4 Espacios vectoriales

EJEMPLO INTRODUCTORIO

Vuelo espacial y sistemas de control

Con sus 12 pisos de altura y un peso de 75 toneladas, el *Columbia* se elevó majestuosamente de la plataforma de lanzamiento en una fresca mañana de Domingo de Ramos en abril de 1981. Producto de 10 años de una intensa labor de investigación y desarrollo, el primer transbordador espacial de Estados Unidos fue un triunfo de diseño de la ingeniería de sistemas de control, en el que participaron varias ramas de la ingeniería: aeronáutica, química, eléctrica, hidráulica y mecánica.

Los sistemas de control de la nave espacial son absolutamente esenciales para el vuelo. Como el transbordador tiene un fuselaje inestable, requiere de una monitorización computarizada constante durante el vuelo atmosférico. El sistema de control de vuelo envía una secuencia de comandos a las superficies de control aerodinámico y a 44 pequeños impulsores de propulsión a chorro. En la figura 1 se muestra un típico sistema de circuito cerrado de retroalimentación que controla el cabeceo del transbordador durante el vuelo. (El cabeceo es el ángulo de elevación del cono de proa). Los símbolos de unión (⊗) indican dónde se añaden las señales de varios sensores a las señales de la computadora que fluyen a través de la parte superior de la figura.

Matemáticamente, las señales de entrada y salida a un sistema de ingeniería son funciones. En las aplicaciones, es importante que estas funciones se puedan sumar, como se muestra en la figura 1, y multiplicarse por escalares. Estas dos operaciones con las funciones tienen propiedades algebraicas que son completamente análogas a las operaciones de suma de vectores en $\mathbb{R}^n$ y a la multiplicación de un vector por un escalar, como se verá en las secciones 4.1 y 4.8. Por esta razón, el conjunto de todas las entradas posibles (funciones) se denomina *espacio vectorial*. Los fundamentos matemáticos para la ingeniería de sistemas se basan en espacios

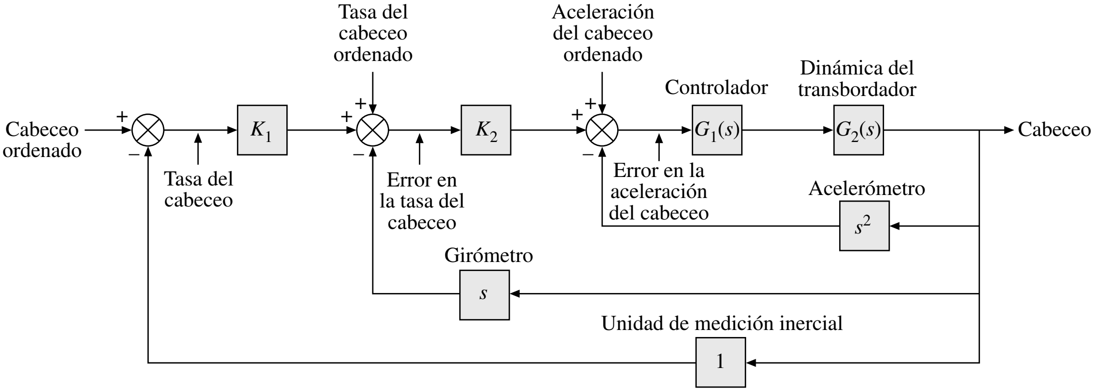


FIGURA 1 Sistema de control del cabeceo para el transbordador espacial. (*Fuente:* Adaptado de *Space Shuttle GN&C Operations Manual*, Rockwell International, © 1988).

vectoriales de funciones, y en este capítulo 4 se amplía la teoría de los vectores en $\mathbb{R}^n$ para incluir dichas funciones. Más adelante, veremos cómo surgen otros espacios vectoriales en ingeniería, física y estadística.

WEB

Las "semillas" matemáticas sembradas en los capítulos 1 y 2 germinan y comienzan a florecer en este capítulo. La belleza y el poder del álgebra lineal se apreciarán más claramente cuando el lector considere a $\mathbb{R}^n$ solo como uno de tantos espacios vectoriales que surgen de forma natural en problemas aplicados. En realidad, el estudio de espacios vectoriales no es muy diferente del propio estudio de $\mathbb{R}^n$, porque usted podrá utilizar su experiencia geométrica con $\mathbb{R}^2$ y $\mathbb{R}^3$ para visualizar muchos conceptos generales.

Comenzando con las definiciones básicas en la sección 4.1, el marco teórico de los espacios vectoriales se desarrolla gradualmente a lo largo del capítulo. Uno de los objetivos de las secciones 4.3 a 4.5 es demostrar cuán estrechamente se parecen otros espacios vectoriales a $\mathbb{R}^n$. En la sección 4.6 se estudiará el rango, uno de los temas más importantes de este capítulo, utilizando la terminología de espacio vectorial para vincular datos importantes de las matrices rectangulares. En la sección 4.8 se aplica la teoría del capítulo a las señales discretas y ecuaciones en diferencias utilizadas en los sistemas de control digital, como en el transbordador espacial. Las cadenas de Markov, en la sección 4.9, representan un cambio de ritmo respecto de las secciones más teóricas del capítulo y dan buenos ejemplos para los conceptos que se introducirán en el capítulo 5.

4.1 ESPACIOS Y SUBESPACIOS VECTORIALES

Gran parte de la teoría en los capítulos 1 y 2 se basa en ciertas propiedades algebraicas sencillas y evidentes de $\mathbb{R}^n$, que se listan en la sección 1.3. De hecho, muchos otros sistemas matemáticos tienen las mismas propiedades. Las propiedades específicas de interés se incluyen en la siguiente definición.

DEFINICIÓN

Un **espacio vectorial** es un conjunto no vacío V de objetos, llamados *vectores*, en el que están definidas dos operaciones, llamadas *suma* y *multiplicación por escalares* (números reales), sujetas a los 10 axiomas (o reglas) que se listan a continuación.[1] Los axiomas deben ser válidos para todos los vectores $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ en V, y para todos los escalares c y d.

1. La suma de $\mathbf{u}$ y $\mathbf{v}$, denotada con $\mathbf{u} + \mathbf{v}$, está en V.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. Hay un vector **cero** ($\mathbf{0}$) en V tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. Para cada $\mathbf{u}$ en V, existe un vector $-\mathbf{u}$ en V tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. El múltiplo escalar de $\mathbf{u}$ por c, que se denota con $c\mathbf{u}$, está en V.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

[1] Técnicamente, V es un *espacio vectorial real*. Toda la teoría en este capítulo también es válida para un *espacio vectorial complejo* en el que los escalares son números complejos. Trataremos brevemente este asunto en el capítulo 5. Hasta entonces, se supone que todos los escalares son reales.

Utilizando tan solo estos axiomas, es posible demostrar que el vector cero en el axioma 4 es único, y el vector $-\mathbf{u}$, llamado el **negativo** de $\mathbf{u}$, en el axioma 5 es único para toda $\mathbf{u}$ en V. Véase los ejercicios 25 y 26. Demostraciones de los hechos que se presentan a continuación se describen en los ejercicios:

Para cada $\mathbf{u}$ en V y escalar c,

$$0\mathbf{u} = \mathbf{0} \tag{1}$$

$$c\mathbf{0} = \mathbf{0} \tag{2}$$

$$-\mathbf{u} = (-1)\mathbf{u} \tag{3}$$

EJEMPLO 1 Los espacios $\mathbb{R}^n$, donde $n \geq 1$, son los ejemplos principales de espacios vectoriales. La intuición geométrica desarrollada para $\mathbb{R}^3$ le ayudará a entender y visualizar muchos conceptos en todo el capítulo. ■

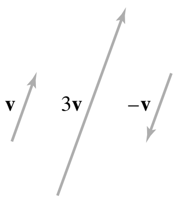


FIGURA 1

EJEMPLO 2 Sea V el conjunto de todas las flechas (segmentos de recta dirigidos) en el espacio tridimensional; se considera que dos flechas son iguales si tienen la misma longitud y la punta en la misma dirección. Defina la suma por la regla del paralelogramo (de la sección 1.3), y para cada $\mathbf{v}$ en V, defina $c\mathbf{v}$ como la flecha cuya longitud es $|c|$ por la longitud de $\mathbf{v}$, apuntando en la misma dirección que $\mathbf{v}$ si $c \geq 0$ y, en caso contrario, apuntando en la dirección opuesta. (Véase la figura 1). Demuestre que V es un espacio vectorial. Este espacio es un modelo común de los problemas físicos de varias fuerzas.

SOLUCIÓN La definición de V es geométrica, utilizando los conceptos de longitud y dirección. Ningún sistema de coordenadas xyz está implicado. Una flecha de longitud cero es un punto único y representa el vector cero. El negativo de $\mathbf{v}$ es $(-1)\mathbf{v}$. Por lo tanto, los axiomas 1, 4, 5, 6 y 10 son evidentes. El resto se comprobará con geometría. Por ejemplo, véase las figuras 2 y 3. ■

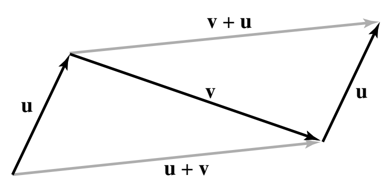


FIGURA 2 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

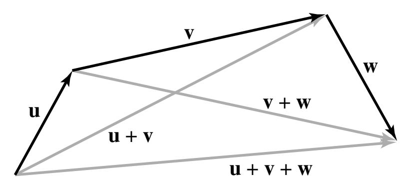


FIGURA 3 $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

EJEMPLO 3 Sea $\mathbb{S}$ el espacio de todas las secuencias doblemente infinitas de números (que generalmente se anotan en una fila más que en una columna):

$$\{y_k\} = (\ldots, y_{-2}, y_{-1}, y_0, y_1, y_2, \ldots)$$

Si $\{z_k\}$ es otro elemento de $\mathbb{S}$, entonces la suma $\{y_k\} + \{z_k\}$ es la sucesión $\{y_k + z_k\}$ formada por la suma de los términos correspondientes de $\{y_k\}$ y $\{z_k\}$. El múltiplo escalar $c\{y_k\}$ es la sucesión $\{cy_k\}$. Los axiomas de espacio vectorial se comprueban en la misma forma que para $\mathbb{R}^n$.

Los elementos de $\mathbb{S}$ surgen en ingeniería, por ejemplo, cuando una señal se mide (o se muestrea) en tiempos discretos. Una señal puede ser eléctrica, mecánica, óptica, etcétera. Los sistemas de control principal del transbordador espacial, que se mencionaron en la introducción del capítulo, usan señales discretas (o digitales). Por conveniencia, llamaremos a $\mathbb{S}$ el espacio de **señales** (discretas de tiempo). Una señal se puede visualizar con una gráfica como la que se ilustra en la figura 4. ■

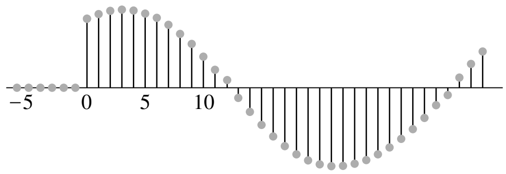


FIGURA 4 Una señal discreta de tiempo.

EJEMPLO 4 Para $n \geq 0$, el conjunto $\mathbb{P}_n$ de polinomios de grado n o menor consiste en todos los polinomios de la forma

$$\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \tag{4}$$

donde los coeficientes $a_0, \ldots, a_n$ y la variable t son números reales. El *grado* de $\mathbf{p}$ es la mayor potencia de t en (4) cuyo coeficiente no es cero. Si $\mathbf{p}(t) = a_0 \neq 0$, el grado de $\mathbf{p}$ es cero. Si todos los coeficientes son iguales a cero, $\mathbf{p}$ se llama el *polinomio cero*. El polinomio cero está incluido en $\mathbb{P}_n$, a pesar de que su grado, por razones técnicas, no esté definido.

Si $\mathbf{p}$ está dada por la ecuación (4) y si $\mathbf{q}(t) = b_0 + b_1 t + \cdots + b_n t^n$, entonces la suma $\mathbf{p} + \mathbf{q}$ se define mediante

$$\begin{aligned}(\mathbf{p} + \mathbf{q})(t) &= \mathbf{p}(t) + \mathbf{q}(t) \\ &= (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n\end{aligned}$$

El múltiplo escalar $c\mathbf{p}$ es el polinomio definido por

$$(c\mathbf{p})(t) = c\mathbf{p}(t) = ca_0 + (ca_1)t + \cdots + (ca_n)t^n$$

Estas definiciones satisfacen los axiomas 1 y 6, ya que $\mathbf{p} + \mathbf{q}$ y $c\mathbf{p}$ son polinomios de grado igual o menor que n. Los axiomas 2, 3, y 7 a 10 son consecuencias de las propiedades de los números reales. Evidentemente, el polinomio cero actúa como vector cero en el axioma 4. Por último, $(-1)\mathbf{p}$ actúa como el negativo de $\mathbf{p}$, por lo que el axioma 5 está satisfecho. Así, $\mathbb{P}_n$ es un espacio vectorial.

Se utilizan espacios vectoriales $\mathbb{P}_n$ para varias n, por ejemplo, en el análisis de tendencia estadística de datos, que se analiza en la sección 6.8. ■

EJEMPLO 5 Sea V el conjunto de todas las funciones de valores reales definidas en un conjunto $\mathbb{D}$. (Por lo general, $\mathbb{D}$ es el conjunto de números reales o algún intervalo en la recta real). Las funciones se suman en la forma habitual: $\mathbf{f} + \mathbf{g}$ es la función cuyo valor en t en el dominio $\mathbb{D}$ es $\mathbf{f}(t) + \mathbf{g}(t)$. Asimismo, para un escalar c y una $\mathbf{f}$ en V, el múltiplo escalar $c\mathbf{f}$ es la función cuyo valor en t es $c\mathbf{f}(t)$. Por ejemplo, si $\mathbb{D} = \mathbb{R}$, $\mathbf{f}(t) = 1 + \operatorname{sen} 2t$, y $\mathbf{g}(t) = 2 + .5t$, entonces

$$(\mathbf{f} + \mathbf{g})(t) = 3 + \operatorname{sen} 2t + .5t \quad \text{y} \quad (2\mathbf{g})(t) = 4 + t$$

Dos funciones en V son iguales si y solo si sus valores son iguales para toda t en $\mathbb{D}$. Por lo tanto, el vector cero en V es la función igual a cero, $\mathbf{f}(t) = 0$ para toda t, y el negativo de $\mathbf{f}$ es $(-1)\mathbf{f}$. Como es evidente, los axiomas 1 y 6 son ciertos, y los otros axiomas se deducen de las propiedades de los números reales, por lo que V es un espacio vectorial. ■

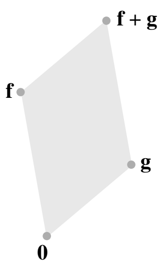


FIGURA 5
La suma de dos vectores (funciones).

Es importante pensar en cada función en el espacio vectorial V del ejemplo 5 como un objeto único, un solo "punto" o un vector en el espacio vectorial. La suma de dos vectores $\mathbf{f}$ y $\mathbf{g}$ (funciones en V, o elementos de *cualquier* espacio vectorial) se puede visualizar como en la figura 5, ya que esto le ayudará a aplicar a un espacio vectorial general la intuición geométrica que ha desarrollado al trabajar con el espacio vectorial $\mathbb{R}^n$. Consulte la *Guía de estudio* como ayuda conforme aprenda a adoptar este punto de vista más general.

Subespacios

En muchos problemas, un espacio vectorial se compone de un subconjunto adecuado de vectores de un espacio vectorial más grande. En este caso, solo se deben comprobar tres de los 10 axiomas de espacio vectorial, y el resto se satisfacen de manera automática.

DEFINICIÓN

Un **subespacio** de un espacio vectorial V es un subconjunto H de V que tiene tres propiedades:

a) El vector cero de V está en H.[2]

b) H es cerrado bajo la suma de vectores. Es decir, por cada $\mathbf{u}$ y $\mathbf{v}$ en H, la suma $\mathbf{u} + \mathbf{v}$ está en H.

c) H es cerrado bajo la multiplicación por escalares. Es decir, para cada $\mathbf{u}$ en H y cada escalar c, el vector $c\mathbf{u}$ está en H.

Las propiedades *a*), *b*) y *c*) garantizan que un subespacio H de V sea en sí mismo un *espacio vectorial* bajo las operaciones de espacio vectorial ya definidas en V. Para comprobar esto, observe que las propiedades *a*), *b*) y *c*) corresponden a los axiomas 1, 4 y 6. Los axiomas 2, 3 y 7 a 10 son automáticamente verdaderos en H, ya que se aplican a todos los elementos de V, incluidos los de H. El axioma 5 también es verdadero en H, ya que si $\mathbf{u}$ está en H, entonces $(-1)\mathbf{u}$ está en H de acuerdo con la propiedad *c*), y sabemos a partir de la ecuación (3) de la página 191 que $(-1)\mathbf{u}$ es el vector $-\mathbf{u}$ en el axioma 5.

Así, cada subespacio es un espacio vectorial. Por el contrario, todo espacio vectorial es un subespacio (de sí mismo y, posiblemente, de otros espacios más grandes). El término *subespacio* se utiliza cuando al menos dos espacios vectoriales están en mente, con uno dentro del otro, y la frase *subespacio de V* identifica a V como el espacio más grande. (Véase la figura 6).

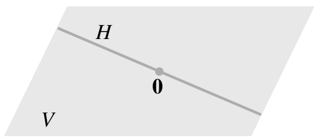


FIGURA 6
Un subespacio de V.

EJEMPLO 6 El conjunto que consta solo del vector cero en un espacio vectorial V es un subespacio de V, llamado **subespacio cero**, y se representa como $\{\mathbf{0}\}$. ■

EJEMPLO 7 Sea $\mathbb{P}$ el conjunto de todos los polinomios con coeficientes reales, con operaciones en $\mathbb{P}$ definidas como en las funciones. De esta forma, $\mathbb{P}$ es un subespacio del espacio de todas las funciones de valores reales definidas en $\mathbb{R}$. Además, para cada $n \geq 0$, $\mathbb{P}_n$ es un subespacio de $\mathbb{P}$, ya que $\mathbb{P}_n$ es un subconjunto de $\mathbb{P}$ que contiene el polinomio cero; la suma de dos polinomios en $\mathbb{P}_n$ también está en $\mathbb{P}_n$, y un múltiplo escalar de un polinomio en $\mathbb{P}_n$ también está en $\mathbb{P}_n$. ■

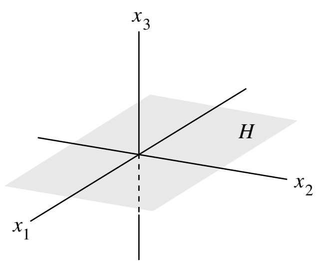


FIGURA 7
El plano x_1x_2 como un subespacio de $\mathbb{R}^3$.

EJEMPLO 8 El espacio vectorial $\mathbb{R}^2$ *no* es un subespacio de $\mathbb{R}^3$ porque $\mathbb{R}^2$ ni siquiera es un subconjunto de $\mathbb{R}^3$. (Todos los vectores en $\mathbb{R}^3$ tienen tres entradas, mientras que los vectores de $\mathbb{R}^2$ tienen solo dos). El conjunto

$$H = \left\{ \begin{bmatrix} s \\ t \\ 0 \end{bmatrix} : s \text{ y } t \text{ son reales} \right\}$$

es un subconjunto de $\mathbb{R}^3$ que se "ve" y "actúa" como $\mathbb{R}^2$, aunque lógicamente es distinto de $\mathbb{R}^2$. Véase la figura 7. Demuestre que H es un subespacio de $\mathbb{R}^3$.

SOLUCIÓN El vector cero está en H, y H es cerrado bajo la suma de vectores y la multiplicación escalar debido a que estas operaciones sobre los vectores de H siempre producen vectores cuyas terceras entradas son iguales a cero (y, por lo tanto, pertenecen a H). Por consiguiente, H es un subespacio de $\mathbb{R}^3$. ■

---

[2] Algunos libros remplazan la propiedad *a*) en esta definición por el supuesto de que H no es vacío. Así, *a*) se puede deducir a partir de *c*) y el hecho de que $0\mathbf{u} = \mathbf{0}$. Pero la mejor manera de someter a prueba un subespacio es buscar primero al vector cero. Si $\mathbf{0}$ está en H, entonces se deben revisar las propiedades *b*) y *c*). Si $\mathbf{0}$ *no* está en H, entonces H no puede ser un subespacio y no se necesita comprobar las otras propiedades.

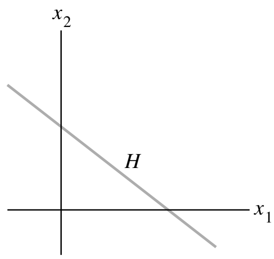


FIGURA 8
Una recta que no es un espacio vectorial.

EJEMPLO 9 Un plano en $\mathbb{R}^3$ que *no* pasa por el origen no es un subespacio de $\mathbb{R}^3$, porque el plano no contiene el vector cero de $\mathbb{R}^3$. Del mismo modo, una recta en $\mathbb{R}^2$ que *no* pasa por el origen, como en la figura 8, *no* es un subespacio de $\mathbb{R}^2$. ■

Un subespacio generado por un conjunto

El siguiente ejemplo ilustra una de las formas más comunes de describir un subespacio. Como en el capítulo 1, el término **combinación lineal** se refiere a cualquier suma de los múltiplos escalares de vectores, y Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ denota el conjunto de todos los vectores que se pueden escribir como combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

EJEMPLO 10 Dados $\mathbf{v}_1$ y $\mathbf{v}_2$ en un espacio vectorial V, sea $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$. Demuestre que H es un subespacio de V.

SOLUCIÓN El vector cero está en H, ya que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. Para demostrar que H es cerrado bajo la suma de vectores, tome dos vectores arbitrarios en H, por ejemplo,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{y} \quad \mathbf{w} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

Por los axiomas 2, 3 y 8 para el espacio vectorial V,

$$\begin{aligned} \mathbf{u} + \mathbf{w} &= (s_1\mathbf{v}_1 + s_2\mathbf{v}_2) + (t_1\mathbf{v}_1 + t_2\mathbf{v}_2) \\ &= (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2 \end{aligned}$$

Así que $\mathbf{u} + \mathbf{w}$ está en H. Además, si c es un escalar cualquiera, entonces por los axiomas 7 y 9,

$$c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$$

lo que demuestra que $c\mathbf{u}$ está en H, y H es cerrado bajo la multiplicación por escalares. Por lo tanto, H es un subespacio de V. ■

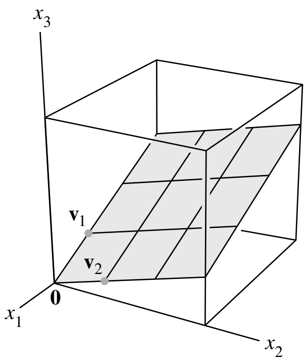


FIGURA 9
Un ejemplo de un subespacio.

En la sección 4.5, veremos que todo subespacio de $\mathbb{R}^3$ distinto de cero, que no sea $\mathbb{R}^3$ mismo, es Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ para algunos $\mathbf{v}_1$ y $\mathbf{v}_2$ linealmente independientes o Gen $\{\mathbf{v}\}$ para $\mathbf{v} \neq \mathbf{0}$. En el primer caso, el subespacio es un plano que pasa por el origen; en el segundo caso, se trata de una recta que pasa por el origen. (Véase la figura 9). Es útil tener en mente estas imágenes geométricas, incluso para un espacio vectorial abstracto.

El argumento en el ejemplo 10 se puede generalizar fácilmente para comprobar el siguiente teorema.

TEOREMA 1

Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en un espacio vectorial V, entonces Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es un subespacio de V.

Se llama Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ al **subespacio generado** por $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Dado un subespacio H de V, un **conjunto generador** para H es un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en H tal que $H = \text{Gen}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

El siguiente ejemplo muestra cómo utilizar el teorema 1.

EJEMPLO 11 Sea H el conjunto de todos los vectores de la forma $(a - 3b, b - a, a, b)$, donde a y b son escalares arbitrarios. Es decir, sea $H = \{(a - 3b, b - a, a, b)\text{: } a \text{ y } b \text{ en } \mathbb{R}\}$. Demuestre que H es un subespacio de $\mathbb{R}^4$.

SOLUCIÓN Represente los vectores en H como vectores columna. Así, un vector arbitrario en H tiene la forma

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}_1} + b\underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}_2}$$

Este cálculo muestra que $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_1$ y $\mathbf{v}_2$ son los vectores indicados anteriormente. Por lo tanto, H es un subespacio de $\mathbb{R}^4$ de acuerdo con el teorema 1. ■

El ejemplo 11 muestra una técnica útil para expresar un subespacio H como el conjunto de combinaciones lineales de una pequeña colección de vectores. Si $H = \text{Gen}\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$, podemos pensar en los vectores $\mathbf{v}_1,\ldots, \mathbf{v}_p$ del conjunto generador como "asas" que nos permiten sujetar el subespacio H. Cálculos con un número infinito de vectores en H con frecuencia se reducen a operaciones con un número finito de vectores en el conjunto generador.

EJEMPLO 12 Determine qué valor(es) de h hará(n) que $\mathbf{y}$ sea un subespacio de $\mathbb{R}^3$ generado por $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, si

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

SOLUCIÓN Esta pregunta es el problema de práctica 2 en la sección 1.3, que aquí se modificó utilizando el término *subespacio* en vez de $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. La solución obtenida ahí muestra que $\mathbf{y}$ está en $\text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si y solo si $h = 5$. Vale la pena revisar esta solución ahora, junto con los ejercicios 11 a 16 y 19 a 21 de la sección 1.3. ■

Aunque muchos espacios vectoriales en este capítulo son subespacios de $\mathbb{R}^n$, es importante considerar que la teoría abstracta se aplica a otros espacios vectoriales. Espacios vectoriales de funciones surgen en muchas aplicaciones, y más adelante se tratarán con más detalle.

PROBLEMAS DE PRÁCTICA

1. Demuestre que el conjunto H de todos los puntos en $\mathbb{R}^2$ de la forma $(3s, 2 + 5s)$ no es un espacio vectorial, al mostrar que no es cerrado bajo la multiplicación escalar. (Encuentre un vector específico $\mathbf{u}$ en H y un escalar c tal que $c\mathbf{u}$ no está en H).
2. Sea $W = \text{Gen}\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$, donde $\mathbf{v}_1,\ldots, \mathbf{v}_p$ se encuentran en un espacio vectorial V. Demuestre que $\mathbf{v}_k$ está en W para $1 \leq k \leq p$. [*Sugerencia:* Escriba primero una ecuación que demuestre que $\mathbf{v}_1$ está en W. Después, ajuste su notación para el caso general].

WEB

4.1 EJERCICIOS

1. Sea V el primer cuadrante en el plano xy; es decir, sea
$$V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \geq 0, y \geq 0 \right\}$$
   a) Si $\mathbf{u}$ y $\mathbf{v}$ están en V, ¿está $\mathbf{u} + \mathbf{v}$ en V? ¿Por qué?
   b) Encuentre un vector específico $\mathbf{u}$ en V y un escalar específico c tal que $c\mathbf{u}$ *no* esté en V. (Esto es suficiente para demostrar que V *no* es un espacio vectorial).

2. Sea W la unión del primer y tercer cuadrantes en el plano xy. Es decir, sea $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : xy \geq 0 \right\}$.
   a) Si $\mathbf{u}$ está en W y c es cualquier escalar, ¿está $c\mathbf{u}$ en W? ¿Por qué?

b) Encuentre dos vectores específicos **u** y **v** tales que $\mathbf{u} + \mathbf{v}$ no esté en W. Esto es suficiente para demostrar que W *no* es un espacio vectorial.

