3 Matrices

Nosotros [Halmos y Kaplansky] compartimos una filosofía acerca del álgebra lineal: pensamos libres de bases, escribimos libres de bases, pero a la hora de la verdad, cerramos la puerta de la oficina y calculamos con matrices como enajenados.

—Irving Kaplansky
en *Paul Halmos: Celebrating 50 Years of Mathematics*
J. H. Ewing y F. W. Gehring, eds.,
Springer-Verlag, 1991, p. 88

3.0 Introducción: matrices en acción

En este capítulo se estudiarán las matrices por derecho propio. Ya se usaron matrices, en forma de matrices aumentadas, para registrar información acerca de los sistemas de ecuaciones lineales y para ayudar a agilizar los cálculos que los involucran. Ahora verá que las matrices tienen propiedades algebraicas propias, lo que permite realizar cálculos con ellas, sujetas a las reglas del álgebra matricial. Más aún, observará que las matrices no son objetos estáticos que registran información y datos; en vez de ello, representan ciertos tipos de funciones que "actúan" sobre los vectores y los transforman en otros vectores. Tales "transformaciones matriciales" comenzarán a tener un papel clave en el estudio del álgebra lineal y arrojarán nueva luz en lo que ya aprendió acerca de los vectores y los sistemas de ecuaciones lineales. Además, las matrices surgen en muchas formas diferentes a las de las matrices aumentadas; al final de este capítulo se explorarán algunas de las muchas aplicaciones de las matrices.

En esta sección se considerarán algunos ejemplos simples para ilustrar cómo las matrices pueden transformar a los vectores. En el proceso tendrá un primer vistazo de la "aritmética matricial".

Considere las ecuaciones

$$\begin{aligned} y_1 &= x_1 + 2x_2 \\ y_2 &= \phantom{x_1 + {}} 3x_2 \end{aligned} \tag{1}$$

Puede considerar que estas ecuaciones describen una transformación del vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ en el vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Si se denota con F la matriz de coeficientes del lado derecho, entonces $F = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ y puede reescribir la transformación como

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

o, de manera más sucinta, $\mathbf{y} = F\mathbf{x}$. (Piense en esta expresión como el análogo de la notación funcional $y = f(x)$ a la que está acostumbrado: $\mathbf{x}$ aquí es la "variable" independiente, $\mathbf{y}$ es la "variable" dependiente y F es el nombre de la "función".)

Por tanto, si $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, entonces las ecuaciones (1) producen

$$\begin{aligned} y_1 &= -2 + 2 \cdot 1 = 0 \\ y_2 &= \, 3 \cdot 1 = 3 \end{aligned} \quad \text{o} \quad \mathbf{y} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Esta expresión puede escribirse como $\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Problema 1 Calcule $F\mathbf{x}$ para los siguientes vectores $\mathbf{x}$:

$$\text{(a)}\ \mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{(b)}\ \mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \text{(c)}\ \mathbf{x} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \qquad \text{(d)}\ \mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Problema 2 Las puntas de los cuatro vectores $\mathbf{x}$ del problema 1 se ubican en las cuatro esquinas de un cuadrado en el plano x_1x_2. Dibuje este cuadrado y marque sus esquinas A, B, C y D, que corresponden a los incisos (a), (b), (c) y (d) del problema 1.

En ejes coordenados separados (marcados y_1 y y_2), dibuje los cuatro puntos determinados por $F\mathbf{x}$ en el problema 1. Marque estos puntos A', B', C' y D'. Haga la suposición (razonable) de que el segmento de recta $\overline{AB}$, se transforma en el segmento de recta $\overline{A'B'}$ y del mismo modo para los otros tres lados del cuadrado $ABCD$. ¿Qué figura geométrica se representa mediante $A'B'C'D'$?

Problema 3 El centro del cuadrado $ABCD$ es el origen $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. ¿Cuál es el centro de $A'B'C'D'$? ¿Qué cálculo algebraico confirma esto?

Ahora considere las ecuaciones

$$\begin{aligned} z_1 &= y_1 - y_2 \\ z_2 &= -2y_1 \end{aligned} \tag{2}$$

que transforman un vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ en el vector $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$. Puede abreviar esta transformación como $\mathbf{z} = G\mathbf{y}$, de donde

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

Problema 4 Va a descubrir cómo G transforma a la figura $A'B'C'D'$. Calcule $G\mathbf{y}$ para cada uno de los cuatro vectores $\mathbf{y}$ que calculó en el problema 1. [Esto es: calcule $\mathbf{z} = G(F\mathbf{x})$. Puede reconocer esta expresión como análoga a la composición de funciones con la que está familiarizado.] Llame a los puntos correspondientes A'', B'', C'' y D'', y bosqueje la figura $A''B''C''D''$ sobre los ejes coordenados z_1z_2.

Problema 5 Al sustituir las ecuaciones (1) en las ecuaciones (2), se obtienen ecuaciones para z_1 y z_2 en términos de x_1 y x_2. Si denota con H la matriz de estas ecuaciones, entonces se tiene $\mathbf{z} = H\mathbf{x}$. Dado que también se tiene $\mathbf{z} = GF\mathbf{x}$, es razonable escribir

$$H = GF$$

¿Puede ver cómo las entradas de H se relacionan con las entradas de F y G?

Problema 6 Repita el proceso anterior en forma inversa: primero transforme el cuadrado $ABCD$, usando G, para obtener la figura $A^*B^*C^*D^*$. Luego transforme la figura resultante, con F, para obtener $A^{**}B^{**}C^{**}D^{**}$. [*Nota:* no se preocupe aquí por las "variables" $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$. Simplemente sustituya las coordenadas de A, B, C y D en las ecuaciones (2) y luego sustituya los resultados en las ecuaciones (1).] ¿$A^{**}B^{**}C^{**}D^{**}$ y

$A''B''C''D''$ son iguales? ¿Qué le dice esto acerca del orden en que se realizan las transformaciones F y G?

Problema 7 Repita el problema 5 con las matrices generales

$$F = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}, \quad G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad \text{y} \quad H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

Esto es: si las ecuaciones (1) y las ecuaciones (2) tienen coeficientes como especifican F y G, encuentre las entradas de H en términos de las entradas de F y G. El resultado será una fórmula para el "producto" $H = GF$.

Problema 8 Repita los problemas 1-6 con las siguientes matrices. (Su fórmula del problema 7 le ayudará a acelerar los cálculos algebraicos). Note cualquier similitud o diferencia que considere significativa.

(a) $F = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, G = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ (b) $F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

(c) $F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$ (d) $F = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

3.1 Operaciones con matrices

Aunque ya se ha topado con las matrices, se comienza por enunciar una definición formal y registrar algunos hechos para futura referencia.

Definición Una ***matriz*** es un arreglo rectangular de números llamados ***entradas***, o ***elementos***, de la matriz.

Aunque los números por lo general se elegirán del conjunto $\mathbb{R}$ de números reales, también pueden tomarse del conjunto $\mathbb{C}$ de números complejos o de $\mathbb{Z}_p$ donde p es primo.

Los siguientes son ejemplos de matrices:

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} \sqrt{5} & -1 & 0 \\ 2 & \pi & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 4 \\ 17 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 5.1 & 1.2 & -1 \\ 6.9 & 0 & 4.4 \\ -7.3 & 9 & 8.5 \end{bmatrix}, \quad \begin{bmatrix} 7 \end{bmatrix}$$

El ***tamaño*** de una matriz es una descripción de los números de renglones y columnas que tiene. Una matriz se llama de $m \times n$ (dígase "m por n") si tiene m renglones y n columnas. Por tanto, los ejemplos anteriores son matrices de tamaños 2×2, 2×3, 3×1, 1×4, 3×3 y 1×1, respectivamente. Una matriz de $1 \times n$ se llama ***matriz renglón*** (o ***vector renglón***), y una matriz de $n \times 1$ se llama ***matriz columna*** (o ***vector columna***).

Técnicamente, hay distinción entre matrices renglón/columna y vectores, pero aquí no se profundizará en dicha distinción. Sin embargo, *sí* se distinguirá entre matrices/vectores *renglón* y matrices/vectores *columna*. Esta distinción es importante, al menos, para cálculos algebraicos, como se demostrará.

Se usa la notación de *doble subíndice* para referirse a las entradas de una matriz A. La entrada de A en el renglón i y la columna j se denota mediante a_{ij}. Por tanto, si

$$A = \begin{bmatrix} 3 & 9 & -1 \\ 0 & 5 & 4 \end{bmatrix}$$

entonces $a_{13} = -1$ y $a_{22} = 5$. (En ocasiones se usa la notación A_{ij} de manera intercambiable con a_{ij}.) Por tanto, de manera compacta, una matriz A se puede denotar mediante $[a_{ij}]$ (o $[a_{ij}]_{m\times n}$ si es importante especificar el tamaño de A, aunque el tamaño por lo general será claro a partir del contexto).

Con esta notación, una matriz general A de $m \times n$ tiene la forma

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Si las columnas de A son los vectores $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, entonces A se puede representar como

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$$

Si los renglones de A son $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m$, entonces A se puede presentar como

$$A = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$$

Las ***entradas diagonales*** de A son $a_{11}, a_{22}, a_{33}, \ldots$, y si $m = n$ (esto es, si A tiene el mismo número de renglones que de columnas), entonces A es una ***matriz cuadrada***. Una matriz cuadrada cuyas entradas no diagonales sean todas cero es una ***matriz diagonal***. Una matriz diagonal cuyas entradas diagonales sean todas iguales es una ***matriz escalar***. Si el escalar en la diagonal es 1, la matriz escalar es una ***matriz identidad***.

Por ejemplo, sean

$$A = \begin{bmatrix} 2 & 5 & 0 \\ -1 & 4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 \\ 4 & 5 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Las entradas diagonales de A son 2 y 4, pero A no es cuadrada; B es una matriz cuadrada de tamaño 2×2, con entradas diagonales 3 y 5; C es una matriz diagonal; D es una matriz identidad de 3×3. La matriz identidad $n \times n$ se denota I_n (o simplemente I si su tamaño se entiende).

Dado que las matrices pueden verse como generalizaciones de vectores (y, de hecho, las matrices pueden y deben considerarse como constituidas tanto por vectores renglón como columna), muchas de las convenciones y operaciones para vectores se realizan sobre las matrices (en una forma obvia).

Dos matrices son ***iguales*** si tienen el mismo tamaño y si sus entradas correspondientes son iguales. Por tanto, si $A = [a_{ij}]_{m\times n}$ y $B = [b_{ij}]_{r\times s}$, entonces $A = B$ si y sólo si $m = r$ y $n = s$ y $a_{ij} = b_{ij}$ para todo i y j.

Ejemplo 3.1

Considere las matrices

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 5 & 3 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 2 & 0 & x \\ 5 & 3 & y \end{bmatrix}$$

Ni A ni B pueden ser iguales a C (sin importar cuáles sean los valores de x y y), pues A y B son matrices de 2×2 y C es de 2×3. Sin embargo, $A = B$ si y sólo si $a = 2$, $b = 0$, $c = 5$ y $d = 3$.

Ejemplo3.2

Considere las matrices

$$R = [1 \quad 4 \quad 3] \quad \text{y} \quad C = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$$

A pesar del hecho de que R y C tienen las mismas entradas en el mismo orden, $R \neq C$, pues R es 1×3 y C es 3×1. (Si R y C se leen en voz alta, suenan igual: "uno, cuatro, tres".) Por tanto, la distinción entre matrices/vectores renglón y matrices/vectores columna es importante.

Suma de matrices y multiplicación por un escalar

Para generalizar la suma de vectores, se define la suma de matrices *por componentes*. Si $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices de $m \times n$, su ***suma*** $A + B$ es la matriz de $m \times n$ que se obtiene al sumar las entradas correspondientes. Por tanto,

$$A + B = [a_{ij} + b_{ij}]$$

[Igualmente se podría definir $A + B$ en términos de suma vectorial al especificar que cada columna (o renglón) de $A + B$ es la suma de las correspondientes columnas (o renglones) de A y B.] Si A y B no son del mismo tamaño, entonces $A + B$ no está definida.

Ejemplo 3.3

Sean

$$A = \begin{bmatrix} 1 & 4 & 0 \\ -2 & 6 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} -3 & 1 & -1 \\ 3 & 0 & 2 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

Entonces

$$A + B = \begin{bmatrix} -2 & 5 & -1 \\ 1 & 6 & 7 \end{bmatrix}$$

pero ni $A + C$ ni $B + C$ están definidas.

La definición por componentes de la multiplicación por un escalar no será sorpresa. Si A es una matriz $m \times n$ y c es un escalar, entonces el ***múltiplo escalar*** cA es la matriz $m \times n$ que se obtiene al multiplicar cada entrada de A por c. De manera más formal, se tiene

$$cA = c[a_{ij}] = [ca_{ij}]$$

[En términos de vectores, podría estipularse de manera equivalente que cada columna (o renglón) de cA es c veces la correspondiente columna (o renglón) de A.]

Ejemplo 3.4

Para la matriz A del ejemplo 3.3,

$$2A = \begin{bmatrix} 2 & 8 & 0 \\ -4 & 12 & 10 \end{bmatrix}, \quad \tfrac{1}{2}A = \begin{bmatrix} \frac{1}{2} & 2 & 0 \\ -1 & 3 & \frac{5}{2} \end{bmatrix} \quad \text{y} \quad (-1)A = \begin{bmatrix} -1 & -4 & 0 \\ 2 & -6 & -5 \end{bmatrix}$$

La matriz $(-1)A$ se escribe como $-A$ y se llama ***negativo*** de A. Como con los vectores, puede usar este hecho para definir la ***diferencia*** de dos matrices: si A y B tienen el mismo tamaño, entonces

$$A - B = A + (-B)$$

Ejemplo 3.5

Para las matrices A y B del ejemplo 3.3,

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

Una matriz cuyas entradas sean todas cero se llama ***matriz cero*** y se denota O (u $O_{m\times n}$ si es importante especificar su tamaño). Debe ser claro que si A es cualquier matriz y O es la matriz cero del mismo tamaño, entonces

$$A + O = A = O + A$$

y

$$A - A = O = -A + A$$

Multiplicación de matrices

La introducción de la sección 3.0 sugirió que hay un "producto" de matrices que es análoga a la composición de funciones. Ahora se hará más precisa esta noción. La definición que está a punto de proporcionarse generaliza lo que debió descubrir en los problemas 5 y 7 de la sección 3.0. A diferencia de las definiciones de suma de matrices y multiplicación por un escalar, la definición del producto de dos matrices *no* es por componentes. Desde luego, no hay nada que impida definir un producto de matrices en forma de componentes; por desgracia, tal definición tiene pocas aplicaciones y no es tan "natural" como la que ahora se propondrá.

En ocasiones, los matemáticos son como el Humpty Dumpty de Lewis Carroll: "Cuando *yo* uso una palabra", dijo Humpty Dumpty, "significa justo lo que quiero que signifique, ni más ni menos" (de *Through the Looking Glass*).

Definición Si A es una matriz de $m \times n$ y B es una matriz de $n \times r$, entonces el ***producto*** $C = AB$ es una matriz de $m \times r$. La entrada (i, j) del producto se calcula del modo siguiente:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Comentarios

- Note que A y B no tienen que ser del mismo tamaño. Sin embargo, el número de *columnas* de A debe ser igual al número de *renglones* de B. Si escribe en orden los tamaños de A, B y AB, podrá ver de inmediato si se satisface este requisito. Más aún, puede predecir el tamaño del producto antes de hacer cálculo alguno, pues el número de *renglones* de AB es igual que el número de renglones de A, mientras que el número de *columnas* de AB es igual que el número de columnas de B, como se muestra a continuación:

$$\begin{array}{ccccc} A & & B & = & AB \\ m \times n & & n \times r & & m \times r \end{array}$$

Igual

Tamaño de AB

- La fórmula para las entradas del producto parece un producto punto, y de hecho lo es. Se dice que la entrada (i, j) de la matriz AB es el producto punto del i-ésimo renglón de A y la j-ésima columna de B:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1r} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nr} \end{bmatrix}$$

Note que en la expresión $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$, los "subíndices exteriores" de cada término ab en la suma siempre son i y j, mientras que los "subíndices interiores" siempre concuerdan y aumentan de 1 a n. Este patrón se ve claramente si escribe c_{ij} usando la notación de suma:

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Ejemplo 3.6

Calcule AB si

$$A = \begin{bmatrix} 1 & 3 & -1 \\ -2 & -1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -4 & 0 & 3 & -1 \\ 5 & -2 & -1 & 1 \\ -1 & 2 & 0 & 6 \end{bmatrix}$$

Solución Dado que A es de 2×3 y B es de 3×4, el producto AB está definido y será una matriz 2×4. El primer renglón del producto $C = AB$ se efectúa al tomar el producto punto del primer renglón de A con cada una de las columnas de B por turno. En consecuencia,

$$\begin{aligned} c_{11} &= 1(-4) + 3(5) + (-1)(-1) = 12 \\ c_{12} &= 1(0) + 3(-2) + (-1)(2) = -8 \\ c_{13} &= 1(3) + 3(-1) + (-1)(0) = 0 \\ c_{14} &= 1(-1) + 3(1) + (-1)(6) = -4 \end{aligned}$$

El segundo renglón de C se calcula al realizar el producto punto del segundo renglón de A con cada una de las columnas de B por turno:

$$\begin{aligned} c_{21} &= (-2)(-4) + (-1)(5) + (1)(-1) = 2 \\ c_{22} &= (-2)(0) + (-1)(-2) + (1)(2) = 4 \\ c_{23} &= (-2)(3) + (-1)(-1) + (1)(0) = -5 \\ c_{24} &= (-2)(-1) + (-1)(1) + (1)(6) = 7 \end{aligned}$$

Por tanto, la matriz producto está dada por

$$AB = \begin{bmatrix} 12 & -8 & 0 & -4 \\ 2 & 4 & -5 & 7 \end{bmatrix}$$

(Con un poco de práctica, deberá hacer estos cálculos mentalmente sin escribir todos los detalles como se hizo aquí. Para ejemplos más complicados, es preferible una calculadora con capacidades matriciales o un sistema algebraico de cómputo.)

Antes de avanzar más, se considerarán dos ejemplos que justifican la definición elegida de multiplicación de matrices.

Ejemplo 3.7

Ann y Bert planean ir a comprar frutas para la siguiente semana. Cada uno quiere comprar algunas manzanas, naranjas y uvas, pero en diferentes cantidades. La tabla 3.1 menciona lo que quieren comprar. Hay dos mercados de frutas cercanos, Sam's y Theo's, y sus precios se proporcionan en la tabla 3.2. ¿Cuánto costará a Ann y Bert hacer sus compras en cada uno de los dos mercados?

Tabla 3.1

	Manzanas	Uvas	Naranjas
Ann	6	3	10
Bert	4	8	5

Tabla 3.2

	Sam's	Theo's
Manzana	\$0.10	\$0.15
Uva	\$0.40	\$0.30
Naranja	\$0.10	\$0.20

Solución Si Ann compra en Sam's, gastará

$$6(0.10) + 3(0.40) + 10(0.10) = \$2.80$$

Si compra en Theo's, gastará

$$6(0.15) + 3(0.30) + 10(0.20) = \$3.80$$

Bert gastará

$$4(0.10) + 8(0.40) + 5(0.10) = \$4.10$$

en Sam's y

$$4(0.15) + 8(0.30) + 5(0.20) = \$4.00$$

en Theo's. (Presumiblemente, Ann comprará en Sam's mientras que Bert irá a Theo's.)

La "forma producto punto" de estos cálculos sugieren que la multiplicación de matrices funciona aquí. Si la información dada se organiza en una matriz demanda D y una matriz precio P, se tiene

$$D = \begin{bmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{bmatrix} \quad \text{y} \quad P = \begin{bmatrix} 0.10 & 0.15 \\ 0.40 & 0.30 \\ 0.10 & 0.20 \end{bmatrix}$$

Los cálculos anteriores son equivalentes a calcular el producto

$$DP = \begin{bmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{bmatrix} \begin{bmatrix} 0.10 & 0.15 \\ 0.40 & 0.30 \\ 0.10 & 0.20 \end{bmatrix} = \begin{bmatrix} 2.80 & 3.80 \\ 4.10 & 4.00 \end{bmatrix}$$

Por tanto, la matriz producto DP dice cuánto costarán las compras de cada persona en cada tienda (tabla 3.3).

Tabla 3.3

	Sam's	Theo's
Ann	\$2.80	\$3.80
Bert	\$4.10	\$4.00

Ejemplo 3.8 Considere el sistema lineal

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 5 \\ -x_1 + 3x_2 + x_3 &= 1 \\ 2x_1 - x_2 + 4x_3 &= 14 \end{aligned} \tag{1}$$

Observe que el lado izquierdo surge del producto matricial

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

de modo que el sistema (1) puede escribirse como

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 14 \end{bmatrix}$$

o $A\mathbf{x} = \mathbf{b}$, donde A es la matriz de coeficientes, $\mathbf{x}$ es el vector (columna) de variables y $\mathbf{b}$ es el vector (columna) de términos constantes.

El lector no debe tener dificultades para ver que *todo* sistema lineal puede escribirse en la forma $A\mathbf{x} = \mathbf{b}$. De hecho, la notación $[A \mid \mathbf{b}]$ para la matriz aumentada de un sistema lineal sólo es una abreviatura para la ecuación matricial $A\mathbf{x} = \mathbf{b}$. Esta forma probará ser tremendamente útil para expresar un sistema de ecuaciones lineales y se le explotará con frecuencia a partir de ahora.

Al combinar esta intuición con el Teorema 2.4, se ve que $A\mathbf{x} = \mathbf{b}$ tiene una solución si y sólo si $\mathbf{b}$ es una combinación lineal de las columnas de A.

Hay otro factor acerca de las operaciones matriciales que también resultará ser bastante útil: la multiplicación de una matriz por un vector unitario estándar puede usarse para "obtener" o "reproducir" una columna o un renglón de una matriz. Sea $A = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}$ y considere los productos $A\mathbf{e}_3$ y $\mathbf{e}_2 A$, con los vectores unitarios $\mathbf{e}_3$ y $\mathbf{e}_2$ elegidos de modo que los productos tengan sentido. Por tanto

$$A\mathbf{e}_3 = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{y} \quad \mathbf{e}_2 A = \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 5 & -1 \end{bmatrix}$$

Note que $A\mathbf{e}_3$ produce la tercera columna de A y $\mathbf{e}_2 A$ produce el segundo renglón de A. El resultado general se registra como teorema.

Teorema 3.1 Sea A una matriz $n \times m$, e_i, un vector unitario estándar $1 \times m$ y $\mathbf{e}_j$ un vector unitario $n \times 1$ estándar. Entonces

a. $\mathbf{e}_i$ es el i-ésimo renglón de A y
b. $A\mathbf{e}_j$ es la j-ésima columna de A.

Demostración Se demuestra (b) y la demostración de (a) se deja como ejercicio 41. Si $\mathbf{a}_1, \ldots, \mathbf{a}_n$ son las columnas de A, entonces el producto $A\mathbf{e}_j$ puede escribirse

$$A\mathbf{e}_j = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 1\mathbf{a}_j + \cdots + 0\mathbf{a}_n = \mathbf{a}_j$$

También podría demostrar (b) mediante cálculo directo:

$$A\mathbf{e}_j = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

pues el 1 en $\mathbf{e}_j$ es la j-ésima entrada.

Matrices particionadas

Con frecuencia será conveniente considerar que una matriz está compuesta de algunas ***submatrices*** más pequeñas. Al introducir rectas verticales y horizontales en una matriz, se le puede ***partir*** en ***bloques***. Hay una forma natural de partir muchas matrices, en particular aquellas que surgen en ciertas aplicaciones. Por ejemplo, considere la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix}$$

Parece natural partir A como

$$\left[\begin{array}{ccc:cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] = \begin{bmatrix} I & B \\ O & C \end{bmatrix}$$

donde I es la matriz identidad de 3×3, B es de 3×2, O es la matriz cero de 2×3 y C es de 2×2. De esta forma, puede ver a A como una matriz de 2×2 cuyas entradas son ellas mismas matrices.

Cuando se multiplican matrices, con frecuencia se gana una ventaja al verlas como matrices particionadas. Esto frecuentemente no sólo revela las estructuras subyacentes, sino que usualmente acelera los cálculos, en especial cuando las matrices son grandes y tienen muchos bloques de ceros. Es evidente que la multiplicación de matrices particionadas es igual que la multiplicación ordinaria de matrices.

Comience por considerar algunos casos especiales de matrices particionadas. Cada una da origen a una forma diferente de ver el producto de dos matrices.

Suponga que A es de $m \times n$ y B es de $n \times r$, de modo que existe el producto AB. Si particiona B en términos de sus vectores columnas, como $B = [\mathbf{b}_1 \vdots \mathbf{b}_2 \vdots \cdots \vdots \mathbf{b}_r]$, entonces

$$AB = A[\mathbf{b}_1 \vdots \mathbf{b}_2 \vdots \cdots \vdots \mathbf{b}_r] = [A\mathbf{b}_1 \vdots A\mathbf{b}_2 \vdots \cdots \vdots A\mathbf{b}_r]$$

Este resultado es una consecuencia inmediata de la definición de multiplicación de matrices. La forma a la derecha se llama ***representación matriz-columna*** del producto.

Ejemplo 3.9

Si

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

entonces

$$A\mathbf{b}_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 13 \\ 2 \end{bmatrix} \quad \text{y} \quad A\mathbf{b}_2 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Por tanto, $AB = [A\mathbf{b}_1 \vdots A\mathbf{b}_2] = \left[\begin{array}{c:c} 13 & 5 \\ 2 & -2 \end{array}\right]$. (Compruebe mediante multiplicación matricial ordinaria.)

Comentario Advierta que la representación matriz-columna de AB permite escribir cada columna de AB como una combinación lineal de las columnas de A con entradas de B como coeficientes. Por ejemplo,

$$\begin{bmatrix} 13 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(Vea los ejercicios 23 y 26.)

Suponga que A es de $m \times n$ y B es de $n \times r$, de modo que existe el producto AB. Si particiona A en términos de sus vectores renglón como

$$A = \left[\begin{array}{c} \mathbf{A}_1 \\ \hdashline \mathbf{A}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{A}_m \end{array}\right]$$

entonces

$$AB = \left[\begin{array}{c} \mathbf{A}_1 \\ \hdashline \mathbf{A}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{A}_m \end{array}\right] B = \left[\begin{array}{c} \mathbf{A}_1 B \\ \hdashline \mathbf{A}_2 B \\ \hdashline \vdots \\ \hdashline \mathbf{A}_m B \end{array}\right]$$

Una vez más, este resultado es consecuencia directa de la definición de multiplicación matricial. La forma a la derecha se conoce como ***representación renglón-matriz*** del producto.

Ejemplo 3.10

Con la representación renglón-matriz, calcule AB para las matrices del ejemplo 3.9.

Solución Calcule

$$\mathbf{A}_1 B = [1 \quad 3 \quad 2]\begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} = [13 \quad 5] \quad \text{y} \quad \mathbf{A}_2 B = [0 \quad -1 \quad 1]\begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} = [2 \quad -2]$$

Por tanto, $AB = \begin{bmatrix} \mathbf{A}_1 B \\ \hdashline \mathbf{A}_2 B \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ \hdashline 2 & -2 \end{bmatrix}$, como antes

La definición del producto matricial AB usa la partición natural de A en renglones y B en columnas; esta forma bien puede llamarse ***representación renglón-columna*** del producto. También puede partir A en columnas y B en renglones; esta forma se llama ***representación columna-renglón*** del producto.

En este caso, se tiene

$$A = [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \cdots \vdots \mathbf{a}_n] \quad \text{y} \quad B = \begin{bmatrix} \mathbf{B}_1 \\ \hdashline \mathbf{B}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{B}_n \end{bmatrix}$$

de modo que
$$AB = [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \cdots \vdots \mathbf{a}_n]\begin{bmatrix} \mathbf{B}_1 \\ \hdashline \mathbf{B}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{B}_n \end{bmatrix} = \mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + \cdots + \mathbf{a}_n\mathbf{B}_n \qquad (2)$$

Note que la suma recuerda una expansión de producto punto; la diferencia es que los términos individuales son matrices, no escalares. Asegúrese de que esto tiene sentido. Cada término $\mathbf{a}_i\mathbf{B}_i$ es el producto de una matriz de $m \times 1$ y una de $1 \times r$. Por tanto, cada $\mathbf{a}_i\mathbf{B}_i$ es una matriz de $m \times r$ del mismo tamaño que AB. Los productos $\mathbf{a}_i\mathbf{B}_i$ se llaman ***productos exteriores*** y (2) se llama ***expansión de producto exterior*** de AB.

Ejemplo 3.11

Calcule la expansión de producto exterior de AB para las matrices del ejemplo 3.9.

Solución Se tiene

$$A = [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \mathbf{a}_3] = \begin{bmatrix} 1 & \vdots & 3 & \vdots & 2 \\ 0 & \vdots & -1 & \vdots & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \mathbf{B}_1 \\ \hdashline \mathbf{B}_2 \\ \hdashline \mathbf{B}_3 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ \hdashline 1 & 2 \\ \hdashline 3 & 0 \end{bmatrix}$$

Los productos exteriores son

$$\mathbf{a}_1\mathbf{B}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}[4 \quad -1] = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{a}_2\mathbf{B}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}[1 \quad 2] = \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}$$

y
$$\mathbf{a}_3\mathbf{B}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}[3 \quad 0] = \begin{bmatrix} 6 & 0 \\ 3 & 0 \end{bmatrix}$$

(Observe que calcular cada producto exterior es exactamente igual que llenar una tabla de multiplicación.) Por tanto, la expansión de producto exterior de AB es

$$\mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + \mathbf{a}_3\mathbf{B}_3 = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix} + \begin{bmatrix} 6 & 0 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ 2 & -2 \end{bmatrix} = AB$$

La expansión de producto exterior se usará en los capítulos 5 y 7 cuando se estudie el Teorema Espectral y la descomposición de valor singular, respectivamente.

Cada una de las particiones siguientes es un caso especial del particionamiento en general. Se dice que una matriz A está particionada si se introdujeron rectas horizontales y verticales que subdividen A en submatrices llamadas bloques. El particionamiento permite que A se escriba como matriz cuyas entradas son sus bloques.

Por ejemplo,

$$A = \left[\begin{array}{ccc:cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] \quad \text{y} \quad B = \left[\begin{array}{cc:cc:c} 4 & 3 & 1 & 2 & 1 \\ -1 & 2 & 2 & 1 & 1 \\ 1 & -5 & 3 & 3 & 1 \\ \hdashline 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 3 \end{array}\right]$$

son matrices particionadas. Tienen las estructuras de bloque

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

Si dos matrices tienen el mismo tamaño y se parten la misma forma, es claro que pueden sumarse y multiplicarse por escalares bloque por bloque. Menos obvio es el hecho de que, con el particionamiento adecuado, las matrices también pueden multiplicarse en forma de bloque. El siguiente ejemplo ilustra este proceso.

Ejemplo 3.12

Considere las matrices A y B anteriores. Si por el momento ignora el hecho de que sus entradas son matrices, entonces A parece ser una matriz de 2×2 y B una matriz de 2×3. En consecuencia, su producto debe ser una matriz de 2×3 dada por

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{bmatrix} \end{aligned}$$

Pero todos los productos de este cálculo en realidad son productos *matriciales*, así que debe asegurarse de que todos estén definidos. Una comprobación rápida revela que de hecho este es el caso, pues los números de *columnas* en los bloques de A (3 y 2) coinciden con los números de *renglones* en los bloques de B. Se dice que las matrices A y B están ***particionadas de conformidad para multiplicación por bloques***.

Al realizar los cálculos indicados se obtiene el producto AB en forma particionada:

$$A_{11}B_{11} + A_{12}B_{21} = I_3B_{11} + A_{12}I_2 = B_{11} + A_{12} = \begin{bmatrix} 4 & 3 \\ -1 & 2 \\ 1 & -5 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 0 & 5 \\ 5 & -5 \end{bmatrix}$$

(Cuando algunos bloques son matrices cero o matrices identidad, como es el caso aquí, dichos cálculos pueden realizarse muy rápidamente.) Los cálculos para los otros cinco bloques de AB son similares. Compruebe que el resultado es

$$\left[\begin{array}{rr|rr|r} 6 & 2 & 1 & 2 & 2 \\ 0 & 5 & 2 & 1 & 12 \\ 5 & -5 & 3 & 3 & 9 \\ \hline 1 & 7 & 0 & 0 & 23 \\ 7 & 2 & 0 & 0 & 20 \end{array}\right]$$

(Observe que el bloque en la esquina superior izquierda es el resultado de los cálculos anteriores.) Compruebe que obtiene la misma respuesta al multiplicar A por B en la forma usual.

Potencias de matrices

Cuando A y B son dos matrices de $n \times n$, su producto AB también será una matriz de $n \times n$. Un caso especial ocurre cuando $A = B$. Tiene sentido definir $A^2 = AA$ y, en general, definir A^k como

$$A^k = \underbrace{AA \cdots A}_{k \text{ factores}}$$

si k es un entero positivo. Por ende, $A^1 = A$ y es conveniente definir $A^0 = I_n$.

Antes de hacer demasiadas suposiciones, debe preguntarse en qué medida las potencias de matrices se comportan como potencias de números reales. Las siguientes propiedades se deducen inmediatamente de las definiciones recién dadas y son los análogos matriciales de las correspondientes propiedades de las potencias de los números reales.

Si A es una matriz cuadrada y r y s son enteros no negativos, entonces

1. $A^rA^s = A^{r+s}$
2. $(A^r)^s = A^{rs}$

En la sección 3.3 se extenderá la definición y las propiedades para incluir potencias de enteros negativos.

Ejemplo 3.13

(a) Si $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, entonces

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \quad A^3 = A^2A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

y, en general,

$$A^n = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix} \quad \text{para todo } n \geq 1$$

El enunciado anterior puede probarse mediante inducción matemática, pues es una colección *infinita* de enunciados, uno para cada número natural n. (El Apéndice B ofrece un breve repaso de inducción matemática.) El paso básico es probar que la fórmula se cumple para $n = 1$. En este caso,

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

como se requiere.

La hipótesis de inducción es suponer que

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

para algún entero $k \geq 1$. El paso de inducción es probar que la fórmula se cumple para $n = k + 1$. Con la definición de potencias de matrices y la hipótesis de inducción, se calcula

$$\begin{aligned} A^{k+1} = A^kA &= \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \\ 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \end{bmatrix} \\ &= \begin{bmatrix} 2^k & 2^k \\ 2^k & 2^k \end{bmatrix} \\ &= \begin{bmatrix} 2^{(k+1)-1} & 2^{(k+1)-1} \\ 2^{(k+1)-1} & 2^{(k+1)-1} \end{bmatrix} \end{aligned}$$

Por tanto, la fórmula se cumple para toda $n \geq 1$ por el principio de inducción matemática.

(b) Si $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, entonces $B^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Al continuar, se encuentra

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

y

$$B^4 = B^3B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Por tanto, $B^5 = B$, y la sucesión de potencias de B se repite en un ciclo de cuatro:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \ldots$$

La transpuesta de una matriz

Hasta el momento, todas las operaciones matriciales definidas son análogas a las operaciones sobre números reales, aunque no siempre pueden comportarse en la misma forma. La siguiente operación no tiene tal análogo.

Definición La ***transpuesta*** de una matriz A de $m \times n$ es la matriz A^T de $n \times m$ que se obtiene al intercambiar las renglones y columnas de A. Esto es: la i-ésima columna de A^T es el i-ésimo renglón de A para toda i.

Ejemplo 3.14

Sea

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 5 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 5 & -1 & 2 \end{bmatrix}$$

Entonces sus transpuestas son

$$A^T = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 1 \end{bmatrix}, \quad B^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad \text{y} \quad C^T = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$$

En ocasiones, la transpuesta se usa para dar una definición alternativa del producto punto de dos vectores en términos de multiplicación de matrices. Si

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

entonces

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \mathbf{u}^T\mathbf{v} \end{aligned}$$

Una definición alternativa útil de la transpuesta se da en forma de componentes:

$$(A^T)_{ij} = A_{ji} \quad \text{para toda } i \text{ y } j$$

En palabras: las entradas en el renglón i y la columna j de A^T son iguales que las entradas en el renglón j y la columna i de A.

La transpuesta también se usa para definir un tipo muy importante de matriz cuadrada: una matriz simétrica.

Definición Una matriz cuadrada A es ***simétrica*** si $A^T = A$; esto es: si A es igual a su propia transpuesta.

Ejemplo 3.15

Sea

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 5 & 0 \\ 2 & 0 & 4 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$$

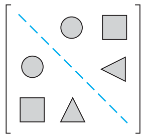


Figura 3.1
Una matriz simétrica

Entonces A es simétrica, pues $A^T = A$; pero B no es simétrica, pues $B^T = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \neq B$.

Una matriz simétrica tiene la propiedad de que es su propia "imagen especular" a través de su diagonal principal. La figura 3.1 ilustra esta propiedad para una matriz de 3×3. Las formas correspondientes representan entradas iguales; las entradas de la diagonal (las que están sobre la recta rayada) son arbitrarias.

Una definición por componentes de una matriz simétrica también es útil. Es simplemente la descripción algebraica de la propiedad de "reflexión".

Una matriz cuadrada A es simétrica si y sólo si $A_{ij} = A_{ji}$ para todo i y j.

Ejercicios 3.1

Sean

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -2 & 1 \\ 0 & 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & -3 \\ -2 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 4 & 2 \end{bmatrix}, \quad F = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

En los ejercicios 1-16, calcule las matrices indicadas (si es posible).

1. $A + 2D$
2. $2D - 5A$
3. $B - C$
4. $B - C^T$
5. AB
6. B^2
7. $D + BC$
8. B^TB
9. $E(AF)$
10. $F(AF)$
11. FE
12. EF
13. $B^TC^T - (CB)^T$
14. $DA - AD$
15. A^3
16. $(I_2 - A)^2$

17. Proporcione un ejemplo de una matriz A de 2×2 distinta de cero tal que $A^2 = 0$.

18. Sea $A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}$. Encuentre matrices de 2×2 B y C tales que $AB = AC$, pero $B \neq C$.

19. Una fábrica elabora tres productos (chismes, cachivaches y chunches) y los embarca a dos almacenes para su cuidado. El número de unidades de cada producto embarcado a cada almacén está dado por la matriz

$$A = \begin{bmatrix} 200 & 75 \\ 150 & 100 \\ 100 & 125 \end{bmatrix}$$

(donde a_{ij} es el número de unidades de producto i enviadas al almacén j y los productos se guardan en orden alfabético). El costo de embarcar una unidad de cada producto por camión es \$1.50 por chisme, \$1.00 por cachivache y \$2.00 por chunche. Los correspondientes costos unitarios para embarcar por tren son \$1.75, \$1.50 y \$1.00. Organice estos costos en una matriz B y luego use la multiplicación de matrices para mostrar cómo la fábrica puede comparar el costo de embarcar sus productos a cada uno de los dos almacenes por camión y por tren.

20. Con referencia al ejercicio 19, suponga que el costo unitario de distribuir los productos a tiendas es el mismo para cada producto, pero varía por almacén debido a las distancias involucradas. Cuesta \$0.75 distribuir una unidad desde el almacén 1 y \$1.00 distribuir una unidad desde el almacén 2. Organice dichos costos en una matriz C y luego use la multiplicación de matrices para calcular el costo total de distribuir cada producto.

En los ejercicios 21-22, escriba el sistema dado de ecuaciones lineales como una ecuación matricial de la forma $A\mathbf{x} = \mathbf{b}$.

21. $$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 0 \\ 2x_1 + x_2 - 5x_3 &= 4 \end{aligned}$$

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

En los ejercicios 23-28, sean

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

y

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

23. Use la representación matriz-columna del producto para escribir cada columna de AB como una combinación lineal de las columnas de A.

24. Use la representación renglón-matriz del producto para escribir cada renglón de AB como una combinación lineal de los renglones de B.

25. Calcule el desarrollo del producto exterior de AB.

26. Use la representación matriz-columna del producto para escribir cada columna de BA como una combinación lineal de las columnas de B.

27. Use la representación renglón-matriz del producto para escribir cada renglón de BA como una combinación lineal de los renglones de A.

28. Calcule el desarrollo del producto exterior de BA.

En los ejercicios 29 y 30, suponga que el producto AB tiene sentido.

29. Demuestre que, si las columnas de B son linealmente dependientes, entonces también lo son las columnas de AB.

30. Demuestre que, si los renglones de A son linealmente dependientes, entonces también lo son los renglones de AB.

En los ejercicios 31-34, calcule AB mediante multiplicación por bloques y use la partición indicada.

31. $A = \left[\begin{array}{cc|cc} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 3 \end{array}\right], \quad B = \left[\begin{array}{cc|c} 2 & 3 & 0 \\ -1 & 1 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 1 \end{array}\right]$

32. $A = \left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right], \quad B = \left[\begin{array}{c|cc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 5 & 4 \\ -2 & 3 & 2 \end{array}\right]$

33. $A = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & -1 \end{array}\right], \quad B = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right]$

34. $A = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ \hline 0 & 0 & 0 & 4 \end{array}\right], \quad B = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & -1 \end{array}\right]$

35. Sea $A = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

(a) Calcule $A^2, A^3, \ldots, A^7$.

(b) ¿Qué es A^{2001}? ¿Por qué?

36. Sea $B = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. Encuentre y justifique B^{2011}.

37. Sea $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Encuentre una fórmula para A^n $(n \geq 1)$ y verifique su fórmula usando inducción matemática.

38. Sea $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$.

(a) Demuestre que $A^2 = \begin{bmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{bmatrix}$.

