - 1\rangle}{\langle x - 1, x - 1\rangle}(x - 1) \\ &= x^2 - 2x + \frac{1}{3} \end{aligned}$$

[9] **Viktor Yakovlevich Bunyakovsky** (1804-1889) nació en Bar, Ucrania, y murió en San Petersburgo, Rusia. Fue profesor en San Petersburgo de 1846 a 1880. Publicó más de 150 trabajos en matemáticas y mecánica. Fue autor de trabajos importantes en teoría de los números, y demostró y publicó la desigualdad de Schwarz en 1859, 25 años antes que Schwarz. También trabajó en geometría y en hidrostática.

Por consiguiente,

$$\{p_1, p_2, p_3\} = \left\{1, x-1, x^2 - 2x + \frac{1}{3}\right\}$$

es una base ortogonal de P_2 con respecto al producto interno dado.

El teorema de la mejor aproximación (sección 8.2) también se generaliza con facilidad a los espacios de productos internos. Dejaremos que el lector elabore los detalles. Esta generalización tiene particular utilidad cuando se trata de aproximar una función usando otras funciones. El tipo de aproximación depende del producto interno que usemos.

■ EJEMPLO 49 **(Mejor aproximación generalizada)** A propósito del ejemplo 48 y de su solución, determine un polinomio $\tilde{p}$ en $P_1 = \text{Gen}\{1, x\} \subseteq P_2$ que se aproxime mejor a $p(x) = 2x^2 - 1$.

SOLUCIÓN De acuerdo con la solución del ejemplo 48, $\{p_0, p_1\} = \{1, x-1\}$ es una base ortogonal de P_1. Ya que

$$\langle 2x^2 - 1, 1\rangle = (-1)\cdot 1 + 1\cdot 1 + 7\cdot 1 = 7$$

$$\langle 2x^2 - 1, x-1\rangle = (-1)\cdot(-1) + 1\cdot 0 + 7\cdot 1 = 8$$

entonces

$$\begin{aligned}\tilde{p} = p_{\text{pr}} &= \frac{\langle p, p_0\rangle}{\langle p_0, p_0\rangle}p_0 + \frac{\langle p, p_1\rangle}{\langle p_1, p_1\rangle}p_1 \\ &= \frac{7}{3} + \frac{8}{2}(x-1) = 4x - \frac{5}{3}\end{aligned}$$

De ahí que $4x - \frac{5}{3}$ de P_1 se aproxima mejor a $2x^2 - 1$ con respecto al producto interno dado (figura 8.19).

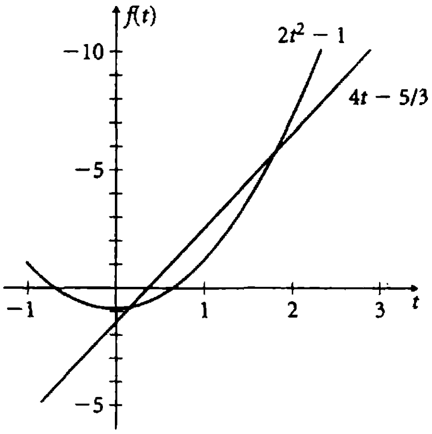


Figura 8.19 $4t - \frac{5}{3}$ es el polinomio que mejor se aproxima a $2t^2 - 1$ cuando se mide la distancia en 0, 1 y 2.

Ejercicios 8.8

Sean $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$. En los ejercicios 1 y 2:

(a) Encuentre $\langle \mathbf{u}, \mathbf{v} \rangle$, $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ y $\|\mathbf{u} + \mathbf{v}\|$.
(b) Calcule la distancia $d(\mathbf{u}, \mathbf{v})$.
(c) Verifique la desigualdad CSB para **u** y **v**.
(d) Demuestre la desigualdad del triángulo para **u** y **v**.
(e) Compruebe la identidad de polarización para **u** y **v**.

1. El producto interno es el del ejemplo 38.
2. El producto interno es el del ejemplo 39, donde $W = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$.
3. En el ejercicio 1, determine un vector ortogonal a **u**.
4. Para el ejercicio 2, determine un vector ortogonal a **u**.

Considere el producto interno del ejemplo 40, si
$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

5. Determine los pares ortogonales entre A, B y C.
6. Para los pares ortogonales del ejercicio 5, verifique el teorema de Pitágoras.
7. Forme pares ortonormales a partir de los ortogonales determinados en el ejercicio 5.
8. ¿Cuál(es) de entre A, B y C son "vectores" unitarios?

En los ejercicios 9 a 15 determine, para la matriz dada A, la función

$$f(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T A \mathbf{v}, \qquad \mathbf{u}, \mathbf{v} \in \mathbf{R}^2$$

defina un producto interno de $\mathbf{R}^2$ como sigue: vea el ejemplo 43 y compruebe si A es definida positiva, y por tanto simétrica. Si no lo es, elabore un axioma de la definición que no se aplica del producto interno.

9. $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$
10. $A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$
11. $A = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$
12. $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$
13. $A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$
14. $A = \begin{bmatrix} -1 & 2 \\ 2 & 2 \end{bmatrix}$
15. $A = \begin{bmatrix} 8 & 7 \\ 1 & 2 \end{bmatrix}$
16. Use la matriz definida positiva 3×3 de su elección para definir su propio producto interno en $\mathbf{R}^3$.
17. Si A es matriz definida positiva (por tanto, simétrica) $n \times n$, demuestre que

$$f(\mathbf{u}, \mathbf{v}) = \mathbf{v}^T A \mathbf{u}, \qquad \mathbf{u}, \mathbf{v} \in \mathbf{R}^n$$

defina un producto interno en $\mathbf{R}^n$.

En los ejercicios 18 a 20, sean $p(x) = 1 - 2x^2$, $q(x) = -2x + x^2$ en P_2.

(a) Determine $\langle p, q \rangle$, $\|p\|$, $\|q\|$ y $\|p + q\|$.
(b) Calcule la distancia $d(q, p)$.
(c) Compruebe la desigualdad CSB para p y q.
(d) Verifique la desigualdad del triángulo para p y q.

18. El producto interno en P_2 es el del ejemplo 41.
19. El producto interno en P_2 es el del ejemplo 42, con $r_0 = -3$, $r_1 = 0$ y $r_2 = 2$.
20. **(Se requiere cálculo)** El producto interno es el del ejemplo 45, con $a = -1$ y $b = 1$.
21. Acerca del ejercicio 18, determine un vector ortogonal a p.
22. Con respecto al ejercicio 19, calcule un vector ortogonal a p.
23. A partir del ejercicio 20, determine un vector ortogonal a p.
24. Suponga que $\mathbf{R}^2$ contiene el producto interno del ejemplo 44. ¿Es ortogonal la base estándar $\{\mathbf{e}_1, \mathbf{e}_2\}$?
25. Termine la demostración del teorema 34.
26. Finalice la demostración del teorema 38.
27. Suponga que $T : \mathbf{R}^n \to \mathbf{R}^n$ es una transformación lineal invertible. Compruebe que la asignación

$$\langle \mathbf{u}, \mathbf{v} \rangle = T(\mathbf{u}) \cdot T(\mathbf{v}), \qquad \mathbf{u}, \mathbf{v} \in \mathbf{R}^n$$

define a un producto interno en $\mathbf{R}^n$.

Proyecciones

NOTA

1. Calcularemos las proyecciones ortogonales con respecto a los productos internos usando la misma forma que antes, cuando el producto punto se reemplaza con el producto interno dado.
2. La distancia de un vector a **u**, un subespacio W en un espacio de producto interno, es la norma $\|\mathbf{u}_c\|$ del componente $\mathbf{u}_c$ ortogonal de **u** con respecto a W.

28. Considere que $\mathbf{R}^2$ contiene el producto interno del ejemplo 44. Determine la proyección ortogonal de (1, 1) sobre la recta l que pasa por $\mathbf{0}$ y por $\mathbf{e}_1$.

29. Retome el ejercicio 28 y calcule la distancia de (1, 1) a l.

30. Para el ejemplo 48, determine la proyección ortogonal (y por tanto la mejor aproximación) de $x^2 + x + 1$, con respecto al subespacio $W = \text{Gen}\{p, q\}$ de P_2, donde

$$p = x - 1, \qquad q = x^2 - 2x + \frac{1}{3}$$

(Observe que $\{p, q\}$ es una base ortogonal de W.)

31. En el ejercicio 30, calcule la distancia de $x^2 + x + 1$ a W.

Gram-Schmidt

32. Suponga que $\mathbf{R}^2$ contiene el producto interno del ejemplo 44. Aplique el proceso de Gram-Schmidt a la base estándar $\{\mathbf{e}_1, \mathbf{e}_2\}$ para determinar una base ortogonal para este producto interior.

33. Aplique el proceso de Gram-Schmidt para construir una base ortogonal de P_2, iniciando con $1, x, x^2$ y empleando el producto interno del ejemplo 42, con $r_0 = -2, r_1 = 0, r_2 = 2$.

Para los ejercicios 34 y 35 se requiere cálculo.

34. Demuestre que los cuatro primeros polinomios de **Legendre** forman una base ortogonal para P_3, para el producto interno del ejemplo 45, con $a = -1$ y $b = 1$.

$$\mathcal{L} = \left\{1, x, \frac{3}{2}x^2 - \frac{1}{2}, \frac{5}{2}x^3 - \frac{3}{2}x\right\}$$

35. Utilice los polinomios de Legendre y los productos internos del ejercicio 34 para construir una base ortonormal para P_3.

8.9 Aplicaciones y temas adicionales

Objetivo del estudiante para esta sección
Conocer algunas de las aplicaciones de los productos punto.

En esta sección describiremos algunas de las muchas aplicaciones de los productos punto, entre ellas la de ortogonalidad, de mínimos cuadrados, etc., las cuales no sólo son muy interesantes, sino también divertidas. Comenzaremos con una forma de evaluar a los (*quarterbacks*) mariscales de campo de la NFL (siglas de la liga estadounidense de fútbol americano).

Evaluación de los mariscales de campo de la NFL

Presentaremos la esencia de un artículo interesante de Roger W. Johnson, publicado en *The College Mathematics Journal* 5, en noviembre de 1993. En su artículo, Johnson deduce la fórmula para evaluar a un mariscal de campo.[10] Aunque seguiremos la explicación de Johnson casi al pie de la letra, usaremos datos un poco más recientes, en parte para verificar si la fórmula sigue siendo válida.

Las tablas 8.1 y 8.2 fueron tomadas del artículo publicado en *The Sports Illustrated 1995 Sports Almanac*, de Peter King. La tabla 8.1 muestra los mejores mariscales de campo de la NFL en 1993 de la Conferencia Americana.

[10] De acuerdo con una tabla de los mejores mariscales de campo en la temporada de 1989-1990.

TABLA 8.1 Conferencia Americana de Fútbol Americano, 1993

Jugador	Int.	Comp.	Yardas	TD	Int.	Calificación
Elway	551	348	4030	25	10	92.8
Montana	298	181	2144	13	7	87.4
Testaverde	230	130	1797	14	9	85.7
Esiason	473	288	3421	16	11	84.5
Mitchell	233	133	1773	12	8	84.2
Hostetler	419	236	3242	14	10	82.5
Kelly	470	288	3382	18	18	79.9
O'Donnell	486	270	3208	14	7	79.5
George	407	234	2526	8	6	76.3
DeBerg	227	136	1707	7	10	75.3

La tabla 8.2 presenta la misma información para la Conferencia Nacional.

TABLA 8.2 Conferencia Nacional de Fútbol Americano, 1993

Jugador	Int.	Comp.	Yardas	TD	Int.	Calificación
Young	462	314	4023	29	16	101.5
Aikman	392	271	3100	15	6	99.0
Simms	400	247	3038	15	9	88.3
Brister	309	181	1905	14	5	84.9
Hebert	430	263	2978	24	17	84.0
Buerlein	418	258	3164	18	17	82.5
McMahon	331	200	1967	9	8	76.2
Favre	522	318	3303	19	24	72.2
Harbaugh	325	200	2002	7	11	72.1
Wilson	388	221	2457	12	15	70.1

Dados los intentos (Int.), los pases completos (Comp.), las yardas, las anotaciones (TD, por siglas en inglés), las intercepciones (Int.) y las calificaciones, deseamos deducir una fórmula para las calificaciones. De acuerdo con Johnson, se sabe que esas calificaciones dependen de los porcentajes de pases completos, las anotaciones y las intercepciones, y también del avance promedio por pase intentado (tomado en cuenta en las tablas 8.3 y 8.4). Sin embargo, parece que no se publicó la fórmula para las calificaciones. Supondremos que éstas dependen *linealmente* de las cuatro cantidades y de una constante, esto es:

$$\text{Calificación} = x_1 + x_2\,(\%\ \text{Comp.}) + x_3\,(\%\ \text{TD}) + x_4\,(\%\ \text{Int.}) + x_5\,(\text{Yard/Intento.}) \tag{8.46}$$

Deseamos calcular los coeficientes desconocidos $\mathbf{x} = (x_1, \ldots, x_5)$ a partir de las tablas 8.3 y 8.4. De esta manera se obtiene un sistema de 20 ecuaciones (10 por tabla) con 5 incógnitas. Si A es la matriz de coeficientes de este sistema, entonces

$$A\mathbf{x} = \mathbf{b} \tag{8.47}$$

donde $\mathbf{b}$ es el vector de todas las calificaciones. Observe que A tiene una primera columna de unos, y las restantes son los porcentajes de las tablas.

Parece razonable usar sólo los datos de 5 jugadores para obtener un sistema cuadrado en $\mathbf{x}$. Sin embargo, si comparamos las soluciones para los primeros 5 jugadores de la Conferencia Americana con las correspondientes para la Conferencia Nacional, las respuestas di-

TABLA 8.3 Conferencia Americana

Jugador	% Comp.	% TDs	% Int.	Yd/Intento
Elway	63.1579	4.5372	1.8149	7.3140
Montana	60.7383	4.3624	2.3490	7.1946
Testaverde	56.5217	6.0870	3.9130	7.8130
Esiason	60.8879	3.3827	2.3256	7.2326
Mitchell	57.0815	5.1502	3.4335	7.6094
Hostetler	56.3246	3.3413	2.3866	7.7375
Kelly	61.2766	3.8298	3.8298	7.1957
O'Donnell	55.5556	2.8807	1.4403	6.6008
George	57.4939	1.9656	1.4742	6.2064
DeBerg	59.9119	3.0837	4.4053	7.5198

TABLA 8.4 Conferencia Nacional

Jugador	% Comp.	% TDs	% Int.	Yd/Intento
Young	67.9654	6.2771	3.4632	8.7078
Aikman	69.1327	3.8265	1.5306	7.9082
Simms	61.7500	3.7500	2.2500	7.5950
Brister	58.5761	4.5307	1.6181	6.1650
Hebert	61.1628	5.5814	3.9535	6.9256
Buerlein	61.7225	4.3062	4.0670	7.5694
McMahon	60.4230	2.7190	2.4169	5.9426
Favre	60.9195	3.6398	4.5977	6.3276
Harbaugh	61.5385	2.1538	3.3846	6.1600
Wilson	56.9588	3.0928	3.8660	6.3325

fieren en (0.82, –0.04, 0, –0.2, 0.28), las cuales se redondearon a 2 cifras decimales. Es claro que el sistema de ecuaciones (8.47) es inconsistente.

Por lo anterior, tendremos que calcular una solución óptima empleando mínimos cuadrados. Las ecuaciones normales son

$$A^TA\mathbf{x} = A^T\mathbf{b}$$

Usando la información de la tabla 8.4 y el MATLAB (formato de notación banco), tenemos el sistema

$$\begin{bmatrix} 20.00 & 1\,209.10 & 78.50 & 58.52 & 142.06 \\ 1\,209.10 & 73\,333.88 & 4\,766.20 & 3\,536.22 & 8\,609.83 \\ 78.50 & 4\,766.20 & 335.31 & 236.62 & 568.22 \\ 58.52 & 3\,536.22 & 236.62 & 192.91 & 417.78 \\ 142.06 & 8\,609.83 & 568.22 & 417.78 & 1\,019.51 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1\,658.90 \\ 100\,652.76 \\ 6\,634.23 \\ 4\,794.16 \\ 11\,871.57 \end{bmatrix}$$

y su solución

$$\mathbf{x} = (2.0589, 0.8321, 3.3178, -4.1666, 4.1884)$$

Por consiguiente, la fórmula para las calificaciones es

$$\text{Calificación} = 2.0589 + 0.8321\ (\%\ \text{Comp.}) + 3.3178\ (\%\ \text{TD})$$
$$-4.1666\ (\%\ \text{Int.}) + 4.1884\ (\text{Yd/Intento})$$

De hecho, si calculamos el producto $A\mathbf{x}$ obtenemos todas las calificaciones hasta la exactitud mostrada.

Esos coeficientes, con las cuatro cifras decimales, no son muy adecuados. Hay una aproximación racional que puede notarse en el artículo de Johnson,

$$\text{Calificación} = \frac{1}{24}\,(50 + 20\ (\%\ \text{Comp.}) + 80\ (\%\ \text{TD}) - 100\ (\%\ \text{Int.}) + 100\ (\text{Yd/Intento}))$$

con la que se obtiene la misma exactitud. Parece razonable suponer que ésta es la fórmula correcta para las calificaciones.

OBSERVACIÓN Hemos usado los datos de ambas conferencias para obtener una aproximación más exacta por mínimos cuadrados.

Análisis de tendencia y polinomios de mínimos cuadrados

En la sección 8.4 vimos cómo determinar una recta por mínimos cuadrados que se ajuste a los puntos en el plano. Sin embargo, no todos los conjuntos de datos pueden aproximarse satisfactoriamente mediante rectas. Con frecuencia hay que usar polinomios cuadráticos o cúbicos, o funciones aún más complicadas. Surge entonces la pregunta: ¿cuál función es adecuada en determinada situación, y cómo se calcula? Éste es el tema de un **análisis de tendencia**. Investigaremos el ajuste polinomial de los datos.

Buscamos un polinomio $q(x)$ con grado máximo $n - 1$ que se aproxime mejor a un conjunto de puntos $(a_1, b_1), \ldots, (a_m, b_m)$. Sea

$$q(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-1}x^{n-1}$$

La evaluación de q en las abscisas de los puntos podría no dar como resultado las ordenadas correspondientes. Supongamos que los errores son $\delta_1, \ldots, \delta_m$. Tenemos lo siguiente:

$$\begin{aligned}
b_1 &= \alpha_0 + \alpha_1 a_1 + \cdots + \alpha_{n-1}a_1^{n-1} + \delta_1\\
b_2 &= \alpha_0 + \alpha_1 a_2 + \cdots + \alpha_{n-1}a_2^{n-1} + \delta_2\\
&\\
b_m &= \alpha_0 + \alpha_1 a_m + \cdots + \alpha_{n-1}a_m^{n-1} + \delta_m
\end{aligned}$$

En notación matricial,

$$\mathbf{b} = A\alpha + \Delta$$

donde

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \qquad \alpha = \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_{n-1} \end{bmatrix}, \qquad \Delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_m \end{bmatrix}$$

y A es la matriz $m \times n$ de coeficientes

$$A = \begin{bmatrix} 1 & a_1 & a_1^2 & & a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_m & a_m^2 & & a_m^{n-1} \end{bmatrix}$$

El objetivo es determinar un vector α que minimice la longitud del vector de error $\|\Delta\| = \|\mathbf{b} - A\alpha\|$. Como vimos en la sección 8.4, esto equivale a despejar α de las ecuaciones normales

$$A^TA\alpha = A^T\mathbf{b}$$

Sabemos que siempre existen las soluciones por mínimos cuadrados, pero ¿qué hay acerca de su unicidad? Después de todo, queremos *un* polinomio que se ajuste mejor a los datos. La respuesta es el teorema siguiente.

TEOREMA 41

Sean los puntos de los datos $(a_1, b_1), \ldots, (a_m, b_m)$ todos ellos con distintas ordenadas (es decir, todas las a_i diferentes). Para cualquier entero positivo $n \leq m$ hay un polinomio *único*

$$q(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-1}x^{n-1}$$

que minimiza $\|\Delta\|$.

DEMOSTRACIÓN Se deja como ejercicio. $\square$

El polinomio único del teorema 41 se llama **polinomio de mínimos cuadrados** de grado $n - 1$ para esos puntos. Un caso especial muy interesante es cuando A es cuadrada y entonces $m = n$.

TEOREMA 42

Si los puntos $(a_1, b_1), \ldots, (a_n, b_n)$ tienen abscisas distintas, entonces $\Delta = \mathbf{0}$ y el polinomio único de mínimos cuadrados

$$q(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-1}x^{n-1}$$

pasa realmente por todos los puntos. Así,

$$q(a_i) = b_i, \qquad i = 1, \ldots, n$$

DEMOSTRACIÓN Se deja como ejercicio. $\square$

El polinomio único del teorema 42 se llama **polinomio de interpolación** para los puntos. Observe que si A es cuadrada, su transpuesta

$$A^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix}$$

es la matriz de Vandermonde que estudiamos en la sección 6.6.

■ EJEMPLO 50 Determine (a) la línea y (b) la forma cuadrática de mínimos cuadrados para las calificaciones C del profesor de álgebra lineal.

Semestre	1	2	3	4	5	6
Porcentaje de calificaciones C	0.20	0.25	0.25	0.35	0.35	0.30

SOLUCIÓN

(a) Para calcular la recta de mínimos cuadrados, tenemos que

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 0.20 \\ 0.25 \\ 0.25 \\ 0.35 \\ 0.35 \\ 0.30 \end{bmatrix}$$

Entonces,

$$A^T A\alpha = \begin{bmatrix} 6 & 21 \\ 21 & 91 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = A^T\mathbf{b} = \begin{bmatrix} 1.7 \\ 6.4 \end{bmatrix}$$

cuya solución es $\alpha = (0.19333, 0.0257)$. Así, la recta de mínimos cuadrados es

$$y = 0.19333 + 0.0257x$$

(b) Para determinar la cuadrática de cuadrados mínimos, sean

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix}$$

Las ecuaciones normales son

$$A^T A\alpha = \begin{bmatrix} 6 & 21 & 91 \\ 21 & 91 & 441 \\ 91 & 441 & 2275 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = A^T\mathbf{b} = \begin{bmatrix} 1.7 \\ 6.4 \\ 28.6 \end{bmatrix}$$

cuya solución única es $\alpha = (0.11, 0.0882, -0.0089)$. Por lo que la curva cuadrática de mínimos cuadrados es

$$y = 0.11 + 0.0882x - 0.0089x^2$$

La figura 8.20 muestra las dos funciones de mínimos cuadrados y los puntos. Si la forma cuadrática es la mejor aproximación (y parece serlo), es de esperar que la cantidad de C disminuya en general. ☐

Mínimos cuadrados continuos (se requiere cálculo)

A estas alturas tenemos experiencia en calcular ajustes por mínimos cuadrados para un conjunto finito de puntos (el caso discreto). ¿Qué hacer si hay un conjunto infinito de puntos y deseamos determinar una curva que se ajuste a ellos?

Supongamos que deseamos encontrar la recta $y = b + mx$ que se aproxime mejor a la función $f(x) = x^2$ en el intervalo $[0, 1]$.

A menos que elijamos una cantidad infinita de puntos, no podremos usar más el producto punto ordinario. Sin embargo, es posible emplear el producto interno para funciones continuas del ejemplo 45 de la sección 8.8.

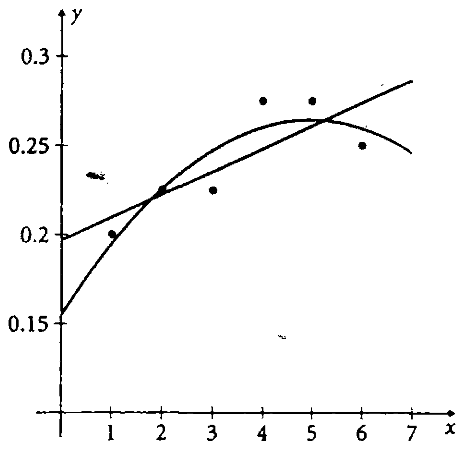


Figura 8.20 Recta y cuadrática de mínimos cuadrados para los mismos datos.

Necesitamos determinar las b y m que minimicen el vector error

$$\Delta = x^2 - (b + mx) \tag{8.48}$$

usando el producto interno integral. Lo que es igual, tenemos que minimizar

$$\|\Delta\|^2 = \int_0^1 \left(x^2 - (b + mx)\right)^2 dx$$

La ecuación (8.48) puede escribirse en notación vectorial como sigue:

$$\Delta = x^2 - A\mathbf{x} \tag{8.49}$$

en la que

$$A = \begin{bmatrix} 1 & x \end{bmatrix} \qquad \text{y} \qquad \mathbf{x} = \begin{bmatrix} b \\ m \end{bmatrix}$$

De acuerdo con la sección 8.4, sabemos que si $\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{b} \\ \tilde{m} \end{bmatrix}$ es una solución del sistema $A\mathbf{x} = x^2$ por mínimos cuadrados, entonces $A\,\tilde{\mathbf{x}}$ debe ser la proyección única de x^2 sobre la Col(A). Y contamos con una fórmula para la proyección, siempre que utilicemos una base ortogonal de la Col(A). Es claro que $\{1, x\}$ produce la base ortogonal $\{1, x - \frac{1}{2}\}$, usando el proceso de Gram-Schmidt. Así,

$$\begin{aligned} A\tilde{\mathbf{x}} = x^2_{\text{pr}} &= \frac{\langle x^2, 1\rangle}{\langle 1, 1\rangle} 1 + \frac{\langle x^2, x - \frac{1}{2}\rangle}{\langle x - \frac{1}{2}, x - \frac{1}{2}\rangle}\left(x - \frac{1}{2}\right) \\ &= -\frac{1}{6} + x \end{aligned}$$

puesto que

$$\int_0^1 x^2 \cdot 1\,dx = \frac{1}{3} \qquad \int_0^1 x^2 \cdot \left(x - \frac{1}{2}\right) dx = \frac{1}{12}$$
$$\int_0^1 1 \cdot 1\,dx = 1 \qquad \int_0^1 \left(x - \frac{1}{2}\right) \cdot \left(x - \frac{1}{2}\right) dx = \frac{1}{12}$$

Por tanto,

$$A\widetilde{\mathbf{x}} = \begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} \widetilde{b} \\ \widetilde{m} \end{bmatrix} = -\frac{1}{6} + x$$

Debido a lo anterior, $\widetilde{b} = -\frac{1}{6}$ y $\widetilde{m} = 1$. Así, la recta de mínimos cuadrados que se aproxima mejor a x^2 en [0, 1] es

$$y = -\frac{1}{6} + x$$

(véase figura 8.21). Observe que la respuesta depende mucho del intervalo que se elija.

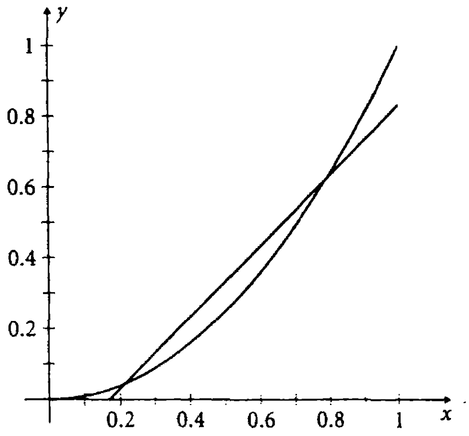


Figura 8.21 La recta de mínimos cuadrados para x^2 sobre [0, 1].

En realidad, no necesitamos ortogonalizar. Podemos usar las "ecuaciones normales" correspondientes

$$A^TA\widetilde{\mathbf{x}} = A^Tx^2$$

En este caso la multiplicación matricial es muy divertida. Usa el producto interno corriente y no el producto punto. Por A^TA se sobreentiende

$$A^TA = \begin{bmatrix} 1 \\ x \end{bmatrix} \begin{bmatrix} 1 & x \end{bmatrix} = \begin{bmatrix} \langle 1,1\rangle & \langle 1,x\rangle \\ \langle x,1\rangle & \langle x,x\rangle \end{bmatrix}$$

y A^Tx^2 es

$$A^Tx^2 = \begin{bmatrix} 1 \\ x \end{bmatrix} \begin{bmatrix} x^2 \end{bmatrix} = \begin{bmatrix} \langle 1,x^2\rangle \\ \langle x,x^2\rangle \end{bmatrix}$$

Al calcular las integrales correspondientes de los elementos se obtiene el sistema

$$\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \tilde{b} \\ \tilde{m} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{4} \end{bmatrix}$$

Con la misma solución $\tilde{b} = -\frac{1}{6}$ y $\tilde{m} = 1$, como antes.

Dejamos al lector la justificación de estas acciones.

Series y polinomios de Fourier (se requiere cálculo)

En muchas aplicaciones se necesita analizar una función en términos de su periodicidad (como la que representa a una onda sonora). Sin embargo, la mayor parte de éstas no son periódicas, y entonces se trata de aproximarlas usando funciones periódicas como el seno y el coseno. Esta idea data de Euler, sin embargo, floreció con el trabajo de Fourier.[11]

Sea $\mathcal{B}$ el conjunto de las siguientes funciones trigonométricas, definidas en $[-\pi, \pi]$.

$$\mathcal{B} = \{1, \cos x, \cos 2x, \ldots, \text{ y } \cos nx, \operatorname{sen} x, \operatorname{sen} 2x, \ldots \operatorname{sen} nx\}$$

Un **polinomio trigonométrico** es una combinación lineal de elementos de $\mathcal{B}$.

$$p(x) = a_0 + a_1 \cos x + \cdots + a_n \cos nx + b_1 \operatorname{sen} x + \cdots + b_n \operatorname{sen} nx$$

Si a_n y b_n no son cero a la vez, se dice que $p(x)$ tiene **orden** n.

Es un hecho básico que cualquier función $f(x)$ en $C[-\pi, \pi]$ puede aproximarse mediante un polinomio trigonométrico. Por aproximar queremos decir que $f(x)$ y algún $p(x)$ están cerca de la norma definida por el producto interno integral del ejemplo 45 de la sección 8.8.

Sea $T_n[-\pi, \pi]$ un subespacio de $C[-\pi, \pi]$ formado por todos los polinomios trigonométricos cuyo orden es cuando mucho n. Entonces $T_n[-\pi, \pi] = \text{Gen}(\mathcal{B})$. El primer hecho básico es el siguiente.

TEOREMA 43

> $\mathcal{B}$ es una base ortogonal de $T_n[-\pi, \pi]$.

DEMOSTRACIÓN Es claro que $\mathcal{B}$ genera a $T_n[-\pi, \pi]$. Dejamos como ejercicio verificar que $\mathcal{B}$ es linealmente independiente. Para comprobar que $\mathcal{B}$ es ortogonal, necesitamos demostrar que dos funciones distintas cualesquiera son ortogonales, es decir,

1. $\langle 1, \cos nx \rangle = 0, n = 1, 2, \ldots$
2. $\langle 1, \sin nx \rangle = 0, n = 1, 2, \ldots$
3. $\langle \cos mx, \cos nx \rangle = 0, m \neq n$
4. $\langle \cos mx, \sin nx \rangle = 0, m, n = 1, 2, \ldots$
5. $\langle \sin mx, \sin nx \rangle = 0, m \neq n$

[11] **Jean-Baptiste Joseph Fourier** (1768-1830), matemático y físico francés. Se hizo famoso por su solución de la ecuación de transmisión de calor. Introdujo la serie de Fourier, herramienta fundamental de la física matemática. Siguió a Napoleón a Egipto y después el emperador le concedió el título de barón.