3. Sea H el conjunto de puntos en el interior y sobre el círculo unitario en el plano xy. Es decir, sea $H = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x^2 + y^2 \leq 1 \right\}$. Encuentre un ejemplo específico —dos vectores o un vector y un escalar—, que demuestre que H no es un subespacio de $\mathbb{R}^2$.

4. Construya una figura geométrica que ilustre por qué una recta en $\mathbb{R}^2$ que *no* pasa por el origen no es cerrada bajo la suma de vectores.

En los ejercicios 5 a 8, determine si el conjunto dado es un subespacio de $\mathbb{P}_n$ para un valor adecuado de n. Justifique sus respuestas.

5. Todos los polinomios de la forma $\mathbf{p}(t) = at^2$, donde a se encuentra en $\mathbb{R}$.

6. Todos los polinomios de la forma $\mathbf{p}(t) = a + t^2$, donde a se encuentra en $\mathbb{R}$.

7. Todos los polinomios de grado 3 como máximo, con números enteros como coeficientes.

8. Todos los polinomios en $\mathbb{P}_n$ tales que $\mathbf{p}(0) = 0$.

9. Sea H el conjunto de todos los vectores de la forma $\begin{bmatrix} -2t \\ 5t \\ 3t \end{bmatrix}$.

Encuentre un vector **v** en $\mathbb{R}^3$ tal que $H = \text{Gen}\,\{\mathbf{v}\}$. ¿Por qué esto demuestra que H es un subespacio de $\mathbb{R}^3$?

10. Sea H el conjunto de todos los vectores de la forma $\begin{bmatrix} 3t \\ 0 \\ -7t \end{bmatrix}$,

donde t es cualquier número real. Demuestre que H es un subespacio de $\mathbb{R}^3$. (Utilice el método del ejercicio 9).

11. Sea W el conjunto de todos los vectores de la forma $\begin{bmatrix} 2b + 3c \\ -b \\ 2c \end{bmatrix}$, donde b y c son arbitrarios. Encuentre los vectores **u** y **v** tales que $W = \text{Gen}\,\{\mathbf{u}, \mathbf{v}\}$. ¿Por qué esto demuestra que W es un subespacio de $\mathbb{R}^3$?

12. Sea W el conjunto de todos los vectores de la forma $\begin{bmatrix} 2s + 4t \\ 2s \\ 2s - 3t \\ 5t \end{bmatrix}$. Demuestre que W es un subespacio de $\mathbb{R}^4$. (Utilice el método del ejercicio 11).

13. Sea $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

a) ¿Está **w** en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Cuántos vectores se encuentran en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

b) ¿Cuántos vectores están en $\text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

c) ¿**w** es el subespacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Por qué?

14. Sea $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ como en el ejercicio 13, y sea $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$.

¿Está **w** en el subespacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Por qué?

En los ejercicios 15 a 18, sea W el conjunto de todos los vectores de la forma indicada, donde a, b y c representan números reales arbitrarios. En cada caso, encuentre un conjunto S de vectores que genere W o dé un ejemplo para demostrar que W *no* es un espacio vectorial.

15. $\begin{bmatrix} 2a + 3b \\ -1 \\ 2a - 5b \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 3a - 5b \\ 3b + 2a \end{bmatrix}$

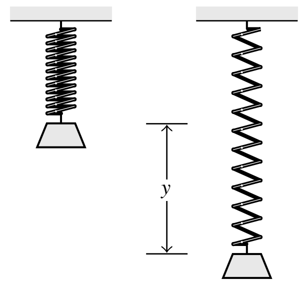
17. $\begin{bmatrix} 2a - b \\ 3b - c \\ 3c - a \\ 3b \end{bmatrix}$

18. $\begin{bmatrix} 4a + 3b \\ 0 \\ a + 3b + c \\ 3b - 2c \end{bmatrix}$

19. Si una masa m se coloca en el extremo de un resorte, y si se tira de la masa hacia abajo y luego se suelta, el sistema masa-resorte comenzará a oscilar. El desplazamiento y de la masa desde su posición de reposo está dado por una función de la forma

$$y(t) = c_1 \cos \omega t + c_2 \operatorname{sen} \omega t \qquad (5)$$

donde ω es una constante que depende del resorte y la masa. (Véase la figura que se muestra a continuación). Demuestre que el conjunto de todas las funciones descritas en la ecuación (5) (con ω fija y c_1 y c_2 arbitrarias) es un espacio vectorial.



20. El conjunto de todas las funciones de valor real continuas y definidas en un intervalo cerrado $[a, b]$ en $\mathbb{R}$ se denota mediante $C[a, b]$. Este conjunto es un subespacio del espacio vectorial de todas las funciones de valor real definidas en $[a, b]$.

a) ¿Qué hechos acerca de las funciones continuas se deberían someter a prueba para demostrar que $C[a, b]$ es en realidad el subespacio que se afirma? (Estos hechos se suelen tratar en una clase de cálculo).

b) Demuestre que $\{\mathbf{f} \text{ en } C[a, b] : \mathbf{f}(a) = \mathbf{f}(b)\}$ es un subespacio de $C[a, b]$.

Para enteros positivos fijos m y n, el conjunto $M_{m\times n}$ de todas las matrices de $m \times n$ es un espacio vectorial, bajo las operaciones habituales de suma de matrices y multiplicación por escalares reales.

21. Determine si el conjunto H de todas las matrices de la forma $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ es un subespacio de $M_{2\times 2}$.

22. Sea F una matriz fija de 3×2, y sea H el conjunto de todas las matrices A en $M_{2\times 4}$ con la propiedad de que $FA = 0$ (la matriz cero en $M_{3\times 4}$). Determine si H es un subespacio de $M_{2\times 4}$.

En los ejercicios 23 y 24, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

23. *a*) Si $\mathbf{f}$ es una función en el espacio vectorial V de todas las funciones de valores reales en $\mathbb{R}$ y si $\mathbf{f}(t) = 0$ para alguna t, entonces $\mathbf{f}$ es el vector cero en V.

b) Un vector es una flecha en el espacio tridimensional.

c) Un subconjunto H de un espacio vectorial V es un subespacio de V si el vector cero está en H.

d) Un subespacio también es un espacio vectorial.

e) Las señales analógicas se utilizan en los sistemas de control principal del transbordador espacial, mencionado en la introducción al capítulo.

24. *a*) Un vector es cualquier elemento de un espacio vectorial.

b) Si $\mathbf{u}$ es un vector en un espacio vectorial V, entonces $(-1)\mathbf{u}$ es igual al negativo de $\mathbf{u}$.

c) Un espacio vectorial es un subespacio.

d) $\mathbb{R}^2$ es un subespacio de $\mathbb{R}^3$.

e) Un subconjunto H de un espacio vectorial V es un subespacio de V si se cumplen las siguientes condiciones: **i.** el vector cero de V está en H, **ii.** $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} + \mathbf{v}$ en H, y **iii.** c es un escalar y $c\mathbf{u}$ está en H.

Los ejercicios 25 a 29 muestran cómo se pueden utilizar los axiomas de un espacio vectorial V para demostrar las propiedades elementales descritas después de la definición de un espacio vectorial. Complete los espacios en blanco con los números del axioma adecuado. De acuerdo con el axioma 2, los axiomas 4 y 5 implican, respectivamente, que $\mathbf{0} + \mathbf{u} = \mathbf{u}$ y $-\mathbf{u} + \mathbf{u} = \mathbf{0}$ para toda $\mathbf{u}$.

25. Complete la siguiente demostración de que el vector cero es único. Suponga que $\mathbf{w}$ en V tiene la propiedad de que $\mathbf{u} + \mathbf{w} = \mathbf{w} + \mathbf{u} = \mathbf{u}$ para toda $\mathbf{u}$ en V. En particular, $\mathbf{0} + \mathbf{w} = \mathbf{0}$. Sin embargo, $\mathbf{0} + \mathbf{w} = \mathbf{w}$, de acuerdo con el axioma ______. Por lo tanto, $\mathbf{w} = \mathbf{0} + \mathbf{w} = \mathbf{0}$.

26. Complete la siguiente demostración de que $-\mathbf{u}$ es el *único vector* en V tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. Suponga que $\mathbf{w}$ satisface $\mathbf{u} + \mathbf{w} = \mathbf{0}$. Sumando $-\mathbf{u}$ en ambos lados, se tiene

$(-\mathbf{u}) + [\mathbf{u} + \mathbf{w}] = (-\mathbf{u}) + \mathbf{0}$

$[(-\mathbf{u}) + \mathbf{u}] + \mathbf{w} = (-\mathbf{u}) + \mathbf{0}$ según el axioma _____ *a*)

$\mathbf{0} + \mathbf{w} = (-\mathbf{u}) + \mathbf{0}$ según el axioma _____ *b*)

$\mathbf{w} = -\mathbf{u}$ según el axioma _____ *c*)

27. Complete los números de axioma que faltan en la siguiente demostración de que $0\mathbf{u} = \mathbf{0}$ para cada $\mathbf{u}$ en V.

$0\mathbf{u} = (0 + 0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$ según el axioma _____ *a*)

Sume el negativo de $0\mathbf{u}$ en ambos lados:

$0\mathbf{u} + (-0\mathbf{u}) = [0\mathbf{u} + 0\mathbf{u}] + (-0\mathbf{u})$

$0\mathbf{u} + (-0\mathbf{u}) = 0\mathbf{u} + [0\mathbf{u} + (-0\mathbf{u})]$ según el axioma _____ *b*)

$\mathbf{0} = 0\mathbf{u} + \mathbf{0}$ según el axioma _____ *c*)

$\mathbf{0} = 0\mathbf{u}$ según el axioma _____ *d*)

28. Complete los números de axioma que faltan en la siguiente demostración de que

$c\mathbf{0} = \mathbf{0}$ para cada escalar c.

$c\mathbf{0} = c(\mathbf{0} + \mathbf{0})$ según el axioma _____ *a*)

$= c\mathbf{0} + c\mathbf{0}$ según el axioma _____ *b*)

Sume el negativo de $c\mathbf{0}$ en ambos lados:

$c\mathbf{0} + (-c\mathbf{0}) = [c\mathbf{0} + c\mathbf{0}] + (-c\mathbf{0})$

$c\mathbf{0} + (-c\mathbf{0}) = c\mathbf{0} + [c\mathbf{0} + (-c\mathbf{0})]$ según el axioma _____ *c*)

$\mathbf{0} = c\mathbf{0} + \mathbf{0}$ según el axioma _____ *d*)

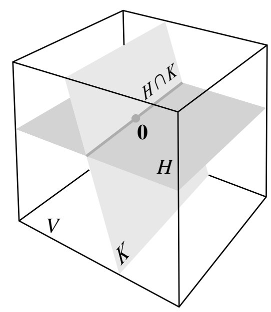
$\mathbf{0} = c\mathbf{0}$ según el axioma _____ *e*)

29. Demuestre que $(-1)\mathbf{u} = -\mathbf{u}$. [*Sugerencia:* Demuestre que $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$. Utilice algunos axiomas y los resultados de los ejercicios 27 y 26].

30. Suponga que $c\mathbf{u} = \mathbf{0}$ para algún escalar c distinto de cero. Demuestre que $\mathbf{u} = \mathbf{0}$. Mencione los axiomas o las propiedades que utilice.

31. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en un espacio vectorial V, y sea H cualquier subespacio de V que contiene a $\mathbf{u}$ y $\mathbf{v}$. Explique por qué H también contiene a Gen $\{\mathbf{u}, \mathbf{v}\}$. Esto demuestra que Gen $\{\mathbf{u}, \mathbf{v}\}$ es el menor subespacio de V que contiene a $\mathbf{u}$ y $\mathbf{v}$.

32. Sean H y K subespacios de un espacio vectorial V. La **intersección** de H y K, que se representa como $H \cap K$, es el conjunto de $\mathbf{v}$ en V que pertenece tanto a H como K. Demuestre que $H \cap K$ es un subespacio de V. (Véase la figura). Dé un ejemplo en $\mathbb{R}^2$ para demostrar que la unión de dos subespacios no es, en general, un subespacio.



33. Dados los subespacios H y K de un espacio vectorial V, la **suma** de H y K, que se escribe como $H + K$, es el conjunto de todos los vectores en V que se pueden representar como la suma de dos vectores, uno en H y el otro en K; es decir,

$$H + K = \{\mathbf{w} : \mathbf{w} = \mathbf{u} + \mathbf{v} \text{ para alguna } \mathbf{u} \text{ en } H \text{ y alguna } \mathbf{v} \text{ en } K\}$$

a) Demuestre que $H + K$ es un subespacio de V.

b) Demuestre que H es un subespacio de $H + K$ y K es un subespacio de $H + K$.

34. Suponga que $\mathbf{u}_1, \ldots, \mathbf{u}_p$ y $\mathbf{v}_1, \ldots, \mathbf{v}_q$ son vectores en un espacio vectorial V, y sea

$H = \text{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ y $K = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$

Demuestre que $H + K = \text{Gen}\{\mathbf{u}_1, \ldots, \mathbf{u}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$.

35. [**M**] Demuestre que $\mathbf{w}$ está en el subespacio de $\mathbb{R}^4$ generado por $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, donde

$$\mathbf{w} = \begin{bmatrix} 9 \\ -4 \\ -4 \\ 7 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 8 \\ -4 \\ -3 \\ 9 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 3 \\ -2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -7 \\ 6 \\ -5 \\ -18 \end{bmatrix}$$

36. [**M**] Determine si $\mathbf{y}$ está en el subespacio de $\mathbb{R}^4$ generado por las columnas de A, donde

$$\mathbf{y} = \begin{bmatrix} -4 \\ -8 \\ 6 \\ -5 \end{bmatrix}, \quad A = \begin{bmatrix} 3 & -5 & -9 \\ 8 & 7 & -6 \\ -5 & -8 & 3 \\ 2 & -2 & -9 \end{bmatrix}$$

37. [**M**] El espacio vectorial $H = \text{Gen}\,\{1, \cos^2 t, \cos^4 t, \cos^6 t\}$ contiene al menos dos funciones interesantes que se utilizarán en un ejercicio posterior:

$$\mathbf{f}(t) = 1 - 8\cos^2 t + 8\cos^4 t$$
$$\mathbf{g}(t) = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$$

Estudie la gráfica de $\mathbf{f}$ para $0 \le t \le 2\pi$, e infiera una fórmula sencilla para $\mathbf{f}(t)$. Verifique su suposición mediante la representación gráfica de la diferencia entre $1 + \mathbf{f}(t)$ y su fórmula para $\mathbf{f}(t)$. (Se espera que vea la función constante 1). Repita el procedimiento para $\mathbf{g}$.

38. [**M**] Repita el ejercicio 37 para las funciones

$\mathbf{f}(t) = 3 \operatorname{sen} t - 4 \operatorname{sen}^3 t$

$\mathbf{g}(t) = 1 - 8 \operatorname{sen}^2 t + 8 \operatorname{sen}^4 t$

$\mathbf{h}(t) = 5 \operatorname{sen} t - 20 \operatorname{sen}^3 t + 16 \operatorname{sen}^5 t$

en el espacio vectorial Gen $\{1, \operatorname{sen} t, \operatorname{sen}^2 t, \ldots, \operatorname{sen}^5 t\}$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Tome cualquier $\mathbf{u}$ en H, digamos, $\mathbf{u} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$, y tome cualquier $c \neq 1$, por ejemplo, $c = 2$. Así, $c\mathbf{u} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$. Si esto se encuentra en H, entonces hay alguna s tal que

$$\begin{bmatrix} 3s \\ 2 + 5s \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

Es decir, $s = 2$ y $s = 12/5$, lo que es imposible. Así que $2\mathbf{u}$ no está en H, y H no es un espacio vectorial.

2. $\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. Esta ecuación expresa a $\mathbf{v}_1$ como una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_p$, por lo que $\mathbf{v}_1$ se encuentra en W. En general, $\mathbf{v}_k$ está en W, ya que

$$\mathbf{v}_k = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{k-1} + 1\mathbf{v}_k + 0\mathbf{v}_{k+1} + \cdots + 0\mathbf{v}_p$$

4.2 ESPACIOS NULOS, ESPACIOS COLUMNA Y TRANSFORMACIONES LINEALES

En las aplicaciones del álgebra lineal, los subespacios de $\mathbb{R}^n$ generalmente se presentan en una de dos maneras: **1.** como el conjunto de todas las soluciones a un sistema de ecuaciones lineales homogéneas, o **2.** como el conjunto de todas las combinaciones lineales de ciertos vectores específicos. En esta sección, se comparan estas dos descripciones de subespacios, lo que nos permitirá practicar el uso del concepto de subespacio. En realidad, como pronto descubrirá, hemos trabajado con subespacios desde la sección 1.3. La principal novedad aquí es la terminología. La sección concluye con un análisis acerca del núcleo y el rango de una transformación lineal.

El espacio nulo de una matriz

Considere el siguiente sistema de ecuaciones homogéneas:

$$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -5x_1 + 9x_2 + x_3 &= 0 \end{aligned} \tag{1}$$

En forma de matriz, este sistema se representa como $A\mathbf{x} = \mathbf{0}$, donde

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

Recordemos que el conjunto de todas las $\mathbf{x}$ que satisfacen (1) se llama el **conjunto solución** del sistema (1). Con frecuencia es conveniente relacionar este conjunto directamente con la matriz A y la ecuación $A\mathbf{x} = \mathbf{0}$. El conjunto de $\mathbf{x}$ que satisfacen $A\mathbf{x} = \mathbf{0}$ se denomina el **espacio nulo** de la matriz A.

DEFINICIÓN

El **espacio nulo** de una matriz A de $m \times n$, que se denota como Nul A, es el conjunto de todas las soluciones de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$. En notación de conjuntos,

$$\text{Nul } A = \{\mathbf{x} : \mathbf{x} \text{ está en } \mathbb{R}^n \text{ y } A\mathbf{x} = \mathbf{0}\}$$

Una descripción más dinámica de Nul A es el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^n$ que se mapean en el vector cero de $\mathbb{R}^m$ a través de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$. Véase la figura 1.

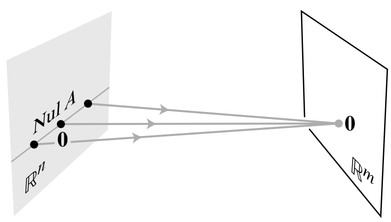


FIGURA 1

EJEMPLO 1 Sea A la matriz en la ecuación (2) anterior, y sea $\mathbf{u} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$. Determine si $\mathbf{u}$ pertenece al espacio nulo de A.

SOLUCIÓN Para probar si $\mathbf{u}$ satisface $A\mathbf{u} = \mathbf{0}$, simplemente se calcula

$$A\mathbf{u} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 - 9 + 4 \\ -25 + 27 - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Por lo tanto, $\mathbf{u}$ está en Nul A. ■

El término *espacio* en el *espacio nulo* es adecuado ya que el espacio nulo de una matriz es un espacio vectorial, como se muestra en el siguiente teorema.

TEOREMA 2

El espacio nulo de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^n$. De manera equivalente, el conjunto de todas las soluciones a un sistema $A\mathbf{x} = \mathbf{0}$ de m ecuaciones lineales homogéneas con n incógnitas es un subespacio de $\mathbb{R}^n$.

DEMOSTRACIÓN Sin duda, Nul A es un subconjunto de $\mathbb{R}^n$ porque A tiene n columnas. Debemos demostrar que Nul A satisface las tres propiedades de un subespacio. Desde luego, $\mathbf{0}$ está en Nul A. Ahora, considere que $\mathbf{u}$ y $\mathbf{v}$ representan los dos vectores de Nul A. De esta forma,

$$A\mathbf{u} = \mathbf{0} \quad \text{y} \quad A\mathbf{v} = \mathbf{0}$$

Para demostrar que $\mathbf{u} + \mathbf{v}$ está en Nul A, se tiene que demostrar que $A(\mathbf{u} + \mathbf{v}) = \mathbf{0}$. Utilizando una propiedad de multiplicación de matrices, calcule

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Por lo tanto, $\mathbf{u} + \mathbf{v}$ está en Nul A, y Nul A es cerrado bajo la suma de vectores. Por último, si c es un escalar cualquiera, entonces

$$A(c\mathbf{u}) = c(A\mathbf{u}) = c(\mathbf{0}) = \mathbf{0},$$

lo que demuestra que $c\mathbf{u}$ está en Nul A. Así, Nul A es un subespacio de $\mathbb{R}^n$. ■

EJEMPLO 2 Sea H el conjunto de los vectores en $\mathbb{R}^4$ cuyas coordenadas a, b, c, d satisfacen las ecuaciones $a - 2b + 5c = d$, y $c - a = b$. Demuestre que H es un subespacio de $\mathbb{R}^4$.

SOLUCIÓN Al reordenar las ecuaciones que describen los elementos de H, y al observar que H es el conjunto de todas las soluciones del siguiente sistema de ecuaciones lineales homogéneas:

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

De acuerdo con el teorema 2, H es un subespacio de $\mathbb{R}^4$. ■

Es importante hacer notar que las ecuaciones lineales que definen el conjunto H son homogéneas. De lo contrario, el conjunto de soluciones *no* será definitivamente un subespacio (porque el vector cero no es una solución de un sistema no homogéneo). Además, en algunos casos, el conjunto de soluciones podría estar vacío.

Una descripción explícita de Nul A

No hay una relación evidente entre los vectores en Nul A y las entradas de A. Se dice que Nul A se define de forma *implícita*, ya que se define por una condición que se debe comprobar. No existe una lista o una descripción explícitas de los elementos de Nul A. Sin embargo, *resolver* la ecuación $A\mathbf{x} = \mathbf{0}$ equivale a hacer una descripción *explícita* de Nul A. El siguiente ejemplo revisa el procedimiento de la sección 1.5.

EJEMPLO 3 Determine un conjunto generador del espacio nulo de la matriz

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

SOLUCIÓN El primer paso es encontrar la solución general de $A\mathbf{x} = \mathbf{0}$ en términos de variables libres. Reduzca por filas la matriz aumentada $[A \quad \mathbf{0}]$ a la forma escalonada *reducida* para escribir las variables básicas en términos de las variables libres:

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

La solución general es $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, con x_2, x_4 y x_5 libres. Luego, hay que descomponer el vector que da la solución general en una combinación lineal de vectores donde *los pesos son las variables libres*. Es decir,

$$\begin{aligned}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} &= \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \mathbf{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}} \\
&= x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} \qquad (3)
\end{aligned}$$

Cada combinación lineal de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ es un elemento de Nul A. Por lo tanto, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es un conjunto generador para Nul A. ■

Hay que hacer notar dos hechos acerca de la solución del ejemplo 3 que se aplican a todos los problemas de este tipo, donde Nul A contiene vectores distintos de cero. Vamos a considerar estos datos más adelante.

1. El conjunto generador producido por el método del ejemplo 3 es linealmente independiente de manera automática debido a que las variables libres son los pesos de los vectores generados. Por ejemplo, vea la segunda, cuarta y quinta entradas en el vector solución en la ecuación (3) y observe que $x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$ puede ser $\mathbf{0}$ solo si los pesos x_2, x_4 y x_5 son todos cero.
2. Cuando Nul A contiene vectores distintos de cero, el número de vectores en el conjunto generador para Nul A es igual al número de variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$.

El espacio columna de una matriz

Otro subespacio importante asociado con una matriz es el espacio columna. A diferencia del espacio nulo, el espacio columna se define de forma explícita a través de combinaciones lineales.

DEFINICIÓN

El **espacio columna** de una matriz A de $m \times n$, que se denota como Col A, es el conjunto de todas las combinaciones lineales de las columnas de A. Si $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, entonces

$$\text{Col } A = \text{Gen } \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$$

Ya que Gen $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ es un subespacio, de acuerdo con el teorema 1, el siguiente teorema se deduce de la definición de Col A y del hecho de que las columnas de A están en $\mathbb{R}^m$.

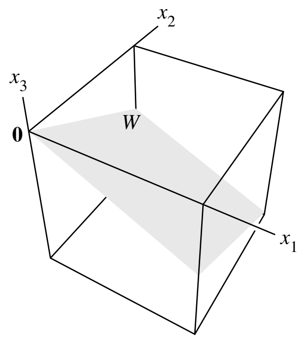
TEOREMA 3

El espacio columna de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^m$.

Observe que un vector típico de Col A se puede escribir como $A\mathbf{x}$ para alguna $\mathbf{x}$, ya que la notación $A\mathbf{x}$ representa una combinación lineal de las columnas de A. Es decir,

$$\boxed{\text{Col } A = \{\mathbf{b} : \mathbf{b} = A\mathbf{x} \text{ para alguna } \mathbf{x} \text{ en } \mathbb{R}^n\}}$$

La notación $A\mathbf{x}$ para los vectores en Col A también muestra que A es el *rango* de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$. Regresaremos a este punto de vista al final de la sección.



EJEMPLO 4 Encuentre una matriz A tal que $W = \text{Col } A$.

$$W = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\}$$

SOLUCIÓN En primer lugar, escriba W como un conjunto de combinaciones lineales.

$$W = \left\{ a \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\} = \text{Gen} \left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

En segundo lugar, utilice los vectores en el conjunto generador como las columnas de A. Sea $A = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix}$. De esta forma, $W = \text{Col } A$, como se desea. ■

Recuerde el teorema 4 de la sección 1.4, según el cual las columnas de A generan a $\mathbb{R}^m$ si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$. Podemos enunciar este hecho con palabras de la siguiente manera:

El espacio columna de una matriz A de $m \times n$ es $\mathbb{R}^m$ si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^m$.

Contraste entre Nul A y Col A

Es natural preguntarse cómo se relacionan el espacio nulo y el espacio columna de una matriz. De hecho, los dos espacios son muy diferentes, como se constata en los ejemplos 5 a 7. Sin embargo, una sorprendente conexión entre el espacio nulo y el espacio columna se verá en la sección 4.6, cuando hayamos estudiado más teoría.

EJEMPLO 5 Sea

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

a) Si el espacio columna de A es un subespacio de $\mathbb{R}^k$, ¿a qué es igual k?

b) Si el espacio nulo de A es un subespacio de $\mathbb{R}^k$, ¿a qué es igual k?

SOLUCIÓN

a) Cada una de las columnas de A tiene tres entradas, por lo que Col A es un subespacio de $\mathbb{R}^k$, donde $k = 3$.

b) Un vector $\mathbf{x}$ tal que $A\mathbf{x}$ esté definido debe tener cuatro entradas, por lo que Nul A es un subespacio de $\mathbb{R}^k$, donde $k = 4$. ■

Cuando una matriz no es cuadrada, como en el ejemplo 5, los vectores en Nul A y Col A viven en "universos" totalmente diferentes. Por ejemplo, ninguna combinación lineal de vectores en $\mathbb{R}^3$ puede producir un vector en $\mathbb{R}^4$. Cuando A es cuadrada, Nul A y Col A tienen el vector cero en común, y en casos especiales es posible que algunos vectores diferentes de cero pertenezcan tanto a Nul A como a Col A.

EJEMPLO 6 Con A como en el ejemplo 5, encuentre un vector diferente de cero en Col A y otro en Nul A.

SOLUCIÓN Es fácil encontrar un vector en Col A. Cualquier columna de A lo permite, por ejemplo, $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$. Para encontrar un vector distinto de cero en Nul A, al reducir por filas la matriz aumentada $[A \quad \mathbf{0}]$ se obtiene

$$[\,A \quad \mathbf{0}\,] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Por lo tanto, si $\mathbf{x}$ satisface $A\mathbf{x} = 0$, entonces $x_1 = -9x_3$, $x_2 = 5x_3$, $x_4 = 0$, y x_3 es libre. Al asignar un valor distinto de cero para x_3, por ejemplo, $x_3 = 1$, se obtiene un vector de Nul A, a saber, $\mathbf{x} = (-9, 5, 1, 0)$. ■

EJEMPLO 7 Con A como en el ejemplo 5, sean $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

a) Determine si $\mathbf{u}$ está en Nul A. ¿Podría $\mathbf{u}$ estar en Col A?

b) Determine si $\mathbf{v}$ está en Col A. ¿Podría $\mathbf{v}$ estar en Nul A?

