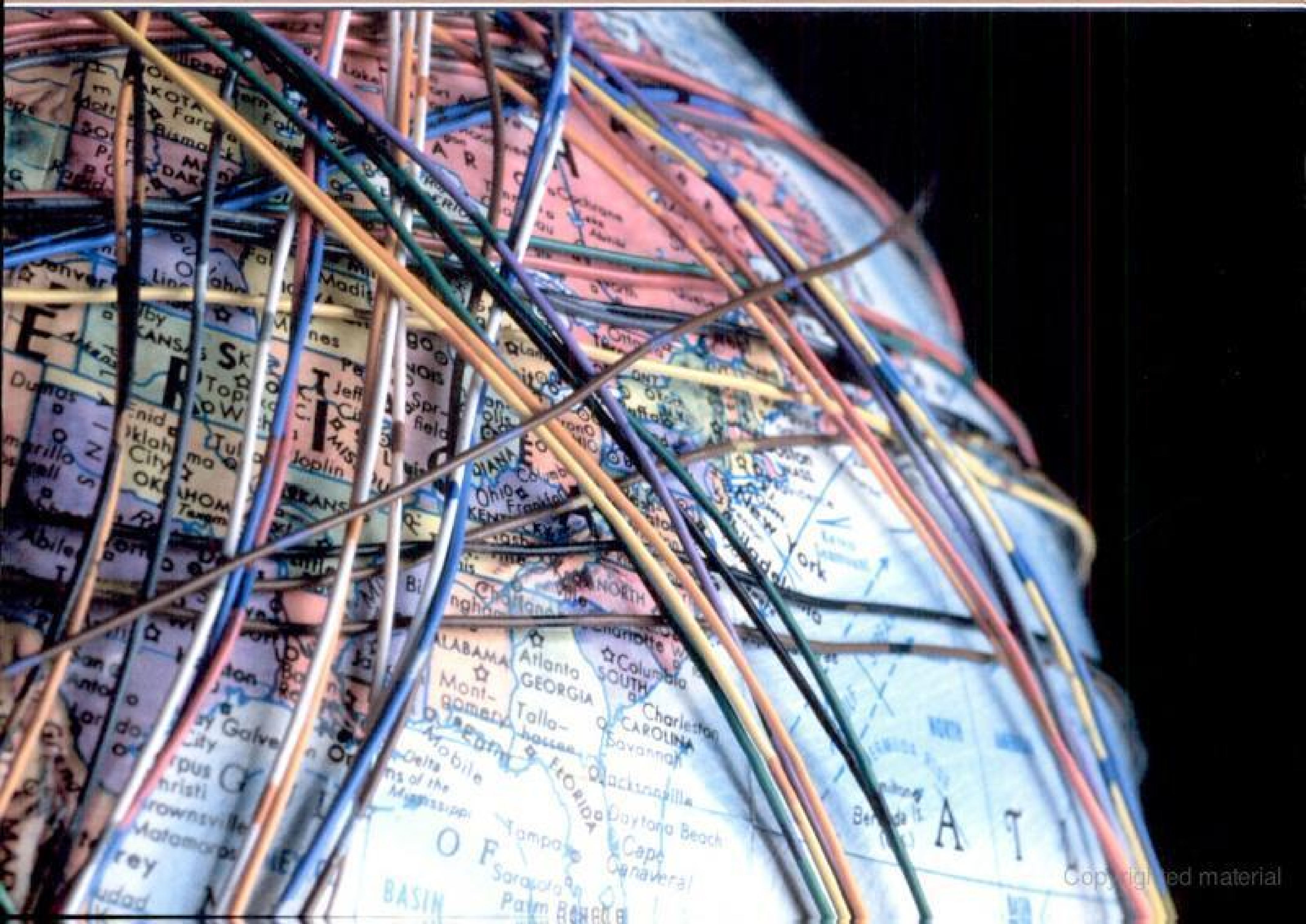
CENGAGE Learning

Segunda edición

ÁLGEBRA LINEAL

Una introducción moderna

David Poole



CENGAGE
Learning™

Álgebra lineal. *Una introducción moderna.*
Segunda edición
David Poole

Presidente de Cengage Learning Latinoamérica:
Javier Arellano Gutiérrez

Director general México y Centroamérica:
Héctor Enrique Galindo Iturribarría

Director editorial Latinoamérica:
José Tomás Pérez Bonilla

Director de producción:
Raúl D. Zendejas Espejel

Editora de desarrollo:
Rocío Cabañas Chávez

Editora de producción:
Abril Vega Orozco

Corporativo Santa Fe
Av. Santa Fe, núm. 505, piso 12
Col. Cruz Manca, Santa Fe
C.P. 05349, México, D.F.

Traducido del libro *Linear Algebra. A Modern Introduction*, 2nd ed.
Publicado en inglés por
Thomson Learning Brooks Cole © 2006
ISBN 0-534-99845-3
Datos para catalogación bibliográfica:
Poole, David
Álgebra lineal. Una introducción moderna.
Segunda edición
ISBN-13: 978-970-686-595-3
ISBN-10: 970-686-595-0

Visite nuestro sitio en:
http://latinoamerica.cengage.com

Impreso en México
1 2 3 4 5 6 7 11 10 09 08

Impreso en Cosegraf; oct. del 2005
Progreso No. 10 Col. Centro
Ixtapaluca Edo. De México

Contenido

Capítulo 3 Matrices 134

Capítulo 4 Eigenvalores y eigenvectores 252

Capítulo 5 Ortogonalidad 363

Características

Estilo claro de exposición

El texto se encuentra escrito en un estilo sencillo, directo y familiar. Tanto como fue posible, empleé "lenguaje coloquial matemático" en lugar de depender excesivamente de la notación matemática. Sin embargo, todas las demostraciones que se proporcionan son completamente rigurosas, y el apéndice A contiene una introducción a la notación matemática para aquellos que deseen racionalizar su propia escritura. Casi siempre ejemplos concretos preceden a los teoremas, seguidos por ejemplos adicionales y aplicaciones. Este flujo (de lo específico a lo general, y a la inversa) es consistente a lo largo del libro.

Conceptos clave presentados desde el principio

Muchos estudiantes encuentran dificultades cuando un curso de álgebra lineal se traslada de lo computacional (resolución de sistemas de ecuaciones lineales, manipulación de vectores y matrices) a lo teórico (conjuntos de espacios generados, independencia lineal, subespacios, base y dimensión). Este libro presenta la totalidad de los conceptos clave del álgebra lineal al principio, en un marco concreto, antes de volverlos a ver con toda generalidad. Los conceptos vectoriales tales como el producto punto, longitud, ortogonalidad y proyección se analizan primero en el capítulo 1 en el entorno concreto de $\mathbb{R}^2$ y $\mathbb{R}^3$, antes de que aparezcan las nociones más generales de producto interno, norma y proyección ortogonal en los capítulos 5 y 7. Del mismo modo, se hace un tratamiento concreto de los conjuntos de espacios generados y la independencia lineal, en el capítulo 2, antes de su generalización a los espacios vectoriales en el capítulo 6. Los conceptos fundamentales de subespacio, base y dimensión aparecen primero en el capítulo 3 cuando se explican los espacios renglón, columna y nulo de una matriz; no es sino hasta el capítulo 6 que se da un tratamiento general a estas ideas. En el capítulo 4, se introducen los eigenvalores y eigenvectores y son explorados para las matrices de 2×2 antes de que aparezcan sus homólogos de $n\times n$. Para el principio del capítulo 4, todos los conceptos fundamentales del álgebra lineal se han presentado, con ejemplos computacionales concretos para apoyarlos. Cuando estas ideas aparecen con toda su generalidad más adelante en el libro, los estudiantes han tenido tiempo de emplearlas y por lo tanto ya no se sienten tan intimidados por ellas.

Énfasis en los vectores y la geometría

Conforme a la filosofía de que el álgebra lineal trata principalmente acerca de vectores, este libro pone énfasis en la intuición geométrica. Por consiguiente, el primer capítulo trata sobre vectores, y desarrolla muchos conceptos que aparecen una y otra vez a lo largo del texto. Conceptos tales como ortogonalidad, proyección y combinación lineal se encuentran en el capítulo 1, como también un detallado tratamiento de las líneas y los planos en $\mathbb{R}^3$ que proporciona una percepción fundamental en la solución de sistemas de ecuaciones lineales. Este énfasis en los vectores, la geometría y la visualización se encuentra en todo el texto. En el capítulo 3, se introducen las transformaciones lineales como transformaciones matriciales, con muchos ejemplos geométricos, antes de que se estudien las transformaciones lineales en el capítulo 6. En el capítulo 4, se presentan los eigenvalores haciendo uso de "eigenimágenes" como auxiliares visuales. La demostración del teorema de Perron se proporciona primero de manera heurística y posteriormente de manera formal, empleando en ambos casos un argumento geométrico. La geometría de los sistemas dinámicos lineales refuerza y resume el material sobre eigenvalores y eigenvectores. En el capítulo 5, las proyecciones ortogonales, los complementos ortogonales de los subespacios y el proceso Gram-Schmidt son todos presentados en el ámbito concreto de $\mathbb{R}^3$ antes de ser generalizados a $\mathbb{R}^n$ y, en el capítulo 7, a los espacios

aparecen en la sección 4.2, donde son utilizados para encontrar los polinomios característicos de matrices pequeñas. Este "curso intensivo" sobre determinantes contiene todo el material esencial que los estudiantes necesitan, incluyendo una demostración opcional pero elemental del teorema de expansión de Laplace. La exploración "Aplicaciones geométricas de los determinantes" contiene varios resultados útiles e interesantes y puede ser un buen proyecto (los profesores que deseen hacer una cobertura más detallada de los determinantes pueden ver una parte de esta exploración en clase). La sección 4.3 presenta la teoría básica de los eigenvalores y los eigenvectores; la sección 4.4 trata sobre el importante tema de la diagonalización. El ejemplo 4.29 acerca de potencias de matrices vale la pena tratarlo en clase. El método de potencias y sus variantes, que se estudian en la sección 4.5, son opcionales, y es recomendable que todos los estudiantes lo conocieran; en un curso aplicado este tema debería tratarse en detalle. El teorema de discos de Gerschgorin puede estudiarse de manera independiente del resto de la sección 4.5. Las cadenas de Markov y el modelo de Leslie de crecimiento poblacional vuelven a aparecer en la sección 4.6. Aunque la demostración del teorema de Perron-Frobenius es opcional, el teorema mismo debería por lo menos ser mencionado (como el más sólido teorema de Perron-Frobenius), puesto que explica *por qué* deberíamos esperar un eigenvalor positivo único con un eigenvector positivo correspondiente en estas aplicaciones. Las aplicaciones sobre relaciones de recurrencia y ecuaciones diferenciales conectan el álgebra lineal con las matemáticas discretas y el cálculo, respectivamente. Si el grupo cuenta con buenos fundamentos de cálculo, es posible abarcar la exponencial de matrices. El tema final de los sistemas dinámicos lineales discretas retoma y resume muchas de las ideas del capítulo 4, observándolas bajo una nueva luz geométrica. A los estudiantes les gustará leer cómo los eigenvectores pueden ayudar en el ranking de los equipos deportivos de los sitios en la Web. Esta viñeta puede dar pie a un proyecto o a una actividad de enriquecimiento.

Véase la página 283

Véanse las páginas 322, 327

Véase la página 353

Capítulo 5: Ortogonalidad

La exploración introductoria, "Sombras sobre una pared", es matemáticas en su mejor expresión: la exploración toma un concepto conocido (la proyección de un vector sobre otro vector) y lo generaliza de una manera útil (la proyección de un vector sobre un subespacio: un plano). Asimismo, descubre previamente algunas propiedades adicionales. La sección 5.1 contiene los resultados básicos acerca de conjuntos de vectores ortogonales y ortonormales que en lo sucesivo serán utilizados una y otra vez. En la sección 5.2 se exponen dos conceptos del capítulo 1: el complemento ortogonal de un subespacio y la proyección ortogonal de un vector sobre un subespacio. El teorema de descomposición ortogonal cobra importancia aquí y ayuda a establecer el proceso de Gram-Schmidt. También nótese la rápida demostración del teorema del Rango. El proceso de Gram-Schmidt se trata en detalle en la sección 5.3, junto con la extremadamente importante factorización *QR*. Las dos exploraciones que siguen dan una idea general de cómo la factorización *QR* se calcula en la práctica y cómo se le puede emplear para aproximar eigenvalores. La sección 5.4 acerca de la diagonalización ortogonal de matrices simétricas (reales) es necesaria para las aplicaciones que siguen. Esta sección también contiene el teorema espectral, uno de los puntos culminantes de la teoría del álgebra lineal. Las aplicaciones de la sección 5.5 incluyen códigos actuales, formas cuadráticas y gráficas de ecuaciones cuadráticas. Siempre incluyo al menos la última en mi curso, puesto que amplía el conocimiento que los estudiantes ya tienen acerca de las secciones cónicas.

Véase la página 363

Véanse las páginas 393, 395

Véanse las páginas 405, 411, 418

Capítulo 6: Espacios vectoriales

La secuencia Fibonacci reaparece en la sección 6.0, sin importar que los alumnos ya la hayan visto. El propósito de esta exploración es mostrar que el concepto de espacio vectorial (sección 3.5) pueden ser usado fructíferamente en un nuevo entorno. Puesto

Véase la página 431

Al estudiante

—Con la venia de su Majestad
—dijo el Conejo—,
¿por dónde empiezo?
—Lo prudente es comenzar por el principio
—dijo el rey con mucha gravedad—,
*y sigue de largo hasta llegar al final [...]**

—Lewis Carroll
Alicia en el país de las maravillas, 1865

El álgebra lineal es una asignatura apasionante. Está llena de resultados interesantes y tiene aplicaciones hacia otras disciplinas y conexiones hacia otras áreas de las matemáticas. El *Student and Study Guide and Solutions Manual* (sólo en inglés) contiene detallados consejos acerca de cómo emplear de la mejor manera este libro: a continuación se exponen algunas sugerencias generales.

El álgebra lineal tiene varias facetas: existen *técnicas computacionales*, *conceptos* y *aplicaciones*. Uno de los objetivos de este libro es ayudarle a dominar todas las facetas del tema y observar la interacción entre ellas. En consecuencia, es importante que usted lea y comprenda cada sección del texto antes de intentar abordar los ejercicios de esa sección. Si usted solamente lee los ejemplos que están relacionados con los ejercicios que hayan sido asignados como tareas, se perderá de mucho. No se preocupe si, para comprender algo, tiene que leerlo más de una vez. Tenga a la mano lápiz y calculadora a medida que esté leyendo. Deténgase para resolver los ejemplos usted mismo o para completar los cálculos faltantes. El icono ➲ en el margen indica un sitio donde debería hacer una pausa y pensar en lo que haya leído hasta ese momento.

Las respuestas a la mayoría de los ejercicios computacionales con numeración impar se encuentran al final del libro. Resista la tentación de consultar una respuesta antes de haber terminado un ejercicio. Y recuerde que incluso si su respuesta difiere de la que se da en el libro, aún puede ser correcta; existe más de una forma correcta para expresar algunas de las soluciones. Por ejemplo, un valor de $1>\sqrt{2}$ también puede expresarse como $\sqrt{2}>2$, y el conjunto de todos los múltiplos del vector $\begin{bmatrix} 3 \\ 1/2 \end{bmatrix}$ es el mismo que el conjunto de todos los múltiplos escalares de $\begin{bmatrix} 6 \\ 1 \end{bmatrix}$.

A medida que usted encuentre nuevos conceptos, intente relacionarlos con los ejemplos que ya conoce. Escriba en su forma completa demostraciones y soluciones a los ejercicios de manera lógica y secuencial, haciendo uso de oraciones completas. Repase lo que haya escrito para ver si tiene sentido. Mejor todavía, si puede, haga que un amigo de su grupo lea lo que usted ha escrito: si no tiene sentido para otro lector, existe la posibilidad de que no tenga sentido, y punto.

*México, Grupo Editorial Tomo, 2002. Traducción de Roberto Mares, página 201.

Problema 5 Escriba la serie de movimientos de Ana por medio de la notación $[a, b]$. Supongamos que ella comienza en el origen $(0, 0)$ sobre los ejes coordenados. Explique cómo se pueden encontrar las coordenadas del punto de la cuadrícula correspondiente a cada uno de sus movimientos *sin observar el papel cuadriculado*. Si los ejes fueran trazados de manera diferente, de modo que el punto de arranque de Ana no fuera el origen sino el punto $(2, 3)$, ¿cuáles serían las coordenadas del punto final de ella?

Aunque es sencillo, este juego introduce varios conceptos que serán útiles en nuestro estudio de los vectores. Las siguientes tres secciones consideran a los vectores desde los puntos de vista geométrico y algebraico, partiendo, como en el juego de la pista de carreras, del plano.

1.1 La geometría y el álgebra de los vectores

Vectores en el plano

Iniciemos tomando en cuenta que el plano cartesiano tiene los ya conocidos ejes x y y. Un ***vector*** es un *segmento de recta dirigido* que corresponde a un *desplazamiento* de un punto A hacia otro punto B (véase la figura 1.2).

El vector que va desde A hasta B se denota por medio de $\overrightarrow{AB}$; al punto A se le conoce como ***punto inicial***, o ***cola***, mientras que al punto B se le denomina ***punto terminal***, o ***cabeza***. Con frecuencia, los vectores se representan con una letra minúscula en negritas; por ejemplo, como $\mathbf{v}$.

El conjunto de todos los puntos en el plano corresponde al conjunto de todos los vectores cuyas colas se encuentran en el origen O. Para cada punto A, corresponde el vector $\mathbf{a} = \overrightarrow{OA}$; a cada vector $\mathbf{a}$ con cola en O, le pertenece una cabeza A (los vectores de esta forma casi siempre se conocen como *vectores de posición*).

Es común representar esos vectores usando coordenadas. Por ejemplo, en la figura 1.3, $A = (3, 2)$ y escribimos el vector $\mathbf{a} = \overrightarrow{OA} = [3, 2]$ con el empleo de corchetes (paréntesis cuadrados). De manera semejante, los otros vectores de la figura 1.3 son

$$\mathbf{b} = [-1, 3] \quad \text{y} \quad \mathbf{c} = [2, -1]$$

Las coordenadas individuales (3 y 2 en el caso de $\mathbf{a}$) son llamadas las ***componentes*** del vector. En ocasiones, se dice que un vector es un *par ordenado* de números reales. El orden es importante puesto que, por ejemplo, $[3, 2] \neq [2, 3]$. En general, dos vectores son iguales si y sólo si sus respectivas componentes son iguales. De este modo, $[x, y] = [1, 5]$ implica que $x = 1$ y que $y = 5$.

Por lo general, se recomienda utilizar ***vectores columna*** en lugar de (o además de) ***vectores renglón***. Otra representación de $[3, 2]$ es $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. (El punto importante es que las

El plano cartesiano recibe ese nombre en honor del filósofo y matemático fránces René Descartes (1596-1650), cuya introducción de las coordenadas permitió que los problemas *geométricos* fueran resueltos con ayuda de técnicas *algebraicas*.

La palabra *vector* proviene de la raíz latina *vector*, *vectoris* que significa "que conduce". Un vector se forma cuando un punto es desplazado (o "conducido") a una distancia determinada en una dirección dada. Vistos de otra manera, los vectores "conducen" dos clases de información: su longitud y su dirección.

Al escribir los vectores a mano, es difícil indicarlos en negritas. Algunas personas prefieren escribir $\vec{v}$ para expresar el vector denotado en la impresión por $\mathbf{v}$, pero en la mayoría de los casos es aceptable el empleo de una v minúscula ordinaria. Por lo general, resultará evidente a partir del contexto los casos en los que se maneje la letra que denote un vector.

La palabra *componente* se deriva de las palabras en latín *co*, que significa "junto con" y *ponere*, que significa "poner". De este modo, un vector se "com-pone" o se "forma al mismo tiempo" que sus componentes.

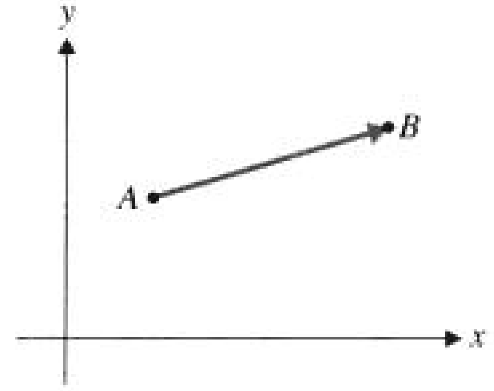


Figura 1.2

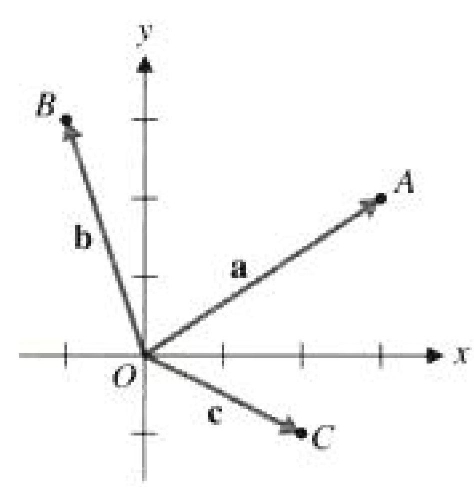


Figura 1.3

componentes están *ordenadas*). En capítulos posteriores, se verá que los vectores columna son mejores desde el punto de vista computacional; por ahora, intente captar el uso de ambas representaciones.

Puede ocurrir que en realidad usted no pueda dibujar el vector $[0, 0] = \overrightarrow{OO}$ desde el origen hasta sí mismo. No obstante, es un vector perfectamente válido y tiene un nombre especial: el ***vector cero***, el cual se denota con un **0**.

$\mathbb{R}^2$ se pronuncia "erre dos".

El conjunto de todos los vectores con dos componentes se indica con $\mathbb{R}^2$ (donde $\mathbb{R}$ denota el conjunto de los números reales del cual se eligen las componentes de los vectores en $\mathbb{R}^2$). De esta manera, $[-1, 3.5]$, $[\sqrt{2}, \pi]$, y $[\frac{5}{3}, 4]$ se encuentran todos en $\mathbb{R}^2$.

Retomemos el juego de la pista de carreras. Tratemos de relacionar los conceptos anteriores con los vectores cuyas colas *no* se encuentran en el origen. La raíz etimológica de la palabra *vector*, del verbo "conducir", nos puede proporcionar una orientación. El vector $[3, 2]$ puede ser interpretado como sigue: comienza en el origen O, viaja 3 unidades hacia la derecha, luego 2 unidades hacia arriba, finalizando en P. El mismo desplazamiento se puede aplicar a otros puntos iniciales. La figura 1.4 muestra dos desplazamientos equivalentes, representados por los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$.

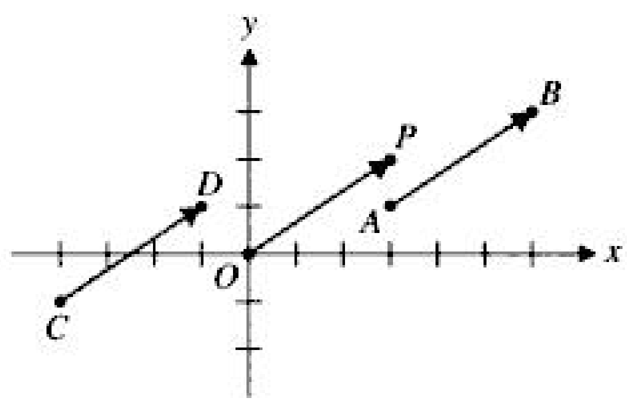


Figura 1.4

Cuando se refiere a los vectores por sus coordenadas, éstos están siendo considerados *analíticamente*.

Definimos dos vectores como *iguales* si tienen la misma longitud y la misma dirección. De este modo, $\overrightarrow{AB} = \overrightarrow{CD}$ en la figura 1.4 (aun cuando tienen diferentes puntos inicial y final, representan el mismo desplazamiento). Geométricamente, dos vectores son iguales si uno puede obtenerse mediante el corrimiento (o *traslación*) del otro de forma paralela a sí mismo hasta que los dos vectores coincidan. En términos de componentes, en la figura 1.4 tenemos que $A = (3, 1)$ y $B = (6, 3)$. Nótese que el vector $[3, 2]$ que registra el desplazamiento es precisamente la diferencia de las respectivas componentes:

$$\overrightarrow{AB} = [3, 2] = [6 - 3, 3 - 1]$$

De manera similar, $\quad \overrightarrow{CD} = [-1 - (-4), 1 - (-1)] = [3, 2]$

y entonces $\overrightarrow{AB} = \overrightarrow{CD}$, como se esperaba.

Se dice que un vector como $\overrightarrow{OP}$ con su cola en el origen se encuentra en ***posición estándar***. La explicación precedente revela que cada vector puede ser dibujado como un vector en posición estándar y, a la inversa, un vector en posición estándar puede volver a trazarse (mediante traslación) de forma que su cola se localice en cualquier punto en el plano.

Ejemplo 1.1

Si $A = (-1, 2)$ y $B = (3, 4)$, encuentre $\overrightarrow{AB}$ y vuelva a trazarlo (a) en posición estándar y (b) con su cola en el punto $C = (2, -1)$.

Solución Calculamos $\overrightarrow{AB} = [3 - (-1), 4 - 2] = [4, 2]$. Si $\overrightarrow{AB}$ se traslada posteriormente a $\overrightarrow{CD}$, donde $C = (2, -1)$, entonces debemos obtener que $D = (2 + 4, -1 + 2) = (6, 1)$ (véase la figura 1.5).

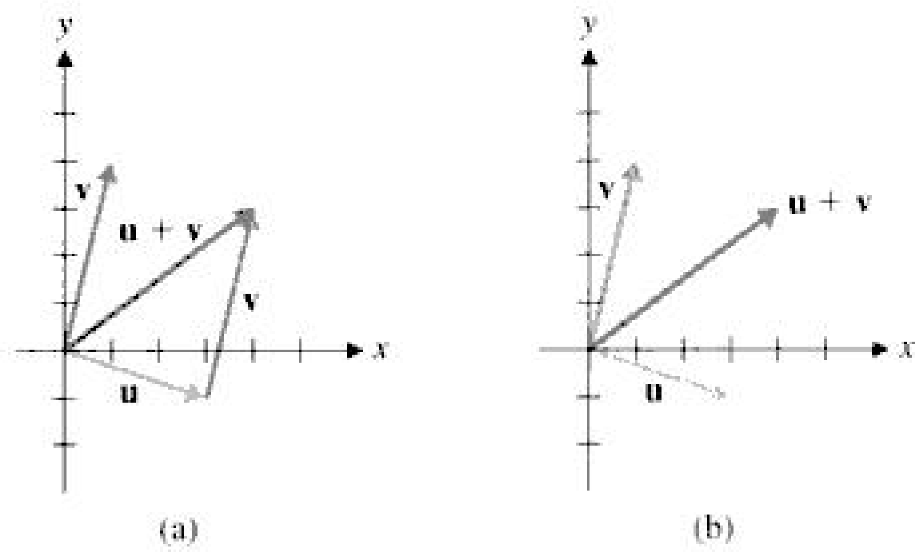


Figura 1.10

La segunda operación básica de vectores es la ***multiplicación por escalares***. Dado un vector **v** y un número real *c*, el ***múltiplo escalar*** *c***v** es el vector originado al multiplicar cada componente de **v** por *c*. Por ejemplo, $3[-2, 4] = [-6, 12]$. En general,

$$c\mathbf{v} = c\,[v_1, v_2] = [cv_1, cv_2]$$

Geométricamente, $c\mathbf{v}$ es una versión "escalada" de **v**.

Ejemplo 1.3 Si $\mathbf{v} = [-2, 4]$, calcule y trace $2\mathbf{v}$, $\frac{1}{2}\mathbf{v}$ y $-2\mathbf{v}$.

Solución Calculamos como se presenta a continuación:

$$\begin{aligned}2\mathbf{v} &= [2(-2), 2(4)] = [-4, 8]\\ \tfrac{1}{2}\mathbf{v} &= [\tfrac{1}{2}(-2), \tfrac{1}{2}(4)] = [-1, 2]\\ -2\mathbf{v} &= [-2(-2), -2(4)] = [4, -8]\end{aligned}$$

Estos vectores se ilustran en la figura 1.11.

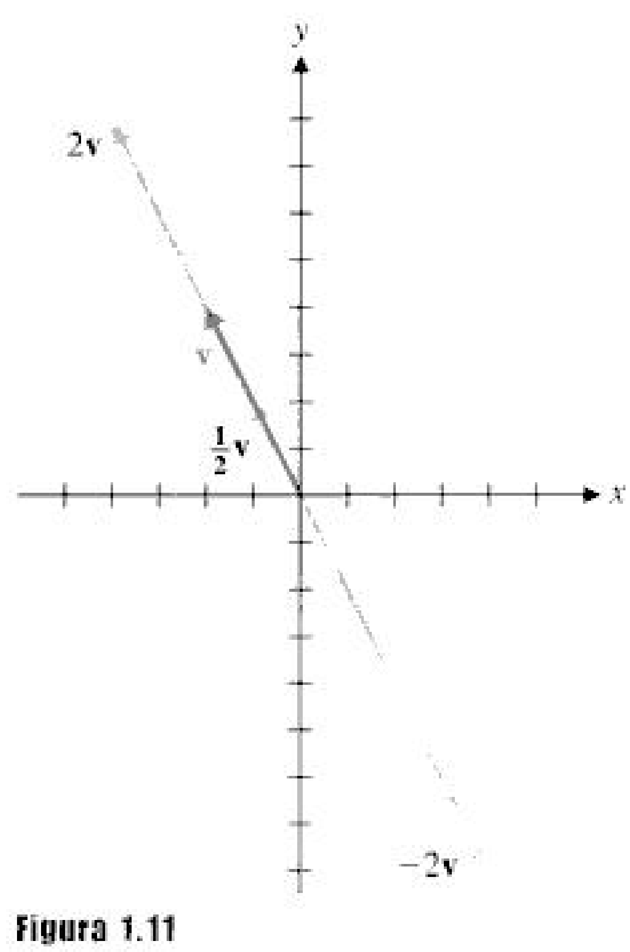


Figura 1.11

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Por la propiedad (b) del teorema 1.1, podemos escribir sin ambigüedad alguna $\mathbf{u} + \mathbf{v} + \mathbf{w}$ sin paréntesis, puesto que podemos agrupar los sumandos en cualquier forma que nos satisfaga. Por (a), también podemos reacomodar los sumandos (por ejemplo, en la forma $\mathbf{w} + \mathbf{u} + \mathbf{v}$) si así lo elegimos. Además, las sumas de cuatro o más vectores pueden ser calculadas sin importar el orden o agrupación. En general, si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ son vectores en $\mathbb{R}^n$, escribiremos esas sumas sin paréntesis:

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k$$

El ejemplo siguiente ilustra el uso del teorema 1.1 al efectuar cálculos algebraicos con vectores.

Ejemplo 1.5

Sean $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{x}$ representaciones de vectores en $\mathbb{R}^n$.

(a) Simplifique $3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a})$.

(b) Si $5\mathbf{x} - \mathbf{a} = 2(\mathbf{a} + 2\mathbf{x})$, resuelva para $\mathbf{x}$ en términos de $\mathbf{a}$.

Solución Daremos ambas soluciones con detalle, haciendo referencia a todas las propiedades del teorema 1.1 que utilicemos. Es una buena práctica justificar todos los pasos las primeras veces que se haga este tipo de cálculo. No obstante, una vez que se dominen las propiedades de los vectores, es aceptable omitir algunos de los pasos intermedios para ahorrar tiempo y espacio.

(a) Comenzamos con la inserción de los paréntesis.

$$\begin{aligned}
3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= (3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a})) + 2(\mathbf{b} - \mathbf{a}) \\
&= (3\mathbf{a} + (-2\mathbf{a} + 5\mathbf{b})) + (2\mathbf{b} - 2\mathbf{a}) && \text{(a), (e)} \\
&= ((3\mathbf{a} + (-2\mathbf{a})) + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(b)} \\
&= ((3 + (-2))\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(f)} \\
&= (1\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) \\
&= ((\mathbf{a} + 5\mathbf{b}) + 2\mathbf{b}) - 2\mathbf{a} && \text{(b), (h)} \\
&= (\mathbf{a} + (5\mathbf{b} + 2\mathbf{b})) - 2\mathbf{a} && \text{(b)} \\
&= (\mathbf{a} + (5 + 2)\mathbf{b}) - 2\mathbf{a} && \text{(f)} \\
&= (7\mathbf{b} + \mathbf{a}) - 2\mathbf{a} && \text{(a)} \\
&= 7\mathbf{b} + (\mathbf{a} - 2\mathbf{a}) && \text{(b)} \\
&= 7\mathbf{b} + (1 - 2)\mathbf{a} && \text{(f), (h)} \\
&= 7\mathbf{b} + (-1)\mathbf{a} \\
&= 7\mathbf{b} - \mathbf{a}
\end{aligned}$$

Es evidente el motivo por el que acordamos omitir algunos de esos pasos. En la práctica, se admite simplificar esta serie de pasos como

$$\begin{aligned}
3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= 3\mathbf{a} + 5\mathbf{b} - 2\mathbf{a} + 2\mathbf{b} - 2\mathbf{a} \\
&= (3\mathbf{a} - 2\mathbf{a} - 2\mathbf{a}) + (5\mathbf{b} + 2\mathbf{b}) \\
&= -\mathbf{a} + 7\mathbf{b}
\end{aligned}$$

o, incluso, hacer la mayor parte del cálculo mentalmente.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

1.2 Longitud y ángulo: El producto punto

Es muy fácil reformular los importantes conceptos geométricos de longitud, distancia y ángulo en términos de vectores. Hacerlo así nos permitirá emplearlos en entornos más generales que $\mathbb{R}^2$ y $\mathbb{R}^3$. En capítulos posteriores, estas sencillas herramientas geométricas serán utilizadas para resolver una amplia variedad de problemas que surgen en las aplicaciones, ¡incluso cuando en apariencia no tengamos geometría en absoluto!