Para demostrar la tercera identidad,

$$\begin{aligned}\langle \cos mx, \cos nx\rangle &= \int_{-\pi}^{\pi} \cos mx \cos nx\, dx\\ &= \frac{1}{2}\int_{-\pi}^{\pi} (\cos(m+n)x + \cos(m-n)x)\, dx\\ &= \frac{1}{2}\left[\frac{\operatorname{sen}(m+n)x}{m+n} + \frac{\operatorname{sen}(m-n)x}{m-n}\right]_{-\pi}^{\pi} = 0\end{aligned}$$

En el segundo paso usamos una identidad trigonométrica, y en el último, el hecho de que sen $k\pi = 0$ para todo k entero. El resto de las identidades se verifican en forma parecida. $\square$

Es fácil calcular las normas de las funciones de $\mathcal{B}$. Por ejemplo, usando la fórmula de la mitad del ángulo

$$\begin{aligned}\|\cos kx\|^2 &= \langle \cos kx, \cos kx\rangle\\ &= \int_{-\pi}^{\pi} \cos^2 kx\, dx\\ &= \frac{1}{2}\int_{-\pi}^{\pi} (1 + \cos 2kx)\, dx\\ &= \frac{1}{2}\left[x + \frac{\operatorname{sen} 2kx}{2k}\right]_{-\pi}^{\pi} = \pi\end{aligned}$$

En forma similar se calculan $\|1\|^2$ y $\|\operatorname{sen} kx\|^2$ con lo que se obtiene

$$\|1\| = \sqrt{2\pi}, \qquad \|\cos kx\| = \sqrt{\pi}, \qquad \|\operatorname{sen} kx\| = \sqrt{\pi}$$

Ahora, para aproximar f tan sólo necesitamos la proyección ortogonal f_{pr} de f sobre $T_n[-\pi, \pi]$ usando la base ortogonal $\mathcal{B}$. Supongamos que

$$f_{\text{pr}}(x) = a_0 + a_1 \cos x + \cdots + a_n \cos nx + b_1 \operatorname{sen} x + \cdots + b_n \operatorname{sen} nx \tag{8.50}$$

entonces, los coeficientes de Fourier se expresan (igual que en el caso del producto punto) por medio de

$$a_0 = \frac{\langle f, 1\rangle}{\langle 1, 1\rangle}, \quad a_k = \frac{\langle f, \cos kx\rangle}{\langle \cos kx, \cos kx\rangle}, \quad b_k = \frac{\langle f, \operatorname{sen} kx\rangle}{\langle \operatorname{sen} kx, \operatorname{sen} kx\rangle}$$

Por consiguiente, para $k \geq 1$,

$$\begin{aligned}a_0 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, dx\\ a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos kx\, dx\\ b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \operatorname{sen} kx\, dx\end{aligned} \tag{8.51}$$

Éstas son las famosas **fórmulas de Euler** que usó Fourier para resolver la ecuación de transmisión de calor. El polinomio trigonométrico que aproxima a la f definida por las ecuaciones

(8.50) y (8.51) se llama **polinomio de Fourier** (o **aproximación de Fourier**) de orden n de f en el intervalo $[-\pi, \pi]$.

■ EJEMPLO 51 Determinar el polinomio de Fourier de orden n de $f(x) = x$ en $[-\pi, \pi]$.

SOLUCIÓN Tenemos que

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x\,dx = \frac{1}{2\pi}\left.\frac{x^2}{2}\right|_{-\pi}^{\pi} dx = 0$$

Para $k \geq 1$, con integración por partes,

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x\cos kx\,dx = \frac{1}{\pi}\left[\frac{\cos kx}{k^2} + \frac{x\,\mathrm{sen}\,kx}{k}\right]_{-\pi}^{\pi} = 0$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x\,\mathrm{sen}\,kx\,dx = \frac{1}{\pi}\left[\frac{\mathrm{sen}\,kx}{k^2} - \frac{x\cos kx}{k}\right]_{-\pi}^{\pi} = \frac{2(-1)^{k+1}}{k}$$

porque $\cos k\pi = (-1)^k$ para todo entero k. Por tanto, la aproximación de Fourier p_n de f es

$$p_n(x) = \mathrm{sen}\,x - \mathrm{sen}\,2x + \frac{2}{3}\,\mathrm{sen}\,3x + \cdots + \frac{2(-1)^{n+1}}{n}\,\mathrm{sen}\,nx$$

La figura 8.22 muestra $f(x) = x$ trazada junto con $p_2(x)$ y $p_3(x)$ en $[-\pi, \pi]$. □

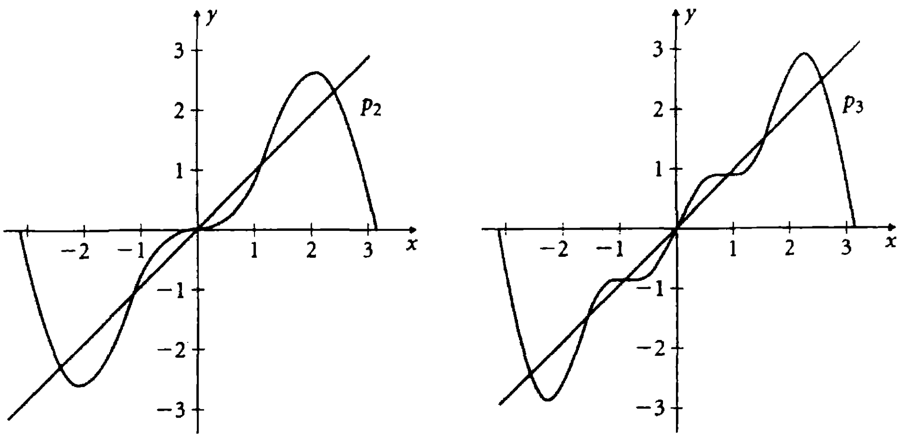


Figura 8.22 Aproximaciones de Fourier de órdenes 2 y 3 para x en $[-\pi, \pi]$.

A medida que crece n, los polinomios p_n se acercan más a f. Tomando el límite cuando $n \to \infty$ se obtiene una serie infinita y se escribe como sigue:

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n\,\mathrm{sen}\,nx)$$

El lado derecho se llama **serie de Fourier** de f en $[-\pi, \pi]$.

Ondulaciones (se requiere cálculo)

Nuestra última pero no menos importante aplicación de productos internos es en la teoría de las ondulaciones,[12] y cada día se utiliza más. El objetivo de esta teoría es atacar muchos de los problemas que se han tratado de resolver con los polinomios de Fourier. Por lo general, en éstos intervienen ondas, frecuencias, amplitudes, etc. En muchos casos, los resultados obtenidos al usar ondulaciones son mucho más favorables, en comparación con los que emplean análisis de Fourier. Ilustraremos algo de lo más destacado de la teoría. En el proyecto 1 de la sección 8.10 presentaremos más información.

Primero definiremos la **ondulación madre** $\psi(x)$ por medio de

$$\psi(x) = \psi_{0.0}(x) = \begin{cases} 1, & \text{si } 0 \le x \le \frac{1}{2} \\ -1, & \text{si } \frac{1}{2} < x \le 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Véase la figura 8.23.

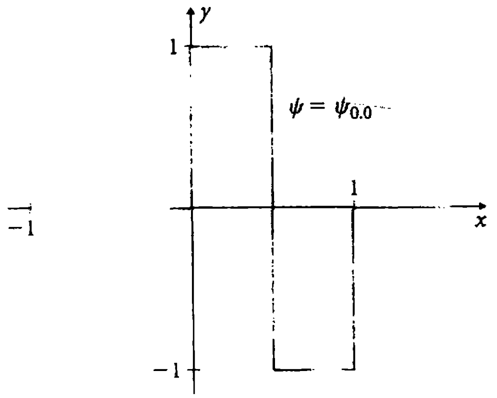


Figura 8.23 La ondulación madre.

A continuación, para cada par de enteros m y n definiremos las **ondulaciones de Haar** (o **básicas**) $\psi_{m,n}(x)$ en función de la ondulación madre como sigue:

$$\psi_{m,n}(x) = 2^{-m/2}\psi(2^{-m}x - n)$$

Como vemos en el proyecto 1, esto equivale a la definición completa

$$\psi_{m,n}(x) = \begin{cases} 2^{-m/2}, & \text{si } 2^m n \le x \le 2^m(n + \frac{1}{2}) \\ -2^{-m/2}, & \text{si } 2^m(n + \frac{1}{2}) < x \le 2^m(n+1) \\ 0, & \text{en cualquier caso.} \end{cases}$$

La figura 8.24 muestra las ondulaciones básicas $\psi_{-2,-3}$, $\psi_{0,1}$, $\psi_{1,2}$ y $\psi_{2,2}$.

[12] George Nakos tiene una deuda con el profesor P. R. Turner por permitirle leer y usar sus notas sobre este tema.

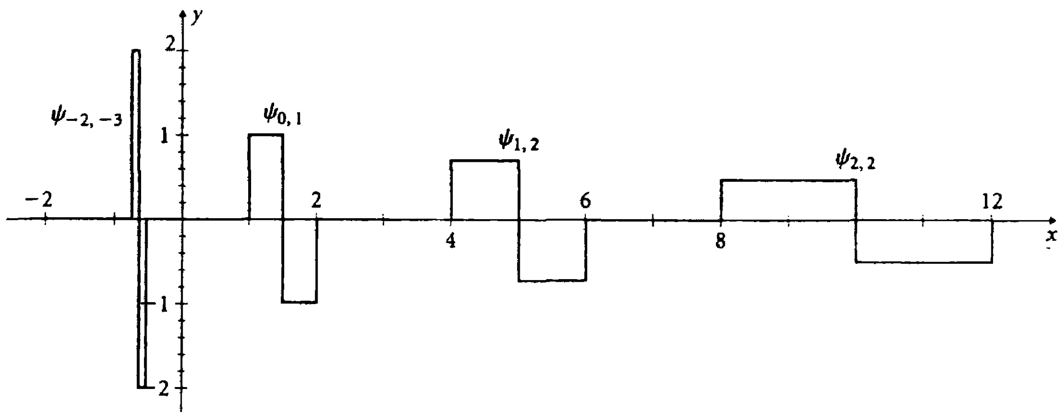


Figura 8.24 Algunas ondulaciones básicas.

El intervalo $I_{m,n} = [2^m n, 2^m(n + 1)]$, que es el único conjunto sobre el cual $\psi_{m,n}$ es distinto de cero, se llama **soporte** de la ondulación. En general, el soporte de una función f es el conjunto de puntos x tales que $f(x) \neq 0$. Por ejemplo, el soporte de $\psi_{-2,-3}$ es $[-\frac{3}{4}, -\frac{1}{2}]$, mientras que el de $\psi_{2,2}$ es $[8, 12]$.

Para las funciones f y g consideraremos el producto interno usual, pero en este caso integraremos sobre la *recta real completa*,

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx \tag{8.52}$$

La integral indefinida no siempre está definida. Sin embargo, puede demostrarse que es definida para funciones con norma finita,

$$\|f\| = \left(\int_{-\infty}^{\infty} f(x)^2\,dx\right)^{1/2} < \infty \tag{8.53}$$

El conjunto de funciones que satisface esta condición, representado por L_2, es un espacio vectorial bajo la suma y multiplicación por escalar normales de las funciones. También es un espacio de producto interno con la ecuación (8.52) porque es la que define dicho producto. Las funciones de L_2 se llaman **funciones L_2**. Las ondulaciones básicas son funciones L_2.

El primer hecho interesante es que todas las ondulaciones básicas son unitarias, es decir,

$$\|\psi_{m,n}\| = 1$$

para todos los enteros m y n. Puesto que

$$\begin{aligned}
\|\psi_{m,n}\|^2 &= \int_{-\infty}^{\infty} \psi_{m,n}(x)^2\,dx \\
&= \int_{-\infty}^{2^m n} \psi_{m,n}(x)^2\,dx + \int_{2^m n}^{2^m(n+1/2)} \psi_{m,n}(x)^2\,dx \\
&\quad + \int_{2^m(n+1/2)}^{2^m(n+1)} \psi_{m,n}(x)^2\,dx + \int_{2^m(n+1)}^{\infty} \psi_{m,n}(x)^2\,dx
\end{aligned}$$

$$
\begin{aligned}
&= 0 + \int_{2^m n}^{2^m(n+1/2)} 2^{-m}\,dx + \int_{2^m(n+1/2)}^{2^m(n+1)} 2^{-m}\,dx + 0 \\
&= 2^{-m}x\Big|_{2^m n}^{2^m(n+1/2)} + 2^{-m}x\Big|_{2^m(n+1/2)}^{2^m(n+1)} \\
&= \frac{1}{2} + \frac{1}{2} = 1
\end{aligned}
$$

También, dos ondulaciones básicas cualesquiera son *ortogonales*. Así, para $(m_1, n_1) \neq (m_2, n_2)$

$$\langle \psi_{m_1,n_1}, \psi_{m_2,n_2} \rangle = \int_{-\infty}^{\infty} \psi_{m_1,n_1}(x)\, \psi_{m_2,n_2}(x)\, dx = 0 \tag{8.54}$$

La demostración de lo anterior se analizará en el proyecto 1. Veamos el teorema siguiente.

TEOREMA 44

Todas las ondulaciones básicas $\psi_{m,n}$ forman un conjunto ortonormal.

De igual manera que las funciones pueden aproximarse con polinomios trigonométricos, esto también puede hacerse con combinaciones lineales de ondulaciones básicas. Y éste es el caso de todas las funciones de L_2. Si f es cualquier función L_2 y V es el generador de ondulaciones finitas de Haar, entonces la proyección f_{pr} de f sobre V es una combinación lineal

$$f_{\text{pr}}(x) = \sum_{m,n} c_{m,n} \psi_{m,n}(x)$$

en la que m y n asumen valores de dos conjuntos finitos. Los coeficientes $c_{m,n}$ se calculan como siempre mediante

$$c_{m,n} = \frac{\langle f, \psi_{m,n} \rangle}{\langle \psi_{m,n}, \psi_{m,n} \rangle} = \int_{-\infty}^{\infty} f(x)\psi_{m,n}(x)\, dx$$

porque $\langle \psi_{m,n}, \psi_{m,n} \rangle = \|\psi_{m,n}\|^2 = 1$. Esta vez la integral no es indefinida, porque el soporte de $\psi_{m,n}$ es un intervalo finito. Tenemos que

$$
\begin{aligned}
c_{m,n} &= \int_{2^m n}^{2^m(n+1)} f(x)\psi_{m,n}(x)\, dx \\
&= \int_{2^m n}^{2^m(n+1/2)} f(x)2^{-m/2}\, dx + \int_{2^m(n+1/2)}^{2^m(n+1)} f(x)(-2^{-m/2})\, dx
\end{aligned}
$$

Así, podemos escribir

$$
\begin{aligned}
c_{m,n} &= A_{m,n} - B_{m,n} \\
A_{m,n} &= 2^{-m/2} \int_{2^m n}^{2^m(n+1/2)} f(x)\, dx \\
B_{m,n} &= 2^{-m/2} \int_{2^m(n+1/2)}^{2^m(n+1)} f(x)\, dx
\end{aligned}
\tag{8.55}
$$

Las fórmulas (8.55) dan como resultado los coeficientes c_{mn} de f_{pr} y son una combinación lineal de $\psi_{m,n}$. Son análogas a las fórmulas (8.51) con las que se calculaban los coeficientes en la aproximación polinomial trigonométrica de f. Para aproximar una función f en forma adecuada, necesitamos tomar en cuenta los coeficientes de *todas* las ondulaciones $\psi_{m,n}$ de Haar, y una cantidad infinita de ellas puede ser distinta de cero. Así, al igual que con la serie de Fourier, escribimos f como una serie infinita en función de $\psi_{m,n}$.

$$f(x) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n}\psi_{m,n}(x)$$

■ EJEMPLO 52 Sea

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

(figura 8.25), y sea V_k el generador de las ondulaciones de Haar

$$V_k = \text{Gen}\{\psi_{1,0}, \psi_{2,0}, \ldots, \psi_{k,0}\}$$

Aproxime f calculando f_{pr} con respecto a V_k.

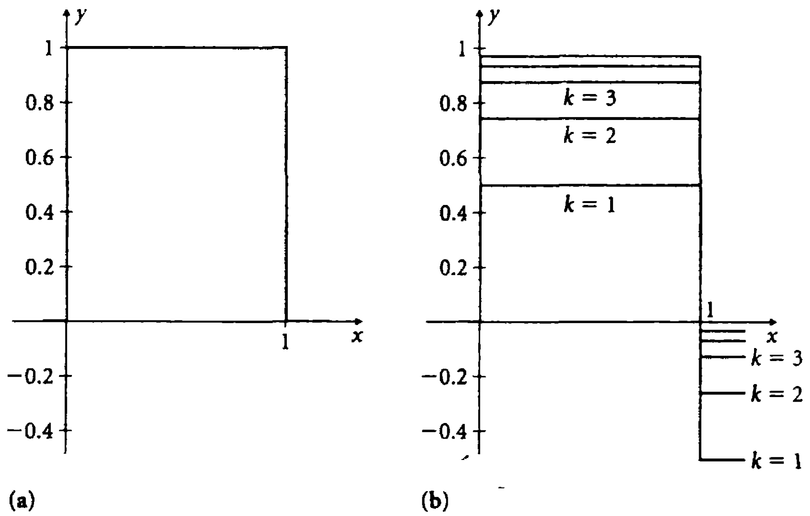


Figura 8.25 Aproximación por ondulaciones de Haar.

SOLUCIÓN Sean $c_{1,0}, \ldots, c_{k,0}$ escalares tales que, para toda x,

$$f_{pr}(x) = c_{1,0}\psi_{1,0}(x) + \cdots + c_{k,0}\psi_{k,0}(x)$$

En el proyecto 1 veremos que para $m = 1, \ldots, k$.

$$c_{m,0} = 2^{-m/2}$$

Por consiguiente,

$$f_{pr}(x) = 2^{-1/2}\psi_{1,0}(x) + 2^{-2/2}\psi_{2,0}(x) + \cdots + 2^{-k/2}\psi_{k,0}(x)$$

La figura 8.25(b) muestra las gráficas de f_{pr} para $k = 1, \ldots, 5$. Es evidente que a medida que k crece, f_{pr} tiende a f con mucha rapidez. □

Ejercicios 8.9

1. (**Se necesita computadora**) Calcule las calificaciones de la NFL usando sólo la tabla 8.1. ¿Es satisfactoria la respuesta?
2. (**Se necesita computadora**) Repita el ejercicio 1 con la tabla 8.2.

En los ejercicios 3 a 5 determine por mínimos cuadrados la función cuadrática $q(x)$ que pase por los puntos dados. A continuación evalúe $q(6)$.

3. $(1,-3),(2,0),(3,4),(4,13),(5,20)$
4. $(1,-1),(2,0),(3,1),(4,4),(5,8)$
5. $(-2,1),(-1,2),(0,4),(1,6),(2,9)$

En los ejercicios 6 y 7 calcule por mínimos cuadrados la curva cúbica $q(x)$ que pase por los puntos. A continuación evalúe $q(3)$.

6. $(-2,-8),(-1,-2),(0,0),(1,8),(2,12)$
7. $(-2,-2),(-1,0),(0,1),(1,-3),(2,7)$
8. Demuestre el teorema 41.
9. Compruebe el teorema 42.
10. Determine la recta de mínimos cuadrados que se aproxime a $f(x) = x^2$ en $[0, 2]$.
11. Obtenga la recta de mínimos cuadrados que se aproxime a $f(x) = x^2$ en $[1, 2]$.
12. Construya la recta de mínimos cuadrados que se aproxime a $f(x) = x^3$ en $[0, 1]$.
13. Encuentre la recta de mínimos cuadrados que se aproxime a $f(x) = x^3$ en $[0, 2]$.

En los ejercicios 14 y 15 use el método de las ecuaciones normales con $A = [1 \;\; x \;\; x^2]$

14. Determine por mínimos cuadrados la forma cuadrática que se aproxime a $f(x) = x^3$ en $[0, 1]$.
15. Mediante mínimos cuadrados obtenga la forma cuadrática que se aproxime a $f(x) = x^3$ en $[0, 2]$.

Partiendo de la demostración del teorema 43 demuestre lo siguiente.

16. La relación 1.
17. La relación 2.
18. La relación 4.
19. La relación 5.

En los ejercicios 20 a 22 calcule los coeficientes de Fourier a_0, a_i y b_i de $f(x)$.

20. $f(x)=\begin{cases}-1, & \text{si } -\pi < x < 0\\ 1, & \text{si } 0 < x < \pi\end{cases}$

21. $f(x)=\begin{cases}0, & \text{si } -\pi < x < 0\\ 1, & \text{si } 0 < x < \pi\end{cases}$

22. $f(x)=\begin{cases}0, & \text{si } -\pi < x < 0\\ 1, & \text{si } 0 < x < \frac{\pi}{2}\\ 0, & \text{si } \frac{\pi}{2} < x < \pi\end{cases}$

En los ejercicios 23 a 27 demuestre que el conjunto dado es ortogonal, usando el producto interno integral en el intervalo dado.

23. $\{\operatorname{sen}(x), \operatorname{sen}(2x), \ldots, \operatorname{sen}(nx)\}$, $[0, \pi]$
24. $\{1, \cos(x), \cos(2x), \ldots, \cos(nx)\}$, $[0, 2\pi]$
25. $\{\operatorname{sen}(\pi x), \operatorname{sen}(2\pi x), \ldots, \operatorname{sen}(n\pi x)\}$, $[-1, 1]$
26. $\{1, \cos(\pi x), \cos(2\pi x), \ldots, \cos(n\pi x)\}$, $[-1, 1]$
27. $\{1, \cos(\pi x), \cos(2\pi x), \ldots, \cos(n\pi x)\}$; $[0, 2]$
28. Sea

$$f(x)=\begin{cases}1, & \text{si } -1 \le x \le 0\\ 0, & \text{en cualquier otro caso}\end{cases}$$

Escriba f_{pr} en la forma

$$f_{pr}(x) = \sum_{m=1}^{k} c_{m,-1}\psi_{m,-1}(x)$$

y demuestre que

$$c_{m,-1} = -2^{-m/2}$$

Trace la gráfica de f y de la f_{pr} para (a) $k = 2$, (b) $k = 3$.

29. Sea

$$f(x)=\begin{cases}-1, & \text{si } 0 \le x \le 1\\ 0, & \text{en cualquier otro caso}\end{cases}$$

Escriba f_{pr} en la forma

$$f_{pr}(x) = \sum_{m=1}^{k} c_{m,0}\psi_{m,0}(x)$$

y compruebe que

$$c_{m,0} = -2^{-m/2}$$

Trace las gráficas de f y de f_{pr} para (a) $k = 2$, (b) $k = 3$.

8.10 Miniproyectos

1 ■ Ondulaciones

En este proyecto lo guiaremos para demostrar algunas afirmaciones que hace la teoría de las ondulaciones, mencionadas en la sección 8.9.

Problema A

Compruebe que la definición de las ondulaciones básicas, en función de la ondulación madre

$$\psi_{m,n}(x) = 2^{-m/2}\psi(2^{-m}x - n)$$

equivale a la definición completa

$$\psi_{m,n}(x) = \begin{cases} 2^{-m/2}, & \text{si } 2^m n \le x \le 2^m\left(n + \dfrac{1}{2}\right) \\ -2^{-m/2}, & \text{si } 2^m\left(n + \dfrac{1}{2}\right) < x \le 2^m(n+1) \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Problema B

Siga estos pasos para demostrar que las ondulaciones básicas son *ortogonales*, es decir, para $(m_1, n_1) \neq (m_2, n_2)$,

$$\langle \psi_{m_1,n_1}, \psi_{m_2,n_2} \rangle = \int_{-\infty}^{\infty} \psi_{m_1,n_1}(x)\, \psi_{m_2,n_2}(x)\, dx = 0 \qquad (8.56)$$

Recuerde que $I_{m,n} = [2^m n, 2^m(n+1)]$ es el soporte de $\psi_{m,n}$.

1. Si $m_1 = m_2$ y $n_1 \neq n_2$, demuestre que la intersección $I_{m_1,n_1} \cap I_{m_2,n_2}$ contiene cuando mucho un punto.
2. Si $m_1 > m_2$, compruebe que $I_{m_1,n_1} \cap I_{m_2,n_2}$ contiene cuando mucho un punto, o bien I_{m_2,n_2} está contenido en I_{m_1,n_1}.
3. Si $(m_1, n_1) \neq (m_2, n_2)$, y si I_{m_2,n_2} está contenido en I_{m_1,n_1}, demuestre que I_{m_2,n_2} está contenido ya sea en $[2^{m_1}n_1, 2^{m_1}(n_1 + 1/2)]$ o bien en $[2^{m_1}(n_1 + 1/2), 2^{m_1}(n_1 + 1)]$.
4. Use las partes 1, 2 y 3 y demuestre la ecuación (8.56) para $(m_1, n_1) \neq (m_2, n_2)$.

Problema C

Sea

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Sea $V_k = \text{Gen}\{\psi_{1,0}, \psi_{2,0}, \ldots, \psi_{k,0}\}$ y V cualquier generador de ondulaciones de Haar que contenga a V_k. Siga los pasos descritos abajo para demostrar que la proyección f_{pr} con respecto a V se expresa como sigue:

$$f_{\text{pr}}(x) = 2^{-1/2}\psi_{1,0}(x) + 2^{-2/2}\psi_{2,0}(x) + \cdots + 2^{-k/2}\psi_{k,0}(x)$$

Primero, sea

$$c_{m,n} = \int_{-\infty}^{\infty} f(x)\psi_{m,n}(x)\,dx = \int_0^1 \psi_{m,n}(x)\,dx$$

1. Si $m \geq 0$, $n \neq 0$, demuestre que $c_{m,n} = 0$.
2. Si $m = 0$, $n = 0$, compruebe que $c_{0,0} = 0$.
3. Si $m \leq 0$, $n \neq 0$, verifique que la intersección $I_{m,n} \cap [0, 1]$ es cuando mucho un punto, o bien $I_{m,n}$. Llegue a la conclusión de que $c_{m,n} = 0$.
4. Si $m > 0$, $n = 0$, compruebe que

$$c_{m,0} = 2^{-m/2}$$

Nos interesa f porque si podemos demostrar su aproximación con ondulaciones, también podremos aproximar las funciones constantes en intervalos. Esas funciones son *densas* en L_2, es decir, pueden acercar a cualquier función L_2. Así, las ondulaciones básicas podrían aproximar cualquier función L_2. La relación entre f y las $\psi_{k,0}$ es fuerte. Se puede comprobar que para toda x,

$$f(x) = \sum_{m=1}^{\infty} c_{m,0}\psi_{m,0}(x)$$

2 ■ Productos internos complejos

NOTA Para esta parte se requiere material del apéndice A. Es necesaria cierta familiaridad con la aritmética de los números complejos.

En este proyecto estudiará los productos internos sobre espacios vectoriales *complejos*. Los espacios vectoriales complejos son aquellos con números complejos como escalares. Se describieron en la sección 4.8 del proyecto 2.

DEFINICIÓN

Un **producto interno** (complejo) sobre un espacio vectorial complejo V es una función que asocia un número complejo $\langle \mathbf{u}, \mathbf{v}\rangle$, a cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ de V, y satisface las siguientes propiedades o **axiomas**.

Para todo vector $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ y para cualquier escalar complejo c:

1. $\langle \mathbf{u}, \mathbf{v}\rangle = \overline{\langle \mathbf{v}, \mathbf{u}\rangle}$;
2. $\langle \mathbf{u} + \mathbf{w}, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle$;
3. $\langle c\mathbf{u}, \mathbf{v}\rangle = c\langle \mathbf{u}, \mathbf{v}\rangle$;
4. $\langle \mathbf{u}, \mathbf{u}\rangle > 0$ if $\mathbf{u} \neq \mathbf{0}$.

Problema A

Sean $\mathbf{u} = (u_1, \ldots, u_n)$ y $\mathbf{v} = (v_1, \ldots, v_n)$ en $\mathbf{C}^n$. Demuestre que el producto **punto** (complejo)

$$\mathbf{u} \cdot \mathbf{v} = u_1\overline{v_1} + \cdots + u_n\overline{v_n}$$

define un producto interno complejo sobre $\mathbf{C}^n$.

Problema B

1. Use un producto interno complejo para definir las nociones de

   (a) Par ortogonal de vectores.
   (b) Conjunto ortogonal.
   (c) Conjunto ortonormal.

2. Cite ejemplos de esos conceptos.
3. Demuestre que las columnas de una matriz unitaria (véase la definición en el apéndice A) forman un conjunto ortonormal con respecto al producto punto complejo del problema A.

Problema C

Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores cualesquiera en un espacio de producto interno complejo, y sea c cualquier escalar complejo. Demuestre las siguientes propiedades:

1. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
2. $\langle \mathbf{u}, c\mathbf{v} \rangle = \bar{c}\langle \mathbf{u}, \mathbf{v} \rangle$

3 ■ El espín de Pauli y las matrices de Dirac

NOTA Para este párrafo se requiere material del apéndice A. Las definiciones de matrices hermitianas y unitarias se encuentran en ese mismo apéndice. Es necesaria cierta familiaridad con la aritmética de los números complejos.

En este proyecto exploraremos las propiedades básicas de ciertas matrices con elementos complejos, que desempeñan un papel importante en la física nuclear y en la mecánica cuántica.

W. Pauli[13] introdujo las tres matrices siguientes, llamadas **matrices de espín de Pauli**, para calcular el espín del electrón

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

En 1927, P. A. M. Dirac,[14] al trabajar en mecánica cuántica, generalizó las matrices de espín de Pauli a las siguientes, llamadas **matrices de Dirac**:

[13] **Wolfgang Joseph Pauli** (1900-1958) nació en Viena, Austria. Estudió física en Munich y realizó trabajos de posgrado con Bohr en Copenhague. Dio clases en Hamburgo y en Zurich. Es famoso por su trabajo sobre partículas elementales, como la predicción del neutrino y el descubrimiento del principio de exclusión que lleva su apellido. También usó 3 matrices complejas 2×2, llamadas matrices de espín de Pauli, para describir su teoría del espín de partículas elementales en 1927.

[14] **Paul Adrien Maurice Dirac** nació en Bristol, Inglaterra, en 1902. En 1926 obtuvo un doctorado de la Universidad de Cambridge. A continuación estudió con Bohr en Copenhague y con Born en Göttingen. En 1932 fue profesor lucasiano de matemáticas en Cambridge, puesto que alguna vez ocupó Newton. Ganó el Premio Nobel de Física en 1933. Se le conoce por su trabajo en mecánica cuántica, partículas elementales y la teoría de la antimateria.

$$\alpha_x = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \qquad \alpha_y = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}$$

$$\alpha_z = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \qquad \beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Las tres primeras son matrices de bloques en las matrices de espín de Pauli.

$$\alpha_x = \begin{bmatrix} \mathbf{0} & \sigma_x \\ \sigma_x & \mathbf{0} \end{bmatrix}, \qquad \alpha_y = \begin{bmatrix} \mathbf{0} & \sigma_y \\ \sigma_y & \mathbf{0} \end{bmatrix}, \qquad \alpha_z = \begin{bmatrix} \mathbf{0} & \sigma_z \\ \sigma_z & \mathbf{0} \end{bmatrix}$$

Una matriz cuadrada A es **involutoria** si $A^{-1} = A$ o, lo que es igual, si $A^2 = I$. Por ejemplo, $-I$ es involutoria.

Problema A

Demuestre que las matrices de espín de Pauli y las matrices de Dirac son

1. hermitianas;
2. unitarias;
3. involutorias.

Problema B

Compruebe que las matrices de espín de Pauli satisfacen las relaciones

$$\sigma_x \sigma_y = i\sigma_z \qquad \sigma_y \sigma_z = i\sigma_x \qquad \sigma_z \sigma_x = i\sigma_y$$

y que *anticonmutan*, es decir, que

$$\sigma_x \sigma_y = -\sigma_y \sigma_x \qquad \sigma_y \sigma_z = -\sigma_z \sigma_y \qquad \sigma_x \sigma_z = -\sigma_z \sigma_x$$

Problema C

Demuestre que las matrices de Dirac satisfacen las relaciones

$$\alpha_x \beta = -\beta \alpha_x \qquad \alpha_y \beta = -\beta \alpha_y \qquad \alpha_z \beta = -\beta \alpha_z$$

y que *anticonmutan*, es decir, que

$$\alpha_x \alpha_y = -\alpha_y \alpha_x \qquad \alpha_y \alpha_z = -\alpha_z \alpha_y \qquad \alpha_x \alpha_z = -\alpha_z \alpha_x$$

8.11 Ejercicios en computadora

Esta sección tiene por objeto familiarizar al lector con las instrucciones básicas de su programa para el material del capítulo 8. También lo ayudará a repasar este material.