SOLUCIÓN

a) Una descripción explícita de Nul A no es necesaria aquí. Basta con calcular el producto $A\mathbf{u}$.

$$A\mathbf{u} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Evidentemente, $\mathbf{u}$ *no* es una solución de $A\mathbf{x} = \mathbf{0}$, por lo que $\mathbf{u}$ no está en Nul A. Además, con cuatro entradas, $\mathbf{u}$ no podría estar en Col A, ya que Col A es un subespacio de $\mathbb{R}^3$.

b) Se reduce $[A \quad \mathbf{v}]$ a una forma escalonada.

$$[\,A \quad \mathbf{v}\,] = \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ 0 & 1 & -5 & -4 & -2 \\ 0 & 0 & 0 & 17 & 1 \end{bmatrix}$$

En este punto, es claro que la ecuación $A\mathbf{x} = \mathbf{v}$ es consistente, por lo que $\mathbf{v}$ está en Col A. Con solo tres entradas, $\mathbf{v}$ no podría estar en Nul A, ya que Nul A es un subespacio de $\mathbb{R}^4$. ■

La tabla de la página 204 resume lo que hemos aprendido acerca de Nul A y Col A. El punto 8 es una nueva formulación de los teoremas 11 y 12*a*) de la sección 1.9.

Núcleo y rango de una transformación lineal

Otros subespacios de espacios vectoriales diferentes de $\mathbb{R}^n$ con frecuencia se describen en términos de una transformación lineal y no de una matriz. Con la finalidad de precisar esto, generalizamos la definición que dimos en la sección 1.8.

Comparación entre Nul A y Col A para una matriz A de $m \times n$

Nul A	Col A
1. Nul A es un subespacio de $\mathbb{R}^n$.	**1.** Col A es un subespacio de $\mathbb{R}^m$.
2. Nul A se define implícitamente, es decir, solo estableciendo una condición ($A\mathbf{x} = \mathbf{0}$) que los vectores de Nul A deben satisfacer.	**2.** Col A se define explícitamente, es decir, se indica cómo construir vectores en Col A.
3. Se necesita tiempo para encontrar vectores de Nul A. Se requiere efectuar operaciones de fila en $[A \quad \mathbf{0}]$.	**3.** Es fácil encontrar vectores en Col A. Las columnas de A se muestran, y las otras se forman a partir de ellas.
4. No hay una relación evidente entre Nul A y las entradas de A.	**4.** Existe una relación evidente entre Col A y las entradas de A, ya que cada columna de A está en Col A.
5. Un vector típico $\mathbf{v}$ de Nul A tiene la propiedad de que $A\mathbf{v} = \mathbf{0}$.	**5.** Un vector típico $\mathbf{v}$ en Col A tiene la propiedad de que la ecuación $A\mathbf{x} = \mathbf{v}$ es consistente.
6. Dado un vector específico $\mathbf{v}$, es fácil saber si $\mathbf{v}$ está en Nul A. Solo hay que calcular $A\mathbf{v}$.	**6.** Dado un vector específico $\mathbf{v}$, puede tomar algún tiempo saber si $\mathbf{v}$ está en Col A. Se requieren operaciones de fila en $[A \quad \mathbf{v}]$.
7. Nul $A = \{\mathbf{0}\}$ si y solo si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.	**7.** Col $A = \mathbb{R}^m$ si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ en $\mathbb{R}^m$.
8. Nul $A = \{\mathbf{0}\}$ si y solo si la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es de uno a uno.	**8.** Col $A = \mathbb{R}^m$ si y solo si la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$.

DEFINICIÓN

Una **transformación lineal** T de un espacio vectorial V en un espacio vectorial W es una regla que asigna a cada vector $\mathbf{x}$ en V un único vector $T(\mathbf{x})$ en W, tal que

i. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para toda $\mathbf{u}$, $\mathbf{v}$ en V, y

ii. $T(c\mathbf{u}) = cT(\mathbf{u})$ para toda $\mathbf{u}$ en V y todo escalar c.

El **núcleo** (o **espacio nulo**) de dicha T es el conjunto de todas las $\mathbf{u}$ en V tales que $T(\mathbf{u}) = \mathbf{0}$ (el vector cero en W). El **rango** de T es el conjunto de todos los vectores en W de la forma $T(\mathbf{x})$ para alguna $\mathbf{x}$ en V. Si ocurre que T surge como una transformación matricial, por ejemplo, $T(\mathbf{x}) = A\mathbf{x}$ para alguna matriz A, entonces el núcleo y la imagen de T son solo el espacio nulo y el espacio columna de A, como se definió anteriormente.

No es difícil demostrar que el núcleo de T es un subespacio de V. La demostración es, en esencia, la misma que para el teorema 2. Además, el rango de T es un subespacio de W. Véase la figura 2 y el ejercicio 30.

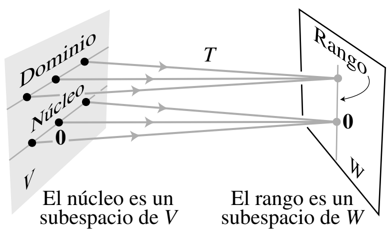


FIGURA 2 Subespacios asociados con una transformación lineal.

En aplicaciones, un subespacio suele surgir ya sea como el núcleo, o bien, como el rango de una transformación lineal adecuada. Por ejemplo, el conjunto de todas las soluciones de una ecuación diferencial lineal homogénea resulta ser el núcleo de una transformación lineal.

En general, una transformación lineal de este tipo se describe en términos de una o más derivadas de una función. Explicar esto en detalle nos llevaría demasiado lejos en este momento, por lo que consideramos solo dos ejemplos. El primero explica por qué la operación de diferenciación es una transformación lineal.

EJEMPLO 8 (Requiere conocimientos de cálculo) Sea V el espacio vectorial de todas las funciones de valores reales f definidas en un intervalo $[a, b]$ con la propiedad de que son diferenciables, y sus derivadas son funciones continuas en $[a, b]$. Sea W el espacio vectorial $C[a, b]$ de todas las funciones continuas en $[a, b]$, y sea $D : V \to W$ la transformación que cambia a f en V en su derivada f'. En cálculo, dos reglas sencillas de la diferenciación son

$$D(f + g) = D(f) + D(g) \quad \text{y} \quad D(cf) = cD(f)$$

Es decir, D es una transformación lineal. Es posible demostrar que el núcleo de D es el conjunto de las funciones constantes en $[a, b]$ y el rango de D es el conjunto W de todas las funciones continuas en $[a, b]$. ■

EJEMPLO 9 (Requiere conocimientos de cálculo) La ecuación diferencial

$$y'' + \omega^2 y = 0, \tag{4}$$

donde ω es una constante, se utiliza para describir una variedad de sistemas físicos, tales como la vibración de un resorte del que pende un peso, el movimiento de un péndulo, y el voltaje en un circuito eléctrico de inductancia-capacitancia. El conjunto de soluciones de (4) es precisamente el núcleo de la transformación lineal que mapea una función $y = f(t)$ en la función de $f''(t) + \omega^2 f(t)$. Encontrar una descripción explícita de este espacio vectorial es un problema de ecuaciones diferenciales. El conjunto solución resulta ser el espacio descrito en el ejercicio 19 de la sección 4.1. ■

PROBLEMAS DE PRÁCTICA

1. Sea $W = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a - 3b - c = 0 \right\}$. Demuestre de dos maneras diferentes que W es un subespacio de $\mathbb{R}^3$. (Utilice dos teoremas).

2. Sea $A = \begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$. Suponga que sabe que las ecuaciones $A\mathbf{x} = \mathbf{v}$ y $A\mathbf{x} = \mathbf{w}$ son consistentes. ¿Qué puede decir acerca de la ecuación $A\mathbf{x} = \mathbf{v} + \mathbf{w}$?

4.2 EJERCICIOS

1. Determine si $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$ está en Nul A, donde

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

2. Determine si $\mathbf{w} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ está en Nul A, donde

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

En los ejercicios 3 a 6, encuentre una descripción explícita de Nul A, haciendo una lista de vectores que generan el espacio nulo.

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

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

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

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

En los ejercicios 7 a 14, utilice un teorema adecuado para demostrar que el conjunto dado, W, es un espacio vectorial, o bien, encuentre un ejemplo específico de lo contrario.

7. $\left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a + b + c = 2 \right\}$ **8.** $\left\{ \begin{bmatrix} r \\ s \\ t \end{bmatrix} : 3r - 2 = 3s + t \right\}$

9. $\left\{ \begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix} : \begin{matrix} p - 3q = 4s \\ 2p = s + 5r \end{matrix} \right\}$ **10.** $\left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} 3a + b = c \\ a + b + 2c = 2d \end{matrix} \right\}$

11. $\left\{ \begin{bmatrix} s - 2t \\ 3 + 3s \\ 3s + t \\ 2s \end{bmatrix} : s, t \text{ reales} \right\}$ **12.** $\left\{ \begin{bmatrix} 3p - 5q \\ 4q \\ p \\ q + 1 \end{bmatrix} : p, q \text{ reales} \right\}$

13. $\left\{ \begin{bmatrix} c - 6d \\ d \\ c \end{bmatrix} : c, d \text{ reales} \right\}$ **14.** $\left\{ \begin{bmatrix} -s + 3t \\ s - 2t \\ 5s - t \end{bmatrix} : s, t \text{ reales} \right\}$

En los ejercicios 15 y 16, encuentre A tal que el conjunto dado sea Col A.

15. $\left\{ \begin{bmatrix} 2s + t \\ r - s + 2t \\ 3r + s \\ 2r - s - t \end{bmatrix} : r, s, t \text{ reales} \right\}$

16. $\left\{ \begin{bmatrix} b - c \\ 2b + 3d \\ b + 3c - 3d \\ c + d \end{bmatrix} : b, c, d \text{ reales} \right\}$

Para las matrices de los ejercicios 17 a 20, *a*) encuentre k tal que Nul A sea un subespacio de $\mathbb{R}^k$, y *b*) encuentre k tal que Col A sea un subespacio de $\mathbb{R}^k$.

17. $A = \begin{bmatrix} 6 & -4 \\ -3 & 2 \\ -9 & 6 \\ 9 & -6 \end{bmatrix}$ **18.** $A = \begin{bmatrix} 5 & -2 & 3 \\ -1 & 0 & -1 \\ 0 & -2 & -2 \\ -5 & 7 & 2 \end{bmatrix}$

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

20. $A = [1 \quad -3 \quad 2 \quad 0 \quad -5]$

21. Con A como en el ejercicio 17, encuentre un vector distinto de cero en Nul A y otro en Col A.

22. Con A como en el ejercicio 18, encuentre un vector distinto de cero en Nul A y otro en Col A.

23. Sea $A = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Determine si $\mathbf{w}$ está en Col A. ¿Está $\mathbf{w}$ en Nul A?

24. Sea $A = \begin{bmatrix} 10 & -8 & -2 & -2 \\ 0 & 2 & 2 & -2 \\ 1 & -1 & 6 & 0 \\ 1 & 1 & 0 & -2 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 2 \end{bmatrix}$. Determine si $\mathbf{w}$ está en Col A. ¿Está $\mathbf{w}$ en Nul A?

En los ejercicios 25 y 26, A denota una matriz de $m \times n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

25. *a*) El espacio nulo de A es el conjunto solución de la ecuación $A\mathbf{x} = \mathbf{0}$.

b) El espacio nulo de una matriz de $m \times n$ está en $\mathbb{R}^m$.

c) El espacio columna de A es el rango del mapeo $\mathbf{x} \mapsto A\mathbf{x}$.

d) Si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces Col A es $\mathbb{R}^m$.

e) El núcleo de una transformación lineal es un espacio vectorial.

f) Col A es el conjunto de todos los vectores que se pueden escribir como $A\mathbf{x}$ para alguna $\mathbf{x}$.

26. *a*) Un espacio nulo es un espacio vectorial.

b) El espacio columna de una matriz de $m \times n$ está en $\mathbb{R}^m$.

c) Col A es el conjunto de todas las soluciones de $A\mathbf{x} = \mathbf{b}$.

d) Nul A es el núcleo del mapeo $\mathbf{x} \mapsto A\mathbf{x}$.

e) El rango de una transformación lineal es un espacio vectorial.

f) El conjunto de todas las soluciones de una ecuación diferencial lineal homogénea es el núcleo de una transformación lineal.

27. Es posible demostrar que una solución del sistema que se muestra a continuación es $x_1 = 3$, $x_2 = 2$, y $x_3 = -1$. Con base en este hecho y en la teoría de esta sección, explique por qué otra solución es $x_1 = 30$, $x_2 = 20$, y $x_3 = -10$. (Observe cómo están relacionadas las soluciones, pero no realice otros cálculos).

$$\begin{aligned} x_1 - 3x_2 - 3x_3 &= 0 \\ -2x_1 + 4x_2 + 2x_3 &= 0 \\ -x_1 + 5x_2 + 7x_3 &= 0 \end{aligned}$$

28. Considere los siguientes dos sistemas de ecuaciones:

$$\begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 1 \\ 4x_1 + x_2 - 6x_3 &= 9 \end{aligned} \qquad \begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 5 \\ 4x_1 + x_2 - 6x_3 &= 45 \end{aligned}$$

Es posible demostrar que el primer sistema tiene una solución. Con base en este hecho y en la teoría de esta sección, explique por qué el segundo sistema también debe tener una solución. (No realice operaciones de fila).

29. Demuestre el teorema 3 de la siguiente manera: dada una matriz A de $m \times n$, un elemento de Col A tiene la forma $A\mathbf{x}$ para alguna $\mathbf{x}$ en $\mathbb{R}^n$. Sean $A\mathbf{x}$ y $A\mathbf{w}$ dos vectores cualesquiera en Col A.

a) Explique por qué el vector cero está en Col A.

b) Demuestre que el vector $A\mathbf{x} + A\mathbf{w}$ está en Col A.

c) Dado un escalar c, demuestre que $c(A\mathbf{x})$ está en Col A.

30. Sea $T: V \to W$ una transformación lineal de un espacio vectorial V en un espacio vectorial W. Demuestre que el rango de T es un subespacio de W. [*Sugerencia:* Considere que los elementos típicos del rango tienen la forma $T(\mathbf{x})$ y $T(\mathbf{w})$ para alguna $\mathbf{x}$, $\mathbf{w}$ en V].

31. Defina $T: \mathbb{P}_2 \to \mathbb{R}^2$ mediante $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$. Por ejemplo, si $\mathbf{p}(t) = 3 + 5t + 7t^2$, entonces $T(\mathbf{p}) = \begin{bmatrix} 3 \\ 15 \end{bmatrix}$.

a) Demuestre que T es una transformación lineal. [*Sugerencia:* Para polinomios arbitrarios $\mathbf{p}$, $\mathbf{q}$ en $\mathbb{P}_2$, calcule $T(\mathbf{p} + \mathbf{q})$ y $T(c\mathbf{p})$].

b) Encuentre un polinomio $\mathbf{p}$ en $\mathbb{P}_2$ que genere el núcleo de T y describa el rango de T.

32. Defina una transformación lineal $T: \mathbb{P}_2 \to \mathbb{R}^2$ mediante $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(0) \end{bmatrix}$. Determine los polinomios $\mathbf{p}_1$ y $\mathbf{p}_2$ en $\mathbb{P}_2$ que generan el núcleo de T, y describa el rango de T.

33. Sea $M_{2\times2}$ el espacio vectorial de todas las matrices de 2×2, y defina $T: M_{2\times2} \to M_{2\times2}$ mediante $T(A) = A + A^T$, donde $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

a) Demuestre que T es una transformación lineal.

b) Sea B cualquier elemento de $M_{2\times2}$ tal que $B^{\mathrm{T}} = B$. Determine una A en $M_{2\times2}$ tal que $T(A) = B$.

c) Demuestre que el rango de T es el conjunto de B en $M_{2\times2}$ con la propiedad de que $B^{\mathrm{T}} = B$.

d) Describa el núcleo de T.

34. (*Se requiere cálculo*) Defina a $T: C[0, 1] \to C[0, 1]$ como sigue: Para $\mathbf{f}$ en $C[0, 1]$, sea $T(\mathbf{f})$ la antiderivada $\mathbf{F}$ de $\mathbf{f}$ tal que $\mathbf{F}(0) = 0$. Demuestre que T es una transformación lineal, y describa el núcleo de T. (Véase la notación en el ejercicio 20 de la sección 4.1).

35. Sean V y W espacios vectoriales, y sea $T: V \to W$ una transformación lineal. Dado un subespacio U de V, sea $T(U)$ el conjunto de todas las imágenes de la forma $T(\mathbf{x})$, donde $\mathbf{x}$ está en U. Demuestre que $T(U)$ es un subespacio de W.

36. Dada $T: V \to W$ como en el ejercicio 35, y dado un subespacio Z de W, sea U el conjunto de todas las $\mathbf{x}$ en V tales que $T(\mathbf{x})$ está en Z. Demuestre que U es un subespacio de V.

37. [**M**] Determine si $\mathbf{w}$ está en el espacio columna de A, en el espacio nulo de A, o en ambos, donde,

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -3 \end{bmatrix}, \quad A = \begin{bmatrix} 7 & 6 & -4 & 1 \\ -5 & -1 & 0 & -2 \\ 9 & -11 & 7 & -3 \\ 19 & -9 & 7 & 1 \end{bmatrix}$$

38. [**M**] Determine si $\mathbf{w}$ está en el espacio columna de A, en el espacio nulo de A, o en ambos, donde,

$$\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$$

39. [**M**] Considere que $\mathbf{a}_1, \ldots, \mathbf{a}_5$ denotan las columnas de la matriz A, donde

$$A = \begin{bmatrix} 5 & 1 & 2 & 2 & 0 \\ 3 & 3 & 2 & -1 & -12 \\ 8 & 4 & 4 & -5 & 12 \\ 2 & 1 & 1 & 0 & -2 \end{bmatrix}, \quad B = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_4\,]$$

a) Explique por qué $\mathbf{a}_3$ y $\mathbf{a}_5$ están en el espacio columna de B.

b) Encuentre un conjunto de vectores que genere a Nul A.

c) Sea $T: \mathbb{R}^5 \to \mathbb{R}^4$ definida por $T(\mathbf{x}) = A\mathbf{x}$. Explique por qué T no es uno a uno ni sobre.

40. [**M**] Sea $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$ y $K = \text{Gen}\,\{\mathbf{v}_3, \mathbf{v}_4\}$, donde

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 \\ -12 \\ -28 \end{bmatrix}.$$

Entonces, H y K son subespacios de $\mathbb{R}^3$. De hecho, H y K son planos en $\mathbb{R}^3$ que pasan por el origen, y que se cruzan en una recta que pasa por $\mathbf{0}$. Encuentre un vector $\mathbf{w}$ distinto de cero que genere esa recta. [*Sugerencia*: $\mathbf{w}$ se puede escribir como $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ y también como $c_3\mathbf{v}_3 + c_4\mathbf{v}_4$. Para construir $\mathbf{w}$, resuelva la ecuación $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_3\mathbf{v}_3 + c_4\mathbf{v}_4$ para las incógnitas c_j].

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *Primer método:* W es un subespacio de $\mathbb{R}^3$ de acuerdo con el teorema 2, ya que W es el conjunto de todas las soluciones de un sistema de ecuaciones lineales homogéneas (donde el sistema solamente tiene una ecuación). De manera equivalente, W es el espacio nulo de la matriz de 1×3, $A = [1 \quad -3 \quad -1]$.

Segundo método: Resuelva la ecuación $a - 3b - c = 0$ para la variable principal a en términos de las variables libres b y c. Cualquier solución tiene la forma $\begin{bmatrix} 3b + c \\ b \\ c \end{bmatrix}$, donde b y c son arbitrarias, y

$$\begin{bmatrix} 3b + c \\ b \\ c \end{bmatrix} = b\underbrace{\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}_1} + c\underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}_2}$$

Este cálculo muestra que $W = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$. Por lo tanto, W es un subespacio de $\mathbb{R}^3$ de acuerdo con el teorema 1. También se podría despejar b o c de la ecuación $a - 3b - c = 0$ y obtener descripciones alternativas de W como un conjunto de combinaciones lineales de dos vectores.

2. Tanto $\mathbf{v}$ como $\mathbf{w}$ están en Col A. Ya que Col A es un espacio vectorial, $\mathbf{v} + \mathbf{w}$ debe estar en Col A. Es decir, la ecuación $A\mathbf{x} = \mathbf{v} + \mathbf{w}$ es consistente.

4.3 CONJUNTOS LINEALMENTE INDEPENDIENTES; BASES

En esta sección identificaremos y estudiaremos los subgrupos que generan un espacio vectorial V o un subespacio H tan "eficientemente" como sea posible. La idea clave es la de independencia lineal, definida como en $\mathbb{R}^n$.

Se dice que un conjunto indexado de vectores $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en V es **linealmente independiente** si la ecuación vectorial

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{1}$$

tiene *solamente* la solución trivial, $c_1 = 0,\ldots, c_p = 0$.[1]

Se dice que el conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ es **linealmente dependiente** si (1) tiene una solución no trivial, es decir, si hay algunos pesos, $c_1,\ldots, c_p$, *no todos cero*, tales que la ecuación (1) sea válida. En tal caso, la ecuación (1) se llama una **relación de dependencia lineal** entre $\mathbf{v}_1,\ldots, \mathbf{v}_p$.

Al igual que en $\mathbb{R}^n$, un conjunto que contiene un único vector $\mathbf{v}$ es linealmente independiente si y solo si $\mathbf{v} \neq 0$. Además, un conjunto de dos vectores es linealmente dependiente si y solo si uno de los vectores es un múltiplo del otro. Y cualquier conjunto que contenga al vector cero es linealmente dependiente. El siguiente teorema tiene la misma demostración que el teorema 7 de la sección 1.7.

TEOREMA 4 Un conjunto indexado $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ de dos o más vectores, con $\mathbf{v}_1 \neq \mathbf{0}$, es linealmente dependiente si y solo si alguna $\mathbf{v}_j$ (con $j > 1$) es una combinación lineal de los vectores anteriores, $\mathbf{v}_1,\ldots, \mathbf{v}_{j-1}$.

La principal diferencia entre la dependencia lineal en $\mathbb{R}^n$ y en un espacio vectorial general es que cuando los vectores no son n-adas, la ecuación homogénea (1), por lo general, no se puede escribir como un sistema de n ecuaciones lineales. Es decir, los vectores no se pueden convertir en las columnas de una matriz A con la finalidad de estudiar la ecuación $A\mathbf{x} = \mathbf{0}$. En vez de ello, debemos utilizar la definición de dependencia lineal y el teorema 4.

EJEMPLO 1 Sea $\mathbf{p}_1(t) = 1$, $\mathbf{p}_2(t) = t$ y $\mathbf{p}_3(t) = 4 - t$. Entonces, $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ es linealmente dependiente en $\mathbb{P}$ debido a $\mathbf{p}_3 = 4\mathbf{p}_1 - \mathbf{p}_2$. ■

[1] Es conveniente utilizar $c_1,\ldots, c_p$ en la ecuación (1) para los escalares en vez de $x_1,\ldots, x_p$, como lo hicimos en el capítulo 1.

EJEMPLO 2 El conjunto {sen t, cos t} es linealmente independiente en $C[0, 1]$, el espacio de todas las funciones continuas en $0 \le t \le 1$, porque sen t y cos t no son múltiplos entre sí *como vectores en* $C[0, 1]$. Es decir, no hay escalar c tal que cos $t = c \cdot$ sen t para toda t en $[0, 1]$. (Véase las gráficas de sen t y cos t). Sin embargo, {sen t cos t, sen $2t$} es linealmente dependiente debido a la identidad: sen $2t = 2$ sen t cos t, para toda t. ■

DEFINICIÓN

Sea H un subespacio de un espacio vectorial V. Un conjunto indexado de vectores $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ en V es una **base** de H si

i. $\mathcal{B}$ es un conjunto linealmente independiente, y

ii. el subespacio generado por $\mathcal{B}$ coincide con H, es decir,

$$H = \text{Gen}\,\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$$

La definición de una base se aplica al caso en que $H = V$, ya que cualquier espacio vectorial es un subespacio de sí mismo. Así, una base de V es un conjunto linealmente independiente que genera a V. Observe que cuando $H \neq V$, la condición **ii** incluye el requisito de que cada uno de los vectores $\mathbf{b}_1, \ldots, \mathbf{b}_p$ debe pertenecer a H, porque Gen $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ contiene a $\mathbf{b}_1, \ldots, \mathbf{b}_p$, como se muestra en la sección 4.1.

EJEMPLO 3 Sea A una matriz invertible de $n \times n$, por ejemplo, $A = \{\mathbf{a}_1 \cdots \mathbf{a}_n]$. Entonces, las columnas de A forman una base para $\mathbb{R}^n$, ya que son linealmente independientes y generan a $\mathbb{R}^n$, por el teorema de la matriz invertible. ■

EJEMPLO 4 Sean $\mathbf{e}_1, \ldots, \mathbf{e}_n$, las columnas de la matriz identidad I_n de $n \times n$. Es decir,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

El conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ se llama la **base estándar** para $\mathbb{R}^n$ (figura 1). ■

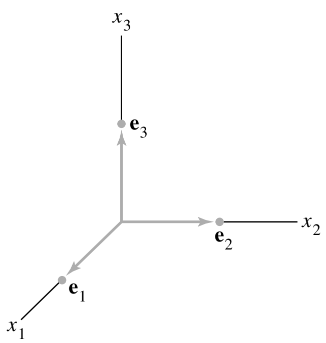


FIGURA 1
La base estándar para $\mathbb{R}^3$.

EJEMPLO 5 Sean $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$. Determine si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para $\mathbb{R}^3$.

SOLUCIÓN Puesto que hay exactamente tres vectores aquí en $\mathbb{R}^3$, se puede utilizar cualquiera de los diversos métodos para determinar si la matriz $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$ es invertible. Por ejemplo, dos remplazos de fila revelan que A tiene tres posiciones pivote. Por lo tanto, A es invertible. Como en el ejemplo 3, las columnas de A forman una base para $\mathbb{R}^3$. ■

EJEMPLO 6 Sea $S = \{1, t, t^2, \ldots, t^n\}$. Compruebe que S es una base para $\mathbb{P}_n$. Esta base se llama la **base estándar** para $\mathbb{P}_n$.

SOLUCIÓN Sin duda, S genera a $\mathbb{P}_n$. Para demostrar que S es linealmente independiente, suponga que $c_0, \ldots, c_n$ satisfacen

$$c_0 \cdot 1 + c_1 t + c_2 t^2 + \cdots + c_n t^n = \mathbf{0}(t) \tag{2}$$

Esta igualdad significa que el polinomio de la izquierda tiene los mismos valores que el polinomio cero de la derecha. Un teorema fundamental del álgebra dice que el único polinomio en $\mathbb{P}_n$, con más de n ceros es el polinomio cero. Es decir, la ecuación (2) se cumple para todas las t solo si $c_0 = \cdots = c_n = 0$. Esto demuestra que S es linealmente independiente y, por lo tanto, es una base para P_n. Véase la figura 2. ■

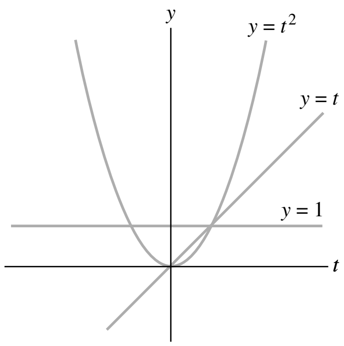


FIGURA 2
La base estándar para $\mathbb{P}_2$.