El producto punto

Las versiones vectoriales de longitud, distancia y ángulo pueden describirse a través del empleo de la noción del producto punto de dos vectores.

Definición 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 el ***producto punto*** $\mathbf{u} \cdot \mathbf{v}$ de $\mathbf{u}$ y $\mathbf{v}$ está definido por

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

En otras palabras, $\mathbf{u} \cdot \mathbf{v}$ es la suma de los productos de las componentes correspondientes de $\mathbf{u}$ y $\mathbf{v}$. Es importante observar un par de cosas acerca de este "producto" que acabamos de definir: en primer lugar, $\mathbf{u}$ y $\mathbf{v}$ deben tener el mismo número de componentes. En segundo, el producto punto $\mathbf{u} \cdot \mathbf{v}$ es un *número*, no otro vector. (Esto es por lo que, en ocasiones, $\mathbf{u} \cdot \mathbf{v}$ se denomina el ***producto escalar*** de $\mathbf{u}$ y $\mathbf{v}$.) El producto punto de los vectores en $\mathbb{R}^n$ es un caso especial e importante de la noción más general del ***producto interno***, el cual se explicará en el capítulo 7.

Ejemplo 1.8

Calcule $\mathbf{u} \cdot \mathbf{v}$ cuando $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}$.

Solución $\mathbf{u} \cdot \mathbf{v} = 1 \cdot (-3) + 2 \cdot 5 + (-3) \cdot 2 = 1$

Nótese que si hubiéramos calculado $\mathbf{v} \cdot \mathbf{u}$ en el ejemplo 1.8, habríamos obtenido

$$\mathbf{v} \cdot \mathbf{u} = (-3) \cdot 1 + 5 \cdot 2 + 2 \cdot (-3) = 1$$

En general, es evidente que $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, puesto que los productos individuales de las componentes se conmutan. Esta propiedad de conmutatividad es una de las propiedades el producto punto que usaremos de manera constante. Las propiedades principales del producto punto se encuentran resumidas en el teorema 1.2.

Observe en la figura 1.25 que estos vectores sirven para localizar los ejes coordenados positivos en $\mathbb{R}^2$ y $\mathbb{R}^3$.

En general, en $\mathbb{R}^n$, definimos vectores unitarios $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, donde $\mathbf{e}_i$ tiene un 1 en su *i*-ésima componente y ceros en todas las demás partes. Estos vectores aparecen en forma constante en el álgebra lineal y se les conoce como ***vectores unitarios estándar***.

Ejemplo 1.12

Normalice el vector $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

Solución $\|\mathbf{v}\| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14}$, de manera que un vector unitario con la misma dirección de $\mathbf{v}$ está dado por

$$\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v} = (1/\sqrt{14})\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{14} \\ -1/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$$

Debido a que la propiedad (b) del teorema 1.3 describe cómo se comporta la longitud con respecto a la multiplicación por escalares, una curiosidad natural puede llevarnos a preguntar si la longitud y la suma vectorial son compatibles. Sería muy útil que tuviéramos una identidad como $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$, pero para casi cualquier selección de vectores $\mathbf{u}$ y $\mathbf{v}$ esto resulta ser falso. [Véase el ejercicio 42(a).] Sin embargo, no todo está perdido, porque si reemplazamos el signo = por el signo $\leq$, la desigualdad resultante es verdadera. La demostración de este famoso y trascendente resultado (la desigualdad del triángulo) depende de otra importante desigualdad (la desigualdad de Cauchy-Schwarz), que demostraremos y analizaremos con más detalle en el capítulo 7.

Teorema 1.4 **La desigualdad de Cauchy-Schwarz**

Para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$,

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$$

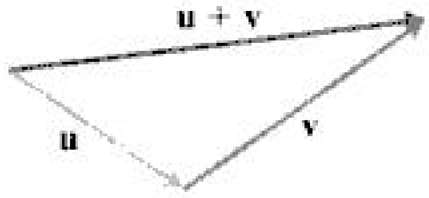


Figura 1.26
Desigualdad del triángulo

Véanse los ejercicios 55 y 56 por los enfoques algebraico y geométrico para la demostración de esta desigualdad.

En $\mathbb{R}^2$ o $\mathbb{R}^3$, donde podemos hacer uso de la geometría, es claro en un diagrama como el de la figura 1.26 que $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ para todos los vectores $\mathbf{u}$ y $\mathbf{v}$. A continuación, de manera más general, demostraremos que ello es verdad.

Teorema 1.5 **La desigualdad del triángulo**

Para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

calculadora o computadora para aproximar el ángulo deseado por medio de la función coseno inverso.

- La derivación de la fórmula para el coseno del ángulo entre dos vectores es válida sólo en $\mathbb{R}^2$ o $\mathbb{R}^3$, puesto que depende de un hecho geométrico: la ley de los cosenos. En cambio, en $\mathbb{R}^n$, para $n > 3$, la fórmula se puede considerar como una *definición*. Esto tiene sentido, puesto que la desigualdad de Cauchy-Schwartz implica que $\left|\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right| \le 1$, de modo que $\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$ abarca desde -1 hasta 1, precisamente como lo hace la función coseno.

Vectores ortogonales

La palabra *ortogonal* se deriva de los vocablos griegos *orthos*, que significa "derecho", y *gonía*, "ángulo". Por lo tanto, ortogonal quiere decir literalmente "en ángulo recto". El equivalente en latín es *rectangular*.

El concepto de perpendicularidad es fundamental para la geometría. Cualquiera que estudie esta disciplina se percata con rapidez de la importancia y utilidad de los ángulos rectos. Ahora, generalizaremos la idea de perpendicularidad a los vectores en $\mathbb{R}^n$, donde se le conoce como ***ortogonalidad***.

En $\mathbb{R}^2$ o $\mathbb{R}^3$, dos vectores distintos de cero $\mathbf{u}$ y $\mathbf{v}$ son perpendiculares si el ángulo θ entre ellos es un ángulo recto; es decir, si $\theta = \pi/2$ radianes, o 90°. Así, $\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = \cos 90° = 0$, y se sigue que $\mathbf{u}\cdot\mathbf{v} = 0$. Esto origina la siguiente definición.

Definición Dos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ son ***ortogonales*** entre sí, si $\mathbf{u}\cdot\mathbf{v} = 0$.

Puesto que $\mathbf{0}\cdot\mathbf{v} = 0$ para todo vector $\mathbf{v}$ en $\mathbb{R}^n$, el vector cero es ortogonal a todo vector.

Ejemplo 1.16 En $\mathbb{R}^3$, $\mathbf{u} = [1, 1, -2]$ y $\mathbf{v} = [3, 1, 2]$ son ortogonales, ya que $\mathbf{u}\cdot\mathbf{v} = 3 + 1 - 4 = 0$.

Si se aplica la notación de la ortogonalidad, se consigue una prueba sencilla del teorema de Pitágoras, válida en $\mathbb{R}^n$.

Teorema 1.6 **Teorema de Pitágoras**

Para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ si y sólo si $\mathbf{u}$ · $\mathbf{v}$ son ortogonales.

Demostración Del ejemplo 1.9, tenemos $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2$ para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$. Inmediatamente se sigue que $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ si y sólo si $\mathbf{u}\cdot\mathbf{v} = 0$. Véase la figura 1.32.

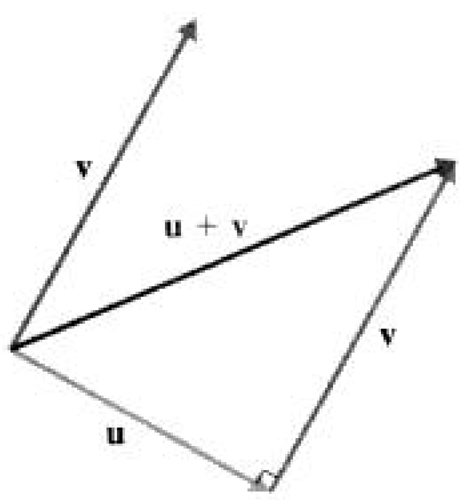


Figura 1.32
El teorema de Pitágoras

El concepto de ortogonalidad es uno de los más importantes y útiles del álgebra lineal, y, con frecuencia, se presenta de maneras sorprendentes. Aunque en el capítulo 5 se da un tratamiento detallado del tema, lo encontraremos en muchas ocasiones antes de éste. Un problema en el cual es obvio que desempeña un papel es en el de hallar la distancia desde un punto a una línea, donde "trazar una perpendicular" es un paso común.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

la forma trigonométrica del área de un triángulo: $\mathcal{A} = \frac{1}{2}\|\mathbf{u}\|\|\mathbf{v}\|\operatorname{sen}\theta$ (Podemos utilizar la identidad $\operatorname{sen}\theta = \sqrt{1 - \cos^2\theta}$ para encontrar sen θ).

En los ejercicios 40 y 41, calcule el área del triángulo con los vértices dados con base en ambos métodos.

40. $A = (1, -1), B = (2, 2), C = (4, 0)$

41. $A = (3, -1, 4), B = (4, -2, 6), C = (5, 0, 2)$

En los ejercicios 42 y 43, halle todos los valores del escalar k para los cuales los dos vectores son ortogonales.

42. $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} k+1 \\ k-1 \end{bmatrix}$ **43.** $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} k^2 \\ k \\ -3 \end{bmatrix}$

44. Describa todos los vectores $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ que sean ortogonales a $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

45. Describa todos los vectores $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ que sean ortogonales a $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$.

46. ¿En qué condiciones se cumple lo siguiente para los vectores **u** y **v** en $\mathbb{R}^2$ o $\mathbb{R}^3$?

(a) $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$ **(b)** $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| - \|\mathbf{v}\|$

47. Demuestre el teorema 1.2(b).

48. Demuestre el teorema 1.2(d).

En los ejercicios 49 a 51, demuestre la propiedad enunciada acerca de la distancia entre vectores.

49. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$ para todos los vectores **u** y **v**

50. $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w})$ para todos los vectores **u**, **v** y **w**

51. $d(\mathbf{u}, \mathbf{v}) = 0$ si y sólo si $\mathbf{u} = \mathbf{v}$

52. Demuestre que $\mathbf{u} \cdot c\mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$ para todos los vectores **u** y **v** en $\mathbb{R}^n$ y todos los escalares c.

53. Demuestre que $\|\mathbf{u} - \mathbf{v}\| \geq \|\mathbf{u}\| - \|\mathbf{v}\|$ para todos los vectores **u** y **v** en $\mathbb{R}^n$. (*Sugerencia:* reemplace **u** por $\mathbf{u} - \mathbf{v}$ en la desigualdad del triángulo.)

54. Supongamos que sabemos que $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$. ¿Esto implica que $\mathbf{v} = \mathbf{w}$? Si es así, proporcione una demostración que sea válida en $\mathbb{R}^n$; de otra forma, ofrezca un *contraejemplo* (es decir, un conjunto *específico* de vectores **u**, **v** y **w** para los cuales $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$, pero $\mathbf{v} \neq \mathbf{w}$).

55. Compruebe que $(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$ para todos los vectores **u** y **v** en $\mathbb{R}^n$.

56. **(a)** Demuestre que $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$ para todos los vectores **u** y **v** en $\mathbb{R}^n$.

(b) Dibuje un diagrama que represente **u**, **v**, $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ en $\mathbb{R}^2$ y utilice (a) para deducir un resultado acerca de paralelogramos.

57. Demuestre que $\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$ para todos los vectores **u** y **v** en $\mathbb{R}^n$.

58. **(a)** Demuestre que $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\|$ si y sólo si **u** y **v** son ortogonales.

(b) Dibuje un diagrama que muestre **u**, **v**, $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ en $\mathbb{R}^2$ y emplee el inciso (a) para deducir un resultado acerca de paralelogramos.

59. **(a)** Compruebe que $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ son ortogonales en $\mathbb{R}^n$ si y sólo si $\|\mathbf{u}\| = \|\mathbf{v}\|$. (*Sugerencia:* véase el ejercicio 47.)

(b) Dibuje un diagrama que represente **u**, **v**, $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ en $\mathbb{R}^2$ y (a) para deducir un resultado acerca de paralelogramos.

60. Si $\|\mathbf{u}\| = 2$, $\|\mathbf{v}\| = \sqrt{3}$, y $\mathbf{u} \cdot \mathbf{v} = 1$, encuentre $\|\mathbf{u} + \mathbf{v}\|$.

61. Demuestre que no hay vectores **u** y **v** tales que $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$, y $\mathbf{u} \cdot \mathbf{v} = 3$.

62. **(a)** Demuestre que si **u** es ortogonal tanto a **v** como a **w**, entonces **u** es ortogonal a $\mathbf{v} + \mathbf{w}$.

(b) Verifique que si **u** es ortogonal tanto a **v** como a **w**, entonces **u** es ortogonal a $s\mathbf{v} + t\mathbf{w}$ para todo escalar s y t.

63. Compruebe que **u** es ortogonal a $\mathbf{v} - \operatorname{proy}_{\mathbf{u}}(\mathbf{v})$ para todos los vectores **u** y **v** en $\mathbb{R}^n$, donde $\mathbf{u} \neq \mathbf{0}$.

64. **(a)** Pruebe que $\operatorname{proy}_{\mathbf{u}}(\operatorname{proy}_{\mathbf{u}}(\mathbf{v})) = \operatorname{proy}_{\mathbf{u}}(\mathbf{v})$.

(b) Demuestre que $\operatorname{proy}_{\mathbf{u}}(\mathbf{v} - \operatorname{proy}_{\mathbf{u}}(\mathbf{v})) = \mathbf{0}$.

(c) Explique (a) y (b) de manera geométrica.

65. La desigualdad de Cauchy-Schwarz $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$ es equivalente a la desigualdad que obtenemos al elevar al cuadrado ambos lados: $(\mathbf{u} \cdot \mathbf{v})^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2$.

(a) En $\mathbb{R}^2$, con $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, la desigualdad se convierte en

$$(u_1v_1 + u_2v_2)^2 \leq (u_1^2 + u_2^2)(v_1^2 + v_2^2)$$

Demuestre esta desigualdad algebraicamente. (*Sugerencia:* reste el miembro izquierdo del miembro derecho y muestre que la diferencia debe ser necesariamente no negativa.)

(b) Compruebe el análogo de (a) en $\mathbb{R}^3$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

1.3 Líneas y planos

Todos estamos familiarizados con la ecuación de una línea recta en el plano cartesiano. Ahora, queremos considerar las líneas en $\mathbb{R}^2$ desde un punto de vista vectorial. Las ideas que obtendremos de este enfoque nos permitirán hacer una generalización a las líneas en $\mathbb{R}^3$ y posteriormente a los planos en $\mathbb{R}^3$. Gran parte del álgebra lineal que consideraremos en capítulos posteriores tiene su origen en la geometría simple de las líneas y los planos; la capacidad de visualizar éstos y de pensar geométricamente acerca de un problema nos será de gran utilidad.

Líneas en $\mathbb{R}^2$ y $\mathbb{R}^3$

En el plano *xy*, la forma general de la ecuación de una recta es $ax + by = c$. Si $b \neq 0$, entonces la ecuación puede reexpresarse como $y = -(a/b)x + c/b$, lo cual tiene la forma $y = mx + k$. [Ésta es la forma con intercepción al origen; m es la pendiente de la recta, y el punto con coordenadas $(0, k)$ es la intercepción en y.] Para entender los vectores en este contexto, consideremos un ejemplo.

Ejemplo 1.19

La recta ℓ con ecuación $2x + y = 0$ se muestra en la figura 1.50. Es una recta con pendiente -2 que pasa por el origen. El lado izquierdo de la ecuación se encuentra en la forma de un producto punto; de hecho, si hacemos $\mathbf{n} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, entonces tenemos que la ecuación se transforma en $\mathbf{n} \cdot \mathbf{x} = 0$. El vector $\mathbf{n}$ es perpendicular a la recta, es decir, es *ortogonal* a cualquier vector $\mathbf{x}$ que sea paralelo a la recta (figura 1.51), y se le conoce como ***vector normal*** a la línea recta. La ecuación $\mathbf{n} \cdot \mathbf{x} = 0$ es la *forma normal* de la ecuación de ℓ.

Otra manera de pensar acerca de la recta es imaginarse una partícula moviéndose a lo largo de ella. Supongamos que la partícula se ubica inicialmente en el origen al tiempo $t = 0$, y que se mueve a lo largo de la recta de manera que su coordenada x cambia en 1 unidad por segundo. Entonces, para $t = 1$ la partícula se localiza en $(1, -2)$, para $t = 1.5$ se encuentra en $(1.5, -3)$ y, si permitimos que haya valores negativos de t (es decir, consideramos dónde estuvo la partícula en el pasado), para $t = -2$ se halla (o se

La palabra latina *norma* se refiere a una escuadra de carpintero utilizada para dibujar ángulos rectos. Por lo tanto, un vector *normal* es aquel que es perpendicular a alguna cosa, casi siempre, a un plano.

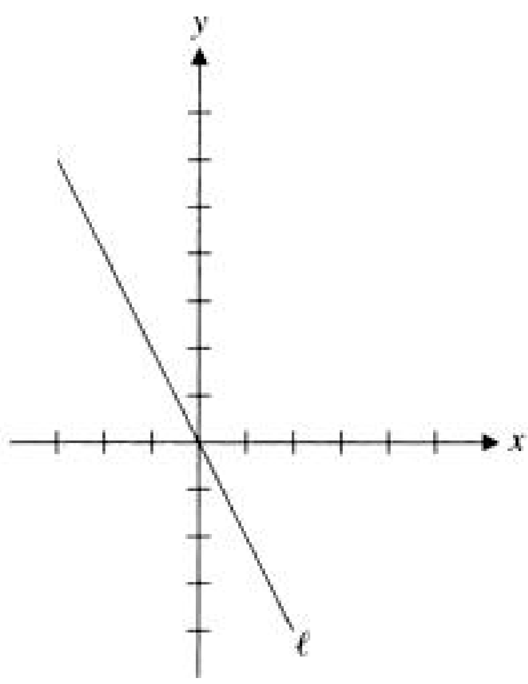


Figura 1.50
La recta $2x + y = 0$

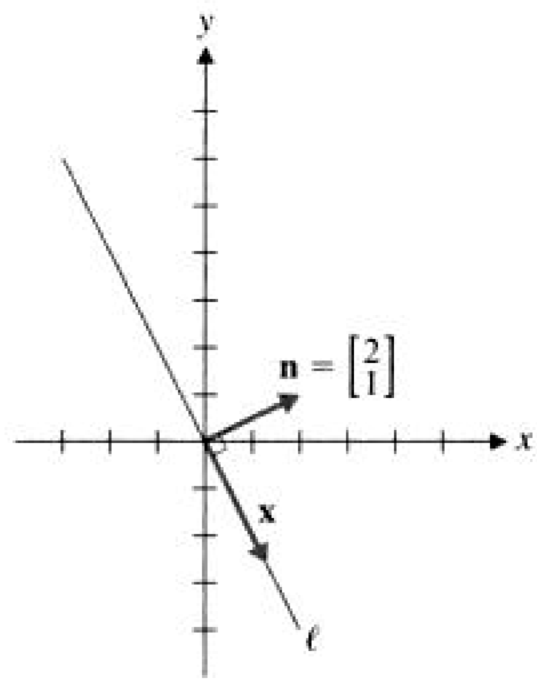


Figura 1.51
Vector normal $\mathbf{n}$

- Intuitivamente, sabemos que una línea es un objeto *unidimensional.* La idea de "dimensión" será aclarada en los capítulos 3 y 6, pero por el momento asumamos que este concepto parece concordar con el hecho de que la forma vectorial de la ecuación de una línea requiere de *un* parámetro.

Ejemplo 1.22

A menudo se escucha la expresión "un par de puntos determinan una recta". Encuentre una ecuación vectorial de la línea ℓ en $\mathbb{R}^3$ determinada por los puntos $P = (-1, 5, 0)$ y $Q = (2, 1, 1)$

Solución Podemos seleccionar cualquier punto sobre ℓ para $\mathbf{p}$, de modo que empleamos P (Q también podría ser una buena opción).

Un vector de dirección conveniente es $\mathbf{d} = \overrightarrow{PQ} = \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$ (o cualquier múltiplo escalar de éste). De esta manera, obtenemos

$$\begin{aligned} \mathbf{x} &= \mathbf{p} + t\mathbf{d} \\ &= \begin{bmatrix} -1 \\ 5 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} \end{aligned}$$

Planos en $\mathbb{R}^3$

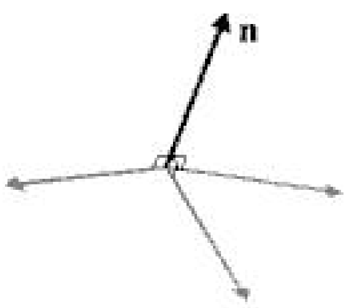


Figura 1.56
$\mathbf{n}$ es ortogonal a un número infinito de vectores

La siguiente pregunta que deberíamos responder es: ¿cómo se generaliza a $\mathbb{R}^3$ la forma general de la ecuación de una recta? Podemos conjeturar razonablemente que si $ax + by = c$ es la forma general de la ecuación de una recta en $\mathbb{R}^2$, entonces $ax + by + cz = d$ podría representar una recta en $\mathbb{R}^3$. En forma normal, esta ecuación sería $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$, donde $\mathbf{n}$ es un vector normal a la recta y $\mathbf{p}$ corresponde a un punto sobre la misma.

Para comprobar si ésta es una hipótesis razonable, pensemos acerca del caso especial de la ecuación $ax + by + cz = 0$. En forma normal, se transforma en $\mathbf{n} \cdot \mathbf{x} = 0$, donde $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Sin embargo, el conjunto de todos los vectores $\mathbf{x}$ que satisfacen esta ecuación es el conjunto de todos los vectores ortogonales a $\mathbf{n}$. Como se muestra en la figura 1.56, esta propiedad la tienen vectores en un número infinito de direcciones, determinando una familia de *planos* paralelos. Por lo tanto, nuestra conjetura fue incorrecta: parece que $ax + by + cz = d$ es la ecuación de un plano (no de una recta) en $\mathbb{R}^3$.

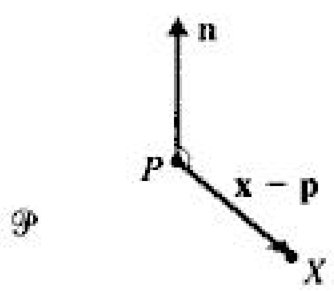


Figura 1.57
$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$

Expliquemos lo anterior en términos más precisos. Todo plano $\mathcal{P}$ en $\mathbb{R}^3$ puede ser determinado al especificar un punto $\mathbf{p}$ sobre $\mathcal{P}$ y un vector $\mathbf{n}$ distinto de cero normal a $\mathcal{P}$ (figura 1.57). De este modo, si $\mathbf{x}$ representa un punto arbitrario sobre $\mathcal{P}$, tenemos $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$ o $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$. Si $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, entonces, en términos de componentes, la ecuación se convierte en $ax + by + cz = d$ (donde $d = \mathbf{n} \cdot \mathbf{p}$).

Definición La ***forma normal de la ecuación de un plano*** $\mathcal{P}$ en $\mathbb{R}^3$ es

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0 \quad \text{o} \quad \mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$$

donde $\mathbf{p}$ es un punto específico sobre $\mathcal{P}$ y $\mathbf{n} \neq \mathbf{0}$ es un vector normal para $\mathcal{P}$.

La ***forma general de la ecuación de*** $\mathcal{P}$ es $ax + by + cz = d$, donde $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ es un vector normal para $\mathcal{P}$.

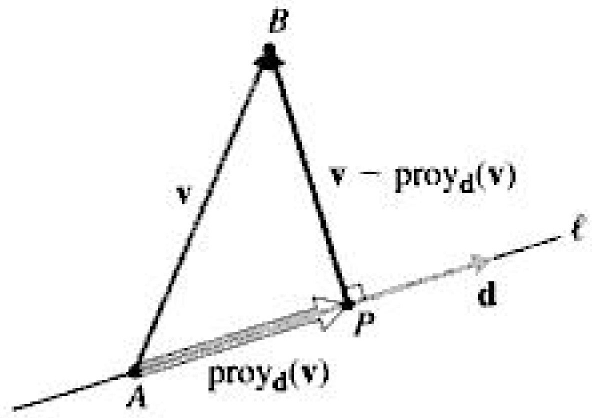


Figura 1.61
$d(B, \ell) = \|\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v})\|$

Paso 2: La proyección de **v** sobre **d** es

$$\begin{aligned}
\text{proy}_{\mathbf{d}}(\mathbf{v}) &= \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\
&= \left(\frac{(-1)\cdot(-2) + 1\cdot(-1) + 0\cdot 1}{(-1)^2 + 1 + 0}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\
&= \tfrac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}
\end{aligned}$$

Paso 3: El vector que queremos es

$$\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v}) = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ -\frac{3}{2} \\ 1 \end{bmatrix}$$

Paso 4: La distancia $d(B, \ell)$ desde B hasta ℓ es

$$\|\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v})\| = \left\|\begin{bmatrix} -\frac{3}{2} \\ -\frac{3}{2} \\ 1 \end{bmatrix}\right\|$$

Si utilizamos el teorema 1.3(b) para simplificar el cálculo, obtenemos que

$$\begin{aligned}
\|\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v})\| &= \tfrac{1}{2}\left\|\begin{bmatrix} -3 \\ -3 \\ 2 \end{bmatrix}\right\| \\
&= \tfrac{1}{2}\sqrt{9 + 9 + 4} \\
&= \tfrac{1}{2}\sqrt{22}
\end{aligned}$$

Nota

- En términos de nuestra anotación anterior, $d(B, \ell) = d(\mathbf{v}, \text{proy}_{\mathbf{d}}(\mathbf{v}))$.

La figura 1.64 sugiere una manera de emplear los vectores para localizar el punto R sobre $\mathcal{P}$ que esté más cercano a Q.

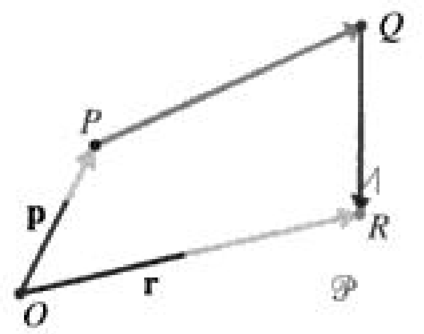


Figura 1.64
$\mathbf{r} = \mathbf{p} + \overrightarrow{PQ} + \overrightarrow{QR}$

33. Establezca el punto R sobre $\mathcal{P}$ que esté más cercano a Q en el ejercicio 29.

34. Encuentre el punto R sobre $\mathcal{P}$ más cercano a Q en el ejercicio 30.

En los ejercicios 35 y 36, halle la distancia entre las líneas paralelas.

35. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} -2 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix} + t\begin{bmatrix} -2 \\ 3 \end{bmatrix}$

36. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 37 y 38, determine la distancia entre los planos paralelos.

37. $2x + y - 2z = 0$ y $2x + y - 2z = 5$

38. $x + y + z = 1$ y $x + y + z = 3$

39. Compruebe la ecuación 3 de la página 40.

40. Demuestre la ecuación 4 de la página 41.

41. Verifique que, en $\mathbb{R}^2$, la distancia entre las rectas paralelas con ecuaciones $\mathbf{n} \cdot \mathbf{x} = c_1$ y $\mathbf{n} \cdot \mathbf{x} = c_2$ está dada por $\dfrac{|c_1 - c_2|}{\|\mathbf{n}\|}$.

42. Demuestre que la distancia entre los planos paralelos con ecuaciones $\mathbf{n} \cdot \mathbf{x} = d_1$ y $\mathbf{n} \cdot \mathbf{x} = d_2$ está dada por $\dfrac{|d_1 - d_2|}{\|\mathbf{n}\|}$.

Si dos planos no paralelos $\mathcal{P}_1$ y $\mathcal{P}_2$ tienen vectores normales $\mathbf{n}_1$ y $\mathbf{n}_2$ y θ es el ángulo entre $\mathbf{n}_1$ y $\mathbf{n}_2$, entonces definimos el ángulo entre $\mathcal{P}_1$ y $\mathcal{P}_2$ como θ o $180° - \theta$, el que sea un ángulo agudo (figura 1.65)

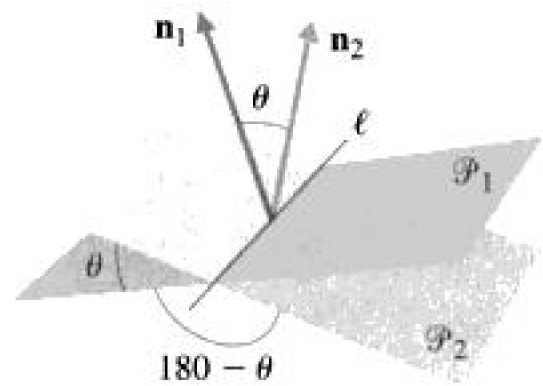


Figura 1.65

En los ejercicios 43 y 44, encuentre el ángulo agudo entre los planos con las ecuaciones dadas.

43. $x + y + z = 0$ y $2x + y - 2z = 0$

44. $3x - y + 2z = 5$ y $x + 4y - z = 2$

En los ejercicios 45 y 46, demuestre que el plano y la línea con las ecuaciones dadas se interceptan, y luego localice el ángulo agudo de la intersección entre ellos.

45. El plano dado por $x + y + 2z = 0$ y la línea dada por

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

46. El plano dado por $4x - y - z = 6$ y la línea dada por

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

En los ejercicios 47 y 48 explore un método para el problema de encontrar la proyección de un vector sobre un plano. Como muestra la figura 1.66, si $\mathcal{P}$ es un plano a través del origen en $\mathbb{R}^3$ con vector normal

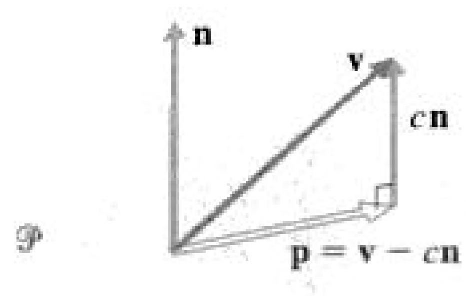


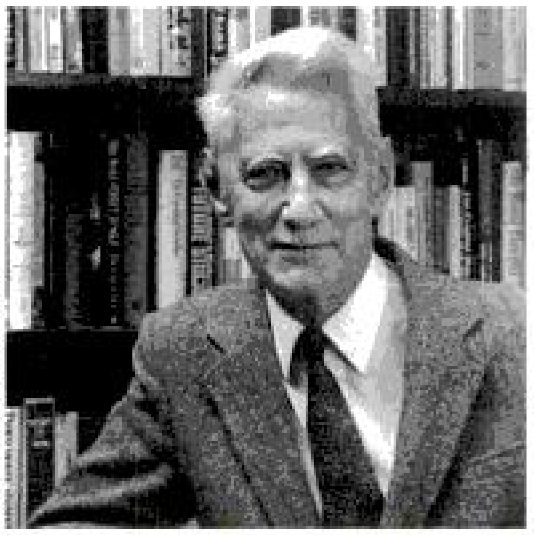
Figura 1.66
Proyección sobre un plano

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

1.4 Vectores de código y aritmética modular

A través de la historia, las personas han transmitido información por medio de códigos. En ocasiones, se intenta disfrazar el mensaje que se envía, tal como cuando cada letra de una palabra se reemplaza por una letra diferente de acuerdo con una regla de sustitución. Aunque fascinantes, estos códigos secretos, o claves, no son de interés aquí; son el centro de atención del campo de la *criptografía*. En lugar de ello, nos concentraremos en los códigos que se emplean cuando los datos deben trasmitirse en forma electrónica.