(b) Demuestre, mediante inducción matemática, que

$$A^n = \begin{bmatrix} \cos n\theta & -\operatorname{sen} n\theta \\ \operatorname{sen} n\theta & \cos n\theta \end{bmatrix} \text{ para } n \geq 1$$

39. En cada uno de los siguientes, encuentre la matriz $A = [a_{ij}]$ de 4×4 que satisfaga la condición dada:

(a) $a_{ij} = (-1)^{i+j}$ **(b)** $a_{ij} = j - i$

(c) $a_{ij} = (i-1)^j$ **(d)** $a_{ij} = \operatorname{sen}\left(\dfrac{(i+j-1)\pi}{4}\right)$

40. En cada uno de los siguientes, encuentre la matriz $A = [a_{ij}]$ de 6×6 que satisfaga la condición dada:

(a) $a_{ij} = \begin{cases} i+j & \text{si } i \leq j \\ 0 & \text{si } i > j \end{cases}$ **(b)** $a_{ij} = \begin{cases} 1 & \text{si } |i-j| \leq 1 \\ 0 & \text{si } |i-j| > 1 \end{cases}$

(c) $a_{ij} = \begin{cases} 1 & \text{si } 6 \leq i+j \leq 8 \\ 0 & \text{de otro modo} \end{cases}$

41. Demuestre el Teorema 3.1(a).

3.2 Álgebra matricial

En ciertos aspectos, la aritmética de matrices generaliza la de los vectores. No se esperan sorpresas con respecto a la suma y la multiplicación por un escalar, y de hecho no las hay. Esto permitirá extender a las matrices varios conceptos que ya son familiares a partir del trabajo con vectores. En particular, combinaciones lineales, conjuntos generadores e independencia lineal se trasladan a las matrices sin dificultad.

Sin embargo, las matrices tienen otras operaciones, como la multiplicación matricial, que no poseen los vectores. No debe esperar que la multiplicación matricial se comporte como la multiplicación de números reales, a menos que pueda probar que lo hace; de hecho, no lo hace. En esta sección se resumen y demuestran algunas de las propiedades principales de las operaciones matriciales y se comienza por desarrollar un álgebra de matrices.

Propiedades de la suma y la multiplicación por un escalar

Todas las propiedades algebraicas de la suma y la multiplicación por un escalar para vectores (Teorema 1.1) se trasladan a las matrices. En el siguiente teorema se resumen dichas propiedades.

Teorema 3.2 **Propiedades algebraicas de la suma y la multiplicación escalar matriciales**
Sean A, B y C matrices del mismo tamaño, y sean c y d escalares. Entonces

a. $A + B = B + A$ Conmutatividad
b. $(A + B) + C = A + (B + C)$ Asociatividad
c. $A + O = A$
d. $A + (-A) = O$
e. $c(A + B) = cA + cB$ Distributividad
f. $(c + d)A = cA + dA$ Distributividad
g. $c(dA) = (cd)A$
h. $1A = A$

Las demostraciones de dichas propiedades son análogos directos de las correspondientes pruebas de las propiedades vectoriales y se dejan como ejercicios. Del mismo modo, los comentarios que siguen al Teorema 1.1 son igualmente válidos aquí y usted no debe tener dificultad para usar dichas propiedades en la realización de manipulaciones algebraicas con matrices. (Revise el ejemplo 1.5 y vea los ejercicios 17 y 18 al final de esta sección.)

La propiedad de asociatividad permite sin ambigüedades combinar la multiplicación por un escalar y la suma sin paréntesis. Si A, B y C son matrices del mismo tamaño, entonces

$$(2A + 3B) - C = 2A + (3B - C)$$

y por tanto simplemente puede escribirse $2A + 3B - C$. En general, entonces, si A_1, $A_2, \ldots, A_k$ son matrices del mismo tamaño y $c_1, c_2, \ldots, c_k$ son escalares, puede formarse la ***combinación lineal***

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k$$

A $c_1, c_2, \ldots, c_k$ se les denomina como los ***coeficientes*** de la combinación lineal. Ahora puede plantear y responder preguntas acerca de las combinaciones lineales de matrices.

Ejemplo 3.16

Sean $A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

(a) ¿$B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ es una combinación lineal de A_1, A_2 y A_3?

(b) ¿$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ es una combinación lineal de A_1, A_2 y A_3?

Solución

(a) Se quieren encontrar escalares c_1, c_2 y c_3 tales que $c_1A_1 + c_2A_2 + c_3A_3 = B$. Por tanto,

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$$

El lado izquierdo de esta ecuación puede reescribirse como

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix}$$

Al comparar entradas y usar la definición de igualdad matricial, se tienen cuatro ecuaciones lineales

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 \quad + c_3 &= 4 \\ -c_1 \quad + c_3 &= 2 \\ c_2 + c_3 &= 1 \end{aligned}$$

La eliminación de Gauss-Jordan produce fácilmente

$$\left[\begin{array}{rrr|r} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

(¡compruébelo!), de modo que $c_1 = 1$, $c_2 = -2$, y $c_3 = 3$. En consecuencia, $A_1 - 2A_2 + 3A_3 = B$, que puede comprobarse fácilmente.

(b) Esta vez se quiere resolver

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Al proceder como en el inciso (a), se obtiene el sistema lineal

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 \quad + c_3 &= 2 \\ -c_1 \quad + c_3 &= 3 \\ c_2 + c_3 &= 4 \end{aligned}$$

La reducción por renglones produce

$$\left[\begin{array}{rrr|r} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 4 \end{array}\right] \xrightarrow{R_4 - R_1} \left[\begin{array}{rrr|r} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 3 \end{array}\right]$$

No es necesario avanzar más: el último renglón implica que no hay solución. Por tanto, en este caso, C no es una combinación lineal de A_1, A_2 y A_3.

Comentario Observe que las columnas de la matriz aumentada contienen las entradas de las matrices que se proporcionan. Si se leen las entradas de cada matriz de izquierda a derecha y de arriba abajo, se obtiene el orden en el que aparecen en las columnas de la matriz aumentada. Por ejemplo, A_1 se lee como "0,1,−1,0", que corresponde a la primera columna de la matriz aumentada. Es como si simplemente se "estiraran" las matrices dadas en vectores columna. Por tanto, habría terminado exactamente con el mismo sistema de ecuaciones lineales que en el inciso (a) si se hubiera preguntado:

$$\text{Es } \begin{bmatrix} 1 \\ 4 \\ 2 \\ 1 \end{bmatrix} \text{ una combinación lineal de } \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \text{ y } \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}?$$

A partir de ahora se encontrarán con frecuencia tales paralelismos. En el capítulo 6 se les explorará con más detalle.

Se puede definir el ***espacio generado*** por un conjunto de matrices como el conjunto de todas las combinaciones lineales de las matrices.

Ejemplo 3.17

Describa el espacio generado por las matrices A_1, A_2 y A_3 en el ejemplo 3.16.

Solución Una forma de hacer esto es simplemente escribir una combinación lineal general de A_1, A_2 y A_3. Por tanto,

$$\begin{aligned} c_1A_1 + c_2A_2 + c_3A_3 &= c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} \end{aligned}$$

(que es análogo a la representación paramétrica de un plano). Pero suponga que se quiere saber cuándo la matriz $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ está en gen(A_1, A_2, A_3). A partir de la representación anterior, se sabe que lo está cuando

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

para alguna elección de escalares c_1, c_2, c_3. Esto da origen a un sistema de ecuaciones lineales cuyo lado izquierdo es exactamente igual que el del ejemplo 3.16, pero cuyo lado derecho es general. La matriz aumentada de este sistema es

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right]$$

y la reducción por renglones produce

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}x - \frac{1}{2}y \\ 0 & 1 & 0 & -\frac{1}{2}x - \frac{1}{2}y + w \\ 0 & 0 & 1 & \frac{1}{2}x + \frac{1}{2}y \\ 0 & 0 & 0 & w - z \end{array}\right]$$

(Compruebe esto cuidadosamente.) La única restricción proviene del último renglón, donde claramente se debe tener $w - z = 0$ para tener una solución. Por tanto, el gen de A_1, A_2 y A_3 consiste de todas las matrices $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ para las cuales $w = z$. Esto es es: $\text{gen}(A_1, A_2, A_3) = \left\{\begin{bmatrix} w & x \\ y & w \end{bmatrix}\right\}$.

Nota Si se hubiese sabido esto *antes* de intentar el ejemplo 3.16, habría visto inmediatamente que $B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ es una combinación lineal de A_1, A_2 y A_3, pues tiene la forma necesaria (tome $w = 1, x = 4$ y $y = 2$), pero $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ no puede ser una combinación lineal de A_1, A_2 y A_3, ya que no tiene la forma adecuada ($1 \neq 4$).

La independencia lineal también tiene sentido para matrices. Se dice que las matrices $A_1, A_2, \ldots, A_k$ del mismo tamaño son ***linealmente independientes*** si la única solución de la ecuación

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k = O \tag{1}$$

es la trivial: $c_1 = c_2 = \cdots = c_k = 0$. Si no hay coeficientes triviales que satisfagan (1), entonces $A_1, A_2, \ldots, A_k$ se llama ***linealmente dependiente***.

Ejemplo 3.18

Determine si las matrices A_1, A_2 y A_3 en el ejemplo 3.16 son linealmente independientes.

Solución Se quiere resolver la ecuación $c_1A_1 + c_2A_2 + c_3A_3 = O$. Al escribir las matrices, se tiene

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

Esta vez se tiene un sistema lineal *homogéneo* cuyo lado izquierdo es el mismo que en los ejemplos 3.16 y 3.17. (¿Comienza a vislumbrar un patrón?) La matriz aumentada se reduce por renglones para obtener

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Por tanto, $c_1 = c_2 = c_3 = 0$, y se concluye que las matrices A_1, A_2 y A_3 son linealmente independientes.

Propiedades de la multiplicación matricial

Siempre que encuentre una nueva operación, como la multiplicación matricial, debe tener cuidado de no suponer demasiado acerca de ella. Sería bueno si la multiplicación matricial se comportara como la multiplicación de números reales. Aunque en muchos aspectos lo hace, existen algunas diferencias significativas.

Ejemplo 3.19

Considere las matrices

$$A = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Al multiplicar se obtiene

$$AB = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -3 & -2 \end{bmatrix} \quad \text{y} \quad BA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

Por tanto, $AB \neq BA$. Así que, en contraste con la multiplicación de números reales, la multiplicación de matrices *no es conmutativa*: ¡el orden de los factores en un producto sí importa!

Es fácil comprobar que $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (¡hágalo!). De este modo, para matrices, la ecuación $A^2 = O$ no implica que $A = O$ (a diferencia de la situación para números reales, donde la ecuación $x^2 = 0$ sólo tiene $x = 0$ como solución).

Por muy tristes que parezcan las cosas después del último ejemplo, la situación realmente no es tan mala: sólo necesita acostumbrarse a trabajar con matrices y recordar constantemente que no son números. El siguiente teorema resume las principales propiedades de la multiplicación de matrices.

Teorema 3.3

Propiedades de la multiplicación de matrices

Sean A, B y C matrices (cuyos tamaños son tales que pueden realizarse las operaciones indicadas) y sea k un escalar. Entonces

a. $A(BC) = (AB)C$	Asociativa
b. $A(B + C) = AB + AC$	Distributiva izquierda
c. $(A + B)C = AC + BC$	Distributiva derecha
d. $k(AB) = (kA)B = A(kB)$	
e. $I_mA = A = AI_n$ si A es $m \times n$	Identidad multiplicativa

Demostración Se demuestran (b) y la mitad de (e). La demostración de la propiedad (a) se difiere hasta la sección 3.6. Las propiedades restantes se consideran en los ejercicios.

(b) Para demostrar $A(B + C) = AB + AC$, los renglones de A se denotan como $\mathbf{A}_i$ y las columnas de B y C como $\mathbf{b}_j$ y $\mathbf{c}_j$. Entonces la columna j-ésima de $B + C$ es $\mathbf{b}_j + \mathbf{c}_j$ (ya que la suma está definida por componentes) y por tanto

$$\begin{aligned}[A(B + C)]_{ij} &= \mathbf{A}_i \cdot (\mathbf{b}_j + \mathbf{c}_j) \\ &= \mathbf{A}_i \cdot \mathbf{b}_j + \mathbf{A}_i \cdot \mathbf{c}_j \\ &= (AB)_{ij} + (AC)_{ij} \\ &= (AB + AC)_{ij}\end{aligned}$$

Dado que esto es cierto para todo i y j, se debe tener $A(B + C) = AB + AC$.
(e) Para demostrar $AI_n = A$, note que la matriz identidad I_n puede partirse en columnas como

$$I_n = [\mathbf{e}_1 \vdots \mathbf{e}_2 \vdots \cdots \vdots \mathbf{e}_n]$$

donde $\mathbf{e}_i$ es un vector unitario estándar. Por tanto,

$$\begin{aligned}AI_n &= [A\mathbf{e}_1 \vdots A\mathbf{e}_2 \vdots \cdots \vdots A\mathbf{e}_n] \\ &= [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \cdots \vdots \mathbf{a}_n] \\ &= A\end{aligned}$$

por el Teorema 3.1(b).

Puede usar estas propiedades para explorar aún más cuán estrechamente la multiplicación de matrices recuerda la multiplicación de números reales.

Ejemplo 3.20

Si A y B son matrices cuadradas del mismo tamaño, ¿$(A + B)^2 = A^2 + 2AB + B^2$?

Solución Con las propiedades de la multiplicación de matrices, se calcula

$$\begin{aligned}(A + B)^2 &= (A + B)(A + B) \\ &= (A + B)A + (A + B)B && \text{por distributividad izquierda} \\ &= A^2 + BA + AB + B^2 && \text{por distributividad derecha}\end{aligned}$$

Por tanto, $(A + B)^2 = A^2 + 2AB + B^2$ si y sólo si $A^2 + BA + AB + B^2 = A^2 + 2AB + B^2$. Al restar A^2 y B^2 de ambos lados se obtiene $BA + AB = 2AB$. Al restar AB de ambos lados se obtiene $BA = AB$. Por tanto, $(A + B)^2 = A^2 + 2AB + B^2$ si y sólo si A y B se conmutan. (¿Puede dar un ejemplo de tal par de matrices? ¿Puede encontrar dos matrices que no satisfagan esta propiedad?)

Propiedades de la transpuesta

Teorema 3.4

Propiedades de la transpuesta

Sean A y B matrices (cuyos tamaños son tales que pueden realizarse las operaciones indicadas) y sea k un escalar. Entonces

a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. $(kA)^T = k(A^T)$
d. $(AB)^T = B^T A^T$
e. $(A^r)^T = (A^T)^r$ para todos los enteros r no negativos

Demostración Las propiedades (a)-(c) son intuitivamente claras y directas de demostrar (vea el ejercicio 30). Demostrar la propiedad (e) es un buen ejercicio de inducción matemática (vea el ejercicio 31). Se demostrará (d), pues no es lo que se hubiera esperado. [¿Usted sospecharía que $(AB)^T = A^TB^T$ puede ser verdadero?]

Primero, si A es de $m \times n$ y B es de $n \times r$, entonces B^T es de $r \times n$ y A^T es de $n \times m$. Por tanto, el producto B^TA^T está definido y es de $r \times m$. Dado que AB es de $m \times r$, $(AB)^T$ es de $r \times m$, y por tanto $(AB)^T$ y B^TA^T tienen el mismo tamaño. Ahora debe demostrar que sus correspondientes entradas son iguales.

Denote el i-ésimo renglón de una matriz X mediante $\text{renglón}_i(X)$ y su j-ésima columna mediante $\text{col}_j(X)$. Al usar estas convenciones, se ve que

$$\begin{aligned}[(AB)^T]_{ij} &= (AB)_{ji} \\ &= \text{renglón}_j(A) \cdot \text{col}_i(B) \\ &= \text{col}_j(A^T) \cdot \text{renglón}_i(B^T) \\ &= \text{renglón}_i(B^T) \cdot \text{col}_j(A^T) = [B^TA^T]_{ij}\end{aligned}$$

(Note que se usó la definición de multiplicación de matrices, la definición de la transpuesta y el hecho de que el producto punto es conmutativo.) Dado que i y j son arbitrarios, este resultado implica que $(AB)^T = B^TA^T$.

Comentario Las propiedades (b) y (d) del Teorema 3.4 pueden generalizarse a sumas y productos de un número finito de matrices:

$$(A_1 + A_2 + \cdots + A_k)^T = A_1^T + A_2^T + \cdots + A_k^T \quad \text{y} \quad (A_1A_2\cdots A_k)^T = A_k^T \cdots A_2^TA_1^T$$

si supone que los tamaños de las matrices son tales que todas las operaciones pueden realizarse. En los ejercicios 32 y 33 se le pide demostrar estos hechos mediante inducción matemática.

Ejemplo 3.21

Sean

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

Entonces $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, de modo que $A + A^T = \begin{bmatrix} 2 & 5 \\ 5 & 8 \end{bmatrix}$, una matriz simétrica.

Se tiene

$$B^T = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix}$$

de modo que

$$BB^T = \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 5 \\ 5 & 14 \end{bmatrix}$$

y

$$B^TB = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 2 & 2 \\ 2 & 10 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

En consecuencia, tanto BB^T como B^TB son simétricas, ¡aun cuando B ni siquiera es cuadrada! (Compruebe que AA^T y A^TA también son simétricas.)

El siguiente teorema dice que los resultados del ejemplo 3.21 son verdaderos en general.

Teorema 3.5 a. Si A es una matriz cuadrada, entonces $A + A^T$ es una matriz simétrica.
b. Para cualquier matriz A, AA^T y A^TA son matrices simétricas.

Demostración Se demuestra (a) y se deja la demostración de (b) como ejercicio 34. Simplemente demuestre que

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$$

(con las propiedades de la transpuesta y la conmutatividad de la suma de matrices). Por tanto, $A + A^T$ es igual a su propia transpuesta y en consecuencia, por definición, es simétrica.

Ejercicios 3.2

En los ejercicios 1-4, resuelva la ecuación para X, dado que $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *y* $B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$.

1. $X - 2A + 3B = O$

2. $3X = A - 2B$

3. $2(A + 2B) = 3X$

4. $2(A - B + 2X) = 3(X - B)$

En los ejercicios 5-8, escriba B como una combinación lineal de las otras matrices, si es posible.

5. $B = \begin{bmatrix} 2 & 5 \\ 0 & 3 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$

6. $B = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$
$A_3 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

7. $B = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$
$A_2 = \begin{bmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

8. $B = \begin{bmatrix} 6 & -2 & 5 \\ -2 & 8 & 6 \\ 5 & 6 & 6 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$
$A_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$
$A_4 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

En los ejercicios 9-12, encuentre la forma general del generador de las matrices indicadas, como en el ejemplo 3.17.

9. gen(A_1, A_2) en el ejercicio 5

10. gen(A_1, A_2, A_3) en el ejercicio 6

11. gen(A_1, A_2, A_3) en el ejercicio 7

12. gen(A_1, A_2, A_3, A_4) en el ejercicio 8

En los ejercicios 13-16, determine si las matrices dadas son linealmente independientes

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

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

15. $\begin{bmatrix} 0 & 1 \\ 5 & 2 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & -1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} -1 & -3 \\ 1 & 9 \\ 4 & 5 \end{bmatrix}$

16. $\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 6 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 3 \\ 0 & 4 & 9 \\ 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{bmatrix},$
$\begin{bmatrix} -1 & 1 & -1 \\ 0 & 0 & -4 \\ 0 & 0 & 0 \end{bmatrix}$

17. Demuestre el Teorema 3.2(a)-(d).

18. Demuestre el Teorema 3.2(e)-(h).

19. Demuestre el Teorema 3.3(c).

20. Demuestre el Teorema 3.3(d).

21. Demuestre la mitad del Teorema 3.3(e) que no se demostró en el texto.

22. Demuestre que, para matrices cuadradas A y B, $AB = BA$ si y sólo si $(A - B)(A + B) = A^2 - B^2$.

En los ejercicios 23-25, si $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, *encuentre las condiciones de a, b, c y d para las que* $AB = BA$.

23. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ **24.** $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ **25.** $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

26. Encuentre las condiciones de a, b, c y d para los que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ conmuta tanto con $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ como con $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

27. Encuentre las condiciones sobre a, b, c y d tales que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ conmuta con toda matriz de 2×2.

28. Demuestre que, si AB y BA están ambas definidas, entonces AB y BA son ambas matrices cuadradas.

Una matriz cuadrada se llama ***triangular superior*** *si todas las entradas abajo de la diagonal principal son cero. Por tanto, la forma de una matriz triangular superior es*

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & * & * \\ 0 & 0 & \cdots & 0 & * \end{bmatrix}$$

donde las entradas marcadas $*$ *son arbitrarias. Una definición más formal de tal matriz* $A = [a_{ij}]$ *es que* $a_{ij} = 0$ *si* $i > j$.

29. Demuestre que el producto de dos matrices triangulares superiores $n \times n$ es triangular superior.

30. Demuestre el Teorema 3.4(a)-(c).

31. Demuestre el Teorema 3.4(e).

32. Mediante inducción, demuestre que para toda $n \geq 1$, $(A_1 + A_2 + \cdots + A_n)^T = A_1^T + A_2^T + \cdots + A_n^T$.

33. Mediante inducción, demuestre que, para todo $n \geq 1$, $(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_1^T$.

34. Demuestre el Teorema 3.5(b).

35. **(a)** Demuestre que si A y B son matrices simétricas de $n \times n$, entonces también lo es $A + B$.

(b) Demuestre que si A es una matriz simétrica de $n \times n$, entonces también lo es kA para cualquier escalar k.

36. **(a)** Proporcione un ejemplo para demostrar que si A y B son matrices simétricas de $n \times n$, entonces AB no necesita ser simétrica.

(b) Demuestre que, si A y B son matrices simétricas de $n \times n$, entonces AB es simétrica si y sólo si $AB = BA$.

Una matriz cuadrada se llama ***antisimétrica*** *si* $A^T = -A$.

37. ¿Cuáles de las siguientes matrices son antisimétricas?

(a) $\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 3 & -1 \\ -3 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix}$ **(d)** $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 5 \\ 2 & 5 & 0 \end{bmatrix}$

38. Proporcione una definición en componentes de una matriz antisimétrica.

39. Demuestre que la diagonal principal de una matriz antisimétrica debe consistir completamente de ceros.

40. Demuestre que, si A y B son matrices de $n \times n$ antisimétricas, entonces también lo es $A + B$.

41. Si A y B son matrices de $n \times n$ antisimétricas, ¿en qué condiciones AB es antisimétrica?

42. Demuestre que, si A es una matriz de $n \times n$, entonces $A - A^T$ es antisimétrica.

43. **(a)** Demuestre que cualquier matriz cuadrada A puede escribirse como la suma de una matriz simétrica y una matriz antisimétrica. [*Sugerencia:* considere el Teorema 3.5 y el ejercicio 42.]

(b) Ejemplifique el inciso (a) con la matriz

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

La ***traza*** *de una matriz* $A = [a_{ij}]$ *de* $n \times n$ *es la suma de las entradas en su diagonal principal y se denota mediante* $\mathrm{tr}(A)$. *Esto es,*

$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

44. Si A y B son matrices de $n \times n$, demuestre las siguientes propiedades de la traza:

(a) $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$

(b) $\mathrm{tr}(kA) = k\mathrm{tr}(A)$, donde k es un escalar

45. Demuestre que si A y B son matrices de $n \times n$, entonces $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

46. Si A es cualquier matriz, ¿a qué es igual $\mathrm{tr}(AA^T)$?

47. Demuestre que no hay matrices de A y B de 2×2 tales que $AB - BA = I_2$.

3.3 La inversa de una matriz

En esta sección se regresa a la descripción matricial $A\mathbf{x} = \mathbf{b}$ de un sistema de ecuaciones lineales y se buscan formas de usar el álgebra matricial para resolver el sistema. Por analogía, considere la ecuación $ax = b$, donde a, b y x representan números reales y se quiere resolver para x. Rápidamente puede suponer que se quiere $x = b/a$ como la solución, pero debe recordar que es cierto sólo si $a \neq 0$. Al proceder con más lentitud y suponer que $a \neq 0$, se llegará a la solución mediante la siguiente secuencia de pasos:

$$ax = b \Rightarrow \frac{1}{a}(ax) = \frac{1}{a}(b) \Rightarrow \left(\frac{1}{a}(a)\right)x = \frac{b}{a} \Rightarrow 1 \cdot x = \frac{b}{a} \Rightarrow x = \frac{b}{a}$$

(¡Este ejemplo muestra cuánto se hace en la cabeza y cuántas propiedades de aritmética y álgebra se dan por sentadas!)

Para imitar este procedimiento para la ecuación matricial $A\mathbf{x} = \mathbf{b}$, ¿qué se necesita? Se necesita encontrar una matriz A' (análoga a $1/a$) tal que $A'A = I$, una matriz identidad (análoga a 1). *Si* tal matriz existe (análoga al requisito de que $a \neq 0$), entonces puede hacer la siguiente secuencia de cálculos:

$$A\mathbf{x} = \mathbf{b} \Rightarrow A'(A\mathbf{x}) = A'\mathbf{b} \Rightarrow (A'A)\mathbf{x} = A'\mathbf{b} \Rightarrow I\mathbf{x} = A'\mathbf{b} \Rightarrow \mathbf{x} = A'\mathbf{b}$$

(¿Por qué se justificaría cada uno de estos pasos?)

La meta en esta sección es determinar precisamente cuándo es posible encontrar tal matriz A'. De hecho, se insistirá un poco más: se quiere no sólo $A'A = I$, sino también $AA' = I$. Este requisito fuerza a A y A' a ser matrices cuadradas. (¿Por qué?)

Definición Si A es una matriz de $n \times n$, una ***inversa*** de A es una matriz A' de $n \times n$ con la siguiente propiedad

$$AA' = I \quad \text{y} \quad A'A = I$$

donde $I = I_n$ es la matriz identidad. Si tal A' existe, entonces A es ***invertible***.

Ejemplo 3.22

Si $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$, entonces $A' = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$ es una inversa de A, pues

$$AA' = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad A'A = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Ejemplo 3.23

Demuestre que las siguientes matrices no son invertibles:

(a) $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

Solución

(a) Es fácil ver que la matriz O no tiene una inversa. Si lo tuviera, entonces sería una matriz O' tal que $OO' = I = O'O$. Pero el producto de la matriz cero con cualquier otra matriz es la matriz cero y por tanto OO' nunca podría ser igual a la matriz identidad I.

(Note que esta prueba no hace alusión al tamaño de las matrices y por tanto es verdadera para matrices de $n \times n$ en general.)

(b) Suponga que B no tiene inversa $B' = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$. La ecuación $BB' = I$ produce

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

de donde se obtienen las ecuaciones

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

Al restar dos veces la primera ecuación de la tercera produce $0 = -2$, que claramente es un absurdo. Por tanto, no hay solución. (La reducción por renglones produce el mismo resultado pero en realidad no es necesaria aquí.) Se deduce que no existe tal matriz B'; esto es: B no es invertible. (De hecho, incluso no tiene una inversa que funcione en un lado, ¡mucho menos en dos!)

Comentarios

- Aun cuando se vio que la multiplicación de matrices, en general, no es conmutativa, A' (si existe) debe satisfacer $A'A = AA'$.
- Los ejemplos anteriores plantean dos preguntas que se responderán en esta sección:

(1) ¿Cómo puede saberse cuándo una matriz tiene inversa?
(2) Si una matriz tiene inversa, ¿cómo puede encontrarla?

- No se reglamentó la posibilidad de que una matriz A pueda tener más de una inversa. El siguiente teorema garantiza que esto no puede suceder.

Teorema 3.6 Si A es una matriz invertible, entonces su inversa es única.

Demostración En matemáticas, una forma estándar de demostrar que sólo hay uno de algo es demostrar que no puede haber más de uno. De este modo, suponga que A tiene dos inversas, por ejemplo, A' y A''. Entonces

$$AA' = I = A'A \qquad \text{y} \qquad AA'' = I = A''A$$

Por tanto, $$A' = A'I = A'(AA'') = (A'A)A'' = IA'' = A''$$

En consecuencia, $A' = A''$, y la inversa es única.

Gracias a este teorema, ahora puede referirse a la inversa de una matriz invertible. A partir de ahora, cuando A sea invertible, su inversa (única) se denotará mediante A^{-1} (pronúnciese "A inversa").

Advertencia ¡No escriba $A^{-1} = \dfrac{1}{A}$! No existe una operación tal como la "división entre una matriz". Incluso si la hubiera, ¿cómo se podría dividir el *escalar* 1 entre la *matriz A*? Si alguna vez se siente tentado a "dividir" entre una matriz, lo que realmente quiere hacer es multiplicar por su inversa.

Ahora puede completar la analogía que se estableció al comienzo de esta sección.

Teorema 3.7 Si A es una matriz invertible de $n \times n$, entonces el sistema de ecuaciones lineales dado por $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$ para cualquier $\mathbf{b}$ en $\mathbb{R}^n$.

Demostración El Teorema 3.7 en esencia formaliza el comentario hecho al comienzo de esta sección. Se revisará nuevamente, ahora con un poco más de cuidado. Se pidió probar dos cosas: que $A\mathbf{x} = \mathbf{b}$ *tiene* una solución y que tiene *una sola* solución. (En matemáticas, tal demostración se conoce como demostración de "existencia y unicidad".)

Para demostrar que existe una solución, sólo necesita verificar que funciona $\mathbf{x} = A^{-1}\mathbf{b}$. Se comprueba que

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$$

De modo que $A^{-1}\mathbf{b}$ satisface la ecuación $A\mathbf{x} = \mathbf{b}$ y por tanto existe *al menos* esta solución.

Para demostrar que esta solución es única, suponga que $\mathbf{y}$ es otra solución. Entonces $A\mathbf{y} = \mathbf{b}$, y al multiplicar ambos lados de la ecuación por A^{-1} a la izquierda, se obtiene la cadena de implicaciones

$$A^{-1}(A\mathbf{y}) = A^{-1}\mathbf{b} \Rightarrow (A^{-1}A)\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow I\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow \mathbf{y} = A^{-1}\mathbf{b}$$

Por tanto, $\mathbf{y}$ es la misma solución que antes, y en consecuencia la solución es única.

De este modo, al regresar a las preguntas planteadas en los comentarios previos al Teorema 3.6, ¿cómo podría decir si una matriz es invertible y cómo puede encontrar su inversa cuando es invertible? Dentro de poco se dará un procedimiento general, pero la situación para las de matrices de 2×2 es suficientemente simple para garantizar que es singular.

Teorema 3.8 Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, entonces A es invertible si $ad - bc \neq 0$, en cuyo caso

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Si $ad - bc = 0$, entonces A no es invertible.

La expresión $ad - bc$ se llama ***determinante*** de A y se denota det A. Por tanto, la fórmula para la inversa de $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (cuando existe) es $\dfrac{1}{\det A}$ multiplicado por la matriz obtenida al intercambiar las entradas en la diagonal principal y cambiar los signos en las otras dos entradas. Además de dar esta fórmula, el Teorema 3.8 dice que una matriz A de 2×2 es invertible si y sólo si $\det A \neq 0$. En el capítulo 4 se verá que el determinante puede definirse para todas las matrices cuadradas y que este resultado sigue siendo cierto, aunque no hay una fórmula simple para la inversa de matrices cuadradas más grandes.

Demostración Suponga que $\det A = ad - bc \neq 0$. Entonces

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = \det A\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

De igual modo,

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det A\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Dado que $\det A \neq 0$, puede multiplicar ambos lados de cada ecuación por $1/\det A$ para obtener

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\left(\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

y

$$\left(\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right)\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[Note que se usó la propiedad (d) del Teorema 3.3.] Por tanto, la matriz

$$\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

satisface la definición de una inversa, de modo que A es invertible. Puesto que la inversa de A es única, por el Teorema 3.6, debe tener

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Por el contrario, suponga que $ad - bc = 0$. Se considerarán por separado los casos donde $a \neq 0$ y donde $a = 0$. Si $a \neq 0$, entonces $d = bc/a$, de modo que la matriz puede escribirse como

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ ac/a & bc/a \end{bmatrix} = \begin{bmatrix} a & b \\ ka & kb \end{bmatrix}$$

donde $k = c/a$. En otras palabras, el segundo renglón de A es un múltiplo del primero. Al referirse al ejemplo 3.23(b), se ve que si A tiene una inversa $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$, entonces

$$\begin{bmatrix} a & b \\ ka & kb \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

y el sistema correspondiente de ecuaciones lineales

$$\begin{aligned} aw \quad\quad + by \quad\quad &= 1 \\ ax \quad\quad + bz &= 0 \\ kaw \quad\quad + kby \quad\quad &= 0 \\ kax \quad\quad + kbz &= 1 \end{aligned}$$

no tiene solución. (¿Por qué?)

Si $a = 0$, entonces $ad - bc = 0$ implica que $bc = 0$, y por tanto b o c es 0. En consecuencia, A es de la forma

$$\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}$$

En el primer caso, $\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ * & * \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. De igual modo, $\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}$ no puede tener una inversa. (Verifique esto.)

En consecuencia, si $ad - bc = 0$, entonces A no es invertible.

Ejemplo 3.24

Encuentre las inversas de $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 12 & -15 \\ 4 & -5 \end{bmatrix}$, si existen.

Solución Se tiene det $A = 1(4) - 2(3) = -2 \neq 0$, de modo que A es invertible, con

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

(Compruebe esto.)

Por otra parte, $B = 12(-5) - (-15)(4) = 0$, de modo que B no es invertible.

Ejemplo 3.25

Use la inversa de la matriz de coeficientes para resolver el sistema lineal

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

Solución La matriz de coeficientes es la matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, cuya inversa se calculó en el ejemplo 3.24. Por el Teorema 3.7, $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$. Aquí se tiene $\mathbf{b} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$; por tanto, la solución al sistema dado es

$$\mathbf{x} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -8 \\ \frac{11}{2} \end{bmatrix}$$

Comentario Resolver un sistema lineal $A\mathbf{x} = \mathbf{b}$ mediante $\mathbf{x} = A^{-1}\mathbf{b}$ parecería ser un buen método. Por desgracia, excepto para matrices de coeficientes de 2×2 y matrices con ciertas formas especiales, casi siempre es más rápido usar eliminación gaussiana o de Gauss-Jordan para encontrar directamente la solución. (Vea el ejercicio 13.) Más aún, la técnica del ejemplo 3.25 funciona sólo cuando la matriz de coeficientes es cuadrada e invertible, mientras que los métodos de eliminación siempre pueden aplicarse.

Propiedades de las matrices invertibles

El siguiente teorema registra algunas de las propiedades más importantes de las matrices invertibles.

Teorema 3.9

a. Si A es una matriz invertible, entonces A^{-1} es invertible y

$$(A^{-1})^{-1} = A$$

b. Si A es una matriz invertible y c es un escalar distinto de cero, entonces cA es una matriz invertible y

$$(cA)^{-1} = \frac{1}{c}A^{-1}$$

c. Si A y B son matrices invertibles del mismo tamaño, entonces AB es invertible y

$$(AB)^{-1} = B^{-1}A^{-1}$$

d. Si A es una matriz invertible, entonces A^T es invertible y

$$(A^T)^{-1} = (A^{-1})^T$$

e. Si A es una matriz invertible, entonces A^n es invertible para todo entero n no negativo y

$$(A^n)^{-1} = (A^{-1})^n$$

Demostración Se demostrarán las propiedades (a), (c) y (e), y las propiedades (b) y (d) se dejarán para su demostración en los ejercicios 14 y 15.

(a) Para demostrar que A^{-1} es invertible, debe argumentar que existe una matriz X tal que

$$A^{-1}X = I = XA^{-1}$$

Pero ciertamente A satisface estas ecuaciones en lugar de X, de modo que A^{-1} es invertible y A es *una* inversa de A^{-1}. Dado que las inversas son únicas, esto significa que $(A^{-1})^{-1} = A$.

(c) Aquí debe demostrar que existe una matriz X tal que

$$(AB)X = I = X(AB)$$

La afirmación es que la sustitución de $B^{-1}A^{-1}$ por X funciona. Se comprueba que

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

donde se usó asociatividad para recorrer los paréntesis. De igual modo, $(B^{-1}A^{-1})(AB) = I$ (¡compruébelo!), de modo que AB es invertible y su inversa es $B^{-1}A^{-1}$.

(e) La idea básica aquí es suficientemente sencilla. Por ejemplo, cuando $n = 2$, se tiene

$$A^2(A^{-1})^2 = AAA^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$

De igual modo, $(A^{-1})^2A^2 = I$. Por ende, $(A^{-1})^2$ es la inversa de A^2. No es difícil ver que un argumento similar funciona para cualquier valor entero superior de n. Sin embargo, la inducción matemática es la forma de realizar la demostración.

El paso básico es cuando $n = 0$, en cuyo caso se pedirá probar que A^0 es invertible y que

$$(A^0)^{-1} = (A^{-1})^0$$

Esto es lo mismo que demostrar que I es invertible y que $I^{-1} = I$, lo que claramente es verdadero. (¿Por qué? Vea el ejercicio 16.)

Ahora suponga que el resultado es verdadero cuando $n = k$, donde k es un entero no negativo específico. Esto es: la hipótesis de inducción es suponer que A^k es invertible y que

$$(A^k)^{-1} = (A^{-1})^k$$

El paso de inducción requiere demostrar que A^{k+1} es invertible y que $(A^{k+1})^{-1} = (A^{-1})^{k+1}$. Ahora se sabe de (c) que $A^{k+1} = A^kA$ es invertible, pues A y (por hipótesis) A^k son ambos invertibles. Más aún,

$$\begin{aligned}(A^{-1})^{k+1} &= (A^{-1})^kA^{-1} \\ &= (A^k)^{-1}A^{-1} && \text{por la hipótesis de inducción} \\ &= (AA^k)^{-1} && \text{por la propiedad (c)} \\ &= (A^{k+1})^{-1}\end{aligned}$$

Por tanto, A^n es invertible para todo entero no negativo n y $(A^n)^{-1} = (A^{-1})^n$ por el principio de inducción matemática.

Comentarios

- Aunque todas las propiedades del Teorema 3.9 son útiles, (c) es la que se debe destacar. Acaso sea la más importante propiedad algebraica de las inversas de matrices. También es la que es más fácil de que salga mal. En el ejercicio 17 se le pedirá proporcionar un contraejemplo para demostrar que, contrario a lo que se quisiera, $(AB)^{-1} \neq A^{-1}B^{-1}$ en general. La propiedad correcta, $(AB)^{-1} = B^{-1}A^{-1}$, en ocasiones se conoce como la regla calcetas-zapatos, porque, aunque uno se pone las calcetas antes que los zapatos, se los quita en el orden inverso.
- La propiedad (c) generaliza los productos de un finito de matrices invertibles: si $A_1, A_2, \ldots, A_n$ son matrices invertibles del mismo tamaño, entonces $A_1A_2 \cdots A_n$ es invertible y

$$(A_1A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$$

(Vea el ejercicio 18.) Por tanto, puede afirmarse:

> El inverso de un producto de matrices invertibles es el producto de sus inversas en el orden inverso.

- Dado que, para números reales, $\dfrac{1}{a+b} \neq \dfrac{1}{a} + \dfrac{1}{b}$, no debe esperar que para matrices cuadradas, $(A + B)^{-1} = A^{-1} + B^{-1}$ (y, de hecho, esto no es cierto en general; vea el ejercicio 19). De hecho, excepto para matrices especiales, no hay fórmula para $(A + B)^{-1}$.
- La propiedad (e) permite definir potencias enteras negativas de una matriz invertible:

> Si A es una matriz invertible y n es un entero positivo, entonces A^{-n} se define por
>
> $$A^{-n} = (A^{-1})^n = (A^n)^{-1}$$

Con esta definición, puede demostrarse que las reglas para exponenciación, $A^rA^s = A^{r+s}$ y $(A^r)^s = A^{rs}$, se sostienen para todos los enteros r y s, siempre que A sea invertible.

Un uso de las propiedades algebraicas de las matrices es ayudar a resolver ecuaciones que involucran matrices. El siguiente ejemplo ilustra el proceso. Note que debe poner particular atención al orden de las matrices en el producto.

Ejemplo 3.26

Resuelva la siguiente ecuación matricial para encontrar X (suponga que las matrices involucradas son tales que todas las operaciones indicadas están definidas):

$$A^{-1}(BX)^{-1} = (A^{-1}B^3)^2$$

Solución Existen muchas formas de proceder aquí. Una solución es

$$\begin{aligned}
A^{-1}(BX)^{-1} = (A^{-1}B^3)^2 &\Rightarrow ((BX)A)^{-1} = (A^{-1}B^3)^2 \\
&\Rightarrow [((BX)A)^{-1}]^{-1} = [(A^{-1}B^3)^2]^{-1} \\
&\Rightarrow (BX)A = [(A^{-1}B^3)(A^{-1}B^3)]^{-1} \\
&\Rightarrow (BX)A = B^{-3}(A^{-1})^{-1}B^{-3}(A^{-1})^{-1} \\
&\Rightarrow BXA = B^{-3}AB^{-3}A \\
&\Rightarrow B^{-1}BXAA^{-1} = B^{-1}B^{-3}AB^{-3}AA^{-1} \\
&\Rightarrow IXI = B^{-4}AB^{-3}I \\
&\Rightarrow X = B^{-4}AB^{-3}
\end{aligned}$$

(¿Puede justificar cada paso?) Note el uso cuidadoso del Teorema 3.9(c) y el desarrollo de $(A^{-1}B^3)^2$. También se usó libremente la asociatividad de la multiplicación de matrices para simplificar la colocación (o eliminación) de los paréntesis.

Matrices elementales

La multiplicación de matrices se usará para dar una perspectiva diferente a la reducción de matrices por renglones. En el proceso, descubrirá muchos nuevos e importantes conocimientos de la naturaleza de las matrices invertibles.

Si

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 5 & 7 \\ -1 & 0 \\ 8 & 3 \end{bmatrix}$$

se encuentra que

$$EA = \begin{bmatrix} 5 & 7 \\ 8 & 3 \\ -1 & 0 \end{bmatrix}$$

En otras palabras, multiplicar A por E (a la izquierda) tiene el mismo efecto que intercambiar los renglones 2 y 3 de A. ¿Qué es lo significativo de E? Simplemente es la matriz que se obtiene al aplicar la misma operación elemental con renglones, $R_2 \leftrightarrow R_3$, a la matriz identidad I_3. Es evidente que esto siempre funciona.

Definición Una ***matriz elemental*** es aquella matriz que puede obtenerse al realizar una operación elemental con renglones sobre una matriz identidad.