Problemas relacionados con la independencia lineal y la generación en $\mathbb{P}_n$ se manejan mejor con una técnica que se analizará en la sección 4.4.

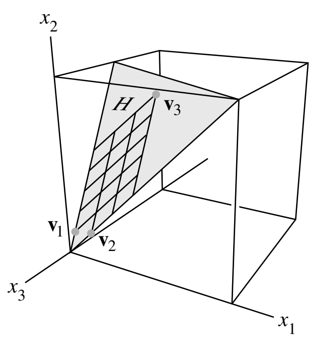
El teorema del conjunto generador

Como se verá, una base es un "eficiente" conjunto generador que no contiene vectores innecesarios. De hecho, una base se construye a partir de un conjunto generador descartando los vectores innecesarios.

EJEMPLO 7 Sean

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix} \quad \text{y} \quad H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}.$$

Considere que $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, y demuestre que Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ = Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$. Luego, encuentre una base para el subespacio H.



SOLUCIÓN Cada vector en Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ pertenece a H porque

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + 0\mathbf{v}_3$$

Ahora sea $\mathbf{x}$ cualquier vector en H, por ejemplo, $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Puesto que $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, podemos sustituir

$$\begin{aligned} \mathbf{x} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(5\mathbf{v}_1 + 3\mathbf{v}_2) \\ &= (c_1 + 5c_3)\mathbf{v}_1 + (c_2 + 3c_3)\mathbf{v}_2 \end{aligned}$$

Así, $\mathbf{x}$ está en Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$, por lo que cada vector en H ya pertenece a Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$. Llegamos a la conclusión de que H y Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ en realidad son el mismo conjunto de vectores. De lo que se deduce que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base de H, ya que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es, sin duda, linealmente independiente. ■

El siguiente teorema generaliza el ejemplo 7.

TEOREMA 5 **El teorema del conjunto generador**

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ un conjunto en V, y sea $H = \text{Gen}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

a) Si uno de los vectores de S, por ejemplo $\mathbf{v}_k$, es una combinación lineal de los vectores restantes en S, entonces el conjunto formado a partir de S al eliminar $\mathbf{v}_k$ aún genera a H.

b) Si $H \neq \{\mathbf{0}\}$, algún subconjunto de S es una base para H.

DEMOSTRACIÓN

a) Al reordenar la lista de vectores en S, si es necesario, podemos suponer que $\mathbf{v}_p$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}$, por ejemplo,

$$\mathbf{v}_p = a_1\mathbf{v}_1 + \cdots + a_{p-1}\mathbf{v}_{p-1} \tag{3}$$

Dada cualquier $\mathbf{x}$ en H, podemos escribir

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_{p-1}\mathbf{v}_{p-1} + c_p\mathbf{v}_p \tag{4}$$

para escalares adecuados $c_1, \ldots, c_p$. Sustituyendo la expresión para $\mathbf{v}_p$ de la ecuación (3) en (4), es fácil ver que $\mathbf{x}$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}$. Por lo tanto, $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}\}$ genera a H, ya que $\mathbf{x}$ es un elemento arbitrario de H.

b) Si el conjunto generador original S es linealmente independiente, entonces ya es una base para H. De lo contrario, uno de los vectores de S depende de los demás y se puede eliminar, de acuerdo con el inciso *a*). Siempre y cuando haya dos o más vectores en el conjunto generado, podemos repetir el proceso hasta que el conjunto generador sea linealmente independiente y, por lo tanto, sea una base para H. Si el conjunto generador finalmente se reduce a un vector, ese vector será distinto de cero (y, por lo tanto, linealmente independiente), ya que $H \neq \{\mathbf{0}\}$. ■

Bases para Nul A y Col A

Ya sabemos cómo encontrar los vectores que generan el espacio nulo de una matriz A. El análisis de la sección 4.2 indicó que nuestro método siempre produce un conjunto linealmente independiente cuando Nul A contiene vectores distintos de cero. Por lo tanto, en este caso, ese método produce una *base* para Nul A.

Los dos ejemplos siguientes describen un sencillo algoritmo que permite encontrar una base para el espacio columna.

EJEMPLO 8 Encuentre una base para Col B, donde

$$B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

SOLUCIÓN Cada columna que no es pivote de B es una combinación lineal de las columnas pivote. De hecho, $\mathbf{b}_2 = 4\mathbf{b}_1$ y $\mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$. De acuerdo con el teorema del conjunto generador, podemos descartar a $\mathbf{b}_2$ y $\mathbf{b}_4$, y $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ aún generará a Col B. Sea

$$S = \{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Ya que $\mathbf{b}_1 \neq 0$ y ningún vector en S es una combinación lineal de los vectores que lo preceden, S es linealmente independiente (teorema 4). Por lo tanto, S es una base de Col B. ■

¿Qué pasa con una matriz A que *no* está en forma escalonada reducida? Recuerde que toda relación de dependencia lineal entre las columnas de A se puede expresar en la forma $A\mathbf{x} = \mathbf{0}$, donde $\mathbf{x}$ es una columna de pesos. (Si algunas columnas no están implicadas en una relación de dependencia en particular, entonces sus pesos son iguales a cero). Cuando A se reduce por filas a una matriz B, las columnas de B con frecuencia son totalmente diferentes de las columnas de A. Sin embargo, las ecuaciones $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen exactamente el mismo conjunto de soluciones. Si $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$ y $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$, entonces las ecuaciones vectoriales

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{0} \quad \text{y} \quad x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n = \mathbf{0}$$

también tienen el mismo conjunto de soluciones. Es decir, las columnas de A tienen *exactamente la misma relación de dependencia lineal* que las columnas de B.

EJEMPLO 9 Es posible demostrar que la matriz

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

es equivalente por filas a la matriz B del ejemplo 8. Encuentre una base para Col A.

SOLUCIÓN En el ejemplo 8 vimos que

$$\mathbf{b}_2 = 4\mathbf{b}_1 \quad \text{y} \quad \mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$$

por lo que podemos esperar que

$$\mathbf{a}_2 = 4\mathbf{a}_1 \quad \text{y} \quad \mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3$$

¡Compruebe que este es el caso! Por lo tanto, podemos descartar a $\mathbf{a}_2$ y $\mathbf{a}_4$ cuando seleccionamos un conjunto generador mínimo de Col A. En efecto, $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ debe ser linealmente independiente porque cualquier relación de dependencia lineal entre $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5$ implicaría una relación de dependencia lineal entre $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$. Pero sabemos que $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ es un conjunto linealmente independiente. Por lo tanto, $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ es una base para Col A. Las columnas que se han utilizado para esta base son las columnas pivote de A. ■

Los ejemplos 8 y 9 ilustran el siguiente hecho que resulta de utilidad.

TEOREMA 6

Las columnas pivote de una matriz A forman una base para Col A.

DEMOSTRACIÓN La demostración general utiliza los argumentos analizados antes. Sea B la forma escalonada reducida de A. El conjunto de columnas pivote de B es linealmente independiente, pues ningún vector del conjunto es una combinación lineal de los vectores que le preceden. Puesto que A es equivalente por filas, también las columnas pivote de A son linealmente independientes, ya que cualquier relación de dependencia lineal entre las columnas de A corresponde a una relación de dependencia lineal entre las columnas de B. Por esta misma razón, todas las columnas que no sean pivote de A son una combinación lineal de las columnas pivote de A. Así, las columnas de A que no son pivote se pueden descartar del conjunto generador de Col A, de acuerdo con el teorema del conjunto generador. Esto deja a las columnas pivote de A como base para Col A. ■

Advertencia: Las columnas pivote de un matriz A son evidentes cuando A se ha reducido solamente a la forma *escalonada*. Sin embargo, tenga cuidado al usar las *columnas pivote de la misma* A como la base de Col A. Las operaciones de fila pueden cambiar el espacio columna de una matriz. Las columnas de una forma escalonada B de A con frecuencia no están en el espacio columna de A. Por ejemplo, todas las columnas de la matriz B en el ejemplo 8 tienen ceros en sus últimas entradas, por lo que no pueden generar el espacio columna de la matriz A en el ejemplo 9.

Dos perspectivas de una base

Cuando se usa el teorema del conjunto generador, la eliminación de los vectores de un conjunto generador se debe detener cuando el conjunto se convierte en linealmente independiente. Si se elimina un vector adicional, no será una combinación lineal de los vectores restantes y, por lo tanto, el conjunto más pequeño ya no generará a V. Así, una base es un conjunto generador que es lo más pequeño posible.

Una base también es un conjunto linealmente independiente lo más grande posible. Si S es una base para V, y si S se amplía con un vector (por ejemplo, $\mathbf{w}$) de V, entonces el nuevo conjunto no puede ser linealmente independiente, ya que S genera a V, y $\mathbf{w}$ es, por lo tanto, una combinación lineal de los elementos en S.

EJEMPLO 10 Los siguientes tres conjuntos de $\mathbb{R}^3$ muestran cómo se puede ampliar un conjunto linealmente independiente a una base y cómo esta ampliación adicional destruye la independencia lineal del conjunto. Además, un conjunto generador se puede reducir a una base, pero una reducción adicional destruye la propiedad de generación.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\}\qquad\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\}\qquad\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

Linealmente independiente, pero no genera a $\mathbb{R}^3$ — Una base para $\mathbb{R}^3$ — Genera a $\mathbb{R}^3$, pero es linealmente dependiente ■

PROBLEMAS DE PRÁCTICA

1. Sea $\mathbf{v}_1 = \begin{bmatrix}1\\-2\\3\end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix}-2\\7\\-9\end{bmatrix}$. Determine si $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para $\mathbb{R}^3$. ¿Es $\{\mathbf{v}_1, \mathbf{v}_2\}$ una base para $\mathbb{R}^2$?

2. Sea $\mathbf{v}_1 = \begin{bmatrix}1\\-3\\4\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}6\\2\\-1\end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix}2\\-2\\3\end{bmatrix}$ y $\mathbf{v}_4 = \begin{bmatrix}-4\\-8\\9\end{bmatrix}$. Determine una base para el subespacio W generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

3. Sea $\mathbf{v}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$ y $H = \left\{\begin{bmatrix}s\\s\\0\end{bmatrix} : s \text{ en } \mathbb{R}\right\}$. Entonces cada vector en H es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ porque

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

¿Es $\{\mathbf{v}_1, \mathbf{v}_2\}$ una base para H?

4.3 EJERCICIOS

Determine si los conjuntos de los ejercicios 1 a 8 son bases para $\mathbb{R}^3$. De los conjuntos que *no* son bases, determine cuáles son linealmente independientes y cuáles generan a $\mathbb{R}^3$. Justifique sus respuestas.

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

2. $\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}$

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

4. $\begin{bmatrix}2\\-1\\1\end{bmatrix},\begin{bmatrix}2\\-3\\2\end{bmatrix},\begin{bmatrix}-8\\5\\4\end{bmatrix}$

5. $\begin{bmatrix}3\\-3\\0\end{bmatrix},\begin{bmatrix}-3\\7\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-3\\5\end{bmatrix}$

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

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

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

Encuentre las bases para los espacios nulos de las matrices en los ejercicios 9 y 10. Consulte los comentarios que siguen el ejemplo 3 en la sección 4.2.

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

10. $\begin{bmatrix}1&1&-2&1&5\\0&1&0&-1&-2\\0&0&-8&0&16\end{bmatrix}$

11. Encuentre una base para el conjunto de vectores en $\mathbb{R}^3$ en el plano $x - 3y + 2z = 0$. [*Sugerencia:* Piense en la ecuación como en un “sistema” de ecuaciones homogéneas].

12. Encuentre una base para el conjunto de vectores de $\mathbb{R}^2$ en la recta $y = -3x$.

En los ejercicios 13 y 14, suponga que A es equivalente por filas a B. Encuentre las bases de Nul A y Col A.

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

14. $A = \begin{bmatrix} 1 & 2 & 3 & -4 & 8 \\ 1 & 2 & 0 & 2 & 8 \\ 2 & 4 & -3 & 10 & 9 \\ 3 & 6 & 0 & 6 & 9 \end{bmatrix}$,

$B = \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & 0 & 3 & -6 & 3 \\ 0 & 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 15 a 18, determine una base para el espacio generado por los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_5$ dados.

15. $\begin{bmatrix} 1 \\ 0 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -8 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 10 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -6 \\ 9 \end{bmatrix}$

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

17. [**M**] $\begin{bmatrix} 2 \\ 0 \\ -4 \\ -6 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 2 \\ -4 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ 0 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} 8 \\ 4 \\ 8 \\ -3 \\ 15 \end{bmatrix}, \begin{bmatrix} -8 \\ 4 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

18. [**M**] $\begin{bmatrix} -3 \\ 2 \\ 6 \\ 0 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -9 \\ 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -4 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \\ -14 \\ 0 \\ 13 \end{bmatrix}, \begin{bmatrix} -6 \\ 3 \\ 0 \\ -1 \\ 0 \end{bmatrix}$

19. Sea $\mathbf{v}_1 = \begin{bmatrix} 4 \\ -3 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 9 \\ -2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 7 \\ 11 \\ 6 \end{bmatrix}$, y también sea $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Es posible comprobar que $4\mathbf{v}_1 + 5\mathbf{v}_2 - 3\mathbf{v}_3 = \mathbf{0}$. Utilice esta información para encontrar una base para H. Hay más de una respuesta.

20. Sea $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \\ -2 \\ -5 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 4 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 5 \\ -6 \\ -14 \end{bmatrix}$. Es posible comprobar que $2\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$. Utilice esta información para encontrar una base para $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

En los ejercicios 21 y 22, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) Un solo vector por sí mismo es linealmente dependiente.

b) Si $H = \text{Gen}\,\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$, entonces $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ es una base para H.

c) Las columnas de una matriz invertible de $n \times n$ forman una base para $\mathbb{R}^n$.

d) Una base es un conjunto generador que es tan grande como sea posible.

e) En algunos casos, las relaciones de dependencia lineal entre las columnas de una matriz se pueden ver afectadas por algunas operaciones elementales de fila en la matriz.

22. *a*) Un conjunto linealmente independiente en un subespacio H es una base para H.

b) Si un conjunto finito S de vectores diferentes de cero genera un espacio vectorial V, entonces algún subconjunto de S es una base para V.

c) Una base es un conjunto linealmente independiente que es lo más grande posible.

d) El método estándar para la obtención de un conjunto generador para Nul A, que se describe en la sección 4.2, a veces falla al producir una base para Nul A.

e) Si B es una forma escalonada de una matriz A, entonces las columnas pivote de B forman una base para Col A.

23. Suponga que $\mathbb{R}^4 = \text{Gen}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$. Explique por qué $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es una base para $\mathbb{R}^4$.

24. Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente en $\mathbb{R}^n$. Explique por qué $\mathcal{B}$ debe ser una base para $\mathbb{R}^n$.

25. Sea $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ y sea H el conjunto de vectores en $\mathbb{R}^3$ cuyas segunda y tercera entradas son iguales. Entonces, todo vector de H tiene una expansión única como una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, porque

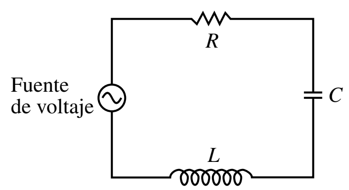
$$\begin{bmatrix} s \\ t \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (t-s)\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

para cualquier s y t. ¿Es $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ una base para H? ¿Por qué?

26. En el espacio vectorial de todas las funciones de valores reales, encuentre una base para el subespacio generado por $\{\text{sen}\,t, \text{sen}\,2t, \text{sen}\,t\cos t\}$.

27. Sea V el espacio vectorial de funciones que describen la vibración de un sistema masa-resorte. (Consulte el ejercicio 19 de la sección 4.1). Encuentre una base para V.

28. (*Circuito RLC*) El circuito de la figura consiste en una resistencia (R, ohms), un inductor (L, henrys), un condensador (C, farads), y una fuente de voltaje inicial. Sea $b = R/(2L)$, y suponga que R, L y C se han seleccionado de manera que b también es igual a $1/\sqrt{LC}$. (Esto se hace, por ejemplo, cuando el circuito se utiliza en un voltímetro). Sea $v(t)$ el voltaje (en volts) en el tiempo t, medido a través del condensador. Es posible demostrar que v está en el espacio nulo H de la transformación lineal que mapea $v(t)$ en $Lv''(t) + Rv'(t) + (1/C)v(t)$, y H se compone de todas las funciones de la forma $v(t) = e^{-bt}(c_1 + c_2 t)$. Encuentre una base para H.



Los ejercicios 29 y 30 indican que cada base de $\mathbb{R}^n$ debe contener exactamente n vectores.

29. Sea $S = \{\mathbf{v}_1,\ldots, \mathbf{v}_k\}$ un conjunto de k vectores en $\mathbb{R}^n$, con $k < n$. Utilice un teorema de la sección 1.4 para explicar por qué S no puede ser una base para $\mathbb{R}^n$.

30. Sea $S = \{\mathbf{v}_1,\ldots, \mathbf{v}_k\}$ un conjunto de k vectores en $\mathbb{R}^n$, con $k > n$. Utilice un teorema del capítulo 1 para explicar por qué S no puede ser una base para $\mathbb{R}^n$.

Los ejercicios 31 y 32 revelan una importante conexión entre la independencia lineal y las transformaciones lineales, y le permiten practicar el uso de la definición de dependencia lineal. Sean V y W espacios vectoriales, sea $T : V \to W$ una transformación lineal, y sea $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ un subconjunto de V.

31. Demuestre que si $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ es linealmente dependiente en V, entonces el conjunto de imágenes, $\{T(\mathbf{v}_1),\ldots, T(v_p)\}$ es linealmente dependiente en W. Este hecho demuestra que si una transformación lineal mapea un conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en un conjunto linealmente *independiente* $\{T(\mathbf{v}_1),\ldots, T(\mathbf{v}_p)\}$, entonces el conjunto original es linealmente independiente también (porque no puede ser linealmente dependiente).

32. Suponga que T es una transformación uno a uno, de modo que la ecuación $T(\mathbf{u}) = T(\mathbf{v})$ siempre implica que $\mathbf{u} = \mathbf{v}$. Demuestre que si el conjunto de imágenes $\{T(\mathbf{v}_1),\ldots, T(\mathbf{v}_p)\}$ es linealmente dependiente, entonces $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ es linealmente dependiente. Este hecho demuestra que *una transformación lineal uno a uno mapea un conjunto linealmente independiente sobre un conjunto linealmente independiente* (porque en este caso el conjunto de las imágenes no puede ser linealmente dependiente).

33. Considere los polinomios $\mathbf{p}_1(t) = 1 + t^2$ y $\mathbf{p}_2(t) = 1 - t^2$. ¿Es $\{\mathbf{p}_1, \mathbf{p}_2\}$ un conjunto linealmente independiente en $\mathbb{P}_3$? ¿Por qué?

34. Considere los polinomios $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, y $\mathbf{p}_3(t) = 2$ (para toda t). Por inspección, escriba una relación de dependencia lineal entre $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$. Después encuentre una base para Gen $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$.

35. Sea V un espacio vectorial que contiene un conjunto linealmente independiente $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$. Explique cómo construir un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ en V tal que $\{\mathbf{v}_1, \mathbf{v}_3\}$ sea una base para Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

36. [**M**] Sea $H = \text{Gen}\ \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y $K = \text{Gen}\ \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, donde

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ -4 \end{bmatrix},$$

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ -2 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 2 \\ -6 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 6 \\ -2 \end{bmatrix}$$

Encuentre las bases para H, K y $H + K$. (Véase los ejercicios 33 y 34 en la sección 4.1).

37. [**M**] Demuestre que $\{t, \text{sen}\ t, \cos 2t, \text{sen}\ t \cos t\}$ es un conjunto linealmente independiente de funciones definidas en $\mathbb{R}$. Comience suponiendo que

$$c_1 \cdot t + c_2 \cdot \text{sen}\ t + c_3 \cdot \cos 2t + c_4 \cdot \text{sen}\ t \cos t = 0 \qquad (5)$$

La ecuación (5) debe ser válida para toda t real, así que elija varios valores específicos de t (es decir, $t = 0, .1, .2$) hasta obtener un sistema de ecuaciones suficientes para determinar que todas las c_j deben ser cero.

38. [**M**] Demuestre que $\{1, \cos t, \cos^2 t,\ldots, \cos^6 t\}$ es un conjunto linealmente independiente de funciones definidas en $\mathbb{R}$. Utilice el método del ejercicio 37. (Este resultado se necesitará en el ejercicio 34 de la sección 4.5).

WEB

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sea $A = [\mathbf{v}_1 \quad \mathbf{v}_2]$. Las operaciones de fila muestran que

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

No todas las filas de A contienen una posición pivote. De manera que las columnas de A no generan a $\mathbb{R}^3$, de acuerdo con el teorema 4 de la sección 1.4. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ no es una base para $\mathbb{R}^3$. Ya que $\mathbf{v}_1$ y $\mathbf{v}_2$ no están en $\mathbb{R}^2$, no pueden ser la base de $\mathbb{R}^2$. Sin embargo, puesto que es evidente que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes, son una base para un subespacio de $\mathbb{R}^3$, a saber, Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$.

2. Construya una matriz cuyo espacio columna sea el espacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, y después reduzca por filas a A para encontrar sus columnas pivote.

$$A = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 20 & 4 & -20 \\ 0 & -25 & -5 & 25 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 5 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las dos primeras columnas de A son las columnas pivote y, por consiguiente, forman una base de Col A = W. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para W. Observe que no se necesita la forma escalonada reducida de A para localizar las columnas pivote.

3. Ni $\mathbf{v}_1$ ni $\mathbf{v}_2$ están en H, por lo que $\{\mathbf{v}_1, \mathbf{v}_2\}$ no puede ser una base para H. De hecho, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para el *plano* de todos los vectores de la forma $(c_1, c_2, 0)$, pero H es solo una *recta*.

4.4 SISTEMAS DE COORDENADAS

Una razón importante para especificar una base $\mathcal{B}$ para un espacio vectorial V es imponer un "sistema de coordenadas" en V. En esta sección se mostrará que si $\mathcal{B}$ contiene n vectores, entonces el sistema de coordenadas hará que V actúe como $\mathbb{R}^n$. Si V es ya $\mathbb{R}^n$ mismo, entonces $\mathcal{B}$ determinará un sistema de coordenadas que da una nueva perspectiva a V.

La existencia de sistemas de coordenadas se basa en el siguiente resultado fundamental.

TEOREMA 7

Teorema de la representación única

Sea $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Así, para cada $\mathbf{x}$ en V, existe un conjunto único de escalares $c_1,\ldots, c_n$ tal que

$$x = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n \tag{1}$$

DEMOSTRACIÓN Puesto que $\mathcal{B}$ genera a V, existen escalares tales que la ecuación (1) es válida. Suponga que $\mathbf{x}$ también tiene la representación

$$\mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_n\mathbf{b}_n$$

para escalares $d_1,\ldots, d_n$. Así, al restar, se tiene

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_n - d_n)\mathbf{b}_n \tag{2}$$

Puesto que $\mathcal{B}$ es linealmente independiente, los pesos en la ecuación (2) deben ser cero. Es decir, $c_j = d_j$ para $1 \le j \le n$. ■

DEFINICIÓN

Suponga que $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ es una base para V y que $\mathbf{x}$ está en V. Las **coordenadas de x respecto de la base $\mathcal{B}$** (o las **$\mathcal{B}$-coordenadas de x**) son los pesos $c_1,\ldots, c_n$ tales que $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$.

Si $c_1,\ldots, c_n$ son las $\mathcal{B}$-coordenadas de $\mathbf{x}$, entonces el vector en $\mathbb{R}^n$

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

es el **vector de coordenadas de x (respecto de $\mathcal{B}$), o el vector de $\mathcal{B}$-coordenadas de x**. El mapeo $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$ es el **mapeo de coordenadas (determinado por $\mathcal{B}$)**.[1]

[1] El concepto de un mapeo de coordenadas supone que la base $\mathcal{B}$ es un conjunto indexado cuyos vectores se listan en un orden fijo asignado con anterioridad. Esta propiedad hace que la definición de $[\,\mathbf{x}\,]_{\mathcal{B}}$ no sea ambigua.

EJEMPLO 1 Considere una base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ para $\mathbb{R}^2$, donde $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Supongamos que una $\mathbf{x}$ en $\mathbb{R}^2$ tiene el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Determine $\mathbf{x}$.

SOLUCIÓN Las coordenadas $\mathcal{B}$ de $\mathbf{x}$ nos dicen cómo construir $\mathbf{x}$ a partir de los vectores en $\mathcal{B}$. Es decir,

$$\mathbf{x} = (-2)\mathbf{b}_1 + 3\mathbf{b}_2 = (-2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix} \qquad \blacksquare$$

EJEMPLO 2 Las entradas en el vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$ son las coordenadas de $\mathbf{x}$ respecto de la *base estándar* $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, ya que

$$\begin{bmatrix} 1 \\ 6 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{e}_1 + 6 \cdot \mathbf{e}_2$$

Si $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, entonces $[\,\mathbf{x}\,]_{\mathcal{E}} = \mathbf{x}$. $\blacksquare$

Una interpretación gráfica de coordenadas

Un sistema de coordenadas en un conjunto se compone de un mapeo uno a uno de los puntos en el conjunto dentro de $\mathbb{R}^n$. Por ejemplo, el papel común para gráficas proporciona un sistema de coordenadas del plano cuando se seleccionan ejes perpendiculares y una unidad de medida en cada eje. La figura 1 muestra la base estándar $\{\mathbf{e}_1, \mathbf{e}_2\}$, los vectores $\mathbf{b}_1$ $(= \mathbf{e}_1)$ y $\mathbf{b}_2$ del ejemplo 1, y del vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$. Las coordenadas 1 y 6 dan la ubicación de $\mathbf{x}$ respecto de la base estándar: 1 unidad en la dirección $\mathbf{e}_1$ y 6 unidades en la dirección $\mathbf{e}_2$.

La figura 2 muestra los vectores $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{x}$ de la figura 1. (Desde el punto de vista geométrico, los tres vectores se encuentran en una recta vertical en ambas figuras). Sin embargo, el sistema de coordenadas estándar se borró y se remplazó por una malla especialmente adaptada a la base $\mathcal{B}$ del ejemplo 1. El vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ da la ubicación de $\mathbf{x}$ en este nuevo sistema de coordenadas: -2 unidades en la dirección $\mathbf{b}_1$ y 3 unidades en la dirección $\mathbf{b}_2$.

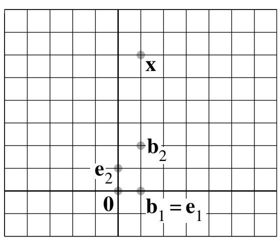


FIGURA 1
Papel cuadriculado estándar.

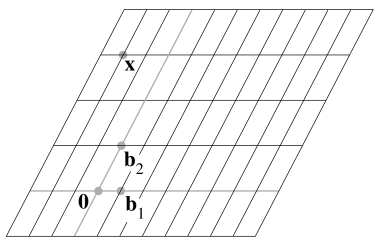


FIGURA 2
Papel para gráficas *B*.

EJEMPLO 3 En cristalografía, la descripción de una red cristalina se mejora eligiendo una base $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ para $\mathbb{R}^3$ que corresponda a tres aristas adyacentes de una "celda unitaria" del cristal. Se construye una red entera apilando muchas copias de una sola celda. Hay 14 tipos básicos de celdas unitarias; en la figura 3, se muestran tres.[2]

[2] Adaptado de *The Science and Engineering of Materials*, 4a. edición, por Donald R. Askeland (Boston: Prindle, Weber & Schmidt © 2002), p. 36.