Lucent Technologies' Bell Labs

La teoría moderna de los códigos se originó con el trabajo del matemático y científico computacional estadounidense Claude Shannon (1916-2001), cuya tesis en 1937 demostró el papel que el álgebra podría desempeñar en el diseño y análisis de circuitos eléctricos. Shannon contribuyó posteriormente a la formación del campo de la *teoría de la información* y proporcionó la base teórica para lo que se conoce ahora como códigos de corrección de errores.

Un ejemplo común de un código de esta naturaleza es el código Morse, con su sistema de puntos y rayas. El advenimiento de las computadoras digitales en el siglo XX generó la necesidad de transmitir cantidades masivas de datos de manera rápida y precisa. Las computadoras están diseñadas para codificar datos como secuencias de "ceros" y "unos". Muchos avances tecnológicos recientes dependen de los códigos y los encontramos cotidianamente sin estar conscientes de ellos: comunicaciones vía satélite, reproductores de discos compactos, el Código Universal de Producto (UPC, Universal Product Code) asociado con los códigos de barras hallados en las mercancías, así como el Número Estándar Internacional de Libros (ISBN, International Standard Book Number) que se halla en cada libro publicado en la actualidad, no son sino algunos ejemplos.

En esta sección utilizaremos vectores para diseñar códigos con el propósito de detectar los errores que pueden presentarse durante la transmisión de datos. En capítulos posteriores, construiremos códigos que no sólo podrán detectar, sino también corregir errores. Los vectores que se manejan en el estudio de códigos no son los vectores de $\mathbb{R}^n$, sino vectores con sólo un número finito de selecciones para cada una de las componentes. Estos vectores dependen de un tipo distinto de aritmética (*la aritmética modular*) la cual se presenta en esta sección y se usa a lo largo del libro.

Códigos binarios

Debido a que las computadoras representan los datos en términos de ceros y unos (lo que puede ser interpretado como on/off, encendido/apagado, cerrado/abierto, falso-/verdadero o no/sí), comenzaremos considerando los códigos *binarios*, que consisten en vectores cuyas componentes pueden ser ya sea un 0 o un 1. En este entorno, las reglas habituales de la aritmética deben modificarse, puesto que el resultado de cada cálculo que involucra escalares debe ser un 0 o un 1. Las reglas modificadas para la suma y la multiplicación se proporcionan a continuación.

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

La única observación importante aquí es la regla de que $1 + 1 = 0$. Este resultado no es tan extraño como parece: si reemplazamos el 0 con la palabra "par" y el 1 con la palabra "impar", estas tablas sólo resumen las conocidas *reglas de paridad* para la suma y la multiplicación de enteros pares e impares. Por ejemplo, $1 + 1 = 0$ expresa el hecho de que la suma de dos enteros impares es un entero par. Con estas reglas, nuestro conjunto de escalares $\{0, 1\}$ es denotado como $\mathbb{Z}_2$, el cual se denomina conjunto de los ***enteros módulo* 2**.

Ejemplo 1.27

En $\mathbb{Z}_2$, $1 + 1 + 0 + 1 = 1$ y $1 + 1 + 1 + 1 = 0$ (nuevamente, estos cálculos ilustran las reglas de paridad: la suma de tres números impares y uno par es impar; la suma de cuatro impares es par).

Se utiliza el término *longitud* de manera distinta de la que se emplea en $\mathbb{R}^n$. Esto no debería resultar confuso, puesto que no existe notación *geométrica* de longitud para vectores binarios.

Con $\mathbb{Z}_2$ como nuestro conjunto de escalares, extendemos ahora las reglas anteriores a los vectores. El conjunto de todas las n-tuplas de ceros y unos (con toda la aritmética aplicada en módulo 2) se denota con $\mathbb{Z}_2^n$. Los vectores en $\mathbb{Z}_2^n$ se denominan ***vectores binarios de longitud n***.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

divide entre 3. Mediante división larga, hallamos que $3548 = 3 \cdot 1182 + 2$, de modo que el residuo es 2. Por lo tanto, 3548 es equivalente a 2 en $\mathbb{Z}_3$.

En cursos de álgebra abstracta y teoría de los números, que exploran este concepto con mayor detalle, la equivalencia anterior se escribe a menudo como $3548 = 2(\text{mod } 3)$ o $3548 \equiv 2(\text{mod } 3)$, donde $\equiv$ se lee "es congruente con". Aquí no se utilizará esa notación o terminología.

Ejemplo 1.34 En $\mathbb{Z}_3$, calcule $2 + 2 + 1 + 2$.

Solución 1 Aplicar los mismos conceptos que en el ejemplo 1.33. La suma ordinaria es $2 + 2 + 1 + 2 = 7$, lo cual es 1 más que 6, de modo que la división entre 3 deja un residuo de 1. De esta manera, $2 + 2 + 1 + 2 = 1$ en $\mathbb{Z}_3$.

Solución 2 Una mejor forma de realizar el cálculo es hacerlo paso a paso sólo en $\mathbb{Z}_3$.

$$\begin{aligned} 2 + 2 + 1 + 2 &= (2 + 2) + 1 + 2 \\ &= 1 + 1 + 2 \\ &= (1 + 1) + 2 \\ &= 2 + 2 \\ &= 1 \end{aligned}$$

En este caso, hemos utilizado paréntesis para agrupar los términos que hemos elegido combinar. Podríamos simplificar el proceso si combinamos de manera simultánea los primeros dos y los últimos dos términos:

$$\begin{aligned} (2 + 2) + (1 + 2) &= 1 + 0 \\ &= 1 \end{aligned}$$

La multiplicación repetida puede ser manejada en forma similar. La idea es emplear las tablas de suma y multiplicación para reducir el resultado de cada cálculo a 0, 1 o 2.

La extensión de estas ideas a los vectores es directa.

Ejemplo 1.35 En $\mathbb{Z}_3^5$, sean $\mathbf{u} = [2, 2, 0, 1, 2]$ y $\mathbf{v} = [1, 2, 2, 2, 1]$. Entonces,

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= 2 \cdot 1 + 2 \cdot 2 + 0 \cdot 2 + 1 \cdot 2 + 2 \cdot 1 \\ &= 2 + 1 + 0 + 2 + 2 \\ &= 1 \end{aligned}$$

Los vectores en $\mathbb{Z}_3^5$, son referidos como ***vectores ternarios de longitud 5.***

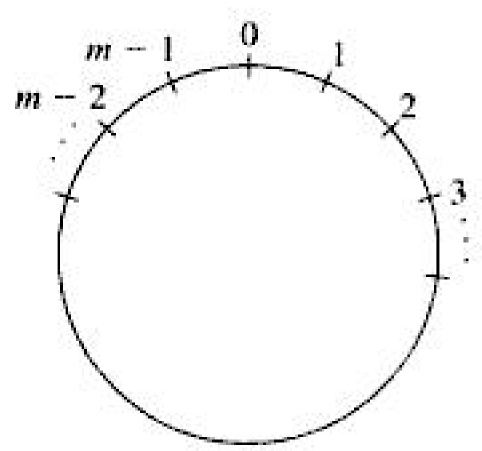


Figura 1.71
Aritmética módulo m

En general, tenemos el conjunto $\mathbb{Z}_m = \{0, 1, 2, \ldots, m - 1\}$ de ***enteros módulo m*** (que corresponden a un reloj de m horas, como se muestra en la figura 1.71) y ***vectores m-arios de longitud n*** denotados por $\mathbb{Z}_m^n$. Los códigos que utilizan vectores m-arios se denominan ***códigos m-arios.***

El ejemplo siguiente es una extensión directa del ejemplo 1.31 para códigos ternarios.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

El sistema Codabar

Toda tarjeta de crédito y débito es identificada de manera única por un número de 16 dígitos que representa un código de un dígito verificador. Al igual que en los ejemplos de esta sección, los primeros 15 dígitos son asignados por el banco emisor de la tarjeta y el último dígito es un dígito verificador determinado por una fórmula que utiliza aritmética modular.

Todos los grandes bancos emplean un sistema llamado Codabar para asignar el dígito verificador. Éste es una variación ligera del método del Código de Producto Universal.

Suponga que los primeros 15 dígitos de su tarjeta son

5412 3456 7890 432

Y que el dígito verificador es d. Éste corresponde al vector

$$\mathbf{x} = [5, 4, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, 4, 3, 2, d]$$

en $\mathbb{Z}_{10}^{16}$. El sistema Codabar utiliza el vector verificador $\mathbf{c} = [2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1]$, pero en lugar de requerir que $\mathbf{c} \cdot \mathbf{x} = 0$ en $\mathbb{Z}_{10}$ un cálculo extra es agregado para incrementar la capacidad de detectar el error del código. Dejemos que h cuente el número de dígitos en las *posiciones impares* que sean *mayores que* 4. En este ejemplo, estos dígitos son 5, 5, 7, y 9, de manera que $h = 4$.

A continuación, se requiere que $\mathbf{c} \cdot \mathbf{x} + h = 0$ en $\mathbb{Z}_{10}$. Por lo tanto en el ejemplo, tenemos, rearreglando y trabajando el módulo 10,

$$\begin{aligned}
\mathbf{c} \cdot \mathbf{x} + h &= (2 \cdot 5 + 4 + 2 \cdot 1 + 2 + 2 \cdot 3 + 4 + 2 \cdot 5 + 6 + 2 \cdot 7 + 8 + 2 \cdot 9 \\
&\quad + 0 + 2 \cdot 4 + 3 + 2 \cdot 2 + d) + 4 \\
&= 2\,(5 + 1 + 3 + 5 + 7 + 9 + 4 + 2) + (4 + 2 + 4 + 6 + 8 + 0 \\
&\quad + 3 + d) + 4 \\
&= 2\,(6) + 7 + d + 4 \\
&= 3 + d
\end{aligned}$$

Así, el dígito verificador d para esta tarjeta debe ser 7, de modo que el resultado del cálculo sea 0 en $\mathbb{Z}_{10}$.

El sistema Codabar es uno de los métodos más eficientes para la detección de errores. Éste identificará todos los errores de un solo dígito y la mayoría de otros errores comunes tales como errores adyacentes de transposición.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Tabla 2.1

Punto	x	y
P_0	0	0
P_1		
P_2		
P_3		
P_4		
P_5		
P_6		

recta desde P_1 hasta la recta 2 y registre este punto como P_2. Después, trace un segmento de recta *horizontal* desde P_2 hasta la recta 1, con lo cual generará el punto P_3. Continúe de este modo, dibujando segmentos verticales hacia la recta 2 seguidos por segmentos horizontales hacia la recta 1. ¿Qué parece estar ocurriendo?

Problema 6 Utilice una calculadora con precisión de dos decimales y determine las coordenadas (aproximadas) de los puntos $P_1, P_2, P_3, \ldots, P_6$. (Le resultará conveniente resolver en primer lugar la primera ecuación para x en términos de y y la segunda ecuación en términos de x.) Registre sus resultados en la tabla 2.1, anotando las coordenadas x y y para cada punto, de manera separada.

Los resultados de estos problemas muestran que la tarea de "resolver" un sistema de ecuaciones lineales puede ser vista de diversas formas. Repita el proceso descrito en los problemas con los siguientes sistemas de ecuaciones:

$$\text{(a)}\ \begin{aligned} 4x - 2y &= 0 \\ x + 2y &= 5 \end{aligned} \qquad \text{(b)}\ \begin{aligned} 3x + 2y &= 9 \\ x + 3y &= 10 \end{aligned} \qquad \text{(c)}\ \begin{aligned} x + y &= 5 \\ x - y &= 3 \end{aligned} \qquad \text{(d)}\ \begin{aligned} x + 2y &= 4 \\ 2x - y &= 3 \end{aligned}$$

¿Todavía son válidas todas sus observaciones de los ejemplos 1 al 6 para estos ejemplos? Observe cualquier similitud o diferencia. En este capítulo, exploraremos estos conceptos con mayor detalle.

2.1 Introducción a los sistemas de ecuaciones lineales

Recordemos que la ecuación general de una recta en $\mathbb{R}^2$ es de la forma

$$ax + by = c$$

y que la ecuación general de un plano en $\mathbb{R}^3$ es de la forma

$$ax + by + cz = d$$

Las ecuaciones de esta forma se denominan ***ecuaciones lineales***.

Definición Una ***ecuación lineal*** en las n variables $x_1, x_2, \ldots, x_n$ es una ecuación que puede escribirse en la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

donde los ***coeficientes*** $a_1, a_2, \ldots, a_n$ y el ***término constante*** b son constantes.

Ejemplo 2.1 Las siguientes ecuaciones son lineales:

$$3x - 4y = -1 \qquad r - \tfrac{1}{2}s - \tfrac{15}{3}t = 9 \qquad x_1 + 5x_2 = 3 - x_3 + 2x_4$$

$$\sqrt{2}x + \frac{\pi}{4}y - \left(\operatorname{sen}\frac{\pi}{5}\right)z = 1 \qquad 3.2x_1 - 0.01x_2 = 4.6$$

Observe que la tercera ecuación es lineal porque puede reescribirse en la forma $x_1 + 5x_2 + x_3 - 2x_4 = 3$. También es importante advertir que, aunque en estos ejemplos (y en la mayoría de las aplicaciones) los coeficientes y términos constantes son números reales, en algunos ejemplos y aplicaciones serán números complejos o miembros de $\mathbb{Z}_p$ para algún número primo p.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Circuitos (redes) eléctricos

Las redes o circuitos eléctricos son un tipo especializado de red que proporciona información acerca de fuentes de poder o alimentación, como las baterías, y los dispositivos energizados por estas fuentes, como las bombillas eléctricas o los motores. Una fuente de poder "obliga" a una corriente de electrones a fluir a través del circuito, donde encuentra varios *resistores*, cada uno de los cuales requiere que una cierta cantidad de fuerza sea aplicada con el fin de que la corriente fluya a través de ellos.

La ley fundamental de la electricidad es la ley de Ohm, que establece exactamente cuánta fuerza E es necesaria para conducir una corriente I a través de un resistor con una resistencia igual a R.

Ley de Ohm

$$\text{fuerza} = \text{resistencia} \times \text{corriente}$$

o

$$E = RI$$

La fuerza se mide en *voltios*, la resistencia en *ohmios* y la corriente en *amperios* (o *amps*, para abreviar). Así, en términos de estas unidades, la ley de Ohm se convierte en "volts = ohms × amps", y nos dice que la "caída de voltaje" se establece cuando una corriente pasa a través de un resistor; es decir, si se sabe cuánto voltaje gasta este último.

La corriente fluye al exterior de la terminal positiva de una batería y fluye de regreso hacia la terminal negativa, viajando alrededor de uno o más circuitos cerrados en el proceso. En un diagrama de un circuito eléctrico, las baterías se representan mediante el símbolo —|⊢ (donde la terminal positiva es la barra vertical más larga) mientras que los resistores están representados con el símbolo —⁄\/\/—.

Las siguientes dos leyes, debidas a Kirchhoff, dirigen los circuitos eléctricos. La primera es una ley de "conservación del flujo" para cada nodo; la segunda es una ley de "balanceo del voltaje" alrededor de cada circuito.

Ley de Kirchhoff

Ley de la corriente (nodos)
La suma de las corrientes que fluyen hacia cualquier nodo es igual a la suma de las corrientes que fluyen hacia afuera de este nodo.

La ley del voltaje (circuitos)
La suma de las caídas de voltaje en cualquier circuito es igual al voltaje total del circuito (proporcionado por las baterías).

La figura 2.12 ilustra las leyes de Kirchhoff. En el inciso (a), la ley de la corriente nos da $I_1 = I_2 + I_3$ (o bien $I_1 - I_2 - I_3 = 0$, como lo escribiremos); el inciso (b) nos da $4I = 10$, donde la ley de Ohm se usa para calcular la caída del voltaje $4I$ en el resistor. Si aplicamos las leyes de Kirchhoff, podemos establecer un sistema de ecuaciones lineales que nos permitirá determinar las corrientes de un circuito eléctrico.

Ejemplo 2.31

Determine las corrientes I_1, I_2 e I_3 del circuito eléctrico que se presenta en la figura 2.13.

Solución Este circuito tiene dos baterías y cuatro resistores. La corriente I_1 fluye a través de la rama superior BCA, la corriente I_2 fluye a través de la rama media o central AB y la corriente I_3 fluye a través de la rama inferior BDA.

En el nodo A, la ley de la corriente nos da $I_1 + I_3 = I_2$, o

$$I_1 - I_2 + I_3 = 0$$

(Observe que obtenemos la misma ecuación en el nodo B.)

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Asimismo, podemos emplear vectores en $\mathbb{Z}_2^5$ para representar la acción de cada interruptor. Si un interruptor cambia el estado de una lámpara, el componente correspondiente es un 1; si no lo hace, es un 0. De acuerdo con esta convención, las acciones de los cinco interruptores están dadas por

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

La situación representada en la figura 2.15(a) corresponde al estado inicial

$$\mathbf{s} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

seguida por

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Es el vector suma (en $\mathbb{Z}_2^5$)

$$\mathbf{s} + \mathbf{a} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Note que este resultado concuerda con la figura 2.15(b).

Comenzando con cualquier configuración inicial $\mathbf{s}$, supongamos que presionamos los interruptores en el orden A, C, D, A, C, B. Ello corresponde a la suma de vectores $\mathbf{s} + \mathbf{a} + \mathbf{c} + \mathbf{d} + \mathbf{a} + \mathbf{c} + \mathbf{b}$. Pero en $\mathbb{Z}_2^5$, , la adición es conmutativa, así que tenemos

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

donde hemos partido del hecho de que $2 = 0$ en $\mathbb{Z}_2$. Por ello, conseguiríamos el mismo resultado pulsando sólo B y D, sin importar el orden. (Verifique que esto es correcto.) Por tanto, en este ejemplo, no necesitamos presionar ningún interruptor más de una vez.

Entonces, para ver si podemos conseguir una configuración objetivo $\mathbf{t}$ a partir de una configuración inicial $\mathbf{s}$, debemos determinar si existen escalares $x_1, \ldots, x_5$ en $\mathbb{Z}_2$ tales que

$$\mathbf{s} + x_1\mathbf{a} + x_2\mathbf{b} + \cdots + x_5\mathbf{e} = \mathbf{t}$$

disponibles 30 k de grano de Colombia, 15 k del de Kenia y 25 k del café tostado de Francia. Si deseamos utilizar la totalidad de los granos, ¿cuántas bolsas de cada tipo de mezcla pueden hacerse?

6. Resuelva de nuevo el ejercicio 5, asumiendo que la mezcla de la casa contiene 300 g de café de Colombia, 50 g del de Kenia y 150 g de café tostado de Francia, mientras que la mezcla *gourmet* contiene 100 g de grano de Colombia, 350 g del de Kenia y 50 g de café tostado de Francia. Esta vez el mercader tiene disponibles 30 k del grano colombiano, 15 k de grano de Kenia y 15 k de grano tostado francés. Supongamos que una bolsa de la mezcla de la casa produce una ganancia de \$0.50, una bolsa de la mezcla especial produce una ganancia de \$1.50 y una bolsa de la mezcla *gourmet* produce una ganancia de \$2.00. ¿Cuántas bolsas de cada tipo debería preparar el comerciante si quiere usar la totalidad de los granos *y además* maximizar sus ganancias? ¿Cuál es la ganancia máxima?

Balanceo de ecuaciones químicas

En los ejercicios 7 a 14, balancee la ecuación química de cada reacción.

7. $FeS_2 + O_2 \longrightarrow Fe_2O_3 + SO_2$
8. $CO_2 + H_2O \longrightarrow C_6H_{12}O_6 + O_2$ (Esta reacción tiene lugar cuando una planta verde convierte el bióxido de carbono y el agua en glucosa y oxígeno durante la fotosíntesis.)
9. $C_4H_{10} + O_2 \longrightarrow CO_2 + H_2O$ (Esta reacción ocurre cuando el butano, C_4H_{10}, se quema en presencia de oxígeno para formar bióxido de carbono y agua.)
10. $C_7H_6O_2 + O_2 \longrightarrow H_2O + CO_2$
11. $C_5H_{11}OH + O_2 \longrightarrow H_2O + CO_2$ (Esta ecuación representa la combustión del alcohol amílico.)
12. $HClO_4 + P_4O_{10} \longrightarrow H_3PO_4 + Cl_2O_7$
13. $Na_2CO_3 + C + N_2 \longrightarrow NaCN + CO$
14. $C_2H_2Cl_4 + Ca(OH)_2 \longrightarrow C_2HCl_3 + CaCl_2 + H_2O$

CAS

Análisis de redes

15. La figura 2.18 muestra una red de tuberías de agua con flujos medidos en litros por minuto.
    (a) Establezca y resuelva un sistema de ecuaciones lineales para encontrar los flujos posibles.
    (b) Si el flujo a través de AB está restringido a 5 L/min, ¿cuáles serán los flujos a través de las otras dos ramas?

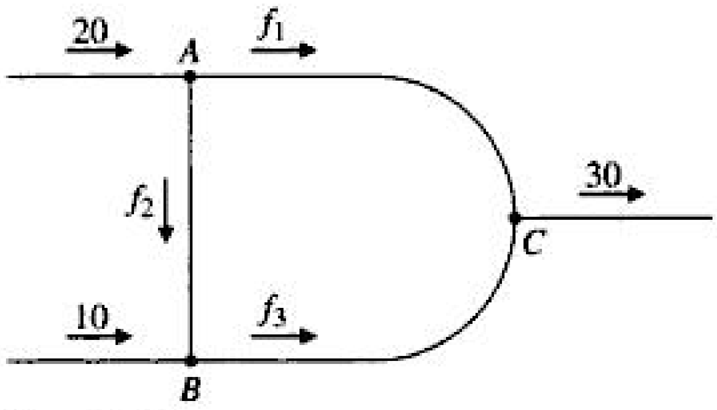


Figura 2.18

    (c) ¿Cuáles son los flujos posibles mínimo y máximo a través de cada rama?
    (d) Hemos estado suponiendo que el flujo siempre es *positivo*. ¿Que significaría un flujo *negativo*, estimando que fuera permitido? Proporcione una ilustración para este ejemplo.

16. El centro de Ciudad Gótica se compone de calles de un solo sentido, y se ha medido el flujo de tráfico en cada intersección. En el área de la ciudad que aparece en la figura 2.19, las cifras representan el número promedio de vehículos por minuto que entran y salen de los puntos de intersección A, B, C y D durante las horas de trabajo.
    (a) Establezca y resuelva un sistema de ecuaciones lineales para hallar los flujos posibles $f_1, \ldots, f_4$.
    (b) Si el tráfico es regulado en CD de manera que $f_4 = 10$ vehículos por minuto, ¿cuáles serán los flujos promedio en las otras calles?
    (c) ¿Cuáles son los flujos posibles mínimo y máximo en cada calle?
    (d) ¿Cómo cambiaría la solución si *todas* las direcciones fueran invertidas

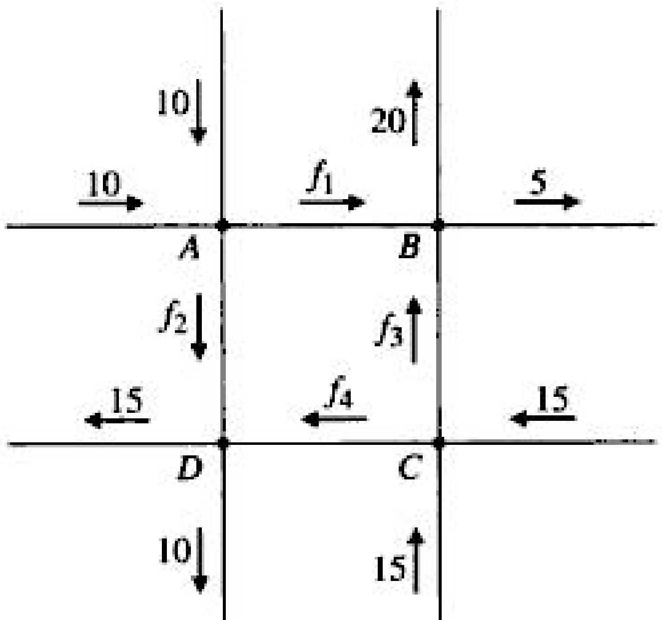


Figura 2.19

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

"Una aplicación de la teoría de matrices", de Paul Glaister en *The Mathematics Teacher*, 85(1992), pp. 220-223.]

(a)

+	a	b
c	2	3
d	4	5

(b)

+	a	b
c	3	6
d	4	5

36. ¿Qué condiciones de w, x, y, z garantizarán que podamos encontrar a, b, c, d, de modo que la siguiente sea una tabla de sumar válida?

+	a	b
c	w	x
d	y	z

37. Describa todos los valores posibles de a, b, c, d, e y f que otorguen validez a cada una de las siguientes tablas de sumar.

(a)

+	a	b	c
d	3	2	1
e	5	4	3
f	4	3	1

(b)

+	a	b	c
d	1	2	3
e	3	4	5
f	4	5	6

38. Si se generaliza el ejercicio 36, encuentre condiciones sobre las entradas de una tabla de sumar de 3×3 que garanticen que podamos resolver a, b, c, d, e y f como en el caso anterior.

39. Con base en la geometría elemental sabemos que existe una recta única que pasa a través de dos puntos cualesquiera de un plano. Menos conocido es el hecho de que existe una parábola única que pasa a través de *tres* puntos cualesquiera no colineales de un plano. Para cada conjunto de puntos que sigue, encuentre una parábola con una ecuación de la forma $y = ax^2 + bx + c$ que pase a través de los puntos dados. (Dibuje la parábola resultante para verificar la validez de su respuesta.)

(a) $(0, 1), (-1, 4)$ y $(2, 1)$

(b) $(-3, 1), (-2, 2)$ y $(-1, 5)$

40. A través de tres puntos cualesquiera no colineales también pasa un círculo único. Determine los círculos (cuyas ecuaciones generales son de la forma $x^2 + y^2 + ax + by + c = 0$) que pasan a través de los conjuntos de puntos del ejercicio 39. (Para verificar la validez de su respuesta, localice el centro y el radio de cada círculo y trace un dibujo.)

El proceso de sumar funciones racionales (razones de polinomios) al colocarlas sobre un denominador común es parecido al de sumar números racionales. El proceso inverso de descomponer una función racional al escribirla como una suma de funciones racionales más simples es útil en diversas áreas de las matemáticas; por ejemplo, aparece en el cálculo cuando necesitamos integrar una función racional y en las matemáticas discretas cuando utilizamos funciones generadoras para resolver relaciones de recurrencia. La descomposición de una función racional como una suma de fracciones parciales genera un sistema de ecuaciones lineales. En los ejercicios 41 a 44, determine la descomposición en fracciones parciales correspondiente a la forma dada. (Las letras mayúsculas representan las constantes.)

41. $\dfrac{3x + 1}{x^2 + 2x - 3} = \dfrac{A}{x - 1} + \dfrac{B}{x + 3}$

42. $\dfrac{x^2 - 3x + 3}{x^3 + 2x^2 + x} = \dfrac{A}{x} + \dfrac{B}{x + 1} + \dfrac{C}{(x + 1)^2}$

CAS **43.** $\dfrac{x - 1}{(x + 1)(x^2 + 1)(x^2 + 4)}$
$= \dfrac{A}{x + 1} + \dfrac{Bx + C}{x^2 + 1} + \dfrac{Dx + E}{x^2 + 4}$

CAS **44.** $\dfrac{x^3 + x + 1}{x(x - 1)(x^2 + x + 1)(x^2 + 1)^3} = \dfrac{A}{x} + \dfrac{B}{x - 1}$
$+ \dfrac{Cx + D}{x^2 + x + 1} + \dfrac{Ex + F}{x^2 + 1} + \dfrac{Gx + H}{(x^2 + 1)^2} + \dfrac{Ix + J}{(x^2 + 1)^3}$

A continuación presentamos dos fórmulas útiles para las sumas de potencias de números naturales consecutivos:

$$1 + 2 + \cdots + n = \frac{n(n + 1)}{2}$$

y

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n + 1)(2n + 1)}{6}$$

La validez de estas fórmulas para todos los valores de $n \geq 1$ (o incluso $n \geq 0$) puede establecerse por medio de la inducción matemática (véase el apéndice B). No obstante, una manera de hacer una observación experimentada sobre lo que son las fórmulas es considerar que podemos volver a escribir las dos fórmulas anteriores como

$$\tfrac{1}{2}n^2 + \tfrac{1}{2}n \quad \text{y} \quad \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n$$

respectivamente. Esto da origen a la conjetura de que la suma de las p potencias de los primeros n números naturales es un polinomio de grado $p + 1$ en la variable n.

45. Suponiendo que $1 + 2 + \ldots + n = an^2 + bn + c$, encuentre a, b y c con la sustitución de tres valores de n y, por tanto, obtenga un sistema de ecuaciones lineales en a, b y c.

46. Asuma que $1^2 + 2^2 + \ldots + n^2 = an^3 + bn^2 + cn + d$. Encuentre a, b, c y d. (*Sugerencia:* es legítimo utilizar $n = 0$. ¿Cuál sería el miembro del lado izquierdo en ese caso?)

47. Demuestre que $1^3 + 2^3 + \ldots + n^3 = (n(n + 1)/2)^2$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

CAS

2.5 Métodos iterativos para resolver sistemas lineales

Los métodos directos para resolver sistemas lineales, utilizando operaciones elementales por renglones, en muchas ocasiones generan soluciones exactas, pero, como hemos visto, están sujetos a errores debido al redondeo y a otros factores. La tercera opción en nuestro "trivium" apunta a una trayectoria muy distinta. En esta sección, exploraremos métodos que avanzan iterativamente generando en forma sucesiva una serie de vectores que aproximan una solución a un sistema lineal. En muchos casos (como sucede cuando una matriz de coeficientes es *dispersa*; es decir, contiene muchas entradas iguales a 0), los métodos iterativos pueden ser más rápidos y más precisos que los métodos directos. Además, los métodos iterativos se pueden detener cuando se quiera que la solución aproximada que generen sea suficientemente precisa. A menudo, los métodos iterativos *se benefician* de la inexactitud: en realidad, el error de redondeo puede acelerar su convergencia hacia una solución.

Exploraremos dos métodos iterativos para resolver sistemas lineales: el ***método de Jacobi*** y un perfeccionamiento del mismo, el ***método de Gauss-Seidel.*** En todos los ejemplos, consideraremos sistemas lineales con el mismo número de variables que de ecuaciones, y supondremos que existe una solución única. Nuestro interés reside en encontrar esta solución con el empleo de métodos iterativos.

Ejemplo 2.35 Considérese el sistema

$$\begin{aligned} 7x_1 - x_2 &= 5 \\ 3x_1 - 5x_2 &= -7 \end{aligned}$$

El método de Jacobi comienza con la resolución de la primera ecuación para x_1 y la segunda ecuación para x_2, obteniendo

$$\begin{aligned} x_1 &= \frac{5 + x_2}{7} \\ x_2 &= \frac{7 + 3x_1}{5} \end{aligned} \tag{1}$$

Ahora, necesitamos una ***aproximación inicial*** a la solución. Resulta que no importa cuál sea esta aproximación inicial, de modo que bien podríamos tomar $x_1 = 0$, $x_2 = 0$. Usemos estos valores en las ecuaciones (1) para conseguir nuevos valores de x_1 y x_2:

$$\begin{aligned} x_1 &= \frac{5 + 0}{7} = \frac{5}{7} \approx 0.714 \\ x_2 &= \frac{7 + 3 \cdot 0}{5} = \frac{7}{5} = 1.400 \end{aligned}$$

Luego, sustituyamos estos valores en (1) para obtener

$$\begin{aligned} x_1 &= \frac{5 + 1.4}{7} \approx 0.914 \\ x_2 &= \frac{7 + 3 \cdot \frac{5}{7}}{5} \approx 1.829 \end{aligned}$$

(aproximados a tres lugares decimales). Repetimos este proceso (mediante la utilización de los valores anteriores de x_2 y x_1 para obtener los nuevos valores de x_1 y x_2), lo cual genera la serie de aproximaciones que se presenta en la tabla 2.7.

Carl Gustav Jacobi (1804-1851) fue un matemático alemán que hizo importantes contribuciones a muchos campos de las matemáticas y la física, incluyendo geometría, teoría de números, análisis, mecánica y mecánica de fluidos. Aunque gran parte de su trabajo se centraba en las matemáticas aplicadas, Jacobi creía en la importancia de hacer matemáticas por el placer de hacerlo, y por beneficio propio. Por todo ello, fue un excelente maestro, dio cátedra en las universidades de Berlín y Königsberg y fue reconocido como uno de los matemáticos más famosos de Europa.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

De modo que, ¿cuándo convergen estos métodos iterativos? Por desgracia, la respuesta a esta cuestión es muy complicada. La responderemos por completo en el capítulo 7, pero por ahora sólo proporcionaremos una respuesta parcial, sin demostración.