Dado que existen tres tipos de operaciones elementales con renglones, existen tres tipos correspondientes de matrices elementales. He aquí algunas de las matrices más elementales.

Ejemplo 3.27 Sean

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}$$

Cada una de esas matrices se obtuvo a partir de la matriz identidad I_4 al aplicar una sola operación elemental con renglones. La matriz E_1 corresponde a $3R_2$, E_2 a $R_1 \leftrightarrow R_3$ y E_3 a $R_4 - 2R_2$. Observe que cuando se multiplica por la izquierda una matriz de $4 \times n$ por una de tales matrices elementales, la correspondiente operación elemental con renglones se realiza sobre la matriz. Por ejemplo, si

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

entonces

$$E_1A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{21} & 3a_{22} & 3a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}, \quad E_2A = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{41} & a_{42} & a_{43} \end{bmatrix},$$

y

$$E_3A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} - 2a_{21} & a_{42} - 2a_{22} & a_{43} - 2a_{23} \end{bmatrix}$$

El ejemplo 3.27 y los ejercicios 24-30 deben convencerlo de que *cualquier* operación elemental con renglones sobre *cualquier* matriz puede lograrse al multiplicar por la izquierda por una matriz elemental adecuada. Este hecho se registra como teorema, del que se omite la demostración.

Teorema 3.10 Sea E la matriz elemental que se obtiene al realizar una operación elemental con renglones sobre I_n. Si la misma operación elemental con renglones se realiza sobre una matriz A de $n \times n$, el resultado es el mismo que la matriz EA.

Comentario Desde un punto de vista computacional, no es buena idea usar matrices elementales para realizar operaciones elementales con renglones, sólo hágalas directamente. Sin embargo, las matrices elementales pueden proporcionar algunas intuiciones valiosas acerca de las matrices invertibles y la solución de sistemas de ecuaciones lineales.

Ya se observó que toda operación elemental con renglones puede "deshacerse" o "invertirse". Esta misma observación aplicada a matrices elementales muestra que son invertibles.

Ejemplo 3.28

Sean

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

Entonces E_1 corresponde a $R_2 \leftrightarrow R_3$, que se deshace al efectuar $R_2 \leftrightarrow R_3$ nuevamente. Por tanto, $E_1^{-1} = E_1$. (Compruebe al demostrar que $E_1^2 = E_1E_1 = I$.) La matriz E_2 proviene de $4R_2$, que se deshace al aplicar $\frac{1}{4}R_2$. Por tanto,

$$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

que puede comprobarse con facilidad. Finalmente, E_3 corresponde a la operación elemental con renglones $R_3 - 2R_1$, que puede deshacerse mediante la operación elemental con renglones $R_3 + 2R_1$. De modo que, en este caso,

$$E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

(Nuevamente, es fácil comprobar esto al confirmar que el producto de esta matriz y E_3, en ambos órdenes, es I.)

Note que no sólo cada matriz elemental es invertible, sino que su inversa es otra matriz elemental del mismo tipo. Este hallazgo se registra como el siguiente teorema.

Teorema 3.11 Cada matriz elemental es invertible y su inversa es una matriz elemental del mismo tipo.

El teorema fundamental de las matrices invertibles

Ahora está en posición de probar uno de los principales resultados de este libro: un conjunto de caracterizaciones equivalentes de lo que significa para una matriz ser invertible. En un sentido, gran parte del álgebra lineal está conectada a este teorema, ya sea en el desarrollo de dichas caracterizaciones o en su aplicación. Como puede esperar, dada esta introducción, este teorema se usará bastante. ¡Hágalo su amigo!

Al Teorema 3.12 se le conocerá como la primera versión del teorema fundamental, pues se le harán agregados en capítulos posteriores. Se le recuerda que, cuando se dice que un conjunto de enunciados acerca de una matriz A son equivalentes, se entiende que, para una A dada, los enunciados son todos verdaderos o todos falsos.

Teorema 3.12 **El teorema fundamental de las matrices invertibles: versión 1**

Sea A una matriz de $n \times n$. Los siguientes enunciados son equivalentes:

a. A es invertible.
b. $A\mathbf{x} = \mathbf{b}$ tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial.
d. La forma escalonada reducida por renglones de A es I_n.
e. A es un producto de matrices elementales.

Demostración El teorema se establecerá al demostrar la cadena circular de implicaciones

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (a)$$

(a) $\Rightarrow$ (b) Ya se demostró que si A es invertible, entonces $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$ para cualquier $\mathbf{b}$ en $\mathbb{R}^n$ (Teorema 3.7).

(b) $\Rightarrow$ (c) Suponga que $A\mathbf{x} = \mathbf{b}$ tiene una solución única para cualquier $\mathbf{b}$ en $\mathbb{R}^n$. Esto implica, en particular, que $A\mathbf{x} = \mathbf{0}$ tiene una solución única. Pero un sistema homogéneo $A\mathbf{x} = \mathbf{0}$ siempre tiene $\mathbf{x} = \mathbf{0}$ como *una* solución. Así que, en este caso, $\mathbf{x} = \mathbf{0}$ debe ser *la* solución.

(c) $\Rightarrow$ (d) Suponga que $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial. El correspondiente sistema de ecuaciones es

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \end{aligned}$$

y se supone que esta solución es

$$\begin{array}{llll} x_1 & & & = 0 \\ & x_2 & & = 0 \\ & & \ddots & \\ & & & x_n = 0 \end{array}$$

En otras palabras, la eliminación de Gauss-Jordan aplicada a la matriz aumentada del sistema produce

$$[A|\mathbf{0}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right] = [I_n|\mathbf{0}]$$

Por tanto, la forma escalonada reducida por renglones de A es I_n.

(d) $\Rightarrow$ (e) Si supone que la forma escalonada reducida por renglones de A es I_n, entonces A puede reducirse a I_n usando una secuencia finita de operaciones elementales con renglones. Por el Teorema 3.10, cada una de dichas operaciones elementales con renglones puede lograrse al multiplicar la izquierda por una matriz elemental adecuada. Si la secuencia adecuada de matrices elementales es $E_1, E_2, \ldots, E_k$ (en ese orden), entonces se tiene

$$E_k \cdots E_2E_1A = I_n$$

De acuerdo con el Teorema 3.11, dichas matrices elementales son todas invertibles. Por tanto, también lo es su producto, y se tiene

$$A = (E_k \cdots E_2E_1)^{-1}I_n = (E_k \cdots E_2E_1)^{-1} = E_1^{-1}E_2^{-1} \cdots E_k^{-1}$$

De nuevo, cada E_i^{-1} es otra matriz elemental, por el Teorema 3.11, así que A se escribió como un producto de matrices elementales, como se requería.

(e) $\Rightarrow$ (a) si A es un producto de matrices elementales, entonces A es invertible, pues las matrices elementales son invertibles y los productos de las matrices invertibles son invertibles.

Ejemplo 3.29 Si es posible, exprese $A = \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix}$ como un producto de matrices elementales.

Solución Se reduce A por renglones del modo siguiente:

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 3 \\ 0 & -3 \end{bmatrix}$$
$$\xrightarrow{R_1 + R_2} \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix} \xrightarrow{-\frac{1}{3}R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Por tanto, la forma escalonada reducida por renglones de A es la matriz identidad, así que el teorema fundamental asegura que A es invertible y puede escribirse como un producto de matrices elementales. Se tiene $E_4E_3E_2E_1A = I$, donde

$$E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad E_4 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{3} \end{bmatrix}$$

son las matrices elementales correspondientes a las cuatro operaciones elementales con renglones utilizadas para reducir A a I. Como en la prueba del teorema, se tiene

$$A = (E_4E_3E_2E_1)^{-1} = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$$

como se requiere.

Comentario Puesto que la secuencia de operaciones elementales con renglones que transforma A en I no es única, tampoco lo es la representación de A como producto de matrices elementales. (Encuentre una forma diferente de expresar A como un producto de matrices elementales).

"Nunca lleve un cañón al escenario en el primer acto, a menos que pretenda dispararlo en el último acto." –Anton Chejov

El teorema fundamental es sorprendentemente poderoso. Para ilustrar su poder, considere dos de sus consecuencias. La primera es que, aunque la definición de una matriz invertible afirma que una matriz A es invertible si existe una matriz B tal que *tanto* $AB = I$ *como* $BA = I$ se satisfacen, sólo es necesario comprobar *una* de dichas ecuaciones. Por tanto, ¡puede reducirse el trabajo a la mitad!

Teorema 3.13 Sea A una matriz cuadrada. Si B es una matriz cuadrada tal que $AB = I$ o $BA = I$, entonces A es invertible y $B = A^{-1}$.

Demostración Suponga $BA = I$. Considere la ecuación $A\mathbf{x} = \mathbf{0}$. Al multiplicar por la izquierda por B, se tiene $BA\mathbf{x} = B\mathbf{0}$. Esto implica que $\mathbf{x} = I\mathbf{x} = \mathbf{0}$. Por tanto, el sistema representado por $A\mathbf{x} = \mathbf{0}$ tiene la solución única $\mathbf{x} = \mathbf{0}$. A partir de la equivalencia de (c) y (a) en el teorema fundamental, se sabe que A es invertible. (Esto es: existe A^{-1} y satisface $AA^{-1} = I = A^{-1}A$.)

Si ahora multiplica por la derecha ambos lados de $BA = I$ por A^{-1}, obtiene

$$BAA^{-1} = IA^{-1} \Rightarrow BI = A^{-1} \Rightarrow B = A^{-1}$$

(La demostración en el caso de $AB = I$ se deja como ejercicio 41.)

La siguiente consecuencia del teorema fundamental es la base de un método eficiente para calcular la inversa de una matriz.

Teorema 3.14 Sea A una matriz cuadrada. Si una secuencia de operaciones elementales con renglones reduce A a I, entonces la misma secuencia de operaciones elementales transforma I en A^{-1}.

Demostración Si A es equivalente por renglones a I, entonces puede lograrse la reducción mediante multiplicación por la izquierda por una secuencia $E_1, E_2, \ldots, E_k$ de matrices elementales. Por tanto, se tiene $E_k \cdots E_2E_1A = I$. Al hacer $B = E_k \cdots E_2E_1$ se obtiene $BA = I$. Por el Teorema 3.13, A es invertible y $A^{-1} = B$. Ahora, aplicar la misma secuencia de operaciones elementales con renglones a I equivale a multiplicar por la izquierda I por $E_k \cdots E_2E_1 = B$. El resultado es

$$E_k \cdots E_2E_1I = BI = B = A^{-1}$$

Por tanto, I se transforma en A^{-1} mediante la misma secuencia de operaciones elementales con renglones.

El método de Gauss-Jordan para calcular la inversa

Es posible realizar operaciones con renglones sobre A e I simultáneamente al construir una "matriz superaumentada" $[A \mid I]$. El Teorema 3.14 muestra que si A es equivalente por renglones a I [lo cual, por el teorema fundamental (d) $\Leftrightarrow$ (a), significa que A es invertible], entonces operaciones elementales con renglones producirán

$$[A \mid I] \longrightarrow [I \mid A^{-1}]$$

Si A no puede reducirse a I, entonces el teorema fundamental garantiza que A no es invertible.

El procedimiento recién descrito es simplemente la eliminación de Gauss-Jordan efectuada sobre una matriz aumentada de $n \times 2n$, en lugar de sobre una de $n \times (n + 1)$. Otra forma de ver este procedimiento es observar el problema de encontrar A^{-1} al resolver la ecuación matricial $AX = I_n$ para una matriz X de $n \times n$. (Esto es suficiente, por el teorema fundamental, pues una inversa correcta de A debe ser una inversa de dos lados.) Si las columnas de X se denotan $\mathbf{x}_1, \ldots, \mathbf{x}_n$, entonces esta ecuación matricial es equivalente a resolver para las columnas de X, una a la vez. Dado que las columnas de I_n son los vectores unitarios estándar $\mathbf{e}_1, \ldots, \mathbf{e}_n$, entonces se tienen n sistemas de ecuaciones lineales, todas con matriz de coeficientes A:

$$A\mathbf{x}_1 = \mathbf{e}_1, \ldots, A\mathbf{x}_n = \mathbf{e}_n$$

Dado que se necesita la misma secuencia de operaciones con renglones para llevar A a la forma escalonada reducida por renglones en cada caso, las matrices aumentadas para estos sistemas, $[A \mid \mathbf{e}_1], \ldots, [A \mid \mathbf{e}_n]$, pueden combinarse como

$$[A \mid \mathbf{e}_1\, \mathbf{e}_2 \cdots \mathbf{e}_n] = [A \mid I_n]$$

Ahora se aplican operaciones con renglones para tratar de reducir A a I_n, lo cual, si tiene éxito, resolverá simultáneamente para las columnas de A^{-1}, lo que transformará I_n en A^{-1}.

Este uso de la eliminación de Gauss-Jordan se ilustra con tres ejemplos.

Ejemplo 3.30 Encuentre la inversa de

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

si existe.

Solución La eliminación de Gauss-Jordan produce

$$[A\,|\,I] = \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3-R_1]{R_2-2R_1} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 6 & -2 & 1 & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{(-\frac{1}{2})R_2} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-R_2} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow[R_2+3R_3]{R_1+R_3} \left[\begin{array}{rrr|rrr} 1 & 2 & 0 & -1 & \frac{1}{2} & 1 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow{R_1-2R_2} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 9 & -\frac{3}{2} & -5 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

Por tanto,

$$A^{-1} = \begin{bmatrix} 9 & -\frac{3}{2} & -5 \\ -5 & 1 & 3 \\ -2 & \frac{1}{2} & 1 \end{bmatrix}$$

(Siempre debe comprobar que $AA^{-1} = I$ mediante multiplicación directa. Por el Teorema 3.13, no necesita comprobar que también $A^{-1}A = I$.)

Comentario Note que se usó la variante de eliminación de Gauss-Jordan que primero introduce todos los ceros *abajo* de los 1 pivote, de izquierda a derecha y de arriba abajo, y luego crea ceros *arriba* de los 1 pivote, de derecha a izquierda y de abajo arriba. Este planteamiento ahorra cálculos, como se anotó en el capítulo 2, pero puede encontrarlo más sencillo, cuando trabaja a mano, para crear *todos* los ceros en cada columna conforme avance. Desde luego, la respuesta será la misma.

Ejemplo 3.31

Encuentre la inversa de

$$A = \begin{bmatrix} 2 & 1 & -4 \\ -4 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix}$$

si existe.

Solución Proceda como en el ejemplo 3.30, adjunte la matriz identidad a A y luego trate de convertir $[A \mid I]$ en $[I \mid A^{-1}]$.

$$[A \mid I] = \left[\begin{array}{rrr|rrr} 2 & 1 & -4 & 1 & 0 & 0 \\ -4 & -1 & 6 & 0 & 1 & 0 \\ -2 & 2 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3+R_1]{R_2+2R_1} \left[\begin{array}{rrr|rrr} 2 & 1 & -4 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & 3 & -6 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-3R_2} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 2 & 1 & 0 \\ 0 & 0 & 0 & -5 & -3 & 1 \end{array}\right]$$

En este punto se ve que no es posible reducir A a I, pues existe un renglón de ceros en el lado izquierdo de la matriz aumentada. En consecuencia, A no es invertible.

Como ilustra el ejemplo siguiente, todo funciona de la misma forma sobre $\mathbb{Z}_p$, donde p es primo.

Ejemplo 3.32

Encuentre la inversa de

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

si existe, sobre $\mathbb{Z}_3$.

Solución 1 Use el método de Gauss-Jordan y recuerde que todos los cálculos son en $\mathbb{Z}_3$.

$$[A \mid I] = \left[\begin{array}{rr|rr} 2 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{2R_1} \left[\begin{array}{rr|rr} 1 & 1 & 2 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_2+R_1} \left[\begin{array}{rr|rr} 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{R_1+2R_2} \left[\begin{array}{rr|rr} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right]$$

Por tanto, $A^{-1} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$ y es fácil comprobar que, sobre $\mathbb{Z}_3$, $AA^{-1} = I$.

Solución 2 Dado que A es una matriz de 2×2, también puede calcular A^{-1} usando la fórmula dada en el Teorema 3.8. El determinante de A es

$$\det A = 2(0) - 2(2) = -1 = 2$$

en $\mathbb{Z}_3$ (ya que $2 + 1 = 0$). Por tanto, A^{-1} existe y está dada por la fórmula del Teorema 3.8. Sin embargo, aquí debe tener cuidado, porque la fórmula introduce la “fracción” $1/\det A$ y en $\mathbb{Z}_3$ no existen fracciones. Debe usar inversos multiplicativos en lugar de división.

En lugar de $1/\det A = 1/2$, use 2^{-1}; esto es, encuentre el número x que satisfaga la ecuación $2x = 1$ en $\mathbb{Z}_3$. Es fácil ver que $x = 2$ es la solución que se quiere: en $\mathbb{Z}_3$, $2^{-1} = 2$, pues $2(2) = 1$. La fórmula para A^{-1} ahora se convierte en

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

que concuerda con la solución anterior.

Ejercicios 3.3

En los ejercicios 1-10, encuentre la inversa de la matriz dada (si existe) usando el Teorema 3.8.

1. $\begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix}$ **2.** $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

3. $\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$ **4.** $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

5. $\begin{bmatrix} \frac{3}{4} & \frac{3}{5} \\ \frac{5}{6} & \frac{2}{3} \end{bmatrix}$ **6.** $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1\sqrt{2} \end{bmatrix}$

7. $\begin{bmatrix} -1.5 & -4.2 \\ 0.5 & 2.4 \end{bmatrix}$ **8.** $\begin{bmatrix} 3.55 & 0.25 \\ 8.52 & 0.60 \end{bmatrix}$

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

10. $\begin{bmatrix} 1/a & 1/b \\ 1/c & 1/d \end{bmatrix}$, donde ni a, b, c ni d son 0.

En los ejercicios 11 y 12, resuelva el sistema dado usando el método del ejemplo 3.25.

11. $\begin{aligned} 2x + y &= -1 \\ 5x + 3y &= 2 \end{aligned}$ **12.** $\begin{aligned} x_1 - x_2 &= 2 \\ x_1 + 2x_2 &= 5 \end{aligned}$

13. Sea $A = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, y $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

(a) Encuentre A^{-1} y úselo para resolver los tres sistemas $A\mathbf{x} = \mathbf{b}_1$, $A\mathbf{x} = \mathbf{b}_2$ y $A\mathbf{x} = \mathbf{b}_3$.

(b) Resuelva los tres sistemas al mismo tiempo mediante la reducción por renglones de la matriz aumentada $[A \mid \mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3]$, usando la eliminación de Gauss-Jordan.

(c) Cuente cuidadosamente el número total de multiplicaciones individuales que realizó en (a) y en (b). Debe descubrir que, incluso para este ejemplo de 2×2, un método usa menos operaciones. Para sistemas más grandes, la diferencia es incluso más pronunciada, y esto explica por qué los sistemas de cómputo no usan uno de dichos métodos para resolver sistemas lineales.

14. Demuestre el Teorema 3.9(b).

15. Demuestre el Teorema 3.9(d).

16. Demuestre que la matriz identidad I_n de $n \times n$ es invertible y que $I_n^{-1} = I_n$.

17. **(a)** Proporcione un contraejemplo para demostrar que $(AB)^{-1} \neq A^{-1}B^{-1}$ en general.

(b) ¿En qué condiciones de A y B, $(AB)^{-1} = A^{-1}B^{-1}$? Pruebe su afirmación.

18. Mediante inducción, demuestre que si $A_1, A_2, \ldots, A_n$ son matrices invertibles del mismo tamaño, entonces el producto $A_1A_2 \cdots A_n$ es invertible y $(A_1A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$.

19. Proporcione un contraejemplo para demostrar que en general $(A + B)^{-1} \neq A^{-1} + B^{-1}$.

En los ejercicios 20-23, resuelva la ecuación matricial dada para X. Simplifique sus respuestas tanto como sea posible. (En palabras de Albert Einstein, "todo debe hacerse tan simple como sea posible, pero no más sencillo".) Suponga que todas las matrices son invertibles.

20. $XA^{-1} = A^3$ **21.** $AXB = (BA)^2$

22. $(A^{-1}X)^{-1} = (AB^{-1})^{-1}(AB^2)$ **23.** $ABXA^{-1}B^{-1} = I + A$

En los ejercicios 24-30, sean

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

$$C = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -1 & 3 \\ 2 & 1 & -1 \end{bmatrix}$$

En cada caso, encuentre una matriz elemental E que satisfaga la ecuación dada.

24. $EA = B$ **25.** $EB = A$ **26.** $EA = C$

27. $EC = A$ **28.** $EC = D$ **29.** $ED = C$

30. ¿Existe una matriz elemental E tal que $EA = D$? ¿Por qué sí o por qué no?

En los ejercicios 31-38, encuentre la inversa de la matriz elemental dada.

31. $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ **32.** $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **34.** $\begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}$

35. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ **36.** $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

37. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}, c \neq 0$ **38.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, c \neq 0$

En los ejercicios 39 y 40, encuentre una secuencia de matrices elementales $E_1, E_2, \ldots, E_k$ tal que $E_k \cdots E_2E_1A = I$. Use esta secuencia para escribir A y A^{-1} como productos de matrices elementales.

39. $A = \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix}$ **40.** $A = \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}$

41. Demuestre el Teorema 3.13 para el caso de $AB = I$.

42. **(a)** Demuestre que si A es invertible y $AB = O$, entonces $B = O$.

(b) Ofrezca un contraejemplo para demostrar que el resultado en el inciso (a) puede fallar si A no es invertible.

43. **(a)** Demuestre que si A es invertible y $BA = CA$, entonces $B = C$.

(b) Ofrezca un contraejemplo para demostrar que el resultado en el inciso (a) puede fallar si A no es invertible.

44. Una matriz cuadrada A se llama ***idempotente*** si $A^2 = A$. (La palabra *idempotente* viene de la palabra latina *idem*, que significa "igual", y *potere*, que significa "tener poder". Por tanto, algo que es idempotente tiene "la misma potencia" cuando se eleva al cuadrado.)

(a) Encuentre tres matrices de 2×2 idempotentes.

(b) Demuestre que la única matriz de $n \times n$ idempotente invertible es la matriz identidad.

45. Demuestre que si A es una matriz cuadrada que satisface la ecuación $A^2 - 2A + I = O$, entonces $A^{-1} = 2I - A$.

46. Demuestre que si una matriz simétrica es invertible, entonces su inversa también es simétrica.

47. Demuestre que si A y B son matrices cuadradas, y AB es invertible, entonces A y B son invertibles.

En los ejercicios 48-63, use el método de Gauss-Jordan para encontrar la inversa de la matriz dada (si existe).

48. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ **49.** $\begin{bmatrix} -2 & 4 \\ 3 & -1 \end{bmatrix}$

50. $\begin{bmatrix} 3 & -4 \\ -6 & 8 \end{bmatrix}$ **51.** $\begin{bmatrix} 1 & a \\ -a & 1 \end{bmatrix}$

52. $\begin{bmatrix} 2 & 0 & -1 \\ 1 & 5 & 1 \\ 2 & 3 & 0 \end{bmatrix}$ **53.** $\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 2 \\ 2 & 3 & -1 \end{bmatrix}$

54. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ **55.** $\begin{bmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{bmatrix}$

56. $\begin{bmatrix} 0 & a & 0 \\ b & 0 & c \\ 0 & d & 0 \end{bmatrix}$ **57.** $\begin{bmatrix} 0 & -1 & 1 & 0 \\ 2 & 1 & 0 & 2 \\ 1 & -1 & 3 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$

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

59. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & b & c & d \end{bmatrix}$ **60.** $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_2$

61. $\begin{bmatrix} 4 & 2 \\ 3 & 4 \end{bmatrix}$ sobre $\mathbb{Z}_5$ **62.** $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_3$

63. $\begin{bmatrix} 1 & 5 & 0 \\ 1 & 2 & 4 \\ 3 & 6 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_7$

Particionar matrices cuadradas grandes en ocasiones puede hacer que sus inversas sean más fáciles de calcular, en particular si los bloques tienen una forma manejable. En los ejercicios 64-68, verifique mediante multiplicación por bloque que la inversa de una matriz, si se particiona como se muestra, es como se afirma. (Suponga que todas las inversas existen como se necesita.)

64. $\begin{bmatrix} A & B \\ O & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ O & D^{-1} \end{bmatrix}$

65. $\begin{bmatrix} O & B \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} -(BC)^{-1} & (BC)^{-1}B \\ C(BC)^{-1} & I - C(BC)^{-1}B \end{bmatrix}$

66. $\begin{bmatrix} I & B \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} (I - BC)^{-1} & -(I - BC)^{-1}B \\ -C(I - BC)^{-1} & I + C(I - BC)^{-1}B \end{bmatrix}$

67. $\begin{bmatrix} O & B \\ C & D \end{bmatrix}^{-1}$

$= \begin{bmatrix} -(BD^{-1}C)^{-1} & (BD^{-1}C)^{-1}BD^{-1} \\ D^{-1}C(BD^{-1}C)^{-1} & D^{-1} - D^{-1}C(BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$

68. $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$, donde $P = (A - BD^{-1}C)^{-1}$, $Q = -PBD^{-1}$, $R = -D^{-1}CP$, y $S = D^{-1} + D^{-1}CPBD^{-1}$

En los ejercicios 69-72, particione la matriz dada de modo que pueda aplicar una de las fórmulas de los ejercicios 64-68, y luego calcule la inversa usando dicha fórmula.

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

70. La matriz del ejercicio 58

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

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

3.4 La factorización *LU*

Así como es natural (e ilustrador) factorizar un número natural en un producto de otros números naturales (por ejemplo, $30 = 2 \cdot 3 \cdot 5$), frecuentemente también es útil factorizar matrices como productos de otras matrices. Cualquier representación de una matriz como producto de dos o más matrices se llama ***factorización de matrices***. Por ejemplo,

$$\begin{bmatrix} 3 & -1 \\ 9 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}$$

es una factorización de una matriz.

No es necesario decir que algunas factorizaciones son más útiles que otras. En esta sección se presenta una factorización matricial que surge en la solución de sistemas de ecuaciones lineales mediante eliminación gaussiana y es particularmente adecuada para su implementación en computadora. En capítulos posteriores se encontrarán otras factorizaciones matriciales igualmente útiles. De hecho, el tema es rico, y a él se han dedicado libros y cursos enteros.

Considere un sistema de ecuaciones lineales de la forma $A\mathbf{x} = \mathbf{b}$, donde A es una matriz de $n \times n$. La meta de este libro es demostrar que la eliminación gaussiana implícitamente factoriza A en un producto de matrices que permiten resolver fácilmente el sistema dado (y cualquier otro sistema con la misma matriz de coeficientes).

El siguiente ejemplo ilustra la idea básica.

Ejemplo 3.33

Sea

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

La reducción por renglones de A se realiza del modo siguiente:

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \xrightarrow[R_3+R_1]{R_2-2R_1} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} \xrightarrow{R_3+2R_2} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U \qquad (1)$$

Las tres matrices elementales E_1, E_2, E_3 que logran esta reducción de A a la forma escalonada U son (en orden):

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

Por tanto,

$$E_3E_2E_1A = U$$

Al despejar para A, se obtiene

$$\begin{aligned} A = E_1^{-1}E_2^{-1}E_3^{-1}U &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}U \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}U = LU \end{aligned}$$

Por tanto, A puede factorizarse como

$$A = LU$$

donde U es una matriz *triangular superior* (vea los ejercicios para la sección 3.2) y L es *triangular inferior unitaria*. Esto es, L tiene la forma

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix}$$

con ceros arriba y números 1 en la diagonal principal.

El gran matemático inglés Alan M. Turing (1912–1954) introdujo la factorización *LU* en 1948, en un ensayo titulado "Rounding-off Errors en Matrix Processes" (*Quarterly Journal of Mechanics and Applied Mathematics, 1* (1948), pp. 287-308). Durante la segunda guerra mundial, Turing fue útil para descifrar el código "Enigma" alemán. Sin embargo, es mejor conocido por su trabajo en lógica matemática que tendió los cimientos teóricos para el desarrollo de las computadoras digitales y el moderno campo de la inteligencia artificial. La "prueba de Turing" que propuso en 1950 todavía se usa como uno de los hitos para abordar la cuestión de si una computadora puede considerarse "inteligente".

El ejemplo anterior motiva la siguiente definición.

Definición Sea A una matriz cuadrada. Una factorización de A como $A = LU$, donde L es triangular inferior unitaria y U es triangular superior, se llama ***factorización LU*** de A.

Comentarios

- Observe que la matriz A en el ejemplo 3.33 tiene una factorización *LU* porque en la reducción por renglones de A *no se necesitaron intercambios de renglón*. Por tanto, todas las matrices elementales que surgieron fueron triangulares inferiores unitarias. Entonces, se garantiza que L es triangular inferior unitaria porque los inversos y productos de matrices triangulares inferiores unitarias también son triangulares inferiores unitarias. (Vea los ejercicios 29 y 30.)

Si apareciera un cero en una posición pivote en algún paso, habría tenido que intercambiar renglones para obtener un pivote distinto de cero. Esto habría resultado en que L ya no fuese triangular inferior unitaria. Más adelante se comentará más acerca de esta observación. (¿Puede encontrar una matriz para la cual será necesario el intercambio de renglones?)

- La noción de una factorización LU puede generalizarse a matrices no cuadradas al simplemente requerir que U sea una matriz en forma escalonada por renglones. (Vea los ejercicios 13 y 14.)
- Algunos libros definen una factorización LU de una matriz cuadrada A como cualquier factorización $A = LU$, donde L es triangular inferior y U es triangular superior.

El primer comentario anterior en esencia es una demostración del siguiente teorema.

Teorema 3.15 Si A es una matriz cuadrada que puede reducirse a forma escalonada sin usar intercambios de renglón, entonces A tiene una factorización LU.

Para ver por qué es útil la factorización LU, considere un sistema lineal $A\mathbf{x} = \mathbf{b}$, donde la matriz de coeficientes tiene una factorización LU, $A = LU$. El sistema $A\mathbf{x} = \mathbf{b}$ puede reescribirse como $LU\mathbf{x} = \mathbf{b}$ o $L(U\mathbf{x}) = \mathbf{b}$. Si ahora se define $\mathbf{y} = U\mathbf{x}$, entonces es posible resolver para $\mathbf{x}$ en dos etapas:

1. Resolver $L\mathbf{y} = \mathbf{b}$ para $\mathbf{y}$ mediante *sustitución hacia adelante* (vea los ejercicios 25 y 26 en la sección 2.1).
2. Resolver $U\mathbf{x} = \mathbf{y}$ para $\mathbf{x}$ mediante *sustitución hacia atrás*.

Cada uno de dichos sistemas lineales tiene solución directa porque las matrices coeficiente L y U son ambas triangulares. El siguiente ejemplo ilustra el método.

Ejemplo 3.34

Use una factorización LU de $A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$ para resolver $A\mathbf{x} = \mathbf{b}$, donde $\mathbf{b} = \begin{bmatrix} 1 \\ -4 \\ 9 \end{bmatrix}$.

Solución En el ejemplo 3.33, se encontró que

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

Como se destaca líneas arriba, para resolver $A\mathbf{x} = \mathbf{b}$ (que es lo mismo que $L(U\mathbf{x}) = \mathbf{b}$), primero se resuelve $L\mathbf{y} = \mathbf{b}$ para $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. Este es justo el sistema lineal

$$\begin{aligned} y_1 \qquad\qquad &= 1 \\ 2y_1 + y_2 \qquad &= -4 \\ -y_1 - 2y_2 + y_3 &= 9 \end{aligned}$$

La sustitución hacia adelante (esto es, al trabajar de arriba abajo) produce

$$y_1 = 1, y_2 = -4 - 2y_1 = -6, y_3 = 9 + y_1 + 2y_2 = -2$$

Por tanto, $\mathbf{y} = \begin{bmatrix} 1 \\ -6 \\ -2 \end{bmatrix}$ y ahora se resuelve $U\mathbf{x} = \mathbf{y}$ para $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Este sistema lineal es

$$\begin{aligned} 2x_1 + \quad x_2 + 3x_3 &= 1 \\ -3x_2 - 3x_3 &= -6 \\ 2x_3 &= -2 \end{aligned}$$

y la sustitución hacia atrás produce rápidamente

$$\begin{aligned} x_3 &= -1, \\ -3x_2 &= -6 + 3x_3 = -9 \quad \text{de manera que } x_2 = 3, \text{ y} \\ 2x_1 &= 1 - x_2 - 3x_3 = 1 \quad \text{de manera que } x_1 = \tfrac{1}{2} \end{aligned}$$

Por tanto, la solución al sistema dado $A\mathbf{x} = \mathbf{b}$ es $\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ 3 \\ -1 \end{bmatrix}$.

Una forma sencilla de encontrar factorizaciones *LU*

En el ejemplo 3.33 se calculó la matriz L como un producto de matrices elementales. Por fortuna, L puede calcularse directamente a partir del proceso de reducción por renglones sin la necesidad de calcular matrices elementales. Recuerde que se supone que A puede reducirse a forma escalonada por renglones sin usar intercambios de renglón. Si este es el caso, entonces todo el proceso de reducción por renglones puede hacerse usando solamente operaciones elementales con renglones de la forma $R_i - kR_j$. (¿Por qué no es necesario usar la operación elemental con los renglones restantes, y multiplicar un renglón por un escalar distinto de cero?) En la operación $R_i - kR_j$, el escalar k se denominará el ***multiplicador***.

En el ejemplo 3.33, las operaciones elementales con renglones que se usaron fueron, en orden,

$$\begin{array}{ll} R_2 - 2R_1 & (\text{multiplicador} = 2) \\ R_3 + R_1 = R_3 - (-1)R_1 & (\text{multiplicador} = -1) \\ R_3 + 2R_2 = R_3 - (-2)R_2 & (\text{multiplicador} = -2) \end{array}$$

¡Los multiplicadores son precisamente las entradas de L que están bajo su diagonal! De hecho,

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

y $L_{21} = 2$, $L_{31} = -1$ y $L_{32} = -2$. Note que la operación elemental con renglones $R_i - kR_j$ tiene su multiplicador k en la entrada (i, j) de L.

Ejemplo 3.35

Encuentre una factorización *LU* de

$$A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix}$$

Solución Al reducir A a forma escalonada por renglones se obtiene

$$A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix} \xrightarrow[\substack{R_3 - R_1 \\ R_4 - (-3)R_1}]{R_2 - 2R_1} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 8 & 7 & -16 \end{bmatrix}$$

$$\xrightarrow[R_4 - 4R_2]{R_3 - \frac{1}{2}R_2} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & -1 & -8 \end{bmatrix}$$

$$\xrightarrow{R_4 - (-1)R_3} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix} = U$$

Los primeros tres multiplicadores son 2, 1 y -3, y van en las entradas subdiagonales de la primera columna de L. De modo que, hasta el momento,

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

Los siguientes dos multiplicadores son $\frac{1}{2}$ y 4, así que se sigue llenando L:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & * & 1 \end{bmatrix}$$

El multiplicador final, -1, sustituye al último * en L para producir

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

Por tanto, una factorización LU de A es

$$A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix} = LU$$

como se comprueba fácilmente.

Comentarios

- Al aplicar este método es importante notar que las operaciones elementales con renglones $R_i - kR_j$ deben realizarse de arriba abajo dentro de cada columna (y usar la entrada diagonal como pivote) y columna por columna de izquierda a derecha. Para ilustrar lo que puede salir mal si no se obedecen estas reglas, considere la siguiente reducción por renglones:

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix} = U$$

Esta vez el multiplicador se colocaría en L del modo siguiente: $L_{32} = 2$, $L_{21} = 1$. Se obtendría

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

pero $A \neq LU$. (¡Compruebe esto! Encuentre una factorización LU correcta de A.)

- Una forma alternativa de construir L es observar que los multiplicadores pueden obtenerse directamente de las matrices obtenidas en los pasos intermedios del proceso de reducción por renglones. En el ejemplo 3.33, examine los pivotes y las correspondientes columnas de las matrices que surgen en la reducción por renglones

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \rightarrow A_1 = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U$$

El primer pivote es 2, que aparece en la primera columna de A. Al dividir por el pivote las entradas de este vector columna que están en la diagonal o debajo de ella , se obtiene

$$\frac{1}{2}\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

El siguiente pivote es -3, que aparece en la segunda columna de A_1. Al dividir por el pivote las entradas de este vector columna que están en la diagonal o debajo de ella, se obtiene

$$\frac{1}{(-3)}\begin{bmatrix} \\ -3 \\ 6 \end{bmatrix} = \begin{bmatrix} \\ 1 \\ -2 \end{bmatrix}$$

El pivote final (que no es necesario usar) es 2, en la tercera columna de U. Al dividir por el pivote las entradas de este vector columna que están en la diagonal o debajo de ella, se obtiene

$$\frac{1}{2}\begin{bmatrix} \\ \\ 2 \end{bmatrix} = \begin{bmatrix} \\ \\ 1 \end{bmatrix}$$

Si los tres vectores columna resultantes se colocan lado a lado en una matriz, se tiene

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

que es exactamente L una vez que las entradas sobre la diagonal se llenan con ceros.

En el capítulo 2 se remarcó que la forma escalonada por renglones de una matriz no es única. Sin embargo, si una matriz *invertible* A tiene una factorización LU, $A = LU$, entonces esta factorización es única.

Teorema 3.16 Si A es una matriz invertible que tiene una factorización LU, entonces L y U son únicas.

Demostración Suponga $A = LU$ y $A = L_1U_1$ son dos factorizaciones LU de A. Entonces $LU = L_1U_1$, donde L y L_1 son triangulares inferiores unitarias y U y U_1 son triangulares superiores. De hecho, U y U_1 son dos formas escalonadas por renglones (posiblemente diferentes) de A.

Por el ejercicio 30, L_1 es invertible. Dado que A es invertible, su forma escalonada reducida por renglones es una matriz identidad I por el teorema fundamental de matrices invertibles. Por tanto, U también se reduce por renglones a I (¿por qué?) y en consecuencia U también es invertible. Por lo cual,

$$L_1^{-1}(LU)U^{-1} = L_1^{-1}(L_1U_1)U^{-1} \quad \text{así} \quad (L_1^{-1}L)(UU^{-1}) = (L_1^{-1}L_1)(U_1U^{-1})$$

En consecuencia,

$$(L_1^{-1}L)I = I(U_1U^{-1}) \quad \text{entonces} \quad L_1^{-1}L = U_1U^{-1}$$

Pero $L_1^{-1}L$ es triangular inferior unitaria por el ejercicio 29 y U_1U^{-1} es triangular superior. (¿Por qué?) Se tiene que $L_1^{-1}L = U_1U^{-1}$ es *tanto* triangular inferior unitaria *como* triangular superior. La única matriz de esta forma es la matriz identidad, de modo que $L_1^{-1}L = I$ y $U_1U^{-1} = I$. Se tiene que $L = L_1$ y $U = U_1$, de modo que la factorización LU de A es única.

La factorización P^TLU

Ahora se explorará el problema de adaptar la factorización LU para manejar casos donde son necesarios los intercambios de renglón durante la eliminación gaussiana. Considere la matriz

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ -1 & 1 & 4 \end{bmatrix}$$

Una reducción por renglón directa produce

$$A \rightarrow B = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & 3 & 3 \end{bmatrix}$$

que no es una matriz triangular superior. Sin embargo, esto se puede convertir fácilmente en forma triangular superior al intercambiar los renglones 2 y 3 de B para obtener

$$U = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix}$$

Alternativamente, primero puede intercambiar los renglones 2 y 3 de A. Para este fin, sea P la matriz elemental

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

que corresponde a intercambiar los renglones 2 y 3, y sea E el producto de las matrices elementales que entonces reduce PA a U (de modo que $E^{-1} = L$ es triangular inferior unitaria). Por tanto, $EPA = U$, de modo que $A = (EP)^{-1}U = P^{-1}E^{-1}U = P^{-1}LU$.

Ahora esto sólo maneja el caso de *un solo* intercambio de renglón. En general, P será el producto $P = P_k \cdots P_2P_1$ de todas las matrices que intercambian renglón $P_1, P_2, \ldots, P_k$ (donde P_1 se realiza primero, etcétera). Tal matriz P se llama ***matriz permutación***. Observe que una matriz permutación surge de permutar los renglones de una matriz identidad en cierto orden. Por ejemplo, las siguientes son todas matrices permutación:

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

Por fortuna, el inverso de una matriz permutación es fácil de calcular; de hecho, ¡no se necesita cálculo alguno!

Teorema 3.17 Si P es una matriz permutación, entonces $P^{-1} = P^T$.

Demostración Debe demostrar que $P^TP = I$. Pero el i-ésimo renglón de P^T es igual que la i-ésima columna de P, y ambos son iguales al mismo vector unitario estándar $\mathbf{e}$, porque P es una matriz permutación. De este modo

$$(P^TP)_{ii} = (i\text{-ésimo renglón de } P^T)(i\text{-ésima columna de } P) = \mathbf{e}^T\mathbf{e} = \mathbf{e}\cdot\mathbf{e} = 1$$

Esto muestra que las entradas diagonales de P^TP son todas 1. Por otra parte, si $j \neq i$, entonces la j-ésima columna de P es un vector unitario estándar *diferente* de $\mathbf{e}$, por decir $\mathbf{e}'$. Por tanto, una entrada típica fuera de la diagonal de P^TP está dado por

$$(P^TP)_{ij} = (i\text{-ésimo renglón de } P^T)(j\text{-ésima columna de } P) = \mathbf{e}^T\mathbf{e}' = \mathbf{e}\cdot\mathbf{e}' = 0$$

En consecuencia, P^TP es una matriz identidad, como se quería demostrar.

Por tanto, en general, puede factorizar una matriz cuadrada A como $A = P^{-1}LU = P^TLU$.

Definición Sea A una matriz cuadrada. Una factorización de A como $A = P^TLU$, donde P es una matriz permutación, L es triangular inferior unitaria y U es triangular superior, se llama ***factorización P^TLU*** de A.

Ejemplo 3.36

Encuentre una factorización P^TLU de $A = \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$.