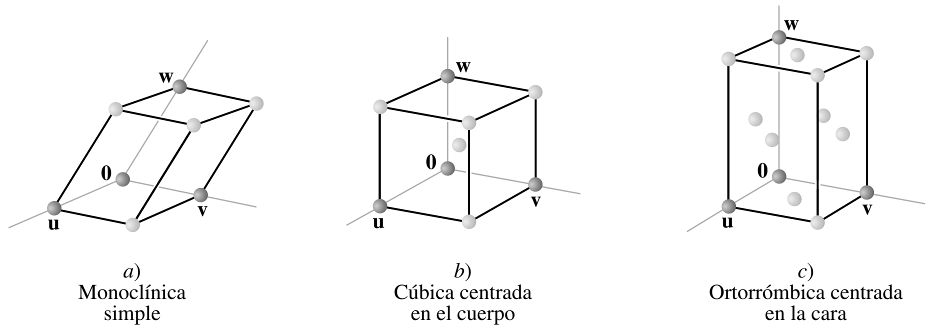


FIGURA 3 Ejemplos de celdas unitarias.

Las coordenadas de los átomos dentro del cristal están dadas respecto de la base de la red. Por ejemplo,

$$\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

identifica el átomo centrado en la cara superior de la celda en la figura 3 *c*). ■

Coordenadas en $\mathbb{R}^n$

Cuando una base $\mathcal{B}$ para $\mathbb{R}^n$ está fija, el vector de coordenadas $\mathcal{B}$ de una $\mathbf{x}$ dada es fácil de determinar, como se muestra en el siguiente ejemplo.

EJEMPLO 4 Sean $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Determine el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$ de $\mathbf{x}$ respecto de $\mathcal{B}$.

SOLUCIÓN Las coordenadas $\mathcal{B}$ de c_1, c_2 de $\mathbf{x}$ satisfacen

$$c_1\underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\mathbf{b}_1} + c_2\underbrace{\begin{bmatrix} -1 \\ 1 \end{bmatrix}}_{\mathbf{b}_2} = \underbrace{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}_{\mathbf{x}}$$

o

$$\begin{array}{cc} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \\ \mathbf{b}_1 \quad \mathbf{b}_2 \qquad\qquad\qquad \mathbf{x} \end{array} \tag{3}$$

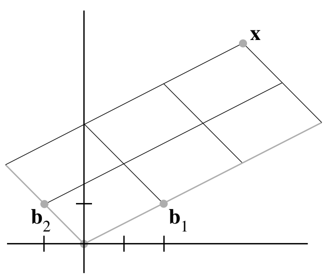


FIGURA 4 El vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$ es (3, 2).

Esta ecuación se puede resolver mediante operaciones de fila en una matriz aumentada o utilizando la inversa de la matriz a la izquierda. En cualquier caso, la solución es $c_1 = 3$, $c_2 = 2$. Por lo tanto, $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$, y

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$ ■

Véase la figura 4.

La matriz en la ecuación (3) cambia las coordenadas $\mathcal{B}$ de un vector $\mathbf{x}$ en las coordenadas estándar para $\mathbf{x}$. Es posible realizar un cambio análogo de coordenadas en $\mathbb{R}^n$ para una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Sea

$$P_{\mathcal{B}} = \{\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n\}$$

Entonces la ecuación vectorial

$$\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_n\mathbf{b}_n$$

es equivalente a

$$\boxed{\mathbf{x} = P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}} \tag{4}$$

Se llama $P_{\mathcal{B}}$ a la **matriz de cambio de coordenadas** de $\mathcal{B}$ a la base estándar en $\mathbb{R}^n$. Multiplicando por la izquierda a $P_{\mathcal{B}}$ se transforma al vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ en $\mathbf{x}$. La ecuación de cambio de coordenadas (4) es importante y será necesaria en varias secciones de los capítulos 5 y 7.

Puesto que las columnas de $P_{\mathcal{B}}$ forman una base para $\mathbb{R}^n$, $P_{\mathcal{B}}$ es invertible (por el teorema de la matriz invertible). Al multiplicar por la izquierda por $P_{\mathcal{B}}^{-1}$ se convierte a $\mathbf{x}$ en su vector de coordenadas $\mathcal{B}$:

$$P_{\mathcal{B}}^{-1}\mathbf{x} = [\mathbf{x}]_{\mathcal{B}}$$

La correspondencia $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$, producida aquí por $P_{\mathcal{B}}^{-1}$, es el mapeo de coordenadas mencionado anteriormente. Ya que $P_{\mathcal{B}}^{-1}$ es una matriz invertible, el mapeo de coordenadas es una transformación lineal uno a uno de $\mathbb{R}^n$ en $\mathbb{R}^n$, de acuerdo con el teorema de la matriz invertible. (Véase también el teorema 12 de la sección 1.9). Esta propiedad del mapeo de coordenadas también es cierto en un espacio vectorial general que tiene una base, como se verá más adelante.

El mapeo de coordenadas

La elección de una base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ de un espacio vectorial V introduce un sistema de coordenadas en V. El mapeo de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ conecta al espacio V, posiblemente desconocido, con el conocido espacio $\mathbb{R}^n$. Véase la figura 5. Los puntos de V ahora se pueden identificar por sus nuevos "nombres".

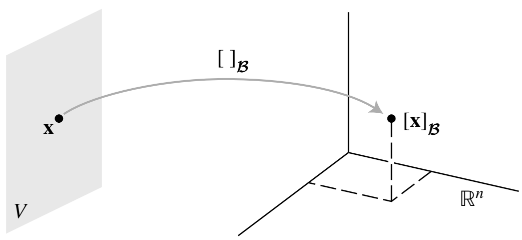


FIGURA 5 El mapeo de coordenadas de V sobre $\mathbb{R}^n$.

TEOREMA 8

> Sea $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Así, el mapeo de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una transformación lineal uno a uno de V en $\mathbb{R}^n$.

DEMOSTRACIÓN Tome dos vectores típicos en V, por ejemplo,

$$\mathbf{u} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$
$$\mathbf{w} = d_1\mathbf{b}_1 + \cdots + d_n\mathbf{b}_n$$

Luego, utilizando las operaciones de vectores,

$$\mathbf{u} + \mathbf{w} = (c_1 + d_1)\mathbf{b}_1 + \cdots + (c_n + d_n)\mathbf{b}_n$$

De ello se sigue que

$$[\mathbf{u}+\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}$$

Por lo tanto, el mapeo de coordenadas conserva la adición. Si r es un escalar cualquiera, entonces

$$r\mathbf{u} = r(c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n) = (rc_1)\mathbf{b}_1 + \cdots + (rc_n)\mathbf{b}_n$$

De esta forma,

$$[r\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[\mathbf{u}]_{\mathcal{B}}$$

Así, el mapeo de coordenadas también conserva la multiplicación escalar, y por consiguiente, es una transformación lineal. Véase los ejercicios 23 y 24 para comprobar que el mapeo de coordenadas es uno a uno y mapea V sobre $\mathbb{R}^n$. ■

La linealidad del mapeo de coordenadas se extiende a las combinaciones lineales, al igual que en la sección 1.8. Si $\mathbf{u}_1, \ldots, \mathbf{u}_p$ están en V y si $c_1, \ldots, c_p$ son escalares, entonces

$$[c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p]_{\mathcal{B}} = c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_p[\mathbf{u}_p]_{\mathcal{B}} \qquad (5)$$

Es decir, (5) nos dice que el vector de $\mathcal{B}$-coordenadas de una combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_p$ es la *misma* combinación lineal de sus vectores de coordenadas.

El mapeo de coordenadas en el teorema 8 es un importante ejemplo de un *isomorfismo* de V en $\mathbb{R}^n$. En general, una transformación lineal uno a uno de un espacio vectorial V en un espacio vectorial W se llama **isomorfismo** de V en W (el término proviene de los vocablos griegos *iso*, que significa "lo mismo", y *morfé*, que significa "forma" o "estructura"). La notación y la terminología para V y W pueden diferir, pero los dos espacios son indistinguibles como espacios vectoriales. *Cada cálculo de espacio vectorial en V se reproduce con exactitud en W, y viceversa*. En particular, cualquier espacio vectorial real con una base de n vectores es indistinguible de $\mathbb{R}^n$. Véase los ejercicios 25 y 26.

EJEMPLO 5 Sea $\mathcal{B}$ la base estándar del espacio $\mathbb{P}_3$ de polinomios; es decir, sea $\mathcal{B} = \{1, t, t^2, t^3\}$. Un elemento típico $\mathbf{p}$ de $\mathbb{P}_3$ tiene la forma

$$\mathbf{p}(t) = a_0 + a_1t + a_2t^2 + a_3t^3$$

Puesto que $\mathbf{p}$ ya se ha desplegado como una combinación lineal de los vectores básicos estándar, concluimos que

$$[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Así, el mapeo de coordenadas $\mathbf{p} \mapsto [\mathbf{p}]_{\mathcal{B}}$ es un isomorfismo de $\mathbb{P}_3$ sobre $\mathbb{R}^4$. Todas las operaciones de espacio vectorial en $\mathbb{P}_3$ corresponden a operaciones en $\mathbb{R}^4$. ■

Si pensamos en $\mathbb{P}_3$ y $\mathbb{R}^4$ como despliegues en dos pantallas de computadora que se conectan a través del mapeo de coordenadas, entonces cada operación de espacio vectorial en $\mathbb{P}_3$ en una pantalla es duplicada exactamente por una operación vectorial correspondiente en $\mathbb{R}^4$ en la otra pantalla. Los vectores en la pantalla $\mathbb{P}_3$ tienen un aspecto diferente respecto de los que aparecen en la pantalla de $\mathbb{R}^4$, pero "actúan" como vectores exactamente de la misma forma. Véase la figura 6.

FIGURA 6 El espacio $\mathbb{P}_3$ es isomorfo a $\mathbb{R}^4$.

EJEMPLO 6 Utilice vectores de coordenadas para comprobar que los polinomios $1 + 2t^2$, $4 + t + 5t^2$, y $3 + 2t$ son linealmente dependientes en $\mathbb{P}_2$.

SOLUCIÓN El mapeo de coordenadas del ejemplo 5 produce los vectores de coordenadas (1, 0, 2), (4, 1, 5) y (3, 2, 0), respectivamente. Al representar estos vectores como las *columnas* de una matriz A, podemos determinar su independencia mediante reducción por filas de la matriz aumentada de $A\mathbf{x} = \mathbf{0}$:

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

Las columnas de A son linealmente dependientes, por lo que los polinomios correspondientes son linealmente dependientes. De hecho, es fácil comprobar que la columna 3 de A es la columna 2 multiplicada por 2, menos la columna 1 multiplicada por 5. La relación correspondiente de los polinomios es

$$3 + 2t = 2(4 + t + 5t^2) - 5(1 + 2t^2) \qquad \blacksquare$$

El último ejemplo se refiere a un plano en $\mathbb{R}^3$ que es isomorfo a $\mathbb{R}^2$.

EJEMPLO 7 Sea

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}.$$

y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Entonces $\mathcal{B}$ es una base para $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$. Determine si $\mathbf{x}$ se encuentra en H y, si lo está, encuentre el vector de coordenadas de $\mathbf{x}$ respecto de $\mathcal{B}$.

SOLUCIÓN Si $\mathbf{x}$ está en H, entonces la siguiente ecuación vectorial es consistente:

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

Los escalares c_1 y c_2, si existen, son las coordenadas $\mathcal{B}$ de $\mathbf{x}$. Usando operaciones de fila, se obtiene

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Así, $c_1 = 2$, $c_2 = 3$ y $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. El sistema de coordenadas en H determinadas por $\mathcal{B}$ se muestra en la figura 7. ■

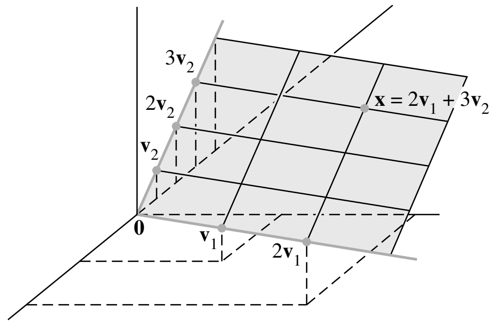


FIGURA 7 Un sistema de coordenadas en un plano H en $\mathbb{R}^3$.

Si se eligiera una base diferente para H, ¿el sistema de coordenadas asociado también haría a H isomorfo a $\mathbb{R}^2$? Sin duda, esto debe ser verdad. Así lo demostraremos en la siguiente sección.

PROBLEMAS DE PRÁCTICA

1. Sea $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} -8 \\ 2 \\ 3 \end{bmatrix}$.

a) Demuestre que el conjunto $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ es una base de $\mathbb{R}^3$.

b) Encuentre la matriz de cambio de las coordenadas de $\mathcal{B}$ a la base estándar.

c) Escriba la ecuación que relaciona $\mathbf{x}$ en $\mathbb{R}^3$ con $[\,\mathbf{x}\,]_{\mathcal{B}}$.

d) Encuentre $[\,\mathbf{x}\,]_{\mathcal{B}}$, para la $\mathbf{x}$ dada anteriormente.

2. El conjunto $\mathcal{B} = \{1 + t, 1 + t^2, t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 6 + 3t - t^2$ en relación con $\mathcal{B}$.

4.4 EJERCICIOS

En los ejercicios 1 a 4, encuentre el vector $\mathbf{x}$ determinado por el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$ y la base $\mathcal{B}$.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$

3. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -3 \\ 0 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$

4. $\mathcal{B} = \left\{ \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 2 \\ -1 \end{bmatrix}$

En los ejercicios 5 a 8, encuentre el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$ de $\mathbf{x}$ respecto de la base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.

5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

6. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -1 \\ -6 \end{bmatrix}$

7. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 9 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 8 \\ -9 \\ 6 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$

En los ejercicios 9 y 10, encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a la base estándar en $\mathbb{R}^n$.

9. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -5 \end{bmatrix} \right\}$

10. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \right\}$

En los ejercicios 11 y 12, utilice una matriz inversa para encontrar $[\,\mathbf{x}\,]_{\mathcal{B}}$ para $\mathbf{x}$ y $\mathcal{B}$.

11. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 5 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$

12. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

13. El conjunto $\mathcal{B} = \{1 + t^2, t + t^2, 1 + 2t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 1 + 4t + 7t^2$ respecto de $\mathcal{B}$.

14. El conjunto $\mathcal{B} = \{1 - t^2, t - t^2, 2 - t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 1 + 3t - 6t^2$ respecto de $\mathcal{B}$.

En los ejercicios 15 y 16, marque cada enunciado como verdadero o falso. Justifique sus respuestas. A menos que se indique lo contrario, $\mathcal{B}$ es una base para un espacio vectorial V.

15. *a*) Si $\mathbf{x}$ está en V y si $\mathcal{B}$ contiene n vectores, entonces el vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$ está en $\mathbb{R}^n$.

b) Si $P_{\mathcal{B}}$ es la matriz del cambio de coordenadas, entonces $[\,\mathbf{x}\,]_{\mathcal{B}} = P_{\mathcal{B}}\mathbf{x}$, para $\mathbf{x}$ en V.

c) Los espacios vectoriales $\mathbb{P}_3$ y $\mathbb{R}^3$ son isomorfos.

16. *a*) Si $\mathcal{B}$ es la base estándar para $\mathbb{R}^n$, entonces el vector de coordenadas $\mathcal{B}$ de una $\mathbf{x}$ en $\mathbb{R}^n$ es $\mathbf{x}$ misma.

b) La correspondencia $[\,\mathbf{x}\,]_{\mathcal{B}} \mapsto \mathbf{x}$ se llama el mapeo de coordenadas.

c) En algunos casos, un plano en $\mathbb{R}^3$ puede ser isomorfo a $\mathbb{R}^2$.

17. Los vectores $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$ generan a $\mathbb{R}^2$, pero no forman una base. Encuentre formas diferentes de expresar $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ como una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

18. Sea $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Explique por qué los vectores de coordenadas $\mathcal{B}$ de $\mathbf{b}_1,\ldots, \mathbf{b}_n$ son las columnas $\mathbf{e}_1,\ldots, \mathbf{e}_n$ de la matriz identidad de $n \times n$.

19. Sea S un conjunto finito en un espacio vectorial V con la propiedad de que cada $\mathbf{x}$ en V tiene una representación única como una combinación lineal de elementos de S. Demuestre que S es una base de V.

20. Suponga que $\{\mathbf{v}_1,\ldots, \mathbf{v}_4\}$ es un conjunto linealmente dependiente que genera un espacio vectorial V. Demuestre que cada $\mathbf{w}$ en V se puede expresar en más de una forma como una combinación lineal de $\mathbf{v}_1,\ldots, \mathbf{v}_4$. [*Sugerencia:* Considere $\mathbf{w} = k_1\mathbf{v}_1 + \cdots + k_4\mathbf{v}_4$ un vector arbitrario en V. Utilice la dependencia lineal de $\{\mathbf{v}_1,\ldots, \mathbf{v}_4\}$ para obtener otra representación de $\mathbf{w}$ como una combinación lineal de $\mathbf{v}_1,\ldots, \mathbf{v}_4$].

21. Sea $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \end{bmatrix} \right\}$. Puesto que el mapeo de coordenadas determinado por $\mathcal{B}$ es una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$, este mapeo se debe implementar mediante alguna matriz A de 2×2. Encuéntrela. [*Sugerencia:* La multiplicación por A debería transformar un vector $\mathbf{x}$ en su vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$].

22. Sea $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ una base para $\mathbb{R}^n$. Obtenga una descripción de una matriz A de $n \times n$ que implemente el mapeo de coordenadas $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$. (Véase el ejercicio 21).

Los ejercicios 23 a 26 se refieren a un espacio vectorial V, una base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$, y el mapeo de coordenadas $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$.

23. Demuestre que el mapeo de coordenadas es uno a uno. (*Sugerencia:* Suponga que $[\,\mathbf{u}\,]_{\mathcal{B}} = [\,\mathbf{w}\,]_{\mathcal{B}}$ para algunas $\mathbf{u}$ y $\mathbf{w}$ en V, y demuestre que $\mathbf{u} = \mathbf{w}$).

24. Demuestre que el mapeo de coordenadas es *sobre* $\mathbb{R}^n$. Es decir, dada cualquier $\mathbf{y}$ en $\mathbb{R}^n$, con entradas $y_1,\ldots, y_n$, obtenga $\mathbf{u}$ en V tal que $[\,\mathbf{u}\,]_{\mathcal{B}} = \mathbf{y}$.

25. Demuestre que un subconjunto $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ en V es linealmente independiente si y solo si el conjunto de vectores de coordenadas $\{[\,\mathbf{u}_1\,]_{\mathcal{B}}, \ldots, [\,\mathbf{u}_p\,]_{\mathcal{B}}\}$ es linealmente independiente en $\mathbb{R}^n$. (*Sugerencia:* Como el mapeo de coordenadas es uno a uno, las siguientes ecuaciones tienen las mismas soluciones, $c_1,\ldots, c_p$).

$c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = \mathbf{0}$ El vector cero en V

$[c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p]_{\mathcal{B}} = [\mathbf{0}]_{\mathcal{B}}$ El vector cero en $\mathbb{R}^n$

26. Dados los vectores $\mathbf{u}_1,\ldots, \mathbf{u}_p$, y $\mathbf{w}$ en V, demuestre que $\mathbf{w}$ es una combinación lineal de $\mathbf{u}_1,\ldots, \mathbf{u}_p$ si y solo si $[\,\mathbf{w}\,]_{\mathcal{B}}$ es una combinación lineal de los vectores de coordenadas $[\,\mathbf{u}_1\,]_{\mathcal{B}}, \ldots, [\,\mathbf{u}_p\,]_{\mathcal{B}}$.

En los ejercicios 27 a 30, utilice los vectores de coordenadas para probar la independencia lineal de los conjuntos de polinomios. Explique su trabajo.

27. $1 + 2t^3,\ 2 + t - 3t^2,\ -t + 2t^2 - t^3$

28. $1 - 2t^2 - t^3,\ t + 2t^3,\ 1 + t - 2t^2$

29. $(1 - t)^2,\ t - 2t^2 + t^3,\ (1 - t)^3$

30. $(2 - t)^3,\ (3 - t)^2,\ 1 + 6t - 5t^2 + t^3$

31. Utilice los vectores de coordenadas para comprobar si los siguientes conjuntos de polinomios generan a $\mathbb{P}_2$. Justifique sus conclusiones.

a) $1 - 3t + 5t^2, -3 + 5t - 7t^2, -4 + 5t - 6t^2, 1 - t^2$

b) $5t + t^2, 1 - 8t - 2t^2, -3 + 4t + 2t^2, 2 - 3t$

32. Sea $\mathbf{p}_1(t) = 1 + t^2, \mathbf{p}_2(t) = t - 3t^2, \mathbf{p}_3(t) = 1 + t - 3t^2$.

a) Utilice vectores de coordenadas para demostrar que estos polinomios forman una base para $\mathbb{P}_2$.

b) Considere de la base $\mathcal{B} = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ para $\mathbb{P}_2$. Encuentre $\mathbf{q}$ en $\mathbb{P}_2$, considerando que $[\mathbf{q}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$.

En los ejercicios 33 y 34, determine si el conjunto de polinomios forman una base para $\mathbb{P}_3$. Justifique sus conclusiones.

33. [**M**] $3 + 7t, 5 + t - 2t^3, t - 2t^2, 1 + 16t - 6t^2 + 2t^3$

34. [**M**] $5 - 3t + 4t^2 + 2t^3, 9 + t + 8t^2 - 6t^3, 6 - 2t + 5t^2, t^3$

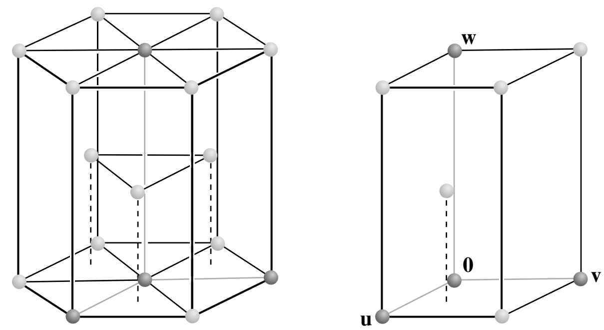
35. [**M**] Sea $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Demuestre que $\mathbf{x}$ está en H y encuentre el vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$, para

$$\mathbf{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}$$

36. [**M**] Sea $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Demuestre que $\mathcal{B}$ es una base para H y que $\mathbf{x}$ está en H, y encuentre el vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$, para

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}$$

[**M**] Los ejercicios 37 y 38 se refieren a la red cristalina del titanio, que tiene la estructura hexagonal que se ilustra a la izquierda en la figura adjunta. Los vectores $\begin{bmatrix} 2.6 \\ -1.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 4.8 \end{bmatrix}$ en $\mathbb{R}^3$ forman una base para la celda unitaria de la derecha. Los números están en unidades *angstrom* (1 Å $= 10^{-8}$ cm). En las aleaciones de titanio, algunos átomos adicionales pueden estar en la celda unitaria en los sitios *octaédricos* y *tetraédricos* (llamados así por los objetos geométricos que forman los átomos en estos lugares).



La red compacta hexagonal y su celda unitaria.

37. Uno de los sitios octaédricos es $\begin{bmatrix} 1/2 \\ 1/4 \\ 1/6 \end{bmatrix}$, respecto de la base de la red. Determine las coordenadas de este sitio en relación con la base estándar de $\mathbb{R}^3$.

38. Uno de los sitios tetraédricos es $\begin{bmatrix} 1/2 \\ 1/2 \\ 1/3 \end{bmatrix}$. Determine las coordenadas de este sitio en relación con la base estándar de $\mathbb{R}^3$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *a*) Es evidente que la matriz $P_{\mathcal{B}} = [\,\mathbf{b}_1 \;\; \mathbf{b}_2 \;\; \mathbf{b}_3\,]$ es equivalente por filas a la matriz de identidad. Por el teorema de la matriz invertible, $P_{\mathcal{B}}$ es invertible y sus columnas forman una base de $\mathbb{R}^3$.

b) Del inciso *a*), la matriz de cambio de coordenadas es $P_{\mathcal{B}} = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}$.

c) $\mathbf{x} = P_{\mathcal{B}}[\,\mathbf{x}\,]_{\mathcal{B}}$

d) Para resolver la ecuación en *c*), es probable que sea más fácil de reducir por filas una matriz aumentada que calcular $P_{\mathcal{B}}^{-1}$:

$$\underbrace{\left[\begin{array}{ccc|c} 1 & -3 & 3 & -8 \\ 0 & 4 & -6 & 2 \\ 0 & 0 & 3 & 3 \end{array}\right]}_{P_{\mathcal{B}} \qquad\quad \mathbf{x}} \sim \underbrace{\left[\begin{array}{ccc|c} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]}_{I \qquad\quad [\,\mathbf{x}\,]_{\mathcal{B}}}$$

De ahí que

$$[\,\mathbf{x}_{\mathcal{B}}\,] = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$$

2. Las coordenadas de $\mathbf{p}(t) = 6 + 3t - t^2$ respecto de $\mathcal{B}$ satisfacen

$$c_1(1 + t) + c_2(1 + t^2) + c_3(t + t^2) = 6 + 3t - t^2$$

Al igualar los coeficientes de potencias de t, se tiene que

$$\begin{aligned} c_1 + c_2 \phantom{{}+c_3} &= 6 \\ c_1 \phantom{{}+c_2} + c_3 &= 3 \\ c_2 + c_3 &= -1 \end{aligned}$$

Al resolver, se encuentra que $c_1 = 5$, $c_2 = 1$, $c_3 = -2$ y $[\,\mathbf{p}\,]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$.

4.5 LA DIMENSIÓN DE UN ESPACIO VECTORIAL

El teorema 8 de la sección 4.4 implica que un espacio vectorial V con una base B que contiene n vectores es isomorfo a $\mathbb{R}^n$. En esta sección veremos que este número n es una propiedad intrínseca (llamada dimensión) del espacio V que no depende de la elección particular de la base. El análisis de la dimensión le ayudará a comprender mejor las propiedades de las bases.

El primer teorema generaliza un resultado bien conocido acerca del espacio vectorial $\mathbb{R}^n$.

TEOREMA 9

Si un espacio vectorial V tiene una base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$, entonces cualquier conjunto en V que contenga más de n vectores debe ser linealmente dependiente.

DEMOSTRACIÓN Sea $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ un conjunto de V con más de n vectores. Los vectores de coordenadas $[\,\mathbf{u}_1\,]_{\mathcal{B}}, \ldots, [\,\mathbf{u}_p\,]_{\mathcal{B}}$ forman un conjunto linealmente dependiente en $\mathbb{R}^n$, porque hay más vectores (p) que entradas (n) en cada vector. Por lo tanto, existen escalares $c_1,\ldots, c_p$, no todos cero, tales que

$$c_1[\,\mathbf{u}_1\,]_{\mathcal{B}} + \cdots + c_p[\,\mathbf{u}_p\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{El vector cero en } \mathbb{R}^n$$

Como el mapeo de coordenadas es una transformación lineal,

$$[\,c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

El vector cero de la derecha muestra los n pesos necesarios para construir el vector $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ de los vectores básicos en $\mathcal{B}$. Es decir, $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = 0 \cdot \mathbf{b}_1 + \cdots + 0 \cdot \mathbf{b}_n = \mathbf{0}$. Puesto que no todas las c_i son cero, $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ es linealmente dependiente.[1] ■

El teorema 9 implica que si un espacio vectorial V tiene una base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$, entonces cada conjunto linealmente independiente ubicado en V no tiene más de n vectores.