Sea A la matriz de $n \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Decimos que A es ***diagonal estrictamente dominante*** si

$$\begin{aligned} |a_{11}| &> |a_{12}| + |a_{13}| + \cdots + |a_{1n}| \\ |a_{22}| &> |a_{21}| + |a_{23}| + \cdots + |a_{2n}| \\ &\vdots \\ |a_{nn}| &> |a_{n1}| + |a_{n2}| + \cdots + |a_{n,n-1}| \end{aligned}$$

Es decir, el valor absoluto de cada entrada diagonal $a_{11}, a_{22}, \ldots, a_{nn}$ es mayor que la suma de los valores absolutos de las entradas restantes en ese renglón

Teorema 2.9 Si un sistema de n ecuaciones lineales con n variables tiene una matriz de coeficientes diagonal estrictamente dominante, entonces tiene una solución única y tanto la del método de Jacobi como la del de Gauss-Seidel convergen en ella.

Observación ¡Sea cauteloso! Este teorema es una implicación de un solo sentido. El hecho de que un sistema *no sea* diagonal estrictamente dominante, *no significa* que los métodos iterativos vayan a ser divergentes. Ellos pueden o no converger. (Véanse los ejercicios 15 a 19.) En realidad, hay ejemplos en los cuales uno de los métodos converge y el otro diverge. Sin embargo, *si* cualquiera de estos métodos converge, entonces debe converger hacia la solución: no puede converger hacia algún otro punto.

Teorema 2.10 Si el método de Jacobi o el de Gauss-Seidel converge para un sistema de n ecuaciones variables en n variables, entonces debe converger hacia la solución del sistema.

Demostración Ilustraremos la idea que hay detrás de la demostración trazando las líneas generales para el caso del método de Jacobi, utilizando el sistema de ecuaciones del ejemplo 2.35. La demostración general es similar.

Convergencia significa que, desde alguna iteración en adelante, los valores de las iteraciones permanecen iguales. Esto significa que x_1 y x_2 convergen a r y s, respectivamente, como se muestra en la tabla 2.12.

Debemos demostrar $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r \\ s \end{bmatrix}$ es la solución del sistema de ecuaciones. En otras palabras, para la $(k + 1)$-ésima iteración, los valores de x_1 y x_2 deben permanecer iguales

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Ejercicios 2.5

En los ejercicios 1 a 6, aplique el método de Jacobi para el sistema dado. Tome el vector cero como la aproximación inicial y trabaje con una exactitud de cuatro dígitos significativos hasta que dos iteraciones sucesivas coincidan dentro del 0.001 de cada variable. En cada caso, compare su respuesta a la solución exacta obtenida con el empleo de cualquier método directo que usted desee.

1. $\begin{aligned} 7x_1 - x_2 &= 6 \\ x_1 - 5x_2 &= -4 \end{aligned}$

2. $\begin{aligned} 2x_1 + x_2 &= 5 \\ x_1 - x_2 &= 1 \end{aligned}$

3. $\begin{aligned} 4.5x_1 - 0.5x_2 &= 1 \\ x_1 - 3.5x_2 &= -1 \end{aligned}$

4. $\begin{aligned} 20x_1 + x_2 - x_3 &= 17 \\ x_1 - 10x_2 + x_3 &= 13 \\ -x_1 + x_2 + 10x_3 &= 18 \end{aligned}$

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

6. $\begin{aligned} 3x_1 - x_2 &= 1 \\ -x_1 + 3x_2 - x_3 &= 0 \\ -x_2 + 3x_3 - x_4 &= 1 \\ -x_3 + 3x_4 &= 1 \end{aligned}$

En los ejercicios 7 a 12, repita el ejercicio dado utilizando el método de Gauss-Seidel. Tome el vector cero como la aproximación inicial y trabaje con una exactitud de cuatro dígitos significativos hasta que dos iteraciones sucesivas coincidan dentro del 0.001 en cada variable. Compare el número de iteraciones requeridas por el método de Jacobi y el de Gauss-Seidel para alcanzar dicha solución aproximada.

7. Ejercicio 1
8. Ejercicio 2
9. Ejercicio 3
10. Ejercicio 4
11. Ejercicio 5
12. Ejercicio 6

En los ejercicios 13 y 14, dibuje diagramas para ilustrar la convergencia del método de Gauss-Seidel con el sistema dado.

13. Sistema en ejercicio 1

14. Sistema en ejercicio 2

En los ejercicios 15 y 16, calcule las primeras cuatro iteraciones, usando el vector cero como la aproximación inicial, para mostrar que el método de Gauss-Seidel diverge. Luego, demuestre que las ecuaciones pueden ser reacomodadas para dar una matriz de coeficientes diagonal estrictamente dominante, y aplique el método de Gauss-Seidel para conseguir una solución aproximada que sea precisa hasta 0.001.

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

16. $\begin{aligned} x_1 - 4x_2 + 2x_3 &= 2 \\ 2x_2 + 4x_3 &= 1 \\ 6x_1 - x_2 - 2x_3 &= 1 \end{aligned}$

17. Dibuje un diagrama para ilustrar la divergencia del método de Gauss-Seidel en el ejercicio 15.

En los ejercicios 18 y 19, la matriz de coeficientes no es diagonal estrictamente dominante, ni las ecuaciones pueden ser reacomodadas para hacerla así. No obstante, tanto el método de Jacobi como el método de Gauss-Seidel convergen de cualquier modo. Pruebe que esto es verdadero con respecto al método de Gauss-Seidel, para lo cual comience con el vector cero como la aproximación inicial y obtenga una solución que sea precisa hasta 0.01.

18. $\begin{aligned} -4x_1 + 5x_2 &= 14 \\ x_1 - 3x_2 &= -7 \end{aligned}$

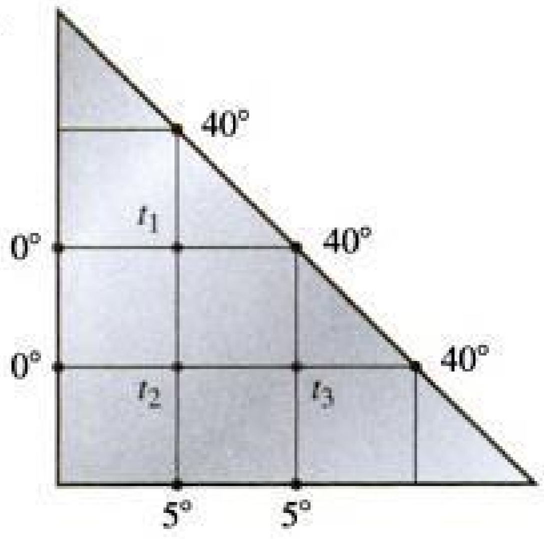
19. $\begin{aligned} 5x_1 - 2x_2 + 3x_3 &= -8 \\ x_1 + 4x_2 - 4x_3 &= 102 \\ -2x_1 - 2x_2 + 4x_3 &= -90 \end{aligned}$

20. Continúe realizando iteraciones en el ejercicio 18 para obtener una solución que sea precisa hasta 0.001.

21. Continúe realizando iteraciones en el ejercicio 19 para obtener una solución que sea precisa hasta 0.001.

En los ejercicios 22 a 24, la placa de metal tiene las temperaturas constantes que muestra en sus límites. Encuentre la temperatura de equilibrio en cada uno de los puntos interiores indicados estableciendo un sistema de ecuaciones lineales y aplicando ya sea el método de Jacobi o el de Gauss-Seidel. Obtenga una solución que sea precisa hasta en un 0.001.

22.



You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

(i) Si un grupo de vectores tiene la propiedad de que dos vectores en el grupo no son múltiplos escalares entre sí, entonces el grupo de vectores es linealmente independiente.

(j) Si existen más vectores en un grupo de vectores que el número de entradas en cada vector, entonces el grupo de vectores es linealmente dependiente.

2. Encuentre el rango de la matriz

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

3. Resuelva el sistema lineal

$$\begin{aligned} x + y - 2z &= 4 \\ x + 3y - z &= 7 \\ 2x + y - 5z &= 7 \end{aligned}$$

4. Resuelva el sistema lineal

$$\begin{aligned} 3w + 8x - 18y + z &= 35 \\ w + 2x - 4y &= 11 \\ w + 3x - 7y + z &= 10 \end{aligned}$$

5. Resuelva el sistema lineal

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

sobre $\mathbb{Z}_7$.

6. Resuelva el sistema lineal

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

sobre $\mathbb{Z}_5$.

7. ¿Para qué valores de k el sistema lineal con matriz aumentada $\left[\begin{array}{cc|c} k & 2 & 1 \\ 1 & 2k & 1 \end{array}\right]$ es inconsistente?

8. Halle las ecuaciones paramétricas para la línea de la intersección de los planos $x + 2y + 3z = 4$ y $5x + 6y + 7z = 8$.

9. Determine el punto de intersección de las siguientes líneas, si es que existe.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + s\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ -4 \end{bmatrix} + t\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

10. Establezca si $\begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}$ está en el espacio generado por $\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$.

11. Encuentre la ecuación general del plano $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$.

12. Determine si $\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 9 \\ -2 \end{bmatrix}$ son linealmente independientes.

13. Determine si $\mathbb{R}^3$= generado($\mathbf{u}, \mathbf{v}, \mathbf{w}$) si:

(a) $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(b) $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

14. Sean $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ vectores linealmente independientes en $\mathbb{R}^3$y $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3]$. ¿Cuál de los siguientes enunciados es verdadero?

(a) La forma escalonada reducida para A es I_3.

(b) El rango de A es 3.

(c) El sistema $[A \mid \mathbf{b}]$ tiene una solución única para cualquier vector $\mathbf{b}$ en $\mathbb{R}^3$.

(d) (a), (b) y (c) son todas verdaderas.

(e) (a) y (b) son verdaderas, pero (c), no.

15. Sean $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ vectores linealmente independientes en $\mathbb{R}^3$, no todos cero, y deje que $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3]$. ¿Cuáles son los posibles valores del rango para A?

16. ¿Cuál es el rango máximo de una matriz de 5×3? ¿Cuál es el rango mínimo de esa misma matriz?

17. Demuestre que si $\mathbf{u}$ y $\mathbf{v}$ son vectores linealmente independientes, entonces también lo son $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$.

18. Compruebe que generado($\mathbf{u}, \mathbf{v}$) = generado($\mathbf{u}, \mathbf{u} + \mathbf{v}$) para cualquier vector $\mathbf{u}$ y $\mathbf{v}$.

19. Con el propósito de que un sistema lineal con matriz aumentada$[A \mid \mathbf{b}]$ sea consistente, ¿qué es lo que tiene que ser verdadero acerca de los rangos A y $[A \mid \mathbf{b}]$?

20. ¿Son las matrices $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & -1 \\ -1 & 4 & 1 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix}$ renglón equivalente? ¿Por qué sí o por qué no?

3 Matrices

Nosotros [Halmos y Kaplansky] compartimos una filosofía acerca del álgebra lineal: pensamos libres de bases, escribimos libres de bases, pero cuando los chips se apagan cerramos la puerta de la oficina y calculamos usando matrices con frenesí.

—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, estudiaremos las matrices por su propio valor. Ya hemos utilizado matrices (en la forma de matrices aumentadas) para registrar información y para ayudar a racionalizar los cálculos que comprenden sistemas de ecuaciones lineales. Ahora veremos que las matrices tienen propiedades algebraicas propias que nos permiten hacer cálculos con ellas, sujetos a las reglas del álgebra de matrices. Además, observaremos que las matrices no son objetos estáticos, que recopilan información y datos; en lugar de ello, representan ciertos tipos de funciones que "actúan" como vectores, transformándolos en otros vectores. Estas "transformaciones matriciales" comenzarán a jugar un papel preponderante en nuestro estudio del álgebra lineal y emitirán una nueva luz sobre lo que ya hemos aprendido acerca de los vectores y sistemas de ecuaciones lineales. Además, las matrices se presentan en muchas otras formas aparte de las versiones aumentadas; exploraremos algunas de las muchas aplicaciones de las matrices al final de este capítulo.

En esta sección, consideraremos unos cuantos ejemplos para ilustrar la forma en que las matrices pueden transformar vectores. En el proceso, usted tendrá su primer contacto con la "aritmética de matrices".

Considere las ecuaciones

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

Podemos ver estas ecuaciones como la descripción de una transformación en el vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ del vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Si denotamos la matriz de coeficientes del lado derecho por F, entonces $F = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ y podemos 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 $y = f(x)$ que usted ya ha utilizado: aquí $\mathbf{x}$ sería la "variable" independiente, $\mathbf{y}$ la "variable" dependiente y F el nombre de la "función".)

De este modo, si $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, entonces las ecuaciones (1) dan

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

podemos escribir esta expresión 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 $\mathbf{x}$:

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

Problema 2 Las cabezas de los cuatro vectores $\mathbf{x}$ del problema 1 se localizan en las cuatro esquinas de un cuadrado en el plano x_1x_2. Dibuje este cuadrado y etiquete sus esquinas como A, B, C y D, correspondientes a las partes a, b, c y d del problema 1.

En ejes de coordenadas separados (etiquetados como y_1 y y_2*), dibuje los cuatro puntos determinados por $F\mathbf{x}$ en el problema 1. Etiquete estos puntos como A', B', C' y D'. Hagamos la (razonable) suposición de que el segmento de recta $\overline{AB}$ se transforma en el segmento de recta $\overline{A'B'}$, y que lo mismo ocurre con los otros tres lados del cuadrado $ABCD$. ¿Cuál figura geométrica es representada por $A'B'C'D'$?

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 cálculos algebraicos confirman esta afirmación?

Ahora consideremos las ecuaciones

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

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

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

Problema 4 Vamos a averiguar cómo G transforma la figura $A'B'C'D'$. Calcule $G\mathbf{y}$ para cada uno de los cuatro vectores $\mathbf{y}$ que se calcularon en el problema 1. [Es decir, calcule $\mathbf{z} = G(F\mathbf{x})$. Usted puede reconocer esta expresión que es análoga a la composición de funciones con la cual usted está familiarizado.] Denomine a los puntos correspondientes $A''B''C''D''$ y haga un esquema de la figura $A''B''C''D''$ en 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 denotamos la matriz de estas ecuaciones como H, entonces tenemos que $\mathbf{z} = H\mathbf{x}$. Puesto que tenemos también que $\mathbf{z} = GF\mathbf{x}$, es razonable escribir

$$H = GF$$

¿Puede usted ver la forma en que las entradas de H se encuentran relacionadas con las entradas de F y G?

Problema 6 Hagamos el proceso anterior a la inversa. Transformemos en primer lugar el cuadrado $ABCD$, utilizando G, para obtener la figura $A^*B^*C^*D^*$. Luego, transformemos la figura resultante, empleando F, para obtener $A^{**}B^{**}C^{**}D^{**}$. [*Nota:* no

* Se lee: "ye subíndice uno y ye subíndice dos".

se preocupe acerca de las "variables" **x**, **y** y **z** aquí. Simplemente sustituya las coordenadas de A, B, C y D en las ecuaciones (2) y después sustituya los resultados en las ecuaciones (1). ¿Son $A^{**}B^{**}C^{**}D^{**}$ y $A''B''C''D''$ lo mismo? ¿Qué le dice esto acerca del orden en el 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 como coeficientes los que se especifican por F y G, encuentre las entradas de H en términos de F y G. El resultado será una fórmula del "producto" $H = GF$.

Problema 8 Repita los problemas 1 al 6 con las matrices siguientes. (Su fórmula del problema 7 puede ayudar a acelerar los cálculos algebraicos.) Advierta cualquier similitud o diferencia que piense que sea de algún significado.

(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 matriciales

Aunque ya hemos encontrado matrices, comencemos por establecer una definición formal y tomar nota de algunos hechos para referencias futuras.

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

Los siguientes son ejemplos de matrices:

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

Aunque estos números por lo regular serán seleccionados del conjunto $\mathbb{R}$ de los números reales, también pueden seleccionarse del conjunto $\mathbb{C}$ de los números complejos o de $\mathbb{Z}_p$, donde p es primo.

El ***tamaño*** de una matriz es una descripción de los números de renglones y columnas que tiene. Se dice que una matriz es de $m \times n$ (que se lee: de "m por n") si tiene m renglones y n columnas. De esta manera, los ejemplos anteriores son matrices de 2×2, 2×3, 3×1, 1×4, 3×3 y 1×1, respectivamente. Una matriz de $1\times m$ se conoce como una ***matriz renglón*** (o ***vector renglón***) y una matriz de $n\times1$ se conoce como una ***matriz columna*** (o ***vector columna***).

Técnicamente, existe una distinción entre matrices vector o columna y vectores, pero no abundaremos en esta distinción. No obstante, *distinguiremos* entre matrices o vectores *renglón* y matrices o vectores *columna*. Esta distinción es importante (cuando menos) para realizar cálculos algebraicos, como demostraremos.

Utilizaremos una notación de *subíndice doble* para hacer referencia a las entradas de una matriz A: la entrada de A en el renglón i y la columna j se denota mediante a_{ij}. De este modo, si

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

entonces $a_{13} = -1$ y $a_{22} = 5$. (La notación A_{ij} se usa, a veces, de manera intercambiable con a_{ij}.) Por lo tanto, podemos denotar una matriz de manera compacta mediante $[a_{ij}]$ (o $[a_{ij}]_{m\times n}$ si es importante especificar el tamaño de A, aunque el tamaño por lo general será más 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 podemos representar A 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 podemos representar A 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 $m = n$ (es decir, si A tiene el mismo número de renglones que columnas), entonces A se conoce como ***matriz cuadrada.*** Una matriz cuadrada cuyas entradas no diagonales sean todas cero se denomina ***matriz diagonal.*** Una matriz diagonal en la cual todas sus entradas diagonales sean las mismas se conoce como ***matriz escalar.*** Si el escalar de la diagonal es 1, la matriz escalar se llama ***matriz de 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 2×2 con entradas diagonales 3 y 5; C es una matriz diagonal; D es la matriz identidad de 3×3. La matriz identidad de $n\times n$ se denota mediante el símbolo I_n (o simplemente I si sus dimensiones se sobreentienden).

Debido a que podemos ver las matrices como generalizaciones de los vectores (y, en efecto, las matrices pueden y deberían ser consideradas como formaciones tanto de vectores renglón como de vectores columna), muchas de las convenciones y operaciones aplicables a los vectores se extienden (de manera obvia) a las matrices.

Dos matrices son ***iguales*** si tienen el mismo tamaño y si sus entradas correspondientes son iguales. De este modo, si $A = [a_{ij}]_{m\times n}$ y $B = [b_{ij}]_{r\times s}$, y entonces $A = B$ si y sólo si $m = r$ y $n = s$ y $a_{ij} = b_{ij}$ para toda 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 (no importa cuáles sean los valores de x y y), puesto que 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 $d = 3$.

Ejemplo 3.2

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, $R \neq C$ puesto que R es de 1×3 y C es de 3×1. (Si leyéramos R y C en voz alta, sonarían iguales: "uno, cuatro, tres".) Por ello, es importante distinguir entre matrices/vectores renglón y matrices/vectores columna.

Adición de matrices y multiplicación por escalares

Generalizando a partir de la adición de vectores, definimos la adición 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 $m\times n$ obtenida mediante la suma de las entradas correspondientes. De esta manera,

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

[De igual manera, bien podríamos haber definido $A + B$ en términos de la adición de vectores especificando que cada columna (o renglón) de $A + B$ fuera la suma de las columnas (o renglones) correspondientes de A y B.] Si A y B no son del mismo tamaño, entonces $A + B$ no está definida.

Ejemplo 3.3 Sea

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

Entonces

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

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

La definición por componentes de la multiplicación por escalares aparece sin ser una sorpresa. Si A es una matriz de $m\times n$ y c es un escalar, entonces el ***múltiplo escalar*** cA es la matriz de $m\times n$ obtenida al multiplicar cada entrada de A por c. De manera más formal, tenemos que

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

[En términos de vectores, podríamos estipular de manera equivalente que cada columna (o renglón) de cA es c veces la columna (o renglón) correspondiente 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 denomina la ***negativa**** de A. Como ocurre con los vectores, podemos usar este hecho para definir la ***diferencia*** de dos matrices: Si A y B son del mismo tamaño, entonces

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

* A también es conocido como el *inverso aditivo* de A.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

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 de matrices

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

de manera 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 independientes.

Usted no debería tener dificultad para ver que *todo* sistema lineal puede expresarse en la forma $A\mathbf{x} = \mathbf{b}$. De hecho, la notación $[A|\mathbf{b}]$ de la matriz aumentada de un sistema lineal es como la expresión taquigráfica o la abreviatura de la ecuación matricial $A\mathbf{x} = \mathbf{b}$. Esta forma, que probará ser formidablemente útil para expresar un sistema de ecuaciones lineales, la explotaremos con frecuencia desde ahora.

Combinando esta idea con el teorema 2.4, observamos 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 hecho acerca de las operaciones matriciales que también probará ser bastante útil: la multiplicación de una matriz por un vector unitario estándar puede emplearse para "extraer" o "reproducir" una columna o 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 $\mathbf{e}_2A$, con los vectores unitarios $\mathbf{e}_3$ y $\mathbf{e}_2$ seleccionados de modo que los productos tengan sentido. Así,

$$A\mathbf{e}_3 = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{y} \quad \mathbf{e}_2A = [0 \quad 1]\begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} = [0 \quad 5 \quad -1]$$

Nótese que $A\mathbf{e}_3$ nos da la tercera columna de A y $\mathbf{e}_2A$ nos proporciona el segundo renglón de A. Anotemos el resultado general como un teorema.

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

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

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

(Observe que calcular cada producto externo es exactamente como completar una tabla de multiplicar.) Por consiguiente, la expansión del producto externo de AB es

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

Utilizaremos la expansión del producto externo en los capítulos 5 y 7 cuando expliquemos el teorema espectral y la descomposición del valor singular, respectivamente.

Cada una de las particiones precedentes es un caso particular de la partición de matrices en general. Se dice que una matriz A se encuentra particionada si se han introducido líneas verticales y horizontales que subdividen a A en submatrices denominadas bloques. La partición permite expresar a A como una 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. Ellas 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 han sido particionadas de la misma manera, es claro que pueden ser sumadas y multiplicadas por escalares bloque a bloque. Menos obvio es el hecho de que, con una partición apropiada, las matrices pueden ser multiplicadas también por bloques. El ejemplo siguiente ilustra este proceso.

Ejemplo 3.12

Considere las matrices A y B anteriores. Si ignoramos por el momento el hecho de que sus entradas son matrices, entonces A parece ser una matriz de 2×2 y B una matriz de 2×3. Su producto debería ser entonces 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 son en realidad productos *matriciales*, de modo que necesitamos estar seguros de que todos ellos se encuentran definidos. Una rápida verificación revela que éste es efectivamente el caso, puesto que los números de *columnas* de los bloques de A (3 y 2) coinciden con los números de los *renglones* de los bloques de B. Se dice que las matrices A y B están ***particionadas de manera conformable para la multiplicación por bloques.***

Si llevamos a cabo los cálculos indicados obtenemos el producto AB en forma particionada, obtenemos:

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

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

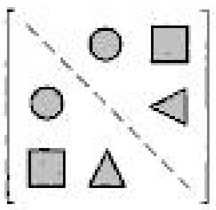


Figura 3.1
Matriz simétrica

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

Una matriz simétrica tiene la propiedad de que es su propia "imagen especular" (imagen en un espejo) con respecto a su diagonal principal. La figura 3.1 ilustra esta propiedad de una matriz dada de 3×3. Las figuras correspondientes representan entradas iguales; las entradas de la diagonal (que se encuentran sobre la línea punteada) son arbitrarias.

Una definición basada en 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 toda i y j.

Ejercicios 3.1

Sean

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

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

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

1. $A + 2D$ **2.** $3D - 2A$

3. $B - C$ **4.** $B - C^T$

5. AB **6.** BD

7. $D + BC$ **8.** B^TB

9. $E(AF)$ **10.** $F(DF)$

11. FE **12.** EF

13. $B^TC^T - (CB)^T$ **14.** $DA - AD$

15. A^3 **16.** $(I_2 - D)^2$

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

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

19. Una fábrica elabora tres productos (cacharros, cachivaches y trebejos) y los envía a dos depósitos para su almacenamiento. El número de unidades de cada producto enviado a cada depósito está dado por la matriz

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

(donde a_{ij} es el número de unidades del producto i enviado al depósito j y los productos se toman en orden alfabético). El costo del envío de una unidad de cada producto por camión es de \$1.50 por cacharro, \$1.00 por cachivache y \$2.00 por trebejo. Los correspondientes costos por unidad en el envío por ferrocarril son \$1.75, \$1.50 y \$1.00, respectivamente. Organice estos costos en una matriz B y después utilice la multiplicación de matrices para mostrar cómo la fábrica puede comparar el costo del envío de sus productos a cada uno de los dos depósitos por camión o por ferrocarril.

20. En referencia al ejercicio 19, supongamos que el costo por unidad para distribuir los productos en las tiendas es el mismo para cada producto, pero varía por depósito debido a las distancias involucradas. Cuesta \$0.75 distribuir una unidad desde el depósito 1 y \$1.00 distribuirla desde el depósito 2. Organice estos costos en una matriz C y a continuación utilice la multiplicación de matrices para calcular el costo total de la distribución de cada producto.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

La reducción por renglón da

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

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

Observación Note que las columnas de la matriz aumentada contienen las entradas de las matrices que estamos dando. Si leemos las entradas de cada matriz de izquierda a derecha y de arriba hacia abajo, obtenemos el orden en el cual las entradas aparecen en las columnas de la matriz aumentada. Por ejemplo, leemos A_1 como "0, 1, -1, 0", lo que corresponde a la primera columna de la matriz aumentada. Es como si sólo "arregláramos" las matrices dadas en vectores columna. Así, acabaríamos con exactamente el mismo sistema de ecuaciones lineales que en el inciso (a) si hubiéramos preguntado

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

Encontraremos paralelismos de esta clase de ahora en adelante. En el capítulo 6, los exploraremos con más detalle.

Podemos 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 del ejemplo 3.16.

Solución Una manera de hacerlo es simplemente escribir una combinación lineal general de A_1, A_2 y A_3. De este modo,

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

(lo cual es análogo a la representación paramétrica de un plano). Pero supongamos que deseamos saber cuándo la matriz $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ se encuentra en el generado (A_1, A_2, A_3). De la representación anterior, sabemos que es 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 selección de los escalares c_1, c_2, c_3. Esto da origen a un sistema de ecuaciones lineales cuyo lado izquierdo es exactamente el mismo que en el ejemplo 3.16, pero

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Demostración La demostración de las propiedades (a) a (c) es intuitivamente clara y directa (véase el ejercicio 30). Demostrar la propiedad (e) es un buen ejercicio para la inducción matemática (véase el ejercicio 31). Demostraremos el inciso (d), puesto que no es lo que usted podría haber esperado. [¿Habría usted sospechado que $(AB)^T = A^T B^T$ podría ser verdadero?]

En primer lugar, 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 $n \times m$. De esta manera, el producto $B^T A^T$ está definido como de $r \times m$. Puesto que AB es de $m \times r$, $(AB)^T$ es de $r \times m$ y así $(AB)^T$ y $B^T A^T$ tienen el mismo tamaño. Ahora debemos probar que sus correspondientes entradas son iguales.

Denotaremos el i-ésimo renglón de una matriz X por $\text{ren}_i(X)$ y su j-ésima columna mediante $\text{col}_j(X)$. Empleando estas convenciones, vemos que

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

(Nótese que hemos utilizado las definiciones de multiplicación de matrices y de transpuesta y el hecho de que el producto punto es conmutativo.) Puesto que i y j son arbitrarios, este resultado implica que $(AB)^T = B^T A^T$.

Observación Las propiedades de los incisos (b) y (d) del teorema 3.4 puede ser generalizadas a sumas finitas y productos finitos de matrices:

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

suponiendo que los tamaños de las matrices sean tales que todas las operaciones puedan ser realizadas. Se le pedirá a usted que demuestre estos hechos por medio de inducción matemática en los ejercicios 32 y 33.

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 manera que $A + A^T = \begin{bmatrix} 2 & 5 \\ 5 & 8 \end{bmatrix}$, una matriz simétrica.

Tenemos

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

de manera 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

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

De este modo, tanto BB^T como $B^T B$ son simétricas, ¡aun cuando B ni siquiera sea cuadrada! (Verifique que AA^T y $A^T A$ sean también simétricas.)

El teorema siguiente nos 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 Demostraremos (a) y dejaremos la demostración de (b) como el ejercicio 34. Simplemente verificamos que

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

(utilizando propiedades de la transpuesta y la conmutatividad de la adición de matrices). Por lo tanto, $A + A^T$ es igual a su propia transpuesta y así, por definición, es simétrica.

Ejercicios 3.2

En los ejercicios 1 a 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. $2X = A - B$

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

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

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

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

6. $B = \begin{bmatrix} 2 & 3 \\ -4 & 2 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

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

8. $B = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 2 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$, $A_4 = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

En los ejercicios 9 a 12, encuentre la forma general del espacio generado por las matrices indicadas, como en el ejemplo 3.17.

9. generado(A_1, A_2) del ejercicio 5

10. generado(A_1, A_2, A_3) del ejercicio 6

11. generado(A_1, A_2, A_3) del ejercicio 7

12. generado(A_1, A_2, A_3, A_4) del ejercicio 8

En los ejercicios 13 a 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. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$

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

matriz identidad I. (Nótese que esta demostración no hace referencia al tamaño de las matrices y por consiguiente es verdadera para matrices de $n \times n$ en general.)

(b) Supongamos que B tiene una inversa $B' = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$. La ecuación $BB' = I$ nos da

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

de lo que obtenemos las ecuaciones

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

Restando dos veces la primera ecuación de la tercera, llegamos a que $0 = -2$, lo que evidentemente es absurdo. Por lo tanto, no existe solución. (La reducción por renglón nos da el mismo resultado, pero no es en realidad necesaria aquí). Deducimos que no existe una matriz B' de ese tipo; es decir, B no es invertible. (¡De hecho, no tiene una inversa que funcione en un lado, y mucho menos en los dos!)

Observaciones

• Aun cuando hemos visto que la multiplicación de matrices no es, en general, conmutativa, A' (si existe) debe satisfacer $A'A = AA'$.

• Los ejemplos anteriores motivan dos preguntas, que responderemos en esta sección:

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

• No hemos descartado la posibilidad de que una matriz A pueda tener más de una inversa. El siguiente teorema nos asegura que esto no puede pasar.

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

Demostración En matemáticas, una manera estándar de demostrar que existe solamente una entidad de algo es mostrar que no puede haber más de una. Así, supongamos que A tiene dos inversas (digamos, A' y A''). Entonces

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

De este modo, $\quad A' = A'I = A'(AA'') = (A'A)A'' = IA'' = A''$

Por lo tanto, $A' = A''$, y la inversa es única.

Gracias a este teorema, podemos hacer referencia a *la* inversa de una matriz invertible. De ahora en adelante, cuando A sea invertible, denotaremos su (única) inversa mediante A^{-1} (pronunciado como "inversa de A").

Advertencia ¡No caiga en la tentación de escribir $A^{-1} = \frac{1}{A}$! No existe una operación tal como la "división entre una matriz". Incluso si la hubiera, ¿cómo rayos podríamos

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 todos los enteros no negativos n y

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

Demostración Demostraremos las propiedades (a), (c) y (e), pero la demostración de (b) y (d) la dejaremos como los ejercicios 14 y 15.

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

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

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

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

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

La afirmación es que sustituir $B^{-1}A^{-1}$ por X funciona. Verificamos que

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

donde hemos utilizado asociatividad para cambiar los paréntesis. De manera semejante, $(B^{-1}A^{-1})(AB) = I$ (¡verifíquelo!), de modo que AB es invertible y su inversa es $B^{-1}A^{-1}$.

(e) La idea básica aquí es bastante fácil. Por ejemplo, cuando $n = 2$, tenemos

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

De manera similar, $(A^{-1})^2 A^2 = I$. Así, $(A^{-1})^2$ es la inversa de A^2. No es difícil apreciar que un argumento semejante surte efecto para cualquier valor entero mayor de n. Sin embargo, la inducción matemática es la manera de llevar a cabo la demostración.

El paso básico es cuando $n = 0$, en cuyo caso estamos obligados a demostrar 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 es evidentemente cierto. (¿Por qué? Véase el ejercicio 16.)

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

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

El paso de inducción requiere que demostremos que A^{k+1} es invertible y que $(A^{k+1})^{-1} = (A^{-1})^{k+1}$. Ahora sabemos por (c) que $A^{k+1} = A^kA$ es invertible, puesto que A y (por hipótesis) A^k son ambas invertibles. Por otra parte,

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

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

proviene de $4R_2$, lo cual es deshecho al efectuar $\frac{1}{4}R_2$. De este modo,

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

lo que puede ser fácilmente verificado. Por último, E_3 corresponde a la operación elemental por renglón $R_3 - 2R_1$, lo que puede ser deshecho mediante la operación elemental por renglón $R_3 + 2R_1$. Por lo tanto, en este caso,

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

(De nueva cuenta, es fácil verificar esta operación mediante la confirmación de que el producto de esta matriz y E_3, en ambos órdenes, es I.)

Nótese que no sólo cada matriz elemental es invertible, sino que su inversa es otra matriz elemental del mismo tipo. Tomemos nota de este hallazgo como el teorema siguiente.