Solución Primero reduzca A a forma escalonada por renglones. Claramente, es necesario al menos un intercambio de renglón.

$$A = \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 2 & 1 & 4 \end{bmatrix} \xrightarrow{R_3 - 2R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 0 & -3 & -2 \end{bmatrix}$$
$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix}$$

Se usaron dos intercambios de renglón ($R_1 \leftrightarrow R_2$ y luego $R_2 \leftrightarrow R_3$), de modo que la matriz permutación requerida es

$$P = P_2P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Ahora se encuentra una factorización LU de PA.

$$PA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 0 & 0 & 6 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix} = U$$

Por tanto, $L_{21} = 2$, y por tanto

$$A = P^TLU = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix}$$

La discusión anterior justifica el siguiente teorema.

Teorema 3.18 Toda matriz cuadrada tiene una factorización $P^T LU$.

Comentario Incluso para una matriz invertible, la factorización $P^T LU$ no es única. En el ejemplo 3.36, un solo intercambio de renglón $R_1 \leftrightarrow R_3$ también habría funcionado, lo que conduciría a una P diferente. Sin embargo, una vez determinada P, L y U son únicas.

Consideraciones de cálculo

Si A es de $n \times n$, entonces el número total de operaciones (multiplicaciones y divisiones) requeridas para resolver un sistema lineal $A\mathbf{x} = \mathbf{b}$ usando una factorización LU de A, es $T(n) \approx n^3/3$, el mismo que se requiere para la eliminación gaussiana. (Vea la Exploración “Operaciones de conteo”, en el capítulo 2.) Esto difícilmente es sorprendente, pues la fase de eliminación hacia adelante produce la factorización LU en $\approx n^3/3$ pasos, mientras que las sustituciones hacia adelante y hacia atrás requieren $\approx n^2/2$ pasos. Por tanto, para valores grandes de n, el término $n^3/3$ es dominante. Desde este punto de vista, entonces, la eliminación gaussiana y la factorización LU son equivalentes.

Sin embargo, la factorización LU tiene otras ventajas:

- Desde un punto de vista de almacenamiento, la factorización LU es muy compacta, porque puede *sobreescribir* las entradas de A con las entradas de L y U *conforme se calculan*. En el ejemplo 3.33 se encontró que

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = LU$$

Esto puede almacenarse como

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

con las entradas colocadas en el oren $(1, 1), (1, 2), (1, 3), (2, 1), (3, 1), (2, 2), (2, 3), (3, 2), (3, 3)$. En otras palabras, las entradas subdiagonales de A se sustituyen con los multiplicadores correspondientes. (¡Compruebe que esto funciona!)

- Una vez calculada la factorización A de LU, puede usarse para resolver tantos sistemas lineales de la forma $A\mathbf{x} = \mathbf{b}$ como se quiera. Sólo necesita aplicar el método del ejemplo 3.34 y variar el vector $\mathbf{b}$ cada vez.
- Para matrices con ciertas formas especiales, sobre todo aquellas con un gran número de ceros (las llamadas matrices "dispersas") concentrados fuera de la diagonal, existen métodos que simplificarán el cálculo de una factorización LU. En estos casos, este método es más rápido que la eliminación gaussiana para resolver $A\mathbf{x} = \mathbf{b}$.
- Para una matriz invertible A, puede usarse una factorización LU de A para encontrar A^{-1}, si es necesario. Más aún, esto puede hacerse en tal forma que simultáneamente produzca una factorización de A^{-1}. (Vea los ejercicios 15-18.)

Comentario Si tiene un CAS (como MATLAB) que tenga incluida la factorización *LU*, puede notar algunas diferencias entre sus cálculos a mano y el resultado de la computadora. Esto se debe a que la mayoría de los CAS automáticamente tratarán de realizar pivoteo parcial para reducir errores de redondeo. (Vea la exploración "Pivoteo parcial" en el capítulo 2.) El ensayo de Turing es una discusión extensa de tales errores en el contexto de las factorizaciones de matrices.

Esta sección sirvió para presentar una de las factorizaciones matriciales más útiles. En capítulos posteriores se encontrarán otras igualmente útiles.

Ejercicios 3.4

En los ejercicios 1-6, resuelva el sistema $A\mathbf{x} = \mathbf{b}$ usando la factorización LU dada de A.

1. $A = \begin{bmatrix} -2 & 1 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 0 & 6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

2. $A = \begin{bmatrix} 4 & -2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ 0 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 8 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 3 & -4 \\ 4 & -3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -\frac{5}{4} & 1 \end{bmatrix}$
$\times \begin{bmatrix} 2 & 1 & -2 \\ 0 & 4 & -6 \\ 0 & 0 & -\frac{7}{2} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 2 & -4 & 0 \\ 3 & -1 & 4 \\ -1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}$
$\times \begin{bmatrix} 2 & -4 & 0 \\ 0 & 5 & 4 \\ 0 & 0 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}$

5. $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 6 & -4 & 5 & -3 \\ 8 & -4 & 1 & 0 \\ 4 & -1 & 0 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 \\ 2 & -1 & 5 & 1 \end{bmatrix}$
$\times \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & -1 & 5 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 4 & 3 & 0 \\ -2 & -5 & -1 & 2 \\ 3 & 6 & -3 & -4 \\ -5 & -8 & 9 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ -5 & 4 & -2 & 1 \end{bmatrix}$
$\times \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 3 & 5 & 2 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ -3 \\ -1 \\ 0 \end{bmatrix}$

En los ejercicios 7-12, encuentre una factorización LU de la matriz dada.

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

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

9. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 7 & 9 \end{bmatrix}$ **10.** $\begin{bmatrix} 2 & 2 & -1 \\ 4 & 0 & 4 \\ 3 & 4 & 4 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 6 & 3 & 0 \\ 0 & 6 & -6 & 7 \\ -1 & -2 & -9 & 0 \end{bmatrix}$

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

Generalice la definición de factorización LU a matrices no cuadradas al simplemente requerir que U sea una matriz en forma escalonada por renglones. Con esta modificación, encuentre una factorización LU de las matrices en los ejercicios 13 y 14.

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

14. $\begin{bmatrix} 1 & 2 & 0 & -1 & 1 \\ -2 & -7 & 3 & 8 & -2 \\ 1 & 1 & 3 & 5 & 2 \\ 0 & 3 & -3 & -6 & 0 \end{bmatrix}$

Para una matriz invertible con una factorización LU, $A = LU$, L y U serán invertibles y $A^{-1} = U^{-1}L^{-1}$. En los ejercicios 15 y 16, encuentre L^{-1}, U^{-1} y A^{-1} para la matriz dada.

15. A en el ejercicio 1 **16.** A en el ejercicio 4

El inverso de una matriz también se puede calcular al resolver varios sistemas de ecuaciones usando el método del ejemplo 3.34. Para una matriz A de $n \times n$, para encontrar su inverso es necesario resolver $AX = I_n$ para la matriz X de $n \times n$. Al escribir esta ecuación como $A[\mathbf{x}_1 \;\; \mathbf{x}_2 \cdots \mathbf{x}_n] = [\mathbf{e}_1 \;\; \mathbf{e}_2 \cdots \mathbf{e}_n]$, y usar la forma matriz-columna de AX, se ve que es necesario resolver n sistemas de ecuaciones lineales. $A\mathbf{x}_1 = \mathbf{e}_1, A\mathbf{x}_2 = \mathbf{e}_2, \ldots, A\mathbf{x}_n = \mathbf{e}_n$. Más aún, puede usar la factorización $A = LU$ para resolver cada uno de dichos sistemas.

En los ejercicios 17 y 18, use el enfoque recién destacado para encontrar A^{-1} para la matriz dada. Compare con el método de los ejercicios 15 y 16.

17. A en el ejercicio 1 **18.** A en el ejercicio 4

En los ejercicios 19-22, escriba la matriz permutación dada como un producto de matrices elementales (intercambio de renglón).

19. $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ **20.** $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

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

En los ejercicios 23-25, encuentre una factorización $P^T LU$ de la matriz A dada.

23. $A = \begin{bmatrix} 0 & 1 & 4 \\ -1 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix}$ **24.** $A = \begin{bmatrix} 0 & 0 & 1 & 2 \\ -1 & 1 & 3 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$

25. $A = \begin{bmatrix} 0 & -1 & 1 & 3 \\ -1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

26. Demuestre que existen exactamente $n!$ matrices permutación de $n \times n$.

En los ejercicios 27-28, resuelva el sistema $A\mathbf{x} = \mathbf{b}$ usando la factorización dada $A = P^T LU$. Puesto que $PP^T = I$, $P^TLU\mathbf{x} = \mathbf{b}$ puede reescribirse como $LU\mathbf{x} = P\mathbf{b}$. Entonces este sistema puede resolverse con el método del ejemplo 3.34.

27. $A = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 3 & 2 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$

$\times \begin{bmatrix} 2 & 3 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & -\frac{5}{2} \end{bmatrix} = P^TLU, \; \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}$

28. $A = \begin{bmatrix} 8 & 3 & 5 \\ 4 & 1 & 2 \\ 4 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}$

$\times \begin{bmatrix} 4 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix} = P^TLU, \; \mathbf{b} = \begin{bmatrix} 16 \\ -4 \\ 4 \end{bmatrix}$

29. Demuestre que un producto de matrices triangulares inferiores unitarias es triangular inferior unitaria.

30. Demuestre que toda matriz triangular inferior unitaria es invertible y que su inversa también es triangular inferior unitaria.

*Una **factorización LDU** de una matriz cuadrada A es una factorización A = LDU, donde L es una matriz triangular inferior unitaria, D es una matriz diagonal y U es una matriz triangular superior unitaria (triangular superior con 1 en su diagonal). En los ejercicios 31 y 32, encuentre una factorización LDU de A.*

31. A en el ejercicio 1 **32.** A en el ejercicio 4

33. Si A es simétrica e invertible y tiene una factorización LDU, demuestre que $U = L^T$.

34. Si A es simétrica e invertible, y $A = LDL^T$ (con L triangular inferior unitaria y D diagonal), demuestre que esta factorización es única. Esto es: demuestre que, si también se tiene $A = L_1D_1L_1^T$ (con L_1 triangular inferior unitaria y D_1 diagonal), entonces $L = L_1$ y $D = D_1$.

3.5 Subespacios, bases, dimensión y rank

Esta sección presenta tal vez las ideas más importantes de todo el libro. Ya se vio que existe una interacción entre geometría y álgebra: con frecuencia es posible usar intuición geométrica y razonar para obtener resultados algebraicos, y la potencia del álgebra con frecuencia permitirá extender los hallazgos más allá de los escenarios geométricos en los que surgieron por primera vez.

En el estudio de los vectores ya encontró informalmente todos los conceptos en esta sección. Aquí, comenzará a volverse más formal al proporcionar definiciones para las ideas clave. Como verá, la noción de *subespacio* es simplemente una generalización algebraica de los ejemplos geométricos de rectas y planos que pasan por el origen. El concepto fundamental de *base* para un subespacio se deriva entonces de la idea de vectores directores para tales rectas y planos. El concepto de base le permitirá dar una definición precisa de *dimensión* que concuerde con una idea geométrica intuitiva del término, aunque es suficientemente flexible para permitir la generalización a otros escenarios.

También comenzará a ver que dichas ideas arrojan más luz a lo que ya sabe acerca de las matrices y la solución de sistemas de ecuaciones lineales. En el capítulo 6 encontrará nuevamente todas estas ideas fundamentales con más detalle. Considere esta sección como una sesión de "familiarización".

Un plano que pasa por el origen en $\mathbb{R}^3$ "parece" una copia de $\mathbb{R}^2$. Intuitivamente, estaría de acuerdo en que ambos son "bidimensionales". Si presiona un poco, también puede decir que cualquier cálculo que pueda hacerse con vectores en $\mathbb{R}^2$ también puede realizarse en un plano que pasa por el origen. En particular, puede sumar y tomar múltiplos escalares (y, más generalmente, formar combinaciones lineales) de vectores en tal plano, y los resultados son otros vectores *en el mismo plano*. Se dice que, como $\mathbb{R}^2$, un plano que pasa por el origen es *cerrado* con respecto a las operaciones de suma y multiplicación escalar. (Vea la figura 3.2.)

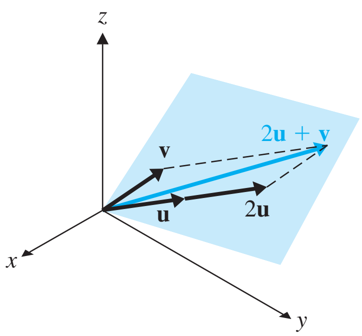


Figura 3.2

Pero los vectores en este plano, ¿son objetos bidimensionales o tridimensionales? Puede argumentar que son tridimensionales porque existen en $\mathbb{R}^3$ y por tanto tienen tres componentes. Por otra parte, pueden describirse como una combinación lineal de sólo dos vectores, vectores dirección para el plano, y por tanto son objetos bidimensionales que existen en un plano bidimensional. La noción de subespacio es la clave para resolver este rompecabezas.

Definición Un ***subespacio*** de $\mathbb{R}^n$ es cualquier colección S de vectores en $\mathbb{R}^n$ tal que:

1. El vector cero $\mathbf{0}$ está en S.
2. Si $\mathbf{u}$ y $\mathbf{v}$ están en S, entonces $\mathbf{u} + \mathbf{v}$ está en S. (S es ***cerrado bajo la suma***.)
3. Si $\mathbf{u}$ está en S y c es un escalar, entonces $c\mathbf{u}$ está en S. (S es ***cerrado bajo la multiplicación por un escalar***.)

También podría combinar las propiedades (2) y (3) y requerir, de manera equivalente, que S sea ***cerrado bajo combinaciones lineales***:

$$\text{Si } \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \text{ están en } S \text{ y } c_1, c_2, \ldots, c_k \text{ son escalares,}$$
$$\text{entonces } c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k \text{ están en } S.$$

Ejemplo 3.37

Toda recta y plano a través del origen en $\mathbb{R}^3$ es un subespacio de $\mathbb{R}^3$. Debe ser claro geométricamente que se satisfacen las propiedades (1) a (3). He aquí una denostración algebraica en el caso de un plano a través del origen. En el ejercicio 9 se le pide dar la demostración correspondiente para una recta.

Sea $\mathcal{P}$ un plano a través del origen con vectores directores $\mathbf{v}_1$ y $\mathbf{v}_2$. Por tanto, $\mathcal{P} = \text{gen}(\mathbf{v}_1,\mathbf{v}_2)$. El vector cero $\mathbf{0}$ está en $\mathcal{P}$, pues $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. Ahora sean

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \quad \text{y} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2$$

dos vectores en $\mathcal{P}$. Entonces

$$\mathbf{u} + \mathbf{v} = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2) + (d_1\mathbf{v}_1 + d_2\mathbf{v}_2) = (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2$$

Por tanto, $\mathbf{u} + \mathbf{v}$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$,y por tanto está en $\mathcal{P}$.

Ahora sea c un escalar. Entonces

$$c\mathbf{u} = c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2$$

que muestra que $c\mathbf{u}$ también es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ y por tanto está en $\mathcal{P}$. Se demostró que $\mathcal{P}$ satisface las propiedades (1) a (3) y en consecuencia es un subespacio de $\mathbb{R}^3$.

Si observa cuidadosamente los detalles del ejemplo 3.37, notará que el hecho de que $\mathbf{v}_1$ y $\mathbf{v}_2$ fueran vectores en $\mathbb{R}^3$ no tuvo papel alguno en la verificación de las propiedades. En consecuencia, el método algebraico utilizado debe generalizarse más allá de $\mathbb{R}^3$ y aplicarse en situaciones donde ya no se pueda visualizar la geometría. Lo hace. Más aún, el método del ejemplo 3.37 puede servir como “plantilla” en escenarios más generales. Cuando el ejemplo 3.37 se generaliza al generador de un conjunto arbitrario de vectores en cualquier $\mathbb{R}^n$, el resultado es suficientemente importante como para llamarse teorema.

Teorema 3.19 Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vectores en $\mathbb{R}^n$. Entonces $\text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ es un subespacio de $\mathbb{R}^n$.

Demostración Sea $S = \text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$. Para demostrar la propiedad (1) de la definición, simplemente se observa que el vector cero $\mathbf{0}$ está en S, ya que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k$.

Ahora sean

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \quad \text{y} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$$

dos vectores en S. Entonces

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) + (d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k)\\ &= (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2 + \cdots + (c_k + d_k)\mathbf{v}_k\end{aligned}$$

Por tanto, $\mathbf{u} + \mathbf{v}$ es una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y por tanto está en S. Esto verifica la propiedad (2).

Para demostrar la propiedad (3), sea c un escalar. Entonces

$$\begin{aligned}c\mathbf{u} &= c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k)\\ &= (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2 + \cdots + (cc_k)\mathbf{v}_k\end{aligned}$$

lo que demuestra que $c\mathbf{u}$ también es una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y por tanto está en S. Se demostró que S satisface las propiedades (1) a (3) y en consecuencia es un subespacio de $\mathbb{R}^n$.

A gen$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ se le llamará ***subespacio generado por*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Con frecuencia podrá ahorrar mucho trabajo al reconocer cuándo puede aplicarse el Teorema 3.19.

Ejemplo 3.38

Demuestre que el conjunto de todos los vectores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisfacen las condiciones $x = 3y$ y $z = -2y$ forman un subespacio de $\mathbb{R}^3$.

Solución La sustitución de las dos condiciones en $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ produce

$$\begin{bmatrix} 3y \\ y \\ -2y \end{bmatrix} = y\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$$

Dado que y es arbitrario, el conjunto dado de vectores es gen$\left(\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}\right)$ y por ende es un subespacio de $\mathbb{R}^3$, por el Teorema 3.19.

Geométricamente, el conjunto de vectores del ejemplo 3.38 representa la recta que pasa por el origen en $\mathbb{R}^3$, con vector director $\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$.

Ejemplo 3.39

Determine si el conjunto de todos los vectores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisface las condiciones $x = 3y + 1$ y $z = -2y$ es un subespacio de $\mathbb{R}^3$.

Solución Esta vez se tienen todos los vectores de la forma

$$\begin{bmatrix} 3y + 1 \\ y \\ -2y \end{bmatrix}$$

El vector cero no es de esta forma. (¿Por qué no? Trate de resolver $\begin{bmatrix} 3y + 1 \\ y \\ -2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.) Por tanto, la propiedad (1) no se cumple, de modo que este conjunto no puede ser un subespacio de $\mathbb{R}^3$.

Ejemplo 3.40

Determine si el conjunto de todos los vectores $\begin{bmatrix} x \\ y \end{bmatrix}$, donde $y = x^2$, es un subespacio de $\mathbb{R}^2$.

Solución Se trata de los vectores de la forma $\begin{bmatrix} x \\ x^2 \end{bmatrix}$, llame S a este conjunto. Esta vez $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ pertenece a S (tome $x = 0$), de modo que se cumple la propiedad (1). Sea $\mathbf{u} = \begin{bmatrix} x_1 \\ x_1^2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} x_2 \\ x_2^2 \end{bmatrix}$ en S. Entonces

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} x_1 + x_2 \\ x_1^2 + x_2^2 \end{bmatrix}$$

que, en general, no está en S, pues no tiene la forma correcta; esto es: $x_1^2 + x_2^2 \neq (x_1 + x_2)^2$. Para ser específico, busque un contraejemplo. Si

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

entonces tanto $\mathbf{u}$ como $\mathbf{v}$ están en S, pero su suma $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ no está en S pues $5 \neq 3^2$.

Por tanto, la propiedad (2) falla y S no es un subespacio de $\mathbb{R}^2$.

Comentario Para que un conjunto S sea un subespacio de algún $\mathbb{R}^n$, debe *demostrar* que las propiedades (1) a (3) se sustentan *en general*. Sin embargo, para que S *falle* en ser subespacio de $\mathbb{R}^n$, es suficiente demostrar que no se cumple *una* de las tres propiedades. La ruta más sencilla por lo general es encontrar un solo *contraejemplo* específico para ilustrar la falla de la propiedad. Una vez hecho esto, no hay necesidad de considerar las otras propiedades.

Subespacios asociados con matrices

En el contexto de las matrices surgen una gran cantidad de ejemplos de subespacios. Ya se encontraron las más importantes en el capítulo 2; ahora se revisarán nuevamente con la noción de subespacio en mente.

Definición Sea A una matriz de $m \times n$.

1. El ***espacio renglón*** de A es el subespacio renglón(A) de $\mathbb{R}^n$ generado por los renglones de A.
2. El ***espacio columna*** de A es el subespacio col(A) de $\mathbb{R}^m$ generado por las columnas de A.

Ejemplo 3.41 Considere la matriz

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

(a) Determine si $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ está en el espacio columna de A.

(b) Determine si $\mathbf{w} = [4 \quad 5]$ está en el espacio renglón de A.

(c) Describa renglón(A) y col(A).

Solución

(a) Por el Teorema 2.4 y el análisis que le precedió, $\mathbf{b}$ es una combinación lineal de las columnas de A si y sólo si el sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente. La matriz aumentada se reduce por renglones del modo siguiente:

$$\left[\begin{array}{rr|r} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 3 & -3 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

Por tanto, el sistema es consistente (y, de hecho, tiene una solución única). Por tanto, $\mathbf{b}$ está en col(A). (De hecho este ejemplo es el 2.18, parafraseado en la terminología de esta sección.)

(b) Como también se vio en la sección 2.3, las operaciones elementales con renglones simplemente crean combinaciones lineales de los renglones de una matriz. Esto es, producen vectores solamente en el espacio renglón de la matriz. Si el vector $\mathbf{w}$ está en renglón(A), entonces $\mathbf{w}$ es una combinación lineal de los renglones de A, de modo que si A se aumenta con $\mathbf{w}$ como $\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right]$, será posible aplicar operaciones elementales con renglones a esta matriz aumentada para reducirla a la forma $\left[\begin{array}{c} A' \\ \hline \mathbf{0} \end{array}\right]$ usando solamente operaciones elementales con renglones de la forma $R_i + kR_j$, donde $i > j$; en otras palabras, *trabajar de arriba abajo en cada columna*. (¿Por qué?)

En este ejemplo se tiene

$$\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right] = \left[\begin{array}{rr} 1 & -1 \\ 0 & 1 \\ 3 & -3 \\ \hline 4 & 5 \end{array}\right] \xrightarrow[R_4 - 4R_1]{R_3 - 3R_1} \left[\begin{array}{rr} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 9 \end{array}\right] \xrightarrow{R_4 - 9R_2} \left[\begin{array}{rr} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 0 \end{array}\right]$$

Por tanto, **w** es una combinación lineal de los renglones de A (de hecho, estos cálculos muestran que $\mathbf{w} = 4[1 \quad -1] + 9[0 \quad 1]$; ¿cómo?), y por tanto **w** está en renglón(A).

(c) Es fácil comprobar que, para cualquier vector $\mathbf{w} = [x \quad y]$, la matriz aumentada $\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right]$ se reduce a

$$\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 0 \end{array}\right]$$

en forma similar. Por tanto, todo vector en $\mathbb{R}^2$ está en renglón(A), y en consecuencia renglón (A) $= \mathbb{R}^2$.

Encontrar col(A) es idéntico a resolver el ejemplo 2.21, en que se determinó que coincide con el plano (a través del origen) en $\mathbb{R}^3$ con ecuación $3x - z = 0$. (Dentro de poco se descubrirán otras formas de responder este tipo de preguntas.)

Comentario También podría responder el inciso (b) y la primera parte del inciso (c) al observar que cualquier pregunta acerca de los *renglones* de A es la correspondiente pregunta acerca de las *columnas* de A^T. De este modo, por ejemplo **w** está en renglón(A) si y sólo si $\mathbf{w}^T$ está en col(A^T). Esto es verdadero si y sólo si el sistema $A^T\mathbf{x} = \mathbf{w}^T$ es consistente. Ahora puede proceder como en el inciso (a). (Vea los ejercicios 21-24.)

Los comentarios hechos acerca de la relación entre operaciones elementales con renglones y el espacio renglón se resumen en el siguiente teorema.

Teorema 3.20 Sea B cualquier matriz que es equivalente por renglones a una matriz A. Entonces renglón(B) = renglón(A).

Demostración La matriz A puede transformarse en B mediante una secuencia de operaciones con renglones. En consecuencia, los renglones de B son combinaciones lineales de los renglones de A; por tanto, las combinaciones lineales de los renglones de B son combinaciones lineales de los renglones de A. (Vea el ejercicio 21 en la sección 2.3.) Se tiene que renglón(B) $\subseteq$ renglón(A).

Por otra parte, invertir dichas operaciones con renglones transforman B en A. Por tanto, el argumento anterior demuestra que renglón(A) $\subseteq$ renglón(B). Combinando estos resultados renglón(A) = renglón(B)

Existe otro importante subespacio que ya se encontró: el conjunto de soluciones de un sistema homogéneo de ecuaciones lineales. Es fácil probar que este subespacio satisface las tres propiedades de los subespacios.

Teorema 3.21 Sea A una matriz de $m \times n$ y sea N el conjunto de soluciones del sistema lineal homogéneo $A\mathbf{x} = \mathbf{0}$. Entonces N es un subespacio de $\mathbb{R}^n$.

Demostración [Note que **x** debe ser un vector (columna) en $\mathbb{R}^n$ para que A**x** esté definido y que $\mathbf{0} = \mathbf{0}_m$ es el vector cero en $\mathbb{R}^m$.] Dado que $A\mathbf{0}_n = \mathbf{0}_m$, $\mathbf{0}_n$ está en N. Ahora sean **u** y **v** en N. Por tanto, $A\mathbf{u} = \mathbf{0}$ y $A\mathbf{v} = \mathbf{0}$. Se tiene que

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

En consecuencia, $\mathbf{u} + \mathbf{v}$ está en N. Finalmente, para cualquier escalar c,

$$A(c\mathbf{u}) = c(A\mathbf{u}) = c\mathbf{0} = \mathbf{0}$$

y en consecuencia $c\mathbf{u}$ también está en N. Se tiene que N es un subespacio de $\mathbb{R}^n$.

Definición Sea A una matriz de $m \times n$. El ***espacio nulo*** de A es el subespacio de $\mathbb{R}^n$ que consiste de las soluciones del sistema lineal homogéneo $A\mathbf{x} = \mathbf{0}$. Se denota mediante nul(A).

El hecho de que el espacio nulo de una matriz sea un subespacio permite probar lo que la intuición y los ejemplos condujeron a entender acerca de las soluciones de los sistemas lineales: no tienen solución, tienen una solución única o un número infinito de soluciones.

Teorema 3.22 Sea A una matriz cuyas entradas son números reales. Para cualquier sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$, exactamente uno de los siguientes es verdadero:

a. No hay solución.
b. Hay una solución única.
c. Hay un número infinito de soluciones.

A primera vista, no es completamente claro cómo se debe proceder para probar este teorema. Un poco de reflexión debe persuadirlo de que lo que realmente se pide probar es que si (a) y (b) no son verdaderas, entonces (c) es la única posibilidad. Esto es, si hay más de una solución, entonces no puede haber sólo dos o incluso muchas en una cantidad finita, sino que debe haber un número infinito.

Demostración Si el sistema $A\mathbf{x} = \mathbf{b}$ no tiene solución o tiene exactamente una solución, está demostrado. Suponga, entonces, que existen al menos dos soluciones distintas de $A\mathbf{x} = \mathbf{b}$, por ejemplo $\mathbf{x}_1$ y $\mathbf{x}_2$. Por tanto,

$$A\mathbf{x}_1 = \mathbf{b} \quad \text{y} \quad A\mathbf{x}_2 = \mathbf{b}$$

con $\mathbf{x}_1 \neq \mathbf{x}_2$. Se tiene que

$$A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Sea $\mathbf{x}_0 = \mathbf{x}_1 - \mathbf{x}_2$. Entonces $\mathbf{x}_0 \neq \mathbf{0}$ y $A\mathbf{x}_0 = \mathbf{0}$. En consecuencia, el espacio nulo de A es no trivial, y dado que nul(A) es cerrado para la multiplicación por un escalar, $c\mathbf{x}_0$ está en nul(A) para todo escalar c. En consecuencia, el espacio nulo de A contiene infinitos vectores (dado que contiene *al menos* a todo vector de la forma $c\mathbf{x}_0$, y existen infinitos de ellos.)

Ahora, considere los vectores (un número infinito) de la forma $\mathbf{x}_1 + c\mathbf{x}_0$, cuando c varía a lo largo del conjunto de los números reales. Se tiene

$$A(\mathbf{x}_1 + c\mathbf{x}_0) = A\mathbf{x}_1 + cA\mathbf{x}_0 = \mathbf{b} + c\mathbf{0} = \mathbf{b}$$

Por tanto, existe un número infinito de soluciones de la ecuación $A\mathbf{x} = \mathbf{b}$.

Bases

Es posible extraer un poco más de la idea intuitiva de que los subespacios son generalizaciones de planos a través del origen en $\mathbb{R}^3$. Un plano se genera mediante cualesquiera dos vectores que son paralelos al plano pero no son paralelos entre ellos. En lenguaje algebraico, dos de tales vectores generan el plano y son linealmente independientes. Menos de dos vectores no funcionarán; más de dos vectores no son necesarios. Esta es la esencia de una *base* para un subespacio.

Definición Una ***base*** para un subespacio S de $\mathbb{R}^n$ es un conjunto de vectores en S que

1. genera S y
2. es linealmente independiente.

Ejemplo 3.42

En la sección 2.3 se vio que los vectores unitarios estándar $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$ en $\mathbb{R}^n$ son linealmente independientes y generan $\mathbb{R}^n$. Por tanto, forman una base para $\mathbb{R}^n$, llamada ***base estándar***.

Ejemplo 3.43

En el ejemplo 2.19 se demostró que $\mathbb{R}^2 = \text{gen}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$. Dado que $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ también son linealmente independientes (ya que no son múltiplos), forman una base para $\mathbb{R}^2$.

Un subespacio puede tener (y tendrá) más de una base. Por ejemplo, recién se vio que $\mathbb{R}^2$ tiene la base estándar $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ y la base $\left\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right\}$. Sin embargo, dentro de poco se probará que el *número* de vectores en una base para un subespacio dado siempre será el mismo.

Ejemplo 3.44

Encuentre una base para $S = \text{gen}(\mathbf{u}, \mathbf{v}, \mathbf{w})$, donde

$$\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \quad \text{y} \quad \mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}$$

Solución Los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ ya generan S, de modo que serán una base de S si también son linealmente independientes. Es fácil determinar que no lo son; de hecho, $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$. Por tanto, puede ignorarse $\mathbf{w}$, pues cualquier combinación lineal que involucre $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ puede reescribirse para involucrar sólo a $\mathbf{u}$ y $\mathbf{v}$. (Vea también el ejercicio 47 de la sección 2.3.) Esto implica que $S = \text{gen}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \text{gen}(\mathbf{u}, \mathbf{v})$, y dado que $\mathbf{u}$ y $\mathbf{v}$ ciertamente son linealmente independientes (¿por qué?), forman una base para S. (Geométricamente, esto significa que $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ yacen todos en el mismo plano, y $\mathbf{u}$ y $\mathbf{v}$ pueden servir como un conjunto de vectores directores para este plano.)

Ejemplo 3.45

Encuentre una base para el espacio renglón de

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

Solución La forma escalonada reducida por renglones de A es

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

Por el Teorema 3.20, renglón(A) = renglón(R), de modo que es suficiente encontrar una base para el espacio renglón de R. Pero renglón(R) claramente es generado por sus renglones distintos de cero, y es fácil comprobar que el patrón en escalera fuerza a los primeros tres renglones de R a ser linealmente independientes. (Este es un hecho general, que necesitará establecer para probar el ejercicio 33.) En consecuencia, una base para el espacio renglón de A es

$$\{[1 \quad 0 \quad 1 \quad 0 \quad -1], [0 \quad 1 \quad 2 \quad 0 \quad 3], [0 \quad 0 \quad 0 \quad 1 \quad 4]\}$$

Puede usar el método del ejemplo 3.45 para encontrar una base para el subespacio generado por un conjunto dado de vectores.

Ejemplo 3.46

Vuelva a trabajar el ejemplo 3.44 usando el método del ejemplo 3.45.

Solución Transponga **u**, **v** y **w** para obtener vectores renglón y luego formar una matriz con dichos vectores como sus renglones:

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

Al proceder como en el ejemplo 3.45, se reduce B a su forma escalonada reducida por renglones

$$\begin{bmatrix} 1 & 0 & \frac{8}{5} \\ 0 & 1 & -\frac{1}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

y use los vectores renglón distintos de cero como base para el espacio renglón. Dado que comenzó con vectores columna, debe transponer nuevamente. Por tanto, una base para gen$(\mathbf{u}, \mathbf{v}, \mathbf{w})$ es

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ \frac{8}{5} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -\frac{1}{5} \end{bmatrix} \right\}$$

Comentarios

- De hecho, no es necesario llegar a la forma escalonada *reducida* por renglones, la forma escalonada es suficiente. Si U es una forma escalonada por renglones de A, entonces los vectores renglón distintos de cero de U formarán una base para renglón(A) (vea el ejercicio 33). Este planteamiento tiene la ventaja de (con frecuencia) permitir evitar las fracciones. En el ejemplo 3.46, B puede reducirse a

$$U = \begin{bmatrix} 3 & -1 & 5 \\ 0 & -5 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

que produce la base

$$\left\{ \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix} \right\}$$

para gen$(\mathbf{u}, \mathbf{v}, \mathbf{w})$.

- Advierta que los métodos usados en el ejemplo 3.44, el ejemplo 3.46 y el comentario anterior por lo general producirán diferentes bases.

Ahora se regresa al problema de encontrar una base para el espacio columna de una matriz A. Un método es simplemente transponer la matriz. Los vectores columna de A se convierten en los vectores renglón de A^T y puede aplicar el método del ejemplo 3.45 para encontrar una base para renglón(A^T). Al transponer estos vectores se produce una base para col(A). (En los ejercicios 21-24 se le pide hacer esto.) Sin embargo, este esquema requiere realizar un nuevo conjunto de operaciones de renglón sobre A^T.

En vez de ello, es preferible un planteamiento que permita usar la forma reducida por renglones de A que ya se calculó. Recuerde que un producto de $A\mathbf{x}$ de una matriz y un vector corresponde a una combinación lineal de las columnas de A con las entradas de $\mathbf{x}$ como coeficientes. Por tanto, una solución no trivial a $A\mathbf{x} = \mathbf{0}$ representa una *relación de dependencia* entre las columnas de A. dado que las operaciones elementales con renglones no afectan al conjunto solución, si A es equivalente por renglones a R, *las columnas de A tienen las mismas relaciones de dependencia que las columnas de R*. Esta importante observación es el fundamento de la técnica que ahora se usa para encontrar una base para col(A).

Ejemplo 3.47

Encuentre una base para el espacio columna de la matriz del ejemplo 3.45,

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

Solución Sea $\mathbf{a}_i$ un vector columna de A y $\mathbf{r}_i$ un vector columna de la forma escalonada reducida

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

Puede ver rápidamente por inspección que $\mathbf{r}_3 = \mathbf{r}_1 + 2\mathbf{r}_2$ y $\mathbf{r}_5 = -\mathbf{r}_1 + 3\mathbf{r}_2 + 4\mathbf{r}_4$. (Compruebe que, como se predijo, los correspondientes vectores columna de A satisfacen las mismas relaciones de dependencia.) Por tanto, $\mathbf{r}_3$ y $\mathbf{r}_5$ no contribuyen a col(R). Los vectores columna restantes, $\mathbf{r}_1$, $\mathbf{r}_2$ y $\mathbf{r}_4$, son linealmente independientes, pues son vectores unitarios estándar. En consecuencia, los enunciados correspondientes son verdaderos para los vectores columna de A.

Por tanto, entre los vectores columna de A, se eliminan los dependientes ($\mathbf{a}_3$ y $\mathbf{a}_5$), y los restantes serán linealmente independientes y en consecuencia forman una base para col(A). ¿Cuál es la forma más rápida de encontrar esta base? Use las columnas de A que correspondan a las columnas de R *que contengan los 1 pivote.* Una base para col(A) es

$$\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

Advertencia ¡Las operaciones elementales con renglones cambian el espacio columna! En el ejemplo, col(A) $\neq$ col(R), ya que todo vector en col(R) tiene su cuarto componente igual a 0, pero esto ciertamente no es verdadero de col(A). De modo que debe regresar a la matriz original A para obtener los vectores columna para una base de col(A). Para ser específico, en el ejemplo 3.47, $\mathbf{r}_1$, $\mathbf{r}_2$ y $\mathbf{r}_4$ *no* forman una base para el espacio columna de A.

Ejemplo 3.48

Encuentre una base para el espacio nulo de la matriz A del ejemplo 3.47.

Solución En realidad no hay nada nuevo aquí, excepto la terminología. Simplemente debe encontrar y describir las soluciones del sistema homogéneo $A\mathbf{x} = 0$. Ya se calculó la forma escalonada reducida por renglones R de A, de modo que todo lo que falta por hacer en la eliminación de Gauss-Jordan es resolver para las variables pivote en términos de las variables libres. La matriz aumentada final es

$$[R\,|\,\mathbf{0}] = \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Si

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

entonces los 1 pivote están en las columnas 1, 2 y 4, así que se resuelve para x_1, x_2 y x_4 en términos de las variables libres x_3 y x_5. Se obtiene $x_1 = -x_3 + x_5$, $x_2 = -2x_3 - 3x_5$ y $x_4 = -4x_5$. Al hacer $x_3 = s$ y $x_5 = t$, se obtiene

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -s + t \\ -2s - 3t \\ s \\ -4t \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} = s\mathbf{u} + t\mathbf{v}$$

Por tanto, $\mathbf{u}$ y $\mathbf{v}$ generan nul(A), y dado que son linealmente independientes, forman una base para nul(A).

A continuación hay un resumen del procedimiento más efectivo para encontrar bases para el espacio renglón, el espacio columna y el espacio nulo de una matriz A.

1. Encuentre la forma escalonada reducida por renglones R de A.
2. Use los vectores renglón distintos de cero de R (que contengan los 1 pivote) para formar una base para renglón(A).
3. Use los vectores columna de A que correspondan a las columnas de R que contengan los 1 pivote (las columnas pivote) para formar una base para col(A).
4. Resuelva para las variables pivote de $R\mathbf{x} = \mathbf{0}$ en términos de las variables libres, iguale las variables libres a parámetros, sustituya de vuelta en $\mathbf{x}$ y escriba el resultado como una combinación lineal de f vectores (donde f es el número de variables libres). Estos vectores f forman una base para nul(A).

Si no necesita encontrar el espacio nulo, entonces es más rápido simplemente reducir A a la forma escalonada por renglones para encontrar bases para los espacios renglón y columna. Los pasos 2 y 3 anteriores siguen siendo válidos (con la sustitución de la palabra "pivotes" en vez de "1 pivotes").

Dimensión y rank

Ya observó que, aunque un subespacio tendrá diferentes bases, cada una tiene el mismo número de vectores. Este hecho fundamental será de vital importancia en este libro a partir de ahora.

Teorema 3.23 **El Teorema de la Base**

Sea S un subespacio de $\mathbb{R}^n$. Entonces cualesquiera dos bases para S tienen el mismo número de vectores.

Demostración Sean $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ y $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ bases de S. Necesita demostrar que $r = s$. Esto se hace al demostrar que no puede ocurrir ninguna de las otras dos posibilidades, $r < s$ o $r > s$.

Suponga que $r < s$. Se demostrará que esto fuerza a $\mathcal{C}$ a ser un conjunto linealmente dependiente de vectores. Para este fin, sea

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s = \mathbf{0} \tag{1}$$

Sherlock Holmes apuntó: "Cuando se ha eliminado lo imposible, cualquier cosa que quede, *aunque sea improbable*, debe ser la verdad" (de *El signo de los cuatro*, de sir Arthur Conan Doyle).

Dado que $\mathcal{B}$ es una base para S, puede escribir cada $\mathbf{v}_i$ como una combinación lineal de los elementos $\mathbf{u}_j$:

$$\begin{aligned} \mathbf{v}_1 &= a_{11}\mathbf{u}_1 + a_{12}\mathbf{u}_2 + \cdots + a_{1r}\mathbf{u}_r \\ \mathbf{v}_2 &= a_{21}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + \cdots + a_{2r}\mathbf{u}_r \\ &\vdots \\ \mathbf{v}_s &= a_{s1}\mathbf{u}_1 + a_{s2}\mathbf{u}_2 + \cdots + a_{sr}\mathbf{u}_r \end{aligned} \tag{2}$$

Al sustituir las ecuaciones (2) en la ecuación (1) se obtiene

$$c_1(a_{11}\mathbf{u}_1 + \cdots + a_{1r}\mathbf{u}_r) + c_2(a_{21}\mathbf{u}_1 + \cdots + a_{2r}\mathbf{u}_r) + \cdots + c_s(a_{s1}\mathbf{u}_1 + \cdots + a_{sr}\mathbf{u}_r) = \mathbf{0}$$

Al reagrupar se tiene

$$(c_1a_{11} + c_2a_{21} + \cdots + c_sa_{s1})\mathbf{u}_1 + (c_1a_{12} + c_2a_{22} + \cdots + c_sa_{s2})\mathbf{u}_2$$
$$+ \cdots + (c_1a_{1r} + c_2a_{2r} + \cdots + c_sa_{sr})\mathbf{u}_r = \mathbf{0}$$

Ahora, dado que $\mathcal{B}$ es una base, los términos $\mathbf{u}_j$ son linealmente independientes. De modo que cada una de las expresiones entre paréntesis debe ser cero:

$$\begin{aligned} c_1a_{11} + c_2a_{21} + \cdots + c_sa_{s1} &= 0 \\ c_1a_{12} + c_2a_{22} + \cdots + c_sa_{s2} &= 0 \\ &\vdots \\ c_1a_{1r} + c_2a_{2r} + \cdots + c_sa_{sr} &= 0 \end{aligned}$$

Este es un sistema homogéneo de r ecuaciones lineales con las s variables $c_1, c_2, \ldots, c_s$. (El hecho de que las variables aparezcan a la izquierda de los coeficientes no hace diferencia.) Dado que $r < s$, se sabe del Teorema 2.3 que existe un número infinito de soluciones. En particular, existe una solución no trivial, lo que da una relación de dependencia no trivial en la ecuación (1). Por tanto, $\mathcal{C}$ es un conjunto linealmente dependiente de vectores. Pero este hallazgo contradice el hecho de que $\mathcal{C}$ fue dado como base y en consecuencia es linealmente independiente. Se concluye que $r < s$ no es posible. De igual modo (al intercambiar los roles de $\mathcal{B}$ y $\mathcal{C}$), se descubre que $r > s$ conduce a una contradicción. Por tanto, debe tener $r = s$, como se deseaba.