[1] El teorema 9 también se aplica a los conjuntos infinitos en V. Se dice que un conjunto infinito es linealmente dependiente si algún subconjunto finito es linealmente dependiente; de lo contrario, el conjunto es linealmente independiente. Si S es un conjunto infinito de V, tome cualquier subconjunto $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ de S, con $p > n$. La demostración anterior indica que este subconjunto es linealmente dependiente y, por lo tanto, también lo es S.

TEOREMA 10

> Si un espacio vectorial V tiene una base de n vectores, entonces toda base de V debe consistir exactamente en n vectores.

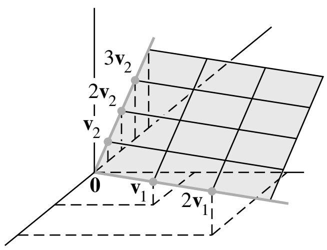
DEMOSTRACIÓN Sea $\mathcal{B}_1$ una base de n vectores y $\mathcal{B}_2$ cualquier otra base (de V). Ya que $\mathcal{B}_1$ es una base y $\mathcal{B}_2$ es linealmente independiente, $\mathcal{B}_2$ no tiene más de n vectores, de acuerdo con el teorema 9. Además, puesto que $\mathcal{B}_2$ es una base y $\mathcal{B}_1$ es linealmente independiente, $\mathcal{B}_2$ tiene al menos n vectores. Así, $\mathcal{B}_2$ se compone de exactamente n vectores. ■

Si un espacio vectorial V distinto de cero es generado por un conjunto finito S, entonces un subconjunto de S es una base para V, de acuerdo con el teorema del conjunto generador. En este caso, el teorema 10 asegura que la siguiente definición tiene sentido.

DEFINICIÓN

> Si V es generado por un conjunto finito, entonces se dice que V tiene **dimensión finita**, y la **dimensión** de V, representada como dim V, es el número de vectores en una base para V. La dimensión del espacio vectorial cero $\{\mathbf{0}\}$ se define como cero. Si V no es generado por un conjunto finito, entonces se dice que V tiene **dimensión infinita**.

EJEMPLO 1 La base estándar para $\mathbb{R}^n$ contiene n vectores, por lo que dim $\mathbb{R}^n = n$. La base polinomial estándar $\{1, t, t^2\}$ indica que dim $\mathbb{P}_2 = 3$. En general, dim $\mathbb{P}_n = n + 1$. El espacio $\mathbb{P}$ de todos los polinomios es de dimensión infinita (ejercicio 27). ■



EJEMPLO 2 Sea $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Entonces H es el plano estudiado en el ejemplo 7 de la sección 4.4. Una base para H es $\{\mathbf{v}_1, \mathbf{v}_2\}$, ya que $\mathbf{v}_1$ y $\mathbf{v}_2$ no son múltiplos y, por consiguiente, son linealmente independientes. Por lo tanto, dim $H = 2$. ■

EJEMPLO 3 Encuentre la dimensión del subespacio

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a, b, c, d \text{ en } \mathbb{R} \right\}$$

SOLUCIÓN Es fácil ver que H es el conjunto de todas las combinaciones lineales de los vectores

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Claramente, $\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2$ no es un múltiplo de $\mathbf{v}_1$, pero $\mathbf{v}_3$ es un múltiplo de $\mathbf{v}_2$. De acuerdo con el teorema del conjunto generador, podemos descartar a $\mathbf{v}_3$ y aún así tener un conjunto generador H. Finalmente, $\mathbf{v}_4$ no es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ es linealmente independiente (de acuerdo con el teorema 4 de la sección 4.3) y, en consecuencia, es una base para H. Por consiguiente, dim $H = 3$. ■

EJEMPLO 4 Los subespacios de $\mathbb{R}^3$ se pueden clasificar por dimensiones. Véase la figura 1.

Subespacios de dimensión 0. Solo el subespacio cero.

Subespacios de dimensión 1. Cualquier subespacio generado por un solo vector distinto de cero. Tales subespacios son rectas que pasan por el origen.

Subespacios de dimensión 2. Cualquier subespacio generado por dos vectores linealmente independientes. Tales subespacios son planos que pasan por el origen.

Subespacios de dimensión 3. Solo el propio $\mathbb{R}^3$. Cualesquiera tres vectores linealmente independientes en $\mathbb{R}^3$ generan todo $\mathbb{R}^3$, de acuerdo con el teorema de la matriz invertible. ■

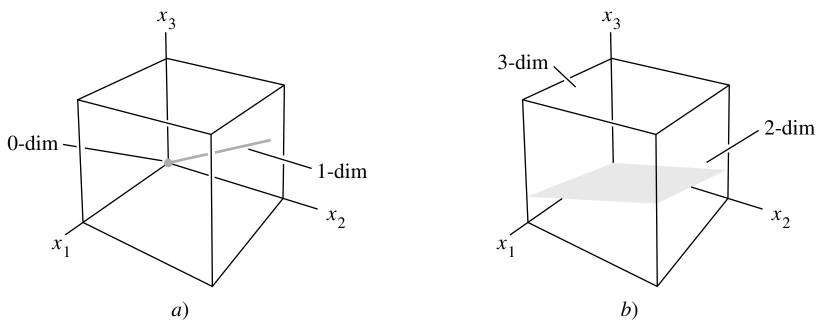


FIGURA 1 Subespacios muestra de $\mathbb{R}^3$.

Subespacios de un espacio de dimensión finita

El siguiente teorema es una contraparte natural del teorema del conjunto generador.

TEOREMA 11

Sea H un subespacio de un espacio vectorial de dimensión finita V. Cualquier conjunto linealmente independiente en H se puede expandir, si es necesario, a una base de H. Además, H tiene dimensión finita y

$$\dim H \leq \dim V$$

DEMOSTRACIÓN Si $H = \{\mathbf{0}\}$, entonces, sin duda, $\dim H = 0 \leq \dim V$. De lo contrario, sea $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ cualquier conjunto linealmente independiente en H. Si S genera a H, entonces S es una base para H. De lo contrario, existe algún $\mathbf{u}_{k+1}$ en H que no está en el generado por S. Pero entonces $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ será linealmente independiente, ya que ningún vector en el conjunto puede ser una combinación lineal de los vectores que le preceden (de acuerdo con el teorema 4).

En tanto que el nuevo conjunto no genere a H, podemos continuar con este proceso de expansión de S a un conjunto más amplio linealmente independiente en H. Sin embargo, el número de vectores en una expansión linealmente independiente de S nunca podrá superar la dimensión de V, de acuerdo con el teorema 9. Así, finalmente la expansión de S generará a H y, por lo tanto, será una base para H, y $\dim H \leq \dim V$. ■

Cuando se conoce la dimensión de un espacio o subespacio vectorial, la búsqueda de una base se simplifica con el siguiente teorema, el cual dice que si un conjunto tiene el número correcto de elementos, entonces solo se tiene que demostrar ya sea que el conjunto es linealmente independiente o que este genera el espacio. El teorema es de importancia fundamental en muchos problemas de aplicación (por ejemplo, los que implican ecuaciones diferenciales o ecuaciones en diferencias), donde la independencia lineal es mucho más fácil de comprobar que la generación.

TEOREMA 12

El teorema base

Sea V un espacio vectorial de dimensión p, donde $p \geq 1$. Cualquier conjunto linealmente independiente de exactamente p elementos en V es, de forma automática, una base para V. Cualquier conjunto de exactamente p elementos que genera a V es, de manera automática, una base para V.

DEMOSTRACIÓN De acuerdo con el teorema 11, un conjunto linealmente independiente S de p elementos se puede ampliar a una base para V. Sin embargo, esa base debe contener exactamente p elementos, ya que dim $V = p$. Así que S debe ser ya una base para V. Ahora suponga que S tiene p elementos y genera a V. Puesto que V es diferente de cero, el teorema del conjunto generador implica que un subconjunto S' de S es una base de V. Como dim $V = p$, S' debe contener p vectores. Por lo tanto, $S = S'$. ■

Las dimensiones de Nul A y Col A

Como las columnas pivote de una matriz A forman una base para Col A, conocemos la dimensión de Col A tan pronto como se conocen las columnas pivote. Tal vez parezca que la dimensión de Nul A requiere de más trabajo, ya que encontrar una base para Nul A, por lo general, toma más tiempo que una base para Col A. Pero, ¡hay un atajo!

Sea A una matriz de $m \times n$, y supongamos que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene k variables libres. De la sección 4.2, sabemos que el método estándar para encontrar un conjunto generador para Nul A producirá exactamente k vectores linealmente independientes, por ejemplo, $\mathbf{u}_1, \ldots, \mathbf{u}_k$, uno para cada variable libre. Por lo tanto, $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es una base para Nul A, y el número de variables libres determina el tamaño de la base. Hagamos un resumen de estos hechos para referencia futura.

> La dimensión de Nul A es el número de variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$, y la dimensión de Col A es el número de columnas pivote de A.

EJEMPLO 5 Determine las dimensiones del espacio nulo y el espacio columna de

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

SOLUCIÓN Reduzca por filas la matriz aumentada $[A \quad \mathbf{0}]$ a la forma escalonada:

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

Hay tres variables libres: x_2, x_4 y x_5. Por consiguiente, la dimensión de Nul A es 3. Por otra parte, dim Col $A = 2$ porque A tiene dos columnas pivote. ■

PROBLEMAS DE PRÁCTICA

Determine si cada enunciado es verdadero o falso, y dé una razón para cada respuesta. Aquí, V es un espacio vectorial de dimensión finita diferente de cero.

1. Si dim $V = p$, y si S es un subconjunto linealmente dependiente de V, entonces S contiene más que p vectores.
2. Si S genera a V y si T es un subconjunto de V que contiene más vectores que S, entonces T es linealmente dependiente.

4.5 EJERCICIOS

Para cada subespacio en los ejercicios 1 a 8, *a*) encuentre una base para el subespacio, y *b*) indique la dimensión.

1. $\left\{\begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s, t \text{ en } \mathbb{R}\right\}$

2. $\left\{\begin{bmatrix} 2a \\ -4b \\ -2a \end{bmatrix} : a, b \text{ en } \mathbb{R}\right\}$

3. $\left\{\begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a, b, c \text{ en } \mathbb{R}\right\}$

4. $\left\{\begin{bmatrix} p+2q \\ -p \\ 3p-q \\ p+q \end{bmatrix} : p, q \text{ en } \mathbb{R}\right\}$

5. $\left\{\begin{bmatrix} p-2q \\ 2p+5r \\ -2q+2r \\ -3p+6r \end{bmatrix} : p, q, r \text{ en } \mathbb{R}\right\}$

6. $\left\{\begin{bmatrix} 3a-c \\ -b-3c \\ -7a+6b+5c \\ -3a+c \end{bmatrix} : a, b, c \text{ en } \mathbb{R}\right\}$

7. $\{(a,b,c) : a-3b+c=0, b-2c=0, 2b-c=0\}$

8. $\{(a,b,c,d) : a-3b+c=0\}$

9. Encuentre la dimensión del subespacio de todos los vectores en $\mathbb{R}^3$ cuyas entradas primera y tercera sean iguales.

10. Encuentre la dimensión del subespacio H de $\mathbb{R}^2$ generado por

$$\begin{bmatrix} 1 \\ -5 \end{bmatrix}, \begin{bmatrix} -2 \\ 10 \end{bmatrix}, \begin{bmatrix} -3 \\ 15 \end{bmatrix}$$

En los ejercicios 11 y 12, encuentre la dimensión del subespacio generado por los vectores dados.

11. $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$

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

Determine las dimensiones de Nul A y Col A de las matrices que se muestran en los ejercicios 13 a 18.

13. $A = \begin{bmatrix} 1 & -6 & 9 & 0 & -2 \\ 0 & 1 & 2 & -4 & 5 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 2 & -4 & 3 & -2 & 6 & 0 \\ 0 & 0 & 0 & 1 & 0 & -3 & 7 \\ 0 & 0 & 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

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

16. $A = \begin{bmatrix} 3 & 2 \\ -6 & 5 \end{bmatrix}$

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

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

En los ejercicios 19 y 20, V es un espacio vectorial. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

19. *a*) El número de columnas pivote de una matriz es igual a la dimensión de su espacio columna.

b) Un plano en $\mathbb{R}^3$ es un subespacio de dimensión 2 de $\mathbb{R}^3$.

c) La dimensión del espacio vectorial $\mathbb{P}_4$ es 4.

d) Si dim $V = n$ y S es un conjunto linealmente independiente en V, entonces S es una base para V.

e) Si un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ genera un espacio vectorial V de dimensión finita y si T es un conjunto de más de p vectores en V, entonces T es linealmente dependiente.

20. *a*) $\mathbb{R}^2$ es un subespacio de dimensión 2 de $\mathbb{R}^3$.

b) El número de variables en la ecuación $A\mathbf{x} = \mathbf{0}$ es igual a la dimensión de Nul A.

c) Un espacio vectorial es de dimensión infinita si es generado por un conjunto infinito.

d) Si dim $V = n$, y si S genera a V, entonces S es una base de V.

e) El único subespacio de dimensión 3 de $\mathbb{R}^3$ es el propio $\mathbb{R}^3$.

21. Los primeros cuatro polinomios de Hermite son $1, 2t, -2+4t^2$, y $-12t+8t^3$. Estos polinomios surgen de forma natural en el estudio de ciertas ecuaciones diferenciales importantes en física matemática.[2] Demuestre que los primeros cuatro polinomios de Hermite forman una base de $\mathbb{P}_3$.

22. Los primeros cuatro polinomios de Laguerre son 1, $1-t$, $2-4t+t^2$, y $6-18t+9t^2-t^3$. Demuestre que estos polinomios forman una base de $\mathbb{P}_3$.

23. Sea B la base de $\mathbb{P}_3$ que consta de los polinomios de Hermite en el ejercicio 21, y sea $\mathbf{p}(t) = -1+8t^2+8t^3$. Encuentre el vector de coordenadas de $\mathbf{p}$ respecto de $\mathcal{B}$.

24. Sea $\mathcal{B}$ la base de $\mathbb{P}_2$ que consiste en los tres primeros polinomios de Laguerre listados en el ejercicio 22, y sea $\mathbf{p}(t) = 5+5t-2t^2$. Encuentre el vector de coordenadas de $\mathbf{p}$ respecto de $\mathcal{B}$.

25. Sea S un subconjunto de un espacio vectorial V de dimensión n, y suponga que S contiene menos de n vectores. Explique por qué S no puede generar a V.

26. Sea H un subespacio de dimensión n de un espacio vectorial V de dimensión n. Demuestre que $H = V$.

27. Explique por qué el espacio $\mathbb{P}$ de todos los polinomios es un espacio de dimensión infinita.

[2] Véase *Introduction to Functional Analysis*, 2a. edición, por A. E. Taylor y David C. Lay (Nueva York: John Wiley & Sons, 1980), pp. 92-93. También se analizan otros conjuntos de polinomios.

28. Demuestre que el espacio $C(\mathbb{R})$ de todas las funciones continuas definidas en la recta real es un espacio de dimensión infinita.

En los ejercicios 29 y 30, V es un espacio vectorial de dimensión finita diferente de cero, y los vectores mencionados pertenecen a V. Marque cada enunciado como verdadero o falso. Justifique sus respuestas. (Estas preguntas tienen mayor grado de dificultad que las de los ejercicios 19 y 20).

29. *a*) Si existe un conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ que genera a V, entonces $\dim V \leq p$.

b) Si existe un conjunto linealmente independiente $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en V, entonces $\dim V \geq p$.

c) Si $\dim V = p$, entonces existe un conjunto generador de $p + 1$ vectores en V.

30. *a*) Si existe un conjunto linealmente dependiente $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en V, entonces $\dim V \leq p$.

b) Si cada conjunto de p elementos en V no genera a V, entonces $\dim V > p$.

c) Si $p \geq 2$ y $\dim V = p$, entonces cada conjunto de $p - 1$ vectores distintos de cero es linealmente independiente.

Los ejercicios 31 y 32 se refieren a espacios vectoriales V y W de dimensión finita y a una transformación lineal $T : V \to W$.

31. Sea H un subespacio distinto de cero de V, y sea $T(H)$ el conjunto de imágenes de vectores en H. Entonces, $T(H)$ es un subespacio de W, de acuerdo con el ejercicio 35 en la sección 4.2. Demuestre que $\dim T(H) \leq \dim H$.

32. Sea H un subespacio distinto de cero de V, y suponga que T es un mapeo uno a uno (lineal) de V en W. Demuestre que $\dim T(H) = \dim H$. Si resulta que T es un mapeo uno a uno de V *sobre* W, entonces $\dim V = \dim W$. Espacios vectoriales isomorfos de dimensión finita tienen la misma dimensión.

33. [**M**] De acuerdo con el teorema 11, un conjunto linealmente independiente $\{\mathbf{v}_1,\ldots, \mathbf{v}_k\}$ en $\mathbb{R}^n$ se puede expandir a una base para $\mathbb{R}^n$. Una manera de hacer esto es crear $A = [\mathbf{v}_1 \cdots \mathbf{v}_k \;\; \mathbf{e}_1 \cdots \mathbf{e}_n]$, siendo $\mathbf{e}_1,\ldots, \mathbf{e}_n$ las columnas de la matriz identidad; las columnas pivote de A forman una base para $\mathbb{R}^n$.

a) Utilice el método descrito para ampliar los siguientes vectores a una base para $\mathbb{R}^5$:

$$\mathbf{v}_1 = \begin{bmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{bmatrix}$$

b) Explique por qué funciona en general el método: ¿por qué están los vectores originales $\mathbf{v}_1,\ldots, \mathbf{v}_k$ incluidos en la base encontrada para Col A? ¿Por qué es Col $A = \mathbb{R}^n$?

34. [**M**] Sea $\mathcal{B} = \{1, \cos t, \cos^2 t,\ldots, \cos^6 t\}$ y $\mathcal{C} = \{1, \cos t, \cos 2t,\ldots, \cos 6t\}$. Suponga las siguientes identidades trigonométricas (véase el ejercicio 37 de la sección 4.1).

$$\begin{aligned} \cos 2t &= -1 + 2\cos^2 t \\ \cos 3t &= -3\cos t + 4\cos^3 t \\ \cos 4t &= 1 - 8\cos^2 t + 8\cos^4 t \\ \cos 5t &= 5\cos t - 20\cos^3 t + 16\cos^5 t \\ \cos 6t &= -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t \end{aligned}$$

Sea H el subespacio de funciones generado por las funciones en $\mathcal{B}$. Entonces $\mathcal{B}$ es una base para H, de acuerdo con el ejercicio 38 de la sección 4.3.

a) Escriba los vectores de $\mathcal{B}$-coordenadas de los vectores en $\mathcal{C}$, y utilícelos para demostrar que $\mathcal{C}$ es un conjunto linealmente independiente en H.

b) Explique por qué $\mathcal{C}$ es una base para H.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Falso. Considere el conjunto $\{\mathbf{0}\}$.
2. Verdadero. De acuerdo con el teorema del conjunto generador, S contiene una base para V; llamémosla base S'. Así, T contendrá más vectores que S'. De acuerdo con el teorema 9, T es linealmente dependiente.

4.6 RANGO

Con la ayuda de los conceptos de espacio vectorial, esta sección ofrece una perspectiva desde *el interior* de una matriz y revela varias relaciones interesantes y útiles, ocultas en sus filas y columnas.

Por ejemplo, imagine que se colocan 2000 números aleatorios en una matriz A de 40×50 y después se determina el número máximo de columnas linealmente independientes de A y el número máximo de columnas linealmente independientes de A^T (filas de A). De manera sorprendente, los dos números son iguales. Como pronto veremos, su valor común es el *rango* de la matriz. Para explicar por qué, necesitamos examinar el subespacio generado por las filas de A.

El espacio fila

Si A es una matriz de $m \times n$, cada fila de A tiene n entradas y, por lo tanto, se puede identificar con un vector en $\mathbb{R}^n$. El conjunto de todas las combinaciones lineales de los vectores fila se denomina **espacio fila** de A y se denota como Fila A. Cada fila tiene n entradas, por lo que Fila A es un subespacio de $\mathbb{R}^n$. Ya que las filas de A se identifican con las columnas de A^T, también podríamos escribir Col A^T en lugar de Fila A.

EJEMPLO 1 Sea

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \quad \text{y} \quad \begin{aligned} \mathbf{r}_1 &= (-2, -5, 8, 0, -17) \\ \mathbf{r}_2 &= (1, 3, -5, 1, 5) \\ \mathbf{r}_3 &= (3, 11, -19, 7, 1) \\ \mathbf{r}_4 &= (1, 7, -13, 5, -3) \end{aligned}$$

El espacio fila de A es el subespacio de $\mathbb{R}^5$ generado por $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. Es decir, Fila $A = \text{Gen}\,\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. Es natural representar vectores fila de forma horizontal; sin embargo, también es posible representarlos como vectores columna si resulta más conveniente. ■

Si supiéramos algo de las relaciones de dependencia lineal entre las filas de la matriz A del ejemplo 1, podríamos usar el teorema del conjunto generador para reducir el tamaño del conjunto generador a una base. Por desgracia, las operaciones de fila en A no nos dan esa información, porque las operaciones de fila cambian las relaciones de dependencia de filas. Pero la reducción de filas de A, ¡sin duda vale la pena, como muestra el siguiente teorema!

TEOREMA 13

> Si dos matrices A y B son equivalentes por filas, entonces sus espacios fila son iguales. Si B está en forma escalonada, las filas de B diferentes de cero forman una base para el espacio fila de A, así como para el de B.

DEMOSTRACIÓN Si B se obtiene a partir de A mediante operaciones de fila, las filas de B son combinaciones lineales de las filas de A. De ello se desprende que cualquier combinación lineal de las filas de B es automáticamente una combinación lineal de las filas de A. Así, el espacio fila de B está contenido en el espacio fila de A. Ya que las operaciones de fila son reversibles, el mismo argumento indica que el espacio fila de A es un subconjunto del espacio fila de B. De manera que los dos espacios fila son iguales. Si B está en forma escalonada, sus filas diferentes de cero son linealmente independientes porque ninguna fila diferente de cero es una combinación lineal de las filas distintas de cero debajo de esta. (Aplique el teorema 4 para las filas diferentes de cero de B en orden inverso, con la primera fila como la última). Así, las filas diferentes de cero de B forman una base del espacio fila (común) de B y A. ■

El resultado principal de esta sección implica los tres espacios: Fila A, Col A y Nul A. El siguiente ejemplo prepara el camino para este resultado y muestra cómo *una* secuencia de operaciones de fila de A conduce a las bases para los tres espacios.

EJEMPLO 2 Encuentre bases para el espacio fila, el espacio columna y el espacio nulo de la matriz

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

SOLUCIÓN Para encontrar las bases para el espacio fila y el espacio columna, reduzca A por filas a una forma escalonada:

$$A \sim B = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

De acuerdo con el teorema 13, las tres primeras filas de B forman una base para el espacio fila de A (así como para el espacio fila de B). Así,

$$\text{Base para Fila } A\text{: } \{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$$

Para el espacio columna, observe a partir de B que los pivotes están en las columnas 1, 2 y 4. Por lo tanto, las columnas 1, 2 y 4 de A (no de B) forman una base para Col A:

$$\text{Base para Col } A\text{: } \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

Considere que cualquier forma escalonada de A proporciona (en sus filas diferentes de cero) una base para Fila A y también identifica las columnas pivote de A para Col A. Sin embargo, para Nul A, se necesita la *forma escalonada reducida*. Otras operaciones de fila sobre B dan como resultado

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

La ecuación $A\mathbf{x} = \mathbf{0}$ es equivalente a $C\mathbf{x} = \mathbf{0}$, es decir,

$$\begin{aligned} x_1 + \quad x_3 \quad + \ x_5 &= 0 \\ x_2 - 2x_3 \quad + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

Así, $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, con x_3 y x_5 como variables libres. Los cálculos usuales (que se analizan en la sección 4.2) demuestran que

$$\text{Base para Nul } A\text{: } \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Observe que, a diferencia de la base para Col A, las bases para Fila A y Nul A no tienen una relación sencilla con las propias entradas de A.[1] ■

---

[1] Es posible encontrar una base para Fila A que utiliza filas de A. En primer lugar se determina A^T, y luego se reduce por filas hasta que se encuentren las columnas pivote de A^T. Estas columnas pivote de A^T son filas de A, y forman una base para el espacio fila de A.

Advertencia: A pesar de que las tres primeras filas de B en el ejemplo 2 son linealmente independientes, es erróneo concluir que las tres primeras filas de A son linealmente independientes. (De hecho, la tercera fila de A es la primera fila multiplicada por 2, más la segunda fila multiplicada por 7). Las operaciones de fila pueden cambiar las relaciones de dependencia lineal entre las *filas* de una matriz.

El teorema del rango

WEB

El siguiente teorema describe las relaciones fundamentales entre las dimensiones de Col A, Fila A y Nul A.

DEFINICIÓN

El **rango** de A es la dimensión del espacio columna de A.

Puesto que Fila A es igual que Col A^T, la dimensión del espacio fila de A es el rango de A^T. La dimensión del espacio nulo a veces se llama la **nulidad** de A, aunque no utilizaremos este término.

Tal vez un lector atento ya haya descubierto la totalidad o parte del siguiente teorema, mientras trabajaba con los ejercicios de la sección 4.5 o al leer el ejemplo 2 anterior.

TEOREMA 14

El teorema del rango

Las dimensiones del espacio columna y del espacio fila de una matriz A de $m \times n$ son iguales. Esta dimensión común, el rango de A, también es igual al número de posiciones pivote en A y satisface la ecuación

$$\text{rango } A + \dim \text{Nul } A = n$$

DEMOSTRACIÓN De acuerdo con el teorema 6 de la sección 4.3, rango A es el número de columnas pivote de A. De manera equivalente, rango A es el número de posiciones pivote en una forma escalonada B de A. Además, puesto que B tiene una fila diferente de cero para cada pivote, y como estas filas forman una base para el espacio fila de A, el rango de A también es la dimensión del espacio fila.

A partir de la sección 4.5, la dimensión de Nul A es igual al número de variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$. Dicho de otra manera, la dimensión de Nul A es el número de columnas de A que *no* son columnas pivote. (Es el número de estas columnas, no las columnas mismas, lo que se relaciona con Nul A). Como es evidente,

$$\left\{\begin{matrix}\text{número de}\\ \text{columnas pivote}\end{matrix}\right\} + \left\{\begin{matrix}\text{número de columnas}\\ \text{que no son pivote}\end{matrix}\right\} = \left\{\begin{matrix}\text{número de}\\ \text{columnas}\end{matrix}\right\}$$

Esto demuestra el teorema. ■

Las ideas que sustentan el teorema 14 se pueden distinguir en los cálculos del ejemplo 2. Las tres posiciones pivote en la forma escalonada B determinan las variables básicas e identifican los vectores básicos para Col A y Fila A.

EJEMPLO 3

a) Si A es una matriz de 7×9 con un espacio nulo de dimensión 2, ¿cuál es el rango de A?

b) ¿Podría una matriz de 6×9 tener un espacio nulo de dimensión 2?

SOLUCIÓN

a) Puesto que A tiene 9 columnas, (rango A) + 2 = 9; por lo tanto, rango $A = 7$.

b) No. Si una matriz de 6×9, llamémosla B, tuviera un espacio nulo de dimensión 2, tendría rango 7, de acuerdo con el teorema del rango. Sin embargo, las columnas de B son vectores en $\mathbb{R}^6$, de manera que la dimensión de Col B no puede ser superior a 6; es decir, el rango de B no puede ser mayor que 6. ■

El siguiente ejemplo proporciona una buena forma de visualizar los subespacios que hemos estudiado. En el capítulo 6 se verá que Fila A y Nul A tienen solo el vector cero en común y, en realidad, son "perpendiculares" entre sí. Este mismo hecho se aplicará a Fila A^T (= Col A) y Nul A^T. Por lo tanto, la figura 1, que acompaña al ejemplo 4, crea una buena imagen mental para el caso general. (El valor de estudiar A^T junto con A se demuestra en el ejercicio 29).