Sean

$$A = \begin{bmatrix} 1 & 4 & 69 \\ 2 & -3 & 28 \\ -3 & 2 & -37 \\ 4 & 2 & -59 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 4 & 4 \end{bmatrix}, \qquad S = \begin{bmatrix} -1 & -1 & 1 \\ -1 & 2 & 4 \\ 1 & 4 & 2 \end{bmatrix}$$

1. Compruebe que A tiene columnas ortogonales como sigue: (a) usando el producto punto, (b) verificando la relación (8.4) del teorema 4, sección 8.1.
2. Sea A_1 la matriz que se obtuvo al agregar $\mathbf{e}_4 \in \mathbf{R}^4$ como última columna de A. Aplique el proceso de Gram-Schmidt a las columnas de A_1 para ortogonalizarlas. Forme una matriz A_2 con ellas y compruebe la relación (8.4).
3. Ortonormalice las columnas de la matriz A_2 del ejercicio 2, para obtener una matriz ortogonal A_3. Compruebe que A_3 sea ortogonal.
4. Verifique la desigualdad de Bessel con el conjunto de vectores ortonormales formados por las tres primeras columnas de A_3 del ejercicio 3, y con $\mathbf{u} = \mathbf{e}_4 \in \mathbf{R}^4$.
5. Determine una matriz cuyas columnas sean ortogonales y generen al espacio de columnas de B.
6. Escriba un programa corto que calcule la proyección ortogonal de un vector $\mathbf{u}$ sobre el generador de un conjunto S de vectores ortogonales.
7. Pruebe su programa del ejercicio 6 con $\mathbf{u} = (1, -1, 2)$ y $S = \{(-1, 4, 1), (5, 1, 1)\}$ (véase el ejemplo 11 de la sección 8.2).
8. Modifique su programa del ejercicio 6 para calcular la proyección ortogonal de un vector sobre el generador de cualquier conjunto finito de vectores linealmente independientes, pero no necesariamente ortogonales.
9. Pruebe su programa del ejercicio 8 con $\mathbf{u} = (1, 1, 1)$ y $S = \{(1, -1, 2), (-1, 1, 4)\}$.
10. Calcule la factorización QR de C y compruebe su respuesta.
11. Determine la recta de mínimos cuadrados que pasa por los 10 puntos de x^2 cuyas abscisas son 0, 1, . . ., 9. En la misma gráfica trace esa recta y también x^2 sobre [0, 9].
12. Deduzca y grafique la cuadrática de mínimos cuadrados para los puntos (−2, 2), (−1, 0), (2, 4), (3, 7) y (4, 9).
13. Diagonalice ortogonalmente la matriz simétrica S.
14. Determine una descomposición de Schur para B.
15. Encuentre los valores singulares y la descomposición de A en valores singulares. Compruebe que $A = U\Sigma V^T$. Repita lo anterior con C.
16. Calcule A^+ y C^+. Verifique las propiedades de Moore-Penrose para (A, A^+) y (C, C^+).
17. Considere el producto interno del ejemplo 45, en la sección 8.8, sobre [0, 1]. Use su programa para calcular los valores $\langle x^2, \operatorname{sen}(x)\rangle$, $\|x^2\|$, $\|\operatorname{sen}(x)\|$. Compruebe la desigualdad de Cauchy-Schwarz-Bunyakovski (CSB).
18. Verifique la desigualdad CSB usando el producto interno del ejemplo 42, sección 8.8, con $r_0 = -10$, $r_1 = 3$, $r_2 = 15$ para $\mathbf{u} = x^2 + x - 1$ y $\mathbf{v} = -x^2 + 2x$.
19. Escriba y pruebe una función con argumentos $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ que calcule el producto punto ponderado $\langle \mathbf{u}, \mathbf{v}\rangle$ con un factor de ponderación $\mathbf{w}$ (véase el ejemplo 39 de la sección 8.8).
20. Con su programa, verifique la fórmula de calificación de los mariscales de campo de la sección 8.9.

Soluciones seleccionadas con Maple

Comandos definite, dotprod, GramSchmidt, innerprod, leastsqrs, norm, normalize, orthog, QRdecomp, singularvals.

```
with(linalg):
A := matrix ([[1,4,69], [2,-3,28], [-3, 2, -37], [4, 2, -59]];     # Todos los datos.
B := matriz ([[1, 2, 3], [4, 5, 6], [7, 8, 9]]);
C := matriz ([[1, 2, 3, 4], [2, 2, 3, 4], [3, 3, 3, 4], [4, 4, 4, 4]]);
DD := matriz ([[-1, -1, 1], [-1, 2, 4], [1, 4, 2]]);     # Maple ya usa D.

# Ejercicio 1

dotprod(col(A, 1), col(A, 2));     # Etc. . Repetir con los otros dos pares.
evalm(transpose(A)&*A);            # (A^T)A es diagonal cuyos elementos diag. son normas^2 de
norm(col(A,1), 2)^2; norm(col(A,2),2)^2; norm(col(A,3),2)^2;     # las columnas de A.

# Ejercicio 2

e4 :=vector]); #([0, 0, 0, 1])          # e_4.
A1 := concat(A,e4);                     A1. A continuación Gram-Schmidt sobre
GramSchmidt ([col(A1, 1). col(A1, 2), col(A1, 3), col(A1, 4)]);     # las columnas de A1.
A2 := concat("[1], "[2], "[3], "[4]);      # A2.
evalm(transpose(A2) &* A2);             # Etc. . Misma verificación.

# Ejercicio 3

seq(col(A2, i)/norm(col(A2, i),2), i=1..4);     # Todas las columnas de A2 se ortonormalizan
A3 := concat(");                                # y se ponen en una matriz.
orthog(A3);                                     # Prueba de ortogonalidad.
evalm(transpose(A3)&* A3);                      # Otra forma. (A3^T) resulta en I.

# Ejercicio 4                                   # e_4 se define arriba.

Norm(e4)^2-(dotprod(e4,col(A3,1))^2\            # S derecha e izquierda de Bessel >0. Correcto.
+ dotprod(e4,col(A3,2)))^2 + dotprod(e4, col(A3,3))^2);

# Ejercicio 5

GramSchmidt ([col(B, 1,). col(B, 2), col(B, 3)]     # GS sobre las columnas de B para formar
B1 :=concat("[1], "[2]);                            # una nueva matriz B1 cuyas columnas
concat(B1, B)                                       # generan Col(B) porque las 2 primeras
rref(");                                            # columnas de [B1,B] son columnas pivote.

# Ejercicio 6

proj := proc(u, lis) local i, s;                    # Función proy para la proyección.
s := [seq(0,i=1..nops(u))]                          # 0 lista convertida a vector en el bucle.
for i from 1 to nops(lis) do
s := evalm(s + (dotprod(u, lis[i])/dotprod(lis[i], lis[i]*lis[i]) od:
evalm(s) end:

# Ejercicio 7

proj([1,-1,2], [[-1,4,1], [5,1,1]]);                # Comprobación con ejemplo 11.

# Ejercicio 10

R :=QRdecomp(C, Q='q'                               # Factorización QR.
evalm(C - q &* R);                                  # La diferencia es la matriz cero,
evalm(transpose(q)&*q);                             # y Q es ortogonal, porque (Q^T)Q=I.
```

```
# Ejercicio 11

A := matrix(10, 2, [1,0,1,1,1,2,1,3,1,4,1,5,1,6,1,7,1,8,1,9]);
b := vector([0,1,4,9,16,25,36,49,64,81]);     # No se necesitan las ecuaciones
leastsqrs(A,b);                                # normales; mínimos cuadrados lo hace en un paso.
plot({x^2,9*x-12},x=0..9);                     # Se grafican la recta de mínimos cuadrados y x^2.

# Ejercicio 13

eigsys :=eigenvects(S);                                            # Eigenvalores y eigenvectores.
D1 := diag(eigsis[1] [1], eigsys [2] [1] eigsys [3] [1]);          # Diagonal de eigenvalores.
eves := eigsys [1] [3] [1], eigsys [2] [3] [1], eigsys[3] [3] [1]; # Los eigenvec. son auto-
Q :=concat(eves[1]/norm(eves[1],2), eves[2]norm(eves[2],2)         # máticamente orto-
eves[3]/norm(eves[3],2));                # gonales, porque corresponden a distintos
orthog(");                               # eigenvals. Se divide entre las normas para obtener Q
evalm(transpose(Q) &* S &* Q);           # que es ortogonal y (Q^T)SQ=D1. Correcto.

# Ejercicio 14

# Hasta ahora no está disponible la descomposición de Schur en una etapa, pero se puede
# tratar de calcular una con los pasos de la demostración del teorema de Schur.

# Ejercicio 15 - parcial

singularvals(A); evalf(")     # Los valores singulares, exactos y aproximados.
sv := evalf(Svd(A, U, V));    # sv son valores singulares también y U y V.
evalm(U); evalm(V);           # Revisar U y V y demostrar.
evalm(transpose(U) &* U); evalm(transpose(V) &* V);    # ortogonalidad
evalm(transpose(U) &* A &* V);    # Verificar que U'AV es Sigma - respuesta de los valores singulares

# Ejercicio 16 - parcial     # No hay paso para calcular la seudoinversa, pero podemos calcularla

diag(sv[1]^(-1),sv[2]^(-1), sv[3]^(-1));          # con facilidad. En la notación del ejercicio 15,
sigplus :=concat(",[0,0,0]);                      # se forma sigma más y se multiplica para obtener
#psA := evalm(V &* sigplus &* transpose(U));      # la seudoinversa
evalm(A &* psA &* A - A);                         # Comprobación de todas
evalm(psA &* A &* psA - psA);                     # las condiciones
evalm( transpose(A &* psA) - A &* psA);           # Moore-Penrose. Todas las matrices
evalm( transpose(psA &* A) - psA &* A);           # son aprox. cero.

# Ejercicio 17

int(x^2*sin(x),x=0..1);           # <x^2, sen(x)>.
Sqrt(int(x^2,x=0..1));            # Norma (x^2).
Sqrt(int(sin(x),x=0..1));         # Norma (sen(x)).
Evalf(" "*"-" " ");        # Norma (x^2)*Norma (sen(x)) - <x^2, sen(x)> es > 0. Correcto.
```

Soluciones seleccionadas con Mathematica

Comandos Dot, Outer, QRDecomposition, SchurDecomposition, SingularValues. Y del paquete 'Orthogonalization' del Álgebra Lineal: GramSchidt, InnerProduct, Normalize, Normalized, Projection.

```
A = {{1, 3, 69}, {2, -3, 28}, {-3, 2, -37}, {4, 2, -59}}     (* Todos los datos. *)
B = {{1, 2, 3}, {4,5,6}, {7, 8, 9}}
CC = {{1, 2, 3, 4}, {2, 2, 3, 4}, {3, 3, 3, 4}, {4, 4, 4, 4}}   (* Mathematica ya usa C. *)
DD = {{-1, -1, 1}, {-1, 2, 4}, {1, 4, 2}}                        (* Mathematica ya usa D. *)
```

```
* Ejercicio 1 *
AT=Transpose[A];            (* Se puede entrar a las columnas más fácilmente por transposición. *)
Dot[AT[[1]],AT[[2]]]        (* Etc. . Repetir con los otros dos pares. *)
AT · A (* A^T)A es diagonal cuyos elementos diag. son normas^2 de las columnas de A. *)
sqnorm[lis_]:Plus@@ (lis^2)   (* Esta pequeña función calcula los cuadrados de
                                 normas de vectores. *)
sqnorm[AT[[1]]]             (* Norma^2 de la columna 1 de A, etc..*)
(* Otra forma: <<LinearAlgebra'Orthogonalization' a continuación *)
(* AT[[1]] [[1]]/Normalize[AT[[1]]] [[1]]  ,   Etcétera *)
(* Ejercicios 2 y 3  *)
e4 = {{0}}, {0}, {0}, {1}}                          (* e_4  *)
<<LinearAlgebra'MatrixManipulation'                 (* El paquete de manipulación de matrices.  *)
A1 = AppendRows [A, eA]; TA1=Transpose[A1]          (* A1 y su transpuesta.  *)
<<LinearAlgebra'Orthogonalization'                  (* Paquete de ortogonalización.  *)
A2 = Transpose [GramSchmidt[TA1]]                   (* Gram-Schmidt sobre las columnas de A.  *)
Transpose[A2] · A2    (* ¡La matriz ya es ortogonal! GS también la normaliza. *)
(* Nota: GramSchmidt[Vec_list, Normalized->False] es GS sin normalización. *)
(* Ejercicio 4 *)
ee4 = {0, 0 0, 1}                     (* Izquierda-derecha de Bessel > 0.  Correcto. Fijarse.  *)
sqnorm[ee4] - Sum[Dot[ee4,Flatten[TakeColumns[A2,{i}]]]^2,{i,1,3}]  (* Aplanamiento.  *)
(* Ejercicio 5 *) (* Al reducir B^T con operaciones de renglón vemos que tiene columnas dependientes. *)
GramSchmidt[{Transpose[B] [[1]], Transpose[B] [[2]]}]  (* Como Mathematica calcula *)
B1 = Transpose [%]       (* el proc. GS de vectores independientes sólo se usan las 2 primeras cols.  *)
AppendRows [B1, B]       (* Las cols. de B1 generan a Col(B) porque las 2 primeras cols. de  *)
RowReduce [%]            (* [B1,B] son columnas pivote.  *)
(* Ejercicio 6  *)
proj [u_ ,lis_] :=
Sum[Dot[u, lis[[i]]]/Dot[lis[[i]], lis[[i]]]*lis[[y]], {i,1,Length[lis]}]
(* Ejercicio 7  *)
proj [{1, -1, 2}, {{-1, 4, 1}, {5, 1, 1}}]     (* Prueba con el ejemplo 11.  *)
(* Ejercicio 10  *)           (* Factorización QR.      *)
QRDecomposition[N[CC]]        (* Primero es necesario evaluar numéricamente CC. *)
Q = %[[1]]; R=%[[2]];         (* Advertencia: Q es tal que (Q^T)R es CC.  *)
Transpose[Q] . R - CC         (* La diferencia es aproximadamente la matriz cero,  *)
Transpose[Q] . Q              (* y Q es ortogonal, porque (Q^T)Q=I.  *)
(* Ejercicio 11  *)
A = {{1, 0}, {1,1}, {1,2}, {1,3}, {1,4}, {1,5}, {1,6}, {1,7}, {1,8}, {1,9}}
AT = Transpose[A]
b = {{0}, {1}, {4}, {9}, {16}, {25}, {36}, {49}, {64}, {81}}
LinearSolve [AT.A,AT.b]        (* Resolver las ecuaciones normales para obtener 9x-12. *)
Plot[{x^2, 9x-12}, {x, 0, 9}]  (* Se grafican la recta de mínimos cuadrados y x^2. *)
(* Ejercicio 13  *)
D1=Diagonal Matrix[Eigenvalues[S]]      (* Matriz diagonal cuyos elem. diag. son los  *)
eves=Eigenvectores[S]                   (* eigenvalores que son distintos, así que los eigenvectores  *)
<<Linear Algebra 'Orthogonalization'    (* ya son ortogonales y sólo  *)
Q=Transpose [Map[Normalize, eves]]      (* necesitan normalización. Q y D1 *)
```

```
Transpose[Q] . S . Q                (* diagonalizan ortogonalmente a S. *)
Transpose[Q] . Q                    (* Prueba de la ortogonalidad para Q. *)
(* Ejercicio 14 *)
Eigenvalues[B]   (* Primero comprobar los eigenvalores. Todos reales. Correcto. *)
sd = Schur Despomposition [N[B]];   (* Descomposición de Schur. *)
Q=sd[[1]]                           (* Q y *)
T=sd[[2]]                           (* T. *)
Transpose[Q] . Q                    (* Q es ortogonal. *)
Q . T . Transpose[Q]                (* Comprobación y se obtuvo B. *)
(* Ejercicio 15 - parcial *)
{Ut,sig,V}=SingularValues[N[A]]  (* La descomposición SVD tiene una notación ligeramente distinta. *)
                                 (* U no es de 4×4, por tanto no es ortogonal, sino de 3×4 y *)
                                 (* sigma es cuadrada con los valores singulares en la diagonal.*)
Transpose[Ut] . Ut               (* Ut no es ortogonal, pero ((Ut)^T)Ut = I_3 *)
Transpose[V] . V                 (* V es ortogonal. *)
Transpose[Ut] . DiagonalMatrix[sig] . V   (* El producto resultante es A. *)
(* Ejercicio 16 - parcial *)
psA = PseudoInverse[A]           (* Cálculo de seudoinversa en un paso. *)
N[A]                             (* Aproximado. *)
A . psA . A - A                  (* Comprobación de todas *)
psA . A . psA - psA              (* las condiciones de *)
Transpose[A . psA] - A . psA     (* Moore-Penrose. Todas las matrices *)
Transpose[psA . A] - psA . A     (* son cero. *)
(* Ejercicio 17 *)
Integrate[x^2, Sin[x], {x, 0, 1}]     (* <x^2, sen(x)>. *)
Sqrt[Integrate[x^2, {x, 0, 1}]]       (* Norma(x^2). *)
Sqrt[Integrate[Sin[x], {x, 0, 1}]]    (* Norma(sen(x)). *)
N[%% %- %%%]   (* Norma(x^2)*Norma(sen(x))-<x^2, sen(x) es > 0. Correcto. *)
```

Soluciones seleccionadas con *MATLAB*

Comandos: lscov, nnls, orth, norm, normest, pinv, qr, qrdelete, qrinsert, rcond, schur, svd. Y mediante las herramientas simbólicas: singvals.

```
A = [1 4 69; 2 -3 28; -3 2 -37; 4 2 -59]     % Todos los datos.
B = [1 2 3; 4 5 6; 7 8 9]
C = [1 2 3 4; 2 2 3 4; 3 3 3 4; 4 4 4 4]
D = [-1 -1 1; -1 2 4; 1 4 2]
% Ejercicio 1
dot(A(:,1), a(:,2))                 % Etc.. Repetir con los otros dos pares.
A.'*a                 % (A^T)A es diagonal con normas^2 de las columnas de A
norm(A(:, 1))^2, norm(A(:, 2))^2, norm(A(:, 3))^2    % en los elementos diagonales.
% Ejercicios 2 y 3
e4 = [0;0;0;1]        % e4.
A1 = [A e4]           % A1.
A2=orth(A1)           % Proc. Gram-Schmidt sobre las columnas de A.
```

```
A2.'*A2         % ¡La matriz ya es ortogonal! GS también normaliza.
% Ejercicio 4   % Izquierda-derecha de desig. Bessel > 0. Correcto.
norm(e4)^2-(dot(e4, A2(:, 1))^2+dot(e4, A2(:, 2))^2+dot(e4, A2(:,3))^2)
% Ejercicio 5
B1 = orth(B)    % GS sobre las columnas de B para formar una nueva matriz B1
[B1 B]          % cuyas columnas generan Col(B), porque las 2 primeras
rref(ans)       % columnas de [B1, B] son columnas pivote.
% Ejercicio 6
function [A] = proj(u,lis)   % En un archivo m llamado proy.m teclear
     [m,n]=size(lis);        % el programa de la izquierda, para
     s = zeros(1,n);         % la proyección ortogonal.
     for i=1:m
       s = s + dot(u,lis(i,:))/dot(lis(i,:),lis(i,:))*lis(i,:);
     end
     A = s;
% Ejercicio 7
proj ([1 -1 2], [-1 4 1; 5 1 1])
% Ejercicio 10
[Q,R] = qr(C)    % Factorización QR.
C - Q*R          % La diferencia es aproximadamente la matriz cero,
Q.' * Q          % y Q es ortogonal, porque (Q^T)Q=I.
% Ejercicio 11
A = [1 0; 1 1; 1 2; 1 3; 1 4; 1 5; 1 6; 1 7; 1 8; 1 9]
b = [0; 1; 4; 9; 16; 25; 36; 49; 64; 81]   % No se necesitan ecuaciones normales; lscov lo hace
lscov(A, b, diag(ones)10,1)))              % en un paso. Revisarlo. También revisar nnls.
x = 0:.1:9;      % Grafica la recta de mínimos cuadrados
y1 = 9*x-12;     % 9x-12 y x^2 en una figura
y2 = x.^2;       % sobre [0, 9].
plot(x,y1,x,y2);
% Ejercicio 13
[D1,Q]=eig(D)              % La normalización ortogonal se hace en un paso.
                           % D1 es diagonal con los eigenv. en la diagonal,
Q' * Q                     % y Q es ortogonal.
Q' * D * Q                 % El producto es D, correcto.
% Ejercicio 14
eig[B]                     % Primero se revisan los eigenvalores. Todos son reales. Correcto.
[Q, T] = schur(B)          % Descomposición de Schur, Q y T.
Q' * Q                     % Q es ortogonal.
Q * T * Q'                 % Se revisó y obtuvo B.
% Ejercicio 15 - parcial
[U,S,V] = svd(A)           % Descomposición en SVD en un paso.   Revisando...
U' * U, V' * V             % U y V son ortogonales.
U*S*V'                     % El producto es A.
% Ejercicio 16 - parcial
psA = pinv(A)              % Cálculo de la seudoinversa en un paso.
```

```
A * psA * A - A            % Comprobando todas las
psA * A * psA - psA        % condiciones de Moore-
(A * psA)' - A * psA       % Penrose. Todas las matrices
(psA * A)' - psA * A       % son aproximadamente cero.

% Ejercicio 17

ff = eval(int('x^2*sin(x)',0,1))        % (ST) <x^2, sen(x)>.
f1 = sqrt(eval(int('x^2',0,1)))% (ST) Norma(x^2).
f2 = sqrt(eval(int('sin(x)',0,1)))      % (ST) Norma(sen(x)).
f1*f2-ff                          % Norma(x^2)*Norma(sen(x))-<x^2, sen(x)> es > 0. Correcto.
```

A

Álgebra lineal con números complejos

En la actualidad, casi en todos los campos de matemáticas, física e ingeniería se usan los números complejos. Fueron inventados por Cardano[1] y mencionados por primera vez en su libro *Ars Magna*[2] (en 1545). Sin embargo, de acuerdo con G. H. Hardy, Gauss "fue el primer matemático que usó números complejos en una forma realmente confiable y científica".

En este apéndice *esbozaremos* el procedimiento para incorporar los números complejos en el álgebra lineal. Todos los procesos son iguales a los anteriores, excepto que se usará aritmética compleja.

Aritmética con números complejos

La **unidad imaginaria** i, o $\sqrt{-1}$, se define por la propiedad

$$i^2 = -1$$

Por consiguiente,

$$i^3 = -i, \qquad i^4 = 1, \qquad i^5 = i$$

■ EJEMPLO 1 Calcular i^{1246}.

SOLUCIÓN $i^{1246} = i^{311\cdot 4+2} = (i^4)^{311}\, i^2 = 1^{311}\,(-i) = -i$ $\square$

---

[1] **Girolamo Cardano** (1501-1576) nació en Pavia, Italia, y fue hijo ilegítimo de un abogado. Pasó su niñez en la extrema pobreza y con mala salud; después estudió medicina y finalmente fue contratado como profesor de matemáticas en Milán. Se hizo famoso como matemático, médico y astrólogo. En su libro *Ars magna* aparecieron por primera vez las soluciones completas de las ecuaciones cúbicas y cuádricas.

[2] En él, Cardano divide a 10 en partes cuyo producto es 40. La respuesta es $5 + 15i$ y $5 - 15i$, como puede comprobarse al resolver la ecuación cuadrática resultante. Cardano escribe la respuesta en la forma 5 p: Rm: 15 y 5 m: Rm: 15.

Un **número complejo** z es una expresión de la forma $z = a + bi$, en la que a y b son números reales. El conjunto de los números complejos está representado por **C**. La **parte real**, $\text{Re}(z)$, de z es a. La **parte imaginaria**, $\text{Im}(z)$, de z es b. Si $b = 0$, z es un número real. Si $a = 0$, z es número imaginario puro. El **complejo conjugado** de z es $\bar{z} = a - ib$.

■ EJEMPLO 2

$$\text{Re}(1 - 2i) = 1, \qquad \text{Im}(5 - 2i) = -2, \qquad \overline{1 - i} = 1 + i, \qquad \overline{-3} = -3$$

Dos números complejos son **iguales** si sus partes reales e imaginarias respectivas lo son también. Por ejemplo, $5 + xi = y - 4i$ si y sólo si $y = 5$ y $x = -4$.

El **valor absoluto** $|z| = |a + bi|$ de un número complejo z es el número real no negativo $\sqrt{a^2 + b^2}$.

■ EJEMPLO 3

$$|-2 + 3i| = \sqrt{(-2)^2 + 3^2} = \sqrt{13}$$

Observamos que

$$z\bar{z} = |z|^2$$

La **suma, diferencia** y **producto** de los números complejos se lleva a cabo como en los números reales, con las siguientes observaciones: se calculan todas las potencias de i. Se agrupan los términos de modo que el resultado final tenga la forma $a + ib$ con a y b reales.

■ EJEMPLO 4

$$(1 - 2i) - (2 + 3i)(-1 + i) = (1 - 2i) - (-5 - i) = 6 - i$$

El **cociente**, z/w, de dos números complejos $z = a + bi$ y $w = c + di$ siendo $c + di \neq 0$, es el número

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}i$$

Es fácil comprobar que $w(\frac{z}{w}) = z$.

■ EJEMPLO 5

$$\frac{2 + 3i}{1 + 2i} = \frac{(2 + 3i)(1 - 2i)}{(1 + 2i)(1 - 2i)} = \frac{8 - i}{5} = \frac{8}{5} - \frac{1}{5}i$$

Se deja como ejercicio comprobar las propiedades siguientes.

$$z + \bar{z} = 2\text{Re}(z) \qquad z - \bar{z} = 2\text{Im}(z)i$$

$$\overline{z + w} = \bar{z} + \bar{w}, \qquad \overline{z - w} = \bar{z} - \bar{w}, \qquad \overline{zw} = \bar{z}\,\bar{w}, \qquad \overline{z/w} = \bar{z}/\bar{w}$$

Interpretación geométrica de los números complejos

Todo número complejo $z = a + ib$ puede representarse con el vector (o el punto) (a, b) en el plano. En este contexto, los ejes x y y se llaman eje **real** y eje **imaginario**, respectivamente. El opuesto $-z$ es la reflexión de z con respecto al origen, y el conjugado es la reflexión en el eje real. La suma de dos números complejos corresponde a la suma vectorial en $\mathbf{R}^2$ y la multiplicación por un número **real** corresponde a la multiplicación por escalar en $\mathbf{R}^2$ (figura A.1).

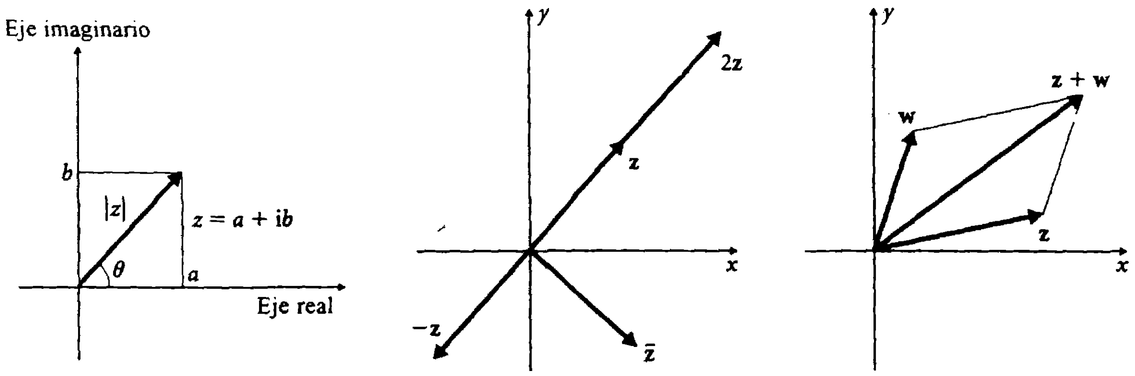


Figura A.1 Números complejos como vectores-z.

Geométricamente, el valor absoluto $|z|$ es la longitud del vector z. El ángulo θ que forman el eje real positivo y el vector (a, b) que representa a $z = a + ib$ se llama **argumento** de z. Como

$$a = |z| \cos \theta, \qquad b = |z| \operatorname{sen} \theta$$

entonces

$$z = |z| (\cos \theta + i \operatorname{sen} \theta) \tag{A.1}$$

A la ecuación (A.1) se le llama **representación polar** de z.

■ EJEMPLO 6 Determine la representación polar de $-1 + i$.

SOLUCIÓN En primer lugar $|-1 + i| = \sqrt{2}$. El argumento de $-1 + i$ puede calcularse a partir de

$$\sqrt{2} \cos \theta = -1, \qquad \sqrt{2} \operatorname{sen} \theta = 1$$

que implican $\theta = 3\pi/4$. Por consiguiente,

$$-1 + i = \sqrt{2}\left(\cos \frac{3\pi}{4} + i \operatorname{sen} \frac{3\pi}{4}\right)$$

Las representaciones polares de los números complejos son muy útiles, en especial en asuntos relacionados con la multiplicación y la división compleja. Por ejemplo, si $w = |w| (\cos \phi + i \operatorname{sen} \phi)$, el producto zw y el cociente z/w tienen las representaciones polares siguientes:

$$zw = |z|\,|w|\,(\cos(\theta + \phi) + i\ \text{sen}(\theta + \phi))$$

$$\frac{z}{w} = \frac{|z|}{|w|}(\cos(\theta - \phi) + i\ \text{sen}\ (\theta - \phi))$$

Estas identidades pueden demostrarse mediante las identidades trigonométricas normales que expresan el seno y el coseno de la suma o diferencia de dos ángulos. También las representaciones polares de las potencias pueden determinarse con facilidad:

$$z^n = |z|^n\,(\cos n\theta + i\ \text{sen}\ n\theta)$$

■ EJEMPLO 7 Escriba $(-1 + i)^{10}$ en la forma $a + ib$.

SOLUCIÓN Tenemos que

$$\begin{aligned}(-1+i)^{10} &= \left(\sqrt{2}\right)^{10}\left(\cos\left(10\cdot\frac{3\pi}{4}\right) + i\ \text{sen}\left(10\cdot\frac{3\pi}{4}\right)\right)\\ &= 32\left(\cos\frac{15\pi}{2} + i\ \text{sen}\ \frac{15\pi}{2}\right)\\ &= -32i\end{aligned}$$

Sistemas de ecuaciones con números complejos

La solución de sistemas de ecuaciones con números complejos se hace igual que con la interpretación geométrica.

■ EJEMPLO 8 Resuelva el sistema siguiente para determinar z y w.

$$3iz + 4w = 5 + 15i$$

$$(5 - i)z + (3 - 4i)w = 24 + 5i$$

SOLUCIÓN Con eliminación de Gauss,

$$\begin{bmatrix} 3i & 4 & 5+15i \\ 5-i & 3-4i & 24+5i \end{bmatrix} \sim \begin{bmatrix} 3i & 4 & 5+15i \\ 0 & \frac{13}{3}+\frac{8}{3}i & \frac{2}{3}+\frac{55}{3}i \end{bmatrix}$$

$$\begin{bmatrix} 3i & 4 & 5+15i \\ 0 & 1 & 2+3i \end{bmatrix} \sim \begin{bmatrix} 3i & 0 & -3+3i \\ 0 & 1 & 2+3i \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1 & 2+3i \end{bmatrix}$$

Así, la solución del sistema es $z = 1 + i$ y $w = 2 + 3i$.

Matrices de números complejos

Toda la aritmética y los teoremas matriciales que hemos expuesto se aplican a matrices cuyos elementos son números complejos.

■ EJEMPLO 9 Calcule $A^TA-(1+i)B$, siendo

$$A=\begin{bmatrix} i & 2 \\ 1-i & 2i \\ 2 & -i \end{bmatrix}, \qquad B=\begin{bmatrix} 1 & 0 \\ -i & 1-i \end{bmatrix}$$

SOLUCIÓN

$$\begin{bmatrix} 3-2i & 2+2i \\ 2+2i & -1 \end{bmatrix}-\begin{bmatrix} 1+i & 0 \\ 1-i & 2 \end{bmatrix}=\begin{bmatrix} 2-3i & 2+2i \\ 1+3i & -3 \end{bmatrix}$$ □

■ EJEMPLO 10 Calcule A^{-1} mediante reducción con operaciones de renglón.

$$A=\begin{bmatrix} i & 1 \\ 1 & 1-i \end{bmatrix}$$

SOLUCIÓN $[A:I]$ se reduce como sigue:

$$\begin{bmatrix} i & 1 & 1 & 0 \\ 0 & 1 & i & 1 \end{bmatrix}\sim\begin{bmatrix} i & 0 & 1-i & -1 \\ 0 & 1 & i & 1 \end{bmatrix}\sim\begin{bmatrix} 1 & 0 & -1-i & i \\ 0 & 1 & i & 1 \end{bmatrix}$$

Asi, $A^{-1}=\begin{bmatrix} -1-i & i \\ i & 1 \end{bmatrix}$. □

La **compleja conjugada** $\overline{A}$ de A es la matriz cuyos elementos son los complejos conjugados de los elementos correspondientes de A.