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

Teorema fundamental de las matrices invertibles

Ahora estamos en posición de demostrar uno de los principales resultados en este libro: un conjunto de caracterizaciones equivalentes a lo que significa que una matriz sea invertible. En cierto sentido, gran parte del álgebra lineal se encuentra conectada con este teorema, ya sea en el desarrollo de estas caracterizaciones o bien en su aplicación. Como podría esperarse, dada esta introducción, aplicaremos mucho este teorema. ¡Conviértalo en uno de sus amigos!

Haremos referencia al teorema 3.12 como la primera versión del teorema fundamental, puesto que lo iremos perfeccionando en capítulos subsiguientes. Recuerde que le advertimos que, cuando afirmamos que un conjunto de declaraciones acerca de una matriz A es equivalente, lo que queremos decir es que para una A dada, las declaraciones son todas verdaderas o todas falsas.

Teorema 3.12 **Teorema fundamental de las matrices invertibles: primera versión**

Sea A una matriz de $n \times n$. Las declaraciones siguientes son equivalentes:

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

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Solución La eliminación de Gauss-Jordan produce

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

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

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

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

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

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

Por consiguiente,

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

(Usted siempre debería verificar que $AA^{-1} = I$ mediante multiplicación directa. Según el teorema 3.13, no necesitamos verificar también que $A^{-1}A = I$.)

Observación Nótese que hemos utilizado la variante de la eliminación de Gauss-Jordan que primero introduce todos los ceros *por debajo* de los "unos" principales, de izquierda a derecha y de arriba hacia abajo, y posteriormente crea "ceros" *por arriba* de los "unos" principales, de derecha a izquierda y de abajo hacia arriba. Este enfoque ahorra cálculos, como advertimos en el capítulo 2, pero tal vez le resulte más fácil, cuando trabaje a mano, crear *todos* los ceros en cada columna a medida que avance. La respuesta, por supuesto, 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 es que existe.

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 a 72, particione la matriz dada de modo que usted pueda aplicar una de las fórmulas de los ejercicios 64 a 68, y entonces calcule la inversa utilizando esa 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 iluminador) factorizar un número natural en un producto de otros números naturales, por ejemplo, $30 = 2 \cdot 3 \cdot 5$ —suele ser de ayuda factorizar las matrices como productos de otras matrices—. Cualquier representación de una matriz como un producto de dos o más matrices se conoce como **factorización matricial**. 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 matricial.

No es necesario decir que unas factorizaciones son más útiles que otras. En esta sección introducimos una factorización matricial que surge en la solución de sistemas de ecuaciones lineales por eliminación gaussiana y que es útil particularmente en la implementación de computadoras. En los capítulos siguientes encontraremos otras factorizaciones matriciales útiles. De hecho, este tema es tan vasto que se han hecho cientos de libros y cursos sobre él.

Consideremos un sistema de ecuaciones lineales de la forma $A\mathbf{x} = \mathbf{b}$, donde A es una matriz $n \times n$. Nuestra meta es mostrar que la eliminación gaussiana factoriza de manera implícita A en un producto de matrices que nos permite resolver con facilidad el sistema dado (y cualquier otro con la misma matriz de coeficientes).

El siguiente ejemplo ilustra esta idea básica.

Ejemplo 3.33 Sea

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

Solución Reduciendo A a la forma escalonada por renglones, tenemos:

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

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

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

Los tres primeros multiplicadores son 2, 1 y -3 y éstos van en las entradas de la subdiagonal de la primera columna de L. Por lo tanto,

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

Los siguientes dos multiplicadores son $\frac{1}{2}$ y 4, de manera que continuaremos completando 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, reemplaza el último * en L para dar

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

De esta manera, 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 puede verificar fácilmente.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Definición Un ***subespacio*** de $\mathbb{R}^n$ es cualquier colección S de vectores en $\mathbb{R}^n$ tales que

1. El vector cero $\mathbf{0}$ se encuentra en S.
2. Si $\mathbf{u}$ y $\mathbf{v}$ se encuentran en S, entonces $\mathbf{u} + \mathbf{v}$ está en S. (S es ***cerrada bajo la adición.***)
3. Si $\mathbf{u}$ se encuentra en S y c es un escalar, entonces $c\mathbf{u}$ está en S. (S es ***cerrada bajo la multiplicación por escalares.***)

Podríamos haber combinado las propiedades (2) y (3) y exigir, de manera equivalente, que S se encuentre ***cerrada bajo las combinaciones lineales***:

Si $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ están en S y $c_1, c_2, \ldots, c_k$ son escalares,
entonces $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k$ está en S.

Ejemplo 3.37

Cada recta y plano a través del origen en $\mathbb{R}^3$ es un subespacio de $\mathbb{R}^3$. Debería ser geométricamente claro que las propiedades (1) a (3) se satisfacen. Aquí tenemos una prueba algebraica del caso de un plano que pasa a través del origen. Se le pedirá proporcionar la prueba correspondiente para una recta en el ejercicio 9.

Sea $\mathcal{P}$ un plano que pasa a través del origen con vectores de dirección $\mathbf{v}_1$ y $\mathbf{v}_2$. Por lo tanto, $\mathcal{P}$ = generado($\mathbf{v}_1, \mathbf{v}_2$). El vector cero $\mathbf{0}$ se encuentra en $\mathcal{P}$, puesto que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. 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

$$\mathbf{u} + \mathbf{v} = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2) + (d_1\mathbf{v}_1 + d_2\mathbf{v}_2) = (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2$$

De esta manera, $\mathbf{u} + \mathbf{v}$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ y por lo tanto se encuentra en $\mathcal{P}$.

Ahora, sea c un escalar. Entonces

$$c\mathbf{u} = c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2$$

lo que demuestra que $c\mathbf{u}$ también es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ y por consiguiente se encuentra en $\mathcal{P}$. Hemos demostrado que $\mathcal{P}$ satisface las propiedades (1) a (3) y por lo tanto es un subespacio de $\mathbb{R}^3$.

Si usted examina cuidadosamente los detalles del ejemplo 3.37, advertirá que el hecho de que $\mathbf{v}_1$ y $\mathbf{v}_2$ fueran vectores en $\mathbb{R}^3$ no desempeñó en absoluto papel alguno en la verificación de las propiedades. Así, el método algebraico que utilizamos debería generalizarse más allá de $\mathbb{R}^3$ y aplicarse en situaciones donde ya no podemos visualizar la geometría. Así ocurre. Además, el método del ejemplo 3.37 puede servir como una "plantilla" en entornos más generales. Cuando generalizamos el ejemplo 3.37 al espacio de un conjunto arbitrario de vectores en cualquier $\mathbb{R}^n$, el resultado es suficientemente importante para ser llamado un teorema.

Teorema 3.19 Sea $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vectores en $\mathbb{R}^n$. Entonces el generado($\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$) es un subespacio de $\mathbb{R}^n$.

Demostración Sea S = generado($\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$). Para verificar la propiedad (1) de la definición, simplemente observamos que el vector cero $\mathbf{0}$ se encuentra en S, puesto que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Reagrupamos y tenemos que

$$(c_1a_{11} + c_2a_{21} + \cdots + c_sa_{s1})\mathbf{u}_1 + (c_1a_{12} + c_2a_{22} + \cdots + c_sa_{s2})\mathbf{u}_2 + \cdots + (c_1a_{1r} + c_2a_{2r} + \cdots + c_sa_{sr})\mathbf{u}_r = \mathbf{0}$$

Ahora, puesto que $\mathcal{B}$ es una base, las $\mathbf{u}_j$ son linealmente independientes. De manera que cada una de las expresiones entre paréntesis debe ser cero:

$$\begin{aligned} c_1a_{11} + c_2a_{21} + \cdots + c_sa_{s1} &= 0 \\ c_1a_{12} + c_2a_{22} + \cdots + c_sa_{s2} &= 0 \\ &\vdots \\ c_1a_{1r} + c_2a_{2r} + \cdots + c_sa_{sr} &= 0 \end{aligned}$$

Éste es un sistema homogéneo de r ecuaciones lineales en las s variables $c_1, c_2, \ldots, c_s$. (El hecho de que las variables aparezcan a la izquierda de los coeficientes no implica ninguna diferencia.) Puesto que $r < s$, sabemos, según el teorema 2.3 que hay un número infinito de soluciones. En particular, existe una solución no trivial, que genera una relación de dependencia no trivial en la ecuación (1). De este modo, $\mathcal{C}$ es un conjunto de vectores linealmente dependiente. Pero este descubrimiento contradice el hecho de que $\mathcal{C}$ se postuló como una base, por lo cual es linealmente independiente. Concluimos entonces que $r < s$ no es posible. De manera semejante (intercambiando los papeles de $\mathcal{B}$ y $\mathcal{C}$), encontramos que $r > s$ nos conduce a una contradicción. Por lo tanto, debe ocurrir que $r = s$, como deseábamos.

Puesto que todas las bases de un subespacio dado deben tener el mismo número de vectores, podemos asociarle un nombre a este número.

Definición Si S es un subespacio de $\mathbb{R}^n$, entonces el número de vectores de una base de S se conoce como la ***dimensión*** de S, denotada como dim S.

Observación El vector cero $\mathbf{0}$ por sí mismo siempre es un subespacio de $\mathbb{R}^n$. (¿Por qué?) Incluso cualquier conjunto que contenga al vector cero (y, en particular, $\{\mathbf{0}\}$) es linealmente dependiente, de modo que $\{\mathbf{0}\}$ no puede tener una base. Definimos la dim $\{\mathbf{0}\}$ igual a 0.

Ejemplo 3.49

Debido a que la base estándar de $\mathbb{R}^n$ tiene n vectores, dim $\mathbb{R}^n = n$. (Advierta que este resultado concuerda con nuestra comprensión intuitiva de dimensión de $n \le 3$.)

Ejemplo 3.50

En los ejemplos 3.45 a 3.48, encontramos que ren(A) tiene una base con tres vectores, col(A) tiene una base con tres vectores y nul(A) una base con dos vectores. Por lo tanto, dim (ren(A)) = 3, dim (col(A)) = 3 y dim (nul(A)) = 2.

Un ejemplo aislado no es suficiente para hacer conjeturas, pero el hecho de que los espacios renglón y columna del ejemplo 3.50 tengan la misma dimensión no es accidental. Tampoco lo es el hecho de que la suma de dim (col(A)) y dim (nul(A)) sea 5, el número de columnas de A. Ahora demostraremos que estas relaciones son ciertas en general.

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 Según 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 rango 3. Realizaremos las operaciones por renglones suficientes para determinar lo que acabamos de afirmar:

$$A = \begin{bmatrix}1 & -1 & 4\\2 & 0 & 9\\3 & 1 & 7\end{bmatrix} \longrightarrow \begin{bmatrix}1 & -1 & 4\\0 & 2 & 1\\0 & 0 & -7\end{bmatrix}$$

Observamos que A tiene rango 3, por lo cual los vectores dados son una base para $\mathbb{R}^3$ por la equivalencia de (f) y (j).

El teorema siguiente es una aplicación tanto del teorema del rango como del teorema fundamental. Requeriremos de este resultado en los capítulos 5 y 7.

Teorema 3.28

Sea A una matriz de $m \times n$. Entonces

a. $\text{rango}(A^TA) = \text{rango}(A)$
b. La matriz A^TA de $n \times n$ es invertible si y sólo si $\text{rango}(A) = n$.

Demostración

(a) Puesto que A^TA es de $n \times n$, tiene el mismo número de columnas que A. El teorema del rango nos dice entonces que

$$\text{rango}(A) + \text{nulidad}(A) = n = \text{rango}(A^TA) + \text{nulidad}(A^TA)$$

Por consiguiente, para demostrar que $\text{rango}(A) = \text{rango}(A^TA)$, es suficiente comprobar que $\text{nulidad}(A) = \text{nulidad}(A^TA)$. Lo haremos luego de establecer que los espacios nulos de A y A^TA son los mismos.

Para este fin, sea $\mathbf{x}$ un elemento de $\text{nul}(A)$, así que $A\mathbf{x} = \mathbf{0}$. Entonces $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$, y de esta manera $\mathbf{x}$ está en $\text{nul}(A^TA)$. A la inversa, sea $\mathbf{x}$ un elemento en $\text{nul}(A^TA)$. Entonces $A^TA\mathbf{x} = \mathbf{0}$, de modo que $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{0} = \mathbf{0}$. Pero entonces

$$(A\mathbf{x}) \cdot (A\mathbf{x}) = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x} = 0$$

y por consiguiente $A\mathbf{x} = \mathbf{0}$, según el teorema 1.2(d). Por lo tanto, $\mathbf{x}$ se encuentra en $\text{nul}(A)$, así que $\text{nul}(A) = \text{nul}(A^TA)$, como se requería.

(b) Según el teorema fundamental, la matriz A^TA de $n \times n$ es invertible si y sólo si $\text{rango}(A^TA) = n$. Sin embargo, según (a) tenemos que ello es así si y sólo si $\text{rango}(A) = n$.

Coordenadas

Ahora volveremos a una de las cuestiones planteadas muy al principio de esta sección: ¿Cómo deberíamos ver los vectores en $\mathbb{R}^3$ que permanecen en un plano que pasa a través del origen? ¿Son bidimensionales o tridimensionales? Las nociones de base y dimensión ayudarán a aclarar las cosas.

En los ejercicios 49 y 50, muestre que **w** *se encuentra en el generado($\mathcal{B}$) y encuentre el vector de coordenadas* $[\mathbf{w}]_{\mathcal{B}}$.

49. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}, \mathbf{w} = \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}$

50. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ 6 \end{bmatrix} \right\}, \mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$

En los ejercicios 51 a 54, calcule el rango y nulidad de las matrices dadas sobre el $\mathbb{Z}_p$ *indicado.*

51. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_2$ **52.** $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 2 \\ 2 & 0 & 0 \end{bmatrix}$ sobre $\mathbb{Z}_3$

53. $\begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 3 & 0 & 1 \\ 1 & 0 & 4 & 0 \end{bmatrix}$ sobre $\mathbb{Z}_5$

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

55. Si A es de $m \times n$, demuestre que todo vector de nul(A) es ortogonal a todo vector de ren(A).

56. Si A y B son matrices de $n \times n$ de rango n, demuestre que AB tiene rango n.

57. (a) Demuestre que rango(AB) $\leq$ rango (B). (*Sugerencia:* repase el ejercicio 29 de la sección 3.1.)
(b) Proporcione un ejemplo en el cual rango(AB) < rango (B).

58. (a) Demuestre que rango(AB) $\leq$ rango (A). (*Sugerencia:* repase el ejercicio 30 de la sección 3.1 o utilice transpuestas y el ejercicio 57(a).)
(b) Proporcione un ejemplo en el cual rango(AB) < rango (A).

59. (a) Demuestre que, si U es invertible, entonces rango (UA) = rango (A). [*Sugerencia:* $A = U^{-1}(UA)$.]
(b) Demuestre que, si V es invertible, entonces rango(AV) = rango(A).

60. Demuestre que una matriz A de $m \times n$ tiene rango 1 si y sólo si A puede ser escrita como el producto externo $\mathbf{u}\mathbf{v}^T$ de un vector $\mathbf{u}$ de $\mathbb{R}^m$ y $\mathbf{v}$ en $\mathbb{R}^n$.

61. Si una matriz A de $m \times n$ tiene rango r, demuestre que A puede ser escrita como la suma de r matrices, cada una de las cuales tenga rango 1. (*Sugerencia:* encuentre una manera de emplear el ejercicio 60.)

62. Demuestre que, para matrices A y B de $m \times n$, rango($A + B$) $\leq$ rango(A) + rango(B).

63. Sea A una matriz $n \times n$ tal que $A^2 = O$. Demuestre que rango (A) $\leq n/2$. [*Sugerencia:* demuestre que col(A) $\subseteq$ nulidad(A) y utilice el teorema del rango.]

64. Sea A una matriz antisimétrica $n \times n$. (Vea los ejercicios en la sección 3.2.)
(a) Demuestre que $\mathbf{x}^T A\mathbf{x} = 0$ para toda $\mathbf{x}$ 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, comenzaremos a explorar uno de los temas de la introducción a este capítulo. Allí vimos que las matrices pueden ser utilizadas para transformar vectores cuando actúan 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 haremos más precisa esta noción y examinaremos varios ejemplos de tales transformaciones matriciales, para llegar al concepto de *transformación lineal*: una idea poderosa que encontraremos repetidamente a partir de ahora.

Comencemos por recordar algunos de los conceptos básicos asociados con las funciones. Usted estará familiarizado con la mayor parte de estas ideas de otros cursos en los cuales haya encontrado funciones de la forma $f: \mathbb{R} \to \mathbb{R}$ [tal como $f(x) = x^2$] que transforman números reales en números reales. Lo que es nuevo aquí es que se encuentran vectores involucrados y estamos interesados solamente en funciones que sean "compatibles" con las operaciones vectoriales de adición y multiplicación por escalares.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Demostración Sean **u** y **v** vectores de $\mathbb{R}^n$ 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 lo tanto, T_A es una transformación lineal.

Ejemplo 3.56

Sea f: $\mathbb{R}^2 \to \mathbb{R}^2$ la transformación que manda a cada punto hacia su punto de reflexión sobre el eje x. Muestre que F es una transformación lineal.

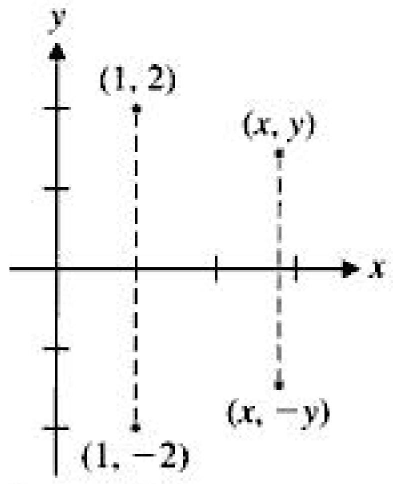


Figura 3.4
Reflexión en el eje x.

Solución Con base en la figura 3.4, se hace evidente que F envía el punto (x, y) al punto $(x, -y)$. De este modo, podemos describir

$$F\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

Podríamos proceder a verificar que F es lineal, como en el ejemplo 3.55 (¡esto es aún más fácil de verificar!), 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 lo 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 manera que F es una transformación matricial. Se sigue ahora, según el teorema 3.30, que F es una transformación lineal.

Ejemplo 3.57

Sea $R : \mathbb{R}^2 \to \mathbb{R}^2$ a transformación que gira cada punto a 90° en el sentido contrario de las manecillas del reloj con respecto al origen. Muestre que R es una transformación lineal.

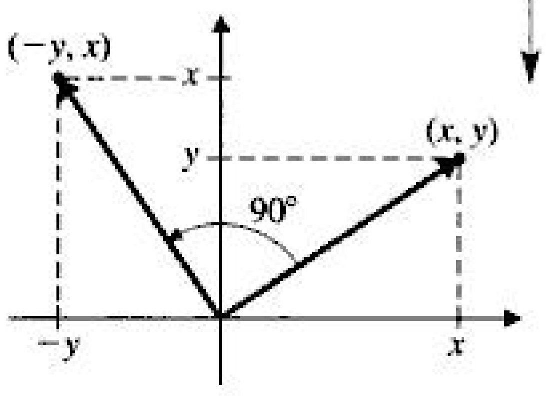


Figura 3.5
Una rotación de 90°.

Solución Como lo muestra la figura 3.5, R manda el punto (x, y) al punto $(-y, x)$. De esta manera, tenemos que

$$R\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} = x\begin{bmatrix} 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Por consiguiente, R es una transformación matricial y por lo tanto es lineal.

Observe que si multiplicamos una matriz por vectores de la base estándar, obtenemos las columnas de la matriz. Por ejemplo,

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \\ e \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \\ f \end{bmatrix}$$

Podemos utilizar esta observación para demostrar que *toda* transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ surge como una transformación matricial.

Para encontrar la matriz estándar de P_ℓ, aplicamos el teorema 3.31. Si $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$, entonces

$$P_\ell(\mathbf{e}_1) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_1}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_1}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 \\ d_1 d_2 \end{bmatrix}$$

y

$$P_\ell(\mathbf{e}_2) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_2}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_2}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1 d_2 \\ d_2^2 \end{bmatrix}$$

Así, la matriz estándar de la proyección es

$$A = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 & d_1 d_2 \\ d_1 d_2 & d_2^2 \end{bmatrix} = \begin{bmatrix} d_1^2/(d_1^2 + d_2^2) & d_1 d_2/(d_1^2 + d_2^2) \\ d_1 d_2/(d_1^2 + d_2^2) & d_2^2/(d_1^2 + d_2^2) \end{bmatrix}$$

Como verificación, advierta que en el inciso (a) podríamos tomar $\mathbf{d} = \mathbf{e}_1$ como un vector de dirección para el eje x. Por consiguiente, $d_1 = 1$ y $d_2 = 0$, y obtenemos que $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, como en el caso anterior.

Nuevas transformaciones lineales a partir de las antiguas

Si $T : \mathbb{R}^m \to \mathbb{R}^n$ y $S : \mathbb{R}^n \to \mathbb{R}^p$ son transformaciones lineales, entonces podemos poner S enseguida de T para formar la ***composición*** de las dos transformaciones, denotada como $S \circ T$. Nótese que, a fin 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 desde el dominio de T hasta el 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 la composición de transformaciones se toma directamente de esta figura y es la misma que la correspondiente definición de la composición de funciones ordinarias:

$$(S \circ T)(\mathbf{v}) = S(T(\mathbf{v}))$$

Naturalmente, nos gustaría que $S \circ T$ fuera también una transformación lineal, y felizmente encontramos que así es. Esta afirmación la podemos demostrar si comprobamos que $S \circ T$ satisface la definición de una transformación lineal (lo cual realizaremos en el capítulo 6), pero, puesto que por el momento suponemos que las transformaciones lineales y las transformaciones matriciales son la misma cosa, será suficiente demostrar que $S \circ T$ es una transformación matricial. Utilizaremos la notación $[T]$ para la matriz estándar de una transformación lineal T.

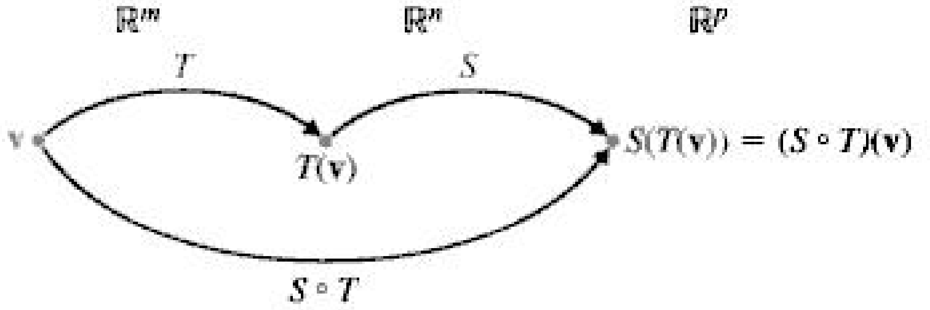


Figura 3.12
Composición de transformaciones

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

(En el ejercicio 29, se le pedirá que verifique este resultado al establecer que

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

y sustituir estos valores en la definición de S, calculando por ello $(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ de manera directa.)

Ejemplo 3.61

Encuentre la matriz estándar de la transformación que primero gira un punto a 90° en dirección contraria a las manecillas del reloj con respecto al origen y luego refleja el resultado en el eje x.

Solución La rotación R y la reflexión F fueron expuestas en los ejemplos 3.57 y 3.56, respectivamente, donde hallamos sus matrices estándar $[R] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ y $[F] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Se sigue que la composición $F \circ R$ tiene para su matriz

$$[F \circ R] = [F][R] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

(Verifique que este resultado es correcto al considerar el efecto de $F \circ R$ sobre los vectores de la base estándar $\mathbf{e}_1$ y $\mathbf{e}_2$. Advierta la importancia del *orden* de las transformaciones: R se realiza antes que F, pero escribimos $F \circ R$. En este caso, $R \circ F$ también tiene sentido. ¿Es $R \circ F = F \circ R$?)

Inversas de transformaciones lineales

Considere el efecto de una rotación de 90° en sentido contrario al de las manecillas del reloj con respecto al origen, seguida por una rotación de 90° en sentido de las manecillas del reloj con respecto al origen. Evidentemente esto deja cada punto de $\mathbb{R}^2$ sin modificación. Si denotamos estas transformaciones medianteR_{90} y R_{-90} (recuerde que la medida de un ángulo negativo corresponde a la dirección en sentido de las manecillas del reloj), entonces podemos expresar esto como $(R_{90} \circ R_{-90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^2$. Advierta que, en este caso, si realizamos la transformación en el otro orden, obtendremos el mismo resultado: $(R_{-90} \circ R_{90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^2$.

De este modo, $R_{90} \circ R_{-90}$ (y también $R_{-90} \circ R_{90}$) es una transformación lineal que deja cada vector de $\mathbb{R}^2$ sin cambio. Una transformación de esta clase se conoce como ***transformación identidad***. Por lo general, tenemos una transformación así para todo $\mathbb{R}^n$, a saber, $I : \mathbb{R}^n \to \mathbb{R}^n$, tal que $I(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ de $\mathbb{R}^n$. (Si es importante seguir la pista de la dimensión del espacio, podemos escribir I_n para mayor claridad).

Así, con esta notación, tenemos que $R_{90} \circ R_{-90} = I = R_{-90} \circ R_{90}$. Un par de transformaciones que se encuentren relacionadas entre sí de esta manera recibe el nombre de ***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 ocurre que $S \circ T = I_n$ y $T \circ S = I_n$

Observación Debido a que esta definición es simétrica con respecto a S y T, afirmaremos que, cuando esta situación ocurra, S será la inversa de T y T será la inversa de S. Además, diremos que S y T son ***invertibles***.

En términos de matrices, vemos 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 identidad es la matriz estándar de la transformación identidad?) También debemos tener que $[T][S] = [T \circ S] = [I] = I$. Esto muestra que $[S]$ y $[T]$ son matrices inversas. También muestra algo más: si una transformación lineal T es invertible, entonces su matriz estándar $[T]$ debe ser invertible, y puesto que las matrices inversas son únicas, esto significa que la inversa de T también es única. Por consiguiente, podemos utilizar sin ambigüedad la notación T^{-1} para hacer referencia a *la* inversa de T. De este modo, podemos reescribir las ecuaciones anteriores como $[T][T^{-1}] = I = [T^{-1}][T]$, la que muestra que la matriz de T^{-1} es la matriz inversa de $[T]$. Hemos terminado de probar 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

$$[T^{-1}] = [T]^{-1}$$

Observación Expresemos esta fórmula también en palabras: "la matriz de la inversa es la inversa de la matriz". ¡Es fabuloso!

Ejemplo 3.62

Encuentre la matriz estándar de una rotación de 60° en sentido de las manecillas del reloj con respecto al origen en $\mathbb{R}^2$.

Solución Con anterioridad calculamos la matriz de una rotación de 60° en sentido contrario a las manecillas del reloj con respecto al origen como

$$[R_{60}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$$

Puesto que una rotación de 60° en el sentido de las manecillas del reloj es la inversa de una rotación de 60° en sentido contrario, podemos aplicar el teorema 4 para obtener

$$[R_{-60}] = [(R_{60})^{-1}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}^{-1} = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{bmatrix}$$

(Verifique el cálculo de la inversa de la matriz. La manera más rápida es utilizar el atajo de 2×2 del teorema 3.8. También, verifique que la matriz resultante tiene el efecto correcto sobre la base estándar de $\mathbb{R}^2$ dibujando 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 $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, la cual no es invertible puesto que su determinante es cero. Por lo tanto, P tampoco lo es.

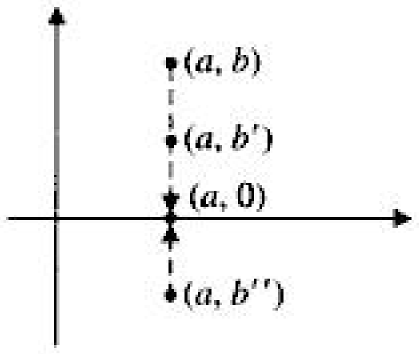


Figura 3.13
Las proyecciones no son invertibles

Observación La figura 3.13 da alguna idea de las razones por las cuales P, en el ejemplo 3.63, no es invertible. La proyección "colapsa" a $\mathbb{R}^2$ sobre el eje x. Para que P sea invertible, deberíamos tener una manera de "deshacerla", para recuperar el punto (a, b) con el que comenzamos. Sin embargo, existe un número infinito de candidatos para la imagen de $(a, 0)$ bajo una "inversa" hipotética de esa clase. ¿Cuál deberíamos utilizar? No podemos decir simplemente que P^{-1} debe enviar $(a, 0)$ hacia (a, b), puesto que esto no puede ser una *definición* cuando no tenemos manera de saber lo que b debería ser. (Véase el ejercicio 42.)

Asociatividad

El teorema 3.3(a) de la sección 3.2 estableció la propiedad asociativa de la multiplicación de matrices: $A(BC) = (AB)C$. (Si usted no intentó demostrarlo entonces, hágalo ahora. Incluso con todas las matrices restringidas a 2×2, usted percibirá la complejidad de la notación involucrada en una demostración "por elementos", lo que debería hacerle apreciar la demostración que estamos por hacer.)

Nuestro enfoque para demostrar lo que decimos se basará en las transformaciones lineales. Hemos visto que toda matriz A de $m\times n$ da origen a una transformación lineal $T_A : \mathbb{R}^n \to \mathbb{R}^m$; recíprocamente, toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ tiene una matriz $[T]$ de $m\times n$ correspondiente. Las dos correspondencias se encuentran inversamente relacionadas; es decir, dada A, $[T_A] = A$, y dada T, $T_{[T]} = T$.

Sean $R = T_A$, $S = T_B$ 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 demostraremos la última identidad. Sea $\mathbf{x}$ en el dominio de T (y por ello en el dominio tanto de $R \circ (S \circ T)$ como en el de $(R \circ S) \circ T$ ¿por qué?). Para demostrar que $R \circ (S \circ T) = (R \circ S) \circ T$, es suficiente probar que tienen el mismo efecto sobre $\mathbf{x}$. Mediante repetidas aplicaciones de la definición de composición, tenemos que

$$\begin{aligned}(R \circ (S \circ T))(\mathbf{x}) &= R((S \circ T)(\mathbf{x})) \\ &= R(S(T(\mathbf{x}))) \\ &= (R \circ S)(T(\mathbf{x})) = ((R \circ S) \circ T)(\mathbf{x})\end{aligned}$$

como se requería. (Verifique cuidadosamente que la definición de composición se ha utilizado cuatro veces.)

Esta sección ha servido como una introducción a las transformaciones lineales. En el capítulo 6, daremos otra mirada más detallada y más general a este tema. Los ejercicios que siguen también contienen algunas exploraciones adicionales de este importante concepto.

Ejercicio 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 $T_A : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación matricial correspondiente a $A = \begin{bmatrix} 4 & 0 & -1 \\ -2 & 1 & 3 \end{bmatrix}$. Encuentre $T_A(\mathbf{u})$ y $T_A(\mathbf{v})$, donde $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}$.

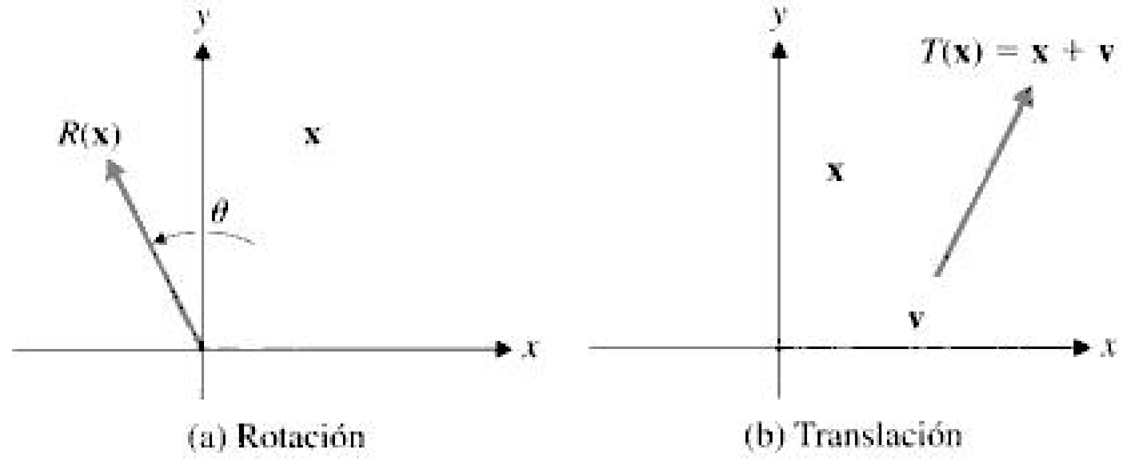


Figura 3.16

Desafortunadamente, la translación no es una transformación lineal $T(\mathbf{0}) \neq \mathbf{0}$. Sin embargo, existe un truco para rodear el problema. Podemos 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 el vector trasladado $T(\mathbf{x})$ en coordenadas homogéneas.