EJEMPLO 4 Sea $A = \begin{bmatrix} 3 & 0 & -1 \\ 3 & 0 & -1 \\ 4 & 0 & 5 \end{bmatrix}$. Se comprueba rápidamente que Nul A es el eje x_2, Fila A es el plano x_1x_3, Col A es el plano cuya ecuación es $x_1 - x_2 = 0$, y Nul A^T es el conjunto de todos los múltiplos de $(1, -1, 0)$. La figura 1 muestra Nul A y Fila A en el dominio de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$; el rango de este mapeo, Col A, se muestra en una copia separada de $\mathbb{R}^3$, junto con Nul A^T. ■

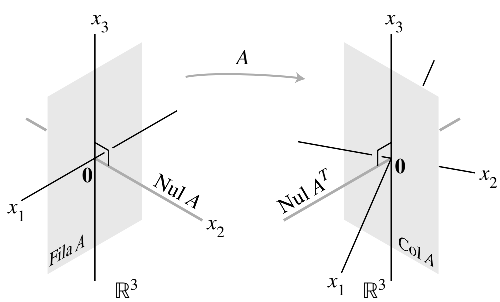


FIGURA 1 Subespacios determinados por una matriz A.

Aplicaciones para sistemas de ecuaciones

El teorema del rango es una poderosa herramienta para el procesamiento de información sobre los sistemas de ecuaciones lineales. El siguiente ejemplo simula la forma como se plantearía un problema de la vida real utilizando ecuaciones lineales, sin mencionar explícitamente términos de álgebra lineal, como matriz, subespacio y dimensión.

EJEMPLO 5 Un científico encontró dos soluciones para un sistema homogéneo de 40 ecuaciones con 42 variables. No son múltiplos las dos soluciones, y todas las demás soluciones se pueden desarrollar sumando múltiplos adecuados de estas dos soluciones. ¿Puede el científico estar *seguro* de que un sistema no homogéneo asociado (con los mismos coeficientes) tiene una solución?

SOLUCIÓN Sí. Sea A una matriz de coeficientes de 40×42 del sistema. La información dada implica que las dos soluciones son linealmente independientes y generan Nul A. Así que dim Nul $A = 2$. De acuerdo con el teorema del rango, dim Col $A = 42 - 2 = 40$. Como $\mathbb{R}^{40}$ es el único subespacio de $\mathbb{R}^{40}$ cuya dimensión es 40, Col A debe ser todo de $\mathbb{R}^{40}$. Esto significa que cada ecuación no homogénea $A\mathbf{x} = \mathbf{b}$ tiene una solución. ■

El rango y el teorema de la matriz invertible

Los distintos conceptos de espacios vectoriales asociados a una matriz proporcionan varios enunciados adicionales al teorema de la matriz invertible. Los nuevos enunciados que se listan a continuación se deducen del teorema de la matriz invertible original de la sección 2.3.

TEOREMA

El teorema de la matriz invertible (continuación)

Sea A una matriz de $n \times n$. Cada uno de los siguientes enunciados es equivalente a la afirmación de que A es una matriz invertible.

m) Las columnas de A forman una base de $\mathbb{R}^n$.

n) $\text{Col } A = \mathbb{R}^n$

o) $\dim \text{Col } A = n$

p) $\text{rango } A = n$

q) $\text{Nul } A = \{\mathbf{0}\}$

r) $\dim \text{Nul } A = 0$

DEMOSTRACIÓN El enunciado *m*) es lógicamente equivalente a los enunciados *e*) y *h*) en relación con la independencia lineal y la generación. Los otros cinco enunciados están vinculados a los anteriores del teorema mediante la siguiente cadena de implicaciones casi triviales:

$$g) \Rightarrow n) \Rightarrow o) \Rightarrow p) \Rightarrow r) \Rightarrow q) \Rightarrow d)$$

El enunciado *g*), el cual dice que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para cada $\mathbf{b}$ en $\mathbb{R}^n$, implica a *n*), porque Col A es precisamente el conjunto de todas las $\mathbf{b}$ tales que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente. Las implicaciones $n) \Rightarrow o) \Rightarrow p)$ se deducen de las definiciones de dimensión y rango. Si el rango de A es n, el número de columnas de A, entonces $\dim \text{Nul } A = 0$, de acuerdo con el teorema del rango, y $\text{Nul } A = \{\mathbf{0}\}$. Por lo tanto, $p) \Rightarrow r) \Rightarrow q)$. Además, *q*) implica que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial, que es el enunciado *d*). Como ya se sabe que los enunciados *d*) y *g*) son equivalentes al enunciado de que A es invertible, la demostración está completa. ■

Nos hemos abstenido de agregar al teorema de la matriz invertible enunciados evidentes acerca del espacio fila de A, ya que el espacio fila es el espacio columna de A^T. Recuerde el enunciado (1) del teorema de la matriz invertible, el cual afirma que A es invertible si y solo si A^T es invertible. Por lo tanto, cada enunciado en el teorema de la matriz invertible también se puede establecer respecto de A^T. Pero hacer esto duplicaría la longitud del teorema y se obtendría una lista ¡de más de 30 enunciados!

NOTA NUMÉRICA

Muchos algoritmos analizados en este libro son útiles para comprender conceptos y efectuar cálculos sencillos a mano. Sin embargo, los algoritmos a menudo son inadecuados para los grandes problemas que surgen en la vida real.

La determinación de un rango es un buen ejemplo. Tal vez parezca fácil reducir una matriz a la forma escalonada y contar los pivotes. Pero, a menos que se realice aritmética exacta en una matriz cuyas entradas se especifiquen con exactitud, las operaciones de fila pueden cambiar el rango aparente de una matriz. Por ejemplo, si el valor de x en la matriz $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ no se almacena exactamente como 7 en una computadora, entonces el rango podría ser 1 o 2, dependiendo de si la computadora trata o no a $x - 7$ como cero.

En las aplicaciones prácticas, el rango efectivo de una matriz A con frecuencia se determina a partir de la descomposición en valores singulares de A, que se analizará en la sección 7.4. Esta descomposición también es una fuente confiable de bases para Col A, Fila A, Nul A y Nul A^T.

WEB

PROBLEMAS DE PRÁCTICA

Las matrices que se muestran a continuación son equivalentes por filas.

$$A = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. Determine rango A y dim Nul A.
2. Encuentre bases para Col A y Fila A.
3. ¿Cuál es el siguiente paso a realizar con la finalidad de encontrar una base para Nul A?
4. ¿Cuántas columnas pivote están en una forma escalonada por filas de A^T?

4.6 EJERCICIOS

En los ejercicios 1 a 4, suponga que la matriz A es equivalente por filas a B. Sin hacer cálculos, liste rango de A y dim Nul A. Luego, encuentre las bases para Col A, Fila A y Nul A.

1. $A = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

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

3. $A = \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ -2 & -3 & 6 & -3 & 0 & -6 \\ 4 & 9 & -12 & 9 & 3 & 12 \\ -2 & 3 & 6 & 3 & 3 & -6 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ 0 & 3 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 & -2 & 0 & 1 & -2 \\ 1 & 2 & -3 & 0 & -2 & -3 \\ 1 & -1 & 0 & 0 & 1 & 6 \\ 1 & -2 & 2 & 1 & -3 & 0 \\ 1 & -2 & 1 & 0 & 2 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & -2 & 0 & 1 & -2 \\ 0 & 1 & -1 & 0 & -3 & -1 \\ 0 & 0 & 1 & 1 & -13 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

5. Si una matriz A de 4×7 tiene rango 3, determine dim Nul A, dim Fila A y el rango de A^T.

6. Si una matriz A de 7×5 tiene rango 2, determine dim Nul A, dim Fila A y el rango de A^T.

7. Suponga que una matriz A de 4×7 tiene cuatro columnas pivote. ¿Es Col $A = \mathbb{R}^4$? ¿Es Nul $A = \mathbb{R}^3$? Explique sus respuestas.

8. Supongamos que una matriz A de 6×8 tiene cuatro columnas pivote. ¿Cuál es la dim Nul A? ¿Es Col $A = \mathbb{R}^4$? ¿Por qué?

9. Si el espacio nulo de una matriz A de 4×6 es de dimensión 3, ¿cuál es la dimensión del espacio columna de A? ¿Es Col $A = \mathbb{R}^3$? ¿Por qué?

10. Si el espacio nulo de una matriz A de 8×7 es de dimensión 5, ¿cuál es la dimensión del espacio columna de A?

11. Si el espacio nulo de una matriz A de 8×5 es de dimensión 3, ¿cuál es la dimensión del espacio fila de A?

12. Si el espacio nulo de una matriz A de 5×4 es de dimensión 2, ¿cuál es la dimensión del espacio fila de A?

13. Si A es una matriz de 7×5, ¿cuál es el mayor rango posible de A? Si A es una matriz de 5×7, ¿cuál es el mayor rango posible de A? Explique sus respuestas.

14. Si A es una matriz de 5×4, ¿cuál es la mayor dimensión posible del espacio fila de A? Si A es una matriz de 4×5, ¿cuál es la máxima dimensión posible del espacio fila de A? Explique sus respuestas.

15. Si A es una matriz de 3×7, ¿cuál es la menor dimensión posible de Nul A?

16. Si A es una matriz de 7×5, ¿cuál es la menor dimensión posible de Nul A?

En los ejercicios 17 y 18, A es una matriz de $m \times n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

17. *a*) El espacio fila de A es el mismo que el espacio columna de A^T.

b) Si B es cualquier forma escalonada de A, y si B tiene tres filas diferentes de cero, entonces las tres primeras filas de A forman una base para Fila A.

c) Las dimensiones del espacio fila y el espacio columna de A son iguales, incluso si A no es cuadrada.

d) La suma de las dimensiones del espacio fila y el espacio nulo de A es igual al número de filas en A.

e) En una computadora, las operaciones de fila pueden modificar el rango aparente de una matriz.

18. *a*) Si B es cualquier forma escalonada de A, entonces las columnas pivote de B forman una base para el espacio columna de A.

b) Las operaciones de fila preservan las relaciones de dependencia lineal entre las filas de A.

c) La dimensión del espacio nulo de A es el número de columnas de A que *no* son columnas pivote.

d) El espacio fila de A^T es igual que el espacio columna de A.

e) Si A y B son equivalentes por filas, entonces sus espacios fila son iguales.

19. Suponga que todas las soluciones de un sistema homogéneo de cinco ecuaciones lineales con seis incógnitas son múltiplos de una solución diferente de cero. ¿El sistema necesariamente tendrá una solución para cada posible elección de las constantes en el lado derecho de las ecuaciones? Explique su respuesta.

20. Suponga que un sistema no homogéneo de seis ecuaciones lineales con ocho incógnitas tiene una solución, con dos variables libres. ¿Es posible cambiar algunas constantes en los miembros derechos de las ecuaciones para hacer que el nuevo sistema sea inconsistente? Explique su respuesta.

21. Suponga que un sistema no homogéneo de nueve ecuaciones lineales con 10 incógnitas tiene una solución para todas las posibles constantes de los miembros derechos de las ecuaciones. ¿Es posible encontrar dos soluciones diferentes de cero del sistema homogéneo asociado que *no* sean múltiplos una de la otra? Analice.

22. ¿Es posible que todas las soluciones de un sistema homogéneo de 10 ecuaciones lineales con 12 variables sean múltiplos de una solución fija diferente de cero? Analice.

23. Un sistema homogéneo de 12 ecuaciones lineales con ocho incógnitas tiene dos soluciones fijas que no son múltiplos una de la otra, y todas las demás soluciones son combinaciones lineales de estas dos soluciones. ¿Puede describirse el conjunto de todas las soluciones con menos de 12 ecuaciones lineales homogéneas? Si es así, ¿con cuántas? Analice.

24. ¿Es posible que un sistema no homogéneo de siete ecuaciones con seis incógnitas tenga una solución única para algún conjunto de constantes del miembro derecho? ¿Es posible que este sistema tenga una solución única para cada miembro derecho? Explique sus respuestas.

25. Un científico resuelve un sistema no homogéneo de 10 ecuaciones lineales con 12 incógnitas y encuentra que tres de las incógnitas son variables libres. ¿Puede el científico estar seguro de que, si se cambia el lado derecho de las ecuaciones, el nuevo sistema no homogéneo tendrá una solución? Analice.

26. En teoría estadística, un requisito común es que una matriz sea de *rango completo*. Es decir, el rango debe ser tan grande como sea posible. Explique por qué una matriz de $m \times n$ con más filas que columnas tiene rango completo si y solo si sus columnas son linealmente independientes.

Los ejercicios 27 a 29 se refieren a una matriz A de $m \times n$ y a lo que con frecuencia se denomina los *subespacios fundamentales* determinados por A.

27. ¿Cuál de los subespacios Fila A, Col A, Nul A, Fila A^T, Col A^T y Nul A^T están en $\mathbb{R}^m$ y cuáles están en $\mathbb{R}^n$? ¿Cuántos subespacios distintos están en esta lista?

28. Justifique las siguientes igualdades:

a) dim Fila A + dim Nul $A = n$ Número de columnas de A

b) dim Col A + dim Nul $A^T = m$ Número de filas de A

29. Con base en el ejercicio 28, explique por qué la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ en $\mathbb{R}^m$ si y solo si la ecuación $A^T\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.

30. Suponga que A es $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$. ¿Qué tiene que ser verdad acerca del rango de $[A \quad \mathbf{b}]$ y el rango A para que la ecuación $A\mathbf{x} = \mathbf{b}$ sea coherente?

Las matrices con rango 1 son importantes en algunos algoritmos de computadora y varios contextos teóricos, incluyendo la descomposición en valores singulares del capítulo 7. Es posible demostrar que una matriz A de $m \times n$ tiene rango 1 si y solo si se trata de un producto externo, es decir, $A = \mathbf{u}\mathbf{v}^T$ por alguna $\mathbf{u}$ en $\mathbb{R}^m$ y $\mathbf{v}$ en $\mathbb{R}^n$. Los ejercicios 31 a 33 sugieren por qué esta propiedad es verdadera.

31. Compruebe que el rango de $\mathbf{u}\mathbf{v}^T \leq 1$ si $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

32. Sea $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Encuentre $\mathbf{v}$ en $\mathbb{R}^3$ tal que $\begin{bmatrix} 1 & -3 & 4 \\ 2 & -6 & 8 \end{bmatrix} = \mathbf{u}\mathbf{v}^T$.

33. Sea A cualquier matriz de 2×3 tal que rango $A = 1$. Sea $\mathbf{u}$ la primera columna de A, y suponga que $\mathbf{u} \neq \mathbf{0}$. Explique por qué hay un vector $\mathbf{v}$ en $\mathbb{R}^3$ tal que $A = \mathbf{u}\mathbf{v}^T$. ¿Cómo podría modificarse esta construcción si la primera columna de A fuera cero?

34. Sea A una matriz de $m \times n$ de rango $r > 0$, y sea U una forma escalonada de A. Explique por qué existe una matriz invertible E tal que $A = EU$, y utilice esta factorización para escribir A como la suma de r matrices con rango 1. [*Sugerencia*: Véase el teorema 10 de la sección 2.4].

35. [**M**] Sea $A = \begin{bmatrix} 7 & -9 & -4 & 5 & 3 & -3 & -7 \\ -4 & 6 & 7 & -2 & -6 & -5 & 5 \\ 5 & -7 & -6 & 5 & -6 & 2 & 8 \\ -3 & 5 & 8 & -1 & -7 & -4 & 8 \\ 6 & -8 & -5 & 4 & 4 & 9 & 3 \end{bmatrix}$.

a) Construya matrices C y N, cuyas columnas son las bases para Col A y Nul A, respectivamente, y construya una matriz R cuyas *filas* formen una base para Fila A.

b) Construya una matriz M cuyas columnas formen una base para Nul A^T, construya las matrices $S = [R^T \quad N]$ y $T = [C \quad M]$, y explique por qué S y T deberían ser cuadradas. Compruebe que tanto S como T son invertibles.

36. [**M**] Repita el ejercicio 35 para una matriz aleatoria A de 6×7 con valores enteros y rango de 4 o menor. Una manera de construir A es crear una matriz aleatoria J de 6×4 con valores enteros y una matriz aleatoria K de 4×7 con valores enteros, y establecer que $A = JK$. (Véase el ejercicio complementario 12 al final del capítulo; véase también la *Guía de estudio* para los programas de generación de matrices).

37. [**M**] Sea A la matriz del ejercicio 35. Construya una matriz C cuyas columnas sean las columnas pivote de A, y construya una matriz R cuyas filas sean las filas diferentes de cero de la forma escalonada reducida de A. Calcule CR, y analice sus hallazgos.

38. [**M**] Repita el ejercicio 37 para tres matrices aleatorias A de 5×7 con valores enteros y rangos de 5, 4 y 3. Haga una conjetura acerca de cómo se relaciona CR con A para cualquier matriz A. Demuestre su suposición.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. A tiene dos columnas pivote, por lo que rango $A = 2$. Puesto que A tiene 5 columnas en total, dim Nul $A = 5 - 2 = 3$.
2. Las columnas pivote de A son las dos primeras columnas. Por lo tanto, una base para Col A es

$$\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

Las filas diferentes de cero de B forman una base para Fila A, a saber, $\{(1, -2, -4, 3, -2), (0, 3, 9, -12, 12)\}$. En este ejemplo en particular, resulta que dos filas cualesquiera de A forman una base para el espacio fila, ya que el espacio fila es de dimensión 2 y ninguna de las filas de A es un múltiplo de otra fila. En general, las filas diferentes de cero de una forma escalonada de A se deberían utilizar como base para Fila A, y no las filas de A.

3. Para Nul A, el siguiente paso es llevar a cabo operaciones de fila en B para obtener la forma escalonada reducida de A.
4. Rango A^T = rango A, de acuerdo con el teorema del rango, ya que Col A^{T} = Fila A. De manera que A^T tiene dos posiciones pivote.

4.7 CAMBIO DE BASE

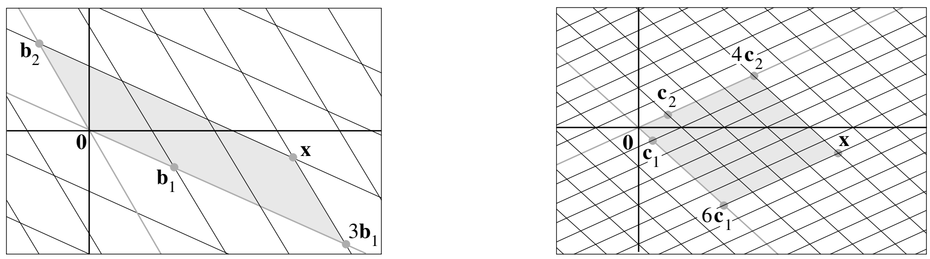
Cuando se elige una base $\mathcal{B}$ para un espacio vectorial V de dimensión n, el mapeo de coordenadas asociado a $\mathbb{R}^n$ proporciona un sistema de coordenadas para V. Se identifica cada $\mathbf{x}$ en V únicamente por su vector de $\mathcal{B}$-coordenadas, $[\,\mathbf{x}\,]_{\mathcal{B}}$.[1]

En algunas aplicaciones, se describe un problema inicialmente usando una base $\mathcal{B}$, pero la solución del problema se facilita cambiando $\mathcal{B}$ a una nueva base $\mathcal{C}$. (Se darán ejemplos en los capítulos 5 y 7). A cada vector se le asigna un nuevo vector de $\mathcal{C}$-coordenadas. En esta sección se estudia cómo $[\,\mathbf{x}\,]_{\mathcal{C}}$ y $[\,\mathbf{x}\,]_{\mathcal{B}}$ se relacionan para toda $\mathbf{x}$ en V.

Para visualizar el problema, considere los dos sistemas de coordenadas de la figura 1. En la figura 1*a*), $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$, mientras que en la figura 1*b*), la misma $\mathbf{x}$ se muestra como $\mathbf{x} = 6\mathbf{c}_1 + 4\mathbf{c}_2$. Es decir,

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{y} \quad [\,\mathbf{x}\,]_{\mathcal{C}} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Nuestro problema es encontrar la conexión entre los dos vectores de coordenadas. El ejemplo 1 muestra cómo hacer esto, siempre que se conozca cómo se forman $\mathbf{b}_1$ y $\mathbf{b}_2$ a partir de $\mathbf{c}_1$ y $\mathbf{c}_2$.



a) *b*)

FIGURA 1 Dos sistemas de coordenadas para el mismo espacio vectorial.

EJEMPLO 1 Considere dos bases $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ de un espacio vectorial V, de manera que

$$\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2 \quad \text{y} \quad \mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2 \tag{1}$$

Suponga que

$$\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2 \tag{2}$$

Es decir, suponga que $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Encuentre $[\,\mathbf{x}\,]_{\mathcal{C}}$.

SOLUCIÓN Aplique el mapeo de coordenadas determinado por C a $\mathbf{x}$ en la ecuación (2). Puesto que el mapeo de coordenadas es una transformación lineal,

$$\begin{aligned} [\,\mathbf{x}\,]_{\mathcal{C}} &= [\,3\mathbf{b}_1 + \mathbf{b}_2\,]_{\mathcal{C}} \\ &= 3[\,\mathbf{b}_1\,]_{\mathcal{C}} + [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{aligned}$$

Podemos escribir esta ecuación vectorial como una ecuación matricial, utilizando los vectores en la combinación lineal como las columnas de una matriz:

$$[\,\mathbf{x}\,]_{\mathcal{C}} = \begin{bmatrix} [\,\mathbf{b}_1\,]_{\mathcal{C}} & [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{3}$$

[1] Piense en $[\,\mathbf{x}\,]_{\mathcal{B}}$ como un "nombre" de $\mathbf{x}$ que lista los pesos utilizados para construir $\mathbf{x}$ como una combinación lineal de los vectores básicos en $\mathcal{B}$.

Esta fórmula da $[\mathbf{x}]_{\mathcal{C}}$, una vez que se conocen las columnas de la matriz. A partir de (1),

$$[\mathbf{b}_1]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad \text{y} \quad [\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

Por lo tanto, la ecuación (3) da la solución:

$$[\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Las $\mathcal{C}$-coordenadas de $\mathbf{x}$ coinciden con las de $\mathbf{x}$ en la figura 1. ■

El argumento que se utiliza para deducir la fórmula (3) se puede generalizar para obtener el siguiente resultado. (Véase los ejercicios 15 y 16).

TEOREMA 15

Sean $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ las bases de un espacio vectorial V. Entonces, existe una única matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ de $n \times n$ tal que

$$[\mathbf{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}} \tag{4}$$

Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son los vectores de $\mathcal{C}$-coordenadas de los vectores en la base $\mathcal{B}$. Es decir,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} & \cdots & [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix} \tag{5}$$

La matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ en el teorema 15 se denomina **matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$**. La multiplicación por $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ convierte a las $\mathcal{B}$-coordenadas en las $\mathcal{C}$-coordenadas.[2] La figura 2 ilustra la ecuación de cambio de coordenadas (4).

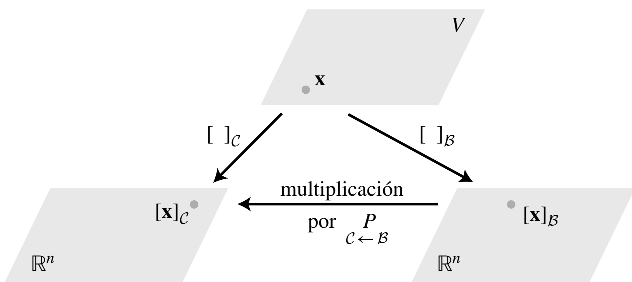


FIGURA 2 Dos sistemas de coordenadas para V.

Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son linealmente independientes porque son los vectores de coordenadas del conjunto linealmente independiente $\mathcal{B}$. (Véase el ejercicio 25 de la sección 4.4). Puesto que $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es cuadrada, debe ser invertible, de acuerdo con el teorema de la matriz invertible. Multiplicando por la izquierda ambos lados de la ecuación (4) por $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$ se obtiene

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}[\mathbf{x}]_{\mathcal{C}} = [\mathbf{x}]_{\mathcal{B}}$$

[2] Para recordar cómo se construye la matriz, piense en $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}}$ como en una combinación lineal de las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. El producto matriz-vector es un vector de $\mathcal{C}$-coordenadas, de modo que las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ también deberían ser los vectores de $\mathcal{C}$-coordenadas.

Por lo tanto, $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$ es la matriz que convierte las $\mathcal{B}$-coordenadas en $\mathcal{C}$-coordenadas. Es decir,

$$\boxed{(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}} \tag{6}$$

Cambio de base en $\mathbb{R}^n$

Si $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{E}$ es la *base estándar* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ en $\mathbb{R}^n$, entonces $[\mathbf{b}_1]_{\mathcal{E}} = \mathbf{b}_1$, y lo mismo para los otros vectores en $\mathcal{B}$. En este caso, $\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}$ es igual a la matriz de cambio de coordenadas $P_{\mathcal{B}}$ que se presentó en la sección 4.4, a saber,

$$P_{\mathcal{B}} = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n]$$

Para cambiar las coordenadas entre dos bases que no son estándar en $\mathbb{R}^n$, se necesita el teorema 15. El teorema demuestra que para resolver el problema de cambio de base, se necesitan los vectores de coordenadas de la antigua base respecto de la nueva base.

EJEMPLO 2 Sean $\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, y considere las bases de $\mathbb{R}^2$ dadas por $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$. Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

SOLUCIÓN La matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ implica a los vectores de $\mathcal{C}$-coordenadas de $\mathbf{b}_1$ y $\mathbf{b}_2$. Sean $[\,\mathbf{b}_1\,]_{\mathcal{C}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ y $[\,\mathbf{b}_2\,]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Así, por definición,

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \quad \text{y} \quad \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{b}_2$$

Para resolver ambos sistemas simultáneamente, se aumenta la matriz de coeficientes con $\mathbf{b}_1$ *y* $\mathbf{b}_2$, y se reduce por filas:

$$\left[\begin{array}{cc:cc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{b}_1 & \mathbf{b}_2 \end{array}\right] = \left[\begin{array}{cc:cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right] \sim \left[\begin{array}{cc:cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right] \tag{7}$$

Así,

$$[\,\mathbf{b}_1\,]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \quad \text{y} \quad [\,\mathbf{b}_2\,]_{\mathcal{C}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

La matriz del cambio de coordenadas deseada es, por consiguiente,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\,\mathbf{b}_1\,]_{\mathcal{C}} & [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$ ■

Observe que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ del ejemplo 2 ya apareció en la ecuación (7). Esto no es sorprendente, ya que la primera columna de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ resulta de reducir por filas $[\mathbf{c}_1 \quad \mathbf{c}_2 \,\vdots\, \mathbf{b}_1]$ a $[I \,\vdots\, [\mathbf{b}_1]_{\mathcal{C}}]$, y de manera similar para la segunda columna de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. Por lo tanto,

$$\boxed{[\,\mathbf{c}_1 \quad \mathbf{c}_2 \,\vdots\, \mathbf{b}_1 \quad \mathbf{b}_2\,] \sim [\,I \,\vdots\, \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\,]}$$

Un procedimiento análogo funciona para encontrar la matriz de cambio de coordenadas entre dos bases cualesquiera en $\mathbb{R}^n$.