■ EJEMPLO 11

$$\overline{\begin{bmatrix} i & 0 \\ -1 & 1+i \end{bmatrix}}=\begin{bmatrix} -i & 0 \\ -1 & 1-i \end{bmatrix}$$ □

Es fácil comprobar las siguientes propiedades:

$$\overline{\overline{A}}=A, \qquad \overline{A\pm B}=\overline{A}\pm\overline{B}, \qquad \overline{AB}=\overline{A}\,\overline{B}, \qquad \overline{A^{-1}}=\left(\overline{A}\right)^{-1}, \qquad \overline{A^T}=\left(\overline{A}\right)^T$$

La **parte real**, Re(A) y la **parte imaginaria**, Im(A), de una matriz A son las matrices cuyos elementos son las partes reales e imaginarias de los elementos correspondientes de A.

■ EJEMPLO 12 Sea

$$A=\begin{bmatrix} i & 2 \\ 1-i & -i \end{bmatrix}$$

Entonces

$$\operatorname{Re}(A)=\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}, \qquad \operatorname{Im}(A)=\begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}$$ □

El cálculo de los determinantes se hace de la misma manera que con elementos reales.

■ EJEMPLO 13

$$\begin{vmatrix} i & 1+i & 1 \\ 2i & 0 & i \\ -2i & 2 & -i \end{vmatrix} = -(1+i)\begin{vmatrix} 2i & i \\ -2i & -i \end{vmatrix} - 2\begin{vmatrix} i & 1 \\ 2i & i \end{vmatrix} = 2+4i$$

□

Algunas matrices cuadradas especiales

Cierto tipo de matrices, que hoy se denominan hermitianas, fueron introducidas por el matemático francés Hermite.[3] Son útiles en ingeniería, matemáticas y física, en especial en física atómica.

DEFINICIÓN

Sea A una matriz cuadrada. A es **hermitiana** si cumple la igualdad $\overline{A}^T = A$. Y es **antihermitiana** si $\overline{A}^T = -A$.

■ EJEMPLO 14 A es hermitiana y B es antihermitiana.

$$A = \begin{bmatrix} -1 & 2i & 3 \\ -2i & -2 & -i \\ 3 & i & -3 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 2-i \\ -2-i & i \end{bmatrix}$$

□

El teorema siguiente resume las propiedades básicas de las matrices hermitianas y antihermitianas. Se deja como ejercicio demostrarlo.

TEOREMA 1

1. La diagonal principal de una matriz hermitiana está formada por números reales.
2. La diagonal principal de una matriz antihermitiana está formada por ceros o por números imaginarios puros.
3. Una matriz que al mismo tiempo es hermitiana y antihermitiana es una matriz cero.
4. Si A y B son hermitianas, también los son $A + B$, $A - B$ y cA para cualquier escalar real c.
5. Si A y B son antihermitianas, también lo son $A + B$, $A - B$ y cA, para cualquier escalar real c.

Veamos el teorema siguiente.

TEOREMA 2

Toda matriz cuadrada A puede escribirse en forma única como la suma de una matriz hermitiana S y una matriz antihermitiana R. Para ser más precisos,

$$A = H + R \quad \text{siendo} \quad H = \frac{1}{2}(A + \overline{A}^T), \qquad R = \frac{1}{2}(A - \overline{A}^T)$$

[3] **Charles Hermite** nació en Dieuze, Lorena, Francia, en 1822. Estudió matemáticas por cuenta propia. Se le conoce por ser autor de muchos maravillosos resultados, aunque relativamente técnicos. Demostró que el número e es trascendente y que las raíces de los polinomios arbitrarios de quinto grado pueden expresarse en términos de funciones modulares elípticas.

También se deja como ejercicio su demostración.

■ EJEMPLO 15

$$A = \begin{bmatrix} 4 & 1-i \\ -1-3i & i \end{bmatrix} = \begin{bmatrix} 4 & i \\ -i & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1-2i \\ -1-2i & i \end{bmatrix} = H + R \qquad \square$$

DEFINICIÓN

Una matriz cuadrada A es unitaria si $\overline{A}^T = A^{-1}$ o, de manera equivalente, si $\overline{A}^T A = I$.

■ EJEMPLO 16 Demuestre que las siguientes matrices son unitarias.

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2}i \\ -\frac{\sqrt{3}}{2}i & \frac{1}{2} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & i & 0 \\ 0 & 0 & i \\ 1 & 0 & 0 \end{bmatrix}$$

SOLUCIÓN

$$\overline{A}^T A = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2}i \\ \frac{\sqrt{3}}{2}i & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2}i \\ -\frac{\sqrt{3}}{2}i & \frac{1}{2} \end{bmatrix} = I_2$$

De igual manera, $\overline{B}^T B = I_3$. $\square$

■ EJEMPLO 17 Calcular la inversa de la matriz unitaria.

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

SOLUCIÓN

$$A^{-1} = \overline{A}^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \square$$

OBSERVACIÓN Si una matriz cuadrada es real, entonces $\overline{A} = A$. En este caso, las relaciones $\overline{A}^T = A$ y $\overline{A}^T = A^{-1}$ se reducen a $A^T = A$ y $A^T = A^{-1}$. Por consiguiente,

- Una matriz hermitiana real es simétrica.
- Una matriz real unitaria es ortogonal.

No pierda de vista que para una matriz antihermitiana *real* A, se cumple $\overline{A} = -A$. Esa matriz se llama **antisimétrica**.

B

Instrucciones de álgebra lineal

En este apéndice hemos reunido **la mayor parte** de las instrucciones del álgebra lineal que se usan en Maple, Mathematica y MATLAB.

Maple

Casi todas las instrucciones de álgebra lineal en Maple se encuentran en el paquete linalg, que se puede cargar con:

```
with(linalg);
```

Los comandos del paquete linalg

GramSchmidt	JordanBlock	LUdecomp	QRdecomp	addcol
addrow	adjoint	angle	augment	backsub
band	basis	bezout	blockmatrix	charmat
charpoly	cholesky	col	coldim	colspace
colspan	companion	cond	copyinto	crossprod
curl	definite	delcols	delrows	det
diag	diverge	dotprod	eigenvals	eigenvects
entermatrix	equal	exponential	extend	ffgausselim
fibonacci	forwardsub	frobenius	gausselim	gaussjord
geneqns	genmatrix	grad	hadamard	hermite
hessian	hilbert	htranspose	ihermite	indexfunc
innerprod	intbasis	inverse	ismith	issimilar
iszero	jacobian	jordan	kernel	laplacian
leastsqrs	linsolve	matadd	matrix	minor
minpoly	mulcol	multiply	norm	normalize
orthog	permanent	pivot	potential	randmatrix
randvector	rank	references	row	rowdim
rowspace	rowspan	scalarmul	singularvals	smith
stack	submatrix	subvector	sumbasis	swapcol
swaprow	sylvester	toeplitz	trace	transpose
vandermonde	vecpotent	vectdim	vector	wronskian

Véanse también array, evalm, identity, list, with, range, table &*, ^, y &^.

Mathematica

Comandos de operaciones matriciales

Det	Dot	Eigensystem	Eigenvalues
Eigenvectors	Inverse	LatticeReduce	LinearProgramming
LinearSolve	MatrixExp	MatrixPower	Minors
NullSpace	Outer	PseudoInverse	QRDecomposition
RowReduce	SchurDecomposition	SingularValues	Transpose

Paquetes de álgebra lineal

Package Name	Commands
LinearAlgebra'Cholesky'	CholeskyDecomposition
LinearAlgebra'CrossProduct'	Cross
LinearAlgebra'GaussianElimination'	LU, LUFactor, LUSolve
LinearAlgebra'MatrixManipulation'	AppendColumns, AppendRows, BlockMatrix, HankelMatrix, HilbertMatrix, LowerDiagonalMatrix, SquareMatrixQ, SubMatrix, TakeColumns, TakeMatrix, TakeRows, TridiagonalMatrix, UpperDiagonalMatrix, ZeroMatrix
LinearAlgebra'Orthogonalization'	GramSchmidt, InnerProduct, Normalize, Normalized, Projection
LinearAlgebra'Tridiagonal'	TridiagonalSolve

Véanse también Array, ColumnForm, DiagonalMatrix, Dimensions, IdentityMatrix, Length, List, MatrixForm, MatrixPower, Range, Table, y . (Dot).

Un paquete como el de LinearAlgebra'Cholesky' se puede cargar con

<<LinearAlgebra'Cholesky'

MATLAB

Álgebra lineal numérica

\	/	balance	cdf2rdf	chol
cond	det	diag	eig	eye
fliplr	flipud	hess	inv	linspace
logspace	lscov	lu	meshgrid	nnls
norm	null	ones	orth	pinv
poly	polyeig	qr	qrdelete	qrinsert
qz	rand	randn	rank	rcond
reshape	rot90	rref	rsf2csf	schur
svd	trace	tril	triu	zeros

Herramientas de matemática simbólica (álgebra lineal y operaciones)

`charpoly`	`colspace`	`determ`	`eigensys`
`inverse`	`jordan`	`linsolve`	`nullspace`
`numeric`	`singvals`	`solve`	`subs`
`sym`	`sym2poly`	`symadd`	`symdiv`
`symmul`	`symop`	`sympow`	`symsize`
`symsub`	`symvar`	`transpose`	

Véanse también poly, polyval, polyvalm y roots.

Respuestas a los ejercicios seleccionados

Capítulo 1

Sección 1.1

1. (a)–(e) Lineal. (f) No lineal

   Sólo (c) es homogénea.

3. **(a1)** x, y **(a2)** y **(a3)** x **(b1)** x, y **(b2)** Ninguna **(b3)** x, y **(c1)** x, y, z **(c2)** x **(c3)** y, z **(d1)** x, y, z **(d2)** x **(d3)** y, z **(e1)** x, y, z, w, t **(e2)** x **(e3)** y, z, w, t

5. Sólo S está en el plano.

7. (a) Si $a = -2$, número infinito de soluciones; si $a = 2$, no hay soluciones; si $a \neq \pm 2$, la única solución es $x = 1/(a-2)$.

   (b) Si $a = \pm 2$, no hay soluciones; si $a \neq \pm 2$, la solución única es $x = 3/(a^2 - 4)$.

   (c) Si $a = \pm 2$, número infinito de soluciones; si $a \neq \pm 2$, la única solución es $x = 0$.

   (d) Si $a = 0$, número infinito de soluciones expresadas por $x = r, y = t$; si $a \neq 0$, tiene un número infinito de soluciones expresadas por $x = 3 + ar, y = r$.

9. La forma canónica es

$$\begin{aligned} 2x + 4z &= -1 \\ -x + 2z + 2w &= 2 \\ -2x - z + 3w &= -3 \\ y + z + t - w &= 4 \end{aligned}$$

(a)–(c) La matriz de coeficientes y el vector de constantes son:

$$\begin{bmatrix} 2 & 0 & 4 & 0 & 0 \\ -1 & 0 & 2 & 0 & 2 \\ -2 & 0 & -1 & 0 & 3 \\ 0 & 1 & 1 & 1 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 2 \\ -3 \\ 4 \end{bmatrix}$$

La matriz aumentada es

$$\left[\begin{array}{ccccccc} 2 & 0 & 4 & 0 & 0 & : & -1 \\ -1 & 0 & 2 & 0 & 2 & : & 2 \\ -2 & 0 & -1 & 0 & 3 & : & -3 \\ 0 & 1 & 1 & 1 & -1 & : & 4 \end{array}\right]$$

(d) El sistema homogéneo asociado es

$$\begin{aligned} 2x + 4z &= 0 \\ -x + 2z + 2w &= 0 \\ -2x - z + 3w &= 0 \\ y + z + t - w &= 0 \end{aligned}$$

11. Solución general: $(-2t - 2r - s, t, 2r - 1, s/2, s, r)$ para $r, s, t, \in \mathbf{R}$.

13. $x_6 = r, x_5 = r, x_3 = r, x_4 = s, x_2 = t, x_1 = 1 + t - 6r + 5s$ para todas r, s y t.

15. $P(\frac{1}{6}, \frac{5}{6})$.

17. En el segundo sistema la primera ecuación es un múltiplo de la primera ecuación del primer sistema. La última ecuación es la suma de las tres ecuaciones del primer sistema.

19. $x = -1, y = 2, z = 2$

21. No hay soluciones.

23. $(r - 1, 2r + 3, r)$, $r \in \mathbf{R}$

25. $(-\frac{r}{3} - 9, \frac{r}{3} + 3, -6, r)$, $r \in \mathbf{R}$

27. $(1 - r, r, 1 - r, r)$, $r \in \mathbf{R}$

29. $(\frac{37}{4}, \frac{17}{4}, \frac{11}{4})$
31. $(1, 0, 3)$
33. $(3, -2, -4)$

35. $\theta = (2k - 1)\pi$, siendo k cualquier entero.
41. 32
43. 16 días.
45. Los ángulos son 100°, 80°, 100°, 80°.
47. $\frac{14}{3}$ y $\frac{20}{3}$, o 6 y 4.
49. 14 renglones.
51. Los botes recorrieron 24 millas.

Sección 1.2

1. **(a)** No está en forma de escalón. **(b)** Forma reducida de escalón. **(c)** No está en forma de escalón. **(d)** Forma escalonada de renglón, pero no reducida.
3. **(a)** Forma escalonada de renglón, pero no reducida. **(b)** Forma escalonada de renglón, pero no reducida. **(c)** Forma escalonada reducida. **(d)** No es forma escalonada.
5. **(a)** No es forma escalonada. **(b)** Forma escalonada de renglón, pero no reducida. **(c)** Forma escalonada reducida. **(d)** Forma escalonada reducida.
9. Tanto A como B tienen I_3 como forma escalonada reducida. Por lo que, $A \sim I$ y $B \sim I$. Entonces $A \sim I$ e $I \sim B$, de acuerdo con el ejercicio 8. Por tanto, según el ejercicio 7, $A \sim B$.
11. Sí. La segunda matriz es la forma escalonada reducida de la primera.
13. Cierto, porque las formas escalonadas reducidas, $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix}$ no son iguales.
15. Como $\cos\theta\cos\theta - (-\operatorname{sen}\theta)(\operatorname{sen}\theta) = \cos^2\theta + \operatorname{sen}^2\theta = 1 \neq 0$, la matriz se reduce a I, de acuerdo con el ejercicio 14.
17. **(a)** $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

    (b) $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
19. $\begin{bmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$
21. $\begin{bmatrix} 1 & 4 & 0 & 5 & 0 & 6 \\ 0 & 0 & 1 & 4 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$
23. $x = 1, y = 2, z = -2, w = -4$
25. $x_1 = -4t - 5s, x_2 = 6 - 2t - 6s, x_3 = t, x_4 = s, x_5 = -3$
27. No hay solución.
29. $x = 1, y = 2, z = -2, w = -4$
31. $x = r_1 - 2, y = -r_2 + 1, z = r_1 - r_2 - 4, w = r_2, t = r_1$
33. $x = 0, y = 0, z = 0, w = 0$
35. $x = \frac{1}{2}a + \frac{1}{2}b, y = \frac{1}{2}a - \frac{1}{2}b, z = -b$
37. $x_1 = r - 1, x_2 = r, x_3 = 1, x_4 = -2$
39. $x_1 = 6 + r, x_2 = r - 1, x_3 = 4, x_4 = r, x_5 = -1$
41. $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0, x_5 = 0$
43. Los sistemas son equivalentes, porque ambas matrices aumentadas tienen la misma forma escalonada reducida:
$$\begin{bmatrix} 1 & 4 & 0 & 5 & 0 & 6 \\ 0 & 0 & 1 & 4 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$
45. **(a)** La última columna no es pivote, de modo que hay 3 soluciones. Como la tercera columna no es pivote tiene un número infinito de soluciones. **(b)** La última columna es pivote, de manera que no hay soluciones.
47. **(a)** Si la última columna es pivote, no hay soluciones. En cualquier otro caso hay un número infinito de soluciones. **(b)** No hay soluciones. **(c)** Si la última columna es pivote, no existen soluciones. En caso contrario se tiene exactamente una solución. **(d)** No hay soluciones.
53. **(a)** Si $a = 4$, hay un número infinito de soluciones. Si $a \neq 4$, sólo hay una solución.

    (b) Si $a = 3$, hay un número infinito de soluciones. Si $a = -3$, no hay soluciones. Si $a \neq -3, 3$, sólo hay una.

Sección 1.3

1. Sí.
3. No.
5. $5x + y = 14, x - 2y = -6$.
7. $x = 1.99, y = 3.97$; la solución exacta es $x = 2, y = 4$.
9. $x = 1.0011, y = 4.9990, z = -2.0022$
11. $x = 2.0176, y = -1.9952, z = 2.9936$
13. $x = 3.9984, y = -1.0003, z = 2.0005$
15. **(a)** Las 5 primeras iteraciones de Gauss-Seidel son (2,0), (2, −2), (0, −2), (0,0) y (2,0). Como el primer y último

resultados son idénticos, esos valores se repiten a medida que k crece. Por consiguiente, la iteración diverge.

(b) La cuarta y quinta iteraciones de Gauss-Seidel son $(-0.5977, 0.5977)$ y $(-0.6006, 0.6006)$. Por tanto, la iteración converge a $(-0.6, 0.6)$ cuando menos con dos cifras decimales.

17. $x = 3.1, y = 1.1$.

19. $x = 2, y = -1, z = 1$.

Sección 1.4

1. La relación es $B = 3A - 1\,000$, que es lineal.

3. 80 000 yenes, 900 francos y 1 200 marcos.

5. Boston 30°, Nueva York 36°, Montreal 24°.

7. Los volúmenes de las soluciones que contienen A, B y C son 2.0 cm^3, 3.5 cm^3 y 1.8 cm^3.

19. $i_1 = \frac{3}{2} = 1.5$, $i_2 = 1$, $i_3 = \frac{1}{2} = 0.5$ A.

11. $x_1 = \frac{10}{7} = x_3$ y $x_2 = \frac{12}{7}$.

13. El sistema se balancea si $w_1 = 8r$, $w_2 = 2r$, $w_3 = 5r$, $w_4 = r$, siendo r cualquier número positivo real.

15. $y = x^3 + 2x^2 - 3x + 1$.

17. $a = \frac{1}{2}$, $b = -\frac{3}{2}$, y $c = 1$.

19. $A = -\frac{1}{3}$ y $B = \frac{1}{3}$.

Capítulo 2

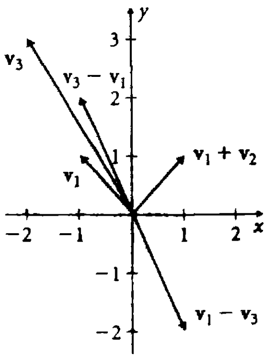
Sección 2.1

1. (a) $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ (b) $\begin{bmatrix} -5 \\ -9 \end{bmatrix}$

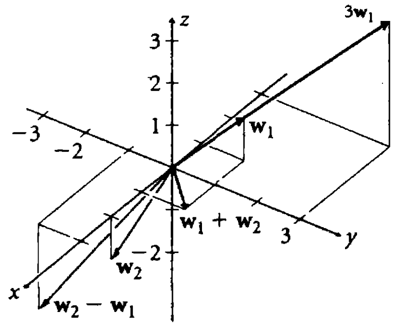
3. (a) $\begin{bmatrix} 4 \\ -9 \\ 4 \end{bmatrix}$ (b) $\begin{bmatrix} 9 \\ -10 \\ 7 \end{bmatrix}$

5. (a) $\begin{bmatrix} -39 \\ 32 \end{bmatrix}$ (b) $\begin{bmatrix} 40 \\ 5 \end{bmatrix}$

7.



9.



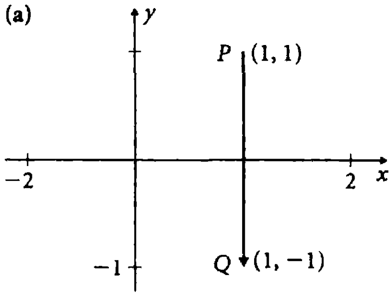
11. (a) $a = 0$, $b = 0$. (b) No hay valores.

13. $\mathbf{x} = \frac{3}{2}\mathbf{a} + 2\mathbf{b}$.

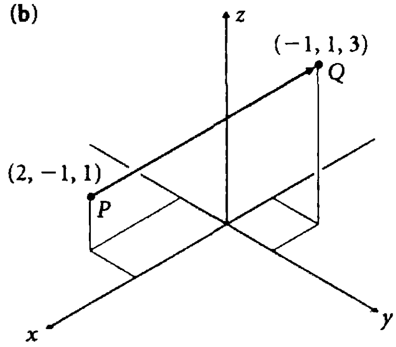
15. $\mathbf{x} = -\mathbf{a} - 3\mathbf{b}$, $\mathbf{y} = -2\mathbf{a} - 4\mathbf{b}$.

17. (a) $PQ = (0, -2)$. (b) $PQ = (-3, 2, 2)$.

(a)



(b)



19. Sí.

21. No.

23. No.

25. Sí.

27. Sí es.

29. No, no es.

31. Sí es.

33. Cierto, según el teorema 2, porque $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & -1 & 1 \end{bmatrix}$ tiene 3 pivotes.

35. $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

37. $\mathbf{R}^2$

39. $\left\{ \begin{bmatrix} r \\ s \\ 0 \end{bmatrix}, r, s \in \mathbf{R} \right\}$

41. $\mathbf{R}^3$

43. Cualquier k distinta de cero.

45. **(a)** $x \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$

(b) $x \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + z \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

47. **(a)** No **(b)** Sí

51. **(a)** Cierto. **(b)** Cierto. **(c)** Falso. **(d)** Cierto. **(e)** Falso. **(f)** Falso. **(g)** Falso. **(h)** Cierto.

53. (10, 20) mi/h (si las direcciones hacia el este y hacia el norte están a lo largo de los ejes x y y positivos).

55. **(a)** (1 100, 1 800, 2 400); un total de \$1 100, \$1 800, \$2 400 se gastan en alimentos por viaje para primera, segunda y tercera clase, respectivamente.

(b) (150, 100, 200); la aerolínea gasta \$150, \$100 y \$200 más para el avión 3 que para el 2 por viaje en primera, segunda y tercera clase, respectivamente.

(c) (4 500, 7 000, 9 000); la aerolínea gasta \$4 500, \$7 000 y \$9 000 en 10 viajes del avión 3 para alimentos en primera, segunda y tercera clase, respectivamente.

(d) (8 900,14 600,19 300); la aerolínea gasta \$8 900, \$14 600 y \$19 300 en alimentos para primera, segunda y tercera clase, respectivamente, en 7 viajes del primer avión, 8 del segundo y 9 del tercero.

Sección 2.2

1. $\|\mathbf{u}\| = 3$, $\|\mathbf{v}\| = 5\sqrt{2}$, $\|\mathbf{w}\| = 2\sqrt{5}$, $\|\mathbf{u}+\mathbf{v}\| = \sqrt{19}$, $\|\mathbf{u}-\mathbf{v}\| = 3\sqrt{11}$, $\|\mathbf{u}-\mathbf{v}+\mathbf{w}\| = \sqrt{139}$, $\|\mathbf{d}\| = 3$, $\|10\mathbf{d}\| = 30$, $\|\|\mathbf{d}\|\,\mathbf{d}\| = 9$

3. $\mathbf{u}\cdot\mathbf{v} = -20$, $\mathbf{w}\cdot\mathbf{u} = 0$, $\mathbf{u}\cdot(\mathbf{v}+\mathbf{w}) = -20$, $\mathbf{v}\cdot\mathbf{u}+\mathbf{w}\cdot\mathbf{u} = -20$, $\mathbf{d}\cdot\mathbf{d} = 9$, $(\mathbf{d}\cdot\mathbf{d})\,\mathbf{d} = \left(-9, -18, 9, 9\sqrt{3}\right)$

5. **(a)** $\left(-\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right)$ **(b)** $\left(\frac{2\sqrt{2}}{5}, -\frac{3\sqrt{2}}{10}, \frac{1\sqrt{2}}{2}\right)$
(c) $\left(-\frac{2\sqrt{5}}{5}, -\frac{\sqrt{5}}{5}, 0\right)$ **(d)** $\left(-\frac{1}{3}, -\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\sqrt{3}\right)$

7. $\left(\frac{2}{3}, -\frac{4}{3}, \frac{4}{3}\right)$

9. $\left(-3, -6, 3, 3\sqrt{3}\right)$

11. **(a)**, **(b)** y **(d)**

13. $\left(\frac{\sqrt{5}}{5}, 0, \frac{2\sqrt{5}}{5}\right)$

15. **(a)** 0.24871. **(b)** 0.3681 rad. **(c)** $\frac{2}{3}\pi$

17. **(a)** $\left(\frac{8}{5}, \frac{16}{5}\right)$ **(b)** $\left(\frac{33}{35}, \frac{64}{35}, \frac{9}{7}\right)$ **(c)** $\left(-2, -1, \frac{3}{10}, \frac{1}{10}\right)$

21. No

25. El círculo de radio 1 centrado en el origen del plano xy.

29. $\|\mathbf{v}\| = \|c\mathbf{u}\| = |c|\,\|\mathbf{u}\| = c1 = c$, porque $c > 0$ y $\mathbf{u}$ es unitaria. Por consiguiente $\mathbf{u} = \frac{1}{c}\mathbf{v} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = (\cos\alpha, \cos\beta, \cos\gamma)$, según la ecuación (2.10).

31. **(a)** $(1, 1, 1) = \sqrt{3}\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)$. Cada coseno director es $\frac{\sqrt{3}}{3}$ y cada ángulo director es de 0.95532 rad.

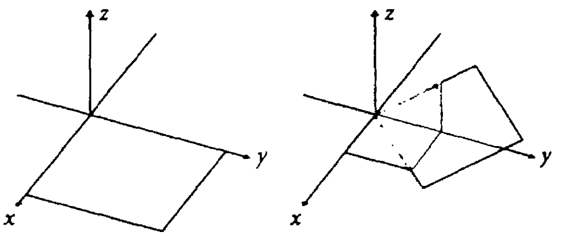
(b) $(-1, 1, -1) = \sqrt{3}\left(-\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}\right)$ con los cosenos directores $-\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}$ y ángulos directores de 2.1863, 0.95532 y 2.1863 rad.

(c) $(\sqrt{3}, 1, 0) = 2\left(\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right)$, con cosenos directores $\frac{\sqrt{3}}{2}, \frac{1}{2}$, y ángulos directores $\frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}$,

(d) $(2, -2, 0) = 2\sqrt{2}\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0\right)$ con cosenos directores $\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}$, 0 y ángulos directores $\frac{\pi}{4}, \frac{3\pi}{4}, \frac{\pi}{2}$.

Sección 2.3

1. (a) Falso. (b) Cierto. (c) Cierto. (d) Cierto.
3. (a) Sí. (b) Sí. (c) Sí. (d) Sí. (e) Sí. (f) Sí. (g) Sí.
5. (a) Sí. (b) No. (c) No.
7. (a) Sí. (b) No.
9. (a) Sí. (b) No.
11. (a) No. (b) No.
13. (a) Sí.
15. (a) Falso (b) Cierto (c) Falso (d) Cierto (e) Cierto (f) Falso (g) Cierto.
21. (a) No. (b) Sí.
23. $\left\{\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}\right\}$
25. Que el sistema $[A : \mathbf{b}]$ es consistente.
27. Toda $x \neq -1$.
29. Los generadores correspondientes son los planos sombreados en las figuras siguientes.



(a) (b)

Sección 2.4

1. Linealmente independiente.
3. Linealmente independiente.
5. Linealmente independiente.
7. Linealmente dependiente.
9. Linealmente independiente.
11. Linealmente dependiente.
13. Linealmente dependiente.
15. Linealmente dependiente.
17. Linealmente independiente.
19. $a = -1, 2$.
21. Sí.
23. $c_1 = -1, c_2 = -1, c_3 = -2$

31. La intersección es una recta que pasa por el origen si los planos son distintos, o es el plano común si los planos coinciden.

Sección 2.5

1. $A\mathbf{u} = \begin{bmatrix} -10 \\ -4 \\ 23 \end{bmatrix}$ $A\mathbf{v}$ y $A\mathbf{w}$ son indefinidos.
3. Si $\mathbf{x} = (x_1, x_2)$, el sistema es $-3x_1 - 2x_2 = 100$, $-x_1 = 200$, $5x_1 - 3x_2 = 300$.
5. $\begin{bmatrix} 1 & -7 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \end{bmatrix}$.
7. $n = 2$; el conjunto es todo $\mathbf{R}^2$.
9. $n = 2$; el conjunto es todo $\mathbf{R}^2$.
11. (33, 44, 40); habría 33 alumnos en 1999, 44 en la primavera del año 2000 y 40 en el otoño de ese mismo año.
13. 6 y 18.
15. **u**.
17. El conjunto solución del sistema es $S = \left\{\begin{bmatrix} 10 + 3r \\ r \end{bmatrix}, r \in \mathbf{R}\right\}$. $\mathbf{p} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$ es una solución particular. El espacio nulo de la matriz de coeficientes es $N = \left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}, r \in \mathbf{R}\right\}$. Es claro que $S = \mathbf{p} + N$, como afirma el teorema 19.

Sección 2.6

1. $(4, -3, -5)$, $(-10, 20, -20)$,
   $(-80, -40, 0)$, $(6, -12, -10)$,
   $(6, -12, -10)$, $(9, -18, -22)$
3. $\begin{matrix} -5 & 5 \\ 5 & -5 \end{matrix}$
5. $\|\mathbf{u} \times \mathbf{v}\|^2 = \|(-29, -18, -15)\| = 1390$, y $\|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 = 26(61) - 196 = 1390$
7. $\dfrac{1}{\sqrt{2249}}(16, 12, 43)$
9. $\sqrt{6}$
11. 40
13. (a) Sí son. (b) No son.
15. No.

Sección 2.7

1. P y Q.
3. (5, –4, 2), (2, –2, 1), (8, –6, 3)

5. l_1 y l_2

7. El punto de intersección es (11, −8, 4).

9. l_1: $\dfrac{x-5}{-3} = \dfrac{y+4}{2} = \dfrac{z-2}{-1}$;

l_2: $\dfrac{x-1}{6} = \dfrac{y-6}{-4} = \dfrac{z-8}{2}$;

l_3: $\dfrac{x-5}{-1} = \dfrac{y+7}{-1} = \dfrac{z-11}{1}$;

l_4: $\dfrac{x-14}{1} = \dfrac{y+2}{2} = \dfrac{z-13}{3}$

11. $x = 5 + 4t, y = -4 - 3t, z = 2 + t$

13. $x = 5 - 3t, y = -4 + 2t, z = 2 - t$

15. $\dfrac{x-3}{-4} = \dfrac{y+1}{3} = \dfrac{z+2}{7}$

17. P y Q

19. $-6(x-5) + 4(y+4) + 5(z-2) = 0$

21. $-6x + 4y + 5z + 36 = 0$

23. $x + 2y + z + 1 = 0$

25. $x + y - z + 1 = 0$

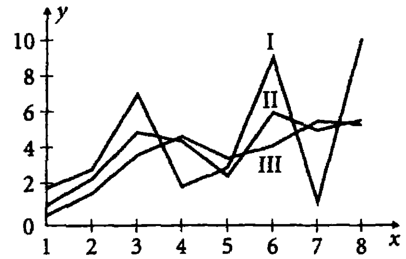
27. $x - 6y + 2z - 21 = 0$

29. $\dfrac{6}{\sqrt{114}}$

31. $-x_1 + 3x_2 - 2x_3 + 8x_4 + 4x_5 + 3 = 0$

Sección 2.8

1. Al promediar una vez se obtiene $(1, \frac{5}{2}, 5, \frac{9}{2}, \frac{5}{2}, 6, 5, \frac{11}{2})$, y al promediar dos veces se obtiene $(\frac{1}{2}, \frac{7}{4}, \frac{15}{4}, \frac{19}{4}, \frac{7}{2}, \frac{17}{4}, \frac{11}{2}, \frac{21}{4})$. En la figura siguiente, la línea I es la original, la II es promediando una vez, y la línea III es promediando dos veces.