Podemos tratar las rotaciones en coordenadas homogéneas también. La multiplicación de matrices

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \\ 1 \end{bmatrix}$$

representa el vector rotado $R(\mathbf{x})$ en coordenadas homogéneas. La composición $T \circ R$ que da la rotación R seguida de la translación T está representada ahora por el producto

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & a \\ \operatorname{sen}\theta & \cos\theta & b \\ 0 & 0 & 1 \end{bmatrix}$$

(Observe que $R \circ T \neq T \circ R$.)

Para modelar un brazo robótico, le damos a cada cadena su propio sistema de coordenadas (denominado *marco*) y examinamos cómo se mueve una cadena en relación con aquellos con los cuales está directamente conectada. Para ser específicos, sean X_i y Y_i los ejes coordenados para la cadena $\mathcal{A}_i$, con el eje X_i alineado con la cadena. La longitud de $\mathcal{A}_i$ se denota por a_i, y el ángulo entre X_i y X_{i-1} se denota mediante θ_i. La unión entre $\mathcal{A}_i$ y $\mathcal{A}_{i-1}$ está en el punto $(0, 0)$ en relación con A_i y $(a_{i-1}, 0)$ respecto de $\mathcal{A}_{i-1}$. Por lo tanto, en relación con $\mathcal{A}_{i-1}$, el sistema de coordenadas para $\mathcal{A}_i$, ha sido

Podemos resumir estas dos ecuaciones en una sola ecuación matricial:

$$\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 120 \\ 80 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$$

Llamaremos a la matriz P y denominaremos a los vectores $\mathbf{x}_0 = \begin{bmatrix} 120 \\ 80 \end{bmatrix}$ y $\mathbf{x}_1 = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$. (Advierta que las componentes de cada vector son los números de los usuarios de la marca A y de la marca B, en ese orden, después del número de meses indicado por el subíndice.) De este modo, tenemos $\mathbf{x}_1 = P\mathbf{x}_0$.

Extenderemos la notación, sea $\mathbf{x}_k$ el vector cuyas componentes registran la distribución de los usuarios de pasta dental después de k meses. Para determinar el número de usuarios de cada marca después de haber transcurrido dos meses, simplemente aplicamos el mismo razonamiento, comenzando con $\mathbf{x}_1$ en lugar de $\mathbf{x}_0$. Obtenemos que

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 90 \\ 110 \end{bmatrix}$$

de lo cual vemos que ahora hay 90 usuarios de la marca A y 110 usuarios de la marca B.

Los vectores $\mathbf{x}_k$ en el ejemplo 3.64 son conocidos como ***vectores de estado*** de la cadena de Markov, mientras que la matriz P se conoce como su ***matriz de transición.*** Acabamos de ver que una cadena de Markov satisface la relación

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{para } k = 0, 1, 2, \ldots$$

De este resultado se sigue que podemos calcular un vector de estado arbitrario *de manera iterativa*, una vez que conocemos $\mathbf{x}_0$ y P. En otras palabras, una cadena de Markov se encuentra *completamente determinada* por sus probabilidades de transición y por su estado inicial.

Observaciones

- Supongamos, en el ejemplo 3.64, que no queremos conocer los números *reales* de los usuarios de pastas dentales sino los números *relativos* de los que utilizan cada marca. Podríamos convertir los datos en porcentajes o fracciones dividiendo entre 200, el número total de usuarios. De este modo, comenzaríamos 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, la proporción marca A-marca B es, inicialmente, 60%-40%. Verifique mediante cálculo directo que $P\mathbf{x}_0 = \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}$, lo que puede tomarse entonces como $\mathbf{x}_1$ (en concordancia con la proporción 50-50 que calculamos anteriormente). Los vectores de esta clase, con componentes no negativos que suman 1, se denominan ***vectores de probabilidad.***

- Observe la forma en que las probabilidades de transición se acomodan dentro de la matriz de transición P. Podemos pensar en las columnas como los estados *presentes* y en los renglones como los estados *siguientes*:

$$\begin{array}{cc} & \textit{Presente} \\ & \begin{array}{cc} \text{A} & \text{B} \end{array} \\ \textit{Siguiente} \begin{array}{c} \text{A} \\ \text{B} \end{array} & \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \end{array}$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Observación Advierta que la matriz de adyacencia de un grafo es necesariamente una matriz simétrica (¿por qué?). Nótese también que una entrada diagonal a_{ii} de A es cero a menos que exista un lazo, bucle o circuito en el vértice i. En algunas situaciones, un grafo puede tener más de una línea entre un par de vértices. En tal caso, puede tener sentido modificar la definición de la matriz de adyacencia de modo que a_{ij} sea igual al *número* de líneas entre los vértices i y j.

Definimos una ***trayectoria*** en un grafo como una sucesión de líneas que nos permiten viajar de un vértice a otro de manera continua. La ***longitud*** de una trayectoria es el número de líneas que contiene, además, haremos referencia a una trayectoria con k líneas como una ***k-trayectoria***. Por ejemplo, en el grafo de la figura 3.25, $v_1v_3v_2v_1$ es una 3-trayectoria, mientras que $v_4v_1v_2v_2v_1v_3$ es una 5-trayectoria. Nótese que la primera de éstas es una trayectoria *cerrada* (comienza y finaliza en el mismo vértice); una trayectoria así se conoce como ***circuito***. La segunda utiliza la línea entre v_1 y v_2 dos veces; una trayectoria que *no* incluye la misma línea más de una vez se conoce como una trayectoria ***simple***.

Podemos utilizar las potencias de una matriz de adyacencia de un grafo para obtener información acerca de las trayectorias de diversas longitudes en el grafo. Considere el cuadrado de la matriz de adyacencia de 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? Examine la entrada (2, 3). Con base en la definición de multiplicación de matrices, sabemos 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 por lo menos uno de los productos $a_{2k}\,a_{k3}$ que forman 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 diferentes de cero, lo que significa que existe una línea entre v_2 y v_k, además de una línea entre v_k y v_3. De este modo, habrá una 2-trayectoria entre los vértices 2 y 3 (vía el vértice k). En nuestro ejemplo, esto ocurre para $k = 1$ y $k = 2$, de manera que

$$\begin{aligned}(A^2)_{23} &= a_{21}a_{13} + a_{22}a_{23} + a_{23}a_{33} + a_{24}a_{43} \\ &= 1\cdot 1 + 1\cdot 1 + 1\cdot 0 + 0\cdot 0 \\ &= 2\end{aligned}$$

lo que nos dice que existen dos 2-trayectorias entre los vértices 2 y 3. (Verifique para ver que las entradas restantes de A^2 proporcionan correctamente 2-trayectorias en el grafo.) El argumento que acabamos de dar puede ser generalizado para producir el resultado siguiente, cuya demostración dejamos como el ejercicio 54.

Si A es la matriz de adyacencia de un grafo G, entonces la entrada(i, j) entrada de A^k es igual al número de *k-trayectorias* entre los vértices i y j.

Ejemplo 3.67 ¿Cuántas 3-trayectorias existen entre v_1 y v_2 en la figura 3.25?

Solución Necesitamos 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 nos da

$$(A^3)_{12} = 3\cdot 1 + 2\cdot 1 + 1\cdot 1 + 0\cdot 0 = 6$$

así que hay seis trayectorias 3 entre los vértices 1 y 2, lo que puede verificarse fácilmente.

recibido como $\mathbf{c}' = [1 \quad 0 \quad 1]^T$. Calculamos

$$P\mathbf{c}' = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \mathbf{0}$$

de modo que sabemos que $\mathbf{c}'$ no puede ser un vector de código. ¿Dónde está el error? Nótese que $P\mathbf{c}'$ es la segunda columna de la matriz de verificación de paridad P, lo cual nos dice que el error se encuentra en la segunda componente de $\mathbf{c}'$ (lo que demostraremos en el teorema 3.34, más adelante) y nos permite corregir el error. (Naturalmente, en este ejemplo podríamos hallar el error más rápidamente sin utilizar matrices, pero la idea es útil).

Para generalizar las ideas en el último ejemplo, precisemos las definiciones siguientes:

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 denomina una ***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 $P = [B \quad I_{n-k}]$, donde B es una matriz de $(n-k)\times k$ matrix sobre $\mathbb{Z}_2$, se conoce como ***matriz estándar de verificación de paridad.*** Se dice que el código tiene ***longitud*** n y ***dimensión*** k.

Aquí está lo que necesitamos conocer: (a) ¿Cuándo es G la matriz generadora estándar de un código binario de *corrección de errores*? (b) Dada G, ¿cómo encontramos una matriz estándar asociada de verificación de paridad P? Resulta que las respuestas son bastante sencillas, como se muestra mediante el siguiente teorema.

Teorema 3.34 Si $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$ es una matriz generadora estándar y $P = [B \quad I_{n-k}]$ es una matriz estándar de verificación de paridad,, entonces P es la matriz de verificación de paridad asociada con G si y sólo si $A = B$. El código binario (n, k) correspondiente es de corrección del error (simple) si y sólo si las columnas de P no son iguales a cero y son distintas.

Antes que demostremos el teorema, consideremos otro ejemplo menos trivial que lo ilustra.

Ejemplo 3.70 Supongamos que queremos diseñar un código de corrección del error que utilice tres ecuaciones de verificación de paridad. Debido a que estas ecuaciones dan origen a los renglones de P, tenemos que $n - k = 3$ y $k = n - 3$. Los vectores de mensaje vienen de $\mathbb{Z}_2^k$, así que nos gustaría que k (y por lo tanto, n) sean tan grandes como fuera posible a fin de que podamos transmitir tanta información como podamos. Según el teorema 3.34, sabemos que las n columnas de P necesitan ser diferentes de cero y distintas, así que el máximo ocurre cuando se componen de todos los $2^3 - 1 = 7$ vectores distintos de cero de $\mathbb{Z}_2^{n-k} = \mathbb{Z}_2^3$. Un candidato de esta clase es

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

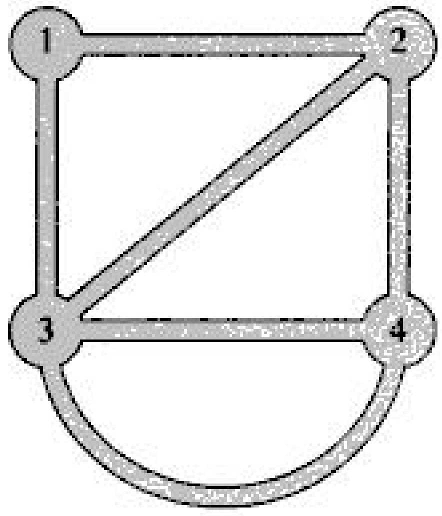


Figura 3.28

(a) Construya la matriz de transición para la cadena de Markov que modela esta situación.

(b) Supongamos que iniciamos con 15 robots en cada cruce. Encuentre la distribución de estado estacionario de los robots. (Suponga que a cada uno de ellos le toma la misma cantidad de tiempo viajar entre dos cruces adyacentes).

13. Denotemos por $\mathbf{j}$ un vector renglón compuesto enteramente por "1". Demuestre que una matriz no negativa P es una matriz estocástica si y sólo si $\mathbf{j}P = \mathbf{j}$.

14. (a) Muestre que el producto de dos matrices estocásticas de 2×2 también es una matriz estocástica.

(b) Demuestre que el producto de dos matrices estocásticas de $n\times n$ también es una matriz estocástica.

Supongamos que queremos conocer el número promedio (o esperado) de pasos necesarios para ir del estado i al estado j en una cadena de Markov. Puede demostrarse que el siguiente cálculo responde esta cuestión: elimine el j-ésimo renglón y la j-ésima columna de la matriz de transición P para obtener una nueva matriz Q. (Mantenga los renglones y las columnas de Q etiquetadas como estaban en P.) El número esperado de pasos desde el estado i hasta el estado j es igual a la suma de las entradas de la columna de $(I - Q)^{-1}$ etiquetada i.

15. En el ejercicio 9, si el lunes es un día sin lluvia, ¿cuál es el número esperado de días que pasarán hasta que llegue un día lluvioso?

16. En el ejercicio 10, ¿cuál es el número esperado de generaciones hasta que una persona de baja estatura tenga un descendiente alto?

17. En el ejercicio 11, si la cosecha de piñón es regular un año, ¿cuál es el número esperado de años hasta que se presente una buena cosecha?

18. En el ejercicio 12, a partir de cada uno de los otros cruces, ¿cuál es el número esperado de movimientos hasta que un robot alcance el cruce 4?

Crecimiento poblacional

19. Una población con tres clases de edades 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 $\mathbf{x}_3$.

20. Una población con 4 clases de edades 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 $\mathbf{x}_3$.

21. Una cierta especie con dos clases de edad de un año de duración tiene una probabilidad de supervivencia de 80% de la clase 1 a la clase 2. La evidencia empírica muestra que, en promedio, cada hembra da a luz a cinco hembras por año. De este modo, son dos posibles matrices de Leslie

$$L_1 = \begin{bmatrix} 0 & 5 \\ 0.8 & 0 \end{bmatrix} \quad \text{y} \quad L_2 = \begin{bmatrix} 4 & 1 \\ 0.8 & 0 \end{bmatrix}$$

(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 de edad con respecto al tiempo (como en la figura 3.23). ¿Qué es lo que sugieren sus gráficas?

22. Supongamos 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 una $\mathbf{x}_0$ arbitraria, determine el comportamiento de esta población.

23. Supongamos 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 que tendrá variar la probabilidad de supervivencia s de los escarabajos menores.

CAS **24.** El caribú del bosque se encuentra principalmente en las provincias occidentales de Canadá y el noroeste de Estados Unidos. La vida promedio de una hembra es de aproximadamente 14 años. Las tasas de nacimiento

57. La matriz de adyacencia en el ejercicio 32

58. La matriz de adyacencia en el ejercicio 31

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

60. **(a)** Demuestre que un grafo es bipartito si y sólo si sus vértices pueden ser etiquetados de modo que su matriz de adyacencia pueda ser particionada como

$$A = \left[\begin{array}{c|c} O & B \\ \hline B^T & O \end{array}\right]$$

(b) Utilizando el resultado del inciso (a), demuestre que un grafo bipartito no tiene circuitos de longitud impar.

Códigos de corrección del error

61. Supongamos que codificamos los cuatro vectores en $\mathbb{Z}_2^2$ repitiendo el vector dos veces. De este modo, tenemos que

$$[0, 0] \to [0, 0, 0, 0]$$
$$[0, 1] \to [0, 1, 0, 1]$$
$$[1, 0] \to [1, 0, 1, 0]$$
$$[1, 1] \to [1, 1, 1, 1]$$

Demuestre que este código no es de corrección del error.

62. Supongamos que codificamos los dígitos binarios 0 y 1 repitiendo cada dígito cinco veces. Así,

$$0 \to [0, 0, 0, 0, 0]$$
$$1 \to [1, 1, 1, 1, 1]$$

Demuestre que este código puede corregir errores dobles.

¿Cuál es el resultado de codificar los mensajes en los ejercicios 63 a 65, utilizando el código Hamming (7, 4) del ejemplo 3.70?

63. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ **64.** $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ **65.** $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Cuando el código Hamming (7, 4) del ejemplo 3.70 se utiliza, supongamos que se reciben los mensajes $\mathbf{c}'$ en los ejercicios 66 a 68. Aplique la matriz estándar de verificación de paridad a $\mathbf{c}'$ para determinar si se ha presentado un error y decodificar correctamente $\mathbf{c}'$ para recobrar el vector de mensaje original $\mathbf{x}$.

66. $\mathbf{c}' = [0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1]^T$

67. $\mathbf{c}' = [1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0]^T$

68. $\mathbf{c}' = [0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0]^T$

69. El código de verificación de paridad del ejemplo 1.31 es un código $\mathbb{Z}_2^6 \to \mathbb{Z}_2^7$.

(a) Encuentre la matriz estándar de verificación de paridad de este código

(b) Encuentre una matriz generadora estándar.

(c) Aplique el teorema 3.34 para explicar por qué este código no es de corrección del error.

70. Defina un código $\mathbb{Z}_2^2 \to \mathbb{Z}_2^5$ empleando la matriz generadora estándar

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

(a) Enumere todas las cuatro palabras del código.

(b) Encuentre la matriz estándar asociada de verificación de paridad de este código. ¿Es este código de corrección del error (simple)?

71. Defina un código $\mathbb{Z}_2^3 \to \mathbb{Z}_2^6$ mediante el empleo de 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) Enumere todas las ocho palabras de código.

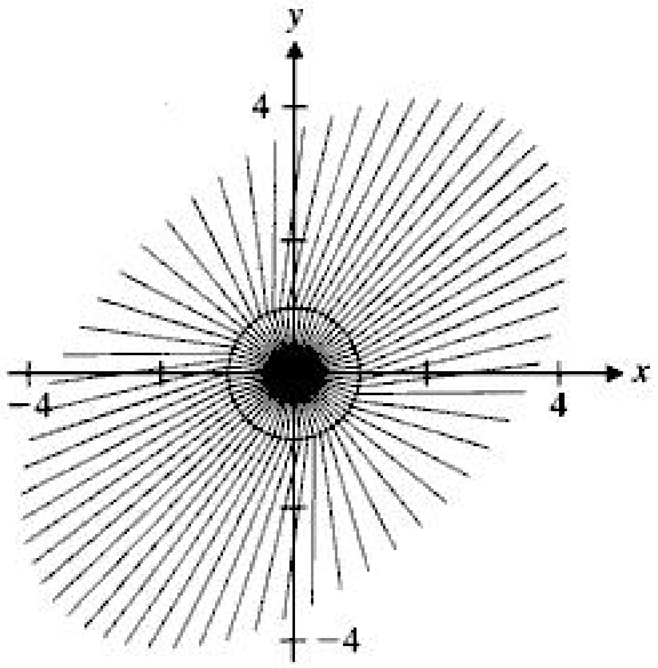
(b) Encuentre la matriz estándar asociada de verificación de paridad para este código. ¿Es este código de corrección de errores (simples)?

72. Demuestre que el código del ejemplo 3.69 es un código Hamming (3, 1).

73. Construya las matrices estándar de verificación de paridad y generadora de un código Hamming (15, 11).

74. En el teorema 3.34, demuestre que si $B = A$, entonces $PG\mathbf{x} = \mathbf{0}$ para todo $\mathbf{x}$ de $\mathbb{Z}_2^k$.

75. En el teorema 3.34, demuestre que si $\mathbf{p}_i = \mathbf{p}_j$, no podemos determinar si ocurre un error en la i-ésima o j-ésima componente del vector recibido.



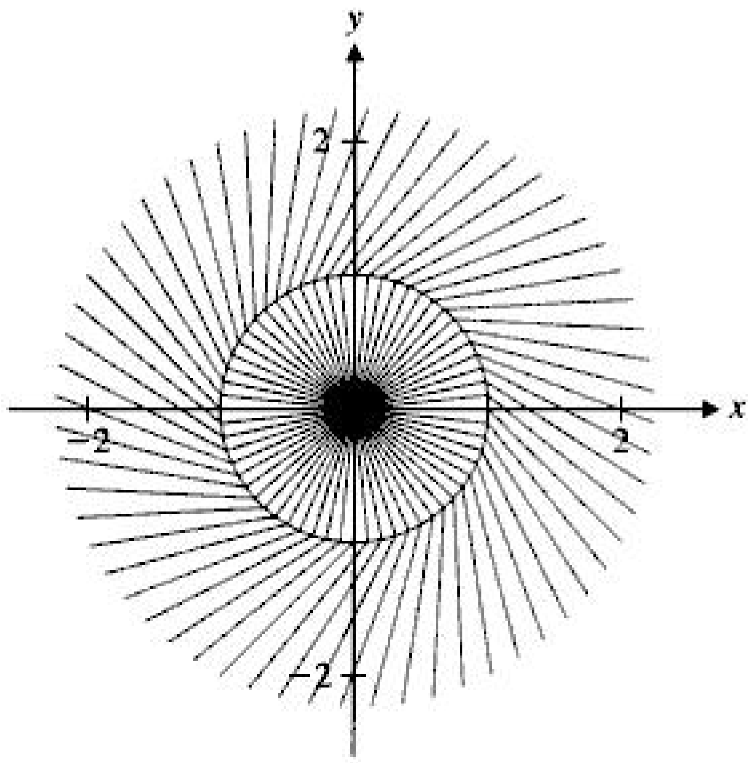


Figura 4.7

En la figura 4.7(b), observamos lo que ocurre cuando utilizamos la matriz $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. ¡No hay eigenvectores!

Ahora sabemos cómo encontrar los eigenvectores una vez que tenemos los eigenvalores correspondientes, y tenemos una interpretación geométrica de los mismos, pero queda una pregunta: ¿cómo encontramos primero los eigenvalores de una matriz dada? La respuesta es por la observación de que λ es un eigenvalor de A si y sólo si el espacio nulo de $A - \lambda I$ es no trivial.

Recordemos de la sección 3.3 que el determinante de una matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ de expresión 2×2 es la expresión $\det A = ad - bc$, y A es invertible si y sólo si $\det A$ es distinto de cero. Además, el teorema fundamental de las matrices invertibles garantiza que una matriz tenga un espacio nulo no trivial si y sólo si es no invertible; por lo tanto, si y sólo si su determinante es cero. Si ponemos estos hechos juntos, deducimos que (al menos para matrices de 2×2), λ es un eigenvalor de A si y sólo si $\det(A - \lambda I) = 0$. Este hecho caracteriza los eigenvalores, y pronto lo generalizaremos a matrices cuadradas de tamaño arbitrario. No obstante, por el momento, sólo lo emplearemos con matrices de 2×2.

Ejemplo 4.5

Encuentre todos los eigenvalores de la matriz $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ del ejemplo 4.1.

Solución Las observaciones anteriores muestran que debemos determinar todas las soluciones λ de la ecuación $\det(A - \lambda I) = 0$. Puesto que

$$\det(A - \lambda I) = \det\begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix} = (3-\lambda)(3-\lambda) - 1 = \lambda^2 - 6\lambda + 8$$

necesitamos resolver la ecuación cuadrática $\lambda^2 - 6\lambda + 8 = 0$. Se halla fácilmente que las soluciones a esta ecuación son $\lambda = 4$ y $\lambda = 2$. Éstos son, por lo tanto, los eigenvalores de A.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

diagonales que aparecen abajo. Coloque los signos de más a los productos de las diagonales descendentes y los signos de menos a los productos provenientes de las diagonales ascendentes.

$$\begin{array}{ccccc} & & - & - & - \\ \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right. & & \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{array} \\ & & + & + & + \end{array} \tag{2}$$

Este método nos da

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

En el ejercicio 19 se le pedirá que verifique que este resultado concuerde con el de la ecuación (1) para un determinante de 3×3.

Ejemplo 4.9

Calcule el determinante de la matriz del ejemplo 4.8 usando el método que se presenta en (2).

Solución Adjuntamos a la matriz A las primeras dos columnas y calculamos los seis productos indicados:

$$\begin{array}{ccccc} & & 0 & -10 & -9 \\ \left[\begin{array}{ccc} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{array}\right] & & \begin{array}{cc} 5 & -3 \\ 1 & 0 \\ 2 & -1 \end{array} \\ & & 0 & -12 & -2 \end{array}$$

Si sumamos los tres productos de la parte inferior y restamos los tres productos de la parte superior obtenemos

$$\det A = 0 + (-12) + (-2) - 0 - (-10) - (-9) = 5$$

como antes.

Advertencia Estamos a punto de definir determinantes para matrices cuadradas arbitrarias. Sin embargo, *no* hay análogo del método del ejemplo 4.9 para matrices más grandes. Es válido *solamente* para matrices de 3×3.

Determinantes de matrices de $n \times n$

La definición del determinante de una matriz de 3×3 se extiende en forma natural a matrices cuadradas arbitrarias.

Definición Sea $A = [a_{ij}]$ una matriz de $n \times n$, donde $n \geq 2$. Entonces, el ***determinante*** de A es el escalar

$$\begin{aligned} \det A = |A| &= a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} \end{aligned} \tag{3}$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

El ejemplo 4.12 debería convencerle de que el determinante de una matriz triangular es el producto de sus entradas diagonales. Se le pedirá que demuestre este hecho en el ejercicio 21. Registraremos el resultado como un teorema.

Teorema 4.2 El determinante de una matriz triangular es el producto de las entradas sobre su diagonal principal. En concreto, si $A = [a_{ij}]$ es una matriz triangular de $n\times n$, entonces

$$\det A = a_{11}a_{22}\cdots a_{nn}$$

Nota En general (es decir, a menos que la matriz sea triangular o tenga alguna otra forma especial), el cálculo de un determinante por medio del desarrollo por cofactores no es eficiente. Por ejemplo, el determinante de una matriz de 3×3 tiene $6 = 3!$ sumandos, cada uno de los cuales requiere de dos multiplicaciones, y posteriormente es necesario efectuar cinco sumas y restas para finalizar los cálculos. Una matriz de $n\times n$ tendrá $n!$ sumandos, cada uno con $n-1$ multiplicaciones y después $n! - 1$ sumas y restas. El número total de operaciones será entonces de

$$T(n) = (n-1)n! + n! - 1 > n!$$

Incluso la más rápida de las supercomputadoras no puede calcular el determinante de una matriz moderadamente grande por medio del desarrollo por cofactores. Por ejemplo, supongamos que se necesita estimar un determinante de 50×50. (Matrices *mucho más* grandes que una de 50×50 se emplean para almacenar los datos de imágenes digitales como las transmitidas a través de Internet o las que se capturan por medio de una cámara digital.) Para calcular el determinante en forma directa requerimos, en general, más de 50! operaciones, y $50! \approx 3 \times 10^{64}$. Si tuviéramos una computadora que pudiera llevar a cabo un billón (10^{12}) de operaciones por segundo, nos tomaría aproximadamente 3×10^{52} segundos, o casi 10^{45} años, concluir los cálculos. Para poner estas cifras en perspectiva, considere que los astrónomos estiman la edad del universo en, por lo menos, 10 mil millones (10^{10}) de años. De este modo, incluso en una supercomputadora muy rápida, ¡calcular un determinante de 50×50 a través del desarrollo por cofactores implicaría más tiempo que el equivalente 10^{30} la edad del universo!

Por fortuna, existen mejores métodos; y ahora regresaremos para desarrollar medios computacionales más eficaces para encontrar determinantes. En primer lugar, debemos examinar algunas de las propiedades de los determinantes.

Propiedades de los determinantes

La manera más eficiente de calcular determinantes es mediante el empleo de la reducción por renglones. A pesar de ello, no todas las operaciones elementales por renglón dejan el determinante de una matriz sin cambios. El teorema siguiente resume las propiedades principales que usted necesita entender con el propósito de aplicar de manera eficiente la reducción por renglones.

Teorema 4.3 Sea $A = [a_{ij}]$ una matriz cuadrada.

a. Si A tiene un renglón (columna) cero, entonces $\det A = 0$.
b. Si B se obtiene al intercambiar dos renglones (columnas) de A, entonces $\det B = -\det A$.
c. Si A tiene dos renglones (columnas) idénticos, entonces $\det A = 0$.
d. Si B se obtiene al multiplicar un renglón (columna) de A por k, entonces $\det B = k \det A$.
e. Si A, B y C son idénticas excepto que el i-ésimo renglón (columna) de C sea la suma de los i-ésimos renglones (columnas) de A y B, entonces $\det C = \det A + \det B$.
f. Si B se obtiene al sumar un múltiplo de un renglón (columna) de A a otro renglón (columna), entonces $\det B = \det A$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

de modo que $x_3 = t$ es libre y $x_1 = -x_3 = -t$ y $x_2 = 3x_3 = 3t$. En consecuencia,

$$E_{-2} = \left\{ \begin{bmatrix} -t \\ 3t \\ t \end{bmatrix} \right\} = \left\{ t \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right\} = \text{generado}\left(\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right)$$

Se sigue, entonces, que $\lambda_1 = \lambda_2 = 0$ tiene multiplicidad geométrica 2 mientras que $\lambda_3 = -2$ tiene una multiplicidad geométrica 1. (Nótese que la multiplicidad algebraica es igual a la multiplicidad geométrica de cada eigenvalor.)

En algunas situaciones, los eigenvalores de una matriz son fáciles de localizar. Si A es una matriz triangular, entonces también lo es $A - \lambda I$, y el teorema 4.2 nos dice que el $\det(A - \lambda I)$ es precisamente el producto de las entradas de la diagonal, lo cual implica que la ecuación característica de una matriz triangular es

$$(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) = 0$$

de lo que de inmediato se deriva que los eigenvalores son $\lambda_1 = a_{11}, \lambda_2 = a_{22}, \ldots, \lambda_n = a_{nn}$. Resumiremos este resultado como un teorema y lo ilustraremos con un ejemplo.

Teorema 4.15 Los eigenvalores de una matriz triangular son las entradas de su diagonal principal.

Ejemplo 4.20 Los eigenvalores de

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 5 & 7 & 4 & -2 \end{bmatrix}$$

son $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 3, \lambda_4 = -2$, de acuerdo con el teorema 4.15. [En realidad, el polinomio característico es $(2 - \lambda)(1 - \lambda)(3 - \lambda)(-2 - \lambda)$.]

Nótese que las matrices diagonales son un caso especial del teorema 4.15. De hecho, una matriz diagonal es tanto triangular superior como triangular inferior.

Los eigenvalores contienen mucha información esencial acerca del comportamiento de una matriz. Una vez que conocemos los eigenvalores de una matriz, podemos deducir una gran cantidad de derivaciones sin hacer ningún trabajo adicional. El siguiente teorema es uno de los más importantes en este sentido.

Teorema 4.16 Una matriz cuadrada A es invertible si y sólo si 0 *no es* un eigenvalor de A.

Demostración Sea A una matriz cuadrada. De acuerdo con el teorema 4.6, A es invertible si y sólo si $\det A \neq 0$. Pero $\det A \neq 0$ es equivalente a $\det(A - 0I) \neq 0$, lo que nos dice que 0 no es una raíz de la ecuación característica de A (es decir, 0 no es un eigenvalor de A).

Ahora podemos extender el teorema fundamental de las matrices invertibles para incluir los resultados que hemos demostrado en este capítulo.

Así, $AP = PB$ con $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. (Note que no es necesario calcular P^{-1}. Véase la primera de las observaciones anteriores.)

Teorema 4.21 Sean A, B y C matrices de $n\times n$

a. $A \sim A$.
b. Si $A \sim B$, entonces $B \sim A$.
c. Si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

Demostración (a) Esta propiedad se desprende del hecho de que $I^{-1}AI = A$.

(b) Si $A \sim B$, entonces $P^{-1}AP = B$ para alguna matriz invertible P. Como se definió en la primera de las observaciones anteriores, esto es equivalente a $PBP^{-1} = A$. Si establecemos $Q = P^{-1}$, tenemos que $Q^{-1}BQ = (P^{-1})^{-1}BP^{-1} = PBP^{-1} = A$. Por lo tanto, por definición, $B \sim A$.

(c) Se le pedirá que demuestre esta propiedad (c) en el ejercicio 30.

Observación Cualquier relación que satisfaga las tres propiedades en el teorema 4.21 se denomina ***relación de equivalencia.*** Con frecuencia, estas relaciones surgen en las matemáticas y los objetos que están relacionados por medio de ellas por lo regular comparten propiedades importantes. Estamos a punto de comprobar que eso es verdad con respecto a las matrices semejantes.

Teorema 4.22 Sean A y B matrices de $n\times n$ con $A \sim B$. Entonces

a. $\det A = \det B$.
b. A es invertible si y sólo si B también lo es.
c. A y B tienen el mismo rango.
d. A y B tienen el mismo polinomio característico.
e. A y B tienen los mismos eigenvalores.

Demostración Comprobaremos (a) y (d) y dejaremos las propiedades restantes como ejercicios. Si $A \sim B$, entonces $P^{-1}AP = B$ para alguna matriz invertible P.

(a) Si tomamos los determinantes de ambos lados, tenemos que

$$\begin{aligned} \det B &= \det(P^{-1}AP) = (\det P^{-1})(\det A)(\det P) \\ &= \left(\frac{1}{\det P}\right)(\det A)(\det P) = \det A \end{aligned}$$

(d) El polinomio característico de B es

$$\begin{aligned} \det(B - \lambda I) &= \det(P^{-1}AP - \lambda I) \\ &= \det(P^{-1}AP - \lambda P^{-1}IP) \\ &= \det(P^{-1}AP - P^{-1}(\lambda I)P) \\ &= \det(P^{-1}(A - \lambda I)P) = \det(A - \lambda I) \end{aligned}$$

con el último paso de la misma forma que en (a). Así, $\det(B - \lambda I) = \det(A - \lambda I)$; es decir, los polinomios característicos de B y A son los mismos.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

de lo que extraemos la solución $x_n = 3^n - 2^n$. (Para verificar nuestro trabajo, podríamos introducir $n = 1, 2, \ldots, 5$ para comprobar que con esta fórmula se obtienen los mismos términos que calculamos con el empleo de la relación de recurrencia. ¡Inténtelo!)