Dado que todas las bases para un subespacio dado deben tener el mismo número de vectores, puede darse un nombre a este número.

Definición Si S es un subespacio de $\mathbb{R}^n$, entonces el número de vectores en una base para S se llama la ***dimensión*** de S, y se denota dim S.

Comentario El vector cero $\mathbf{0}$ por sí mismo siempre es un subespacio de $\mathbb{R}^n$. (¿Por qué?) Sin embargo, cualquier conjunto que contenga el vector cero (y, en particular, $\{\mathbf{0}\}$) es linealmente dependiente, de modo que $\{\mathbf{0}\}$ no puede tener una base. Se define $\{\mathbf{0}\}$ como 0.

Ejemplo 3.49

Dado que la base estándar para $\mathbb{R}^n$ tiene n vectores, $\mathbb{R}^n = n$. (Note que este resultado concuerda con la comprensión intuitiva de dimensión para $n \leq 3$.)

Ejemplo 3.50

En los ejemplos 3.45 a 3.48, se descubrió que renglón(A) tiene una base con tres vectores, col(A) tiene una base con tres vectores y nul(A) tiene una base con dos vectores. En consecuencia, (renglón(A)) = 3, dim (col(A)) = 3 y dim (nul(A)) = 2.

Un solo ejemplo no es suficiente para especular, pero el hecho de que los espacios renglón y columna del ejemplo 3.50 tengan la misma dimensión no es accidente. Como tampoco lo es el hecho de que la suma de dim(col(A)) y dim(nul(A)) es 5, el número de columnas de A. Ahora se probará que estas relaciones son verdaderas en general.

Teorema 3.24 Los espacios renglón y columna de una matriz A tienen la misma dimensión.

Demostración Sea R la forma escalonada reducida por renglones de A. Por el Teorema 3.20, renglón(A) = renglón(R), de modo que

$$\begin{aligned}\dim(\text{renglón}(A)) &= \dim(\text{renglón}(R))\\ &= \text{número de renglones de } R \text{ distintos de cero}\\ &= \text{número de 1 pivote de } R\end{aligned}$$

Llame a este número r.

Ahora col$(A) \neq$ col(R), pero las columnas de A y R tienen las mismas relaciones de dependencia. Por tanto, dim (col(A)) = dim (col(R)). Dado que existen r 1 pivotes, R tiene r columnas que son vectores unitarios estándar, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$. (Éstos serán vectores en $\mathbb{R}^m$ si A y R son matrices de $m \times n$.) Esos r vectores son linealmente independientes y las columnas restantes de R son combinaciones lineales de ellos. Por tanto, dim (col(R)) $= r$. Se tiene que dim (renglón(A)) $= r =$ dim (col(A)), como se quería demostrar.

El rango de una matriz la definió por primera vez, en 1878, Georg Frobenius (1849–1917), aunque él la definió usando determinantes y no como se hace aquí. (Vea el capítulo 4.) Frobenius fue un matemático alemán que recibió su doctorado en la Universidad de Berlín, donde más tarde dio clases. Mejor conocido por sus aportaciones a la teoría de grupos, Frobenius usó matrices en su trabajo con representaciones de grupo.

Definición El ***rank*** de una matriz A es la dimensión de sus espacios renglón y columna y se denota mediante rango(A).

Por tanto, para el ejemplo 3.50, puede escribir rango$(A) = 3$.

Comentarios

- La definición anterior concuerda con la definición más informal de rank que se dio en el capítulo 2. La ventaja de la nueva definición es que es mucho más flexible.
- El rank de una matriz simultáneamente proporciona información acerca de la dependencia lineal entre los vectores renglón de la matriz *y* entre sus vectores columna. En particular, dice el número de renglones y columnas que son linealmente independientes (¡y este número es el mismo en cada caso!).

Dado que los vectores renglón de A son los vectores columna de A^T, el Teorema 3.24 tiene el siguiente corolario inmediato.

Teorema 3.25 Para cualquier matriz A,

$$\operatorname{rank}(A^T) = \operatorname{rank}(A)$$

Demostración Se tiene

$$\begin{aligned}\operatorname{rank}(A^T) &= \dim(\operatorname{col}(A^T))\\ &= \dim(\text{renglón}(A))\\ &= \operatorname{rank} A\end{aligned}$$

Definición La ***nulidad*** de una matriz A es la dimensión de su espacio nulo y se denota por nulidad(A).

En otras palabras, nulidad(A) es la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$, que es igual que el número de variables libres en la solución. Ahora puede regresar al teorema del rank (Teorema 2.2) y parafrasearlo en términos de las nuevas definiciones.

Teorema 3.26 **El Teorema del rank**

Si A es una matriz de $m \times n$, entonces

$$\text{rank}(A) + \text{nulidad}(A) = n$$

Demostración Sea R la forma escalonada reducida por renglones de A, y suponga que rank(A) $= r$. Entonces R tiene r 1 pivote, de modo que existen r variables pivote y $n - r$ variables libres en la solución de $A\mathbf{x} = \mathbf{0}$. Puesto que dim(nul(A)) $= n - r$, se tiene

$$\begin{aligned}\text{rank}(A) + \text{nulidad}(A) &= r + (n - r)\\ &= n\end{aligned}$$

Con frecuencia, cuando se necesita conocer la nulidad de una matriz, no es necesario conocer la solución real de $A\mathbf{x} = \mathbf{0}$. El teorema del rank es extremadamente útil en tales situaciones, como ilustra el siguiente ejemplo.

Ejemplo 3.51

Encuentre la nulidad de cada una de las siguientes matrices:

$$M = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 4 & 7 \\ 3 & 6 \end{bmatrix} \quad \text{y}$$

$$N = \begin{bmatrix} 2 & 1 & -2 & -1 \\ 4 & 4 & -3 & 1 \\ 2 & 7 & 1 & 8 \end{bmatrix}$$

Solución Dado que las dos columnas de M claramente son linealmente independientes, rank(M) $= 2$. Por tanto, por el teorema del rank, nulidad(M) $= 2 -$ rank(M) $= 2 - 2 = 0$.

No hay dependencia obvia entre los renglones o columnas de N, así que se aplican operaciones con renglones para reducirla a

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

La matriz se redujo lo suficiente (aquí no es necesaria la forma escalonada *reducida* por renglones, pues no se busca una base para el espacio nulo). Se ve que sólo existen dos renglones distintos de cero, de modo que rank(N) $= 2$. En consecuencia, nulidad (N) $=$ $4 -$ rank(N) $= 4 - 2 = 2$.

Los resultados de estas acciones permiten extender el teorema fundamental de las matrices invertibles (Teorema 3.12).

Teorema 3.27 **El teorema fundamental de las matrices invertibles: versión 2**

Sea A una matriz de $n \times n$. Los siguientes enunciados son equivalentes:

a. A es invertible.
b. $A\mathbf{x} = \mathbf{b}$ tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial.
d. La forma escalonada reducida por renglones de A es I_n.
e. A es un producto de matrices elementales.
f. $\text{rank}(A) = n$
g. $\text{nulidad}(A) = 0$
h. Los vectores columna de A son linealmente independientes.
i. Los vectores columna de A generan $\mathbb{R}^n$.
j. Los vectores columna de A forman una base para $\mathbb{R}^n$.
k. Los vectores renglón de A son linealmente independientes.
l. Los vectores renglón de A generan $\mathbb{R}^n$.
m. Los vectores renglón de A forman una base para $\mathbb{R}^n$.

La nulidad de una matriz la definió en 1884 James Joseph Sylvester (1814–1887), quien estaba interesado en las *invariantes*: propiedades de las matrices que no cambian en ciertos tipos de transformaciones. Originario de Inglaterra, Sylvester se convirtió en el segundo presidente de la Sociedad Matemática de Londres. En 1878, mientras impartía clases en la Johns Hopkins University en Baltimore, fundó el *American Journal of Mathematics*, la primera revista matemática en Estados Unidos.

Demostración Ya se estableció la equivalencia desde (a) hasta (e). Falta por demostrar que los enunciados (f) al (m) son equivalentes a los primeros cinco enunciados.

(f) $\Leftrightarrow$ (g) Dado que $\text{rank}(A) + \text{nulidad}(A) = n$ cuando A es una matriz de $n \times n$, se tiene del teorema del rank que $\text{rank}(A) = n$ si y sólo si $\text{nulidad}(A) = 0$.

(f) $\Rightarrow$ (d) $\Rightarrow$ (c) $\Rightarrow$ (h) Si $\text{rank}(A) = n$, entonces la forma escalonada reducida por renglón de A tiene n 1 pivotes, y por tanto es I_n. De (d) $\Rightarrow$ (c) se sabe que $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial, lo que implica que los vectores columna de A son linealmente independientes, pues $A\mathbf{x}$ es justo una combinación lineal de los vectores columna de A.

(h) $\Rightarrow$ (i) Si los vectores columna de A son linealmente independientes, entonces $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial. Por tanto, por (c) $\Rightarrow$ (b), $A\mathbf{x} = \mathbf{b}$, tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$. Esto significa que todo vector $\mathbf{b}$ en $\mathbb{R}^n$ puede escribirse como una combinación lineal de los vectores columna de A, lo que establece (i)

(i) $\Rightarrow$ (j) Si los vectores columna de A generan $\mathbb{R}^n$, entonces $\text{col}(A) = \mathbb{R}^n$ por definición, de modo que $\text{rank}(A) = \dim(\text{col}(A)) = n$. Esto es (f), y ya se estableció que (f) $\Rightarrow$ (h). Se concluye que los vectores columna de A son linealmente independientes y por tanto forman una base para $\mathbb{R}^n$, pues, por suposición, también generan $\mathbb{R}^n$.

(j) $\Rightarrow$ (f) Si los vectores columna de A forman una base para $\mathbb{R}^n$, entonces, en particular, son linealmente independientes. Se tiene que la forma escalonada reducida por renglones de A contiene n 1 pivotes, y por tanto $\text{rank}(A) = n$.

El análisis anterior demuestra que (f) $\Rightarrow$ (d) $\Rightarrow$ (c) $\Rightarrow$ (h) $\Rightarrow$ (i) $\Rightarrow$ (j) $\Rightarrow$ (f) $\Leftrightarrow$ (g). Ahora recuerde que, por el Teorema 3.25, $\text{rank}(A^T) = \text{rank}(A)$, de modo que lo recién demostrado da los resultados correspondientes acerca de los vectores columna de A^T. Estos son entonces resultados acerca de los vectores *renglón* de A, lo que incluye en (k), (l) y (m) a la red de equivalencias y completa la demostración.

Los teoremas como el teorema fundamental no son simplemente de interés teórico. También son tremendos dispositivos que ahorran trabajo. El teorema fundamental ya permitió recortar a la mitad el trabajo necesario para comprobar que dos matrices cuadradas son inversas. También simplifica la tarea de demostrar que ciertos conjuntos de vectores son bases para $\mathbb{R}^n$. De hecho, cuando se tiene un conjunto de n vectores en $\mathbb{R}^n$, dicho conjunto será una base para $\mathbb{R}^n$ si es verdadera *alguna* de las propiedades necesarias de independencia lineal o de conjunto generador. El siguiente ejemplo muestra cuán fáciles pueden ser los cálculos.

Ejemplo 3.52

Demuestre que los vectores

$$\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}4\\9\\7\end{bmatrix}$$

forman una base para $\mathbb{R}^3$.

Solución De acuerdo con el Teorema Fundamental, los vectores formarán una base para $\mathbb{R}^3$ si y sólo si una matriz con estos vectores como sus columnas (o renglones) tiene rank 3. Se realizan sólo suficientes operaciones con renglones para determinar esto:

$$A = \begin{bmatrix}1 & -1 & 4\\2 & 0 & 9\\3 & 1 & 7\end{bmatrix} \longrightarrow \begin{bmatrix}1 & -1 & 4\\0 & 2 & 1\\0 & 0 & -7\end{bmatrix}$$

Se ve que A tiene rank 3, de modo que los vectores dados son una base para $\mathbb{R}^3$, por la equivalencia de (f) y (j).

El siguiente teorema es una aplicación tanto del teorema del rank como del teorema fundamental. Este resultado se requerirá en los capítulos 5 y 7.

Teorema 3.28

Sea A una matriz de $m \times n$. Entonces:

a. $\operatorname{rank}(A^TA) = \operatorname{rank}(A)$
b. La matriz A^TA de $n \times n$ es invertible si y sólo si $\operatorname{rank}(A) = n$.

Demostración

(a) Dado que A^TA es de $n \times n$, tiene el mismo número de columnas que A. El teorema del rank dice entonces que

$$\operatorname{rank}(A) + \operatorname{nulidad}(A) = n = \operatorname{rank}(A^TA) + \operatorname{nulidad}(A^TA)$$

Por tanto, para demostrar que $\operatorname{rank}(A) = \operatorname{rank}(A^TA)$, es suficiente demostrar que nulidad(A) = nulidad(A^TA). Se hará así al establecer que los espacios nulos de A y A^TA son iguales.

Para este fin, sea $\mathbf{x}$ en nul(A), de modo que $A\mathbf{x} = \mathbf{0}$. Entonces $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$, y por tanto $\mathbf{x}$ está en nul(A^TA). A la inversa, sea $\mathbf{x}$ en nul(A^TA). Entonces $A^TA\mathbf{x} = \mathbf{0}$, de modo que $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$. Pero entonces

$$(A\mathbf{x}) \cdot (A\mathbf{x}) = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x} = 0$$

y en consecuencia $A\mathbf{x} = 0$, por el Teorema 1.2(d). Por tanto, $\mathbf{x}$ está en nul(A), de modo que nul(A) = nul(A^TA), como se requería.

(b) Por el teorema fundamental, la matriz A^TA de $n \times n$ es invertible si y sólo si $\operatorname{rank}(A^TA) = n$. Pero, por (a), esto se cumple si y sólo si $\operatorname{rank}(A) = n$.

Coordenadas

Ahora regrese a una de las preguntas planteadas al comienzo de esta sección: ¿cómo se verían los vectores en $\mathbb{R}^3$ que existen en un plano que pasa por el origen? ¿Son bidimensionales o tridimensionales? Las nociones de base y dimensión ayudarán a clarificar las cosas.

Un plano que pasa por el origen es un subespacio bidimensional de $\mathbb{R}^3$, con cualquier conjunto de dos vectores dirección que funcionan como base. Los vectores base ubican ejes coordenados en el plano-subespacio y a su vez permiten ver el plano como una "copia" de $\mathbb{R}^2$. Antes de ilustrar este enfoque, se demostrará un teorema que garantiza que las "coordenadas" surgidas de esta forma son únicas.

Teorema 3.29 Sea S un subespacio de $\mathbb{R}^n$ y sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base para S. Para todo vector $\mathbf{v}$ en S, existe exactamente una forma de escribir $\mathbf{v}$ como una combinación lineal de los vectores base en $\mathcal{B}$:

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

Demostración Dado que $\mathcal{B}$ es una base, genera a S, de modo que $\mathbf{v}$ puede escribirse en *al menos una* forma como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Sea una de dichas combinaciones lineales

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

Su tarea es demostrar que ésta es la *única* forma de escribir $\mathbf{v}$ como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Para este fin, suponga que también se tiene

$$\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$$

Entonces $$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$$

Al reordenar (con las propiedades del álgebra vectorial), se obtiene

$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \cdots + (c_k - d_k)\mathbf{v}_k = \mathbf{0}$$

Dado que $\mathcal{B}$ es una base, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ son linealmente independientes. Por tanto,

$$(c_1 - d_1) = (c_2 - d_2) = \cdots = (c_k - d_k) = 0$$

En otras palabras, $c_1 = d_1, c_2 = d_2, \ldots, c_k = d_k$, y las dos combinaciones lineales en realidad son iguales. Por ende, existe exactamente una forma de escribir $\mathbf{v}$ como una combinación lineal de los vectores base en $\mathcal{B}$.

Definición Sea S un subespacio de $\mathbb{R}^n$ y sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base para S. Sea $\mathbf{v}$ un vector en S, y escriba $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$. Entonces $c_1, c_2, \ldots, c_k$ se llaman las ***coordenadas de* v *con respecto a*** $\mathcal{B}$, y el vector columna

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$$

se llama ***vector coordenada de* v *con respecto a*** $\mathcal{B}$.

Ejemplo 3.53 Sea $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base estándar de $\mathbb{R}^3$. Encuentre el vector coordenada de

$$\mathbf{v} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$$

con respecto a $\mathcal{E}$.

Solución Dado que $\mathbf{v} = 2\mathbf{e}_1 + 7\mathbf{e}_2 + 4\mathbf{e}_3$,

$$[\mathbf{v}]_{\mathcal{E}} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$$

Debe ser claro que el vector coordenada de todo vector (columna) en $\mathbb{R}^n$ con respecto a la base estándar es justo el vector mismo.

Ejemplo 3.54

En el ejemplo 3.44 se vio que $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}$ son tres vectores en el mismo subespacio (plano que pasa por el origen) S de $\mathbb{R}^3$ y que $\mathcal{B} = \{\mathbf{u}, \mathbf{v}\}$ es una base para S. Dado que $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$, se tiene

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

Vea la figura 3.3.

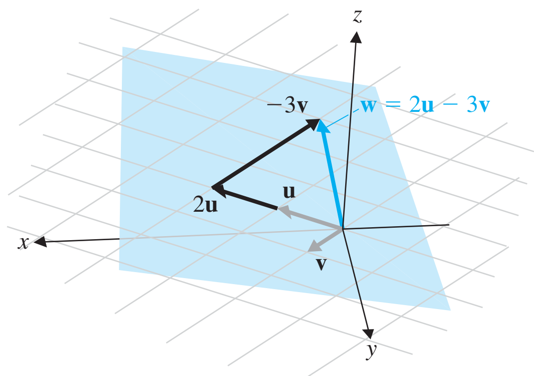


Figura 3.3
Las coordenadas de un vector con respecto a una base

Ejercicios 3.5

En los ejercicios 1-4, sea S la colección de vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ *en* $\mathbb{R}^2$ *que satisfacen la propiedad dada. En cada caso, demuestre que S forma un subespacio de* $\mathbb{R}^2$ *u ofrezca un contraejemplo para demostrar que no lo hace.*

1. $x = 0$ **2.** $x \geq 0, y \geq 0$

3. $y = 2x$ **4.** $xy \geq 0$

En los ejercicios 5-8, sea S la colección de vectores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *en* $\mathbb{R}^3$ *que satisfacen la propiedad dada. En cada caso, demuestre que S forma un subespacio de* $\mathbb{R}^3$ *u ofrezca un contraejemplo para demostrar que no lo hace.*

5. $x = y = z$ **6.** $z = 2x, y = 0$

7. $x - y + z = 1$ **8.** $|x - y| = |y - z|$

9. Demuestre que toda recta que pasa por el origen en $\mathbb{R}^3$ es un subespacio de $\mathbb{R}^3$.

10. Suponga que S consiste de todos los puntos en $\mathbb{R}^2$ que están sobre el eje x o el eje y (o ambos). (S se llama la *unión* de los dos ejes.) ¿S es un subespacio de $\mathbb{R}^2$? ¿Por qué sí o por qué no?

En los ejercicios 11 y 12, determine si **b** *está en col(A) y si* **w** *está en renglón(A), como en el ejemplo 3.41.*

11. $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $\mathbf{w} = [-1 \quad 1 \quad 1]$

12. $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & 0 \\ 3 & -1 & -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{w} = [1 \quad -3 \quad -3]$

13. En el ejercicio 11, determine si **w** está en renglón(A), usando el método descrito en el comentario que sigue al ejemplo 3.41.

14. En el ejercicio 12, determine si **w** está en renglón(A), usando el método descrito en el comentario que sigue al ejemplo 3.41.

15. Si A es la matriz del ejercicio 11, ¿$\mathbf{v} = \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$ está en nul(A)?

16. Si A es la matriz del ejercicio 12, ¿$\mathbf{v} = \begin{bmatrix} 7 \\ -1 \\ 2 \end{bmatrix}$ está en nul(A)?

En los ejercicios 17-20, proporcione bases para renglón(A), col(A) y nul(A).

17. $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ **18.** $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & 1 \\ 1 & -1 & -2 \end{bmatrix}$

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

20. $A = \begin{bmatrix} 2 & -4 & 0 & 2 & 1 \\ -1 & 2 & 1 & 2 & 3 \\ 1 & -2 & 1 & 4 & 4 \end{bmatrix}$

En los ejercicios 21-24, encuentre bases para renglón(A) y col(A) en los ejercicios dados usando A^T.

21. Ejercicio 17 **22.** Ejercicio 18

23. Ejercicio 19 **24.** Ejercicio 20

25. Explique cuidadosamente por qué sus respuestas a los ejercicios 17 y 21 son ambas correctas, aun cuando parezca haber diferencias.

26. Explique cuidadosamente por qué sus respuestas a los ejercicios 18 y 22 son ambas correctas, aun cuando parezca haber diferencias.

En los ejercicios 27-30, encuentre una base para el generador de los vectores dados.

27. $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ **28.** $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

29. $[2 \quad -3 \quad 1], [1 \quad -1 \quad 0], [4 \quad -4 \quad 1]$

30. $[3 \quad 1 \quad -1 \quad 0], [0 \quad -1 \quad 2 \quad -1], [4 \quad 3 \quad 8 \quad 3]$

Para los ejercicios 31 y 32, encuentre bases para los generadores de los vectores en los ejercicios dados de entre los vectores mismos.

31. Ejercicio 29 **32.** Ejercicio 30

33. Demuestre que si R es una matriz en forma escalonada, entonces una base para renglón(R) consiste de los renglones distintos de cero de R.

34. Demuestre que si las columnas de A son linealmente independientes, entonces deben formar una base para col(A).

Para los ejercicios 35-38, proporcione el rank y la nulidad de las matrices en los ejercicios dados.

35. Ejercicio 17

36. Ejercicio 18

37. Ejercicio 19

38. Ejercicio 20

39. Si A es una matriz de 3×5, explique por qué las columnas de A deben ser linealmente dependientes.

40. Si A es una matriz de 4×2, explique por qué los renglones de A deben ser linealmente dependientes.

41. Si A es una matriz de 3×5, ¿cuáles son los posibles valores de nulidad(A)?

42. Si A es una matriz de 4×2, ¿cuáles son los posibles valores de nulidad(A)?

En los ejercicios 43 y 44, encuentre todos los posibles valores de rank(A) cuando a varía.

43. $A = \begin{bmatrix} 1 & 2 & a \\ -2 & 4a & 2 \\ a & -2 & 1 \end{bmatrix}$ **44.** $A = \begin{bmatrix} a & 2 & -1 \\ 3 & 3 & -2 \\ -2 & -1 & a \end{bmatrix}$

Responda los ejercicios 45-48 al considerar la matriz con los vectores dados como sus columnas.

45. ¿$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ forma una base para $\mathbb{R}^3$?

46. ¿$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ forma una base para $\mathbb{R}^3$?

47. ¿$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ forma una base para $\mathbb{R}^4$?

48. ¿ $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\-1\end{bmatrix}$ forma una base para $\mathbb{R}^4$?

49. ¿ $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}$ forma una base para $\mathbb{Z}_2^3$?

50. ¿ $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}$ forma una base para $\mathbb{Z}_3^3$?

En los ejercicios 51 y 52, demuestre que $\mathbf{w}$ *está en gen*$(\mathcal{B})$ *y encuentre el vector coordenada* $[\mathbf{w}]_{\mathcal{B}}$.

51. $\mathcal{B} = \left\{\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-1\end{bmatrix}\right\}, \mathbf{w} = \begin{bmatrix}1\\6\\2\end{bmatrix}$

52. $\mathcal{B} = \left\{\begin{bmatrix}3\\1\\4\end{bmatrix}, \begin{bmatrix}5\\1\\6\end{bmatrix}\right\}, \mathbf{w} = \begin{bmatrix}1\\3\\4\end{bmatrix}$

En los ejercicios 53-56, calcule el rank y la nulidad de las matrices dadas sobre el $\mathbb{Z}_p$ *indicado.*

53. $\begin{bmatrix}1&1&0\\0&1&1\\1&0&1\end{bmatrix}$ sobre $\mathbb{Z}_2$

54. $\begin{bmatrix}1&1&2\\2&1&2\\2&0&0\end{bmatrix}$ sobre $\mathbb{Z}_3$

55. $\begin{bmatrix}1&3&1&4\\2&3&0&1\\1&0&4&0\end{bmatrix}$ sobre $\mathbb{Z}_5$

56. $\begin{bmatrix}2&4&0&0&1\\6&3&5&1&0\\1&0&2&2&5\\1&1&1&1&1\end{bmatrix}$ sobre $\mathbb{Z}_7$

57. Si A es de $m \times n$, demuestre que todo vector en nul(A) es ortogonal a todo vector en renglón(A).

58. Si A y B son matrices de $n \times n$ de rank n, demuestre que AB tiene rank n.

59. **(a)** Demuestre que rank$(AB) \leq$ rank(B). [*Sugerencia:* revise el ejercicio 29 de la sección 3.1.]
(b) Ofrezca un ejemplo en el que rank$(AB) <$ rank(B).

60. **(a)** Demuestre que rank$(AB) \leq$ rank(A). [*Sugerencia:* Revise el ejercicio 30 de la sección 3.1 o use transpuestas y el ejercicio 59(a).]
(b) Ofrezca un ejemplo en el que rank$(AB) =$ rank(A).

61. **(a)** Demuestre que si U es invertible, entonces rank$(UA) =$ rank(A). [*Sugerencia:* $A = U^{-1}(UA)$.]
(b) Demuestre que si V es invertible, entonces rank$(AV) =$ rank(A).

62. Demuestre que una matriz A de $m \times n$ tiene rank 1 si y sólo si A puede escribirse como el producto exterior $\mathbf{u}\mathbf{v}^T$ de un vector $\mathbf{u}$ en $\mathbb{R}^m$ y $\mathbf{v}$ en $\mathbb{R}^n$.

63. Si una matriz A de $m \times n$ tiene rank r, demuestre que A puede escribirse como la suma de r matrices, cada una de las cuales tiene rank 1. [*Sugerencia:* encuentre una forma de usar el ejercicio 62.]

64. Demuestre que, para matrices A y B de $m \times n$, rank$(A + B) \leq$ rank(A) + rank(B).

65. Sea A una matriz $n \times n$ tal que $A^2 = O$. Demuestre que rank$(A) \leq n/2$. [*Sugerencia:* demuestre que col$(A) \subseteq$ nul(A) y use el teorema del rank.]

66. Sea A una matriz antisimétrica $n \times n$. (Vea la página 168.)
(a) Demuestre que $\mathbf{x}^T A\mathbf{x} = 0$ para todo $\mathbf{x}$ en $\mathbb{R}^n$.
(b) Demuestre que $I + A$ es invertible. [*Sugerencia:* demuestre que nul$(I + A) = \{\mathbf{0}\}$.]

3.6 Introducción a las transformaciones lineales

En esta sección comienza la exploración de uno de los temas de la introducción a este capítulo. Ahí se vio que la matrices pueden usarse para transformar vectores y actuar como un tipo de "función" de la forma $\mathbf{w} = T(\mathbf{v})$, donde la variable independiente $\mathbf{v}$ y la variable dependiente $\mathbf{w}$ son vectores. Ahora esta noción se hará más precisa y se observarán varios ejemplos de tales transformaciones matriciales, lo que conduce al concepto de *transformación lineal*, una poderosa idea que se encontrará repetidamente a partir de ahora.

Comience por recordar algunos de los conceptos básicos asociados con las funciones. Se familiarizará con más de estas ideas en otros cursos en los que encontrará funciones de la forma $f: \mathbb{R} \to \mathbb{R}$ [como $f(x) = x^2$] que transforman números reales en números reales. Lo que es novedoso aquí es que están involucrados vectores y que sólo se está interesado en funciones que son "compatibles" con las operaciones vectoriales de suma y multiplicación escalar.

Considere un ejemplo. Sean

$$A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Entonces

$$A\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$

Esto demuestra que A transforma $\mathbf{v}$ en $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$.

Esta transformación puede describirse de manera más general. La ecuación matricial

$$\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

proporciona una fórmula que muestra cómo A transforma un vector arbitrario $\begin{bmatrix} x \\ y \end{bmatrix}$ en $\mathbb{R}^2$ en el vector $\begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$ en $\mathbb{R}^3$. Esta transformación se denota T_A y se escribe

$$T_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

(Aunque técnicamente desordenada, omitir los paréntesis en definiciones como esta es una convención común que ahorra cierta escritura. La descripción de T_A se convierte en

$$T_A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

con esta convención.)

Con este ejemplo en mente, ahora considere alguna terminología. Una ***transformación*** (o ***mapeo*** o ***función***) T de $\mathbb{R}^n$ en $\mathbb{R}^m$ es una regla que asigna a cada vector $\mathbf{v}$ en $\mathbb{R}^n$ un vector único $T(\mathbf{v})$ en $\mathbb{R}^m$. El ***dominio*** de T es $\mathbb{R}^n$, y el ***codominio*** de T es $\mathbb{R}^m$. Esto se indica al escribir $T: \mathbb{R}^n \to \mathbb{R}^m$. Para un vector $\mathbf{v}$ en el dominio de T, el vector $T(\mathbf{v})$ en el codominio se conoce como la ***imagen*** de $\mathbf{v}$ bajo (la acción de) T. Al conjunto de todas las posibles imágenes $T(\mathbf{v})$ (conforme $\mathbf{v}$ varía a lo largo del dominio de T) se le llama ***rango*** de T.

En el ejemplo, el dominio de T_A es $\mathbb{R}^2$ y su codominio es $\mathbb{R}^3$, así que se escribe $T_A: \mathbb{R}^2 \to \mathbb{R}^3$. La imagen de $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ es $\mathbf{w} = T_A(\mathbf{v}) = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$. Cuál es el rango de T_A?

Consiste de todos los vectores en el codominio $\mathbb{R}^3$ que son de la forma

$$T_A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

que describe al conjunto de todas las combinaciones lineales de los vectores columna $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$ de A. En otras palabras, ¡el rango de T es el espacio columna de A! (Más adelante se hablará más acerca de esto; por el momento, simplemente se indicará como una comentario interesante.) Geométricamente, esto demuestra que el rango de T_A es el plano que pasa por el origen en $\mathbb{R}^3$ con vectores directores dados por los vectores columna de A. Note que el rango de T_A es estrictamente menor que el codominio de T_A.

Transformaciones lineales

El ejemplo T_A anterior es un caso especial de un tipo más general de transformación llamada *transformación lineal*. En el capítulo 6 se considerará la definición general, pero la esencia de ella es que se trata de transformaciones que "conservan" las operaciones vectoriales de suma y multiplicación por un escalar.

Definición Una transformación $T : \mathbb{R}^n \to \mathbb{R}^m$ se llama ***transformación lineal*** si

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para todo $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ y
2. $T(c\mathbf{v}) = cT(\mathbf{v})$ para todo $\mathbf{v}$ en $\mathbb{R}^n$ y todo escalar c.

Ejemplo 3.55

Considere una vez más la transformación $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

Compruebe que T es una transformación lineal. Para verificar (1), sea

$$\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

Entonces

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ 2(x_1 + x_2) - (y_1 + y_2) \\ 3(x_1 + x_2) + 4(y_1 + y_2) \end{bmatrix} \\ &= \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 2x_2 - y_1 - y_2 \\ 3x_1 + 3x_2 + 4y_1 + 4y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ (2x_1 - y_1) + (2x_2 - y_2) \\ (3x_1 + 4y_1) + (3x_2 + 4y_2) \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ 2x_1 - y_1 \\ 3x_1 + 4y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ 2x_2 - y_2 \\ 3x_2 + 4y_2 \end{bmatrix} = T\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + T\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = T(\mathbf{u}) + T(\mathbf{v}) \end{aligned}$$

Para demostrar (2), sea $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ y sea c un escalar. Entonces

$$\begin{aligned} T(c\mathbf{v}) &= T\left(c\begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} cx \\ cy \end{bmatrix}\right) \\ &= \begin{bmatrix} cx \\ 2(cx) - (cy) \\ 3(cx) + 4(cy) \end{bmatrix} = \begin{bmatrix} cx \\ c(2x - y) \\ c(3x + 4y) \end{bmatrix} \\ &= c\begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = cT\begin{bmatrix} x \\ y \end{bmatrix} = cT(\mathbf{v}) \end{aligned}$$

Por tanto, T es una transformación lineal.

Comentario La definición de una transformación lineal puede simplificarse al combinar (1) y (2) como se muestra a continuación.

$T: \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal si

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) \quad \text{para todo } \mathbf{v}_1, \mathbf{v}_2 \text{ en } \mathbb{R}^n \text{ y escalares } c_1, c_2$$

En el ejercicio 53 se le pedirá demostrar que el enunciado anterior es equivalente a la definición original. En la práctica, esta formulación equivalente puede ahorrar alguna escritura, ¡inténtelo!

Aunque la transformación lineal T en el ejemplo 3.55 originalmente surgió como una transformación *matricial* T_A, es asunto sencillo recuperar la matriz A a partir de la definición de T dada en el ejemplo. Observe que

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

de modo que $T = T_A$, donde $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$. (Note que, cuando las variables x y y se alinean, la matriz A es justo su matriz de coeficientes.)

Es importante reconocer que una transformación es una transformación matricial, pues, como muestra el siguiente teorema, todas las transformaciones matriciales son transformaciones lineales.

Teorema 3.30 Sea A una matriz de $m \times n$. Entonces la transformación matricial $T_A: \mathbb{R}^n \to \mathbb{R}^m$ definida por

$$T_A(\mathbf{x}) = A\mathbf{x} \qquad (\text{para } \mathbf{x} \text{ en } \mathbb{R}^n)$$

es una transformación lineal.

Demostración Sean **u** y **v** vectores en $\mathbb{R}^n$ y sea c un escalar. Entonces

$$T_A(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T_A(\mathbf{u}) + T_A(\mathbf{v})$$

y
$$T_A(c\mathbf{v}) = A(c\mathbf{v}) = c(A\mathbf{v}) = cT_A(\mathbf{v})$$

Por tanto, T_A es una transformación lineal.

Ejemplo 3.56

Sea $F: \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que envía cada punto a su reflexión en el eje x. Demuestre que F es una transformación lineal.

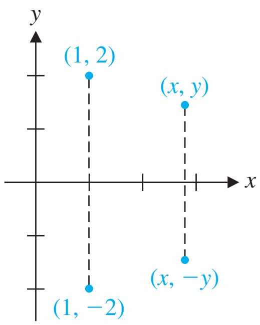


Figura 3.4
Reflexión en el eje x

Solución En la figura 3.4 es claro que F envía el punto (x, y) al punto $(x, -y)$. Por tanto, puede escribir

$$F\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

Podría proceder para comprobar que F es lineal, como en el ejemplo 3.55 (¡esto incluso es más sencillo de comprobar!), pero es más rápido observar que

$$\begin{bmatrix} x \\ -y \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Por tanto, $F\begin{bmatrix} x \\ y \end{bmatrix} = A\begin{bmatrix} x \\ y \end{bmatrix}$, donde $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, de modo que F es una transformación matricial. Ahora se tiene, por el Teorema 3.30, que F es una transformación lineal.

Ejemplo 3.57

Sea $R: \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que rota cada punto 90° en sentido contrario al de las manecillas del reloj en torno al origen. Demuestre que R es una transformación lineal.

Solución Como muestra la figura 3.5, R envía el punto (x, y) al punto $(-y, x)$. Por tanto, se tiene

$$R\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} = x\begin{bmatrix} 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

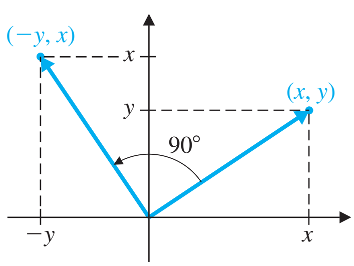


Figura 3.5
Una rotación de 90°

En consecuencia, R es una transformación matricial y por lo tanto es lineal.

Observe que, si multiplica una matriz por vectores base estándar, obtiene las columnas de la matriz. Por ejemplo,

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \\ e \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \\ f \end{bmatrix}$$

Puede usar esta observación para demostrar que *toda* transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ surge como una transformación matricial.

Teorema 3.31 Sea $T:\mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces T es una transformación matricial. De manera más específica, $T = T_A$, donde A es la matriz de $m \times n$

$$A = [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots \cdots \vdots T(\mathbf{e}_n)]$$

Demostración Sean $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ los vectores base estándar en $\mathbb{R}^n$ y sea $\mathbf{x}$ un vector en $\mathbb{R}^n$. Puede escribir $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$ (donde las x_i son los componentes de $\mathbf{x}$). También se sabe que $T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)$ son vectores (columna) en $\mathbb{R}^m$. Sea $A = [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots \cdots \vdots T(\mathbf{e}_n)]$ la matriz de $m \times n$ con dichos vectores como sus columnas. Entonces

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots \cdots \vdots T(\mathbf{e}_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{x} \end{aligned}$$

como se requiere.

La matriz A en el Teorema 3.31 se llama ***matriz estándar de la transformación lineal*** T.

Ejemplo 3.58 Demuestre que una rotación en torno al origen, a través de un ángulo θ, define una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ y encuentre su matriz estándar.

Solución Sea R_θ la rotación. Se dará un argumento geométrico para establecer el hecho de que R_θ es lineal. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en $\mathbb{R}^2$. Si no son paralelos, entonces la figura 3.6(a) muestra la regla del paralelogramo que determina $\mathbf{u} + \mathbf{v}$. Si ahora se aplica R_θ, todo el paralelogramo rota un ángulo θ, como se muestra en la figura 3.6(b). Pero la diagonal de este paralelogramo debe ser $R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$, de nuevo, por la regla del paralelogramo. Por tanto, $R_\theta(\mathbf{u} + \mathbf{v}) = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$. (¿Qué ocurre si $\mathbf{u}$ y $\mathbf{v}$ son paralelos?)

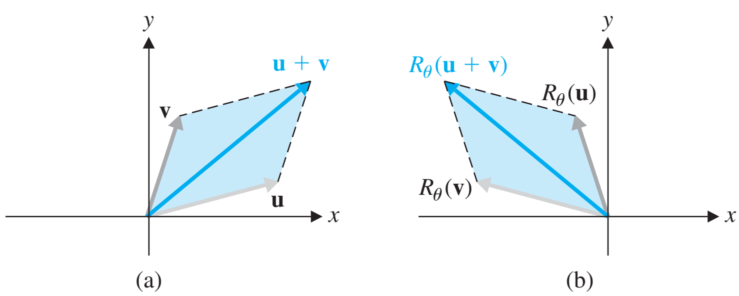


Figura 3.6

De igual modo, si se aplica R_θ a $\mathbf{v}$ y $c\mathbf{v}$, se obtiene $R_\theta(\mathbf{v})$ and $R_\theta(c\mathbf{v})$, como se muestra en la figura 3.7. Pero, dado que la rotación no afecta las longitudes, entonces se debe tener $R_\theta(c\mathbf{v}) = cR_\theta(\mathbf{v})$, como se requiere. (Dibuje diagramas para el caso $0 < c < 1$, $-1 < c < 0$ y $c < -1$.)

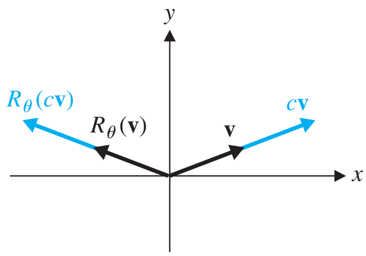


Figura 3.7

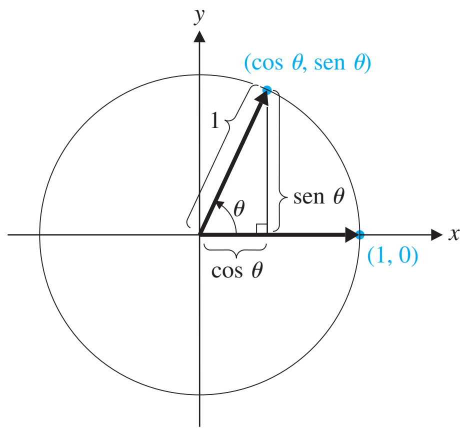


Figura 3.8
$R_\theta(\mathbf{e}_1)$

Por tanto, R_θ es una transformación lineal. De acuerdo con el Teorema 3.31, se puede encontrar su matriz al determinar su efecto sobre los vectores base estándar $\mathbf{e}_1$ y $\mathbf{e}_2$ de $\mathbb{R}^2$.

Ahora, como muestra la figura 3.8, $R_\theta\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}\cos\theta\\ \operatorname{sen}\theta\end{bmatrix}$.

Puede encontrar $R_\theta\begin{bmatrix}0\\1\end{bmatrix}$ de igual forma, pero es más rápido observar que $R_\theta\begin{bmatrix}0\\1\end{bmatrix}$ debe ser perpendicular (en sentido contrario al de las manecillas del reloj) a $R_\theta\begin{bmatrix}1\\0\end{bmatrix}$ y por tanto, por el ejemplo 3.57, $R_\theta\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}-\operatorname{sen}\theta\\ \cos\theta\end{bmatrix}$ (figura 3.9).

En consecuencia, la matriz estándar de R_θ es $\begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}$.