3. $\begin{bmatrix} Y_{k+1} \\ A_{k+1} \end{bmatrix} = \begin{bmatrix} 2 & 10 \\ \frac{4}{5} & 0 \end{bmatrix} \begin{bmatrix} Y_k \\ A_k \end{bmatrix}$; condición inicial: $\begin{bmatrix} Y_0 \\ A_0 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$;

después de tres unidades de tiempo: $\begin{bmatrix} Y_3 \\ A_3 \end{bmatrix} = \begin{bmatrix} 16\,000 \\ 2\,560 \end{bmatrix}$

5. $\begin{bmatrix} Y_{k+1} \\ A_{k+1} \end{bmatrix} = \begin{bmatrix} 4 & 10 \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} Y_k \\ A_k \end{bmatrix}$; condición inicial: $\begin{bmatrix} Y_0 \\ A_0 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$;

después de tres unidades de tiempo: $\begin{bmatrix} Y_3 \\ A_3 \end{bmatrix} = \begin{bmatrix} 30\,900 \\ 2\,500 \end{bmatrix}$.

7. $\begin{bmatrix} Y_{k+1} \\ A_{k+1} \end{bmatrix} = \begin{bmatrix} 3 & 12 \\ \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} Y_k \\ A_k \end{bmatrix}$; condición inicial: $\begin{bmatrix} Y_0 \\ A_0 \end{bmatrix} = \begin{bmatrix} 300 \\ 100 \end{bmatrix}$;

después de tres unidades de tiempo: $\begin{bmatrix} Y_3 \\ A_3 \end{bmatrix} = \begin{bmatrix} 30\,900 \\ 2\,500 \end{bmatrix}$.

9. $\begin{bmatrix} A_{k+1} \\ B_{k+1} \\ C_{k+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{5}{2} & \frac{3}{2} \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} A_k \\ B_k \\ C_k \end{bmatrix}$; condición inicial: $\begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} =$

$\begin{bmatrix} 4\,800 \\ 4\,800 \\ 4\,800 \end{bmatrix}$; después de seis semanas: $\begin{bmatrix} A_3 \\ B_3 \\ C_3 \end{bmatrix} = \begin{bmatrix} 15\,975 \\ 2\,625 \\ 1\,700 \end{bmatrix}$.

11. 0; el punto está en el plano.

13. $\dfrac{2\sqrt{14}}{7}$

15. $\dfrac{4\sqrt{146}}{73}$

19. $(0, \frac{3}{4}, 0)$

21. $(-\frac{2}{3}, \frac{29}{12}, -\frac{25}{12})$

23. $\mathbf{R} \cdot d = (1, 1, 1) \cdot (1, 9, -7) = 3$

25. $25 \cos 45° = \dfrac{25\sqrt{2}}{2}$ lb

Capítulo 3

Sección 3.1

1. A: renglones [−1 0], [2 3], [−2 1]; columnas $\begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$; el tamaño es 3×2; el elemento (2, 2) es 3; el elemento (3, 1) es −2; B: renglones [−1 0 −2], [2 2 1]; columnas $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$; el tamaño es 2×3; el elemento (2, 2) es 2; el elemento (2, 3) es 1.

3. (a) El sistema $x = 1, y - 2 = 0, x - y = 0$ es inconsistente.
(b) El sistema $x + y = 1, -y + z = 1, x + z = 3$ es inconsistente.

5. (c) y (e) son imposibles debido a sus tamaños incompatibles.

(a) $\begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} -3 & -3 & 3 \\ -3 & -3 & 3 \end{bmatrix}$

(d) $\begin{bmatrix} -9 & 9 \\ 9 & -9 \\ -9 & 9 \end{bmatrix}$

(f) $\begin{bmatrix} -12 & 26 \\ -40 & -18 \end{bmatrix}$

7. (a) $\begin{bmatrix} -4 & \frac{1}{2} \\ \frac{1}{2} & -1 \end{bmatrix}$

(b) $\begin{bmatrix} \frac{17}{2} & -4 \\ -\frac{25}{2} & 8 \end{bmatrix}$

(c) $\begin{bmatrix} 4 & -16 \\ -52 & 32 \end{bmatrix}$

11. $\begin{bmatrix} 13 & 12 & 9 & 8 \end{bmatrix}$

13. 1

15. $A^8 = \begin{bmatrix} 1 & 8 \\ 0 & 1 \end{bmatrix}$; $A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$

19. $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

21. $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$

23. Como $AB = BA$, entonces

$$\begin{aligned}(AB)^2 &= (AB)(AB) = A(BA)B \\ &= A(AB)B = (AA)(BB) = A^2B^2\end{aligned}$$

27. $\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}$

31. $\begin{bmatrix} 64 & 32 & 24 \\ 80 & 40 & 32 \\ 48 & 24 & 16 \end{bmatrix}$

Sección 3.2

1. $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{3}{4} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & -\frac{5}{8} \\ -\frac{1}{2} & \frac{7}{8} \end{bmatrix}$

3. $\begin{bmatrix} 4 & -\frac{3}{2} \\ -1 & \frac{1}{2} \end{bmatrix}$

5. $A^{-1} = \dfrac{1}{a^2 + b^2} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

7. Según el teorema 7,

$$\begin{aligned}(2A)^3 &= 2^3A^3 \\ &= 8\begin{bmatrix} 1 & 1 \\ -5 & -2 \end{bmatrix} = \begin{bmatrix} 8 & 8 \\ -40 & -16 \end{bmatrix}\end{aligned}$$

Por consiguiente,

$$\begin{aligned}(2A)^{-3} &= \begin{bmatrix} 8 & 8 \\ -40 & -16 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} -\frac{1}{12} & -\frac{1}{24} \\ \frac{5}{24} & \frac{1}{24} \end{bmatrix}\end{aligned}$$

9. $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} \frac{1}{5} & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}$

13. (a) $\begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{4} \\ 0 & -1 & 1 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$

(c) No invertible.

(d) $\begin{bmatrix} -1 & -1 & 1 \\ 0 & -1 & 1 \\ -1 & -2 & 1 \end{bmatrix}$

15. $\begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ -1 & 1 & 1 & 2 \\ -1 & 1 & 1 & 1 \end{bmatrix}$

17. (a) $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ -1 \\ -6 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \\ 2 \end{bmatrix}$

19. $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$

21. $A^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{c_3} \\ 0 & \frac{1}{c_2} & 0 \\ \frac{1}{c_1} & 0 & 0 \end{bmatrix}$

$$B^{-1} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{c_4} \\ 0 & 0 & \frac{1}{c_3} & 0 \\ 0 & \frac{1}{c_2} & 0 & 0 \\ \frac{1}{c_1} & 0 & 0 & 0 \end{bmatrix}$$

23. $A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $A^{-2} = I$. Por consiguiente, $A^{-2k} = I$ y $A^{-(2k-1)} = A^{-1}$ para cualquier entero positivo k. En particular, $A^{-24} = I$ y $A^{-25} = A^{-1}$.

25. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$

Sección 3.3

1. A, C y D son elementales. Las operaciones correspondientes son $R_1 + R_2 \to R_1$, $R_1 - 2R_3 \to R_1$ y $R_2 \leftrightarrow R_4$.

3. $R_1 \leftrightarrow R_2, 2R_2 \to R_2, R_1 - 5R_3 \to R_1$, y $-R_1 + R_3 \to R_3$

5. (a) $R_1 \leftrightarrow R_3$ (b) $R_1 - 5R_3 \to R_1$ (c) $2R_4 \to R_4$ (d) $-10R_1 + R_3 \to R_3$

7. $$A = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$
$$A^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}$$

9. $$A = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$
$$B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$
$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

11. La matriz es singular, porque su forma escalonada reducida $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ tiene un renglón de ceros.

13. Es cierto, de acuerdo con el teorema 15, porque la matriz de coeficientes es invertible. Específicamente,

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

15. $$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c_1 \end{bmatrix}$$
$$A^{-1} = \begin{bmatrix} 0 & 0 & c_3^{-1} \\ 0 & c_2^{-1} & 0 \\ c_1^{-1} & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c_1^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_2^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} c_3^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

19. **(a)** Falso. **(b)** Cierto. **(c)** Falso. **(d)** Falso. **(e)** Cierto. **(f)** Cierto.

21. $n \times m$.

Sección 3.4

1. $\mathbf{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$

3. $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$

5. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & -7 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 7 & -3 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & -5 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 5 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 1 & 2 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 5 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

15. $LU = \begin{bmatrix} 1 & 0 \\ 7 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & -5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -2 \\ 10 \end{bmatrix}$

17. $LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 2 & 1 & 0 \\ 0 & 5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 2 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$,

$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

19. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A$

$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 & -4 \\ 0 & 1 & 1 \\ 0 & 0 & -6 \end{bmatrix}$

21. Una factorización $PA = LU$ se expresa como sigue:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & -2 \end{bmatrix}$$

Por lo que el método LU para el sistema

$$PA\mathbf{x} = P\mathbf{b} = \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix}$$

da como resultado $\mathbf{x} = (-2, 0, 3)$.

23. Una factorización $PA = LU$ se expresa como sigue:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -5 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 3 \end{bmatrix}$$

Por lo que el método LU para el sistema

$$PA\mathbf{x} = P\mathbf{b} = \begin{bmatrix} -8 \\ 4 \\ 2 \end{bmatrix}$$

da como resultado $\mathbf{x} = (1, 2, 0)$.

Sección 3.5

1. $M_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, M_2 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$,

$M_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

3. $A(G_1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A(G_2) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$,

$A(G_3) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

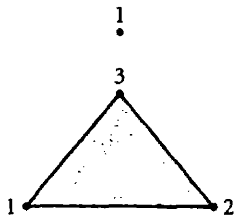
5. $A(G_7) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, A(G_7)^2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

La cantidad de caminatas de longitud 2 de (**a**) de 1 a 1 es 1, (**b**) de 1 a 2 es 0, (**c**) de 1 a 3 es 1. Esos números concuerdan con los elementos (1, 1), (1, 2) y (1, 3) de $A(G_7)^2$.

7. $I(G_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad I(G_2) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$,

$I(G_3)$ no existe.

9.



$L(G_1)$

$L(G_2)$

No tiene gráfica $L(G_3)$

11. $A(L(G_7)) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad I(G_7) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$

Por consiguiente,

$$I(G_7)^T I(G_7) - 2I_2$$
$$= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = A(L(G_7))$$

13. Sólo C es estocástica.
15. G y H son doblemente estocásticas.
17. Todas las potencias de las estocásticas son doblemente estocásticas. Por tanto, la matriz dada también lo es porque

$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ es estocástica.

19. Para A: $x = 0.8, y = 0.8$; para B: $x = \frac{3}{4}, y = \frac{1}{4}$.

21. (a) En vista de que

$$(I - C)^{-1} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.6 \end{bmatrix}\right)^{-1}$$
$$= \begin{bmatrix} 0.5 & -0.4 \\ -0.1 & 0.4 \end{bmatrix}^{-1}$$
$$= \begin{bmatrix} 2.5 & 2.5 \\ 0.625 & 3.125 \end{bmatrix}$$

existe, C es productiva.

(b) El vector de producción se calcula como sigue:

$$X = (I - C)^{-1}D$$
$$= \begin{bmatrix} 2.5 & 2.5 \\ 0.625 & 3.125 \end{bmatrix} \begin{bmatrix} 10 \\ 20 \end{bmatrix} = \begin{bmatrix} 75 \\ 68.75 \end{bmatrix}$$

Capítulo 4

Sección 4.1

1. Cierto, porque tanto $\begin{bmatrix} a \\ 0 \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} = \begin{bmatrix} a+b \\ 0 \end{bmatrix}$ como $c\begin{bmatrix} a \\ 0 \end{bmatrix} = \begin{bmatrix} ca \\ 0 \end{bmatrix}$ están en V. Es por eso que, V es un subespacio vectorial de $\mathbf{R}^2$.

3. Tenemos que

$$\begin{bmatrix} a \\ 0 \\ -2a \end{bmatrix} + \begin{bmatrix} b \\ 0 \\ -2b \end{bmatrix} = \begin{bmatrix} a+b \\ 0 \\ -2a-2b \end{bmatrix}$$
$$= \begin{bmatrix} a+b \\ 0 \\ -2(a+b) \end{bmatrix} \in V$$

y

$$c\begin{bmatrix} a \\ 0 \\ -2a \end{bmatrix} = \begin{bmatrix} ca \\ 0 \\ -2(ca) \end{bmatrix} \in V$$

Por tanto, V es un subespacio vectorial de $\mathbf{R}^3$.

5. Tenemos que

$$\begin{bmatrix} a-b \\ b-c \\ c-d \\ d-a \end{bmatrix} + \begin{bmatrix} a'-b' \\ b'-c' \\ c'-d' \\ d'-a' \end{bmatrix} = \begin{bmatrix} a-b+a'-b' \\ b-c+b'-c' \\ c-d+c'-d' \\ d-a+d'-a' \end{bmatrix}$$
$$= \begin{bmatrix} (a+a')-(b+b') \\ (b+b')-(c+c') \\ (c+c')-(d+d') \\ (d+d')-(a+a') \end{bmatrix}$$
$$\in V$$

y

$$k\begin{bmatrix} a-b \\ b-c \\ c-d \\ d-a \end{bmatrix} = \begin{bmatrix} (ka)-(kb) \\ (kb)-(kc) \\ (kc)-(kd) \\ (kd)-(ka) \end{bmatrix} \in V$$

Por tanto, V es un subespacio vectorial de $\mathbf{R}^4$.

7. Sea V el conjunto de todos los vectores-4 cuyos tres primeros componentes son cero, y estén $\mathbf{v}_1$ y $\mathbf{v}_2$ en V. Entonces, la suma $\mathbf{v}_1 + \mathbf{v}_2$ tiene los tres primeros componentes iguales a cero y en consecuencia está en V. También, $c\mathbf{v}_1$ tiene sus tres primeros componentes iguales a cero, así que está en V. De manera que V es un subespacio vectorial de $\mathbf{R}^4$.

9. Sean $\mathbf{v}_1 = (1, -1, 0)$ y $\mathbf{v}_2 = (2, 0, 2)$. Sea V el conjunto de todas las combinaciones lineales de $\mathbf{v}_1$ y $\mathbf{v}_2$. Entonces, si $\mathbf{u}_1 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ y $\mathbf{u}_2 = k_1\mathbf{v}_1 + k_2\mathbf{v}_2$,

$$\mathbf{u}_1 + \mathbf{u}_2 = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2) + (k_1\mathbf{v}_1 + k_2\mathbf{v}_2)$$
$$= (c_1 + k_1)\mathbf{v}_1 + (c_2 + k_2)\mathbf{v}_2 \in V$$

y

$$c\mathbf{u}_1 = c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2 \in V$$

Por consiguiente, V es un subespacio de $\mathbf{R}^3$.

11. Sí.

13. No.

15. No.

17. Sí.

19. No.

21. Sí.

23. No.

25. No.

27. Sí.

29. No.

31. $V_1 \cap V_2$ es una recta o un plano (si $V_1 = V_2$) que pasan por el origen; de modo que es un subespacio de $\mathbf{R}^3$.

33. Sí.

35. No.

37. Sí.

39. Sí.

41. $\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$

43. $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}\right\}$, o aun la base estándar de $\mathbf{R}^3$.

45. Sí.

47. No

49. No

51. (**a**) No (**b**) No

53. (**a**) No (**b**) No

55. $\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}\right\}$

57. $\left\{\begin{bmatrix} -5 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 7 \end{bmatrix}\right\}$

61. $\mathbf{x} = \begin{bmatrix} -7 \\ 6 \\ 2 \end{bmatrix}$

63. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 7 \\ -3 \end{bmatrix}$

65. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 4 \\ -1 \end{bmatrix}$

67. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} a-b \\ a+b \end{bmatrix}$

Sección 4.2

1. Si f y g son dos polinomios de grado $\le n$, también lo son la suma $f+g$ y el producto por escalar cf. Así, los axiomas (A1) y (M1) son válidos, y también (A2), (A3), (M2) a (M5) lo son porque las propiedades correspondientes son válidas para los coeficientes de los polinomios. Ahora bien, P_n contiene el polinomio cero, por definición, y es claro que $f+\mathbf{0}=\mathbf{0}+f=f$. Además, el grado de $-f$ es $\le n$, y $f+(-f)=\mathbf{0}$. Por lo que P_n es un espacio vectorial.

3. Sea V el conjunto dado y sean $f = ap + bq$ y $f' = a'p + b'q$ dos elementos de V. Entonces

$$\begin{aligned} f + f' &= ap + bq + a'p + b'q \\ &= (a + a')p + (b + b')q \in V \end{aligned}$$

y

$$cf = c(ap + bq) = (ca)p + (cb)q \in V$$

Por consiguiente, (A1) y (M1) son válidos. El resto de los axiomas son válidos, porque son válidos en el conjunto mayor P.

5. No, el axioma (M5) falla: $1(1, 1) = (0, 0) \ne (1, 1)$.

7. Cierto, porque la suma de dos matrices invertibles puede no ser invertible.

9. No, vea el coeficiente delantero de $2p$.

11. No. (Demuestre que para f y g dentro del conjunto, la suma $f+g$ no se incluye en éste.)

13. Sí. Es igual que en el ejercicio 8.

15. No. La suma $(1 + x^2) + (x + x^2) = 2x^2 + x + 1$ no está en el conjunto.

17. Sí, porque si V está en el conjunto, entonces

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} a' & -b' \\ b' & a' \end{bmatrix} = \begin{bmatrix} a+a' & -(b+b') \\ b+b' & a+a' \end{bmatrix} \in V$$

y

$$c\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} ca & -(cb) \\ cb & ca \end{bmatrix} \in V$$

19. No, la matriz cero no está en el conjunto.

21. Sí.

23. Sí.

25. Sí. Si f y g son pares, también lo son $f+g$ y cf.

29. Si $r = 0$, está bien. Si $r \ne 0$, entonces r^{-1} existe y

$$\begin{aligned} r\mathbf{u} &= r\mathbf{v} &&\Rightarrow \\ r^{-1}(r\mathbf{u}) &= r^{-1}(r\mathbf{v}) &&\Rightarrow \\ (r^{-1}r)\mathbf{u} &= (r^{-1}r)\mathbf{v} &&\Rightarrow \\ 1\mathbf{u} &= 1\mathbf{v} &&\Rightarrow \\ \mathbf{u} &= \mathbf{v} \end{aligned}$$

Sección 4.3

1. (**a**) Sí. (**b**) No (**c**) No.

3. No

5. Sí

7. Sí

9. No

11. Sí

13. Sí

15. (**a**) Falso. (**b**) Cierto. (**c**) Falso. (**d**) Cierto. (**e**) Cierto. (**f**) Falso. (**g**) Cierto.

19. No.

21. Sí.

23. (**a**) Falso. (**b**) Falso. (**c**) Cierto. (**d**) Falso.

27. No, por ejemplo, suponga que $p = x + 1$, $q = 1$, $r = -x - 1$.

29. No.

31. En el ejercicio 18, de la sección 4.2 vimos que V es un subespacio vectorial de M_{22}. Si suponemos $a = 1$, $b = c = 0$, comprobaremos que $E_{11} - E_{22}$ está en V. También, partiendo

de $a = c = 0$ y $b = 1$, veremos que E_{12} está en V. Por último, suponiendo que $a = b = 0$, $c = 1$ indica que E_{21} está en V. $\mathcal{B}$ genera a V, porque

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$\mathcal{B}$ es linealmente independiente, porque cuando se iguala a cero la combinación lineal del lado derecho, implica que la matriz del lado izquierdo es cero. De ahí se infiere que $a = b = c = 0$. En consecuencia, $\mathcal{B}$ es una base de V.

33. Cierto, porque

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

tiene exactamente 4 pivotes.

35. **(a)** $\mathcal{B} = \{2 + x + 2x^2, x^2, 1 - x - x^2\}$
(b) $V = P_2$

37. **(a)** $\mathcal{B} = \{x^2, 1 + x, -1 + x^2\}$
(b) $V = P_2$

39. **(a)** $\mathcal{B} = \{-x + x^2, -5 + x, -x^2\}$
(b) $V = P_2$

41. $\{-x + x^2, x + x^2, 1\}$

43. $\{1 + x, -1 + x^2, x^2\}$

45. $\{-x + x^2, -5 + x, x^2\}$

Sección 4.4

1. $\dim(V) = 1$

3. $\dim(V) = 1$

5. $\dim(V) = 2$

7. $\dim(V) = 1$

9. La dimensión es 2.

11. La dimensión es 3.

13. $\dim(V) = 3$

15. $\dim(V) = 2$

17. $\dim(V) = 2$

19. **(a)** Falso **(b)** Cierto. **(c)** Falso. **(d)** Cierto. **(e)** Falso. **(f)** Cierto. **(g)** Falso.

21. La dimensión es 2.

23. La dimensión es 2.

25. $\{\mathbf{0}\}$, $\mathbf{R}^4$, todas las rectas pasan por el origen, lo mismo que todos los planos y los hiperplanos.

Sección 4.5

1. $p = -3 + 24x$

3. $p = (2a - 3b) + (2a + 3b)x$

5. $[p]_{\mathcal{B}} = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$

7. $[p]_{\mathcal{B}} = \begin{bmatrix} a - b \\ a + b \end{bmatrix}$

9. $(1, -2, 1, -1)$

11. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

13. $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{bmatrix}$

17. $\begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{3}{4} \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{bmatrix}$

19. $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 & \frac{3}{2} \\ 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$

21. La matriz de transición es $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ y las nuevas coordenadas de $(1,1)$ se expresan por $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$.

Sección 4.6

1. **(a)** Base: $\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$; la nulidad es 1.

(b) Base: $\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$; la nulidad es 1.

3. **(a)** Base: $\left\{\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right\}$;

la nulidad es 3.

(b) Base: vacía. La nulidad es 0.

5. **(a)** Base: $\left\{\begin{bmatrix}-1\\1\\1\\0\end{bmatrix},\begin{bmatrix}2\\1\\0\\1\end{bmatrix}\right\}$;

la nulidad es 2.

(b) Base: $\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$; la nulidad es 1.

7. Base: $\left\{\begin{bmatrix}-1\\-1\\0\\0\\1\end{bmatrix},\begin{bmatrix}-4\\-2\\1\\0\\0\end{bmatrix}\right\}$; la nulidad es 2.

Sean N la nulidad, P la cantidad de columnas pivote y C la cantidad de columnas de una matriz. Entonces:

9. **(a)** $N = 2, P = 2, 2 + 2 = 4 = C$

(b) $N = 1, P = 2, 1 + 2 = 3 = C$

11. $\mathbf{b} = \mathbf{0}$

13. **a**, **b** y **c** están en el espacio de columnas.

15. **u** y **w** están en el espacio de columnas.

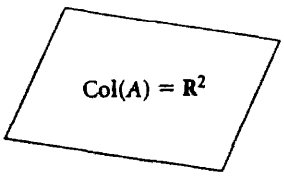
17. Ninguno de los vectores está en el espacio de columnas.

19. $\left\{\begin{bmatrix}2\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$

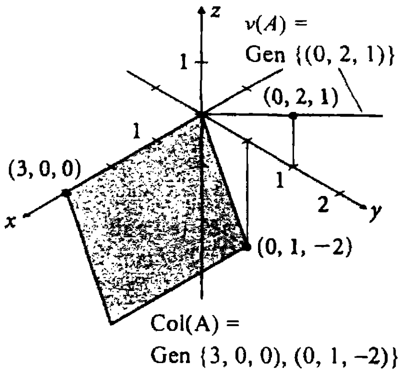
21. $\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}-2\\-2\\2\\0\end{bmatrix},\begin{bmatrix}1\\2\\-4\\0\end{bmatrix},\begin{bmatrix}0\\4\\2\\2\end{bmatrix}\right\}$

23. $\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}-2\\-2\\0\\0\end{bmatrix},\begin{bmatrix}0\\4\\2\\2\end{bmatrix}\right\}$

25. El espacio de columnas es $\mathbf{R}^2$.



27. El espacio nulo está generado por $\{(0, 2, 1)\}$ y el espacio de columnas está generado por el conjunto $\{(3, 0, 0), (0, 1, -2)\}$.



29. $A = \begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$ y $B = \begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}$; $A \sim B$, pero la Col(A) contiene a todos los múltiplos de $\begin{bmatrix}1\\1\end{bmatrix}$, mientras que la Col($B$) contiene todos los múltiplos de $\begin{bmatrix}1\\0\end{bmatrix}$

31. $\left\{\begin{bmatrix}-1\\2\\3\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$

33. $\left\{\begin{bmatrix}1\\-2\\-3\\1\end{bmatrix},\begin{bmatrix}0\\1\\2\\0\end{bmatrix}\right\}$

35. $\left\{\begin{bmatrix}-1\\0\\1\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$

37. $\left\{\begin{bmatrix}-1\\0\\1\\0\end{bmatrix},\begin{bmatrix}1\\-1\\0\\0\end{bmatrix},\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$

39. Base: $\left\{\begin{bmatrix}1\\2\\2\\-1\end{bmatrix},\begin{bmatrix}0\\-1\\2\\3\end{bmatrix}\right\}$; el rango es 2.

41. Base: $\left\{\begin{bmatrix}1\\2\\2\end{bmatrix},\begin{bmatrix}0\\-1\\2\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$; el rango es 3.

43. $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$

45. $\left\{\begin{bmatrix}1\\0\\-2\end{bmatrix},\begin{bmatrix}0\\1\\-4\end{bmatrix}\right\}$

47. $B \sim \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, rango$(A) = 2$.

$B^T \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, rango$(B^T) = 2$.

49. La nulidad de B es 2 (una base de espacio nulo: $\{(-6, 0, 2, 1), (-1, 1, 0, 0)\}$). El rango es 2. $4 = 2 + 2$ y equivale a la cantidad de columnas.

51. $B \sim \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, de modo que el rango de B es 2. Y

$[B\colon \mathbf{b}] \sim \begin{bmatrix} 1 & 1 & 2 & 2 & 1 \\ 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, así que el rango de $[B : \mathbf{b}]$ es nuevamente 2.

53. Sí. A tiene 450 columnas y nulidad 50, de modo que su rango es 400, de acuerdo con el teorema del rango. Por consiguiente, la dimensión del espacio de columnas es 400. Entonces, el espacio de columnas es todo $\mathbf{R}^{400}$. Así, todo vector-400 $\mathbf{b}$ está generado por las columnas de A. En consecuencia, $A\mathbf{x} = \mathbf{b}$ es consistente para todos los vectores-400 $\mathbf{b}$.

Sección 4.7

1. (a) $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ (d) $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

3. $A^2 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$; $A^3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

5. No, porque $\mathbf{u} + \mathbf{v} + \mathbf{w} = \mathbf{0}$.

7. $\{(1, 1, 1)\}$ es una base del espacio nulo de A sobre Z_2. Los únicos elementos del espacio nulo sobre Z_2 son $(1, 1, 1)$ y $(0, 0, 0)$. Sobre $\mathbf{R}$, el espacio nulo de A es $\{\mathbf{0}\}$, así que la única base es el conjunto vacío.

9. Como las matrices conmutan, $AB = BA$. Tenemos que,

$$\begin{aligned}(A + B)^2 &= (A + B)(A + B) \\ &= A^2 + AB + BA + B^2 \\ &= A^2 + 2AB + B^2 \\ &= A^2 + B^2\end{aligned}$$

porque $2AB$ es la matriz cero sobre Z_2.

11. 0101010

13. (a) $(1, 1, 1, 1)$ (b) $(0, 1, 1, 1)$

15. (a) $(0, 1, 1, 0)$ (b) $(0, 1, 1, 0)$

17. (a) $(1, 1, 1, 0)$ (b) $(1, 1, 1, 0)$

Capítulo 5

Sección 5.1

1. Es una transformación. Su dominio es $\mathbf{R}^2$ y su codominio es $\mathbf{R}^2$.

3. Es una transformación. Su dominio es $\mathbf{R} \times \mathbf{R}^+$ ($\mathbf{R}^+$ es el conjunto de los números reales positivos), y su codominio es $\mathbf{R}^2$.

5. No es una transformación. (El producto matricial está indefinido.)

7. No son iguales; tienen distintos codominios.

9. **(a)** $n = 2$, $m = 2$ **(b)** $\mathbf{R}^2$ y $\mathbf{R}^2$ **(c)** $\{0\}$ **(d)** $\mathbf{R}^2$

11. **(a)** $n = 3$, $m = 2$ **(b)** $\mathbf{R}^3$ y $\mathbf{R}^2$ **(c)** Gen$\{1, -1, 1)\}$ **(d)** $\mathbf{R}^2$

13. **(a)** $n = 3$, $m = 3$ **(b)** $\mathbf{R}^3$ y $\mathbf{R}^3$

(c) Gen $\{(0, 1, 0), (1, 0, 1)\}$ **(d)** Gen$\{(1, -1, 1)\}$

15. **(a)** $n = 4$, $m = 2$ **(b)** $\mathbf{R}^4$ y $\mathbf{R}^2$

(c) $\{(-1, -1, 1, 0), (2, 0, 0, 1)\}$ **(d)** $\mathbf{R}^2$

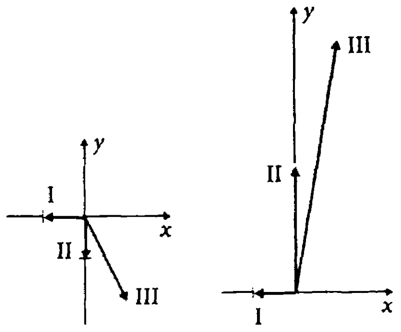
17. $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$

19. $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$

21. $\left\{\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}\right\}$

23. **(a)** $\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}$; rotación de 180°.

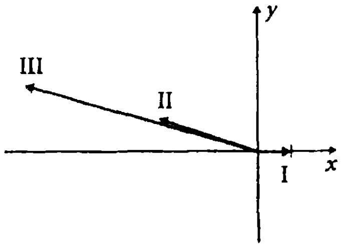
(b) $\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \end{bmatrix}$; ninguna.



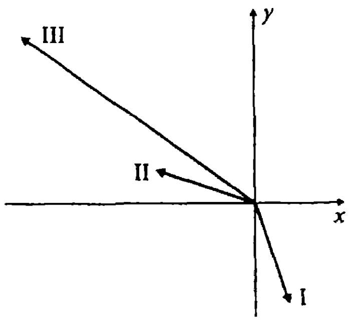
(a) (b)

25. (a) $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}-3\\1\end{bmatrix}, \begin{bmatrix}-7\\2\end{bmatrix}$; deslizamiento por un factor de 3 a lo largo de la dirección x opuesta.

(b) $\begin{bmatrix}1\\-3\end{bmatrix}, \begin{bmatrix}-3\\1\end{bmatrix}, \begin{bmatrix}-7\\5\end{bmatrix}$; ninguna.



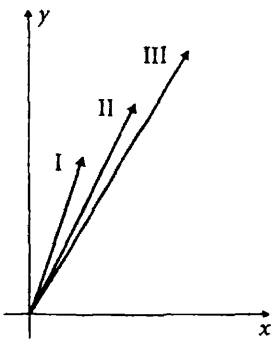
(a)



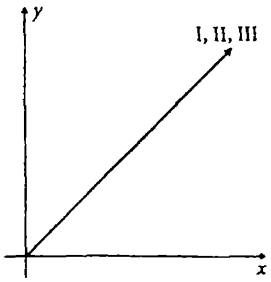
(b)

27. (a) $\begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix}, \begin{bmatrix}3\\5\end{bmatrix}$; ninguna.

(b) $\begin{bmatrix}3\\3\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}$; ninguna.



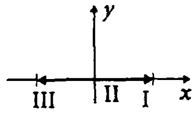
(a)



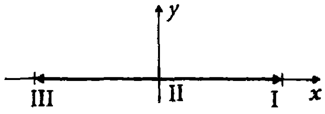
(b)

29. (a) $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\end{bmatrix}$; proyección sobre el eje x.

(b) $\begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}-3\\0\end{bmatrix}$; ninguna.



(a)



(b)

31. $T(x, y) = (-y, -x)$. $A = \begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix}$

33. $R(x, y) = \begin{bmatrix}(\cos\theta)\,x + (\operatorname{sen}\theta)\,y\\-(\operatorname{sen}\theta)\,x + (\cos\theta)\,y\end{bmatrix}$

$$A = \begin{bmatrix}\cos\theta & \operatorname{sen}\theta\\-\operatorname{sen}\theta & \cos\theta\end{bmatrix}$$

37. R_θ^x y R_θ^y con sus matrices respectivas:

$$\begin{bmatrix}1 & 0 & 0\\0 & \cos\theta & -\operatorname{sen}\theta\\0 & \operatorname{sen}\theta & \cos\theta\end{bmatrix}$$

y

$$\begin{bmatrix}\cos\theta & 0 & -\operatorname{sen}\theta\\0 & 1 & 1\\\operatorname{sen}\theta & 0 & \cos\theta\end{bmatrix}$$

Sección 5.2

1. $\mathbf{R}^2$ y $\mathbf{R}^2$; lineal.
3. $\mathbf{R}^2$ y $\mathbf{R}^2$; no lineal.
5. $\mathbf{R}^2$ y $\mathbf{R}^2$; lineal.
7. $\mathbf{R}^2$ y $\mathbf{R}^2$; lineal.
9. $\mathbf{R}^2$ y $\mathbf{R}^3$; lineal.
11. $\mathbf{R}^3$ y $\mathbf{R}^2$; lineal.