Observe que x_n es una combinación lineal de potencias de los eigenvalores. Ése es necesariamente el caso cuando los eigenvalores son distintos [como el teorema 4.38(a) lo hará explícito]. Con base en esta observación, podemos ahorrarnos algo de trabajo. Una vez que hemos calculado los eigenvalores $\lambda_1 = 3$ y $\lambda_2 = 2$, podemos apuntar de inmediato

$$x_n = c_1 3^n + c_2 2^n$$

donde c_1 y c_2 deberán ser determinados. Si se utilizan las condiciones iniciales,

$$1 = x_1 = c_1 3^1 + c_2 2^1 = 3c_1 + 2c_2$$

cuando $n = 1$ y

$$5 = x_2 = c_1 3^2 + c_2 2^2 = 9c_1 + 4c_2$$

cuando $n = 2$. A continuación resolvemos el sistema

$$\begin{aligned} 3c_1 + 2c_2 &= 1 \\ 9c_1 + 4c_2 &= 5 \end{aligned}$$

para c_1 y c_2 para obtener $c_1 = 1$ y $c_2 = -1$. De este modo, $x_n = 3^n - 2^n$, como antes.

Ese es el método que usaremos en la práctica. Ahora ejemplificaremos su uso para hallar una fórmula explícita para los números de Fibonacci.

Ejemplo 4.41

Resuelva la recurrencia de Fibonacci $f_0 = 0, f_1 = 1$ y $f_n = f_{n-1} + f_{n-2}$ para $n \geq 2$.

Solución Si escribimos la recurrencia como $f_n - f_{n-1} - f_{n-2} = 0$, detectamos que la ecuación característica es $\lambda^2 - \lambda - 1 = 0$, así que los eigenvalores son

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} \quad \text{y} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}$$

De la exposición anterior se desprende que la solución de la relación de recurrencia tiene la forma

$$f_n = c_1\lambda_1^n + c_2\lambda_2^n = c_1\left(\frac{1 + \sqrt{5}}{2}\right)^n + c_2\left(\frac{1 - \sqrt{5}}{2}\right)^n$$

para escalares c_1 y c_2.

Usando las condiciones iniciales, encontramos que

$$0 = f_0 = c_1\lambda_1^0 + c_2\lambda_2^0 = c_1 + c_2$$

y

$$1 = f_1 = c_1\lambda_1^1 + c_2\lambda_2^1 = c_1\left(\frac{1 + \sqrt{5}}{2}\right) + c_2\left(\frac{1 - \sqrt{5}}{2}\right)$$

Resolviendo para c_1 y c_2, obtenemos $c_1 = 1/\sqrt{5}$ y $c_2 = -1/\sqrt{5}$. Por lo tanto, una fórmula explícita del n-ésimo número de Fibonacci es

$$f_n = \frac{1}{\sqrt{5}}\left(\frac{1 + \sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1 - \sqrt{5}}{2}\right)^n \tag{5}$$

Jacques Binet (1786-1856) realizó contribuciones a la teoría de matrices, la teoría de números, la física y la astronomía. Descubrió la regla para la multiplicación de matrices en 1812. En realidad, la fórmula de Binet para los números de Fibonacci se debe a Euler, quien la publicó en 1765; sin embargo, fue olvidada hasta que Binet publicó su versión en 1843. Al igual que Cauchy, Binet fue realista y perdió su posición universitaria cuando Carlos X abdicó en 1830. Recibió muchos honores por su trabajo, incluyendo su nombramiento, en 1843, para la Académie des Sciences.

$a + bi$ **Ejemplo 4.45** Los petirrojos y las lombrices coexisten en un ecosistema. Los primeros se comen a las segundas, que son su única fuente de alimento. Las poblaciones de petirrojos y lombrices en el tiempo t años se denotan con $r(t)$ y $w(t)$, respectivamente. Las ecuaciones que gobiernan el crecimiento de las dos poblaciones son

$$\begin{aligned} r'(t) &= w(t) - 12 \\ w'(t) &= -r(t) + 10 \end{aligned} \tag{8}$$

Si inicialmente 6 petirrojos y 20 lombrices ocupan el ecosistema, determine el comportamiento de las dos poblaciones a lo largo del tiempo.

Solución Lo primero que observamos con respecto a este ejemplo es la presencia de las constantes extra, -12 y 10, en las dos ecuaciones. Por fortuna, podemos deshacernos de ellas con un simple cambio de variables. Si hacemos que $r(t) = x(t) + 10$ y $w(t) = y(t) + 12$, entonces $r'(t) = x'(t)$ y $w'(t) = y'(t)$. Sustituyendo en las ecuaciones (8), tenemos que

$$\begin{aligned} x'(t) &= y(t) \\ y'(t) &= -x(t) \end{aligned} \tag{9}$$

con las cuales es más fácil trabajar. Las ecuaciones (9) tienen la forma $\mathbf{x}' = A\mathbf{x}$, donde $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Nuestras nuevas condiciones iniciales son

$$x(0) = r(0) - 10 = 6 - 10 = -4 \quad \text{y} \quad y(0) = w(0) - 12 = 20 - 12 = 8$$

de modo que $\mathbf{x}(0) = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$.

Si procedemos como en el último ejemplo, encontraremos los eigenvalores y eigenvectores de A. El polinomio característico es $\lambda^2 + 1$, el cual no tiene raíces reales. ¿Qué debemos hacer? No tenemos otra elección más que utilizar las raíces complejas, que son $\lambda_1 = i$ y $\lambda_2 = -i$. Los eigenvectores correspondientes también son complejos, a saber, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$. Con base en el teorema 4.40, nuestra solución tiene la forma

$$\mathbf{x}(t) = C_1e^{it}\mathbf{v}_1 + C_2e^{-it}\mathbf{v}_2 = C_1e^{it}\begin{bmatrix} 1 \\ i \end{bmatrix} + C_2e^{-it}\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

De $\mathbf{x}(0) = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$, obtenemos

$$C_1\begin{bmatrix} 1 \\ i \end{bmatrix} + C_2\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$$

cuya solución es $C_1 = -2 - 4i$ y $C_2 = -2 + 4i$. Así que la solución del sistema (9) es

$$\mathbf{x}(t) = (-2 - 4i)e^{it}\begin{bmatrix} 1 \\ i \end{bmatrix} + (-2 + 4i)e^{-it}\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

¿Qué haremos con esta solución? Los petirrojos y las lombrices existen en un mundo real, ¡pero nuestra solución involucra números complejos! Apliquemos sin temor alguno la fórmula de Euler

$$e^{it} = \cos t + i \operatorname{sen} t$$

donde el vector $\mathbf{x}_k$ registra el estado del sistema en el "momento" k y A es una matriz cuadrada. Como se señaló, el comportamiento a largo plazo de estos sistemas se relaciona con los eigenvalores y los eigenvectores de la matriz A. El método de potencias explota la naturaleza iterativa de esos sistemas dinámicos para aproximarse a los eigenvalores y eigenvectores, y el teorema de Perron-Frobenius proporciona información específica sobre el comportamiento a largo plazo de un sistema dinámico lineal discreto cuyo coeficiente matricial A es no negativo.

Cuando A es una matriz de 2×2, podemos describir la evolución de un sistema dinámico en forma geométrica. La ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ es una colección infinita de ecuaciones. Empezando con un vector inicial $\mathbf{x}_0$, tenemos:

$$\begin{aligned}\mathbf{x}_1 &= A\mathbf{x}_0\\ \mathbf{x}_2 &= A\mathbf{x}_1\\ \mathbf{x}_3 &= A\mathbf{x}_2\\ &\vdots\end{aligned}$$

El conjunto $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ recibe el nombre de ***trayectoria*** del sistema. (Para propósitos de graficación, identificaremos cada vector en una trayectoria con su cabeza de manera que podamos graficarlo como un punto.) Observe que $\mathbf{x}_k = A^k\mathbf{x}_0$.

Ejemplo 4.48

Sea $A = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.8 \end{bmatrix}$. Para el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, grafique los primeros cinco puntos en las trayectorias con los siguientes vectores iniciales:

(a) $\mathbf{x}_0 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ (b) $\mathbf{x}_0 = \begin{bmatrix} 0 \\ -5 \end{bmatrix}$ (c) $\mathbf{x}_0 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ (d) $\mathbf{x}_0 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$

Solución (a) Calculamos $\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 2.5 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 1.25 \\ 0 \end{bmatrix}$, $\mathbf{x}_3 = A\mathbf{x}_2 = \begin{bmatrix} 0.625 \\ 0 \end{bmatrix}$, $\mathbf{x}_4 = A\mathbf{x}_3 = \begin{bmatrix} 0.3125 \\ 0 \end{bmatrix}$. Éstos se grafican en la figura 4.22, y los puntos se conectan para resaltar la trayectoria. Cálculos similares producen las trayectorias marcadas como (*b*), (*c*) y (*d*) en la figura 4.22.

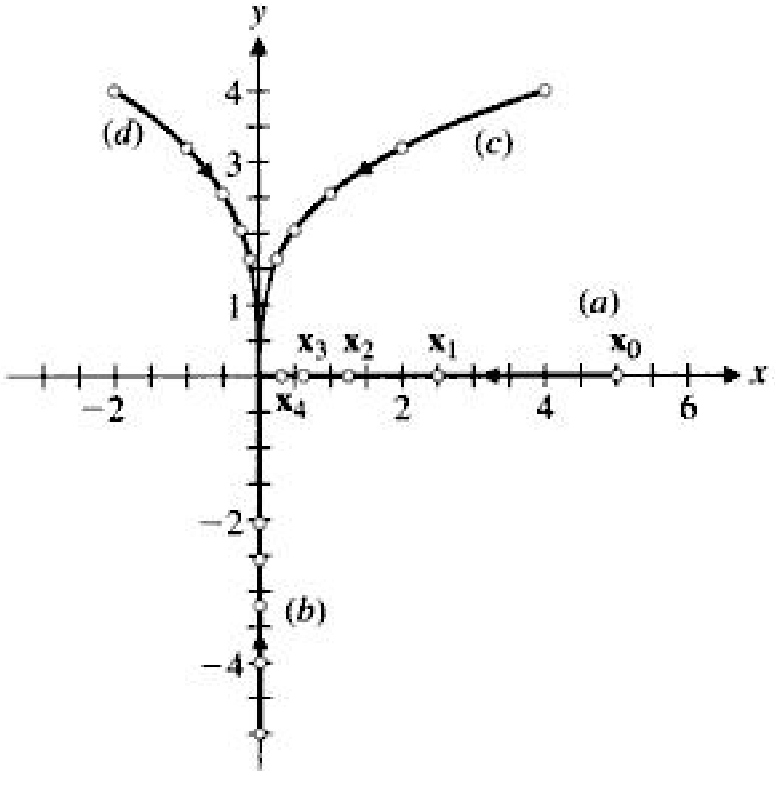


Figura 4.22

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Teorema de Perron-Frobenius

En los ejercicios 28 a 31, encuentre la raíz de Perron y el eigenvector de Perron correspondiente de A.

28. $A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$ **29.** $A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$

30. $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ **31.** $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

Puede demostrarse que una matriz no negativa de $n \times n$ es irreducible si y sólo si $(I + A)^{n-1} > O$. En los ejercicios 32 a 35, aplique este criterio para determinar si la matriz A es irreducible. Si lo es, establezca una permutación de sus renglones y columnas que transforme esa matriz en la forma de bloques

$$\left[\begin{array}{c|c} B & C \\ \hline O & D \end{array}\right]$$

32. $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ **33.** $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

34. $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$ **35.** $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

36. **(a)** Si A es la matriz de adyacencia de una gráfica G, verifique que A es irreducible si y sólo si G está conectada. (Una gráfica está ***conectada*** si existe una trayectoria entre cada par de vértices.)
(b) ¿Cuál de las gráficas de la sección 4.0 tiene una matriz de adyacencia irreducible? ¿Cuál tiene una matriz de adyacencia primitiva?

37. Sea G una gráfica bipartita con matriz de adyacencia A.
(a) Compruebe que A no es primitiva.
(b) Demuestre que si λ es un eigenvalor de A, también lo es $-\lambda$. [*Sugerencia:* emplee el ejercicio 60 de la sección 3.7 y particione un eigenvector de λ de modo que sea compatible con esta partición de A. Utilice esta partición para determinar un eigenvector para $-\lambda$.]

38. Una gráfica se denomina ***k-regular*** si k lados se encuentran en cada vértice. Sea G una gráfica k-regular.
(a) Justifique que la matriz de adyacencia A de G tiene $\lambda = k$ como un eigenvalor. (*Sugerencia:* adapte el teorema 4.30.)
(b) Demuestre que si A es primitiva, entonces los otros eigenvalores son menores que k en valor absoluto. (*Sugerencia:* aplique el teorema 4.31.)

39. Explique los resultados de su análisis de la sección 4.0 con base en los ejercicios 36 al 38 y la sección 4.5.

40. Sean A, B, C y D matrices de $n \times n$, $\mathbf{x}$ en $\mathbb{R}^n$, y c un escalar. Compruebe las siguientes desigualdades matriciales:

(a) $|cA| = |c||A|$ **(b)** $|A + B| \le |A| + |B|$
(c) $|A\mathbf{x}| \le |A||\mathbf{x}|$ **(d)** $|AB| \le |A||B|$
(e) Si $A \ge B \ge O$ y $C \ge D \ge O$, entonces $AC \ge BD \ge O$.

Relaciones de recurrencia lineal

En los ejercicios 41 a 44, escriba los primeros seis términos de la sucesión definida por la relación de recurrencia con las condiciones iniciales dadas.

41. $x_0 = 1, x_n = 2x_{n-1}$ para $n \ge 1$

42. $a_1 = 128, a_n = a_{n-1}/2$ para $n \ge 2$

43. $y_0 = 0, y_1 = 1, y_n = y_{n-1} - y_{n-2}$ para $n \ge 2$

44. $b_0 = 1, b_1 = 1, b_n = 2b_{n-1} + b_{n-2}$ para $n \ge 2$

En los ejercicios 45 a 50, resuelva la relación de recurrencia con las condiciones iniciales dadas.

45. $x_0 = 0, x_1 = 5, x_n = 3x_{n-1} + 4x_{n-2}$ para $n \ge 2$

46. $x_0 = 0, x_1 = 1, x_n = 4x_{n-1} - 3x_{n-2}$ para $n \ge 2$

47. $y_1 = 1, y_2 = 6, y_n = 4y_{n-1} - 4y_{n-2}$ para $n \ge 3$

48. $a_0 = 4, a_1 = 1, a_n = a_{n-1} - a_{n-2}/4$ para $n \ge 2$

49. $b_0 = 0, b_1 = 1, b_n = 2b_{n-1} + 2b_{n-2}$ para $n \ge 2$

50. La relación de recurrencia del ejercicio 43. Demuestre que su solución concuerda con la respuesta al ejercicio 43.

51. Complete la demostración del teorema 4.38(a) y verifique que si la relación de recurrencia $x_n = ax_{n-1} + bx_{n-2}$ tiene eigenvalores distintos $\lambda_1 \ne \lambda_2$, entonces la solución será de la forma

$$x_n = c_1\lambda_1^n + c_2\lambda_2^n$$

(*Sugerencia:* muestre que el método del ejemplo 4.40 es útil en general.)

52. Compruebe que para cualquier selección de condiciones iniciales $x_0 = r$ y $x_1 = s$, pueden localizarse los escalares c_1 y c_2, como se estableció en los incisos (a) y (b) del teorema 4.38.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Figura 5.14
Cónicas no degeneradas

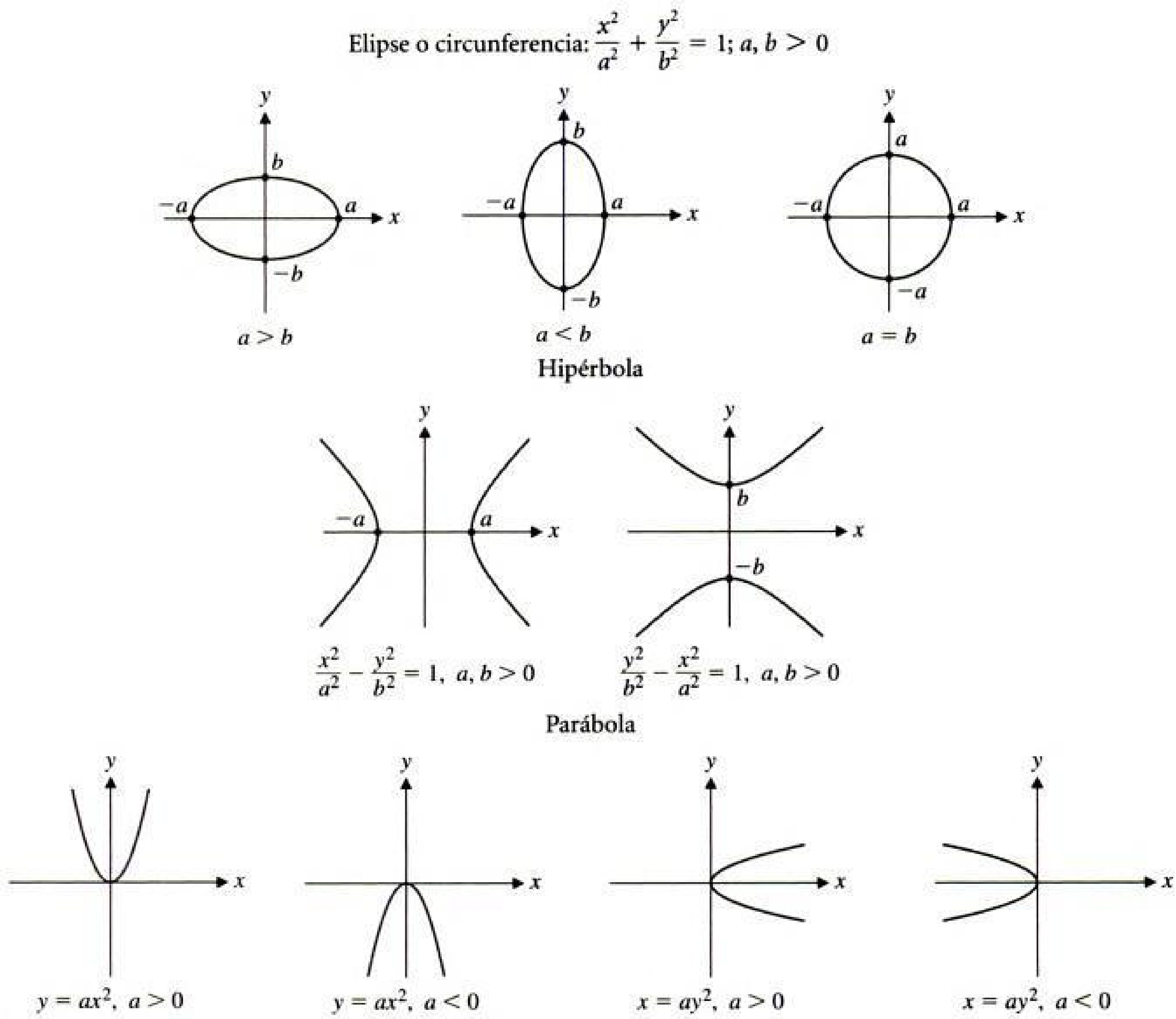


Figura 5.15
Cónicas no degeneradas en posición estándar

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

cuadrática (cuádrica) necesitamos situarla en la posición estándar. Algunas cuadráticas en posición estándar se muestran en la figura 5.19; otras son obtenidas mediante la permutación de las variables.

Ejemplo 5.34 Identifique la superficie cuadrática cuya ecuación es

$$5x^2 + 11y^2 + 2z^2 + 16xy + 20xz - 4yz = 36$$

Solución La ecuación puede ser expresada en forma matricial como $\mathbf{x}^TA\mathbf{x} = 36$, donde

$$A = \begin{bmatrix} 5 & 8 & 10 \\ 8 & 11 & -2 \\ 10 & -2 & 2 \end{bmatrix}$$

Encontramos que los eigenvalores de A son 18, 9 y -9, con eigenvectores ortogonales correspondientes

$$\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$$

respectivamente. Los normalizamos para obtener

$$\mathbf{q}_1 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \quad \text{y} \quad \mathbf{q}_3 = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

y formamos la matriz ortogonal

$$Q = [\mathbf{q}_1 \ \ \mathbf{q}_2 \ \ \mathbf{q}_3] = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$$

Advierta que a fin de que Q sea la matriz de una rotación, requerimos que $\det Q = 1$, lo que es verdadero en este caso. (De otra manera, $\det Q = -1$, e intercambiando dos columnas se cambia el signo del determinante.) Por consiguiente,

$$Q^TAQ = D = \begin{bmatrix} 18 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -9 \end{bmatrix}$$

y, con el cambio de variable $\mathbf{x} = Q\mathbf{x}'$, obtenemos $\mathbf{x}^TA\mathbf{x} = (\mathbf{x}')D\mathbf{x}' = 36$, así que

$$18(x')^2 + 9(y')^2 - 9(z')^2 = 36 \quad \text{o} \quad \frac{(x')^2}{2} + \frac{(y')^2}{4} - \frac{(z')^2}{4} = 1$$

Según la figura 5.19, reconocemos esta ecuación como la ecuación de un hiperboloide de una hoja. Los ejes x', y' y z' están en las direcciones de los eigenvectores $\mathbf{q}_1$, $\mathbf{q}_2$ y $\mathbf{q}_3$, respectivamente. La gráfica se muestra en la figura 5.20.

6 Espacios vectoriales

El álgebra es generosa; a menudo nos da más de lo que le pedimos.
—Jean le Rond d'Alembert
(1717-1783)
citado en la obra de Carl B. Boyer
A History of Mathematics
Wiley, 1968, p. 481

6.0 Introducción: Fibonacci en el espacio (vectorial)

La sucesión Fibonacci se introdujo en la sección 4.6; es la sucesión

$$0, 1, 1, 2, 3, 5, 8, 13, \ldots$$

de enteros no negativos con la propiedad de que después de los dos primeros términos, cada término es la suma de los dos términos anteriores. Por tanto $0 + 1 = 1$, $1 + 1 = 2$, $1 + 2 = 3$, $2 + 3 = 5$, y así sucesivamente...

Si denotamos los términos de la sucesión Fibonacci mediante $f_0, f_1, f_2, \ldots$, entonces toda la sucesión se determina totalmente al especificar que

$$f_0 = 0, f_1 = 1 \quad \text{y} \quad f_n = f_{n-1} + f_{n-2} \quad \text{para } n \geq 2$$

Por analogía con la notación vectorial, escribamos una sucesión $x_0, x_1, x_2, x_3, \ldots$

$$\mathbf{x} = [x_0, x_1, x_2, x_3, \ldots)$$

Entonces la secuencia Fibonacci se vuelve

$$\mathbf{f} = [f_0, f_1, f_2, f_3, \ldots) = [0, 1, 1, 2, \ldots)$$

Ahora generalizamos esta noción.

Definición Una **secuencia del tipo Fibonacci** es *cualquier* sucesión $\mathbf{x} = [x_0, x_1, x_2, x_3, \ldots)$ donde x_0 y x_1 son números reales y $x_n = x_{n-1} + x_{n-2}$ para $n \geq 2$.

Por ejemplo, $[1, \sqrt{2}, 1 + \sqrt{2}, 1 + 2\sqrt{2}, 2 + 3\sqrt{2}, \ldots)$ es una secuencia del tipo Fibonacci.

Problema 1 Escriba los primeros cinco términos de tres secuencias Fibonacci. De nuevo, por analogía con los vectores, definamos la *suma* de dos sucesiones $\mathbf{x} = [x_0, x_1, x_2, \ldots)$ y $\mathbf{y} = [y_0, y_1, y_2, \ldots)$ como la sucesión

$$\mathbf{x} + \mathbf{y} = [x_0 + y_0, x_1 + y_1, x_2 + y_2, \ldots)$$

Si c es un escalar, podemos definir el múltiplo escalar de una sucesión mediante

$$c\mathbf{x} = [cx_0, cx_1, cx_2, \ldots)$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

En general, para cualquier $n \geq 0$, fija, el conjunto $\mathcal{P}_n$ de todos los polinomios con grado menor o igual a n es un espacio vectorial, como lo es el conjunto $\mathcal{P}$ de *todos* los polinomios.

Ejemplo 6.4

Denotemos como $\mathcal{F}$ el conjunto de todas las funciones con valores reales definidas sobre la recta de los números reales. Si f y g son dos funciones de este tipo y c es un escalar, entonces $f + g$ y cf están definidos mediante

$$(f + g)(x) = f(x) + g(x) \quad \text{y} \quad (cf)(x) = cf(x)$$

En otras palabras, el *valor* de $f + g$ en x es obtenido al sumar los valores de f y g en x [(figura 61(a)]. De manera semejante, el valor de cf en x es precisamente el valor de f en x multiplicado por el escalar c [figura 6.1(b)]. El vector nulo de $\mathcal{F}$ es la función constante f_0 que es idénticamente cero; es decir, $f_0(x) = 0$ para toda x. El negativo de una función f es la función $-f$ definida mediante $(-f)(x) = -f(x)$ [figura 6.1(c)].

Los axiomas 1 y 6 son obviamente verdaderos. La verificación de los axiomas restantes se deja como ejercicio 13. De esta manera, $\mathcal{F}$ es un espacio vectorial.

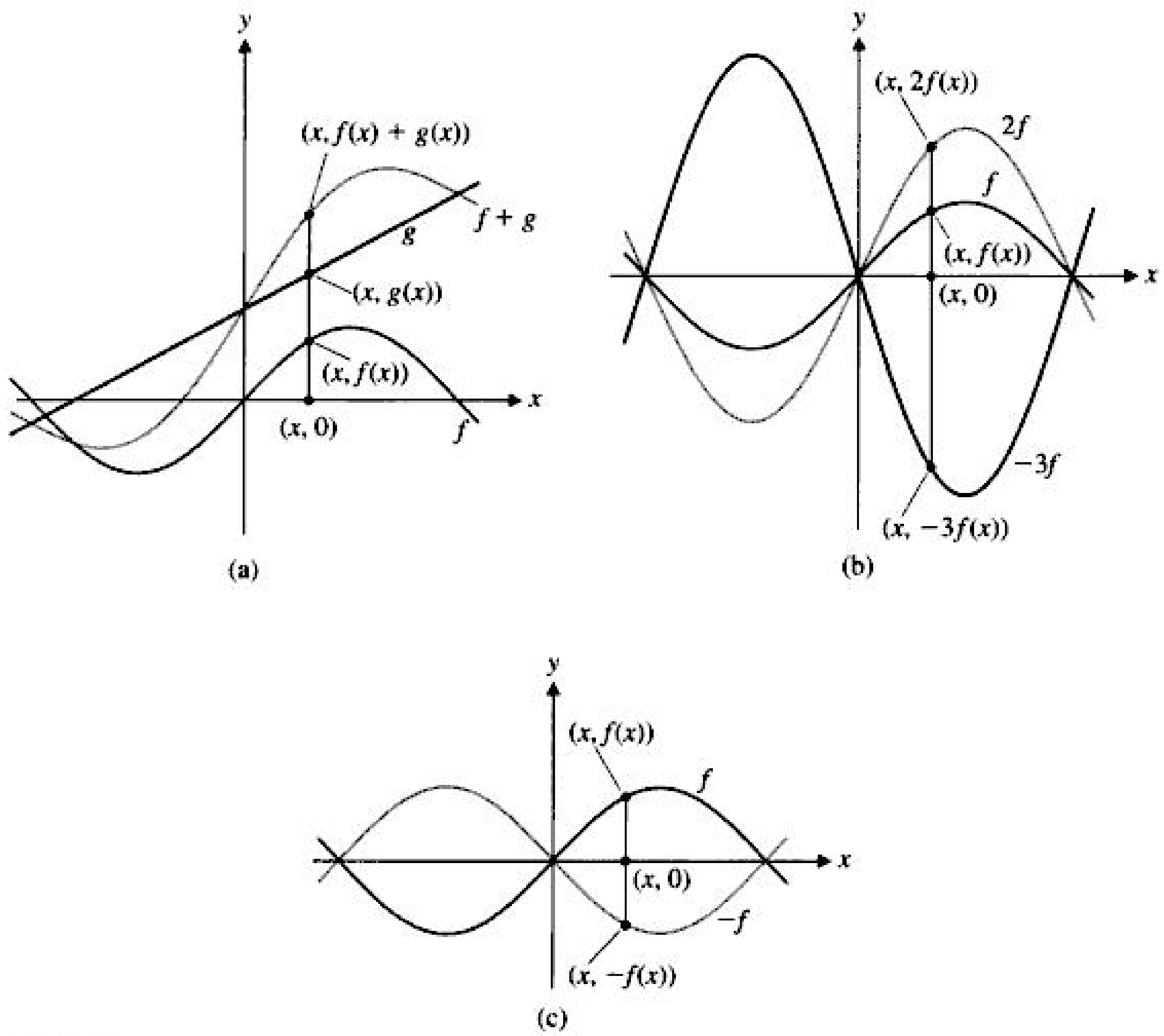


Figura 6.1

Las gráficas de (a) f, g y $f + g$, (b) f, $2f$ y $-3f$, y (c) f y $-f$

Un examen de la demostración del teorema 6.2 revela un útil hecho:

Si W es un subespacio de un espacio vectorial V, entonces W contiene el vector nulo $\mathbf{0}$ de V.

Esto es congruente con el hecho, y análogo a él, de que las rectas y planos son subespacios de $\mathbb{R}^3$ si y sólo si contienen al origen. En ocasiones, el requisito de que todo subespacio debe contener a $\mathbf{0}$ es útil cuando se trata de demostrar que un conjunto *no es* un subespacio.

Ejemplo 6.15 Sea W el conjunto de todas las matrices de 2×2 de la forma

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

¿Es W un subespacio de M_{22}?

Solución Cada matriz en W tiene la propiedad de que su entrada (1, 2) es una más que su entrada (1, 1). En razón de que la matriz cero

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

no tiene esta propiedad, no se encuentra en W. Por lo tanto, W no es un subespacio de M_{22}.

Ejemplo 6.16 Sea W el conjunto de todas las matrices de 2×2 con determinante cero. ¿Es W un subespacio de M_{22}? (Debido a que $\det O = 0$, la matriz cero se encuentra en W, de modo que el método del ejemplo 6.15 no es útil para nosotros.)

Solución Sea

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

Entonces $\det A = \det B = 0$, de manera que A y B se encuentran en W. Sin embargo,

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

de modo que $\det(A + B) = 1 \neq 0$, y por lo tanto $A + B$ no se encuentra en W. De esta forma, W no es cerrado bajo la adición y por ello no es un subespacio de M_{22}.

Conjuntos generadores

La noción de un conjunto generador de vectores se lleva fácilmente desde $\mathbb{R}^n$ a los espacios vectoriales generales.

Definición Si $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto de vectores en un espacio vectorial V, entonces el conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ se conoce como ***espacio generado por*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y se denota mediante la expresión generado$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ o generado(S). Si $V =$ generado(S), entonces S se denomina ***conjunto generador*** para V y se dice que V es ***generado*** por S.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

50. Sea W un subespacio de un espacio vectorial V. Demuestre que $\Delta = \{(\mathbf{w}, \mathbf{w}) : \mathbf{w} \text{ está en } W\}$ es un subespacio de $V \times V$.

En los ejercicios 51 y 52, sea $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ *y* $B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$. *Determine si C se encuentra en generado(A, B).*

51. $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ **52.** $C = \begin{bmatrix} 3 & -5 \\ 5 & -1 \end{bmatrix}$

En los ejercicios 53 y 54, sea $p(x) = 1 - 2x$, $q(x) = x - x^2$, *y* $r(x) = -2 + 3x + x^2$. *Determine si* $s(x)$ *se encuentra en generado*$(p(x), q(x), r(x))$.

53. $s(x) = 3 - 5x - x^2$ **54.** $s(x) = 1 + x + x^2$

En los ejercicios 55 a 58, sean $f(x) = \operatorname{sen}^2 x$ *y* $g(x) = \cos^2 x$. *Determine si* $h(x)$ *se encuentra en generado*$(f(x), g(x))$.

55. $h(x) = 1$ **56.** $h(x) = \cos 2x$

57. $h(x) = \operatorname{sen} 2x$ **58.** $h(x) = \operatorname{sen} x$

59. ¿Es M_{22} generado por $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$?

60. ¿Es M_{22} generado por $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$?

61. ¿Es $\mathscr{P}_2$ generado por $1 + x, x + x^2, 1 + x^2$?

62. ¿Es $\mathscr{P}_2$ generado por $1 + x + 2x^2, 2 + x + 2x^2, -1 + x + 2x^2$?

63. Demuestre que todo espacio vectorial tiene un vector cero único.

64. Demuestre que para todo vector $\mathbf{v}$ en un espacio vectorial V, existe un $\mathbf{v}'$ único en V tal que $\mathbf{v} + \mathbf{v}' = \mathbf{0}$.