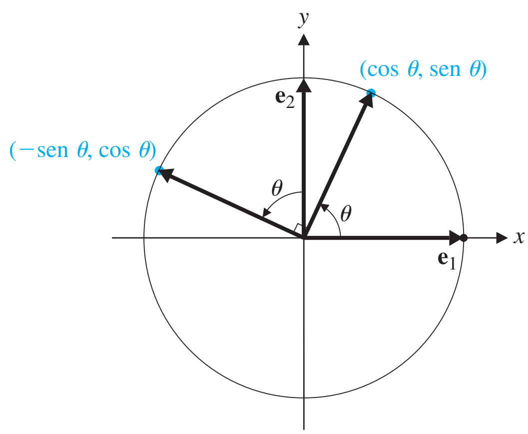


Figura 3.9
$R_\theta(\mathbf{e}_2)$

Ahora puede usar el resultado del ejemplo 3.58 para calcular el efecto de cualquier rotación. Por ejemplo, suponga que quiere rotar 60° el punto $(2, -1)$ en torno al origen. (La convención es que un ángulo positivo corresponde a una rotación en sentido con-

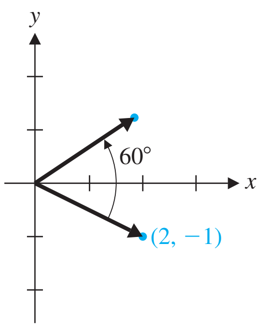


Figura 3.10
Una rotación de 60°

trario al de las manecillas del reloj, mientras que un ángulo negativo es en el sentido de las manecillas del reloj.) Dado que $\cos 60° = 1/2$ y $\text{sen } 60° = \sqrt{3}/2$, se calcula

$$R_{60}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} \cos 60° & -\text{sen } 60° \\ \text{sen } 60° & \cos 60° \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} (2+\sqrt{3})/2 \\ (2\sqrt{3}-1)/2 \end{bmatrix}$$

Por tanto, la imagen del punto $(2, -1)$ bajo esta rotación es el punto $((2+\sqrt{3})/2, (2\sqrt{3}-1)/2) \approx (1.87, 1.23)$, como se muestra en la figura 3.10.

Ejemplo 3.59

(a) Demuestre que la transformación $P : \mathbb{R}^2 \to \mathbb{R}^2$ que proyecta un punto sobre el eje x es una transformación lineal y encuentre su matriz estándar.
(b) De manera más general, si ℓ es una recta que pasa por el origen en $\mathbb{R}^2$, demuestre que la transformación $P_\ell : \mathbb{R}^2 \to \mathbb{R}^2$ que proyecta un punto sobre ℓ es una transformación lineal y encuentre su matriz estándar.

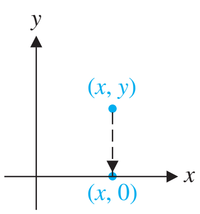


Figura 3.11
Una proyección

Solución (a) Como muestra la figura 3.11, P envía el punto (x, y) al punto $(x, 0)$. Por tanto,

$$P\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Se tiene que P es una transformación matricial (y en consecuencia una transformación lineal) con matriz estándar $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

(b) Sea la recta ℓ que tiene vector director $\mathbf{d}$ y sea $\mathbf{v}$ un vector arbitrario. Entonces P_ℓ está dado por $\text{proy}_{\mathbf{d}}(\mathbf{v})$, la proyección de $\mathbf{v}$ sobre $\mathbf{d}$, que de la sección 1.2 recordará tiene la fórmula

$$\text{proy}_{\mathbf{d}}(\mathbf{v}) = \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d}$$

Por tanto, para demostrar que P_ℓ es lineal, se procede del modo siguiente:

$$\begin{aligned} P_\ell(\mathbf{u}+\mathbf{v}) &= \left(\frac{\mathbf{d}\cdot(\mathbf{u}+\mathbf{v})}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}+\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}}{\mathbf{d}\cdot\mathbf{d}} + \frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} + \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = P_\ell(\mathbf{u}) + P_\ell(\mathbf{v}) \end{aligned}$$

De igual modo, $P_\ell(c\mathbf{v}) = cP_\ell(\mathbf{v})$ para cualquier escalar c (ejercicio 52). En consecuencia, P_ℓ es una transformación lineal.

Para encontrar la matriz estándar de P_ℓ, aplique el Teorema 3.31. Si se hace $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$, entonces

$$P_\ell(\mathbf{e}_1) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_1}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_1}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 \\ d_1d_2 \end{bmatrix}$$

y
$$P_\ell(\mathbf{e}_2) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_2}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_2}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1d_2 \\ d_2^2 \end{bmatrix}$$

Por tanto, la matriz estándar de la proyección es

$$A = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 & d_1d_2 \\ d_1d_2 & d_2^2 \end{bmatrix} = \begin{bmatrix} d_1^2/(d_1^2 + d_2^2) & d_1d_2/(d_1^2 + d_2^2) \\ d_1d_2/(d_1^2 + d_2^2) & d_2^2/(d_1^2 + d_2^2) \end{bmatrix}$$

Como comprobación, note que en el inciso (a) podría tomar $\mathbf{d} = \mathbf{e}_1$ como un vector director para el eje x. Por tanto, $d_1 = 1$ y $d_2 = 0$, y se obtiene $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, como antes.

Nuevas transformaciones lineales a partir de otras anteriores

Si $T : \mathbb{R}^m \to \mathbb{R}^n$ y $S : \mathbb{R}^n \to \mathbb{R}^p$ son transformaciones lineales, entonces puede seguir T por S para formar la ***composición*** de las dos transformaciones, denotada $S \circ T$. Note que, para que $S \circ T$ tenga sentido, el codominio de T y el dominio de S deben coincidir (en este caso, ambos son $\mathbb{R}^n$) y la transformación compuesta resultante $S \circ T$ va del dominio de T al codominio de S (en este caso, $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$). La figura 3.12 muestra de manera esquemática cómo funciona esta composición. La definición formal de composición de transformaciones se toma directamente de esta figura y es igual que la correspondiente definición de composición de funciones ordinarias:

$$(S \circ T)(\mathbf{v}) = S(T(\mathbf{v}))$$

Desde luego, le gustaría que $S \circ T$ también fuese una transformación lineal, y felizmente se encuentra que lo es. Esto se puede demostrar al mostrar que $S \circ T$ satisface la definición de transformación lineal (lo que se hará en el capítulo 6), pero, dado que por el momento se supone que las transformaciones lineales y las transformaciones matriciales son la misma cosa, es suficiente demostrar que $S \circ T$ es una transformación matricial. Se usará la notación $[T]$ para la matriz estándar de una transformación lineal T.

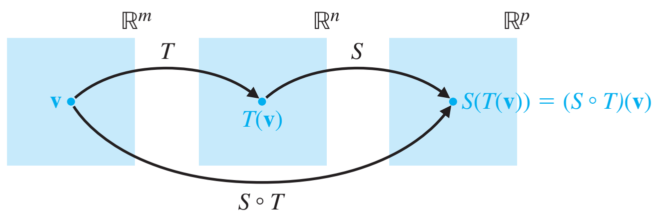


Figura 3.12

La composición de transformaciones

Teorema 3.32 Sean $T : \mathbb{R}^m \to \mathbb{R}^n$ y $S : \mathbb{R}^n \to \mathbb{R}^p$ transformaciones lineales. Entonces, $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$ es una transformación lineal. Más aún, sus matrices estándar están relacionadas mediante

$$[S \circ T] = [S][T]$$

Demostración Sean $[S] = A$ y $[T] = B$. (Note que A es de $p \times n$ y B es de $n \times m$.) Si $\mathbf{v}$ es un vector en $\mathbb{R}^m$, entonces simplemente se calcula

$$(S \circ T)(\mathbf{v}) = S(T(\mathbf{v})) = S(B\mathbf{v}) = A(B\mathbf{v}) = (AB)\mathbf{v}$$

(Note aquí que las dimensiones de A y B garantizan que el producto AB tenga sentido.) Por tanto, se ve que el efecto de $S \circ T$ es multiplicar vectores por AB, de donde se tiene inmediatamente que $S \circ T$ es una transformación matricial (por tanto, lineal) con $[S \circ T] = [S][T]$.

¿No es un gran resultado? Dicho con palabras: "la matriz de la composición es el producto de la matrices". ¡Vaya fórmula adorable!

Ejemplo 3.60 Considere la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ del ejemplo 3.55, definida por

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

y la transformación lineal $S : \mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2y_1 + y_3 \\ 3y_2 - y_3 \\ y_1 - y_2 \\ y_1 + y_2 + y_3 \end{bmatrix}$$

Encuentre $S \circ T : \mathbb{R}^2 \to \mathbb{R}^4$.

Solución Se ve que las matrices estándar son

$$[S] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad [T] = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$$

de modo que el Teorema 3.32 produce

$$[S \circ T] = [S][T] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}$$

Se tiene que

$$(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5x_1 + 4x_2 \\ 3x_1 - 7x_2 \\ -x_1 + x_2 \\ 6x_1 + 3x_2 \end{bmatrix}$$

(En el ejercicio 29 se le pedirá comprobar este resultado al hacer

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

y sustituir estos valores en la definición de S, y de este modo calcular directamente $(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.)

Ejemplo 3.61 Encuentre la matriz estándar de la transformación que primero rota un punto 90° en sentido contrario al de las manecillas del reloj en torno al origen y luego refleja el resultado en el eje x.

Solución La rotación R y la reflexión F se discutieron en los ejemplos 3.57 y 3.56, respectivamente, de donde se encuentra que sus matrices estándar son $[R] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ y $[F] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Se tiene que la composición $F \circ R$ tiene por su matriz

$$[F \circ R] = [F][R] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

(Compruebe que este resultado es correcto al considerar el efecto de $F \circ R$ sobre los vectores base estándar $\mathbf{e}_1$ y $\mathbf{e}_2$. Note la importancia del *orden* de las transformaciones: R se realiza antes que F, pero se escribe $F \circ R$. En este caso, $R \circ F$ también tiene sentido. ¿$R \circ F = F \circ R$?)

Inversas de transformaciones lineales

Considere el efecto de una rotación de 90° en sentido contrario al de las manecillas del reloj en torno al origen, seguida por una rotación de 90° en el sentido de las manecillas del reloj. Claramente, esto deja todo punto en $\mathbb{R}^2$ sin cambios. Si estas transformaciones se denotan R_{90} y R_{-90} (recuerde que un ángulo negativo corresponde a dirección contraria a la de las manecillas del reloj), entonces puede expresar esto como $(R_{90} \circ R_{-90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^2$. Note que, en este caso, si realiza las transformaciones en el otro orden, obtiene el mismo resultado final: $(R_{-90} \circ R_{90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^2$.

Por tanto, $R_{90} \circ R_{-90}$ (y también $R_{-90} \circ R_{90}$) es una transformación lineal que deja todo vector en $\mathbb{R}^2$ sin cambio. Tal transformación se conoce como ***transformación identidad***. En general, se tiene una de tales transformaciones para todo $\mathbb{R}^n$; a saber, $I : \mathbb{R}^n \to \mathbb{R}^n$ tal que $I(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^n$. (Si es importante seguir la pista de la dimensión del espacio, puede escribir I_n por claridad.)

De este modo, con esta notación, se tiene $R_{90} \circ R_{-90} = I = R_{-90} \circ R_{90}$. Un par de transformaciones que se relacionan mutuamente de esta forma se llaman ***transformaciones inversas***.

Definición Sean S y T transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^n$. Entonces S y T son ***transformaciones inversas*** si $S \circ T = I_n$ y $T \circ S = I_n$.

Comentario Dado que esta definición es simétrica con respecto a S y T, se dirá que, cuando ocurre esta situación, S es el inverso de T y T es el inverso de S. Más aún, se dirá que S y T son ***invertibles***.

En términos de matrices, se ve inmediatamente que si S y T son transformaciones inversas, entonces $[S][T] = [S \circ T] = [I] = I$, donde la última I es la *matriz* identidad. (¿Por qué la matriz estándar de la transformación identidad es la matriz identidad?) También se debe tener $[T][S] = [T \circ S] = [I] = I$. Esto demuestra que $[S]$ y $[T]$ son matrices inversas. Demuestra algo más: si una transformación lineal T es invertible, entonces su matriz estándar $[T]$ debe ser invertible, y dado que los inversos matriciales son únicos, esto significa que la inversa de T también es única. Por tanto, puede usar sin ambigüedad la notación T^{-1} para referirse a *la* inversa de T. Por tanto, puede reescribir las ecuaciones anteriores como $[T][T^{-1}] = I = [T^{-1}][T]$, lo que muestra que la matriz de T^{-1} es la matriz inversa de $[T]$. Se acaba de demostrar el siguiente teorema.

Teorema 3.33 Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal invertible. Entonces su matriz estándar $[T]$ es una matriz invertible y

$$[T^{-1}] = [T]^{-1}$$

Comentario También enuncie esto con palabras: "la matriz de la inversa es la inversa de la matriz". ¡Fabuloso!

Ejemplo 3.62 Encuentre la matriz estándar de una rotación de 60° en el sentido de las manecillas del reloj en torno al origen en $\mathbb{R}^2$.

Solución Anteriormente se calculó la matriz de una rotación de 60° en sentido contrario al de las manecillas del reloj en torno al origen como

$$[R_{60}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$$

Dado que una rotación de 60° en el sentido de las manecillas del reloj es el inverso de una rotación de 60° en sentido contrario al de las manecillas del reloj, puede aplicar el Teorema 3.33 para obtener

$$[R_{-60}] = [(R_{60})^{-1}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}^{-1} = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{bmatrix}$$

(Compruebe el cálculo de la inversa matricial. La forma más rápida es usar la abreviatura 2×2 del Teorema 3.8. También compruebe que la matriz resultante tiene el efecto correcto sobre la base estándar en $\mathbb{R}^2$ al dibujar un diagrama.)

Ejemplo 3.63 Determine si la proyección sobre el eje x es una transformación invertible y, si lo es, encuentre su inversa.

Solución La matriz estándar de esta proyección P es $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, que no es invertible pues su determinante es 0. Por tanto, P no es invertible tampoco.

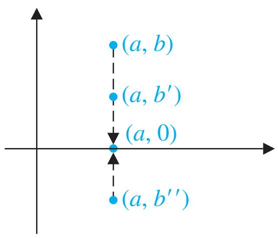


Figura 3.13
Las proyecciones no son invertibles

Comentario La figura 3.13 da alguna idea de por qué P en el ejemplo 3.63 no es invertible. La proyección "colapsa" $\mathbb{R}^2$ sobre el eje x. Para que P sea invertible, habría que tener una forma de "deshacerlo" para recuperar el punto (a, b) con el que se comenzó. Sin embargo, existe un número infinito de candidatos para la imagen de $(a, 0)$ bajo tal "inverso" hipotético. ¿Cuál usaría? No puede decir simplemente que P^{-1} debe enviar $(a, 0)$ a (a, b), pues esto no puede ser una *definición* cuando no hay forma de saber cuál debe ser b. (Vea el ejercicio 42.)

Asociatividad

El Teorema 3.3(a) de la sección 3.2 enuncia la propiedad de asociatividad para la multiplicación matricial: $A(BC) = (AB)C$. (Si no trató de demostrarlo entonces, hágalo ahora. Incluso con todas las matrices restringidas de 2×2 , obtendrá cierto sentimiento de la complejidad notacional involucrada en una prueba "por elementos", que debe hacerlo apreciar la demostración que está a punto de darse.)

El planteamiento de la demostración es vía transformaciones lineales. Se vio que toda matriz A de $n \times n$ da lugar a una transformación lineal $T_A : \mathbb{R}^n \to \mathbb{R}^m$ por el contrario, toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ tiene una correspondiente matriz $[T]$ de $m \times n$. Las dos correspondencias están inversamente relacionadas; esto es, dada A, $[T_A] = A$, y dada T, $T_{[T]} = T$.

Sea $R = T_A$, $S = T_B$ y $T = T_C$. Entonces, por el Teorema 3.32,

$$A(BC) = (AB)C \quad \text{si y sólo si} \quad R \circ (S \circ T) = (R \circ S) \circ T$$

Ahora se demuestra la última identidad. Sea $\mathbf{x}$ en el dominio de T (y por tanto en el dominio tanto de $R \circ (S \circ T)$ como de $(R \circ S) \circ T$; ¿por qué?). Para demostrar que $R \circ (S \circ T) = (R \circ S) \circ T$, es suficiente demostrar que tienen el mismo efecto sobre $\mathbf{x}$. Mediante la aplicación repetida de la definición de composición, se tiene

$$\begin{aligned}(R \circ (S \circ T))(\mathbf{x}) &= R((S \circ T)(\mathbf{x})) \\ &= R(S(T(\mathbf{x}))) \\ &= (R \circ S)(T(\mathbf{x})) = ((R \circ S) \circ T)(\mathbf{x})\end{aligned}$$

como se requería. (Compruebe cuidadosamente cómo se usó cuatro veces la definición de composición.)

Esta sección sirvió como introducción a las transformaciones lineales. En el capítulo 6 se observarán con más detalle y más generalidad estas transformaciones. Los ejercicios que siguen también contienen algunas exploraciones adicionales de este importante concepto.

Ejercicios 3.6

1. Sea $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación matricial correspondiente a $A = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$. Encuentre $T_A(\mathbf{u})$ y $T_A(\mathbf{v})$, donde $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

2. Sea $T_A : \mathbb{R}^2 \to \mathbb{R}^3$ la transformación matricial correspondiente a $A = \begin{bmatrix} 3 & -1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}$. Encuentre $T_A(\mathbf{u})$ y $T_A(\mathbf{v})$, donde $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

En los ejercicios 3-6, pruebe que la transformación dada es una transformación lineal, usando la definición (o el Comentario que sigue al ejemplo 3.55).

3. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ x - y \end{bmatrix}$

4. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ -x \\ 3x - 7y \end{bmatrix}$

5. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ 2x + y - 3z \end{bmatrix}$

6. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ x + y \\ x + y + z \end{bmatrix}$

En los ejercicios 7-10, proporcione un contraejemplo para demostrar que la transformación dada no es una transformación lineal.

7. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x^2 \end{bmatrix}$

8. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} |x| \\ |y| \end{bmatrix}$

9. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} xy \\ x + y \end{bmatrix}$

10. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 1 \\ y - 1 \end{bmatrix}$

En los ejercicios 11-14, encuentre la matriz estándar de la transformación lineal en el ejercicio dado.

11. Ejercicio 3

12. Ejercicio 4

13. Ejercicio 5

14. Ejercicio 6

En los ejercicios 15-18, demuestre que la transformación dada de $\mathbb{R}^2$ a $\mathbb{R}^2$ es lineal, al demostrar que es una transformación matricial.

15. F refleja un vector en el eje y.

16. R rota un vector 45° en sentido contrario al de las manecillas del reloj en torno al origen.

17. D estira un vector por un factor de 2 en el componente x y un factor de 3 en el componente y.

18. P proyecta un vector sobre la recta $y = x$.

19. Los tres tipos de matrices elementales dan lugar a cinco tipos de matrices de 2×2 con uno de las formas siguientes:

$$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \text{ o } \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$

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

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \text{ o } \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

Cada una de estas matrices elementales corresponde a una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$. Realice dibujos para ilustrar el efecto de cada una sobre el cuadrado unitario con vértices en $(0, 0)$, $(1, 0)$, $(0, 1)$ y $(1, 1)$.

En los ejercicios 20-25, encuentre la matriz estándar de la transformación lineal dada de $\mathbb{R}^2$ a $\mathbb{R}^2$.

20. Rotación en sentido contrario al de las manecillas del reloj de 120° en torno al origen.

21. Rotación en el sentido de las manecillas del reloj 30° en torno al origen.

22. Proyección sobre la recta $y = 2x$

23. Proyección sobre la recta $y = -x$

24. Reflexión en la recta $y = x$

25. Reflexión en la recta $y = -x$

26. Sea ℓ una recta que pasa por el origen en $\mathbb{R}^2$, P_ℓ la transformación lineal que proyecta un vector sobre ℓ y F_ℓ la transformación que refleja un vector en ℓ.

(a) Dibuje diagramas para demostrar que F_ℓ es lineal.

(b) La figura 3.14 sugiere una forma de encontrar la matriz de F_ℓ usando el hecho de que las diagonales de un paralelogramo se bisectan mutuamente. Demuestre que $F_\ell(\mathbf{x}) = 2P_\ell(\mathbf{x}) - \mathbf{x}$ y use este resultado para demostrar que la matriz estándar de F_ℓ es

$$\frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 - d_2^2 & 2d_1d_2 \\ 2d_1d_2 & -d_1^2 + d_2^2 \end{bmatrix}$$

(donde el vector director de ℓ es $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$).

(c) Si el ángulo entre ℓ y el eje x positivo es θ, demuestre que la matriz de F_ℓ es

$$\begin{bmatrix} \cos 2\theta & \operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & -\cos 2\theta \end{bmatrix}$$

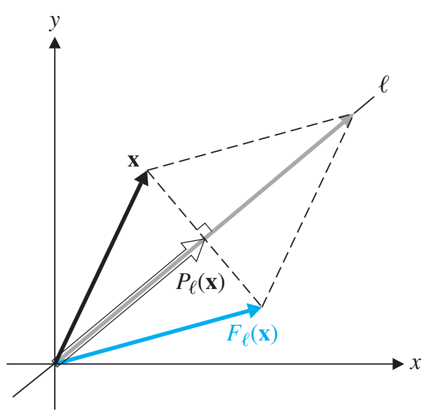


Figura 3.14

En los ejercicios 27 y 28, aplique el inciso (b) o (c) del ejercicio 26 para encontrar la matriz estándar de la transformación.

27. Reflexión en la recta $y = 2x$.

28. Reflexión en la recta $y = \sqrt{3}x$

29. Compruebe la fórmula para $S \circ T$ en el ejemplo 3.60 al realizar la sustitución directa sugerida.

En los ejercicios 30-35, verifique el Teorema 3.32 al encontrar la matriz de $S \circ T$ (a) mediante sustitución directa y (b) mediante multiplicación matricial de $[S][T]$.

30. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2y_1 \\ -y_2 \end{bmatrix}$

31. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ -3x_1 + x_2 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + 3y_2 \\ y_1 - y_2 \end{bmatrix}$

32. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + 3y_2 \\ 2y_1 + y_2 \\ y_1 - y_2 \end{bmatrix}$

33. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 - x_3 \\ 2x_1 - x_2 + x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 4y_1 - 2y_2 \\ -y_1 + y_2 \end{bmatrix}$

34. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 2x_2 - x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ y_1 + y_2 \\ -y_1 + y_2 \end{bmatrix}$

35. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ y_2 - y_3 \\ -y_1 + y_3 \end{bmatrix}$

En los ejercicios 36-39, encuentre la matriz estándar de la transformación compuesta de $\mathbb{R}^2$ a $\mathbb{R}^2$.

36. Rotación en sentido contrario al de las manecillas del reloj de 60°, seguida por reflexión en la recta $y = x$

37. Reflexión en el eje y, seguida por rotación de 30° en sentido de las manecillas del reloj

38. Rotación de 45° en sentido de las manecillas del reloj, seguida por proyección sobre el eje y, seguida por rotación de 45° en el sentido de las manecillas del reloj

39. Reflexión en la recta $y = x$, seguida por rotación de 30° en sentido contrario al de las manecillas del reloj, seguida por reflexión en la recta $y = -x$

En los ejercicios 40-43, use matrices para probar los enunciados dados acerca de las transformaciones de $\mathbb{R}^2$ en $\mathbb{R}^2$.

40. Si R_θ denota una rotación (en torno al origen) en un ángulo θ, entonces $R_\alpha \circ R_\beta = R_{\alpha+\beta}$.

41. Si θ es el ángulo entre las rectas ℓ y m (que pasan por el origen), entonces $F_m \circ F_\ell = R_{+2\theta}$. (Vea el ejercicio 26.)

42. **(a)** Si P es una proyección, entonces $P \circ P = P$.
(b) La matriz de una proyección nunca puede ser invertible.

43. Si ℓ, m y n son tres rectas que pasan por el origen, entonces $F_n \circ F_m \circ F_\ell$ es también una reflexión que pasa por el origen.

44. Sea T una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^2$ (o de $\mathbb{R}^3$ a $\mathbb{R}^3$). Demuestre que T mapea una recta en una recta o un punto. [*Sugerencia:* use la forma vectorial de la ecuación de una recta.]

45. Sea T una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^2$ (o de $\mathbb{R}^3$ a $\mathbb{R}^3$). Pruebe que T mapea rectas paralelas a rectas paralelas, una sola recta, un par de puntos, o un solo punto.

En los ejercicios 46-51, sea ABCD el cuadrado con vértices $(-1, 1), (1, 1), (1, -1)$ *y* $(-1, -1)$. *Use los resultados en los ejercicios 44 y 45 para encontrar y dibujar la imagen de ABCD de acuerdo con la transformación dada.*

46. T en el ejercicio 3

47. D en el ejercicio 17

48. P en el ejercicio 18

49. La proyección en el ejercicio 22

50. T en el ejercicio 31

51. La transformación en el ejercicio 37

52. Demuestre que $P_\ell(c\mathbf{v}) = cP_\ell(\mathbf{v})$ para cualquier escalar c [Ejemplo 3.59(b)].

53. Demuestre que $T: \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal si y sólo si

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

para todo $\mathbf{v}_1, \mathbf{v}_2$ en $\mathbb{R}^n$ y escalares c_1, c_2.

54. Demuestre que (como se anotó al comienzo de esta sección) el rango de una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ es el espacio columna de su matriz $[T]$.

55. Si A es una matriz invertible de 2×2, ¿qué afirma el teorema fundamental de las matrices invertibles acerca de la correspondiente transformación lineal T_A a la luz del ejercicio 19?

Robótica

En 1981, el transbordador espacial estadounidense *Columbia* despegó con un dispositivo llamado Shuttle Remote Manipulador System (SRMS, sistema manipulador remoto del transbordador). Este brazo robótico, conocido como Canadarm, resultó ser una herramienta vital en todas las misiones posteriores del transbordador espacial, pues proporcionó una manipulación firme, aunque precisa y delicada, de sus cargas (vea la figura 3.15).

Canadarm se ha usado para poner satélites en su órbita adecuada y para recuperar los que funcionan mal para repararlos, y también ha realizado reparaciones cruciales al mismo transbordador. Notablemente, el brazo robótico fue útil en la reparación exitosa del *Telescopio Espacial Hubble*. Desde 1998, Canadarm tiene un papel importante en el ensamblado de la *Estación Espacial Internacional*.

NASA

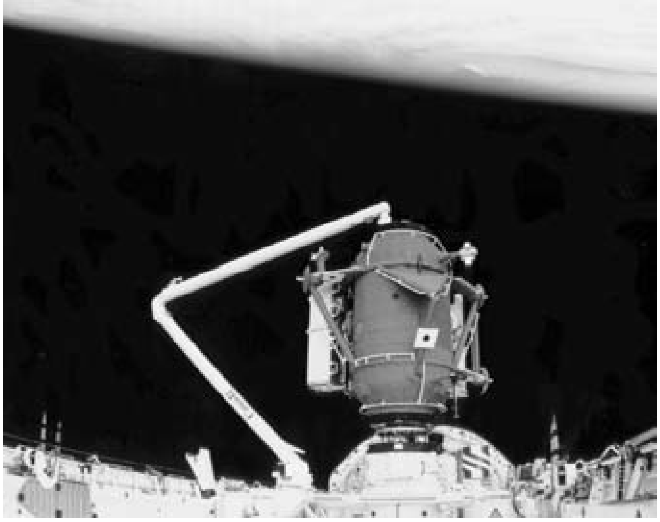

NASA

Figura 3.15
Canadarm

Un brazo robótico consiste de una serie de *vástagos* de longitud fija conectados en *articulaciones* donde pueden rotar. En consecuencia, cada vástago rota en el espacio o (mediante el efecto de los otros vástagos) se traslada paralelo a sí mismo, o se mueve mediante una combinación (composición) de rotaciones y traslaciones. Antes de poder diseñar un modelo matemático para un brazo robótico, es necesario entender cómo funcionan las rotaciones y traslaciones en composición. Para simplificar las cosas, se supondrá que el brazo está en $\mathbb{R}^2$.

En la sección 3.6 se vio que la matriz de una rotación R en torno al origen en un ángulo θ es una transformación lineal con matriz $\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ [figura 3.16(a)]. Si $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$, entonces una ***traslación a lo largo de* v** es la transformación

$$T(\mathbf{x}) = \mathbf{x} + \mathbf{v} \quad \text{o de manera equivalente,} \quad T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + a \\ y + b \end{bmatrix}$$

[figura 3.16(b)].

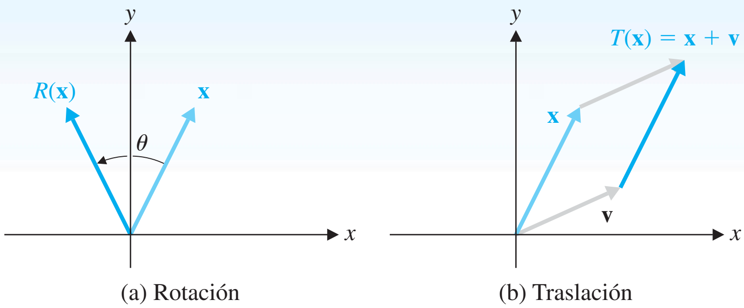


Figura 3.16

Por desgracia, la traslación no es una transformación lineal, porque $T(\mathbf{0}) \neq \mathbf{0}$. Sin embargo, hay un truco que dará la vuelta a este problema. Puede representar el vector $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ como el vector $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ en $\mathbb{R}^3$. A esto se le llama representar $\mathbf{x}$ en ***coordenadas homogéneas***. Entonces la multiplicación matricial

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + a \\ y + b \\ 1 \end{bmatrix}$$

representa al vector trasladado $T(\mathbf{x})$ en coordenadas homogéneas.

También puede tratar las rotaciones en coordenadas homogéneas. La multiplicación matricial

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \\ 1 \end{bmatrix}$$

representa el vector rotado $R(\mathbf{x})$ en coordenadas homogéneas. La composición $T \circ R$ que produce la rotación R seguida por la traslación T ahora se representa mediante el producto

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & a \\ \operatorname{sen}\theta & \cos\theta & b \\ 0 & 0 & 1 \end{bmatrix}$$

(Note que $R \circ T \neq T \circ R$.)

Para modelar un brazo robótico, proporcione a cada vástago su propio sistema coordenado (llamado *marco*) y examine cómo se mueve un vástago en relación con aquellos con los que está directamente conectado. Para ser específico, sean X_i y Y_i los ejes coordenados del vástago $\mathcal{A}_i$ con el eje X_i alineado con el vástago. La longitud de $\mathcal{A}_i$ se denota mediante a_i y el ángulo entre X_i y X_{i-1} se denota con θ_i. La articulación entre $\mathcal{A}_i$ y $\mathcal{A}_{i-1}$ está en el punto $(0, 0)$ con respecto a $\mathcal{A}_i$ y $(a_{i-1}, 0)$ relativo a $\mathcal{A}_{i-1}$. Por tanto, en relación con $\mathcal{A}_{i-1}$, el sistema coordenado para $\mathcal{A}_i$ rotó un ángulo de θ_i y luego se trasladó a lo largo de $\begin{bmatrix} a_{i-1} \\ 0 \end{bmatrix}$ (figura 3.17). Esta transformación se representa en coordenadas homogéneas mediante la matriz

$$T_i = \begin{bmatrix} \cos\theta_i & -\operatorname{sen}\theta_i & a_{i-1} \\ \operatorname{sen}\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

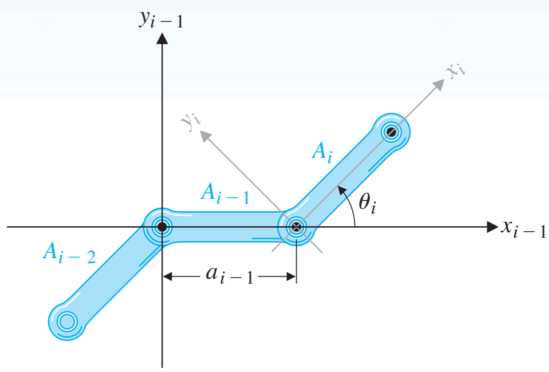


Figura 3.17

Para dar un ejemplo específico, considere la figura 3.18(a). En ella se muestra un brazo con tres vástagos en el que $\mathcal{A}_1$ está en su posición inicial y cada uno de los otros dos vástagos rotó 45° desde el vástago previo. La longitud de cada vástago se considerará en 2 unidades. La figura 3.18(b) muestra $\mathcal{A}_3$ en su marco inicial. La transformación

$$T_3 = \begin{bmatrix} \cos 45 & -\operatorname{sen} 45 & 2 \\ \operatorname{sen} 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 2 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

causa una rotación de 45° y luego una traslación de 2 unidades. Como se muestra en 3.18(c), esto coloca a $\mathcal{A}_3$ en su posición adecuada en relación con el marco de $\mathcal{A}_2$. A continuación, la transformación

$$T_2 = \begin{bmatrix} \cos 45 & -\operatorname{sen} 45 & 2 \\ \operatorname{sen} 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 2 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

se aplica al resultado anterior. Esto coloca tanto a $\mathcal{A}_3$ como a $\mathcal{A}_2$ en su posición correcta en relación con $\mathcal{A}_1$, como se muestra en la figura 3.18(d). Por lo general, al resultado previo se aplicaría una tercera transformación T_1 (una rotación), pero en este caso, T_1 es la transformación identidad porque $\mathcal{A}_1$ permanece en su posición inicial.

Usualmente se quiere conocer las coordenadas del extremo (la "mano") del brazo robótico, dados los parámetros de longitud y ángulo; esto se conoce como *forward kinematics*, cinemática hacia adelante. Al continuar con la secuencia de cálculos anterior y

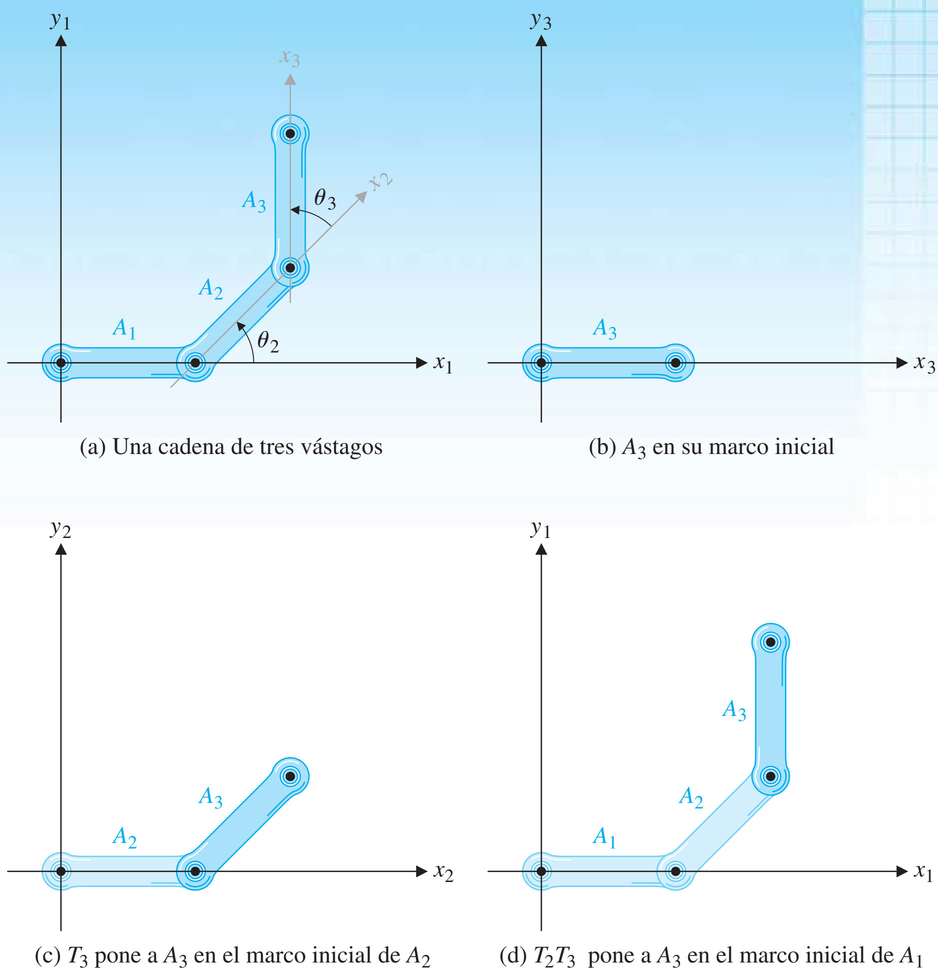


Figure 3.18

consultar la figura 3.18, se ve que es necesario determinar dónde termina el punto $(2, 0)$ después de aplicar T_3 y T_2. Por tanto, la mano del brazo está en

$$T_2T_3\begin{bmatrix}2\\0\\1\end{bmatrix} = \begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2} & 2\\ 1/\sqrt{2} & 1/\sqrt{2} & 0\\ 0 & 0 & 1\end{bmatrix}^2\begin{bmatrix}2\\0\\1\end{bmatrix} = \begin{bmatrix}0 & -1 & 2+\sqrt{2}\\ 1 & 0 & \sqrt{2}\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}2\\0\\1\end{bmatrix}$$

$$= \begin{bmatrix}2+\sqrt{2}\\ 2+\sqrt{2}\\ 1\end{bmatrix}$$

que representa el punto $(2+\sqrt{2}, 2+\sqrt{2})$ en coordenadas homogéneas. Fácilmente se comprueba, a partir de la figura 3.18(a), que esto es correcto.

Los métodos empleados en este ejemplo se generalizan a brazos robóticos en tres dimensiones, aunque en $\mathbb{R}^3$ existen más grados de libertad y por tanto más variables. El método de coordenadas homogéneas también es útil en otras aplicaciones, notablemente en gráficos de computadora.

3.7 Aplicaciones

Cadenas de Markov

Un equipo de investigación de mercado realiza un estudio controlado para determinar qué dentífricos prefieren las personas. La muestra consiste de 200 personas, y a cada una se le pide probar dos marcas de dentífrico durante un periodo de varios meses. Con base en las respuestas de la encuesta, el equipo de investigación compila las siguientes estadísticas acerca de las preferencias de dentífrico.

De quienes usan la marca A en cualquier mes, 70% siguen usándola el mes siguiente, mientras que 30% cambia a la marca B; de quienes usan la marca B en cualquier mes, 80% siguen usándolo el mes siguiente, mientras que 20% cambian a la marca A. Estos hallazgos se resumen en la figura 3.19, en la que los porcentajes se convirtieron en decimales; se pensará en ellos como probabilidades.

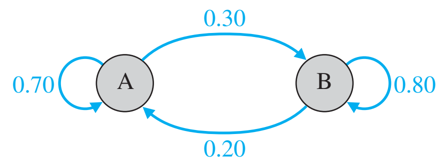


Figura 3.19

Andrei A. Markov (1856–1922) fue un matemático ruso que estudió y más tarde impartió clases en la Universidad de San Petersburgo. Estuvo interesado en teoría de números, análisis y teoría de las fracciones continuas, un campo recientemente desarrollado que Markov aplicó a la teoría de probabilidad. Markov también estuvo interesado en la poesía y uno de los usos que dio a las cadenas de Markov fue el análisis de patrones en poemas y otros textos literarios.

La figura 3.19 es un ejemplo simple de una ***cadena de Markov*** (finita). Representa un proceso evolutivo que consiste de un número finito de *estados*. En cada paso o punto en el tiempo, el proceso puede estar en cualquiera de los estados; en el paso siguiente, el proceso puede permanecer en su estado presente o cambiar a uno de los otros estados. El estado hacia donde avanza el proceso en el siguiente paso y la probabilidad de hacerlo depende *solamente* del estado presente y no de la historia pasada del proceso. Estas probabilidades se conocen como *probabilidades de transición* y se suponen constantes (esto es: la probabilidad de moverse del estado i al estado j siempre es la misma).

Ejemplo 3.64

En el estudio de dentífricos descrito anteriormente, existen sólo dos estados (usar la marca A y usar la marca B) y las probabilidades de transición son las indicadas en la figura 3.19. Suponga que, cuando comienza el estudio, 120 personas usan la marca A y 80 personas usan la marca B. ¿Cuántas personas usarán cada marca 1 mes más tarde? ¿Y 2 meses más tarde?

Solución El número de usuarios de la marca A después de 1 mes será 70% de los que inicialmente usan la marca A (quienes siguen siendo leales a la marca A) más 20% de los usuarios de la marca B (quienes cambian de B a A):

$$0.70(120) + 0.20(80) = 100$$

De igual modo, el número de usuarios de la marca B después de 1 mes será una combinación de quienes cambian a la marca B y quienes siguen usándola:

$$0.30(120) + 0.80(80) = 100$$

Estas dos ecuaciones pueden resumirse en una sola ecuación matricial:

$$\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 120 \\ 80 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$$

Llame P a la matriz y marque los vectores $\mathbf{x}_0 = \begin{bmatrix} 120 \\ 80 \end{bmatrix}$ y $\mathbf{x}_1 = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$. (Note que los componentes de cada vector son el número de usuarios de las marcas A y B, en ese orden, después del número de meses indicado por el subíndice.) Por tanto, se tiene $\mathbf{x}_1 = P\mathbf{x}_0$.

Al extender la notación, sea $\mathbf{x}_k$ el vector cuyos componentes registran la distribución de usuarios de dentífrico después de k meses. Para determinar el número de usuarios de cada marca después de transcurridos 2 meses, simplemente aplique el mismo razonamiento comenzando con $\mathbf{x}_1$ en lugar de $\mathbf{x}_0$. Se obtiene

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 90 \\ 110 \end{bmatrix}$$

de donde se ve que ahora existen 90 usuarios de la marca A y 110 usuarios de la marca B.

Los vectores $\mathbf{x}_k$ en el ejemplo 3.64 se llaman ***vectores de estado*** de la cadena de Markov, y la matriz P se llama ***matriz de transición***. Acaba de ver que una cadena de Markov satisface la relación

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{para} \quad k = 0, 1, 2, \ldots$$

A partir de este resultado se tiene que es posible calcular un vector de estado arbitrario *de manera iterativa* una vez se conoce $\mathbf{x}_0$ y P. En otras palabras, una cadena de Markov está *completamente determinada* por sus probabilidades de transición y su estado inicial.