13.
$$T\left(c_1\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix} + c_2\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\right) = T\begin{bmatrix}c_1x_1 + c_2x_2\\c_1y_1 + c_2y_2\\c_1z_1 + c_2z_2\end{bmatrix}$$
$$= \begin{bmatrix}c_1x_1 + c_2x_2 + c_1y_1 + c_2y_2 + c_1z_1 + c_2z_2\\c_1x_1 + c_2x_2 - c_1y_1 - c_2y_2 - c_1z_1 - c_2z_2\end{bmatrix}$$
$$= c_1\begin{bmatrix}x_1 + y_1 + z_1\\x_1 - y_1 - z_1\end{bmatrix} + c_2\begin{bmatrix}x_2 + y_2 + z_2\\x_2 - y_2 - z_2\end{bmatrix}$$

$$= c_1T\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}\right) + c_2T\left(\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right)$$

15. Lineal.
17. No lineal.
19. No lineal.
21. Sí.
23. No
25. $\begin{bmatrix} \frac{1}{2}x + \frac{1}{2}y + 2z \\ x - y - z \end{bmatrix}$ y $\begin{bmatrix} -\frac{95}{2} \\ 0 \end{bmatrix}$.
27. $-a - b + c + \left(2a + \frac{1}{2}b + \frac{1}{2}c\right)x$ y $\left(-\frac{95}{2}\right)x$

29. Ya que

$$\begin{aligned} T(c_1X + c_2Y) &= C^{-1}(c_1X + c_2Y)C \\ &= C^{-1}(c_1XC + c_2YC) \\ &= C^{-1}(c_1XC) + C^{-1}(c_2YC) \\ &= c_1(C^{-1}XC) + c_2(C^{-1}YC) \\ &= c_1T(X) + c_2T(Y) \end{aligned}$$

T es lineal, de acuerdo con el teorema 3.

31. $(-2, 2)$ está en el generador de $\{(1, -1)\}$, así que no hay manera de especificar T para algún vector que no esté en ese generador. Algunos ejemplos de transformaciones lineales distintas, tales que $T(1, -1) = (3, -1)$, de modo que $T(-2, 2) = (-6, 2)$, son los siguientes:

$$\begin{aligned} T(x,y) &= (-6x - 9y, 2x + 3y) \\ T(x,y) &= (-3x - 6y, x + 2y) \\ T(x,y) &= \left(-\tfrac{3}{2}x - \tfrac{9}{2}y, \tfrac{1}{2}x + \tfrac{3}{2}y\right) \end{aligned}$$

33. $T(x,y) = (xy, x)$
35. $T(x,y) = (x,x), \mathbf{v}_1 = (2,0), \mathbf{v}_2 = (1,1)$
39. El tamaño de A es de $m \times n$ y $\mathbf{b}$ es un vector-n. T es no lineal, porque.

$$T(\mathbf{0}) = A\mathbf{0} + \mathbf{b} = \mathbf{b} \neq \mathbf{0}$$

41. $A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$
43. $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$
45. $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
47. $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$
49. Tienen la forma $T(x) = cx$ para un escalar c fijo.

Sección 5.3

1. Base del núcleo: $\{(1, 1, 1)\}$; base del contradominio: $\{(1, 0, -1), (-1, 1, 0)\}$; la nulidad es 1 y el rango es 2. Teorema de la dimensión: $1+2 = \dim(\mathbf{R}^3)$.

3. Base del núcleo: $\{(2, 1)\}$; base del contradominio: $\{(1, 1, 2, 0)\}$; la nulidad es 1 y el rango es 1. Teorema de la dimensión: $1+1 = \dim(\mathbf{R}^2)$.

5. Base del núcleo: el conjunto vacío; base del contradominio: $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$; la nulidad es 0 y el rango es 3. Teorema de la dimensión: $0 + 3 = \dim(\mathbf{R}^3)$.

7. Base del núcleo: el conjunto vacío; base del contradominio: $\{(1, 0), (0, 1)\}$; la nulidad es 0 y el rango es 2. Teorema de la dimensión: $0 + 2 = \dim(\mathbf{R}^2)$.

9. Base del núcleo: $\{1 + x - x^2 + x^3\}$; base del con-tradominio: $\{1, -1 + x, x^2\}$; la nulidad es 1 y el rango es 3. Teorema de la dimensión: $1 + 3 = \dim(P_3)$.

11. Base del núcleo: el conjunto vacío; base del contradominio: $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$

13. Base del núcleo: $\{1, x\}$; base del contradominio: $\{1\}$.

15. $\dim(\operatorname{Ker}(T)) = 3$.

17. La nulidad es 2. El rango es 2.

19. La nulidad es 2. El rango es 2.

21. El sistema $2x + y = 0, -3x + 4y = 0$ sólo tiene la solución trivial, así que T es biunívoca, de acuerdo con el teorema 9.

23. El sistema $2x - y = 0, x - y = 0, -x + y = 0, x - 2y = 0$ sólo tiene la solución trivial, así que T es biunívoca, de acuerdo con el teorema 9.

25. El sistema $2a - b = 0, -b + c = 0, -3a + c = 0$ sólo tiene la solución trivial, así que T es biunívoca, de acuerdo con el teorema 9.

27. Como la nulidad es cero, el rango es 2, de modo que el rango es $\mathbf{R}^2$; por consiguiente, T es sobre.

29. T es sobre, porque el sistema $x - y + z = a, -x + y + z = b$ puede resolverse para todas a y b. Por ejemplo, $x = \frac{1}{2}a - \frac{1}{2}b$, $y = 0, z = \frac{1}{2}a + \frac{1}{2}b$ es una solución, de modo que para esas elecciones (x, y, z) se transforma en (a, b).

31. T es sobre, porque el sistema $a + b = A, a + c = B$ se puede resolver para todas A y B. Por ejemplo, $a = B, b = A - B, c = 0$ es una solución, de modo que para esas elecciones, $a + bx + cx^2$ se transforma en $A + Bx$.

33. El núcleo es cero, porque $a - 2b = 0$, $-2a + b = 0$ tiene sólo la solución trivial. En consecuencia, T es biunívoca. Entonces es un isomorfismo, según el teorema 11.

35. El núcleo es cero, porque $x - y + z = 0$, $-x + y + z = 0$, $-y + z = 0$ sólo tiene la solución trivial. Por tanto, T es biunívoca. Entonces, es un isomorfismo, según el teorema 11.

37. El núcleo es cero, porque $c = 0$, $b = 0$, $a - b = 0$ sólo tiene la solución trivial. De modo que, T es biunívoca. Entonces es un isomorfismo, según el teorema 11.

39. No es isomorfismo.

41. Isomorfismo.

42. No es isomorfismo.

Sección 5.4

1. $\begin{bmatrix} 1 & -1 \\ 0 & -2 \\ 2 & 2 \end{bmatrix}$

3. $\begin{bmatrix} 3 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 2 \end{bmatrix}$

5. **(a)** $A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ **(b)** $A' = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

(c) Directamente, $T(6x - 2x^2) = 4 + 6x$. Usando A:

$$\begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 6 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix}$$

Así se obtiene $4 + 6x$ en términos de la base estándar. Empleando A': primero $6x - 2x^2$ en términos de $\mathcal{B}'$, tenemos

$$6x - 2x^2 = -2(-x + x^2) + 0(1 + x) + 4(x)$$

Entonces se forma el producto matricial

$$\begin{bmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}$$

para obtener las coordenadas de la imagen con respecto a $\mathcal{B}'$. Así, $T(6x - 2x^2) = 0(-x + x^2) + 4(1 + x) + 2(x) = 4 + 6x$.

7. **(a)** $A = \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -3 & 1 \end{bmatrix}$

(b) Directamente: $T(5 - 2x) = 5 - 5x + 2x^2$; usando A:

$$5 - 2x = \frac{3}{2}(1 + x) - \frac{7}{2}(-1 + x)$$

Así,

$$\begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} \frac{3}{2} \\ -\frac{7}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ -8 \end{bmatrix}$$

Por consiguiente,

$$T(5 - 2x) = 2(-x + x^2) + 5(1 + x) - 8x = 5 - 5x + 2x^2$$

9. El vector coordenado $\begin{bmatrix} 1 & -2 \\ -4 & 3 \end{bmatrix}$ con respecto a $\mathcal{R}$ es $(2, 0, -4, -1)$; en consecuencia,

$$\begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ -4 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -3 \\ 2 \end{bmatrix}$$

y

$$T\begin{bmatrix} 1 & -2 \\ -4 & 3 \end{bmatrix} = -3\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -3 & 5 \end{bmatrix}$$

13. $m = 2, n = 4$

$$T(a + bx + cx^2 + dx^3) = (-9a - b + 8c + 5d) + (11a + 3b - 10c + d)x$$

15. $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{bmatrix}$

A es invertible, de modo que T es un isomorfismo, de acuerdo con el teorema 16.

Sección 5.5

1. $(f + g)(x, y, z) = \begin{bmatrix} z \\ 3x - z \end{bmatrix}$

$(f + g)(-1, 2, 0) = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$

$(-4f)(x, y, z) = \begin{bmatrix} -4x + 4y - 4z \\ -4x - 4y \end{bmatrix}$

$(-4f)(-1, 2, 0) = \begin{bmatrix} 12 \\ -4 \end{bmatrix}$

3. $(f + g)(x, y, z) = \begin{bmatrix} -x - y + z \\ 2x - 3y - z \end{bmatrix}$

$(f + g)(-1, 2, 0) = \begin{bmatrix} -1 \\ -8 \end{bmatrix}$

$$(-4f)(x,y,z) = \begin{bmatrix} 4x+8y \\ -4x+4z \end{bmatrix}$$

$$(-4f)(-1,2,0) = \begin{bmatrix} -12 \\ 4 \end{bmatrix}$$

5. Matriz (f) + Matriz (g)

$$= \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 0 \\ 2 & -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 1 \\ 3 & 0 & -1 \end{bmatrix}$$

= Matriz $(f+g)$ −2 Matriz (f)

$$= \begin{bmatrix} -2 & 2 & -2 \\ -2 & -2 & 0 \end{bmatrix} = \text{matriz } (-2f)$$

7. Matriz (f) + Matriz(g)

$$= \begin{bmatrix} -1 & -2 & 0 \\ 1 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \\ 1 & -3 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -1 & 1 \\ 2 & -3 & -1 \end{bmatrix}$$

= Matriz$(f+g)$ − 2 matriz(f)

$$= \begin{bmatrix} 2 & 4 & 0 \\ -2 & 0 & 2 \end{bmatrix} = \text{Matriz}(-2f)$$

9. $f \circ g(x,y) = (x+4y, 3x-y), f \circ g(-1,-3) = (-13,0)$

11. Matriz (f) · Matriz(g) $= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -5 & 1 \\ 1 & 3 \end{bmatrix}$

$$= \begin{bmatrix} -6 & -2 \\ -4 & 4 \end{bmatrix}$$

= Matriz $(f \circ g)$

13. Matriz (f) · Matriz (g) $= \begin{bmatrix} -1 & -2 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$

$$\cdot \begin{bmatrix} 0 & 1 \\ 1 & -3 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 4 \\ -1 & 1 \end{bmatrix} = \text{Matriz}(f \circ g)$$

15. (a) El codominio de g es $\mathbf{R}^3$, mientras que el dominio de f es $\mathbf{R}^2$. $g \circ f$ es indefinida.
(b) El codominio de g es $\mathbf{R}^2$, mientras que el dominio de f está en $\mathbf{R}^3$. $g \circ f$ es definida.
(c) El codominio de g es $\mathbf{R}^3$, mientras que el dominio de f es $\mathbf{R}^2$; $g \circ f$ es definida.

17. $f^3(x,y) = (-2x-2y, 2x-2y)$. $f^3(1,-1) = (0,4)$

19. f es invertible porque su matriz estándar

$$\begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & -2 \\ 2 & -1 & 0 \end{bmatrix}, \text{ es invertible.}$$

21. f es invertible porque su matriz estándar también lo es, y su inversa es $\begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{4} \\ 0 & -1 & 1 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$. Por consiguiente,

$$f^{-1}(x,y,z) = \left(-\tfrac{1}{2}x + \tfrac{1}{4}z, -y+z, -\tfrac{1}{2}z\right)$$

23. f es invertible, porque su matriz estándar es invertible, y su inversa es $\begin{bmatrix} \frac{5}{3} & -\frac{7}{6} & \frac{1}{6} \\ -1 & 1 & 0 \\ -\frac{1}{3} & -\frac{1}{6} & \frac{1}{6} \end{bmatrix}$. Por consiguiente,

$$f^{-1}(x,y,z) = \left(\tfrac{5}{3}x - \tfrac{7}{6}y + \tfrac{1}{6}z, -x+y, -\tfrac{1}{3}x - \tfrac{1}{6}y + \tfrac{1}{6}z\right)$$

25. La matriz es invertible, así que f es invertible; f^{-1} es la multiplicación izquierda por la inversa de la matriz dada. De modo que, f^{-1}: $\mathbf{R}^4 \to \mathbf{R}^4$ está expresada por

$$f^{-1}(\mathbf{x}) = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ -1 & 1 & 1 & 2 \\ -1 & 1 & 1 & 1 \end{bmatrix} \mathbf{x}$$

35. Como $f^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{5}x + \frac{2}{5}y \\ -\frac{2}{5}x + \frac{1}{5}y \end{bmatrix}$ y

$f^{-2}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{3}{25}x + \frac{4}{25}y \\ -\frac{4}{25}x - \frac{3}{25}y \end{bmatrix}$, $f^{-2} \circ f^{-1}$ está expresada por

$$f^{-2} \circ f^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{11}{125}x - \frac{2}{125}y \\ \frac{2}{125}x - \frac{11}{125}y \end{bmatrix}$$

que también es la fórmula de $f^{-3}(x,y)$.

Sección 5.6

1. $T(\mathbf{0}) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $T(\mathbf{e}_1) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $T(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$

3. $T(\mathbf{0}) = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $T(\mathbf{e}_1) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$,

$T(\mathbf{e}_2) = \begin{bmatrix} 3 \\ 1 \\ -4 \end{bmatrix}$, $T(\mathbf{e}_3) = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

5. $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

7. $\begin{bmatrix} -1 & 3 & 0 \\ 1 & 0 & -1 \\ 1 & -5 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

9. $A = \begin{bmatrix} 0 & 5 \\ 2 & -8 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

11. $A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$

13. $A = \begin{bmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

15. Como $T(\mathbf{0}) = \mathbf{b}$, $\mathbf{b}$ está determinado por $T(\mathbf{0})$. Además, puesto que $L(\mathbf{x}) = T(\mathbf{x}) - \mathbf{b} = A\mathbf{x}$ es una transformación matricial, se puede determinar únicamente por medio de $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$, y así T está definida en forma única por los valores

$$T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n), T(\mathbf{0})$$

17. Los dos conjuntos son iguales.

Capítulo 6

Sección 6.1

1. (a) -3, (b) -3
3. (a) -10, (b) -10
5. (a) 158, (b) -15
7. (a) 1, (b) 0
9. -80
11. 0
13. 76
15. (a) $\det(A) = -2 = \det(A^T)$
    (b) $\det(AB) = 4 = -2(-2) = \det(A)\det(B)$
    (c) $\det(A^{-1}) = -\frac{1}{2} = 1/\det(A)$
17. (a) $M_{11} = -30, \quad C_{11} = -30,$
    $M_{12} = 10, \quad C_{12} = -10,$
    $M_{13} = -35, \quad C_{13} = -35,$
    $M_{21} = 12, \quad C_{21} = -12,$
    $M_{22} = -20, \quad C_{22} = -20,$
    $M_{23} = 14, \quad C_{23} = -14,$
    $M_{31} = -2, \quad C_{31} = -2,$
    $M_{32} = -10, \quad C_{32} = 10,$
    $M_{33} = 11, \quad C_{33} = 11$
    (b) (b_1) $1(-30) - 2(-10) + 2(-35) = -80$
    (b_2) $3(-12) + 5(-20) - 4(-14) = -80$
    (b_3) $7(-2) + 0(10) - 6(11) = -80$
    (b_4) $1(-30) + 3(-12) + 7(-2) = -80$
    (b_5) $-2(-10) + 5(-20) + 0(10) = -80$
    (b_6) $2(-35) - 4(-14) - 6(11) = -80$
19. $\det(A^{-1}) = \frac{1}{2} = 1/\det(A)$

21. $\det(A^{-1}) = \begin{vmatrix} \dfrac{d}{ad-bc} & -\dfrac{b}{ad-bc} \\ -\dfrac{c}{ad-bc} & \dfrac{a}{ad-bc} \end{vmatrix}$
$= \dfrac{ad-bc}{(ad-bc)^2} = \dfrac{1}{ad-bc} = \dfrac{1}{\det(A)}$

23. $\det(E_3) = r = \det(E_4)$

25. El determinante de una matriz elemental E es 1, r o -1, dependiendo de si E proviene de eliminación, escalamiento (con factor r) o intercambio.

27. (a) $\det(E_3A) = r(aei - afh - bdi + cdh + bfg - ceg) = \det(E_3)\det(A) = \det(A_3)$
    (b) $\det(E_4A) = r(aei - afh - bdi + cdh + bfg - ceg) = \det(E_4)\det(A) = \det(A_4)$
29. (a) $\lambda = 2 \pm \sqrt{3}$ (b) $\lambda = 0, 11$
31. $\lambda = 0, 3$
33. $x = 1, x = 2$
35. $a = 2, b = 4$ o $a = -2, b = -4$
37. *Sugerencia:* Si T es invertible, las imágenes de $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ son linealmente independientes; en consecuencia definen un paralelepípedo. Su volumen es $|\det(A)|$, de acuerdo con la sección 2.6. ¿Qué sucede si T no es invertible?

Sección 6.2

1. (a) -50 (b) 0
3. (a) -1 (b) -1
5. (a) 1 (b) 1
7. 24
9. Se usó la propiedad 3 del teorema 1.
11. Se usó la propiedad 4 del teorema 4.
13. Se usó la propiedad 4 del teorema 1, para modificar el segundo renglón (no cambió el determinante), y después se empleó la propiedad 2 para escalar el último renglón.
15. Se usó el teorema 7.
17. $x = 0$ hace que el segundo renglón sea cero, y $x = 2$ hace que los renglones primero y último sean proporcionales. De modo que para cada caso, el determinante debe ser 0.
19. (a) $-\begin{vmatrix} 2 & 2 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -4$
    (b) $-1\begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = -1$

21. $\begin{vmatrix} 2 & -4 & 2 & 8 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & -2 \end{vmatrix} = 28$

23. $\begin{vmatrix} 1 & 0 & -1 & 2 & 1 \\ 0 & 2 & 6 & -2 & 6 \\ 0 & 0 & 3 & 0 & 3 \\ 0 & 0 & 0 & 5 & -1 \\ 0 & 0 & 0 & 0 & \frac{2}{5} \end{vmatrix} = 12$

25. (a) Al desarrollar cada lado se obtiene

$$ka_1b_2c_3 - ka_1b_3c_2 - ka_2b_1c_3 + ka_3b_1c_2 + ka_2b_3c_1 - ka_3b_2c_1$$

(b) Al desarrollar cada lado se obtiene

$$a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_3b_1c_2 + a_2b_3c_1 - a_3b_2c_1$$

29. $\det(C) = -1$, de modo que C es invertible. $\det(D) = -1$, por tanto D es invertible.

31. $\det(G) = 3$, por lo que G es invertible. $\det(H) = 2$, entonces H es invertible.

33. $k = -2, 0, 2$

35. $$\begin{aligned}\det(B^{-1}AB) &= \det(B^{-1})\det(A)\det(B) \\ &= \det(B)^{-1}\det(A)\det(B) \\ &= \det(A)\end{aligned}$$

37. (a) -14 (b) -56
(c) 7 (d) -2
(e) -14 (f) -1134

39. $$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & b^2-a^2 & c^2-a^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & 0 & (c-a)(c-b) \end{vmatrix} = (b-a)(c-b)(c-a)$$

41. $$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & b^3-a^3 & c^3-a^3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & 0 & (c-b)(c-a)(a+c+b) \end{vmatrix} = (b-a)(c-b)(c-a)(a+b+c)$$

Sección 6.3

1. (a) $\frac{1}{-2}\begin{bmatrix} 5 & -4 \\ -3 & 2 \end{bmatrix}$ (b) $\frac{1}{2}\begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix}$

3. $\frac{1}{4}\begin{bmatrix} 0 & 2 & 2 \\ 2 & -2 & 0 \\ 2 & 0 & -2 \end{bmatrix}$

5. $\frac{1}{1}\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{bmatrix}$

7. Se obtiene

$$A^{-1} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 1 & -1 & 1 \\ 3 & -3 & 2 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix}$$

por cualquier método. La reducción por operaciones de renglón es mucho más eficiente.

9. (a) $x = \frac{-2}{-2} = 1$, $y = \frac{0}{-2} = 0$
(b) $x = \frac{2}{-2} = -1$, $z = \frac{-2}{-2} = 1$

11. $x = \frac{4}{1} = 4; y = \frac{-3}{1} = -3; z = \frac{1}{1} = 1$

13. $z = \frac{-21}{-101} = \frac{21}{101}$

15. *Sugerencia:* Aplique determinantes a ambos lados de $A\,\mathrm{Adj}(A) = \det(A)I_n$.

Sección 6.4

1. (2, 1, 3, 4) es impar y su signo es −1; (1, 4, 2, 3) es par y su signo es 1; (1, 5, 2, 4, 3) es par y su signo es −1; (1, 4, 3, 5, 2) es par y su signo es 1.

3. (3, 1, 4, 2) es impar y su signo es −1; (4, 2, 1, 3) es par y su signo es −1; (3, 4, 2, 1, 5) es impar y su signo es −1; (4, 2, 5, 1, 3) es par y su signo es 1.

5. (a) $-2(3)(4) = -24$ (b) $2(3)(4) = 24$

7. (a) $2(3)(4) - 5(3)(6) = -66$ (b) $-2(3)(-4) - 5(3)(-6) = 114$

9. $1(2)(3)(4)(5) = 120$

11. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (1, 2, 3)

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (1, 3, 2)

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad (3,1,2)$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (3,2,1)$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2,1,3)$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (2,3,1)$$

13. Sea $p = (j_1, \ldots, j_n)$ una permutación y sea A la matriz correspondiente de permutación. Así, el i-ésimo renglón de A tiene 1 en su j_i-ésima columna, y ceros los demás elementos. Por consiguiente, el determinante de A sólo consiste en un término de la forma signo $(p)1 \cdots 1$ = signo (p).

Sección 6.5

1. $x + 2y - 3 = 0$
3. $3x + 2y - 2 = 0$
5. No, no son.
7. $x^2 + y^2 + 2y = 0$ o $x^2 + (y+1)^2 = 1$; el centro está en $(0, -1)$ y el radio es 1.
9. $x^2 + y^2 - 6x + 8 = 0$ o $(x-3)^2 + (y-2)^2 = 13$. El centro está en $(3, 2)$ y el radio es .
11. $-2y + 4x^2 - 6x + 8 = 0$ o $y = 2x^2 - 3x + 4$.
13. $-2y + 4x^2 - 22x + 32 = 0$ o $y = 2x^2 - 11x + 16$
15. $x + 2y - 3z + 2 = 0$
17. Primero, se reformula el sistema con y, desconocida, y coeficientes paramétricos en x.

$$y^2 + (x^2 - 1) = 0$$
$$y^2 + (x^2 - 2x - 1) = 0$$

La resultante de Sylvester es

$$\begin{vmatrix} 1 & 0 & x^2 - 1 & 0 \\ 0 & 1 & 0 & x^2 - 1 \\ 1 & -2 & x^2 + 2x + 1 & 0 \\ 0 & 1 & -2 & x^2 + 2x + 1 \end{vmatrix} = 0$$

Por consiguiente, $8x^2 + 8x = 0$. Así, $x = 0$ o $x = -1$. Si $x = 0$, entonces $y = 1$, de acuerdo con la segunda ecuación, y también se satisface la primera. Si $x = -1$, entonces $y = 0$, de acuerdo con la primera ecuación, y la segunda también se satisface. Así llegamos a dos soluciones: $x = 0$, $y = 1$ y $x = -1$, $y = 0$.

19. Partiendo de

$$\begin{vmatrix} x^2 + y^2 + z^2 & x & y & z & 1 \\ 54 & 1 & 2 & 7 & 1 \\ 38 & 5 & 2 & 3 & 1 \\ 46 & 1 & 6 & 3 & 1 \\ 6 & 1 & 2 & -1 & 1 \end{vmatrix} = 0$$

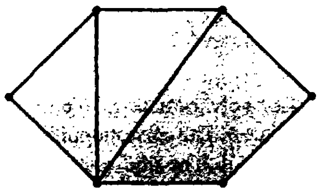
obtenemos, después de simplificar,

$$x^2 + y^2 + z^2 - 2x - 4y - 6z - 2 = 0$$

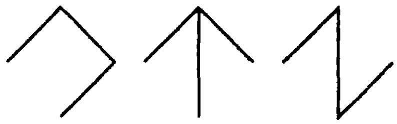
es decir

$$(x-1)^2 + (y-2)^2 + (z-3)^2 = 16$$

Por consiguiente, la esfera tiene su centro en $(1, 2, 3)$ y su radio es 4.

21. 

23. (a) Tenemos los siguientes árboles generadores:



(b) La matriz del árbol es

$$\begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}$$

y el cofactor es

$$C_{11} = (-1)^2 \begin{vmatrix} 1 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 3$$

De manera que hay tres árboles generadores.

Capítulo 7

Sección 7.1

1. $A\mathbf{u} = (-5, 5) = 5\mathbf{u}$. El eigenvalor es 5.
3. $A\mathbf{v} = (0, 0) = 0\mathbf{v}$. El eigenvalor es 0.
5. No, porque $A(\mathbf{u} + \mathbf{v}) = A(1, 4) = (-5, 5)$, que no es un múltiplo de $(1, 4)$.

7. $(-1, 1)$ no es un eigenvector de la forma escalonada reducida de A. Por otro lado, $(2, 3)$ sí es.

9. Sólo $(-2, 1)$ y $(1, -\frac{1}{2})$ son eigenvectores. El eigenvalor correspondiente de ambos es 0.

11. Como el polinomio característico es $-(\lambda - 7) \times (\lambda^2 + 2\lambda - 24)$, los eigenvalores son 7, 4 y -6.

13. A lo largo del eje x, los vectores permanecen igual, de modo que son eigenvectores y su eigenvalor es 1. Sobre el eje y, los vectores van al origen, de modo que son eigenvectores cuyo eigenvalor es 0.

15. No hay eigenvectores (con elementos reales). No queda algún vector distinto de cero en su propia línea después de la rotación.

17. **(a)** El polinomio característico es $\lambda^2 + 2\lambda - 24$. Los eigenvalores son 4 y -6. Las bases correspondientes de eigenvectores son $\{(2, 3)\}$ y $\{(1, -1)\}$. Todas las multiplicidades son 1.
 (b) El polinomio característico es $\lambda^2 + 6\lambda + 9$. El único eigenvalor es -3. La base correspondiente de eigenvectores es $\{(3, 1)\}$. La multiplicidad algebraica es 2. La multiplicidad geométrica es 1.
 (c) El polinomio característico es $\lambda^2 - 2\lambda - 35$. Los eigenvalores son 7 y -5. Las bases correspondientes de eigenvectores son $\{(1, 1)\}$ y $\{(-7, 5)\}$. Todas las multiplicidades son 1.

19. **(a)** El polinomio característico es $-\lambda^3 + \lambda^2 + \lambda - 1$. Los eigenvalores son 1, cuya base de eigenvectores $\{(1, 0, 1), (0, 1, 0)\}$, y -1, con base de eigenvectores $\{(-1, 0, 1)\}$. La multiplicidad algebraica y geométrica de 1 es 2. La multiplicidad algebraica y geométrica de -1 es 1.
 (b) El polinomio característico es $-(\lambda - 1) \times (\lambda - 2)(\lambda - 3)$. Los eigenvalores son 1, 2, 3, con las bases correspondientes de eigenvectores $\{(1, 0, 0)\}$, $\{(1, 1, 0)\}$ y $\{(0, 0, 1)\}$. Todas las multiplicidades son 1.

21. **(a)** El polinomio característico es $-\lambda^3 + 3\lambda^2 + 4\lambda - 12$. Los eigenvalores son 2, 3 y -2, cuyas bases correspondientes de eigenvectores son $\{(1, 1, 0)\}$, $\{(0, 0, 1)\}$ y $\{(-1, 1, 0)\}$. Todas las multiplicidades son 1.
 (b) El polinomio característico es $-(\lambda - 1) \times (\lambda^2 - 1)$. Los eigenvalores son 1, cuyas bases correspondientes de eigenvectores $\{(1, 0, 0), (0, 1, 1)\}$; y para las -1 con base de eigenvectores son $\{(0, -1, 1)\}$. La multiplicidad algebraica y geométrica de 1 es 2. La multiplicidad algebraica y geométrica de -1 es 1.

23. **(a)** El polinomio característico es $-(\lambda - 1) \times (\lambda - 2)(\lambda - 3)$. Los eigenvalores son 1, 2, 3, con las bases correspondientes de eigenvectores $\{(2, -3, 3)\}$, $\{(0, 1, 0)\}$ y $\{(0, 1, 1)\}$. Todas las multiplicidades son 1.
 (b) El polinomio característico es $-(\lambda - 1) \times (\lambda - 2)(\lambda - 3)$. Los eigenvalores son 1, 2, 3, con las bases correspondientes de eigenvectores $\{(1, -2, 2)\}$, $\{(0, 1, 0)\}$ y $\{(0, 1, 1)\}$. Todas las multiplicidades son 1.

25. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ con eigenvalores 1 y 2.

27. **(a)** 1, -2. **(b)** -4, 4.

29. $(1, 1, 1)$, con eigenvalor $a + b + c$.

31. *Sugerencia:* Primero demuestre que $\det((A - \lambda I)^T) = \det(A^T - \lambda I)$.

33. $\frac{1}{2}\mathbf{v}$

35. *Sugerencia:* Primero demuestre que si A es invertible, entonces $\lambda \neq 0$. A continuación multiplique por la izquierda $A\mathbf{v} = \lambda\mathbf{v}$ por $\lambda^{-1}A^{-1}$.

37. *Sugerencia:* Primero demuestre que $\det(P^{-1}AP - \lambda I) = \det(A - \lambda I)$.

39. $A\mathbf{v} = 0 = 0\mathbf{v}$, porque $\mathbf{v} \in \mathcal{N}(A)$. De manera que $\mathbf{v} \neq \mathbf{0}$ es un eigenvector con eigenvalor 0.

41. Suponga que $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

43. *Sugerencia:* Primero observe que $\mathbf{e}_1, \ldots, \mathbf{e}_n$ son eigenvectores. Si $\lambda_1, \ldots, \lambda_n$ son los eigenvalores correspondientes, entonces $A\mathbf{e}_i = \lambda_i\mathbf{e}_i$. Esto demuestra que A es diagonal. A continuación compruebe que todas las λ_i son iguales.

45. *Sugerencia:* 0 es el único eigenvalor de A. La multiplicidad geométrica de 0 es la dimensión de E_0. Ahora use el ejercicio 32.

47. $C(p) = \begin{bmatrix} 0 & 1 \\ 15 & -2 \end{bmatrix}$. El polinomio característico es $\lambda^2 + 2\lambda - 15$.

49. De acuerdo con el último ejercicio sólo es necesario construir un polinomio mónico con raíces 4 y -5, y después tomar su matriz asociada Así,

$$p(x) = (x - 4)(x + 5) = x^2 + x - 20$$

Por consiguiente, $C(p) = \begin{bmatrix} 0 & 1 \\ 20 & -1 \end{bmatrix}$ debe tener los eigenvalores 4 y -5. Sí los tiene.

51. El polinomio $(x - 4)(x + 5)(x + 2) = x^3 + 3x^2 - 18x - 40$ tiene la matriz asociada

$$C(p) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 40 & 18 & -3 \end{bmatrix}$$

Sus eigenvalores son las raíces de su polinomio característico $-\lambda^3 - 3\lambda^2 + 18\lambda + 40$, que son -5 y -2.