6.2 Independencia lineal, base y dimensión

En esta sección extendemos las nociones de independencia lineal, base y dimensión a todos los espacios vectoriales, generalizando los resultados de las secciones 2.3 y 3.5. En la mayoría de los casos, las demostraciones de los teoremas se conservan: simplemente reemplazamos a $\mathbb{R}^n$ con el espacio vectorial V.

Independencia lineal

Definición Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ de un espacio vectorial V es ***linealmente dependiente*** si existen escalares $c_1, c_2, \ldots, c_k$, *al menos uno de los cuales no es 0*, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

Un conjunto de vectores que no es linealmente dependiente se dice que es ***linealmente independiente***.

Como en $\mathbb{R}^n$, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es linealmente independiente en un espacio vectorial V si y sólo si

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0} \quad \text{implica que} \quad c_1 = 0, c_2 = 0, \ldots, c_k = 0$$

También tenemos la siguiente útil formulación alternativa de la dependencia lineal.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Ejemplo 6.33 Demuestre que $\mathcal{B} = \{1, x, x^2, \ldots\}$ es una base para $\mathscr{P}$.

Solución En el ejemplo 6.28, vimos que $\mathcal{B}$ es linealmente independiente. También genera a $\mathscr{P}$, debido a que evidentemente todo polinomio es una combinación lineal de (un número finito de) potencias de x.

Ejemplo 6.34 Encuentre bases para los tres espacios vectoriales del ejemplo 6.13:

(a) $W_1 = \left\{\begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix}\right\}$ (b) $W_2 = \{a + bx - bx^2 + ax^3\}$ (c) $W_3 = \left\{\begin{bmatrix} a & b \\ -b & a \end{bmatrix}\right\}$

Solución Una vez más, trabajaremos los tres ejemplos lado a lado para resaltar las semejanzas entre ellos. En un sentido más estricto, todos ellos son el *mismo* ejemplo, pero esperaremos hasta la sección 6.5 para hacer de esta idea algo perfectamente preciso.

(a) En razón de que

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

tenemos que W_1 = generado($\mathbf{u}$, $\mathbf{v}$), donde

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

Debido a que $\{\mathbf{u}, \mathbf{v}\}$ es, de manera evidente, linealmente independiente, también es una base para W_1.

(b) Puesto que

$$a + bx - bx^2 + ax^3 = a(1 + x^3) + b(x - x^2)$$

tenemos que W_2 = generado($u(x)$, $v(x)$), donde

$$u(x) = 1 + x^3$$

y

$$v(x) = x - x^2$$

Ya que $\{u(x), v(x)\}$ es, de manera evidente, linealmente independiente, también es una base para W_2.

(c) Debido a que

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

tenemos que W_3 = generado(U, V), donde

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad V = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Puesto que $\{U, V\}$ es, de manera evidente, linealmente independiente, también es una base para W_3.

Coordenadas

La sección 3.5 presentó la idea de las coordenadas de un vector con respecto a una base para subespacios de $\mathbb{R}^n$. Ahora extendemos este concepto a espacios vectoriales arbitrarios.

Teorema 6.5 Sea V un espacio vectorial y $\mathcal{B}$ una base para V. Para todo vector $\mathbf{v}$ en V, existe exactamente una manera de expresar $\mathbf{v}$ como una combinación lineal de los vectores básicos en $\mathcal{B}$.

Demostración La demostración es la misma que se aplica al teorema 3.29. Funciona incluso si la base $\mathcal{B}$ es infinita, debido a que las combinaciones lineales son, por definición, finitas.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Teorema 6.9 **El teorema de la base**

Si un espacio vectorial V tiene una base con n vectores, entonces toda base para V tiene exactamente n vectores.

La demostración del teorema 3.23 también funciona aquí, prácticamente palabra por palabra. Sin embargo, es más fácil utilizar el teorema 6.8.

Demostración Sea $\mathcal{B}$ una base para V con n vectores y sea $\mathcal{B}'$ otra base para V con m vectores. Según el teorema 6.8, $m \leq n$; de otro modo, $\mathcal{B}'$ sería linealmente dependiente.

Ahora utilizamos el teorema 6.8 e intercambiaremos los papeles de $\mathcal{B}$ y $\mathcal{B}'$. En razón de que $\mathcal{B}'$ es una base de V con m vectores, el teorema 6.8 implica que cualquier conjunto con más de m vectores en V es linealmente dependiente. Por lo tanto, $n \leq m$, puesto que $\mathcal{B}$ es una base y es, por consiguiente, linealmente independiente.

Debido a que $n \leq m$ y $m \leq n$, debemos tener que $n = m$, como se requería.

En este momento, la definición siguiente tiene sentido, en razón de que el número de vectores en una base (finita) no depende de la selección de la base.

Definición Se dice que un espacio vectorial V es ***de dimensión finita*** si tiene una base determinada por un número finito de vectores. La ***dimensión*** de V, denotada mediante la expresión dim V, es el número de vectores en una base para V. La dimensión del espacio vectorial $\{\mathbf{0}\}$, se define como cero. Un espacio vectorial que no tiene base finita se conoce como de ***dimensión infinita***.

Ejemplo 6.38

Debido a que la base estándar para $\mathbb{R}^n$ tiene n vectores, $\dim \mathbb{R}^n = n$. En el caso de $\mathbb{R}^3$, un subespacio unidimensional es precisamente el espacio generado por un solo vector distinto de cero, es decir, una recta que pasa a través del origen. Un espacio bidimensional es generado por una base compuesta por dos vectores linealmente independientes (es decir, no paralelos) y por lo tanto es un plano que pasa a través del origen. Tres vectores cualquiera linealmente independientes deben generar a $\mathbb{R}^3$, según el teorema fundamental. Los subespacios de $\mathbb{R}^3$ están ahora completamente clasificados según su dimensión, como se muestra en la tabla 6.1.

Tabla 6.1

dim V	V
3	$\mathbb{R}^3$
2	Plano a través del $\mathbf{0}$
1	Recta a través del $\mathbf{0}$
0	$\{\mathbf{0}\}$

Ejemplo 6.39

La base estándar para $\mathcal{P}_n$ contiene $n + 1$ vectores (véase el ejemplo 9), de manera que $\dim \mathcal{P}_n = n + 1$.

Exploración

Cuadrados mágicos

El grabado de la página 465 es la "*Melancolía I*" de Albrecht Dürer o Alberto Durero (1514). Entre los muchos artefactos matemáticos incluidos en él, se aprecia la tabla de números que cuelga sobre la pared, bajo la campana, en la esquina superior derecha. (Véase el detalle a la derecha.) Un arreglo de números de esa clase se conoce como *cuadrado mágico*. Podemos imaginárnoslo como una matriz de 4×4

$$\begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{bmatrix}$$

Observe que la suma de los números en cada renglón, en cada columna y en ambas diagonales es la misma: 34. Advierta, además, que las entradas son los enteros 1, 2, . . . , 16. (Nótese que Durero colocó muy inteligentemente los números 15 y 14 de manera adyacente en el último renglón, lo cual le permitió proporcionar la fecha del grabado.) Estas observaciones nos conducen a la definición siguiente.

Definición Una matriz M de $n \times n$ se denomina ***cuadrado mágico*** si la suma de las entradas en cada renglón, en cada columna y en ambas diagonales es la misma. Esta suma común se conoce como ***peso*** de M, denotado como wt(M). Si M es un cuadrado mágico de $n \times n$ que contiene cada una de las entradas $1, 2, \ldots, n^2$ exactamente una vez, entonces M se conoce como un cuadrado mágico ***clásico***.

1. Si M es un cuadrado mágico clásico de $n \times n$, demuestre que

$$\text{wt}(M) = \frac{n(n^2 + 1)}{2}$$

(*Sugerencia:* utilice el ejercicio 45 de la sección 2.4.)

2. Elabore un cuadrado mágico clásico de 3×3. A continuación, haga uno diferente. ¿Están relacionados sus dos ejemplos de alguna manera?

(Consulte de nuevo la observación que sigue al ejemplo 6.26.) De aquí se sigue que

$$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Para determinar $P_{\mathcal{C}\leftarrow\mathcal{B}}$, podríamos expresar cada vector en $\mathcal{B}$ como una combinación lineal de los vectores en $\mathcal{C}$ (haga esto), pero es mucho más fácil utilizar el hecho de que $P_{\mathcal{C}\leftarrow\mathcal{B}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}$, según el teorema 6.12(c). Encontramos que

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

De aquí se sigue ahora que

$$\begin{aligned} [p(x)]_{\mathcal{C}} &= P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \end{aligned}$$

lo que concuerda con el ejemplo 6.37.

Observación Si no necesitamos $P_{\mathcal{C}\leftarrow\mathcal{B}}$ de manera explícita, podemos hallar $[p(x)]_{\mathcal{C}}$ a partir de $[p(x)]_{\mathcal{B}}$ y $P_{\mathcal{B}\leftarrow\mathcal{C}}$ mediante el empleo de la eliminación gaussiana. La reducción por renglones produce

$$[P_{\mathcal{B}\leftarrow\mathcal{C}} \mid [p(x)]_{\mathcal{B}}] \longrightarrow [I \mid (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}[p(x)]_{\mathcal{B}}] = [I \mid P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}}] = [I \mid [p(x)]_{\mathcal{C}}]$$

(Véase la sección siguiente acerca del uso de la eliminación de Gauss-Jordan.)

Vale la pena repetir la observación del ejemplo 6.46: El cambio *hacia* una base estándar es fácil. Si $\mathcal{E}$ es la base estándar para un espacio vectorial V y $\mathcal{B}$ es cualquier otra base, entonces las columnas de $P_{\mathcal{E}\leftarrow\mathcal{B}}$ son los vectores coordenados de $\mathcal{B}$ con respecto a $\mathcal{E}$, y éstos son, por lo regular, "visibles". Apliquemos esta observación nuevamente en el ejemplo siguiente.

Ejemplo 6.47 En M_{22}, sea $\mathcal{B}$ la base $\{E_{11}, E_{21}, E_{12}, E_{22}\}$ y sea $\mathcal{C}$ la base $\{A, B, C, D\}$, donde

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Encuentre la matriz de cambio de base $P_{\mathcal{C}\leftarrow\mathcal{B}}$ y verifique que $[X]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}}$ para $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Ejercicios 6.3

En los ejercicios 1 a 4:
(a) Encuentre los vectores de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ *y* $[\mathbf{x}]_{\mathcal{C}}$ *de* $\mathbf{x}$ *con respecto a las bases* $\mathcal{B}$ *y* $\mathcal{C}$*, respectivamente.*
(b) Encuentre la matriz de cambio de base $P_{\mathcal{C}\leftarrow\mathcal{B}}$ *de* $\mathcal{B}$ *a* $\mathcal{C}$*.*
(c) Utilice su respuesta al inciso (b) para calcular $[\mathbf{x}]_{\mathcal{C}}$*, y compare su respuesta con la hallada en el inciso(a).*
(d) Encuentre la matriz de cambio de base $P_{\mathcal{B}\leftarrow\mathcal{C}}$ *de* $\mathcal{C}$ *a* $\mathcal{B}$*.*
(e) Utilice sus respuestas a los incisos (c) y (d) para calcular $[\mathbf{x}]_{\mathcal{B}}$*, y compare su respuesta con la que encontró en el inciso (a).*

1. $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ en $\mathbb{R}^2$

2. $\mathbf{x} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}, \mathcal{C} = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ en $\mathbb{R}^2$

3. $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ en $\mathbb{R}^3$

4. $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ en $\mathbb{R}^3$

En los ejercicios 5 a 8, siga las instrucciones para los ejercicios 1 a 4, pero emplee $p(x)$ *en lugar de* $\mathbf{x}$*.*

5. $p(x) = 2 - x, \mathcal{B} = \{1, x\}, \mathcal{C} = \{x, 1 + x\}$ en $\mathscr{P}_1$

6. $p(x) = 1 + 3x, \mathcal{B} = \{1 + x, 1 - x\}, \mathcal{C} = \{2x, 4\}$ en $\mathscr{P}_1$

7. $p(x) = 1 + x^2, \mathcal{B} = \{1 + x + x^2, x + x^2, x^2\},$
$\mathcal{C} = \{1, x, x^2\}$ en $\mathscr{P}_2$

8. $p(x) = 4 - 2x - x^2, \mathcal{B} = \{x, 1 + x^2, x + x^2\},$
$\mathcal{C} = \{1, 1 + x, x^2\}$ en $\mathscr{P}_2$

En los ejercicios 9 y 10, siga las instrucciones para los ejercicios 1 a 4, pero utilice A en lugar de $\mathbf{x}$*.*

9. $A = \begin{bmatrix} 4 & 2 \\ 0 & -1 \end{bmatrix}$, $\mathcal{B}$ = la base estándar,
$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ en M_{22}

10. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$
$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ en M_{22}

En los ejercicios 11 y 12, siga las instrucciones para los ejercicios 1 a 4, pero utilice $f(x)$ *en lugar de* $\mathbf{x}$*.*

11. $f(x) = 2 \operatorname{sen} x - 3 \cos x, \mathcal{B} = \{\operatorname{sen} x + \cos x, \cos x\},$
$\mathcal{C} = \{\operatorname{sen} x, \cos x\}$ en generado$(\operatorname{sen} x, \cos x)$

12. $f(x) = \operatorname{sen} x, \mathcal{B} = \{\operatorname{sen} x + \cos x, \cos x\}, \mathcal{C} = \{\cos x - \operatorname{sen} x, \operatorname{sen} x + \cos x\}$ en generado$(\operatorname{sen} x, \cos x)$

13. Haga rotar los ejes xy en el plano en dirección contraria a las manecillas del reloj, de manera que barran un ángulo $\theta = 60°$ para obtener los nuevos ejes $x'y'$. Utilice los métodos de esta sección para encontrar: (a) las coordenadas $x'y'$ del punto cuyas coordenadas xy son $(3, 2)$ y (b) las coordenadas xy del punto cuyas coordenadas $x'y'$ son $(4, -4)$.

14. Repita el ejercicio 13 con $\theta = 135°$.

15. Sean $\mathcal{B}$ y $\mathcal{C}$ bases para $\mathbb{R}^2$. Si $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ y la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

encuentre $\mathcal{B}$.

16. Sean $\mathcal{B}$ y $\mathcal{C}$ bases para $\mathscr{P}_2$. Si $\mathcal{B} = \{x, 1 + x, 1 - x + x^2\}$ y la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

encuentre $\mathcal{C}$.

En cálculo, aprendimos que un ***polinomio de Taylor de grado n alrededor de a*** *es un polinomio de la forma*

$$p(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_n(x - a)^n$$

donde $a_n \neq 0$*. En otras palabras, es un polinomio que ha sido expandido en términos de potencias de* $x - a$ *en vez de potencias de* x*. Los polinomios de Taylor son muy útiles para aproximar funciones que se "comportan bien" en la vecindad de* $x = a$*.*

encontramos que $c_1 = -7$ y $c_2 = 3$. Por consiguiente,

$$\begin{aligned} T\begin{bmatrix} -1 \\ 2 \end{bmatrix} &= T\left(-7\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \\ &= -7T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3T\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= -7(2 - 3x + x^2) + 3(1 - x^2) \\ &= -11 + 21x - 10x^2 \end{aligned}$$

De manera similar, descubrimos que

$$\begin{bmatrix} a \\ b \end{bmatrix} = (3a - 2b)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

de manera que

$$\begin{aligned} T\begin{bmatrix} a \\ b \end{bmatrix} &= T\left((3a - 2b)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \\ &= (3a - 2b)T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)T\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= (3a - 2b)(2 - 3x + x^2) + (b - a)(1 - x^2) \\ &= (5a - 3b) + (-9a + 6b)x + (4a - 3b)x^2 \end{aligned}$$

(Note que cuando $a = -1$ y $b = 2$, recobramos la solución $T\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -11 + 21x - 10x^2$.)

La demostración del teorema general es bastante directa.

Teorema 6.15 Sea $T: V \rightarrow W$ una transformación lineal y sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto generador para V. Entonces $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ genera la imagen de T.

Demostración La imagen de T es el conjunto de todos los vectores en W que son de la forma $T(\mathbf{v})$, donde $\mathbf{v}$ se encuentra en V. Sea $T(\mathbf{v})$ en la imagen de T. Debido a que $\mathcal{B}$ genera V, existen escalares $c_1, \ldots, c_n$ tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Si aplicamos T y utilizamos el hecho de que es una transformación lineal, observamos que

$$T(\mathbf{v}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n)$$

En otras palabras, $T(\mathbf{v})$ se encuentra en generado $(T(\mathcal{B}))$, como se requiere.

El teorema 6.15 se aplica, en particular, cuando $\mathcal{B}$ es una base de V. Usted puede conjeturar que, en este caso, $T(\mathcal{B})$ sería entonces una base para la imagen de T. Desafortunadamente, éste no es siempre el caso. Abordaremos esta cuestión en la sección 6.5.

Composición de transformaciones lineales

En la sección 3.6 definimos la composición de transformaciones matriciales. De manera obvia, la definición se extiende a las transformaciones lineales en general.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Ejercicios 6.4

En los ejercicios del 1 a 12, determine si T es una transformación lineal.

1. $T: M_{22} \to M_{22}$ definida por
$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ 0 & c+d \end{bmatrix}$$

2. $T: M_{22} \to M_{22}$ definida por
$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & a-d \\ b-c & 1 \end{bmatrix}$$

3. $T: M_{nn} \to M_{nn}$ definida por $T(A) = AB$, donde B es una matriz fija de $n \times n$

4. $T: M_{nn} \to M_{nn}$ definida por $T(A) = AB - BA$, donde B es una matriz fija de $n \times n$

5. $T: M_{nn} \to \mathbb{R}$ definida por $T(A) = \operatorname{tr}(A)$

6. $T: M_{nn} \to \mathbb{R}$ definida por $T(A) = a_{11}a_{22}\cdots a_{nn}$

7. $T: M_{nn} \to \mathbb{R}$ definida por $T(A) = \text{rango}(A)$

8. $T: \mathscr{P}_2 \to \mathscr{P}_2$ definida por $T(a + bx + cx^2) = (a+1) + (b+1)x + (c+1)x^2$

9. $T: \mathscr{P}_2 \to \mathscr{P}_2$ definida por $T(a + bx + cx^2) = a + b(x+1) + b(x+1)^2$

10. $T: \mathscr{F} \to \mathscr{F}$ definida por $T(f) = f(x^2)$

11. $T: \mathscr{F} \to \mathscr{F}$ definida por $T(f) = (f(x))^2$

12. $T: \mathscr{F} \to \mathbb{R}$ definida por $T(f) = f(c)$, donde c es un escalar fijo

13. Demuestre que las transformaciones S y T del ejemplo 6.56 son lineales.

14. Sea $T: \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal para la cual
$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \text{y} \quad T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$$
Encuentre $T\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ y $T\begin{bmatrix} a \\ b \end{bmatrix}$.

15. Sea $T: \mathbb{R}^2 \to \mathscr{P}_2$ una transformación lineal para la cual
$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 - 2x \quad \text{y} \quad T\begin{bmatrix} 3 \\ -1 \end{bmatrix} = x + 2x^2$$
Encuentre $T\begin{bmatrix} -7 \\ 9 \end{bmatrix}$ y $T\begin{bmatrix} a \\ b \end{bmatrix}$.

16. Sea $T: \mathscr{P}_2 \to \mathscr{P}_2$ una transformación lineal para la cual
$$T(1) = 3 - 2x, \quad T(x) = 4x - x^2, \quad T(x^2) = 2 + 2x^2$$
Encuentre $T(6 + x - 4x^2)$ y $T(a + bx + cx^2)$.

17. Sea $T: \mathscr{P}_2 \to \mathscr{P}_2$ una transformación lineal para la cual
$$T(1 + x) = 1 + x^2, \quad T(x + x^2) = x - x^2,$$
$$T(1 + x^2) = 1 + x + x^2$$
Encuentre $T(4 - x + 3x^2)$ y $T(a + bx + cx^2)$.

18. Sea $T: M_{22} \to \mathbb{R}$ una transformación lineal para la cual
$$T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 1, \quad T\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 2,$$
$$T\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = 3, \quad T\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 4$$
Encuentre $T\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$ y $T\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

19. Sea $T: M_{22} \to \mathbb{R}$ una transformación lineal. Demuestre que existen escalares a, b, c y d tales que
$$T\begin{bmatrix} w & x \\ y & z \end{bmatrix} = aw + bx + cy + dz$$
para toda $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ en M_{22}.

20. Demuestre que no existe una transformación lineal $T: \mathbb{R}^3 \to \mathscr{P}_2$ tal que
$$T\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 1 + x, \quad T\begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} = 2 - x + x^2,$$
$$T\begin{bmatrix} 0 \\ 6 \\ -8 \end{bmatrix} = -2 + 2x^2$$

21. Demuestre el teorema 6.14(b).

22. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V y sea $T: V \to V$ una transformación lineal. Demuestre que si $T(\mathbf{v}_1) = \mathbf{v}_1$, $T(\mathbf{v}_2) = \mathbf{v}_2 \ldots$, $T(\mathbf{v}_n) = \mathbf{v}_n$, entonces T es la transformación identidad sobre V.

23. Sea $T: \mathscr{P}_n \to \mathscr{P}_n$ una transformación lineal tal que $T(x^k) = kx^{k-1}$ para $k = 0, 1, \ldots, n$. Demuestre que T debe ser el operador diferencial D.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Demostración Supongamos que T es invertible. Entonces, existe una transformación lineal $T^{-1}: W \to V$ tal que

$$T^{-1} \circ T = I_V \quad \text{y} \quad T \circ T^{-1} = I_W$$

Para demostrar que T es inyectiva, sea $\mathbf{v}$ en el núcleo de T. Entonces, $T(\mathbf{v}) = \mathbf{0}$. Por lo tanto,

$$\begin{aligned} T^{-1}(T(\mathbf{v})) = T^{-1}(\mathbf{0}) &\Rightarrow (T^{-1} \circ T)(\mathbf{v}) = \mathbf{0} \\ &\Rightarrow I(\mathbf{v}) = \mathbf{0} \\ &\Rightarrow \mathbf{v} = \mathbf{0} \end{aligned}$$

lo que establece que $\ker(T) = \{\mathbf{0}\}$. Por consiguiente, T es inyectiva, según el teorema 6.20.

Para demostrar que T es suprayectiva, sea $\mathbf{w}$ en W y sea $\mathbf{v} = T^{-1}(\mathbf{w})$. Entonces,

$$\begin{aligned} T(\mathbf{v}) &= T(T^{-1}(\mathbf{w})) \\ &= (T \circ T^{-1})(\mathbf{w}) \\ &= I(\mathbf{w}) \\ &= \mathbf{w} \end{aligned}$$

lo que demuestra que $\mathbf{w}$ es la imagen de $\mathbf{v}$ bajo T. Debido a que $\mathbf{v}$ se encuentra en V, esto comprueba que T es suprayectiva.

Recíprocamente, supongamos que T es inyectiva y suprayectiva. Eso significa que nulidad$(T) = 0$ y rango$(T) = \dim W$. Necesitamos demostrar que existe una transformación lineal $T': W \to V$ tal que $T' \circ T = I_V$ y $T \circ T' = I_W$.

Sea $\mathbf{w}$ en W. Como T es suprayectiva, existe algún vector $\mathbf{v}$ en V tal que $T(\mathbf{v}) = \mathbf{w}$. Existe solamente un vector $\mathbf{v}$ así, ya que, si $\mathbf{v}'$ es otro vector en V tal que $T(\mathbf{v}') = \mathbf{w}$, entonces $T(\mathbf{v}) = T(\mathbf{v}')$; el hecho de que T sea inyectiva implicaría que $\mathbf{v} = \mathbf{v}'$. Por lo tanto, tiene sentido definir una función $T': W \to$ estableciendo que $T'(\mathbf{w}) = \mathbf{v}$.

De aquí se sigue que

$$(T' \circ T)(\mathbf{v}) = T'(T(\mathbf{v})) = T'(\mathbf{w}) = \mathbf{v}$$

y

$$(T \circ T')(\mathbf{w}) = T(T'(\mathbf{w})) = T(\mathbf{v}) = \mathbf{w}$$

Entonces se sigue que $T' \circ T = I_V$ y $T \circ T' = I_W$. Ahora debemos demostrar que T' es una transformación *lineal*.

Para este fin, sean $\mathbf{w}_1$ y $\mathbf{w}_2$ en W, y c_1 y c_2 escalares. Como antes, sea $T(\mathbf{v}_1) = \mathbf{w}_1$ y $T(\mathbf{v}_2) = \mathbf{w}_2$. Entonces, $\mathbf{v}_1 = T'(\mathbf{w}_1)$ y $\mathbf{v}_2 = T'(\mathbf{w}_2)$ y

$$\begin{aligned} T'(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) &= T'(c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)) \\ &= T'(T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)) \\ &= I(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) \\ &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \\ &= c_1T'(\mathbf{w}_1) + c_2T'(\mathbf{w}_2) \end{aligned}$$

En consecuencia, T' es lineal, de modo que, según el teorema 6.17, $T' = T^{-1}$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

6.6 Matriz de una transformación lineal

Mediante el teorema 6.15 se demostró que una transformación lineal $T: V \to W$ está completamente determinada por su efecto sobre un conjunto generador para V. En particular, si sabemos cómo actúa T sobre una base para V, entonces podemos calcular $T(\mathbf{v})$ para cualquier vector $\mathbf{v}$ de V. El ejemplo 6.55 ilustra el proceso. Implícitamente recurrimos a esta importante propiedad de las transformaciones lineales en el teorema 3.31 para ayudarnos a calcular la matriz estándar de una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$. En esta sección, mostraremos que toda transformación lineal entre espacios vetoriales de dimensión finita puede ser representada como una transformación matricial.

Supongamos que V es un espacio vectorial de dimensión n, W es un espacio vectorial de dimensión m y $T: V \to W$ es una transformación lineal. Sean $\mathcal{B}$ y $\mathcal{C}$ bases para V y W, respectivamente. Entonces, la función que lleva el vector de coordenadas $R(\mathbf{v}) = [\mathbf{v}]_\mathcal{B}$ define un isomorfismo $R: V \to \mathbb{R}^n$. Al mismo tiempo, tenemos un isomorfismo $S: W \to \mathbb{R}^m$ dado por $S(\mathbf{w}) = [\mathbf{w}]_\mathcal{C}$, que nos permite asociar la imagen $T(\mathbf{v})$ con el vector $[T(\mathbf{v})]_\mathcal{C}$ en $\mathbb{R}^m$. La figura 6.11 ilustra estas relaciones.

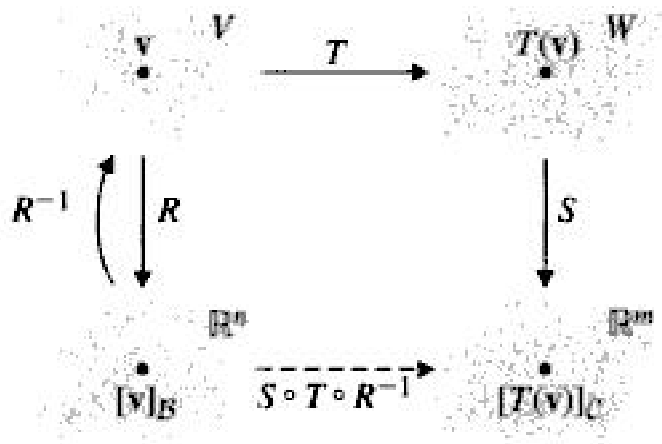


Figura 6.11

En razón de que R es un isomorfismo, es invertible, de manera que podemos formar la función compuesta

$$S \circ T \circ R^{-1}: \mathbb{R}^n \to \mathbb{R}^m$$

que lleva $[\mathbf{v}]_\mathcal{B}$ en $[T(\mathbf{v})]_\mathcal{C}$. Debido a que esta función va de $\mathbb{R}^n$ a $\mathbb{R}^m$, sabemos, por lo que se explicó en el capítulo 3, que es una transformación matricial. Entonces, ¿cuál es la matriz estándar de $S \circ T \circ R^{-1}$? Nos gustaría encontrar la matriz A de $m \times n$ tal que $A[\mathbf{v}]_\mathcal{B} = (S \circ T \circ R^{-1})([\mathbf{v}]_\mathcal{B})$. O, ya que $(S \circ T \circ R^{-1})([\mathbf{v}]_\mathcal{B}) = [T(\mathbf{v})]_\mathcal{C}$, requerimos que

$$A[\mathbf{v}]_\mathcal{B} = [T(\mathbf{v})]_\mathcal{C}$$

Esto resulta ser sorprendentemente fácil de encontrar. La idea básica es la del teorema 3.31. Las columnas de A son las imágenes de los vectores en la base estándar para $\mathbb{R}^n$ bajo $S \circ T \circ R^{-1}$. Pero, si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base para V, entonces

$$\begin{aligned} R(\mathbf{v}_i) &= [\mathbf{v}_i]_\mathcal{B} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \leftarrow i\text{-ésima entrada} \\ &= \mathbf{e}_i \end{aligned}$$

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

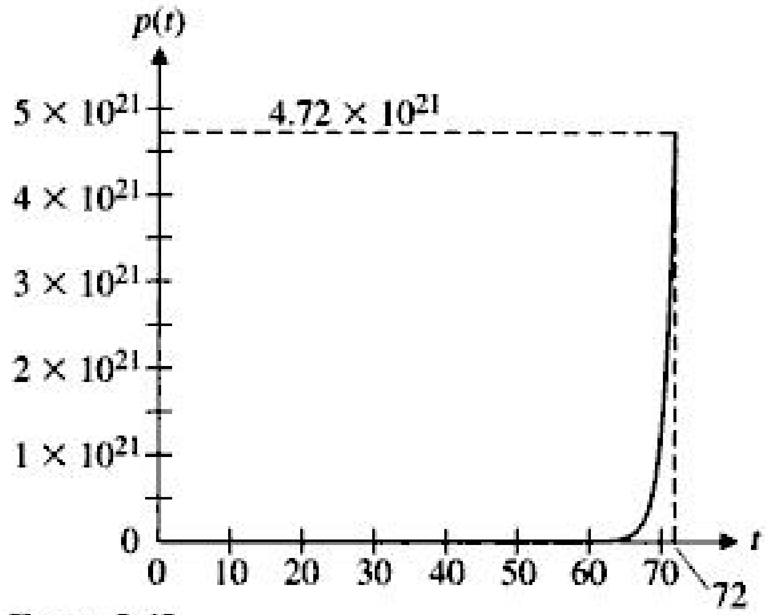


Figura 6.18
Crecimiento exponencial

Las sustancias radiactivas decaen mediante la emisión de radiación. Si $m(t)$ denota la masa de la sustancia en el tiempo t, entonces la tasa de decaimiento es $m'(t)$. Los físicos han determinado que la tasa del decaimiento de una sustancia es proporcional a su masa; es decir,

$$m'(t) = km(t) \quad \text{o} \quad m' - km = 0$$

donde k es una constante negativa. Si Aplicamos el teorema 6.32, tenemos que

$$m(t) = ce^{kt}$$

para alguna constante c. El tiempo requerido para que una sustancia radiactiva decaiga a la mitad se conoce como su ***vida media***.

Ejemplo 6.88

Después de 5.5 días, una muestra de 100 mg de radón 222 decae a 37 mg.

(a) Encuentre la fórmula de $m(t)$, la masa restante después de t días.
(b) ¿Cuál es la vida media del radón 222?
(c) ¿Cuándo quedarán solamente 10 mg?

Solución (a) De $m(t) = ce^{kt}$, tenemos que

$$100 = m(0) = ce^{k\cdot 0} = c \cdot 1 = c$$

de manera que $$m(t) = 100e^{kt}$$

Con el tiempo medido en días, se nos informa que $m(5.5) = 37$. Por consiguiente,

$$100e^{5.5k} = 37$$

de manera que $$e^{5.5k} = 0.37$$

Resolviendo para k, determinamos que

$$5.5k = \ln(0.37)$$

de modo que $$k = \frac{\ln(0.37)}{5.5} \approx -0.18$$

Por lo tanto, $m(t) = 100e^{-0.18t}$.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

El álgebra lineal es una asignatura apasionante. El objetivo principal de este libro es transmitir esta pasión y facilitar la comprensión de los conceptos y resultados interesantes, relacionándolos con abundantes aplicaciones reales de diversas disciplinas y conexiones con otras áreas de las matemáticas.

Este texto fue escrito teniendo en mente la flexibilidad. La gama de temas y aplicaciones que abarca lo hacen adecuado para una gran diversidad de lectores y tipos de curso.

El estudio guiado con este libro ayudará al estudiante a dominar todas las facetas del álgebra lineal y a observar la interacción entre ellas.

ISBN-13: 978-970-686-595-3
ISBN-10: 970-686-595-0

9 789706 865953

CENGAGE Learning

http://latinoamerica.cengage.com