Comentarios

- Suponga que en el ejemplo 3.64 se quiere seguir la pista no de los números *reales* de usuarios de dentífrico, sino de los números *relativos* que usan cada marca. Podría convertir los datos en porcentajes o fracciones al dividir el número total de usuarios entre 200. Por tanto, comenzaría con

$$\mathbf{x}_0 = \begin{bmatrix} \frac{120}{200} \\ \frac{80}{200} \end{bmatrix} = \begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix}$$

para reflejar el hecho de que, inicialmente, la división marca A-marca B es 60%–40%. Compruebe mediante cálculo directo que $P\mathbf{x}_0 = \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}$, que entonces puede tomarse como $\mathbf{x}_1$ (en concordancia con la división 50-50 que se calculó anteriormente). Los vectores como este, con componentes no negativos que suman 1, se llaman ***vectores de probabilidad***.

- Observe cómo se ordenan las probabilidades de transición dentro de la matriz de transición P. Puede considerar a las columnas como etiquetadas con los estados *presente* y los renglones como etiquetados con los estados *siguiente*:

$$\begin{array}{cc} & \textit{Presente} \\ & \begin{array}{cc} \text{A} & \text{B} \end{array} \\ \textit{Siguiente} \begin{array}{c} \text{A} \\ \text{B} \end{array} & \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \end{array}$$

La palabra *estocástico* se deriva del adjetivo griego *stokhastikos*, que significa "capaz de conjeturar" (o adivinar). Llegó a aplicarse a algo que está gobernado por las leyes de la probabilidad en el sentido de que la probabilidad realiza predicciones acerca de la posibilidad de que ocurran las cosas. En teoría de probabilidad, los "procesos estocásticos" forman una generalización de las cadenas de Markov.

Note también que las columnas de P son vectores de probabilidad; cualquier matriz cuadrada con esta propiedad se llama ***matriz estocástica***.

Puede darse cuenta de la naturaleza determinista de las cadenas de Markov en otra forma. Note que puede escribir

$$\mathbf{x}_2 = P\mathbf{x}_1 = P(P\mathbf{x}_0) = P^2\mathbf{x}_0$$

y, en general,

$$\mathbf{x}_k = P^k\mathbf{x}_0 \quad \text{para} \quad k = 0, 1, 2, \ldots$$

Esto conduce a examinar las potencias de una matriz de transición. En el ejemplo 3.64 se tiene

$$P^2 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} = \begin{bmatrix} 0.55 & 0.30 \\ 0.45 & 0.70 \end{bmatrix}$$

¿Qué se hará con las entradas de esta matriz? Lo primero que debe observar es que p^2 es otra matriz estocástica, pues sus columnas suman 1. (Se le pedirá probar esto en el ejercicio 14.) ¿Podría ser que P^2 también sea una matriz de transición del mismo tipo? Considere una de sus entradas, por ejemplo, $(P^2)_{21} = 0.45$. Los tres diagramas de la figura 3.20 clarifican de dónde provienen estas entradas.

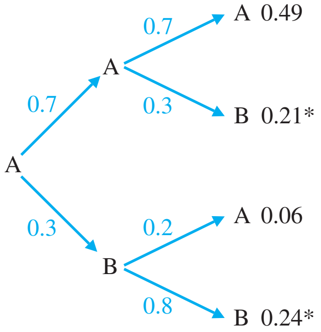


Figura 3.20

Existen cuatro posibles cambios de estado que pueden ocurrir durante 2 meses y corresponden a las cuatro ramas (o rutas) de longitud 2 en el árbol. Alguien que inicialmente use la marca A puede terminar usando la marca B dos meses después en dos formas diferentes (marcadas * en la figura): la persona puede seguir usando A después de 1 mes y luego cambiar a B (con probabilidad $0.7(0.3) = 0.21$), o la persona puede cambiar a B después de 1 mes y luego seguir con B (con probabilidad $0.3(0.8) = 0.24$). La suma de estas probabilidades produce una probabilidad global de 0.45. Observe que dichos cálculos son *exactamente* lo que se hizo cuando se calculó $(P^2)_{21}$.

Se tiene que $(P^2)_{21} = 0.45$ representa la probabilidad de moverse del estado 1 (marca A) al estado 2 (marca B) en dos transiciones. (Note que el orden de los subíndices es el *inverso* del que pudo suponer.) El argumento puede generalizarse para demostrar que

$(P^k)_{ij}$ es la probabilidad de moverse del estado j al estado i en k transiciones.

En el ejemplo 3.64, ¿qué sucederá con la distribución de usuarios de dentífrico a largo plazo? Trabaje con vectores de probabilidad como vectores de estado. Al continuar con los cálculos (redondeados a tres lugares decimales), se encuentra

$$\mathbf{x}_0 = \begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix}, \mathbf{x}_1 = \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}, \mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix},$$

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix} = \begin{bmatrix} 0.425 \\ 0.575 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 0.412 \\ 0.588 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 0.406 \\ 0.594 \end{bmatrix},$$

$$\mathbf{x}_6 = \begin{bmatrix} 0.403 \\ 0.597 \end{bmatrix}, \mathbf{x}_7 = \begin{bmatrix} 0.402 \\ 0.598 \end{bmatrix}, \mathbf{x}_8 = \begin{bmatrix} 0.401 \\ 0.599 \end{bmatrix}, \mathbf{x}_9 = \begin{bmatrix} 0.400 \\ 0.600 \end{bmatrix}, \mathbf{x}_{10} = \begin{bmatrix} 0.400 \\ 0.600 \end{bmatrix},$$

etcétera. Parece que los vectores de estado tienden (o *convergen a*) el vector $\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$, lo que implica que, a la larga, 40% de los usuarios de dentífrico en el estudio usarán la marca A y 60% usarán la marca B. De hecho, es fácil comprobar que, una vez alcanzada esta distribución, nunca cambiará. Simplemente calcule

$$\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

Un vector de estado $\mathbf{x}$ con la propiedad de que $P\mathbf{x} = \mathbf{x}$ se llama ***vector de estado estacionario***. En el capítulo 4 se probará que toda cadena de Markov tiene un vector de estado estacionario único. Por ahora, acepte esto como un hecho y vea cómo puede encontrar tal vector sin hacer iteración alguna.

Comience por reescribir la ecuación matricial $P\mathbf{x} = \mathbf{x}$ como $P\mathbf{x} = I\mathbf{x}$, que a su vez puede reescribirse como $(I - P)\mathbf{x} = \mathbf{0}$. Ahora, esto es justo un sistema homogéneo de ecuaciones lineales con matriz de coeficientes $I - P$, de modo que la matriz aumentada es $[I - P|\mathbf{0}]$. En el ejemplo 3.64 se tiene

$$[I - P \mid \mathbf{0}] = \left[\begin{array}{cc|c} 1 - 0.70 & -0.20 & 0 \\ -0.30 & 1 - 0.80 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 0.30 & -0.20 & 0 \\ -0.30 & 0.20 & 0 \end{array}\right]$$

que se reduce a

$$\left[\begin{array}{cc|c} 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

De modo que, si el vector de estado estacionario es $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, entonces x_2 es una variable libre y la solución paramétrica es

$$x_1 = \tfrac{2}{3}t, \quad x_2 = t$$

Si se requiere que $\mathbf{x}$ sea un vector de probabilidad, entonces debe tener

$$1 = x_1 + x_2 = \tfrac{2}{3}t + t = \tfrac{5}{3}t$$

Por tanto, $x_2 = t = \frac{3}{5} = 0.6$ y $x_1 = \frac{2}{5} = 0.4$, de modo que $\mathbf{x} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$, en concordancia con los anteriores cálculos iterativos. (Si requiere que $\mathbf{x}$ contenga la distribución *real*, entonces en este ejemplo debe tener $x_1 + x_2 = 200$, de donde se tiene que $\mathbf{x} = \begin{bmatrix} 80 \\ 120 \end{bmatrix}$.)

Ejemplo 3.65

Un psicólogo coloca una rata en una jaula con tres compartimentos, como se muestra en la figura 3.21. La rata fue entrenada para seleccionar una puerta al azar siempre que suene una campana y moverse a través de ella hacia el siguiente compartimento.

(a) Si inicialmente la rata está en el compartimiento 1, ¿cuál es la probabilidad de que estará en el compartimiento 2 después de que la campana suene dos veces? ¿Tres veces?

(b) A largo plazo, ¿qué proporción de su tiempo pasará la rata en cada compartimiento?

Solución Sea $P = [p_{ij}]$ la matriz de transición para esta cadena de Markov. Entonces

$$p_{21} = p_{31} = \tfrac{1}{2},\ p_{12} = p_{13} = \tfrac{1}{3},\ p_{32} = p_{23} = \tfrac{2}{3} \ \text{ y } \ p_{11} = p_{22} = p_{33} = 0$$

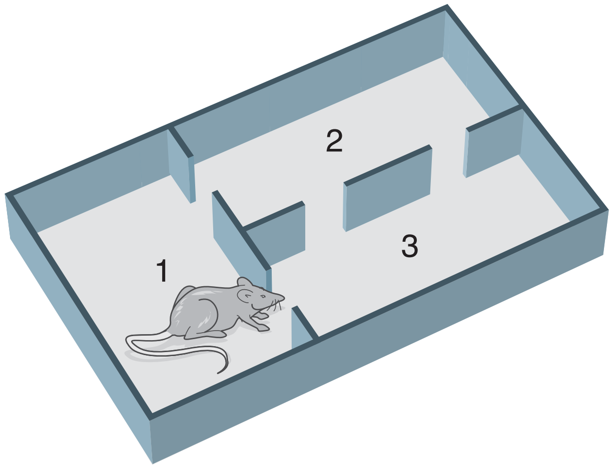


Figura 3.21

(¿Por qué? Recuerde que p_{ij} es la probabilidad de moverse de j a i.) Por tanto

$$P = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}$$

y el vector de estado inicial es

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

(a) Después de un campanazo, se tiene

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix}$$

Al continuar (redondeado a tres lugares decimales), se encuentra

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}\begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \approx \begin{bmatrix} 0.333 \\ 0.333 \\ 0.333 \end{bmatrix}$$

y

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{9} \\ \frac{7}{18} \\ \frac{7}{18} \end{bmatrix} \approx \begin{bmatrix} 0.222 \\ 0.389 \\ 0.389 \end{bmatrix}$$

Por tanto, después de dos campanazos, la probabilidad de que la rata esté en el compartimiento 2 es $\frac{1}{3} \approx 0.333$, y después de tres campanazos, la probabilidad de que la rata esté en el compartimiento 2 es $\frac{7}{18} \approx 0.389$. [Note que estas preguntas también podrían responderse al calcular $(P^2)_{21}$ y $(P^3)_{21}$.]

(b) Esta pregunta pide el vector de estado estacionario $\mathbf{x}$ como un vector de probabilidad. Como se vio anteriormente, $\mathbf{x}$ debe estar en el espacio nulo de $I - P$, de modo que se procede a resolver el sistema

$$[I - P \mid \mathbf{0}] = \left[\begin{array}{ccc|c} 1 & -\frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{2} & 1 & -\frac{2}{3} & 0 \\ -\frac{1}{2} & -\frac{2}{3} & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -\frac{2}{3} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Por tanto, si $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, entonces $x_3 = t$ es libre y $x_1 = \frac{2}{3}t$, $x_2 = t$. Dado que $\mathbf{x}$ debe ser un vector de probabilidad, se necesita $1 = x_1 + x_2 + x_3 = \frac{8}{3}t$. En consecuencia, $t = \frac{3}{8}$ y

$$\mathbf{x} = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{8} \\ \frac{3}{8} \end{bmatrix}$$

lo cual dice que, a largo plazo, la rata pasa $\frac{1}{4}$ de su tiempo en el compartimiento 1 y $\frac{3}{8}$ de su tiempo en cada uno de los otros dos compartimientos.

Modelos económicos lineales

Ahora se visitarán nuevamente los modelos económicos que se encontraron por primera vez en la sección 2.4 y se reformularán en términos de matrices. El ejemplo 2.33 ilustró el modelo cerrado de Leontief. El sistema de ecuaciones que debió resolverse fue

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

En forma matricial, esta es la ecuación $E\mathbf{x} = \mathbf{x}$, donde

$$E = \begin{bmatrix} 1/4 & 1/3 & 1/2 \\ 1/4 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

La matriz E se llama ***matriz de intercambio*** y el vector $\mathbf{x}$ se llama ***vector de precio***. En general, si $E = [e_{ij}]$, entonces e_{ij} representa la fracción (o porcentaje) de output de la industria j que consume la industria i, y x_i es el precio que cobra la industria i por su output.

En una economía cerrada, la suma de cada columna de E es 1. Dado que las entradas de E también son no negativas, E es una matriz estocástica y el problema de encontrar una solución a la ecuación

$$E\mathbf{x} = \mathbf{x} \tag{1}$$

¡es precisamente el mismo que el problema de encontrar el vector de estado estacionario de una cadena de Markov! Por tanto, para encontrar un vector precio $\mathbf{x}$ que satisfaga $E\mathbf{x} = \mathbf{x}$, se resuelve la ecuación homogénea equivalente $(I - E)\mathbf{x} = 0$. Siempre existirá un número infinito de soluciones; se busca una solución donde los precios sean todos no negativos y al menos un precio sea positivo.

El modelo abierto de Leontief es más interesante. En el ejemplo 2.34 fue necesario resolver el sistema

$$\begin{aligned} x_1 &= 0.2x_1 + 0.5x_2 + 0.1x_3 + 10 \\ x_2 &= 0.4x_1 + 0.2x_2 + 0.2x_3 + 10 \\ x_3 &= 0.1x_1 + 0.3x_2 + 0.3x_3 + 30 \end{aligned}$$

En forma matricial, se tiene

$$\mathbf{x} = C\mathbf{x} + \mathbf{d} \quad o \quad (I - C)\mathbf{x} = \mathbf{d} \tag{2}$$

donde

$$C = \begin{bmatrix} 0.2 & 0.5 & 0.1 \\ 0.4 & 0.2 & 0.2 \\ 0.1 & 0.3 & 0.3 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \mathbf{d} = \begin{bmatrix} 10 \\ 10 \\ 30 \end{bmatrix}$$

La matriz C se llama ***matriz de consumo***, $\mathbf{x}$ es el ***vector de producción***, y $\mathbf{d}$ es el ***vector de demanda***. En general, si $C = [c_{ij}]$, $\mathbf{x} = [x_i]$, y $\mathbf{d} = [d_i]$, entonces c_{ij} representa el valor en dólares del output de la industria i que se necesita para producir un dólar de output de la industria j, x_i es el valor monetario (precio) del output de la industria i, y d_i es el valor monetario de la demanda externa para el output de la industria i. Una vez más, se está interesado en encontrar un vector de producción $\mathbf{x}$ con entradas no negativas tales que al menos una entrada sea positiva. A tal vector $\mathbf{x}$ se le llama ***solución factible***.

Ejemplo 3.66

Determine si existe una solución para modelo abierto de Leontief determinado por las siguientes matrices de consumo:

(a) $C = \begin{bmatrix} 1/4 & 1/3 \\ 1/2 & 1/3 \end{bmatrix}$ (b) $C = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 2/3 \end{bmatrix}$

Solución (a) Se tiene

$$I - C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/4 & 1/3 \\ 1/2 & 1/3 \end{bmatrix} = \begin{bmatrix} 3/4 & -1/3 \\ -1/2 & 2/3 \end{bmatrix}$$

de modo que la ecuación $(I - C)\mathbf{x} = \mathbf{d}$ se convierte en

$$\begin{bmatrix} 3/4 & -1/3 \\ -1/2 & 2/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

En la práctica se reduciría por renglón la correspondiente matriz aumentada para determinar una solución. Sin embargo, en este caso es instructivo notar que la matriz de coeficientes $I - C$ es invertible y entonces aplicar el Teorema 3.7. Se calcula

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3/4 & -1/3 \\ -1/2 & 2/3 \end{bmatrix}^{-1} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3/2 & 9/4 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

Dado que d_1, d_2 y todas las entradas de $(I - C)^{-1}$ son no negativas, también lo son x_1 y x_2. Por tanto, puede encontrarse una solución factible para *cualquier* vector de demanda distinto de cero.

(b) En este caso,

$$I - C = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 2/3 \end{bmatrix} \quad \text{y} \quad (I - C)^{-1} = \begin{bmatrix} -4 & -6 \\ -6 & -6 \end{bmatrix}$$

de modo que

$$\mathbf{x} = (I - C)^{-1}\mathbf{d} = \begin{bmatrix} -4 & -6 \\ -6 & -6 \end{bmatrix}\mathbf{d}$$

Dado que todas las entradas de $(I - C)^{-1}$ son negativas, esto no producirá una solución factible para *cualquier* vector de demanda distinto de cero $\mathbf{d}$.

Motivada por el ejemplo 3.66, se tiene la siguiente definición. (Para dos matrices $m \times n$; $A = [a_{ij}]$ y $B = [b_{ij}]$, se escribirá $A \geq B$ si $a_{ij} \geq b_{ij}$ para todo i y j. De igual modo, puede definirse $A > B$, $A \leq B$, etcétera. Una matriz A se llama no negativa si $A \geq O$ y positiva si $A > O$.)

Definición Una matriz de consumo C se llama ***productiva*** si $I - C$ es invertible y $(I - C)^{-1} \geq O$.

Ahora se dan tres resultados que proporcionan criterios para que una matriz de consumo sea productiva.

Teorema 3.34 Sea C una matriz de consumo. Entonces C es productiva si y sólo si existe un vector de producción $\mathbf{x} \geq \mathbf{0}$ tal que $\mathbf{x} > C\mathbf{x}$.

Demostración Suponga que C es productiva. Entonces $I - C$ es invertible y $(I - C)^{-1} \geq O$. Sea

$$\mathbf{j} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Entonces $\mathbf{x} = (I - C)^{-1}\mathbf{j} \geq 0$ y $(I - C)\mathbf{x} = \mathbf{j} > 0$. Por tanto, $\mathbf{x} - C\mathbf{x} > 0$ o, de manera equivalente, $\mathbf{x} > C\mathbf{x}$

Por el contrario, suponga que existe un vector $\mathbf{x} \geq 0$ tal que $\mathbf{x} > C\mathbf{x}$. Dado que $C \geq O$ y $C \neq O$, se tiene $\mathbf{x} > 0$ por el ejercicio 35. Más aún, debe existir un número real λ con $0 < \lambda < 1$ tal que $C\mathbf{x} < \lambda\mathbf{x}$. Pero entonces

$$C^2\mathbf{x} = C(C\mathbf{x}) \leq C(\lambda\mathbf{x}) = \lambda(C\mathbf{x}) < \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}$$

Por inducción, puede demostrarse que $0 \leq C^n\mathbf{x} < \lambda^n\mathbf{x}$ para toda $n \geq 0$. (Escriba los detalles de esta demostración por inducción.) Dado que $0 < \lambda < 1$, λ^n se aproxima a 0 cuando n se vuelve grande. Por tanto, cuando $n \to \infty$, $\lambda^n\mathbf{x} \to 0$ y en consecuencia $C^n\mathbf{x} \to 0$. Puesto que $\mathbf{x} > 0$, debe tener $C^n \to O$ cuando $n \to \infty$.

Ahora considere la ecuación matricial

$$(I - C)(I + C + C^2 + \dots + C^{n-1}) = I - C^n$$

Cuando $n \to \infty$, $C^n \to O$, así que se tiene

$$(I - C)(I + C + C^2 + \dots) = I - O = I$$

Por tanto, $I - C$ es invertible, con su inverso dado por la serie matricial infinita $I + C + C^2 + \dots$. Dado que todos los términos en esta serie son no negativos, también se tiene

$$(I - C)^{-1} = I + C + C^2 + \dots \geq O$$

En consecuencia, C es productiva.

Comentarios

- La serie infinita $I + C + C^2 + \dots$ es el análogo matricial de la serie geométrica $1 + x + x^2 + \dots$. Es posible que el lector esté familiarizado con el hecho de que, para $|x| < 1$, $1 + x + x^2 + \dots = 1/(1 - x)$.
- Dado que el vector $C\mathbf{x}$ representa las cantidades consumidas por cada industria, la desigualdad $\mathbf{x} > C\mathbf{x}$ significa que existe cierto nivel de producción para el cual cada industria produce más de lo que consume.
- Para un planteamiento alternativo a la primera parte de la demostración del Teorema 3.34, vea el ejercicio 42 de la sección 4.6.

COROLARIO 3.35

Sea C una matriz de consumo. Si la suma de cada renglón de C es menor que 1, entonces C es productiva.

La palabra *corolario* proviene de la palabra latina *corollarium*, que se refiere a una guirnalda otorgada como recompensa. Por tanto, un corolario es una pequeña recompensa adicional que se tiene de un teorema.

Demostración Si

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

entonces $C\mathbf{x}$ es un vector que consiste de las sumas de renglón de C. Si cada suma de renglón de C es menor que 1, entonces se satisface la condición $\mathbf{x} > C\mathbf{x}$. Por tanto, C es productiva.

COROLARIO 3.36

Sea C una matriz de consumo. Si la suma de cada columna de C es menor que 1, entonces C es productiva.

Demostración Si cada suma de columna de C es menor que 1, entonces cada suma de columna de C^T es menor que 1. Por tanto, C^T es productiva, por el Corolario 3.35. En consecuencia, por los Teoremas 3.9(d) y 3.4,

$$((I - C)^{-1})^T = ((I - C)^T)^{-1} = (I^T - C^T)^{-1} = (I - C^T)^{-1} \geq O$$

Se tiene que $(I - C)^{-1} \geq O$ también y, por tanto, C es productiva.

En el ejercicio 52 de la sección 7.2 se le pedirá dar demostraciones alternativas a los Corolarios 3.35 y 3.36.

De la definición de matriz de consumo se tiene que la suma de la columna j es el valor monetario total de todas las entradas necesarias para producir un importe monetario de la output de la industria j; esto es, el ingreso de la industria j supera sus gastos. Se dice que tal industria es ***rentable***. En consecuencia, el Corolario 3.36 puede parafrasearse para enunciar que una matriz de consumo es productiva si todas las industrias son rentables.

P. H. Lewis, "On the Use of Matrices en Certain Population Mathematics", *Biometrika* **33** (1945), pp. 183-212.

Crecimiento poblacional

Uno de los modelos más populares de crecimiento poblacional es uno basado en matrices, introducido por primera vez por P. H. Leslie en 1945. El *modelo de Leslie* describe el crecimiento de la porción femenina de una población, la cual se supone tiene una vida máxima. Las hembras se dividen en clases etáreas (por edades), todas las cuales abarcan un número igual de años. Con datos acerca de las tasas de natalidad promedio y probabilidades de supervivencia de cada clase, el modelo es capaz de determinar el crecimiento de la población a lo largo del tiempo.

Ejemplo 3.67

Cierta especie de escarabajo alemán, llamado de Vollmar-Wasserman (o escarabajo VW, para abreviar), vive cuando mucho 3 años. Las hembras del escarabajo VW se dividen en tres clases etáreas de 1 año cada una: jóvenes (0-1 año), juveniles (1-2 años) y adultas (2-3 años). Las jóvenes no ponen huevos; cada juvenil produce un promedio de cuatro escarabajos hembras; y cada adulta produce un promedio de tres hembras.

La tasa de supervivencia para jóvenes es 50% (esto es: la probabilidad de que una joven sobreviva hasta volverse juvenil es 0.5), y la tasa de supervivencia de las juveniles es 25%. Suponga que comienza con una población de 100 escarabajos VW hembras: 40 jóvenes, 40 juveniles y 20 adultas. Prediga la población de escarabajos para cada uno de los siguientes 5 años.

Solución Después de 1 año, el número de jóvenes será el número producido durante dicho año:

$$40 \times 4 + 20 \times 3 = 220$$

El número de juveniles simplemente será el número de jóvenes que sobrevivió:

$$40 \times 0.5 = 20$$

Del mismo modo, el número de adultas será el número de juveniles que sobrevivió:

$$40 \times 0.25 = 10$$

Puede combinar esto en una sola ecuación matricial

$$\begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}\begin{bmatrix} 40 \\ 40 \\ 20 \end{bmatrix} = \begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix}$$

o $L\mathbf{x}_0 = \mathbf{x}_1$, donde $\mathbf{x}_0 = \begin{bmatrix} 40 \\ 40 \\ 20 \end{bmatrix}$ es el vector de distribución de población inicial y $\mathbf{x}_1 = \begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix}$ es la distribución después de 1 año. Se ve que la estructura de la ecuación es exactamente igual que la de las cadenas de Markov: $\mathbf{x}_{k+1} = L\mathbf{x}_k$ para $k = 0, 1, 2, \ldots$ (aunque la interpretación es bastante diferente). Se tiene que es posible calcular por iteración vectores de distribución de población sucesivos. (También se tiene que $\mathbf{x}_k = L^k\mathbf{x}_0$ para $k = 0, 1, 2, \ldots$, como para las cadenas de Markov, pero aquí no se usará este hecho.)

Calcule

$$\mathbf{x}_2 = L\mathbf{x}_1 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}\begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 110 \\ 110 \\ 5 \end{bmatrix}$$

$$\mathbf{x}_3 = L\mathbf{x}_2 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}\begin{bmatrix} 110 \\ 110 \\ 5 \end{bmatrix} = \begin{bmatrix} 455 \\ 55 \\ 27.5 \end{bmatrix}$$

$$\mathbf{x}_4 = L\mathbf{x}_3 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 455 \\ 55 \\ 27.5 \end{bmatrix} = \begin{bmatrix} 302.5 \\ 227.5 \\ 13.75 \end{bmatrix}$$

$$\mathbf{x}_5 = L\mathbf{x}_4 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 302.5 \\ 227.5 \\ 13.75 \end{bmatrix} = \begin{bmatrix} 951.2 \\ 151.2 \\ 56.88 \end{bmatrix}$$

Por tanto, el modelo predice que, después de 5 años, habrá aproximadamente 951 hembras jóvenes de escarabajo VW, 151 juveniles y 57 adultas. (*Nota:* podría argumentar que en cada paso habría que redondear al entero más cercano, por ejemplo, 28 adultos después del paso 3, lo que afectaría las iteraciones posteriores. Se eligió *no* hacer esto, pues, de cualquier forma, los cálculos sólo son aproximaciones y es mucho más fácil usar una calculadora o CAS si no redondea conforme avanza.)

La matriz L en el ejemplo 3.67 se llama ***matriz de Leslie***. En general, si se tiene una población con n clases etáreas de igual duración, L será una matriz de $n \times n$ con la siguiente estructura:

$$L = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{n-1} & b_n \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix}$$

Aquí, $b_1, b_2, \ldots$ son los *parámetros de natalidad* (b_i = número promedio de hembras producidas por cada hembra en la clase i) y $s_1, s_2, \ldots$ son las *probabilidades de supervivencia* (s_i = probabilidad de que una hembra en la clase i sobreviva en la clase $i + 1$).

¿Qué se hará con los cálculos? En general, la población de escarabajos parece aumentar, aunque existen algunas fluctuaciones, como una disminución de 250 a 225 del año 1 al año 2. La figura 3.22 muestra el cambio en la población en cada una de las tres clases etáreas y claramente muestra el crecimiento, con fluctuaciones.

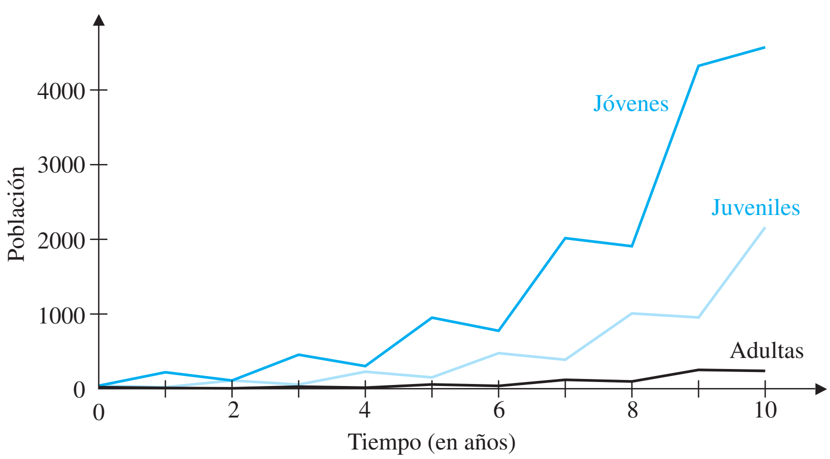


Figura 3.22

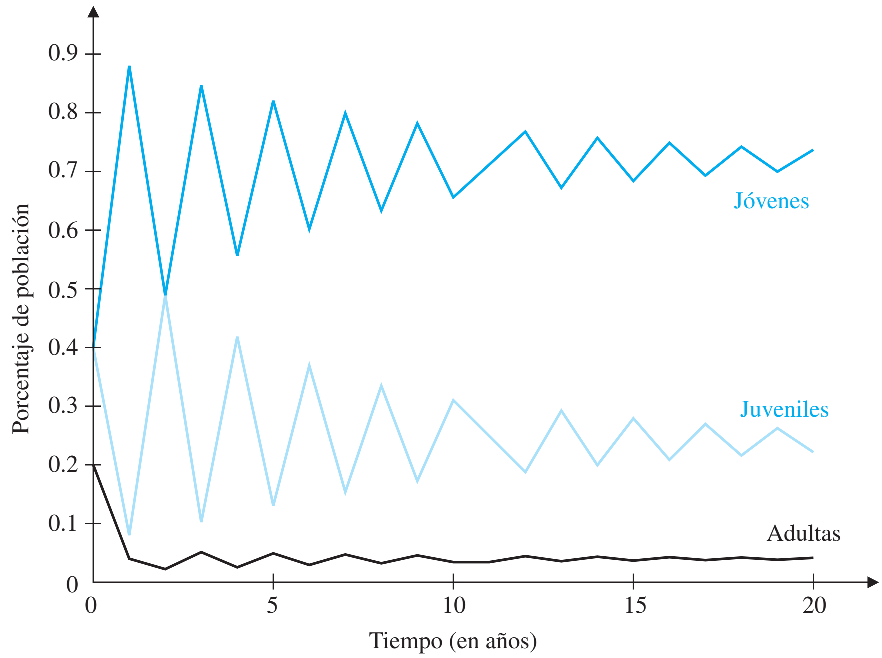


Figura 3.23

Si, en lugar de graficar la población *real* se grafica la población *relativa* en cada clase, surge un patrón diferente. Para hacer esto, es necesario calcular la fracción de la población en cada clase etárea en cada año; esto es, es necesario dividir cada vector de distribución por la suma de sus componentes. Por ejemplo, después de 1 año, se tiene

$$\frac{1}{250}\mathbf{x}_1 = \frac{1}{250}\begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 0.88 \\ 0.08 \\ 0.04 \end{bmatrix}$$

la cual dice que 88% de la población consiste de jóvenes, 8% de juveniles y 4% de adultas. Si este tipo de datos se grafican contra el tiempo, se obtiene una gráfica como la de la figura 3.23, que muestra claramente que la proporción de la población en cada clase se aproxima a un estado estable. Se evidencia que el vector de estado estacionario en este ejemplo es

$$\begin{bmatrix} 0.72 \\ 0.24 \\ 0.04 \end{bmatrix}$$

Esto es, a largo plazo, 72% de la población será de jóvenes, 24% de juveniles y 4% de adultas. (En otras palabras, la población se distribuye entre las tres clases etáreas en la razón 18 : 6 : 1.) En el capítulo 4 se verá cómo determinar esta razón con exactitud.

Grafos y digrafos

Existen muchas situaciones en las que es importante poder modelar las interrelaciones entre un conjunto finito de objetos. Por ejemplo, acaso quiera describir varios tipos de redes (carreteras que conectan ciudades, rutas de aerolíneas que conectan ciudades, ligas de comunicación que conectan satélites, etcétera) o relaciones entre grupos o individuos (relaciones de amistad en una sociedad, relaciones depredador-presa en un ecosistema,

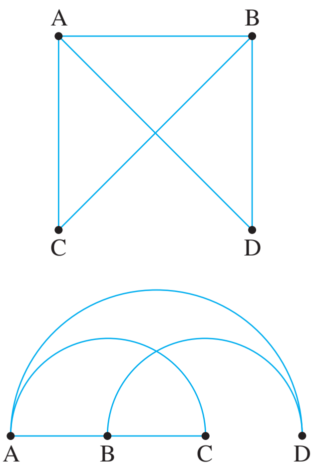


Figura 3.24
Dos representaciones del mismo grafo

relaciones de dominio en un deporte, etcétera). Los grafos son idealmente adecuados para modelar tales redes y relaciones, y es evidente que las matrices son una herramienta útil en su estudio.

Un ***grafo*** consiste de un conjunto finito de puntos (llamados *vértices*) y un conjunto finito de *aristas*, cada una de las cuales conecta dos vértices (no necesariamente distintos). Se dice que dos vértices son *adyacentes* si son los puntos finales de una arista. La figura 3.24 muestra un ejemplo del mismo grafo dibujado en dos formas diferentes. Los grafos son "iguales" en el sentido de que todo por lo que debe preocuparse es por las relaciones de adyacencia que identifican las aristas.

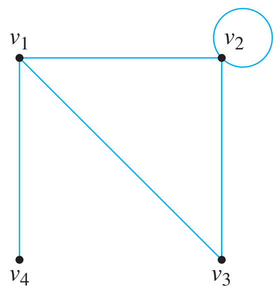
Es posible registrar la información esencial acerca de un grafo en una matriz y usar álgebra matricial para ayudar a responder ciertas preguntas acerca del grafo. Esto es particularmente útil si los grafos son grandes, pues las computadoras pueden manejar los cálculos muy rápidamente.

Definición Si G es un grafo con n vértices, entonces su ***matriz de adyacencia*** es la matriz A [o $A(G)$] de $n \times n$ definida por

$$a_{ij} = \begin{cases} 1 & \text{si existe una arista entre los vértices } i \text{ y } j \\ 0 & \text{de otro modo} \end{cases}$$

El término *vértice* viene del verbo latino *vertere*, que significa "volver". En el contexto de los grafos (y la geometría), un vértice es una esquina, un punto donde una arista "se vuelve" una arista diferente.

La figura 3.25 muestra un grafo y su matriz de adyacencia asociada.



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

Figura 3.25
Un grafo con matriz de adyacencia A

Comentario Observe que la matriz de adyacencia de un grafo necesariamente es una matriz simétrica. (¿Por qué?) Note también que una entrada diagonal a_{ii} de A es cero a menos que haya un bucle en el vértice i. En algunas situaciones, un grafo puede tener más de una arista entre un par de vértices. En tales casos, puede tener sentido modificar la definición de la matriz de adyacencia de modo que a_{ij} sea igual al *número* de aristas entre los vértices i y j.

Una ***trayectoria*** en un grafo es una secuencia de aristas que permiten viajar de un vértice a otro de manera continua. La ***longitud*** de una trayectoria es el número de aristas que contiene, y a una trayectoria con k aristas se le denominará ***k-trayectoria***. Por ejemplo, en el grafo de la figura 3.25, $v_1v_3v_2v_1$ es una 3-trayectoria, y $v_4v_1v_2v_2v_1v_3$ es una 5-trayectoria. Note que la primera de éstas es *cerrada* (comienza y termina en el mismo vértice); a tal trayectoria se le llama ***circuito***. La segunda usa las aristas entre v_1 y v_2 dos veces; una trayectoria que *no* incluye la misma arista más de una vez se llama trayectoria ***simple***.

Es posible usar las potencias de la matriz de adyacencia de un grafo para obtener información acerca de las trayectorias de varias longitudes en el grafo. Considere el cuadrado de la matriz de adyacencia en la figura 3.25:

$$A^2 = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 2 & 3 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

¿Qué representan las entradas de A^2? Observe la entrada (2, 3). A partir de la definición de multiplicación matricial, se sabe que

$$(A^2)_{23} = a_{21}a_{13} + a_{22}a_{23} + a_{23}a_{33} + a_{24}a_{43}$$

La única manera en que esta expresión puede resultar en un número distinto de cero es si al menos uno de los productos $a_{2k}a_{k3}$ que constituyen la suma es distinto de cero. Pero $a_{2k}a_{k3}$ es distinto de cero si y sólo si tanto a_{2k} como a_{k3} son distintos de cero, lo que significa que hay una arista entre v_2 y v_k, así como una arista entre v_k y v_3. Por tanto, habrá una 2-trayectoria entre los vértices 2 y 3 (vía el vértice k). En el ejemplo, esto ocurre para $k =$ 1 y para $k = 2$, de modo que

$$\begin{aligned}(A^2)_{23} &= a_{21}a_{13} + a_{22}a_{23} + a_{23}a_{33} + a_{24}a_{43} \\ &= 1\cdot 1 + 1\cdot 1 + 1\cdot 0 + 0\cdot 0 \\ &= 2\end{aligned}$$

la cual dice que hay dos 2-trayectorias entre los vértices 2 y 3. (Compruebe para ver que las entradas restantes de A^2 correctamente dan 2-trayectorias en el grafo.) El argumento que acaba de darse puede generalizarse para producir el siguiente resultado, cuya demostración se deja como ejercicio 72.

Si A es la matriz de adyacencia de un grafo G, entonces la entrada (i, j) de A^k es igual al número de k-trayectorias entre los vértices i y j.

Ejemplo 3.68 ¿Cuántas 3-trayectorias hay entre v_1 y v_2 en la figura 3.25?

Solución Se necesita la entrada (1, 2) de A^3, que es el producto punto del renglón 1 de A^2 y la columna 2 de A. El cálculo produce

$$(A^3)_{12} = 3\cdot 1 + 2\cdot 1 + 1\cdot 1 + 0\cdot 0 = 6$$

de modo que hay seis 3-trayectorias entre los vértices 1 y 2, lo que puede comprobarse con facilidad.

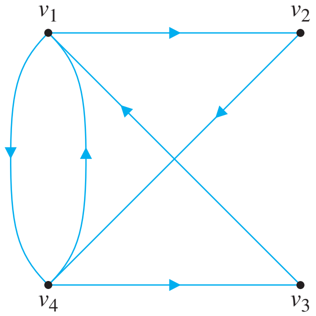


Figura 3.26
Un digrafo

En muchas aplicaciones que pueden modelarse mediante un grafo, los vértices se ordenan mediante algún tipo de relación que impone una dirección a las aristas. Por ejemplo, las aristas dirigidas pueden usarse para representar rutas de una vía en un grafo que modela una red de transporte o relaciones depredador-presa en un grafo que modela un ecosistema. Un grafo con aristas dirigidas se llama ***digrafo***. La figura 3.26 muestra un ejemplo.

Una sencilla modificación de la definición de las matrices de adyacencia permite usarlas con digrafos.

Definición Si G es un digrafo con n vértices, entonces su ***matriz de adyacencia*** es la matriz A [o $A(G)$] de $n \times n$ definida por

$$a_{ij} = \begin{cases} 1 & \text{si existe una arista del vértice } i \text{ al vértice } j \\ 0 & \text{de otro modo} \end{cases}$$

Por tanto, la matriz de adyacencia para el digrafo en la figura 3.26 es

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

No es de sorprender que la matriz de adyacencia de un digrafo no sea simétrica en general. (¿Cuándo lo sería?) El lector no debe tener dificultad para ver que A^k ahora contiene los números de k-trayectorias *dirigidas* entre vértices, donde se insiste que todas las aristas a lo largo de una trayectoria fluyen en la misma dirección. (Vea el ejercicio 72.) El siguiente ejemplo ofrece una aplicación de esta idea.

Ejemplo 3.69 Cinco tenistas (Djokovic, Federer, Nadal, Roddick y Safin) compiten en un torneo con el sistema *round-robin* en el que cada jugador compite con todos los demás una vez. El digrafo en la figura 3.27 resume los resultados. Una arista dirigida del vértice i al vértice j significa que el jugador i venció al jugador j. (Un digrafo en el que existe exactamente una arista dirigida entre cada par de vértices se llama ***torneo***.)

La matriz de adyacencia para el digrafo de la figura 3.27 es

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

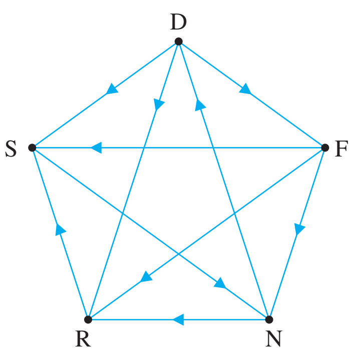


Figura 3.27
Un torneo

donde el orden de los vértices (y por tanto, los renglones y columnas de A) se determina alfabéticamente. Por tanto, Federer corresponde a renglón 2 y columna 2, por ejemplo.

Suponga que se quiere clasificar a los cinco jugadores con base en los resultados de sus partidos. Una forma de hacer esto puede ser contar el número de triunfos de cada jugador. Observe que el número de triunfos de cada jugador es justo la suma de las entradas en el renglón correspondiente; de manera equivalente, el vector que contiene todas las sumas de renglón está dado por el producto $A\mathbf{j}$, donde

$$\mathbf{j} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

En este caso se tiene

$$A\mathbf{j} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

que produce la siguiente clasificación:

Primero: Djokovic, Federer (empatados)
Segundo: Nadal
Tercero: Roddick, Safin (empatados)

¿Los jugadores empatados en esta clasificación son igualmente fuertes? Djokovic puede argumentar que, puesto que venció a Federer, él merece el primer lugar. Roddick usaría el mismo tipo de argumento para romper el empate con Safin. Sin embargo, Safin podría argumentar que él tiene dos victorias “indirectas” porque venció a Nadal, quien venció a otros *dos*; más aún, él puede puntualizar que Roddick sólo tiene *una* victoria indirecta (sobre Safin, quien venció a Nadal).