53. *Sugerencia:* primero verifique la afirmación para $n = 2$. Después suponga que es cierta para $n - 1$ y demuéstrela para n.

55. Primero se determinan los eigenvalores y eigenvectores de la matriz $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ de T con respecto a la base estándar $\{1, x\}$.

Los eigenvalores son 1 y -1, con sus eigenvectores básicos correspondientes $(1, 1)$ y $(-1, 1)$. De manera que T tiene eigenvalores 1 y -1, cuyos eigenvectores básicos correspondientes $x + 1$ y $-x + 1$.

Sección 7.2

1. $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$

3. $P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 4 \\ 1 & 0 & -4 \end{bmatrix}, D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -7 \end{bmatrix}$

5. No es diagonalizable.

7. $A(10,0) = (-10,0)$, de modo que $(10, 0)$ es un eigenvector con eigenvalor -1. $A(6, 5) = (24, 20)$, así que $(6, 5)$ es un eigenvector con eigenvalor 4. Como $(10, 0)$ y $(6, 5)$ pertenecen a distintos eigenvalores, son linealmente independientes. A es diagonalizable, con $P = \begin{bmatrix} 10 & 6 \\ 0 & 5 \end{bmatrix}$ y $\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$.

9. Porque

$$A \begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 0 \\ 0 & 2 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 2 & -2 \\ 4 & 0 & 0 \\ 0 & 4 & 4 \end{bmatrix}$$

se ve que $(0, 2,0), (1, 0, 2), (1, 0, -2)$ son eigenvectores cuyos eigenvalores correspondientes son 2, 2 y -2. Es obvio que los dos primeros eigenvectores son linealmente independientes. Por ello, S es linealmente independiente, porque $(1, 0, -2)$ pertenece a un eigenvalor distinto. A es diagonalizable, con $P = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 0 \\ 0 & 2 & -2 \end{bmatrix}$ y $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$.

11. Sean $P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. Entonces

$$A = PDP^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

13. Sean $P = \begin{bmatrix} 1 & -3 & -2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, D = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Entonces

$$A = PDP^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$$

15. A tiene tres eigenvalores distintos, por consiguiente es diagonalizable, según el teorema 8.

17. $\{(1, 1, 1), (-2, 1, 0), (-2, 0, 1)\}$.

19. A es diagonalizable, porque tiene tres eigenvectores linealmente independientes, que son $(1, 1, 1)$, $(-3, 2, 0)$, $(-2, 0, 1)$.

21. A no es diagonalizable, porque 5, su único eigenvalor, sólo tiene un (< 3) eigenvector básico, $(1, 0, 0)$.

25. La matriz es diagonalizable sobre los números reales, sí y sólo si $a > 0$.

27. $$A^6 = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -4 \end{bmatrix}^6 \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 4096 & 0 \\ 0 & 4096 \end{bmatrix} = \begin{bmatrix} 2^{12} & 0 \\ 0 & 2^{12} \end{bmatrix}$$

$$A^9 = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -4 \end{bmatrix}^9 \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 524,288 \\ 131,072 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2^{19} \\ 2^{17} & 0 \end{bmatrix}$$

29. $P = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ -1 & -1 & 2 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ diagonaliza la matriz. Por tanto,

$$PD^7P^{-1} = \begin{bmatrix} 2 \cdot 5^6 & 2 \cdot 5^6 & 2 \cdot 5^6 \\ 5^6 & 5^6 & 5^6 \\ 2 \cdot 5^6 & 2 \cdot 5^6 & 2 \cdot 5^6 \end{bmatrix} = \begin{bmatrix} 31,250 & 31,250 & 31,250 \\ 15,625 & 15,625 & 15,625 \\ 31,250 & 31,250 & 31,250 \end{bmatrix}$$

31. La identidad es cierta, porque si la matriz base se diagonaliza en la izquierda se obtiene

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}^{k+1} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}^{k+1} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 2^k & 2^k \\ 2^k & 2^k \end{bmatrix}$$

para $k = 0, 1, 2, \ldots$

35. Sea $\mathcal{B} = \{1, x, x^2, x^3\}$ la base estándar de P_3. Entonces la matriz de $T : P_3 \to P_3$

$$T(a + bx + cx^2 + dx^3) = a$$

con respecto a $\mathcal{B}$ es

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A es diagonalizable. Sus eigenvaores son 0 y 1, y sus bases correspondientes de eigenvectores $\{\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ y $\{\mathbf{e}_1\}$. Por ello, T es diagonalizable. $\mathcal{B}$ es una base de P_3 que contiene eigenvectores de T.

37. Sea T la reflexión dada. Entonces, la matriz estándar de T es

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

que tiene los eigenvalores 1 y -1, y las bases correspondientes de eigenvectores $\{(1, 0)\}$, $\{(0, 1)\}$. En consecuencia, A es diagonalizable y T también lo es. La base estándar $\{\mathbf{e}_1, \mathbf{e}_2\}$ es una base de $\mathbf{R}^2$ formada por eigenvectores de T.

39. Sea T la proyección dada. La matriz estándar de T es

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

cuyos eigenvalores son 0, 1, y sus bases correspondientes de eigenvectores son $\{(0, 1)\}$ y $\{(1, 0)\}$. Por consiguiente, A es diagonalizable. Entonces T es diagonalizable. La base estándar $\{\mathbf{e}_1, \mathbf{e}_2\}$ es una base de $\mathbf{R}^2$ que consiste en eigenvectores de T.

41. Sea T la proyección dada. La matriz estándar de T es

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

cuyos eigenvalores son 0, 1, y sus bases correspondientes de eigenvectores son $\{(0, 0, 1)\}$ y $\{(1, 0, 0), (0, 1, 0)\}$. De modo que, A es diagonalizable. Y por ende T también lo es. La base estándar $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ es una base de $\mathbf{R}^3$ formada por eigenvectores de T.

Sección 7.3

1. 5 es un eigenvalor con eigenvector $(1, 1)$.

3. $(2, 1)$ es un eigenvector con el eigenvalor correspondiente 4.

5. $(2, 1)$ es un eigenvector con el eigenvalor correspondiente $\frac{1}{4}$.

7. Para $\mathbf{x} = (1, 2)$

$$A^3\mathbf{x} = \begin{bmatrix} 61 \\ -64 \end{bmatrix} \qquad A^4\mathbf{x} = \begin{bmatrix} -311 \\ 314 \end{bmatrix}$$
$$\lambda_{\text{appr}} = -5.0984 \qquad \lambda = -5$$
$$\mathbf{v}_{\text{appr}} = \begin{bmatrix} -0.99 \\ 1.0 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

9. Para $\mathbf{x} = (1, 2)$

$$A^3\mathbf{x} = \begin{bmatrix} 170 \\ -173 \end{bmatrix} \qquad A^4\mathbf{x} = \begin{bmatrix} -1199 \\ 1202 \end{bmatrix}$$
$$\lambda_{\text{appr}} = -7.0529 \qquad \lambda = -7$$
$$\mathbf{v}_{\text{appr}} = \begin{bmatrix} -0.997 \\ 1.0 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

11. Para $\mathbf{x} = (1, 2)$

$$A^3\mathbf{x} = \begin{bmatrix} 363 \\ -366 \end{bmatrix} \qquad A^4\mathbf{x} = \begin{bmatrix} -3279 \\ 3282 \end{bmatrix}$$
$$\lambda_{\text{appr}} = -9.0331 \qquad \lambda = -9$$
$$\mathbf{v}_{\text{appr}} = \begin{bmatrix} -0.999 \\ 1.0 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

13. Para $\mathbf{x} = (1, 2)$

$$A^3\mathbf{x} = \begin{bmatrix} 364 \\ -148 \end{bmatrix} \qquad A^4\mathbf{x} = \begin{bmatrix} -2924 \\ 1172 \end{bmatrix}$$
$$\lambda_{\text{appr}} = -8.033 \qquad \lambda = -8$$
$$\mathbf{v}_{\text{appr}} = \begin{bmatrix} 1.0 \\ -0.4008 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 1 \\ -\frac{2}{5} \end{bmatrix}$$

15. Para $\mathbf{x} = (1, 2)$

$$A^3\mathbf{x} = \begin{bmatrix} 2395 \\ 1598 \end{bmatrix} \qquad A^4\mathbf{x} = \begin{bmatrix} 26,353 \\ 17,570 \end{bmatrix}$$
$$\lambda_{\text{appr}} = 11.003 \qquad \lambda = 11$$
$$\mathbf{v}_{\text{appr}} = \begin{bmatrix} 1.0 \\ 0.66672 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}$$

17. Para $\mathbf{x} = (1, 2)$

$$A^3\mathbf{x} = \begin{bmatrix} 1412 \\ -316 \end{bmatrix} \qquad A^4\mathbf{x} = \begin{bmatrix} -16,964 \\ 3772 \end{bmatrix}$$
$$\lambda_{\text{appr}} = -12.014 \qquad \lambda = -12$$
$$\mathbf{v}_{\text{appr}} = \begin{bmatrix} 1.0 \\ -0.22235 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 1 \\ -\frac{2}{9} \end{bmatrix}$$

19. Para $\mathbf{x} = (1, 2)$

$$A^3\mathbf{x} = \begin{bmatrix} 1265 \\ -1479 \end{bmatrix} \qquad A^4\mathbf{x} = \begin{bmatrix} -17,729 \\ 20,687 \end{bmatrix}$$

$$\lambda_{\text{appr}} = -14.015 \qquad \lambda = -14$$

$$\mathbf{v}_{\text{appr}} = \begin{bmatrix} -0.85701 \\ 1.0 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} -\frac{6}{7} \\ 1 \end{bmatrix}$$

21. Comenzando en (1, 2), las cuatro primeras iteraciones producen −1.4000, −3.9412, −4.9432, −4.9977. Así, el eigenvalor dominante es −5.

23. Empezando en (1, 2), las cuatro primeras iteraciones producen −1.6000, −6.0690, −6.9776, −6.9995. Así, el eigenvalor dominante es −7.

25. Iniciando en (1, 2), las cuatro primeras iteraciones producen 0.2800, 0.7882, 0.9886, 0.9995. El eigenvalor verdadero más próximo al origen es 1, y el más cercano es 0.9995.

27. Comenzando en (1, 2), las cuatro primeras iteraciones producen 0.2286, 0.8670, 0.9968, 0.9999 El eigenvalor verdadero más cercano al origen es 1, y el más aproximado es 0.9999.

29. Empezando en (1, 2), las cuatro primeras iteraciones producen 0.2000, 0.9111, 0.9988, 1.0000. El eigenvalor verdadero más próximo al origen es 1, y el más cercano es 1.0000.

31. Iniciando en (1, 2), las cuatro primeras iteraciones producen −1.1111, −1.0194, −1.0022, −1.0022. El eigenvalor verdadero más cercano al origen es −1, y el más aproximado es −1.0022.

33. Comenzando en (1, 2), las cuatro primeras iteraciones producen −1.1000, −1.0205, −1.0026, −1.0003. El eigenvalor verdadero más cercano al origen es −1, y el más aproximado es −1.0003.

35. Al aplicar el método de potencia inversa desplazada a $\begin{bmatrix} 0 & 1 \\ -4 & 5 \end{bmatrix}$ y comenzando en (1, 0) se obtiene −1.0139 después de 4 iteraciones. De modo que, la raíz más cercana a 5 es $5 + (-1.0139)^{-1} = 4.0138$.

37. Al emplear el método de potencia inversa desplazada a $\begin{bmatrix} 0 & 1 \\ -6 & 7 \end{bmatrix}$ e iniciando en (1, 0) se obtiene 0.9780 después de 4 iteraciones. Así, la raíz más cercana a 5 es $5 + (0.9780)^{-1} = 6.0225$.

39. Aplicando el método de potencia inversa desplazada a $\begin{bmatrix} 0 & 1 \\ 9 & 8 \end{bmatrix}$ y comenzando en (1, 0) se obtiene −1.0006 después de 4 iteraciones. Así, la raíz más cercana a 10 es $10 + (-1.0006)^{-1} = 9.0006$.

Sección 7.4

1. (a) $\mathbf{x}_k = \frac{3}{2} \cdot 1^k \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{5}{2} \cdot 5^k \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $A\mathbf{x}_0 = \mathbf{x}_1 = \begin{bmatrix} 11 \\ 14 \end{bmatrix}$; $A^2\mathbf{x}_0 = \mathbf{x}_2 = \begin{bmatrix} 61 \\ 64 \end{bmatrix}$

(c) Ninguno

3. (a) $\mathbf{x}_k = \frac{3}{2} \cdot (-7)^k \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{5}{2} \cdot (-1)^k \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $A\mathbf{x}_0 = \mathbf{x}_1 = \begin{bmatrix} 8 \\ -13 \end{bmatrix}$; $A^2\mathbf{x}_0 = \mathbf{x}_2 = \begin{bmatrix} -71 \\ 76 \end{bmatrix}$

(c) Ninguno

5. (a) $\mathbf{x}_k = \frac{3}{2} \cdot \left(\frac{1}{2}\right)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{5}{2} \cdot \left(\frac{5}{2}\right)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $A\mathbf{x}_0 = \mathbf{x}_1 = \begin{bmatrix} \frac{11}{2} \\ 7 \end{bmatrix}$; $A^2\mathbf{x}_0 = \mathbf{x}_2 = \begin{bmatrix} \frac{61}{4} \\ 16 \end{bmatrix}$

(c) Ninguno

7. (a) $\mathbf{x}_k = \frac{3}{2} \cdot (-13)^k \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{5}{2} \cdot (-1)^k \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $A\mathbf{x}_0 = \mathbf{x}_1 = \begin{bmatrix} 17 \\ -22 \end{bmatrix}$; $A^2\mathbf{x}_0 = \mathbf{x}_2 = \begin{bmatrix} -251 \\ 256 \end{bmatrix}$

(c) Ninguno

9. (a) $\mathbf{x}_k = \frac{3}{2} \cdot \left(-\frac{1}{10}\right)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{5}{2} \cdot \left(\frac{1}{2}\right)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $A\mathbf{x}_0 = \mathbf{x}_1 = \begin{bmatrix} \frac{7}{5} \\ \frac{11}{10} \end{bmatrix}$; $A^2\mathbf{x}_0 = \mathbf{x}_2 = \begin{bmatrix} \frac{61}{100} \\ \frac{16}{25} \end{bmatrix}$

(c) Atractor.

11. (a) $\begin{bmatrix} -3 & 2 \\ 2 & -3 \end{bmatrix}^5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

(b) $\mathbf{x}_5 = 0 \cdot (-5)^5 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + 1 \cdot (-1)^5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

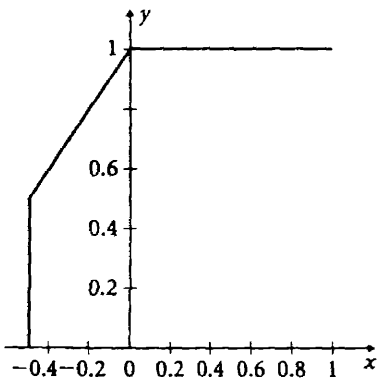
13. (a) $\begin{bmatrix} \frac{3}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{3}{10} \end{bmatrix}^5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{32} \\ \frac{1}{32} \end{bmatrix}$

(b) $\mathbf{x}_5 = 0 \cdot \left(\frac{1}{10}\right)^5 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + 1 \cdot \left(\frac{1}{2}\right)^5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{32} \\ \frac{1}{32} \end{bmatrix}$

15. (a) $\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}^5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 59,049 \\ 59,049 \end{bmatrix}$

(b) $\mathbf{x}_5 = 0 \cdot 1^5 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + 1 \cdot 9^5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 59,049 \\ 59,049 \end{bmatrix}$

17. $\mathbf{x}_1 = (0, 1)$, $\mathbf{x}_2 = (-\frac{1}{2}, \frac{1}{2})$, $\mathbf{x}_3 = (-\frac{1}{2}, 0)$. El origen es un atractor.



19. Es cierto, porque ambos eigenvalores, $\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$, son raíces sextas de 1. Es decir $\left(\frac{1}{2} \pm \frac{\sqrt{3}}{2}i\right)^6 = 1$.

21. En notación matricial:

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix}^k \begin{bmatrix} A_k \\ B_k \end{bmatrix}$$

$$\mathbf{x}_0 = \begin{bmatrix} A_0 \\ B_0 \end{bmatrix} = \begin{bmatrix} 300 \\ 100 \end{bmatrix}$$

(a) Después de la tercera generación hay 311 hembras y 89 machos.

(b) A largo plazo, hay 311.11 hembras y 88.889 machos. Así, al final las hembras predominarán en la población.

Sección 7.5

1. El segundo y el cuarto.
3. (a) La matriz no es regular porque es triangular inferior, debido a esto todas sus potencias también son triangulares inferiores; en consecuencia siempre hay elementos que son 0.
   (b) Como la matriz es triangular superior no es regular, todas sus potencias también son triangulares superiores, y por tanto siempre habrá elementos que son 0.
5. Sea $A = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix}$. Al resolver $[A - I : \mathbf{0}]$ se obtiene $(2r, r)$. Pero se desea $2r + r = 1$. Así, $r = \frac{1}{3}$. Por consiguiente, $\mathbf{v} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$ es el vector de estado estable de A. Sea $B = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{bmatrix}$. Al resolver $[B - I : \mathbf{0}]$ se obtiene $(r/2, r)$. Pero lo que deseamos es $r/2 + r = 1$.

   Así, $r = \frac{2}{3}$. Por consiguiente, $\mathbf{v} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$ es el vector de estado estable de B.

7. (a) $\left(\frac{5}{13}, \frac{2}{13}, \frac{6}{13}\right)$ (b) $\left(\frac{1}{4}, \frac{13}{36}, \frac{7}{18}\right)$
9. Suponga que $\mathbf{u}$ es otro vector de probabilidad tal que $A\mathbf{u} = \mathbf{u}$. Entonces, $\lim_{k\to\infty} A^k\mathbf{u} = \mathbf{v}$. Y como $A\mathbf{u} = \mathbf{u}$, entonces $A^k\mathbf{u} = \mathbf{u}$ para toda k. Por consiguiente, $\mathbf{u} = \mathbf{v}$.
11. M es regular, porque M^2 sólo tiene elementos positivos. El equilibrio de M es $\left(\frac{5}{13}, \frac{2}{13}, \frac{6}{13}\right)$. Así, a largo plazo el plan C es el más popular, y B el menos popular.

Capítulo 8

Sección 8.1

1. $(1, -2, 1) \cdot (4, 2, 0) = 0$
   $(1, -2, 1) \cdot (-1, 2, 5) = 0$
   $(4, 2, 0) \cdot (-1, 2, 5) = 0$
   Este conjunto forma una base ortogonal de $\mathbf{R}^3$.
3. $(1, 1, -1, 1) \cdot (1, 1, 1, -1) = 0$
   $(1, 1, -1, 1) \cdot (0, 0, 1, 1) = 0$
   $(1, 1, 1, -1) \cdot (0, 0, 1, 1) = 0$
   Este conjunto no forma una base ortogonal de $\mathbf{R}^4$.
5. $\mathbf{v}_1 = (1, 1, 2), \mathbf{v}_2 = (-1, 1, 0), \mathbf{v}_3 = (0, 0, 1)$
7. El producto punto de cada par de vectores vale cero; en consecuencia, los vectores son linealmente independientes. Entonces, forman una base ortogonal de $\mathbf{R}^3$.

$$(1, 1, 1) = \frac{3}{19}\mathbf{v}_1 + \frac{1}{19}\mathbf{v}_2 + 1\mathbf{v}_3$$

9. Cada par de vectores tiene el producto punto que vale cero; por tanto, los vectores son linealmente independientes. Entonces, forman una base ortogonal de $\mathbf{R}^3$.

$$(1, 1, 1) = 0\mathbf{v}_1 + \tfrac{1}{7}\mathbf{v}_2 + \tfrac{1}{7}\mathbf{v}_3$$

11. No es ortonormal. Al ortonormalizarlo se obtiene

$$\begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

13. Sí es ortonormal.
15. $\mathcal{B}$ es una base ortonormal para $\mathbf{R}^2$, porque ambos vectores son unitarios y su producto punto es cero.

$$\mathbf{e}_1 = \frac{2}{\sqrt{5}}\mathbf{v}_1 - \frac{1}{\sqrt{5}}\mathbf{v}_2$$

17. Es ortogonal; su inversa es $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

19. Es ortogonal; su inversa es

$$\begin{bmatrix} \frac{3}{\sqrt{14}} & -\frac{2}{\sqrt{14}} & \frac{1}{\sqrt{14}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{21}} & -\frac{1}{\sqrt{21}} & \frac{4}{\sqrt{21}} \end{bmatrix}$$

21. No, posiblemente se tengan columnas cero.

Sección 8.2

1. $\mathbf{u}_{pr} = \dfrac{(-2,1)\cdot(1,2)}{(1,2)\cdot(1,2)}(1,2) = (0,0)$

3. $\mathbf{u}_{pr} = \dfrac{(3,-1,2)\cdot(1,1,-3)}{(1,1,-3)\cdot(1,1,-3)}(1,1,-3)$
$= \left(-\frac{4}{11}, -\frac{4}{11}, \frac{12}{11}\right)$

5. $\mathbf{u} = \left(-\frac{2}{13}, -\frac{3}{13}\right) + \left(-\frac{24}{13}, \frac{16}{13}\right)$

7. $\mathbf{u} = (-1,1) + (2,2)$

9. $\mathbf{u}_{pr} = (0,0,1)$

13. Ortogonal:

$$\left\{ \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{4}{3} \\ \frac{7}{3} \\ -\frac{1}{3} \end{bmatrix}, \begin{bmatrix} -\frac{1}{11} \\ \frac{1}{11} \\ \frac{3}{11} \end{bmatrix} \right\}$$

Ortonormal:

$$\left\{ \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} \frac{4}{\sqrt{66}} \\ \frac{7}{\sqrt{66}} \\ -\frac{1}{\sqrt{66}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{11}} \\ \frac{1}{\sqrt{11}} \\ \frac{3}{\sqrt{11}} \end{bmatrix} \right\}$$

15. $\left\{ \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} \frac{1}{21} \\ \frac{32}{21} \\ \frac{68}{21} \end{bmatrix} \right\}$

17. $\mathbf{u} = \begin{bmatrix} \frac{80}{103} \\ \frac{70}{103} \\ -\frac{37}{103} \end{bmatrix} + \begin{bmatrix} \frac{126}{103} \\ -\frac{70}{103} \\ \frac{140}{103} \end{bmatrix}$

19. $\|\mathbf{u}_c\| = \left\| \begin{bmatrix} \frac{8}{15} \\ -\frac{8}{3} \\ -\frac{16}{15} \end{bmatrix} \right\| = \frac{8}{15}\sqrt{30}$

21. $\|\mathbf{u}_c\| = \left\| \begin{bmatrix} -\frac{7}{10} \\ \frac{3}{2} \\ \frac{7}{5} \\ 1 \end{bmatrix} \right\| = \frac{1}{10}\sqrt{570}$

Sección 8.3

1. $\begin{bmatrix} 0 & 1 \\ +1 & 0 \end{bmatrix} \begin{bmatrix} +1 & +3 \\ 0 & +2 \end{bmatrix}$

3. $\begin{bmatrix} 0 & 0 & +1 \\ 0 & +1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} +1 & 0 & -2 \\ 0 & +1 & +2 \\ 0 & 0 & +4 \end{bmatrix}$

5. $\begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ 0 & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \end{bmatrix} \begin{bmatrix} \sqrt{2} & -\sqrt{2} & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{6} \end{bmatrix}$

7. $\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{5}}{10} \\ \frac{1}{2} & \frac{\sqrt{5}}{10} \\ \frac{1}{2} & \frac{3\sqrt{5}}{10} \\ -\frac{1}{2} & \frac{3\sqrt{5}}{10} \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 0 & 2\sqrt{5} \end{bmatrix}$

9. $\begin{bmatrix} 2 & 0 \\ 0 & \sqrt{6} \end{bmatrix}$

11. Sea $Q = A$. Esa Q es aceptable, porque A tiene columnas ortogonales. Entonces $R = Q^TA = A^TA = I$, porque A es ortonormal. Así, $A = AI$ es una factorización QR de A.

15. Ya que,

$$A_1 = \begin{bmatrix} 8.8916 & 6.1382 \\ 0.13817 & 1.1075 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 8.9869 & 6.0159 \\ 0.015947 & 1.0121 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 8.9970 & 6.0013 \\ 0.0015697 & 1.0013 \end{bmatrix}$$

los eigenvalores estimados son 8.9970, 1.0013. Los eigenvalores reales son 9 y 1. Los errores son 0.003 y −0.0013.

17. Ya que

$$A_1 = \begin{bmatrix} 13.864 & 9.1909 \\ 0.19135 & 1.1367 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 13.992 & 9.0157 \\ 0.015813 & 1.0109 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 14.002 & 8.9943 \\ -0.0054328 & 0.99650 \end{bmatrix}$$

los eigenvalores estimados son 14.002, 0.99650. Los eigenvalores reales son 14 y 1. Los errores son −0.002 y 0.0035.

Sección 8.4

1. Las ecuaciones normales

$$\begin{bmatrix} 3 & -1 \\ -1 & 9 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$$

dan como resultado $\tilde{\mathbf{x}} = \begin{bmatrix} -\frac{6}{13} \\ \frac{8}{13} \end{bmatrix}$.

3. Las ecuaciones normales

$$\begin{bmatrix} 10 & 5 \\ 5 & 6 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

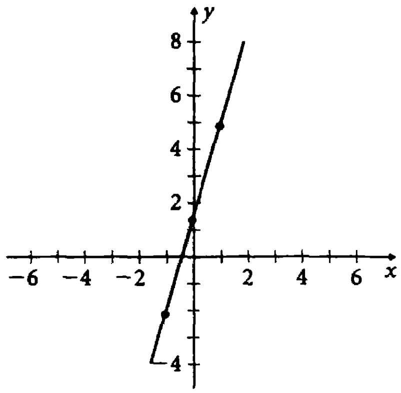
dan como resultado $\tilde{\mathbf{x}} = \begin{bmatrix} -.11864 \\ .18644 \end{bmatrix}$.

5. Todo lo que se necesita es un sistema con matriz invertible de coeficientes. Por ejemplo:

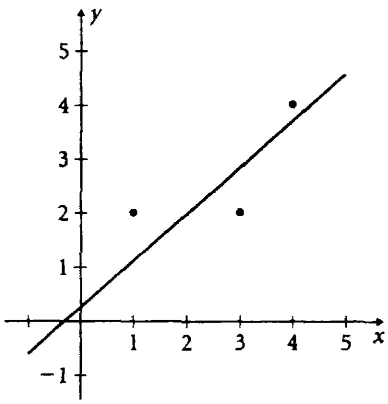
$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 9 \\ 6 \\ 3 \end{bmatrix}$$

cuya solución es (1, 7, −1), por eliminación directa o mediante las ecuaciones normales.

7. $y = \frac{7}{2}x + \frac{4}{3}$



9. $y = \frac{51}{59}x + \frac{14}{59}$



$$\begin{bmatrix} 5 & 1 & -3 \\ 1 & 1 & 1 \\ -3 & 1 & 5 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} -2 \\ 2 \\ 6 \end{bmatrix}$$

se obtiene $\tilde{\mathbf{x}} = (r-1, -2r+3, r)$, $r \in \mathbf{R}$.

13. El sistema $R\tilde{\mathbf{x}} = Q^T\mathbf{b}$,

$$\begin{bmatrix} 5 & 10 \\ 0 & 5 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} -4 \\ \frac{9}{5} \end{bmatrix}$$

da como resultado $\tilde{\mathbf{x}} = (-\frac{38}{25}, \frac{9}{25})$

17. Según el ejercicio 15,

$$\tilde{\mathbf{x}} = \begin{bmatrix} \frac{5}{9} \\ -\frac{1}{9} \end{bmatrix}$$

19. Según el ejercicio 15,

$$\tilde{\mathbf{x}} = \begin{bmatrix} \frac{5/3}{1} \\ \frac{-1/3}{1} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ -\frac{1}{3} \end{bmatrix}$$

Sección 8.5

1. No es simétrica.

3. No es simétrica.

5. $A^2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

7. $A^{-1} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}$ de acuerdo con el ejercicio 6.

9. $Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$

11. $Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, D = \begin{bmatrix} 2 & 0 \\ 0 & 10 \end{bmatrix}$

13. $Q = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

15. $Q = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$

$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

17. *Sugerencia:* Demuestre que Q y D^2 diagonalizan ortogonalmente a A^2.

19. $Q = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, T = \begin{bmatrix} 1 & -6 \\ 0 & 11 \end{bmatrix}$

21. $Q = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, T = \begin{bmatrix} 1 & -7 \\ 0 & 12 \end{bmatrix}$

23. Tenemos que

$$\begin{aligned} A &= QDQ^T \\ &= [\mathbf{v}_1 \cdots \mathbf{v}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix}^T \\ &= [\lambda_1\mathbf{v}_1 \cdots \lambda_n\mathbf{v}_n] \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix}^T \\ &= \lambda_1\mathbf{v}_1\mathbf{v}_1^T + \cdots + \lambda_n\mathbf{v}_n\mathbf{v}_n^T \end{aligned}$$

25. $5 \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}^T + 15 \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}^T$

Sección 8.6

1. $q(\mathbf{x}) = -2x^2 + 3y^2 + 4xy$

3. $q(\mathbf{x}) = x^2 + 5z^2 - 6xy + 4xz + 2yz$

5. $\begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix}$

7. $\begin{bmatrix} -4 & 1 \\ 1 & -4 \end{bmatrix}$

9. $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 5 & -4 & 0 \\ -4 & 3 & 6 \\ 0 & 6 & 1 \end{bmatrix}$

13. La matriz de la forma es $\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$, que se diagonaliza ortogonalmente con

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

Sea $\mathbf{y} = (x', y')$. Entonces

$$\mathbf{y} = Q^T\mathbf{x} = \begin{bmatrix} \frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} \\ -\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} \end{bmatrix}$$

y

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{y}^T \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \mathbf{y} \\ &= 2x'^2 + 4y'^2 \end{aligned}$$

15. La matriz de la forma es $\begin{bmatrix} -4 & 2 \\ 2 & -4 \end{bmatrix}$, que se diagonaliza ortogonalmente con

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad D = \begin{bmatrix} -2 & 0 \\ 0 & -6 \end{bmatrix}$$

Sea $\mathbf{y} = (x', y')$. Entonces

$$\mathbf{y} = Q^T\mathbf{x} = \begin{bmatrix} \frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} \\ -\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} \end{bmatrix}$$

y

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{y}^T \begin{bmatrix} -2 & 0 \\ 0 & -6 \end{bmatrix} \mathbf{y} \\ &= -2x'^2 - 6y'^2 \end{aligned}$$

17. La matriz de la forma es $\begin{bmatrix} 2 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 2 \end{bmatrix}$, que se diagonaliza ortogonalmente con

$$Q = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Sea $\mathbf{y} = (x', y', z')$. Entonces

$$\mathbf{y} = Q^T\mathbf{x} = \begin{bmatrix} y \\ \frac{x}{\sqrt{2}} + \frac{z}{\sqrt{2}} \\ -\frac{x}{\sqrt{2}} + \frac{z}{\sqrt{2}} \end{bmatrix}$$

y

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{y}^T \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{y} \\ &= y'^2 + 3z'^2 \end{aligned}$$

19. La matriz de la forma es $\begin{bmatrix} 5 & -4 & 0 \\ -4 & 3 & 4 \\ 0 & 4 & 1 \end{bmatrix}$, que se diagonaliza ortogonalmente con

$$Q = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$$D = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Sea $\mathbf{y} = (x', y', z')$. Entonces

$$\mathbf{y} = Q^T\mathbf{x} = \begin{bmatrix} \frac{1}{3}x + \frac{2}{3}y - \frac{2}{3}z \\ -\frac{2}{3}x + \frac{2}{3}y + \frac{1}{3}z \\ \frac{2}{3}x + \frac{1}{3}y + \frac{2}{3}z \end{bmatrix}$$

y

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{y}^T \begin{bmatrix} -3 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{y} \\ &= -3x'^2 + 9y'^2 + 3z'^2 \end{aligned}$$

21. La matriz de la forma es $\begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$, que se diagonaliza ortogonalmente con

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}$$

Sea $\mathbf{y} = (x', y')$. Entonces

$$\mathbf{y} = Q^T\mathbf{x} = \begin{bmatrix} \frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} \\ -\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} \end{bmatrix}$$

y

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{y}^T \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix} \mathbf{y} \\ &= x'^2 + 9y'^2 \end{aligned}$$

Por consiguiente, la sección cónica es una elipse.

25. *Sugerencia:* Primero demuestre que Q y D^2 diagonalizan ortogonalmente a A^2.

27. *Sugerencia:* Considere $P = BQ^T$, donde B es la matriz diagonal cuyos elementos diagonales son las raíces cuadradas de los elementos diagonales de D.

29. Tenemos

$$\begin{aligned} q(x,y) &= ax^2 + bxy + cy^2 \\ &= a\left(x^2 + 2x\frac{b}{2a}y + \frac{c}{a}y^2 + \frac{b^2}{4a^2}y^2 - \frac{b^2}{4a^2}y^2\right) \end{aligned}$$

$$= a\left(x + \frac{b}{2a}y\right)^2 + \left(c - \frac{b^2}{4a}\right)y^2$$
$$= aX^2 + By^2$$

donde $X = x + (b/2a)y$ y $B = c - b^2/4a$.

31. Sí, como sigue:

$$\begin{aligned} q(x,y) &= bxy + cy^2 \\ &= c\left(y^2 + 2\frac{b}{2c}xy + \frac{b^2}{4c^2}x^2 - \frac{b^2}{4c^2}x^2\right) \\ &= c\left(y + \frac{b}{2c}x\right)^2 + \frac{-b^2}{4c}x^2 \\ &= cY^2 + Ax^2 \end{aligned}$$

en donde $Y = y + (b/2c)x$ y $A = -b^2/4c^2$.

Sección 8.7

1. 3, 2, 0

3. 2, 1, 0

5. $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}^T$

7. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T$

9. $\begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 1 & 0 \\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} 10 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\times \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 1 & 0 \\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix}^T$

11. $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T$

13. A^TA es complicada, pero AA^T es diagonal. Una descomposición de A^T en valores singulares es

$$\begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}^T$$

Por consiguiente, se obtiene una descomposición de A en valores singulares por transposición:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}^T$$

15. (a) La verificación de las propiedades de Moore-Penrose es directa, dada:

$$\begin{aligned} A^+ &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^T \\ &= \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \end{aligned}$$

(b) La verificación de las propiedades de Moore-Penrose es directa:

$$\begin{aligned} A^+ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^T \\ &= \begin{bmatrix} 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned}$$

17. A^+ fue calculada en el ejercicio 15(**b**), y se obtuvo $\begin{bmatrix} 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 \\ 1 & 0 & 0 \end{bmatrix}$, que también es A^{-1}.

19.

$$AA^+A = \begin{bmatrix} -3 & 0 \\ 0 & 0 \\ 0 & 4 \end{bmatrix}$$

$$A^+AA^+ = \begin{bmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$$

$$(AA^+)^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$AA^+ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(A^+A)^T = I$$
$$A^+A = I$$

21. Basta comprobar las condiciones de Moore-Penrose para (A^+, A), porque entonces A sería la inversa única de A^+. Así, $A^{++} = A$. De acuerdo con el ejercicio 18,

$$A^+AA^+ = A^+, \quad AA^+A = A,$$
$$(A^+A)^T = A^TA, \quad (AA^+)^T = AA^+$$

Por consiguiente, las condiciones son válidas para el par (A^+, A).

23. $A^+\mathbf{b} = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{3}{5} \end{bmatrix}$

25. $A^+\mathbf{b} = \begin{bmatrix} -\frac{2}{9} & \frac{1}{9} & \frac{2}{9} \\ \frac{2}{27} & \frac{2}{27} & \frac{1}{27} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$

27. (a) $A = PQ = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

(b) $A = PQ = \begin{bmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$

Sección 8.8

1. (a) $-12, \sqrt{19}, 2\sqrt{21}, \sqrt{79}$
   (b) $\|(-3,-5)\| = \sqrt{127}$
   (c) $\sqrt{19}\cdot 2\sqrt{21} \cong 39.95 \geq |-12| = 12$
   (d) $\sqrt{79} \cong 8.8882 \leq \sqrt{19} + 2\sqrt{21} \cong 13.524$
   (e) $\frac{1}{4}(79 - 127) = -12$

3. $\left(1, \frac{3}{8}\right)$

5. A y B son ortogonales. También A y C son ortogonales.

7. Los pares ortonormales son (A', B') y (A', C'), siendo

$$A' = \frac{1}{\|A\|}A = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$B' = \frac{1}{\|B\|}B = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$C' = \frac{1}{\|C\|}C = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

9. Es, porque A es positiva definida

.11. Es, porque A es positiva definida.

13. No lo es. f no es positiva definida, porque $f(\mathbf{u}, \mathbf{u}) = 0$ cuando $\mathbf{u} = (-1, 1)$.

15. No lo es. f no es simétrica, porque para $\mathbf{u} = (1, 2)$ y $\mathbf{v} = (2, 1)$, se tiene que $f(\mathbf{u}, \mathbf{v}) = 31$, mientras que $f(\mathbf{v}, \mathbf{u}) = 49$.

17. *Sugerencia:* Compruebe los 4 axiomas.

19. (a) $-255, \sqrt{339}, 15, 3\sqrt{6}$
   (b) $\|1 - 3x^2 + 2x\| = \sqrt{1074}$
   (c) $|-255| = 255 \leq \sqrt{339}\cdot 15 \cong 276.18$
   (d) $3\sqrt{6} \cong 7.3485 \leq \sqrt{339} + 15 \cong 33.412$

21. $2 + x^2$

23. $\frac{1}{5} + x^2$

29. Aprovechando el cálculo $(1, 1)_{pr} = (\frac{2}{3}, 0)$ del ejercicio anterior, la distancia de $(1, 1)$ a l es

$$\begin{aligned}\|(1,1)_c\| &= \|(1,1) - (1,1)_{pr}\| \\ &= \|(1,1) - (\tfrac{2}{3}, 0)\| = \|(\tfrac{1}{3}, 1)\| \\ &= \sqrt{\tfrac{7}{3}}\end{aligned}$$

31. Sea $f(x) = x^2 + x + 1$. De acuerdo con el ejercicio anterior, $f_{pr} = x^2 + x - \frac{8}{3}$. Por consiguiente, la distancia de $x^2 + x + 1$ a W es

$$\|f_c\| = \|f - f_{pr}\| = \left\|\frac{11}{3}\right\| = \frac{11}{\sqrt{3}}$$

33. Sea $\mathbf{u}_1 = 1$. Entonces

$$\mathbf{u}_2 = x - \tfrac{0}{3}\cdot 1 = x$$

y

$$\begin{aligned}\mathbf{u}_3 &= x^2 - \tfrac{8}{3}1 - \tfrac{0}{4}x \\ &= x^2 - \tfrac{8}{3}\end{aligned}$$

Así, $\{1, x, x^2 - \frac{8}{3}\}$ es una base ortogonal de P_2 para este producto interno.

35. Ya que

$$\begin{aligned}\textstyle\int_{-1}^{1} 1^2 dx &= 2 \\ \textstyle\int_{-1}^{1} x^2 dx &= \tfrac{2}{3} \\ \textstyle\int_{-1}^{1} (\tfrac{3}{2}x^2 - \tfrac{1}{2})^2 dx &= \tfrac{2}{5} \\ \textstyle\int_{-1}^{1} (\tfrac{5}{2}x^3 - \tfrac{3}{2}x)^2 dx &= \tfrac{2}{7}\end{aligned}$$

una base ortonormal para P_3 es

$$\mathcal{L}' = \left\{\tfrac{1}{2}, \tfrac{3}{2}x, \tfrac{15}{4}x^2 - \tfrac{5}{4}, \tfrac{35}{4}x^3 - \tfrac{21}{4}x\right\}$$

Sección 8.9

1. Esta vez se obtiene la solución

$$\mathbf{x} = (1.9847, 0.8318, 3.3135, -4.1813, 4.2094)$$

que difiere de la descrita en la sección en

$$(0.0745, 0.0003, 0.0044, 0.0147, -0.0211)$$

La respuesta es satisfactoria, pero todavía no se obtienen los coeficientes racionales correctos.

3. $q(x) = -\frac{22}{5} + \frac{23}{70}x + \frac{13}{14}x^2; q(6) = 31$

5. $q(x) = \frac{134}{35} + 2x + \frac{2}{7}x^2$; $q(6) = \frac{914}{35} \cong 26.114$

7. $q(x) = -\frac{34}{35} - \frac{11}{4}x + \frac{11}{14}x^2 + \frac{5}{4}x^3$; $q(3) = \frac{158}{5} \cong 31.6$

11. Ya que

$$\int_1^2 1dx = 1, \qquad \int_1^2 xdx = \frac{3}{2}$$
$$\int_1^2 x^2dx = \frac{7}{3}, \qquad \int_1^2 x^3dx = \frac{15}{4}$$

entonces

$$\begin{bmatrix} 1 & \frac{3}{2} & \frac{7}{3} \\ \frac{3}{2} & \frac{7}{3} & \frac{15}{4} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{13}{6} \\ 0 & 1 & 3 \end{bmatrix}$$

Así, $y = -\frac{13}{6} + 3x$.

13. Ya que

$$\int_0^2 1\,dx = 2, \qquad \int_0^2 x\,dx = 2$$
$$\int_0^2 x^2\,dx = \frac{8}{3}, \qquad \int_0^2 x^3\,dx = 4$$
$$\int_0^2 x^4\,dx = \frac{32}{5}$$

entonces

$$\begin{bmatrix} 2 & 2 & 4 \\ 2 & \frac{8}{3} & \frac{32}{5} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{8}{5} \\ 0 & 1 & \frac{18}{5} \end{bmatrix}$$

Así, $y = -\frac{8}{5} + \frac{18}{5}x$.

15. Ya que

$$\int_0^2 1\,dx = 2, \qquad \int_0^2 x\,dx = 2$$
$$\int_0^2 x^2\,dx = \frac{8}{3}, \qquad \int_0^2 x^3\,dx = 4$$
$$\int_0^2 x^4\,dx = \frac{32}{5}, \qquad \int_0^2 x^5\,dx = \frac{32}{3}$$

entonces

$$\begin{bmatrix} 2 & 2 & \frac{8}{3} & 4 \\ 2 & \frac{8}{3} & 4 & \frac{32}{5} \\ \frac{8}{3} & 4 & \frac{32}{5} & \frac{32}{3} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \frac{2}{5} \\ 0 & 1 & 0 & -\frac{12}{5} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Así, $y = \frac{2}{5} - \frac{12}{5}x + 3x^2$.

17. $$\langle 1, \operatorname{sen} nx\rangle = \int_{-\pi}^{\pi} 1 \cdot \operatorname{sen} nx\,dx = -\frac{\cos nx}{n}\Big|_{-\pi}^{\pi} = 0$$

19. $$\begin{aligned}\langle \operatorname{sen} mx, \operatorname{sen} nx\rangle &= \int_{-\pi}^{\pi} \operatorname{sen} mx \operatorname{sen} nx\,dx \\ &= \frac{1}{2}\frac{\operatorname{sen}(m-n)x}{m-n} \\ &\quad - \frac{1}{2}\frac{\operatorname{sen}(m+n)x}{m+n}\Big|_{-\pi}^{\pi} = 0\end{aligned}$$

21. $a_0 = \frac{1}{2}, a_n = 0, b_n = \dfrac{1-(-1)^n}{n\pi}$.

23. Se tiene que,

$$\begin{aligned}\langle \operatorname{sen} nx, \operatorname{sen} mx\rangle &= \int_0^{\pi} \operatorname{sen} nx \operatorname{sen} mx\,dx \\ &= \frac{1}{2}\frac{\operatorname{sen}(n-m)x}{n-m} \\ &\quad - \frac{1}{2}\frac{\operatorname{sen}(n+m)x}{n+m}\Big|_0^{\pi} = 0\end{aligned}$$

25. Se tiene que,

$$\begin{aligned}\langle \operatorname{sen} n\pi x, \operatorname{sen} m\pi x\rangle &= \int_{-1}^{1} \operatorname{sen} n\pi x \operatorname{sen} m\pi x\,dx \\ &= \frac{1}{2}\frac{\operatorname{sen}(n\pi - m\pi)x}{n\pi - m\pi} \\ &\quad - \frac{1}{2}\frac{\operatorname{sen}(n\pi + m\pi)x}{n\pi + m\pi}\Big|_{-1}^{1} = 0\end{aligned}$$

27. Se tiene que,

$$\begin{aligned}\langle 1, \cos m\pi x\rangle &= \int_0^2 \cos m\pi x\,dx \\ &= \frac{\operatorname{sen} m\pi x}{m\pi}\Big|_0^2 = 0\end{aligned}$$

y

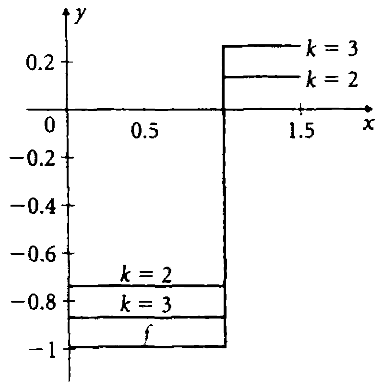
$$\begin{aligned}\langle \cos n\pi x, \cos m\pi x\rangle &= \int_0^2 \cos n\pi x \cos m\pi x\,dx \\ &= \frac{1}{2}\frac{\operatorname{sen}(n\pi - m\pi)x}{n\pi - m\pi} \\ &\quad + \frac{1}{2}\frac{\operatorname{sen}(n\pi + m\pi)x}{n\pi + m\pi}\Big|_0^2 \\ &= 0\end{aligned}$$

29. Ya que para $m = 1, 2, \ldots,$

$$\psi_{m,0} = \begin{cases} 2^{-m/2} & 0 \le x \le 2^{m-1} \\ -2^{-m/2} & 2^{m-1} < x \le 2^m \end{cases}$$

se tiene que

$$c_{m,0} = \int_{-\infty}^{\infty} f(x)\psi_{m,0}(x)dx$$
$$\int_0^1 (-1)\psi_{m,0}(x)dx = \int_0^1 (-2^{-m/2})dx = -2^{-m/2}$$



Índice analítico

Aditividad, 81, 575
Adjunta, 402
Algoritmo
para la base del generador, 277, 280
diagonalización de la matriz simétrica, 548
eliminación de Gauss, 18
generador de imagen fractal, 370
inversión de matrices, 175, 184
para base de espacio nulo, 273
para eigenvalores y eigenvectores, 445
para el código (7,4) de Hamming, 292
proceso de diagonalización, 461
método de la potencia inversa, 475
método de potencia, 471
método de los cocientes de Rayleigh, 473
método de Rayleigh-Ritz, 473
método de la potencia inversa desplazada, 476
método QR, 531
solución de sistema lineal, 23
Análisis de tendencia, 591
Ángulo
entre planos, 126
entre vectores, 84
Anticonmutativa, 115
Apolonio, 61
Aproximaciones a curvas, 54
Árbol generador, 422
Área
del paralelogramo, 118
del triángulo, 118
Arquímedes, 54
Arquímedes
ley de la palanca, 44
problema del ganado, 54
Atractor, 484
Axiomas para espacio vectorial, 236

Balanceo de pesos, 44
Base, 229, 250
cambio de, 268
coordenadas con respecto a, 232
estándar, 230
ordenada, 263
Bisectriz de planos, 142
Bunyakovsky, Víctor, 584

Calor
conducción, 177
transferencia, 74
Cambio de base, 268
Caminata, 199
Campo, 294
Característico(s),(a),(as)
ecuación, 442, 445
matriz, 445
polinomio, 445
raíces de la ecuación secular, 441
valores, 442
Caracterización de matrices invertibles, 185
Cauchy, Augustin Louis, 584
Cayley (Sir) Arthur, 153
Centro de masa, 140
Centroide, 140
Cero
matriz, 155
solución, 6
vector, 64, 236
Circuito, 422
Circuitos eléctricos, 42
Cociente de Rayleigh, 472
Código de Hamming, 290
Código
corrector de errores, 288
de Hamming, 290
detector de errores, 288
lineal, 290
no lineal, 293
Códigos, 211
Cofactor(es), 386
desarrollo por, 386
Columna
matriz, 154
espacio de, 274
Combinación lineal, 67, 237
no trivial, 96
trivial, 96
Composición de transformaciones, 357
Compresión y expansión, 311
Comprobación de paridad
matriz, 290
palabra, 288
Condición inicial, 132
Conjunto generador, 90, 244
reducción del, 92, 247
Conjuntos de sistemas, 53
Contracción, 323
Contradominio, 331
Correspondencia, 105
Cortes, 312
Crecimiento de población, 131, 488
Criterio
para diagonalización, 457
para independencia lineal, 99
para subespacio, 239
Cruce
diagrama, 410
número, 410

Cuadrado(a)
matriz, 33, 154
sistema, 31
Cuadrados mágicos, 49

Decodificación, 212
Definida positiva
matriz, 560
forma cuadrática, 560
Definitivamente positivo, 81
Demanda
función, 40
vector, 209
Desarrollo de Laplace, 386
Descartes, René, 62
Descomposición de Schur, 548
Descomposición polar
de una matriz, 572
de un vector, 89
Desigualdad de Bessel, 511
Desigualdad de Cauchy-Schwarz, 82, 583
Desigualdad de Cauchy-Schwarz-Bunyakovsky, 583
Desigualdad de CSB, 584
Desigualdad del triángulo, 83, 585
Determinante, 383
desarrollo por cofactores de, 386
desarrollo completo de, 407, 409
geometría de, 388
propiedades de, 392
Diagonal principal, 155
Diagonalización
de transformaciones lineales, 463
de matrices, 456
Diagráfica, 200
Diagráfica de dominancia, 203
Diferencia
ecuación, 131, 479
de matrices, 155
Dilatación, 323
Dimensión, 255
Dirac, Paul, 607
Directores
ángulos, 89
cosenos, 89
Distancia euclidiana, 79
Distancia, 79, 80
de un punto a un plano, 133
entre líneas oblicuas, 133
entre vectores, 580
Economía abierta, 209
Economía, 40
Ecuación general del plano, 125
Ecuación lineal, 2
coeficientes de, 2
conjunto solución de, 3
homogénea, 2
homogénea asociada de, 2
incógnitas de, 2
indeterminantes de, 2
parámetros de, 3
solución de, 3
solución general de, 3
solución particular de, 3
término constante de, 2
variables de, 2
variable delantera de, 2
variables libres de, 2
Ecuación no lineal, 3
Ecuación paramétrica de la recta, 121
Eigenespacio, 445
Eigenvalor, 442
dominante, 468
de transformación lineal, 450
Eigenvalores
aproximaciones de, 467
complejos, 448
Eigenvector, 442
de transformaciones lineales, 450
Elemental(es)
operaciones de renglón, 7
operaciones en ecuaciones, 7
matriz, 180
matriz de permutación, 181
Eliminación, 7
teoría, 418
Eliminación de Gauss-Jordan, 25
Enteros
modo 2, 289
modo *p*, 297
Equilibrio de mercado, 177, 500
Equilibrio, 177, 494
condiciones, 138
precio, 177
Equivalentes
matrices, 17
sistemas lineales, 6
Escalamiento, 7
de variables, 37
Escalar
matriz, 155
múltiplo de transformación, 356
multiplicación por, 63, 236
producto, 156
Espacio nulo, 108, 272
Espacio vectorial, 236
axiomas para, 236
complejo, 296
de dimensiones finitas, 255
de dimensiones infinitas, 255
general, 295
racional, 296
real, 296
Espacios vectoriales isomórficos, 340
Espectral
descomposición, 551
teorema, 548
Esquema de Sarrus, 384
Estándar
base, 64, 230, 250
matriz, 309
posición, 554
Existencia de soluciones, 23

Factorización de Crout, 192
Factorización de Dolittle, 192
Fibonacci, 1, 48
números, 213, 498
problema del dinero, 48
Forma cuadrática, 552
degenerada, 558
diagonalización de, 555
definida negativa, 560
definida positiva, 560
indefinida, 560
semidefinida negativa, 560
semidefinida positiva, 560
Forma escalón, 16
Forma escalón reducida de renglones, 16
Forma punto-normal, 124
Fórmula de Euler, 120
Fourier
aproximación, 598
polinomio, 598
serie, 598
Fractal de abeto, 369
Fractales, 368, 372
Función, 307

Galilei, Galileo, 316
Gauss, Karl Friedrich, 15
Gauss,
eliminación, 18
multiplicadores, 192
Generador, 90, 244
Geometría de fractales, 368
Gráfica, 197
dirigida, 200
dominancia, 203

línea, 202
nodos, 197
orillas, 197
simple, 422
vértices, 197
Gráficas en computadora, 313
Gram, Jorgen Pederson, 522
Grassmann, Hermann, 225

Hiperplano, 127
Homogeneidad, 81, 575
Homotecia, 323

Identidad de Jacobi, 120
Identidad de Lagrange, 116
Identidad de polarización, 89
Iguales
matrices, 155
vectores, 63
Inteligencia animal, 53
Intercambio, 7
Inversa derecha de una transformación lineal, 364
Inversa izquierda de transformación lineal, 364
Inversa, 169
cálculo de, 173
de transformación lineal, 361
Inversión, 408
Invertible
matriz, 169
operación de renglón, 181
transformación lineal, 361
Isomorfismo, 340
Iteración de Gauss-Seidel, 32
Iteración de Jacobi, 31
Iteraciones convergentes, 31
Iteraciones divergentes, 31

Jacobi, Karl Gustav, 31
Jordan, Wilhelm, 25
Juegos ejército-marina, 207, 492

Kirchhoff
leyes, 42, 421
de corriente, 42
de voltaje, 42
Kirchhoff, G. R., 421

Lados adyacentes, 197
Laplace, Pierre Simon, 386
Leontief, Wassily, 208
Leontief,
modelo abierto, 209
modelo cerrado, 209
modelos de entrada-salida, 208
Ley asociativa, 65, 156, 160, 236
Ley conmutativa, 65, 156, 236
Ley de los cosenos, 45
Ley del paralelogramo, 88, 582
de suma, 63
Ley distributiva, 65, 156, 236
Ley distributiva derecha, 160
Ley distributiva izquierda, 160
Leyes de anulación, 172
Lineal
mapeo, 320
operador, 320
segmento, 54
transformación, 320
Linealmente dependiente, 96, 247
Linealmente independiente, 97, 247
Líneas, 121
Longitud del vector, 79, 580
LU
descomposición, 188
factorización, 188

Magnitud de un vector, 79, 580
Mapeo, 307
Markov
cadena, 207, 491
proceso, 207
Markov, Andrei, 207
Matrices de Dirac, 606
Matrices de espín de Pauli, 606
Matrices diagonalmente dominantes, 33
Matrices ortogonalmente semejantes, 544
Matrices semejantes, 353
Matriz asociada, 455, 477
Matriz aumentada, 4
Matriz de adyacencia, 198, 200
Matriz de árbol, 422
Matriz de bloques, 162
Matriz de coeficientes, 4
Matriz de incidencia, 201, 203
Matriz de transición, 268
Matriz diagonal, 155
Matriz dispersa, 34
Matriz doblemente estocástica, 204
Matriz estocástica, 204
límite de la, 494
Matriz generadora, 291
Matriz identidad, 106
Matriz involutoria, 607
Matriz productiva, 209
Matriz regular, 493
Matriz triangular inferior, 155
Matriz triangular superior, 155
Matriz, matricial, 4, 154
adjunta, 402
adyacencia, 198, 200
algoritmo de inversión, 175, 184
asociada, 455, 477
cero, 155
columnas de, 154
columna cero de, 16
columna no cero de, 16
columnas pivote de, 20
consumo, 208
cuadrada, 33, 154
con respecto a bases, 346
de bloques, 162
de cofactores, 402
definida positiva, 560
de probabilidades de transición, 206
de una gráfica, 198
de una transformación lineal, 345
diagonal, 155
diagonalizable, 456
diagonalmente dominante, 33
dispersa, 34
doblemente estocástica, 204
elemental, 180
elemento de, 15, 154
elemento delantero de, 16
equivalencia, 17
entrada-salida, 208
escalar, 155
estándar, 309
estocástica, 204
forma escalón de, 16
forma escalón de renglones, 16
forma escalón reducida de, 16
identidad, 106
igualdad de, 155
incidencia, 201, 203
inversa, 169
invertible, 169
involutoria, 607
multiplicación, 158
opuesta, 156
ortogonal, 511
ortogonalmente diagonalizable, 544
permutación, 193
pivote, 20
posiciones pivote de, 20
potencias, 161, 172
productiva, 209

regular, 493
renglón cero de, 16
renglón no cero de, 16
renglones de, 154
resta, 155
simétrica, 175
submatriz de, 162
suma, 155
tamaño de, 15
transformación, 309
transpuesta de, 175
triangular inferior, 155
triangular inferior unitaria, 190
triangular superior, 155
uno delantero de, 16
Menor, 385
Mensajes digitales, 288
Método de autocorrección, 34
Método de la potencia inversa, 474
Método de las potencias, 467
Métodos directos, 30
Métodos iterativos, 31
Mínimos cuadrados, 505, 534
continuos, 593
error de, 537
polinomio, 592
solución por, 536, 538
Momento, 137
Moore-Penrose
condiciones, 570
inversa, 569
Multigráfica, 197
Multiplicidad
algebraica, 445
geométrica, 445

Newton (Sir) Isaac, 382
NFL, evaluación de los *quarterbacks*, 588
Norma de un vector, 79, 580
Normal(es)
ecuaciones, 537
vector, 124
Núcleo, 331
Nulidad, 272, 333
Número complejo
compleja conjugada de un, 545

Ondulación madre, 599
Ondulación
Haar, 599
soporte de, 599
Orilla múltiple, 197
Ortogonal
base, 508
conjunto de vectores, 506
complemento, 516
componente, 518, 520
descomposición, 524
matriz, 511
proyección, 86, 518, 520
transformación, 556
vectores, 78, 580
Ortogonalización de matrices simétricas, 543
Ortonormal
base, 511
conjunto de vectores, 509

Palabra binaria, 288
Palabras de código, 290
Papo, 143
teorema de, 144
Par de torsión, 137
Parámetros, 21
Parcial(es)
fracciones, 47
pivoteo, 20, 36
Paso directo, 19
Paso hacia atrás, 19
Pauli, Wolfgang Joseph, 606
Peano, Giuseppe, 226
Permutación, 408
impar, 408
matriz, 193, 413
método vectorial, 20
par, 408
signo de, 408
Pivote, 20
columnas, 20
posiciones, 20
Pivoteo total, 36
Pivoteo
total, 36
parcial, 20, 36
Plano inclinado, 136
Plano
ecuación general un, 125
forma punto-norma de un, 124
Planos, 124
Polinomio de interpolación, 592
Polinomio trigonométrico, 596
Polinomios de Chebyshev
primera especie, 254, 272
segunda especie, 272
Polinomios de Euler, 272
Polinomios de Hermite, 272
Polinomios de Laguerre, 272
Polinomios de Legendre, 272
Positividad, 575
Proceso de Gram-Schmidt, 522, 585
Producto cruz, 113
para vectores-*n*, 432
Producto de matrices, 158
Producto interno, 575
complejo, 605
espacio, 575
Producto matriz por vector, 105
Producto punto ponderado, 577
Producto punto, 78
ponderado, 577
Promediado de datos, 129
Promediado, 129
Promedio dorado, 499
Propiedad del valor medio para el calor, 43, 74
Proyección ortogonal, 86
Proyecciones, 314
Prueba de independencia lineal, 100, 249
Punto silla, 484

QR
factorización, 528
método, 530
Raíces latentes, 442
Rango, 280, 333
cálculo numérico de, 569
Reacciones químicas, 41
Recta, 126
ecuación paramétrica de, 121
ecuaciones simétricas de, 124
gráfica, 202
Reducción del conjunto generador, 92
Reflexiones, 310
Regla de Cramer, 405
Relación de dependencia lineal, 96
Renglón
espacio, 278
forma escalón, 16
matriz, 154
matrices equivalentes, 17
Repulsor, 484
Resta de vectores, 64
Resultante de Bezout, 431
Resultante, 420, 431
Rotaciones, 313

Schmidt, Erhardt, 522
Schwarz, Karl, 584
Secciones cónicas, 555, 556
Seidel, Philipp Ludwig, 31

Seudoinversa, 569
Simetría, 81, 575
Simétrica(s)
  ecuaciones de la recta, 124
  matriz, 175
Similitud, 371
Singulares
  descomposición en valores, 563
  valores, 564
Sistema derecho, 113
Sistema dinámico, 131, 479
  comportamiento a largo plazo de, 141, 480
  discreto, 131
  desacoplado, 484
  solución de, 131, 480
Sistema izquierdo, 113
Sistema lineal homogéneo, 4
  soluciones de, 25
Sistema lineal, 4
  asociado homogéneo de, 4
  coeficientes de, 4
  conjunto solución de, 5
  consistente, 5
  existencia de soluciones, 23
  homogéneo, 4
  inconsistente, 5
  mal acondicionado, 34
  parámetros de, 21
  solución de, 5, 23
  solución cero de, 6
  solución general de, 5
  solución no trivial de, 6
  solución particular de, 5
  solución trivial de, 6
  términos constantes de, 4
  unicidad de las soluciones de, 24
  variables libres de, 21
  variables delanteras de, 21
Sistema mal acondicionado, 34
Sistemas lineales
  cuadrados, 31
  métodos directos, 30
  métodos iterativos, 31
  soluciones numéricas de, 30
Sobreescritura, 192
Solución de sistemas lineales, 5, 23
Solución no trivial, 6
Solución trivial, 6
Soluciones de productos químicos, 41
Soluciones de sistemas homogéneos, 25
Soluciones numéricas, 30
Steinitz, 255
Suavizamiento de datos, 128
Subespacio, 226, 239
  trivial, 227, 240
  cero, 227, 240
Submatriz, 162
Suma
  de cuadrados, 47
  de matrices, 155
  de transformaciones, 356
Suma
  de matrices, 155
  de transformaciones lineales, 356
  de vectores, 63, 236
Superficies cuádricas, 559
Sustitución hacia atrás, 6
SVD, 563
Sylvester, James Joseph, 153
  resultante, 420

Teorema de Cauchy, 398
  una demostración del, 401
Teorema de Cayley-Hamilton, 497
Teorema de intercambio, 255
  dos demostraciones del, 260
Teorema de la dimensión, 336
  demostración del, 354
Teorema de la mejor aproximación, 521, 586
Teorema de los ejes principales, 555
Teorema de Varignon, 145
Teorema del rango, 281
Teoría de codificación, 287
Trabajo, 136
Transformación afín, 365
Transformación de Galileo, 316
Transformación lineal
  contradominio de, 331
  inversa de, 361
  inversa derecha de, 364
  invertible, 361
  inversa izquierda de, 364
  matriz de, 345
  núcleo de, 331
  nulidad de, 333
  potencias de, 359
  rango de, 333
Transformación, 307
  afín, 365
  biunívoca, 337
  codominio de, 307
  contradominio de, 307
  cero, 323
  dominio de, 307
  identidad, 323
  igualdad, 307
  lineal, 320
  sobre, 337
Transpuesta, 175
Traslación, 365
Trayectoria, 481
Triángulo de Sierpinski, 368

Un número complejo
Unicidad
  de la forma escalón reducida, 20
  de soluciones, 24
Unitario(a)
  círculo, 580
  esfera, 580
  vector, 80, 580

Vacas y campos de Newton, 382, 424
Vandermonde
  determinante, 429
  matriz, 428, 592
Vandermonde, A. T., 383
Variables delanteras, 21
Variables libres, 21
Varignon, 145
Vector coordenado, 263
Vector de estado estable, 494
Vector de posición, 71
Vector de probabilidad, 492
Vector libre, 71
Vector opuesto, 63, 236
Vector salida, 209
Vector(ial), 62
  cero, 64
  componente ortogonal, 86
  componentes de, 62
  de constantes, 4
  diferencia, 64
  igualdad, 63
  libre, 71
  longitud de, 79, 580
  magnitud de, 79, 580
  multiplicación por escalar, 63
  norma de, 79, 580
  normal al plano, 124
  opuesto, 63
  resta, 64
  suma, 63
  tamaño de, 62
  unitario, 80, 580
Vectores perpendiculares, 85
Vectores, 236
  producto cruz de, 113

linealmente dependientes, 96, 247
linealmente independientes, 97, 247
ortogonales, 78, 580
perpendiculares, 85
Vector-n, 62
Vértices incidentes, 197
Volumen del paralelepípedo, 118