Dado que en un grupo de empates puede no haber un jugador que venciera a todos los demás en el grupo, la notación de victorias indirectas parece más útil. Más aún, una victoria indirecta corresponde a una 2-trayectoria en el digrafo, de modo que puede usar el cuadrado de la matriz de adyacencia. Para calcular tanto los triunfos como las victorias indirectas para cada jugador, se necesitan las sumas de renglón de la matriz $A + A^2$, que están dadas por

$$(A + A^2)\mathbf{j} = \left(\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 2 & 1 & 2 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 2 & 2 & 3 \\ 1 & 0 & 2 & 2 & 2 \\ 1 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 6 \\ 2 \\ 3 \end{bmatrix}$$

En consecuencia, los jugadores se clasificarían en el siguiente orden: Djokovic, Federer, Nadal, Safin, Roddick. Por desgracia, este planteamiento no garantiza romper todos los empates.

Códigos de corrección de error

En la sección 1.4 se estudiaron ejemplos de códigos de detección de error. Ahora se aborda el problema de diseñar códigos que puedan *corregir* así como detectar ciertos tipos de errores. El mensaje será un vector $\mathbf{x}$ en $\mathbb{Z}_2^k$ para alguna k, y se le codificará usando una transformación matricial $T : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$ para algún $n > k$. El vector $T(\mathbf{x})$ se llamará ***vector código***. Un ejemplo simple servirá para ilustrar el enfoque a tomar, que es una generalización de los vectores de control de paridad del ejemplo 1.37.

Ejemplo 3.70

Suponga que el mensaje es un solo dígito binario: 0 o 1. Si el mensaje se codifica al simplemente repetirlo dos veces, entonces los vectores código son [0, 0] y [1, 1]. Este código puede detectar errores individuales. Por ejemplo, si se transmite [0, 0] y ocurre un error en el primer componente, entonces se recibe [1, 0] y se detecta un error, porque este no es un vector código legal. Sin embargo, el receptor no puede corregir el error, pues [1, 0] también sería resultado de un error en el segundo componente si se hubiese transmitido [1, 1].

Este problema se puede resolver al hacer los vectores código más largos: repetir el dígito de mensaje tres veces en lugar de dos. Por tanto, 0 y 1 se codifican como [0, 0, 0] y [1, 1, 1], respectivamente. Ahora, si ocurre un solo error, no sólo se puede detectar, sino también corregir. Por ejemplo, si se recibe [0, 1, 0], entonces se sabe que debe ser el resultado de un solo error en la transmisión de [0, 0, 0], pues un solo error en [1, 1, 1] no se habría producido.

Note que el código en el ejemplo 3.70 puede lograrse mediante una transformación matricial, aunque en una particularmente trivial. Sea $G = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y defina $T : \mathbb{Z}_2 \to \mathbb{Z}_2^3$ mediante $T(\mathbf{x}) = G\mathbf{x}$. (Aquí los elementos de $\mathbb{Z}_2$ se consideran matrices de 1×1.) La matriz G se llama ***matriz generadora*** del código.

Para decir si un vector recibido es un vector código, se realiza no una, sino *dos* comprobaciones de paridad. Se requiere que el vector recibido $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$ satisfaga $c_1 = c_2 = c_3$. Estas ecuaciones pueden escribirse como un sistema lineal sobre $\mathbb{Z}_2$:

$$\begin{aligned} c_1 &= c_2 \\ c_1 &= c_3 \end{aligned} \quad \text{o} \quad \begin{aligned} c_1 + c_2 &= 0 \\ c_1 + c_3 &= 0 \end{aligned} \tag{1}$$

Si se hace $P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, entonces (1) es equivalente a $P\mathbf{c} = \mathbf{0}$. La matriz P se llama ***matriz de control de paridad*** para el código. Observe que $PG = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = O$.

Para ver cómo estas matrices entran en juego para la corrección de errores, suponga que se envía 1 como $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = [1 \quad 1 \quad 1]^T$, pero un solo error hace que se reciba como

$\mathbf{c}' = [1 \quad 0 \quad 1]^T$. Se calcula

$$P\mathbf{c}' = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \mathbf{0}$$

de modo que se sabe que $\mathbf{c}'$ no puede ser un vector código. ¿Dónde está el error? Note que $P\mathbf{c}'$ es la segunda columna de la matriz de control de paridad P; ésta dice que el error está en el segundo componente de $\mathbf{c}'$ (lo que se probará adelante en el Teorema 3.37) y permite corregir el error. (Desde luego, en este ejemplo podría encontrar el error más rápidamente sin usar matrices, pero la idea es útil.)

Para generalizar las ideas en el último ejemplo, se hacen las siguientes definiciones.

Definiciones Si $k < n$, entonces cualquier matriz de $n \times k$ de la forma $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$, donde A es una matriz de $(n - k) \times k$ sobre $\mathbb{Z}_2$, se llama ***matriz generadora estándar*** para un ***código binario*** (n, k) $T : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$. Cualquier matriz de $(n - k) \times n$ de la forma $P = [B \; I_{n-k}]$, donde B es una matriz de $(n - k) \times k$ sobre $\mathbb{Z}_2$, se llama ***matriz de comprobación de paridad estándar***. Se dice que el código tiene ***longitud*** n y ***dimensión*** k.

He aquí lo que se necesita saber: (a) ¿Cuándo es G la matriz generadora estándar para un código binario *de corrección de error*? (b) Dada G, ¿cómo se encuentra una matriz de control de paridad estándar asociada P? Es evidente que las respuestas son bastante sencillas, como se muestra mediante el siguiente teorema.

Teorema 3.37 Si $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$ es una matriz generadora estándar y $P = [B \quad I_{n-k}]$ es una matriz de control de paridad estándar, entonces P es la matriz de control de paridad asociada con G, si y sólo si $A = B$. El correspondiente código binario (n, k) es de corrección de error (sencillo) si y sólo si las columnas de P son distintas de cero y diferentes.

Antes de demostrar el teorema, considere otro ejemplo menos trivial que lo ilustra.

Ejemplo 3.71 Suponga que quiere diseñar un código de corrección de error que use tres ecuaciones de control de paridad. Dado que dichas ecuaciones dan lugar a los renglones de P, se tiene $n - k = 3$ y $k = n - 3$. Los vectores de mensaje vienen de $\mathbb{Z}_2^k$, así que se buscaría que k (y por tanto n) fuese tan largo como sea posible, con la finalidad de poder transmitir tanta información como sea posible. Por el Teorema 3.37, las n columnas de P deben ser distintas de cero y diferentes así que el máximo ocurre cuando ellas consisten de todos los $2^3 - 1 = 7$ vectores distintos de cero de $\mathbb{Z}_2^{n-k} = \mathbb{Z}_2^3$. Uno de tales candidatos es

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

Esto significa que

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

y por tanto, por el Teorema 3.37, una matriz generadora estándar para este código es

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

Como ejemplo de cómo funciona la matriz generadora, suponga que se codifica $\mathbf{x} = [0 \quad 1 \quad 0 \quad 1]^T$ para obtener el vector código

$$\mathbf{c} = G\mathbf{x} = [0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0]^T$$

Si se recibe este vector, se ve que es correcto, pues $P\mathbf{c} = \mathbf{0}$. Por otra parte, si se recibe $\mathbf{c}' = [0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0]^T$ se calcula

$$P\mathbf{c}' = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

que se reconoce como la columna 3 de P. Por tanto, el error está en el tercer componente de $\mathbf{c}'$, y al cambiarlo se recupera el vector código correcto $\mathbf{c}$. También se sabe que los primeros cuatro componentes de un vector código son el vector mensaje original, así que en este caso se decodifica $\mathbf{c}$ para obtener el original $\mathbf{x} = [0 \quad 1 \quad 0 \quad 1]^T$.

El código del ejemplo 3.71 se llama código de Hamming (7, 4). Cualquier código binario construido de esta forma se llama ***código de Hamming* (*n*, *k*)**. Observe que, por construcción, un código de Hamming (n, k) tiene $n = 2^{n-k} - 1$.

Demostración del Teorema 3.37 (A lo largo de esta demostración se denota mediante $\mathbf{a}_i$ la i-ésima columna de una matriz A.) Con P y G como en el enunciado del teorema, primero suponga que son matrices de control de paridad estándar y generadora para el mismo código binario. Por tanto, para toda $\mathbf{x}$ en $\mathbb{Z}_2^k$, $PG\mathbf{x} = \mathbf{0}$. En términos de multiplicación de bloques,

$$[B \quad I]\begin{bmatrix} I \\ A \end{bmatrix}\mathbf{x} = \mathbf{0} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{Z}_2^k$$

De manera equivalente, para toda $\mathbf{x}$ en $\mathbb{Z}_2^k$, se tiene

$$B\mathbf{x} + A\mathbf{x} = (B + A)\mathbf{x} = (BI + IA)\mathbf{x} = [B \quad I]\begin{bmatrix} I \\ A \end{bmatrix}\mathbf{x} = \mathbf{0}$$

o
$$B\mathbf{x} = A\mathbf{x}$$

Si ahora toma $\mathbf{x} = \mathbf{e}_i$, el i-ésimo vector base estándar en $\mathbb{Z}_2^k$, se ve que

$$\mathbf{b}_i = B\mathbf{e}_i = A\mathbf{e}_i = \mathbf{a}_i \quad \text{para toda } i$$

Por tanto, $B = A$.

Por el contrario, es fácil comprobar que si $B = A$, entonces $PG\mathbf{x} = \mathbf{0}$ para toda $\mathbf{x}$ en $\mathbb{Z}_2^k$ (vea el ejercicio 92).

Para ver que tal par determina un código de corrección de error si las columnas de P son distintas de cero y diferentes, sea $\mathbf{x}$ un vector mensaje en $\mathbb{Z}_2^k$ y sea el correspondiente vector código $\mathbf{c} = G\mathbf{x}$. Entonces $P\mathbf{c} = \mathbf{0}$. Suponga que hay un error en el i-ésimo componente, que resulta en el vector $\mathbf{c}'$. Se tiene que $\mathbf{c}' = \mathbf{c} + \mathbf{e}_i$. Ahora calcule

$$P\mathbf{c}' = P(\mathbf{c} + \mathbf{e}_i) = P\mathbf{c} + P\mathbf{e}_i = \mathbf{0} + \mathbf{p}_i = \mathbf{p}_i$$

que localiza el error en el i-ésimo componente.

Por otra parte, si $\mathbf{p}_i = \mathbf{0}$, entonces un error en el i-ésimo componente no se detectará (es decir, $P\mathbf{c}' = \mathbf{0}$), y si $\mathbf{p}_i = \mathbf{p}_j$, entonces no se puede determinar si ocurrió un error en el i-ésimo o el j-ésimo componente (ejercicio 93).

Richard W. Hamming (1915–1998) recibió su doctorado en matemáticas de la Universidad de Illinois en Urbana-Champaign en 1942. Sus intereses de investigación en matemáticas fueron los campos de las ecuaciones diferenciales y el análisis numérico. De 1946 a 1976, trabajó en los Laboratorios Bell, luego de lo cual se unió al cuerpo docente en la Escuela de Posgraduados Navales de Estados Unidos, en Monterey, California. En 1950 publicó su ensayo fundamental acerca de los códigos de corrección de errores, lo que proporcionó una construcción explícita para los códigos óptimos que Claude Shannon probó teóricamente posibles en 1948.

A continuación se resumen las ideas principales de esta sección.

1. Para $n > k$, una matriz G de $n \times k$ y una matriz P de $(n - k) \times n$ (con entradas en $\mathbb{Z}_2$) son una matriz generadora estándar y una matriz de control de paridad estándar, respectivamente, para un código binario (n, k) si y sólo si, en forma de bloque,

$$G = \begin{bmatrix} I_k \\ A \end{bmatrix} \quad \text{y} \quad P = [A \quad I_{n-k}]$$

para alguna matriz A de $(n - \mathrm{k}) \times k$ sobre $\mathbb{Z}_2$.
2. G codifica un vector mensaje $\mathbf{x}$ en $\mathbb{Z}_2^k$ como un vector código $\mathbf{c}$ en $\mathbb{Z}_2^n$ vía $\mathbf{c} = G\mathbf{x}$.
3. G es de corrección de error si y sólo si las columnas de P son distintas de cero y diferentes. Un vector $\mathbf{c}'$ en $\mathbb{Z}_2^n$ es un vector código si y sólo si $P\mathbf{c}' = \mathbf{0}$. En este caso, el correspondiente vector mensaje es el vector $\mathbf{x}$ en $\mathbb{Z}_2^k$ que consiste de los primeros k componentes de $\mathbf{c}'$. Si $P\mathbf{c}' \neq \mathbf{0}$, entonces $\mathbf{c}'$ no es un vector código y $P\mathbf{c}'$ es una de las columnas de P. Si $P\mathbf{c}'$ es la i-ésima columna de P, entonces el error está en el i-ésimo componente de $\mathbf{c}'$ y se puede recuperar el vector código correcto (y en consecuencia el mensaje) al cambiar este componente.

Ejercicios 3.7

Cadenas de Markov

En los ejercicios 1-4, sea $P = \begin{bmatrix} 0.5 & 0.3 \\ 0.5 & 0.7 \end{bmatrix}$ *la matriz de transición para una cadena de Markov con dos estados. Sea* $\mathbf{x}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$ *el vector de estado inicial para la población.*

1. Calcule $\mathbf{x}_1$ y $\mathbf{x}_2$.

2. ¿Qué proporción de la población del estado 1 estará en el estado 2 después de dos pasos?

3. ¿Qué proporción de la población del estado 2 estará en el estado 2 después de dos pasos?

4. Encuentre el vector de estado estacionario.

En los ejercicios 5-8, sea $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{3} & 0 \end{bmatrix}$ *la matriz de transición para una cadena de Markov con tres estados. Sea* $\mathbf{x}_0 = \begin{bmatrix} 120 \\ 180 \\ 90 \end{bmatrix}$ *el vector de estado inicial para la población.*

5. Calcule $\mathbf{x}_1$ y $\mathbf{x}_2$.

6. ¿Qué proporción de la población del estado 1 estará en el estado 1 después de dos pasos?

7. ¿Qué proporción de la población del estado 2 estará en el estado 3 después de dos pasos?

8. Encuentre el vector de estado estacionario.

9. Suponga que el clima en una región particular se comporta de acuerdo con una cadena de Markov. Específicamente, suponga que la probabilidad de que mañana será un día húmedo es de 0.662 si hoy es húmedo y de 0.250 si hoy es seco. La probabilidad de que mañana sea un día seco es de 0.750 si hoy es seco y de 0.338 si hoy es húmedo. [Este ejercicio se basa en un estudio real de lluvias en Tel Aviv durante un periodo de 27 años. Vea K. R. Gabriel y J. Neumann, "A Markov Chain Model for Daily Rainfall Occurrence at Tel Aviv", *Quarterly Journal of the Royal Meteorological Society, 88* (1962), pp. 90-95.]
   (a) Escriba la matriz de transición para esta cadena de Markov.
   (b) Si el lunes es un día seco, ¿cuál es la probabilidad de que el miércoles sea húmedo?
   (c) A largo plazo, ¿cuál será la distribución de días húmedos y secos?

10. Se han acumulado datos acerca de las estaturas de niños en relación con sus padres. Suponga que las probabilidades de que un padre alto tenga un hijo alto, de mediana estatura o bajo son 0.6, 0.2 y 0.2, respectivamente; las probabilidades de que un padre de talla media tenga un hijo alto, de estatura media o bajo son 0.1, 0.7 y 0.2, respectivamente; y las probabilidades de que un padre bajo tenga un hijo alto, mediano o bajo son 0.2, 0.4 y 0.4, respectivamente.
   (a) Escriba la matriz de transición para esta cadena de Markov.
   (b) ¿Cuál es la probabilidad de que una persona baja tenga un nieto alto?
   (c) Si 20% de la población actual es alta, 50% es de estatura media y 30% es baja, ¿cuál será la distribución en tres generaciones?
   (d) Si los datos del inciso (c) no cambian con el tiempo, ¿qué proporción de la población será alta, de estatura mediana y baja a largo plazo?

11. Un estudio de cultivos de pino piñonero en el sureste estadounidense, de 1940 a 1947, planteó la hipótesis que la producción de pinos seguía una cadena de Markov. [Vea D. H. Thomas, "A Computer Simulation Model of Great Basin Shoshonean Subsistence and Settlement Patterns", en D. L. Clarke ed., *Models in Archaeology* (Londres: Methuen, 1972).] Los datos sugirieron que si la cosecha de un año era buena, entonces la probabilidades de que el cultivo del siguiente año fuera bueno, adecuado o pobre eran de 0.08, 0.07 y 0.85, respectivamente; si la cosecha de un año era adecuada, entonces las probabilidades de que la cosecha del siguiente año fuera buena, adecuada y pobre eran de 0.09, 0.11 y 0.80, respectivamente; si la cosecha de un año era pobre, entonces las probabilidades de que la cosecha del siguiente año fuera buena, adecuada o pobre eran de 0.11, 0.05 y 0.84, respectivamente.
   (a) Escriba la matriz de transición para esta cadena de Markov.
   (b) Si la cosecha de pino piñonero fue buena en 1940, encuentre las probabilidades de una buena cosecha en los años 1941 a 1945.
   (c) A largo plazo, ¿qué proporción de las cosechas será buena, adecuada y pobre?

12. Se han programado robots para recorrer el laberinto que se muestra en la figura 3.28 y en cada unión eligen cuál camino seguir en forma aleatoria.

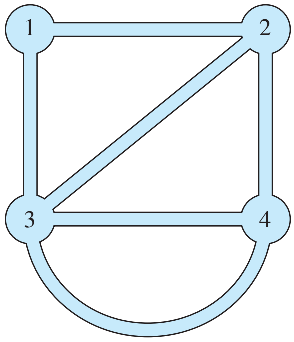


Figura 3.28

(a) Construya la matriz de transición para la cadena de Markov que modela esta situación.

(b) Suponga que comienza con 15 robots en cada unión. Encuentre la distribución de estado estacionario de robots. (Suponga que a cada robot le toma la misma cantidad de tiempo recorrer la distancia entre dos uniones adyacentes.)

13. Sea $\mathbf{j}$ un vector renglón que consiste por completo de números 1. Demuestre que una matriz P no negativa es una matriz estocástica si y sólo si $\mathbf{j}P = \mathbf{j}$.

14. (a) Demuestre que el producto de dos matrices estocásticas de 2×2 también es una matriz estocástica.

(b) Demuesrtre que el producto de dos matrices estocásticas de $n \times n$ también es una matriz estocástica.

(c) Si una matriz estocástica P de 2×2 es invertible, demuestre que P^{-1} también es una matriz estocástica.

Suponga que se quiere conocer el número promedio (o esperado) de pasos que se darán para ir del estado i al estado j en una cadena de Markov. Se puede demostrar que el siguiente cálculo responde esta pregunta: borre el renglón j-ésimo y la columna j-ésima de la matriz de transición P para obtener una nueva matriz Q. (Mantenga los renglones y columnas de Q etiquetadas como lo estaban en P.) El número esperado de pasos del estado i al estado j está dado por la suma de las entradas en la columna de $(I - Q)^{-1}$ marcada i.

15. En el ejercicio 9, si el lunes es un día seco, ¿cuál es el número esperado de días hasta un día húmedo?

16. En el ejercicio 10, ¿cuál es el número esperado de generaciones hasta que una persona baja tenga un descendiente alto?

17. En el ejercicio 11, si la cosecha de pino piñonero es adecuada un año, ¿cuál es el número esperado de años que deben pasar hasta que ocurre una buena cosecha?

18. En el ejercicio 12, a partir de cada una de las otras uniones, ¿cuál es el número esperado de movimientos hasta que un robot llegue a la unión 4?

Modelos económicos lineales

En los ejercicios 19-26, determine cuáles de las matrices son de intercambio. Para las que lo sean, encuentre un vector de precio no negativo que satisfaga la ecuación (1).

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

20. $\begin{bmatrix} 1/3 & 2/3 \\ 1/2 & 1/2 \end{bmatrix}$

21. $\begin{bmatrix} 0.4 & 0.7 \\ 0.6 & 0.4 \end{bmatrix}$

22. $\begin{bmatrix} 0.1 & 0.6 \\ 0.9 & 0.4 \end{bmatrix}$

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

24. $\begin{bmatrix} 1/2 & 1 & 0 \\ 0 & 0 & 1/3 \\ 1/2 & 0 & 2/3 \end{bmatrix}$

25. $\begin{bmatrix} 0.3 & 0 & 0.2 \\ 0.3 & 0.5 & 0.3 \\ 0.4 & 0.5 & 0.5 \end{bmatrix}$

26. $\begin{bmatrix} 0.50 & 0.70 & 0.35 \\ 0.25 & 0.30 & 0.25 \\ 0.25 & 0 & 0.40 \end{bmatrix}$

En los ejercicios 27-30, determine si la matriz de consumo dada es productiva.

27. $\begin{bmatrix} 0.2 & 0.3 \\ 0.5 & 0.6 \end{bmatrix}$

28. $\begin{bmatrix} 0.20 & 0.10 & 0.10 \\ 0.30 & 0.15 & 0.45 \\ 0.15 & 0.30 & 0.50 \end{bmatrix}$

29. $\begin{bmatrix} 0.35 & 0.25 & 0 \\ 0.15 & 0.55 & 0.35 \\ 0.45 & 0.30 & 0.60 \end{bmatrix}$

30. $\begin{bmatrix} 0.2 & 0.4 & 0.1 & 0.4 \\ 0.3 & 0.2 & 0.2 & 0.1 \\ 0 & 0.4 & 0.5 & 0.3 \\ 0.5 & 0 & 0.2 & 0.2 \end{bmatrix}$

En los ejercicios 31-34, se proporcionan una matriz de consumo C y un vector de demanda **d**. *En cada caso, encuentre un vector de producción factible* **x** *que satisfaga la ecuación (2).*

31. $C = \begin{bmatrix} 1/2 & 1/4 \\ 1/2 & 1/2 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

32. $C = \begin{bmatrix} 0.1 & 0.4 \\ 0.3 & 0.2 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

33. $C = \begin{bmatrix} 0.5 & 0.2 & 0.1 \\ 0 & 0.4 & 0.2 \\ 0 & 0 & 0.5 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$

CAS **34.** $C = \begin{bmatrix} 0.1 & 0.4 & 0.1 \\ 0 & 0.2 & 0.2 \\ 0.3 & 0.2 & 0.3 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 1.1 \\ 3.5 \\ 2.0 \end{bmatrix}$

35. Sea A una matriz de $n \times n$ y $A \geq O$. Suponga que $A\mathbf{x} < \mathbf{x}$ para alguna $\mathbf{x}$ en $\mathbb{R}^n$, $\mathbf{x} \geq \mathbf{0}$. Demuestre que $\mathbf{x} > \mathbf{0}$.

36. Sean A, B, C y D matrices de $n \times n$ y $\mathbf{x}, \mathbf{y}$ vectores en $\mathbb{R}^n$. demuestre las siguientes desigualdades:

(a) Si $A \geq B \geq O$ y $C \geq D \geq O$, entonces $AC \geq BD \geq O$.

(b) Si $A > B$ y $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} \neq \mathbf{0}$, entonces $A\mathbf{x} > B\mathbf{x}$.

Crecimiento poblacional

37. Una población con tres clases etáreas tiene una matriz de Leslie $L = \begin{bmatrix} 1 & 1 & 3 \\ 0.7 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$. Si el vector de población inicial es $\mathbf{x}_0 = \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix}$, calcule $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$.

38. Una población con cuatro clases etáreas tiene una matriz de Leslie $L = \begin{bmatrix} 0 & 1 & 2 & 5 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \end{bmatrix}$. Si el vector de población inicial es $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}$, calcule $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$.

39. Cierta especie con dos clases etáreas de 1 año de duración tiene una probabilidad de supervivencia de 80% de la clase 1 a la clase 2. Evidencia empírica demuestra que, en promedio, cada hembra pare cinco hembras por año. En consecuencia, dos posibles matrices de Leslie son

$$L_1 = \begin{bmatrix} 0 & 5 \\ 0.8 & 0 \end{bmatrix} \quad \text{y} \quad L_2 = \begin{bmatrix} 4 & 1 \\ 0.8 & 0 \end{bmatrix}$$

(a) A partir de $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$, calcule $\mathbf{x}_1, \ldots, \mathbf{x}_{10}$ en cada caso.

(b) Para cada caso, grafique el tamaño relativo de cada clase etárea contra el tiempo (como en la figura 3.23). ¿Qué sugieren sus gráficas?

40. Suponga que la matriz de Leslie para el escarabajo VW es $L = \begin{bmatrix} 0 & 0 & 20 \\ 0.1 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$. A partir de un $\mathbf{x}_0$ arbitrario, determine el comportamiento de esta población.

41. Suponga que la matriz de Leslie para el escarabajo VW es $L = \begin{bmatrix} 0 & 0 & 20 \\ s & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$. Investigue el efecto de variar la probabilidad de supervivencia s de los escarabajos jóvenes.

CAS **42.** Los caribúes se encuentran principalmente en las provincias occidentales de Canadá y el noroeste estadounidense. La vida promedio de una hembra es de alrededor de 14 años. Las tasas de natalidad y supervivencia para cada grupo etáreo se proporcionan en la tabla 3.4, que muestra que las hembras de caribú no dan a luz durante sus primeros 2 años y paren aproximadamente una cría por año durante sus años intermedios. La tasa de mortalidad para las crías jóvenes es muy alta.

Tabla 3.4

Edad (años)	Tasa de natalidad	Tasa de supervivencia
0–2	0.0	0.3
2–4	0.4	0.7
4–6	1.8	0.9
6–8	1.8	0.9
8–10	1.8	0.9
10–12	1.6	0.6
12–14	0.6	0.0

Los números de caribúes reportados en el Jasper National Park de Alberta en 1990 se muestran en la tabla 3.5. Con un CAS, prediga la población de caribúes para 1992 y 1994. Luego proyecte la población para los años 2010 y 2020. ¿Qué concluye? (¿Qué su-

Tabla 3.5 Población de caribués en el Jasper National Park, 1990

Edad (años)	Número
0–2	10
2–4	2
4–6	8
6–8	5
8–10	12
10–12	0
12–14	1

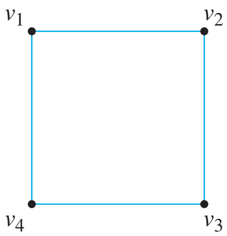
Fuente: World Wildlife Fund Canada

posiciones hace este modelo y cómo podría mejorarlo?)

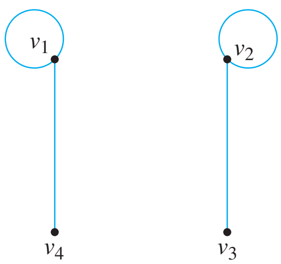
Grafos y digrafos

En los ejercicios 43-46, determine la matriz de adyacencia del grafo dado.

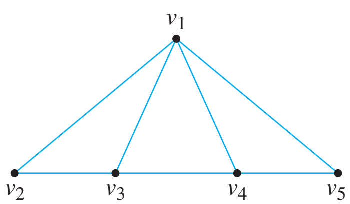
43.



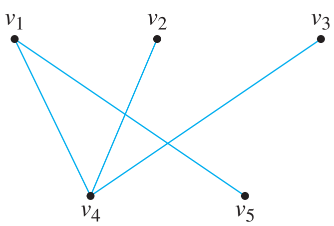
44.



45.



46.



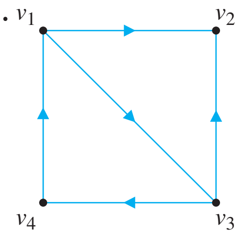
En los ejercicios 47-50, dibuje un grafo que tenga la matriz de adyacencia dada.

47. $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ **48.** $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

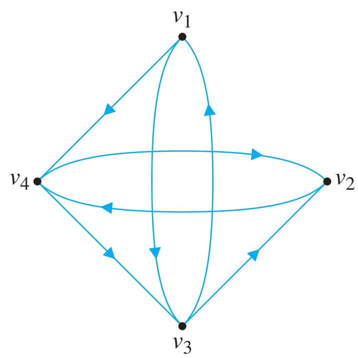
49. $\begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$ **50.** $\begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$

En los ejercicios 51-54, determine la matriz de adyacencia del digrafo dado.

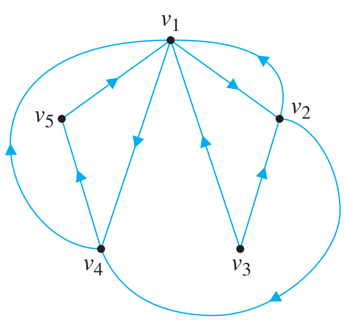
51.



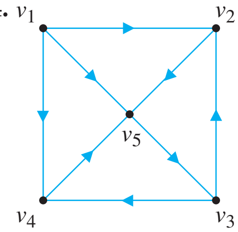
52.



53.



54.



En los ejercicios 55-58, dibuje un digrafo que tenga la matriz de adyacencia dada.

55. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}$ **56.** $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

57. $\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$ **58.** $\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 59-66, use potencias de matrices de adyacencia para determinar el número de trayectorias de la longitud especificada entre los vértices dados.

59. Ejercicio 48, longitud 2, v_1 y v_2

60. Ejercicio 50, longitud 2, v_1 y v_2

61. Ejercicio 48, longitud 3, v_1 y v_3

62. Ejercicio 50, longitud 4, v_1 y v_2

63. Ejercicio 55, longitud 2, v_1 a v_3

64. Ejercicio 55, longitud 3, v_4 a v_1

65. Ejercicio 58, longitud 3, v_4 a v_1

66. Ejercicio 58, longitud 4, v_1 a v_4

67. Sea A la matriz de adyacencia de un grafo G.

(a) Si el renglón i de A es todo ceros, ¿qué implica esto acerca de G?

(b) Si la columna j de A es toda ceros, ¿qué implica esto acerca de G?

68. Sea A la matriz de adyacencia de un digrafo D.

(a) Si el renglón i de A^2 es todo ceros, ¿qué implica esto acerca de D?

(b) Si la columna j de A^2 es toda ceros, ¿qué implica esto acerca de D?

69. La figura 3.29 es el digrafo de un torneo con seis jugadores, P_1 a P_6. Con matrices de adyacencia, clasifique a los jugadores primero al determinar sólo victorias, y luego al usar la noción de victorias combinadas y victorias indirectas, como en el ejemplo 3.69.

70. La figura 3.30 es un digrafo que representa una cadena alimenticia en un pequeño ecosistema. Una arista dirigida de a a b indica que a tiene a b como fuente de alimento. Construya la matriz de adyacencia A para este digrafo y úsela para responder las siguientes preguntas.

(a) ¿Cuál especie tiene las fuentes de alimento más directas? ¿Cómo muestra A esto?

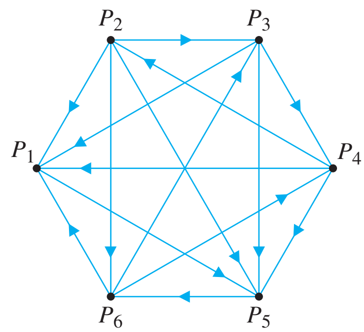


Figura 3.29

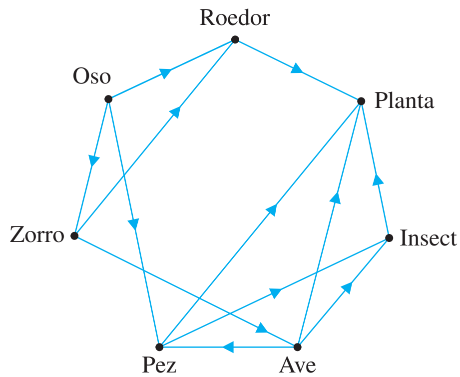


Figura 3.30

(b) ¿Cuál especie es una fuente directa de alimento para la mayoría de las otras especies? ¿Cómo muestra A esto?

(c) Si a come a b y b come a c, se dice que a tiene a c como una fuente indirecta de alimento. ¿Cómo puede usar A para determinar cuál especie tiene más fuentes indirectas de alimento? ¿Cuál especie tiene más fuentes de alimento directas e indirectas combinadas?

(d) Suponga que los contaminantes matan a las plantas en esta cadena alimenticia y quiere determinar el efecto que tendrá este cambio sobre el ecosistema. Construya una nueva matriz de adyacencia A^* a partir de A, al borrar el renglón y la columna correspondientes a planta. Repita los incisos (a) a (c) y determine cuáles especies son las más y menos afectadas por el cambio.

(e) ¿Cuál será el efecto a largo plazo de la contaminación? ¿Qué cálculos matriciales demostrarán esto?

71. Cinco personas se conectan todas mediante correo electrónico. Siempre que una de ellas escucha un trozo de chismorreo interesante, lo envía por correo electrónico a alguien más en el grupo, de acuerdo con la tabla 3.6.

(a) Dibuje el digrafo que modela esta "red de chismorreo" y encuentre su matriz de adyacencia A.

Tabla 3.6

Remitente	Destinatarios
Ann	Carla, Ehaz
Bert	Carla, Dana
Carla	Ehaz
Dana	Ann, Carla
Ehaz	Bert

(b) Defina un *paso* como el tiempo que una persona tarda en enviar correo electrónico a todos los de su lista. (Por tanto, en un paso, el chisme va de Ann tanto a Carla como a Ehaz.) Si Bert escucha un rumor, ¿cuántos pasos transcurrirán para que todos los demás escuchen el rumor? ¿Qué cálculo matricial revela esto?

(c) Si Ann escucha un rumor, ¿cuántos pasos transcurrirán para que los demás escuchen el rumor? ¿Qué cálculo matricial revela esto?

(d) En general, si A es la matriz de adyacencia de un digrafo, ¿cómo puede decir si el vértice i está conectado al vértice j mediante una trayectoria (de cierta longitud)?

[La red de chismorreo en este ejercicio recuerda la noción de "seis grados de separación" (que se encuentra en la obra y la película del mismo nombre), que sugiere que cualesquiera dos personas están conectadas mediante una trayectoria de conocidos cuya longitud es cuando mucho 6. El juego "Six Degrees of Kevin Bacon" asevera más frívolamente que todos los actores están conectados al actor Kevin Bacon en tal forma.]

72. Sea A la matriz de adyacencia de un grafo G.

(a) Por inducción, pruebe que, para toda $n \geq 1$, la entrada (i, j) de A^n es igual al número de n-trayectorias entre los vértices i y j.

(b) ¿Cómo deben modificarse el enunciado y la demostración del inciso (a) si G es un digrafo?

73. Si A es la matriz de adyacencia de un digrafo G, ¿qué representa la entrada (i, j) de AA^T si $i \neq j$?

Un grafo se llama ***bipartita*** *si sus vértices pueden subdividirse en dos conjuntos U y V tales que toda arista tiene un punto final en U y el otro punto final en V. Por ejemplo, el grafo del ejercicio 46 es bipartita con $U = \{v_1, v_2, v_3\}$ y $V = \{v_4, v_5\}$. En los ejercicios 74-77, determine si un grafo con la matriz de adyacencia dada es bipartita.*

74. La matriz de adyacencia en el ejercicio 47

75. La matriz de adyacencia en el ejercicio 50

76. La matriz de adyacencia en el ejercicio 49

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

78. (a) Demuestre que un grafo es bipartita si y sólo si sus vértices pueden etiquetarse de modo que su matriz de adyacencia pueda particionarse como

$$A = \left[\begin{array}{c|c} O & B \\ \hline B^T & O \end{array}\right]$$

(b) Con el resultado del inciso (a), demuestre que un grafo bipartita no tiene circuitos de longitud impar.

Códigos de corrección de error

79. Suponga que los cuatro vectores en $\mathbb{Z}_2^2$ se codifican al repetir el vector dos veces. Por tanto, se tiene

$$\begin{aligned} [0, 0] &\to [0, 0, 0, 0] \\ [0, 1] &\to [0, 1, 0, 1] \\ [1, 0] &\to [1, 0, 1, 0] \\ [1, 1] &\to [1, 1, 1, 1] \end{aligned}$$

Demuestre que este código no es de corrección de error.

80. Suponga que codifica los dígitos binarios 0 y 1 al repetir cada dígito cinco veces. Por tanto,

$$\begin{aligned} 0 &\to [0, 0, 0, 0, 0] \\ 1 &\to [1, 1, 1, 1, 1] \end{aligned}$$

Demuestre que este código puede corregir errores dobles.

¿Cuál es el resultado de codificar los mensajes en los ejercicios 81-83 usando el código de Hamming (7, 4) del ejemplo 3.71?

81. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ **82.** $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ **83.** $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Cuando se usa el código de Hamming (7, 4) del ejemplo 3.71, suponga que se reciben los mensajes $\mathbf{c}'$ en los ejercicios 84-86.

Aplique la matriz de control de paridad estándar a $\mathbf{c}'$ *para determinar si ocurrió un error y decodificar correctamente* $\mathbf{c}'$ *para recuperar el vector mensaje original* $\mathbf{x}$.

84. $\mathbf{c}' = [0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1]^T$

85. $\mathbf{c}' = [1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0]^T$

86. $\mathbf{c}' = [0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0]^T$

87. El código de control de paridad en el ejemplo 1.37 es un código $\mathbb{Z}_2^6 \to \mathbb{Z}_2^7$.

(**a**) Encuentre una matriz de control de paridad estándar para este código.

(**b**) Encuentre una matriz generadora estándar.

(**c**) Aplique el Teorema 3.37 para explicar por qué este código no es de corrección de error.

88. Defina un código $\mathbb{Z}_2^2 \to \mathbb{Z}_2^5$ con el uso de una matriz generadora estándar

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

(**a**) Enliste todas las cuatro palabras código.

(**b**) Encuentre la matriz de control de paridad estándar asociada para este código. ¿Este código es de corrección de error (sencillo)?

89. Defina un código $\mathbb{Z}_2^3 \to \mathbb{Z}_2^6$ con la matriz generadora estándar

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

(**a**) Mencione las ocho palabras código.

(**b**) Encuentre la matriz de control de paridad estándar asociada para este código. ¿Este código es de corrección de error (sencillo)?

90. Demuestre que el código del ejemplo 3.70 es un código de Hamming (3, 1).

91. Construya matrices de comprobación de paridad estándar y generadora para un código de Hamming (15, 11).

92. En el Teorema 3.37, demuestre que si $B = A$, entonces $PG\mathbf{x} = \mathbf{0}$ para todo $\mathbf{x}$ en $\mathbb{Z}_2^k$.

93. En el Teorema 3.37, demuestre que, si $\mathbf{p}_i = \mathbf{p}_j$, entonces no es posible determinar si ocurre un error en el componente i-ésimo o j-ésimo del vector recibido.

Repaso del capítulo

Definiciones y conceptos clave

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso:
   (a) Para cualquier matriz A, tanto AA^T como A^TA están definidas.
   (b) Si A y B son matrices tales que $AB = O$ y $A \neq O$, entonces $B = O$.
   (c) Si A, B y X son matrices invertibles tales que $XA = B$, entonces $X = A^{-1}B$.
   (d) La inversa de una matriz elemental es una matriz elemental.
   (e) La transpuesta de una matriz elemental es una matriz elemental.
   (f) El producto de dos matrices elementales es una matriz elemental.
   (g) Si A es una matriz de $m \times n$, entonces el espacio nulo de A es un subespacio de $\mathbb{R}^n$.
   (h) Todo plano en $\mathbb{R}^3$ es un subespacio bidimensional de $\mathbb{R}^3$.
   (i) La transformación $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(\mathbf{x}) = -\mathbf{x}$ es una transformación lineal.
   (j) Si $T : \mathbb{R}^4 \to \mathbb{R}^5$ es una transformación lineal, entonces existe una matriz A de 4×5 tal que $T(\mathbf{x}) = A\mathbf{x}$ para toda $\mathbf{x}$ en el dominio de T.

En los ejercicios 2-7, sea $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ *y* $B = \begin{bmatrix} 2 & 0 & -1 \\ 3 & -3 & 4 \end{bmatrix}$.

Calcule la matrices indicadas si es posible.

2. A^2B **3.** A^TB^2 **4.** $B^TA^{-1}B$

5. $(BB^T)^{-1}$ **6.** $(B^TB)^{-1}$

7. El desarrollo del producto exterior de AA^T

8. Si A es una matriz tal que $A^{-1} = \begin{bmatrix} 1/2 & -1 \\ -3/2 & 4 \end{bmatrix}$, encuentre A.

9. Si $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & -1 \\ 0 & 1 & 1 \end{bmatrix}$ y X es una matriz tal que

$AX = \begin{bmatrix} -1 & -3 \\ 5 & 0 \\ 3 & -2 \end{bmatrix}$, encuentre X.

10. Si es posible, exprese la matriz $A = \begin{bmatrix} 1 & 2 \\ 4 & 6 \end{bmatrix}$ como un producto de matrices elementales.

11. Si A es una matriz cuadrada tal que $A^3 = O$, demuestre que $(I - A)^{-1} = I + A + A^2$.

12. Encuentre una factorización LU de $A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}$.

13. Encuentre bases para el espacio renglón, espacio columna y espacio nulo de $A = \begin{bmatrix} 2 & -4 & 5 & 8 & 5 \\ 1 & -2 & 2 & 3 & 1 \\ 4 & -8 & 3 & 2 & 6 \end{bmatrix}$.

14. Suponga que las matrices A y B son equivalentes en renglón. ¿Tienen el mismo espacio renglón? ¿Por qué sí o por qué no? ¿A y B tienen el mismo espacio columna? ¿Por qué sí o por qué no?

15. Si A es una matriz invertible, explique por qué A y A^T deben tener el mismo espacio nulo. ¿Esto es verdad si A es una matriz cuadrada no invertible? Explique.

16. Si A es una matriz cuadrada cuyos renglones suman el vector cero, explique por qué A no puede ser invertible.

17. Sea A una matriz de $m \times n$ con columnas linealmente independientes. Explique por qué A^TA debe ser una matriz invertible. ¿AA^T también debe ser invertible? Explique.

18. Encuentre una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $T\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$.

19. Encuentre la matriz estándar de la transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ que corresponda a una rotación de 45° en sentido contrario al de las manecillas del reloj en torno al origen seguida por una proyección sobre la recta $y = -2x$.

20. Suponga que $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal y suponga que $\mathbf{v}$ es un vector tal que $T(\mathbf{v}) \neq \mathbf{0}$, pero $T^2(\mathbf{v}) = \mathbf{0}$ (donde $T^2 = T \circ T$). Demuestre que $\mathbf{v}$ y $T(\mathbf{v})$ son linealmente independientes.