EJEMPLO 3 Sea $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, y considere las bases para $\mathbb{R}^2$ dadas por $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$.

a) Determine la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

b) Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

SOLUCIÓN

a) Considere que $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ se necesita más que $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, y calcule

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \vdots & \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} = \left[\begin{array}{cc:cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array}\right] \sim \left[\begin{array}{cc:cc} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array}\right]$$

Por lo tanto,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$$

b) De acuerdo con el inciso *a*) y la propiedad (6) anterior (con $\mathcal{B}$ y $\mathcal{C}$ intercambiadas),

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = (\underset{\mathcal{B}\leftarrow\mathcal{C}}{P})^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix} \qquad \blacksquare$$

Otra descripción de la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ utiliza las matrices de cambio de coordenadas $P_{\mathcal{B}}$ y $P_{\mathcal{C}}$ que convierten a las coordenadas $\mathcal{B}$ y $\mathcal{C}$, respectivamente, en coordenadas estándar. Recuerde que para cada $\mathbf{x}$ en $\mathbb{R}^n$,

$$P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = \mathbf{x}, \quad P_{\mathcal{C}}[\mathbf{x}]_{\mathcal{C}} = \mathbf{x} \quad \text{y} \quad [\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}\mathbf{x}$$

Por lo tanto,

$$[\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}\mathbf{x} = P_{\mathcal{C}}^{-1}P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$$

En $\mathbb{R}^n$, la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ se puede calcular como $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$. En realidad, para matrices más grandes que 2×2, un algoritmo similar al del ejemplo 3 es más rápido que calcular $P_{\mathcal{C}}^{-1}$ y luego $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$. Véase el ejercicio 12 en la sección 2.2.

PROBLEMAS DE PRÁCTICA

1. Sean $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$ y $\mathcal{G} = \{\mathbf{g}_1, \mathbf{g}_2\}$ bases para un espacio vectorial V, y sea P una matriz cuyas columnas son $[\mathbf{f}_1]_{\mathcal{G}}$ y $[\mathbf{f}_2]_{\mathcal{G}}$. ¿Cuál de las siguientes ecuaciones satisface P para toda $\mathbf{v}$ en V?

   i. $[\mathbf{v}]_{\mathcal{F}} = P[\mathbf{v}]_{\mathcal{G}}$ **ii.** $[\mathbf{v}]_{\mathcal{G}} = P[\mathbf{v}]_{\mathcal{F}}$

2. Sean $\mathcal{B}$ y $\mathcal{C}$ como en el ejemplo 1. Utilice los resultados de ese ejemplo para encontrar la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

4.7 EJERCICIOS

1. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para un espacio vectorial V y suponga que $\mathbf{b}_1 = 6\mathbf{c}_1 - 2\mathbf{c}_2$ y $\mathbf{b}_2 = 9\mathbf{c}_1 - 4\mathbf{c}_2$.

   a) Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

   b) Encuentre $[\mathbf{x}]_{\mathcal{C}}$ para $\mathbf{x} = -3\mathbf{b}_1 + 2\mathbf{b}_2$. Utilice el inciso *a*).

2. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para un espacio vectorial V, y suponga que $\mathbf{b}_1 = -2\mathbf{c}_1 + 4\mathbf{c}_2$ y $\mathbf{b}_2 = 3\mathbf{c}_1 - 6\mathbf{c}_2$.

   a) Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

   b) Determine $[\mathbf{x}]_{\mathcal{C}}$ para $\mathbf{x} = 2\mathbf{b}_1 + 3\mathbf{b}_2$.

3. Sean $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2\}$ y $\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2\}$ bases para V, y sea P una matriz cuyas columnas son $[\mathbf{u}_1]_{\mathcal{W}}$ y $[\mathbf{u}_2]_{\mathcal{W}}$. ¿Cuál de las siguientes ecuaciones satisface P para toda $\mathbf{x}$ en V?

i. $[\mathbf{x}]_{\mathcal{U}} = P[\mathbf{x}]_{\mathcal{W}}$ **ii.** $[\mathbf{x}]_{\mathcal{W}} = P[\mathbf{x}]_{\mathcal{U}}$

4. Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ bases para V, y sea $P[[\mathbf{d}_1]_{\mathcal{A}}\ [\mathbf{d}_2]_{\mathcal{A}}\ [\mathbf{d}_3]_{\mathcal{A}}]$. ¿Cuál de las siguientes ecuaciones satisface P para toda $\mathbf{x}$ en V?

i. $[\mathbf{x}]_{\mathcal{A}} = P[\mathbf{x}]_{\mathcal{D}}$ **ii.** $[\mathbf{x}]_{\mathcal{D}} = P[\mathbf{x}]_{\mathcal{A}}$

5. Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ bases para un espacio vectorial V, y suponga que $\mathbf{a}_1 = 4\mathbf{b}_1 - \mathbf{b}_2$, $\mathbf{a}_2 = -\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$, y $\mathbf{a}_3 = \mathbf{b}_2 - 2\mathbf{b}_3$.

a) Encuentre la matriz de cambio de coordenadas de $\mathcal{A}$ a $\mathcal{B}$.

b) Determine $[\mathbf{x}]_{\mathcal{B}}$ para $\mathbf{x} = 3\mathbf{a}_1 + 4\mathbf{a}_2 + \mathbf{a}_3$.

6. Sean $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ y $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ bases para un espacio vectorial V, y suponga que $\mathbf{f}_1 = 2\mathbf{d}_1 - \mathbf{d}_2 + \mathbf{d}_3$, $\mathbf{f}_2 = 3\mathbf{d}_2 + \mathbf{d}_3$, y $\mathbf{f}_3 = -3\mathbf{d}_1 + 2\mathbf{d}_3$.

a) Encuentre la matriz de cambio de las coordenadas de $\mathcal{F}$ a $\mathcal{D}$.

b) Determine $[\mathbf{x}]_{\mathcal{D}}$ para $\mathbf{x} = \mathbf{f}_1 - 2\mathbf{f}_2 + 2\mathbf{f}_3$.

En los ejercicios 7 a 10, sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para $\mathbb{R}^2$. En cada ejercicio, encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$ y la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

7. $\mathbf{b}_1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 8 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -7 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

9. $\mathbf{b}_1 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

10. $\mathbf{b}_1 = \begin{bmatrix} 6 \\ -12 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 3 \\ 9 \end{bmatrix}$

En los ejercicios 11 y 12, $\mathcal{B}$ y $\mathcal{C}$ son bases para un espacio vectorial V. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

11. *a*) Las columnas de la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son vectores de $\mathcal{B}$-coordenadas de los vectores en $\mathcal{C}$.

b) Si $V = \mathbb{R}^n$, y $\mathcal{C}$ es la base *estándar* para V, entonces $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es igual a la matriz de cambio de coordenadas $P_{\mathcal{B}}$ que se presentó en la sección 4.4.

12. *a*) Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son linealmente independientes.

b) Si $V = \mathbb{R}^2$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, entonces la reducción por filas de $[\mathbf{c}_1\ \ \mathbf{c}_2\ \ \mathbf{b}_1\ \ \mathbf{b}_2]$ a $[I\ \ P]$ produce una matriz P que satisface $[\mathbf{x}]_{\mathcal{B}} = P[\mathbf{x}]_{\mathcal{C}}$ para toda $\mathbf{x}$ en V.

13. En $\mathbb{P}_2$, encuentre la matriz de cambio de coordenadas de la base $\mathcal{B} = \{1 - 2t + t^2, 3 - 5t + 4t^2, 2t + 3t^2\}$ a la base estándar de $\mathcal{C} = \{1, t, t^2\}$. Luego, encuentre el vector de $\mathcal{B}$-coordenadas para $-1 + 2t$.

14. En $\mathbb{P}_2$, encuentre la matriz de cambio de coordenadas de la base $\mathcal{B} = \{1 - 3t^2, 2 + t - 5t^2, 1 + 2t\}$ a la base estándar. Luego, escriba t^2 como una combinación lineal de los polinomios en $\mathcal{B}$.

Los ejercicios 15 y 16 brindan una demostración del teorema 15. Complete la justificación para cada paso.

15. Dada $\mathbf{v}$ en V, existen escalares $x_1, \ldots, x_n$, tales que

$$\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n$$

porque *a*) ______ . Aplique el mapeo de coordenadas determinado por la base $\mathcal{C}$, y obtenga

$$[\mathbf{v}]_{\mathcal{C}} = x_1[\mathbf{b}_1]_{\mathcal{C}} + x_2[\mathbf{b}_2]_{\mathcal{C}} + \cdots + x_n[\mathbf{b}_n]_{\mathcal{C}}$$

porque *b*) ______ . Esta ecuación se puede escribir en la forma

$$[\,\mathbf{v}\,]_{\mathcal{C}} = \begin{bmatrix} [\,\mathbf{b}_1\,]_{\mathcal{C}} & [\,\mathbf{b}_2\,]_{\mathcal{C}} & \cdots & [\,\mathbf{b}_n\,]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{8}$$

de acuerdo con la definición de *c*) ______ . Esto indica que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ que se muestra en la ecuación (5) satisface $[\mathbf{v}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{v}]_{\mathcal{B}}$ para toda $\mathbf{v}$ en V, porque el vector del lado derecho en (8) es *d*) ______ .

16. Suponga que Q es cualquier matriz de manera que

$$[\mathbf{v}]_{\mathcal{C}} = Q[\mathbf{v}]_{\mathcal{B}} \quad \text{para toda } \mathbf{v} \text{ en } V \tag{9}$$

Establezca $\mathbf{v} = \mathbf{b}_1$ en la ecuación (9). Después (9) indica que $[\mathbf{b}_1]_{\mathcal{C}}$ es la primera columna de Q porque *a*) ______ . De manera similar, para $k = 2, \ldots, n$, la k-ésima columna de Q es *b*) _____ porque *c*) _____ . Esto indica que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ definida por (5) en el teorema 15 es la única matriz que satisface la condición (4).

17. **[M]** Sea $\mathcal{B} = \{\mathbf{x}_0, \ldots, \mathbf{x}_6\}$ y $\mathcal{C} = \{\mathbf{y}_0, \ldots, \mathbf{y}_6\}$, donde $\mathbf{x}_k$ es la función $\cos^k t$ y $\mathbf{y}_k$ es la función $\cos kt$. El ejercicio 34 de la sección 4.5 mostró que tanto B como C son bases para el espacio vectorial $H = \text{Gen}\,\{\mathbf{x}_0, \ldots, \mathbf{x}_6\}$.

a) Sea $P = \begin{bmatrix} [\,\mathbf{y}_0\,]_{\mathcal{B}} & \cdots & [\,\mathbf{y}_6\,]_{\mathcal{B}} \end{bmatrix}$, y calcule P^{-1}.

b) Explique por qué las columnas de P^{-1} son los vectores de $\mathcal{C}$-coordenadas de $\mathbf{x}_0, \ldots, \mathbf{x}_6$. Después, utilice estos vectores de coordenadas para escribir identidades trigonométricas que expresen potencias de $\cos t$ en términos de las funciones en $\mathcal{C}$.

18. **[M]** (*Se requiere cálculo*)[3] Recuerde de sus clases de cálculo que las integrales como

$$\int (5\cos^3 t - 6\cos^4 t + 5\cos^5 t - 12\cos^6 t)\, dt \tag{10}$$

son tediosas de calcular. (El método habitual es aplicar integración por partes varias veces y usar la fórmula de la mitad del ángulo). Utilice la matriz P o P^{-1} del ejercicio 17 para transformar (10), y después calcule la integral.

[3] La idea de los ejercicios 17 y 18 y cinco ejercicios relacionados en las secciones anteriores provienen de un documento de Jack W. Rogers, Jr., de la Universidad de Auburn, presentado en una reunión de la International Linear Algebra Society, en agosto de 1995. Véase la sección "Applications of Linear Algebra in Calculus", *American Mathematical Monthly* **104** (1), 1997.

19. [**M**] Sea

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

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -8 \\ 5 \\ 2 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} -7 \\ 2 \\ 6 \end{bmatrix}$$

a) Encuentre una base $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ para $\mathbb{R}^3$ tal que P es la matriz de cambio de coordenadas de $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ a la base $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. [*Sugerencia:* Pregúntese qué representan las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$].

b) Encuentre una base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ para $\mathbb{R}^3$ tal que P es la matriz de cambio de coordenadas de $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

20. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ y bases para un espacio vectorial de dimensión 2.

a) Escriba una ecuación que relacione las matrices $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, $\underset{\mathcal{D}\leftarrow\mathcal{C}}{P}$ y $\underset{\mathcal{D}\leftarrow\mathcal{B}}{P}$. Justifique su resultado.

b) [**M**] Utilice un programa de matrices ya sea para ayudarle a encontrar la ecuación o para comprobar la ecuación que escribió. Trabaje con tres bases de $\mathbb{R}^2$. (Véase los ejercicios 7 a 10).

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Como las columnas de P son vectores de coordenadas $\mathcal{G}$, un vector de la forma $P\mathbf{x}$ debe ser un vector de coordenadas $\mathcal{G}$. Por lo tanto, P satisface la ecuación **ii**.

2. Los vectores de coordenadas que se encuentran en el ejemplo 1 indican que

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left[\ [\mathbf{b}_1]_{\mathcal{C}} \quad [\mathbf{b}_2]_{\mathcal{C}}\ \right] = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}$$

Por lo tanto,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = (\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \frac{1}{10}\begin{bmatrix} 1 & 6 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} .1 & .6 \\ -.1 & .4 \end{bmatrix}$$

4.8 APLICACIONES A LAS ECUACIONES EN DIFERENCIAS

Ahora que se dispone de poderosas computadoras, cada vez más problemas científicos y de ingeniería se analizan utilizando datos discretos, o digitales, en lugar de datos continuos. Las ecuaciones en diferencias con frecuencia son la herramienta adecuada para analizar esos datos. Incluso cuando una ecuación diferencial se utiliza para modelar un proceso continuo, a menudo se obtiene una solución numérica a partir de una ecuación en diferencias relacionada.

En esta sección se ponen de relieve algunas propiedades fundamentales de las ecuaciones en diferencias lineales que se explican mejor utilizando el álgebra lineal.

Señales discretas de tiempo

El espacio vectorial $\mathbb{S}$ de las señales discretas de tiempo se presentó en la sección 4.1. Una **señal** en $\mathbb{S}$ es una función definida solo con números enteros y se visualiza como una secuencia de números, por ejemplo, $\{y_k\}$. La figura 1 muestra tres señales típicas cuyos términos generales son $(.7)^k$, 1^k y $(-1)^k$, respectivamente.

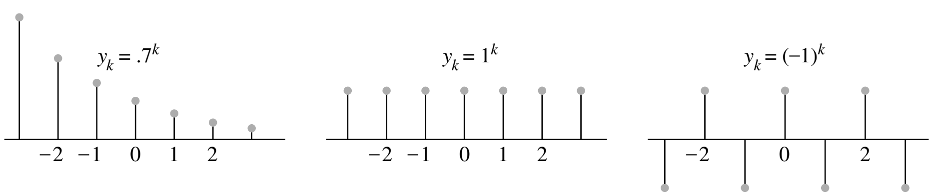


FIGURA 1 Tres señales en $\mathbb{S}$.

Las señales digitales, desde luego, surgen en ingeniería de sistemas eléctricos y de control, pero secuencias de datos discretos también se generan en biología, física, economía, demografía y muchas otras áreas, donde se mide un proceso, o se *muestrea*, en intervalos discretos de tiempo. Cuando se inicia un proceso en un momento específico, a veces es conveniente representar una señal como una secuencia de la forma $(y_0, y_1, y_2, \ldots)$. Se supone que los términos y_k para $k < 0$ son cero o simplemente se omiten.

EJEMPLO 1 El sonido cristalino de un reproductor de discos compactos proviene de música de la que se han tomado muestras a una velocidad de 44,100 veces por segundo. Véase la figura 2. En cada medición, la amplitud de la señal de música se registra como un número, por ejemplo, y_k. La música original está compuesta de muchos sonidos diferentes de diversas frecuencias; sin embargo, la secuencia $\{y_k\}$ contiene suficiente información para reproducir todas las frecuencias en el sonido hasta aproximadamente 20,000 ciclos por segundo, más allá de lo que el oído humano puede percibir. ■

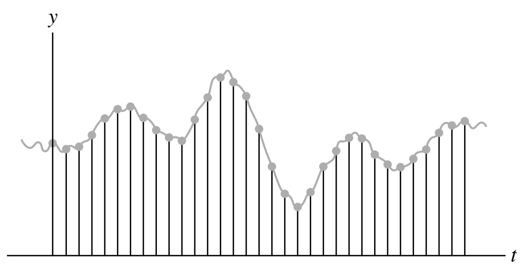


FIGURA 2 Datos muestreados a partir de una señal de música.

Independencia lineal en el espacio $\mathbb{S}$ de las señales

Para simplificar la notación, consideramos un conjunto de solo tres señales en $\mathbb{S}$, por ejemplo, $\{u_k\}$, $\{v_k\}$ y $\{w_k\}$. Son linealmente independientes precisamente cuando la ecuación

$$c_1 u_k + c_2 v_k + c_3 w_k = 0 \quad \text{para toda } k \tag{1}$$

implica que $c_1 = c_2 = c_3 = 0$. La frase "para toda k" significa para todos los enteros: positivos, negativos y cero. También se podría considerar que las señales comienzan con $k = 0$, por ejemplo; en tal caso, "para toda k" significaría para todos los enteros mayores que o iguales a cero ($k \geq 0$).

Suponga que c_1, c_2, c_3 satisfacen la ecuación (1). Entonces la ecuación (1) es válida para cualesquiera tres valores consecutivos de k, por ejemplo, k, $k + 1$ y $k + 2$. Así, la ecuación (1) implica que

$$c_1 u_{k+1} + c_2 v_{k+1} + c_3 w_{k+1} = 0 \quad \text{para toda } k$$

y

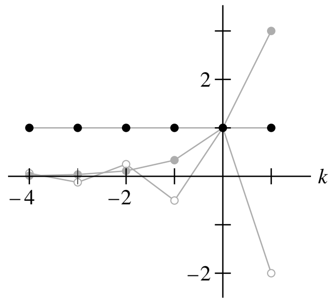
$$c_1 u_{k+2} + c_2 v_{k+2} + c_3 w_{k+2} = 0 \quad \text{para toda } k$$

Por lo tanto c_1, c_2, c_3 satisfacen

$$\begin{bmatrix} u_k & v_k & w_k \\ u_{k+1} & v_{k+1} & w_{k+1} \\ u_{k+2} & v_{k+2} & w_{k+2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{para toda } k \tag{2}$$

La matriz de coeficientes de este sistema se llama la **matriz de Casorati** de las señales, y el determinante de la matriz se denomina **casoratiano** de $\{u_k\}$, $\{v_k\}$ y $\{w_k\}$. Si la matriz de Casorati es invertible, para al menos un valor de k, entonces la ecuación (2) implicará que $c_1 = c_2 = c_3 = 0$, lo que demuestra que las tres señales son linealmente independientes.

EJEMPLO 2 Compruebe que las señales 1^k, $(-2)^k$ y 3^k son linealmente independientes.



Las señales 1^k, $(-2)^k$ y 3^k.

SOLUCIÓN La matriz de Casorati es

$$\begin{bmatrix} 1^k & (-2)^k & 3^k \\ 1^{k+1} & (-2)^{k+1} & 3^{k+1} \\ 1^{k+2} & (-2)^{k+2} & 3^{k+2} \end{bmatrix}$$

Con las operaciones de fila se puede demostrar muy fácilmente que esta matriz siempre es invertible. Sin embargo, es más rápido sustituir k por un valor, por ejemplo, $k = 0$ y reducir por filas la matriz numérica:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 3 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 0 & 10 \end{bmatrix}$$

La matriz de Casorati es invertible para $k = 0$. Por lo tanto, 1^k, $(-2)^k$ y 3^k son linealmente independientes. ■

Si una matriz de Casorati no es invertible, las señales asociadas que se están sometiendo a prueba pueden o no ser linealmente dependientes. (Véase el ejercicio 33). Sin embargo, es posible demostrar que si todas las señales son las soluciones de la *misma* ecuación en diferencias homogénea (que se describe a continuación), entonces la matriz de Casorati es invertible para toda k y las señales son linealmente independientes, o bien, la matriz de Casorati no es invertible para toda k y las señales son linealmente dependientes. En la *Guía de estudio* se presenta una excelente demostración utilizando transformaciones lineales.

Ecuaciones lineales en diferencias

Considerando los escalares $a_0, \ldots, a_n$, con a_0 y a_n diferentes de cero, y dada una señal $\{z_k\}$, la ecuación

$$a_0 y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{3}$$

se llama una **ecuación lineal en diferencias** (o **relación de recurrencia lineal**) **de orden n**. Para simplificar, a_0 con frecuencia se considera igual a 1. Si $\{z_k\}$ es la secuencia cero, la ecuación es **homogénea**; de lo contrario, la ecuación es **no homogénea**.

EJEMPLO 3 En el procesamiento de señal digital, una ecuación en diferencias tal como la ecuación (3) describe un **filtro lineal**, y $a_0, \ldots, a_n$ se denominan los **coeficientes de filtro**. Si $\{y_k\}$ se trata como la entrada y $\{z_k\}$ como la salida, entonces las soluciones de la ecuación homogénea asociada son las señales que se filtran *hacia fuera* y se transforman en señal cero. Vamos a alimentar dos señales diferentes en el filtro

$$.35y_{k+2} + .5y_{k+1} + .35y_k = z_k$$

Aquí, .35 es una abreviatura de $\sqrt{2}/4$. La primera señal se crea mediante el muestreo de la señal continua $y = \cos(\pi t/4)$ para valores enteros de t, como en la figura 3a). La señal discreta es

$$\{y_k\} = \{\ldots, \cos(0), \cos(\pi/4), \cos(2\pi/4), \cos(3\pi/4), \ldots\}$$

Para simplificar, se escribe $\pm.7$ en lugar de $\pm\sqrt{2}/2$, de manera que

$$\{y_k\} = \{\ldots, \underset{\substack{\uparrow \\ k=0}}{1}, .7, 0, -.7, -1, -.7, 0, .7, 1, .7, 0, \ldots\}$$

La tabla 1 muestra el cálculo de la secuencia de salida $\{z_k\}$, donde .35(.7) es una abreviatura de $(\sqrt{2}/4)(\sqrt{2}/2) = .25$. La salida es $\{y_k\}$, corrida un término.

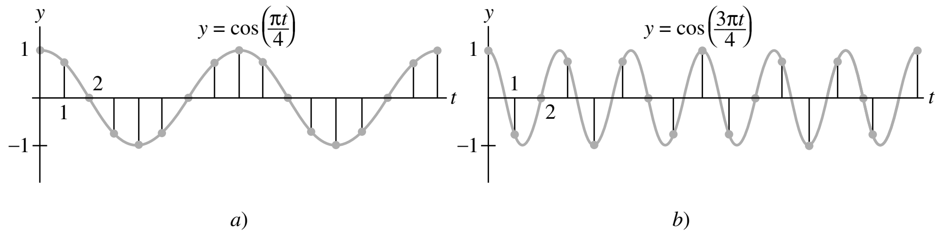


FIGURA 3 Señales discretas con diferentes frecuencias.

TABLA 1 Cálculo de la salida de un filtro

k	y_k	y_{k+1}	y_{k+2}	$.35y_k$	$+ .5y_{k+1}$	$+ .35y_{k+2}$	$=$	z_k
0	1	.7	0	.35(1)	+ .5(.7)	+ .35(0)	=	.7
1	.7	0	−.7	.35(.7)	+ .5(0)	+ .35(−.7)	=	0
2	0	−.7	−1	.35(0)	+ .5(−.7)	+ .35(−1)	=	−.7
3	−.7	−1	−.7	.35(−.7)	+ .5(−1)	+ .35(−.7)	=	−1
4	−1	−.7	0	.35(−1)	+ .5(−.7)	+ .35(0)	=	−.7
5	−.7	0	.7	.35(−.7)	+ .5(0)	+ .35(.7)	=	0
⋮	⋮							⋮

Una señal de entrada diferente se produce a partir de la mayor frecuencia de la señal $y = \cos(3\pi t/4)$, que se muestra en la figura 3b). El muestreo a la misma tasa de antes produce una nueva secuencia de entrada:

$$\{w_k\} = \{\ldots, \underset{\substack{\uparrow \\ k=0}}{1}, -.7,\ 0,\ .7,\ -1,\ .7,\ 0,\ -.7,\ 1,\ -.7,\ 0, \ldots\}$$

Cuando se alimenta al filtro con $\{\omega_k\}$, la salida es la secuencia cero. El filtro, conocido como *filtro pasa bajos*, hace que $\{y_k\}$ pase, pero detiene a las señales de mayor frecuencia $\{\omega_k\}$. ■

En muchas aplicaciones, se especifica una secuencia $\{z_k\}$ para el lado derecho de una ecuación en diferencias (3), y una $\{y_k\}$ que satisface la ecuación (3) se considera una **solución** de la ecuación. El siguiente ejemplo ilustra cómo encontrar soluciones para una ecuación homogénea.

EJEMPLO 4 Soluciones de una ecuación homogénea en diferencias a menudo tienen la forma $y_k = r^k$ para alguna r. Encuentre algunas soluciones de la ecuación

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k \tag{4}$$

SOLUCIÓN Se sustituye y^k por r^k en la ecuación y se factoriza el lado izquierdo:

$$r^{k+3} - 2r^{k+2} - 5r^{k+1} + 6r^k = 0 \tag{5}$$

$$r^k(r^3 - 2r^2 - 5r + 6) = 0$$

$$r^k(r-1)(r+2)(r-3) = 0 \tag{6}$$

Como (5) es equivalente a (6), r^k satisface la ecuación en diferencias (4) si y solo si r satisface (6). Así, 1^k, $(-2)^k$ y 3^k son todas soluciones de (4). Por ejemplo, para comprobar que 3^k es una solución de (4), calcule

$$\begin{aligned} &3^{k+3} - 2\cdot 3^{k+2} - 5\cdot 3^{k+1} + 6\cdot 3^k \\ &= 3^k(27 - 18 - 15 + 6) = 0 \quad \text{para toda } k \end{aligned}$$

■

En general, una señal distinta de cero r^k satisface la ecuación homogénea en diferencias

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

si y solo si r es una raíz de la **ecuación auxiliar**

$$r^n + a_1 r^{n-1} + \cdots + a_{n-1} r + a_n = 0$$

No vamos a considerar el caso en que r es una raíz repetida de la ecuación auxiliar. Cuando la ecuación auxiliar tiene una *raíz compleja*, la ecuación en diferencias tiene soluciones de la forma $s^k \cos k\omega$ y s^k sen $k\omega$, para las constantes s y ω. Esto sucedió en el ejemplo 3.

Conjuntos de soluciones de ecuaciones lineales en diferencias

A partir de $a_1, \ldots, a_n$, considere el mapeo $T : \mathbb{S} \to \mathbb{S}$ que transforma una señal $\{y_k\}$ en una señal $\{\omega_k\}$ dada por

$$w_k = y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k$$

Es fácil comprobar que T es una transformación *lineal*. Esto implica que el conjunto solución de la ecuación homogénea

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

es el núcleo de T (el conjunto de las señales que T mapea en la señal cero) y, por lo tanto, el conjunto solución es un *subespacio* de $\mathbb{S}$. Cualquier combinación lineal de soluciones es de nuevo una solución.

El siguiente teorema, un resultado sencillo, pero básico, conducirá a mayor información acerca de los conjuntos solución de ecuaciones en diferencias.

TEOREMA 16

Si $a_n \neq 0$ y si $\{z_k\}$ está dada, la ecuación

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{7}$$

tiene una única solución siempre que se especifican $y_0, \ldots, y_{n-1}$.

DEMOSTRACIÓN Si $y_0, \ldots, y_{n-1}$ se especifican, utilice la ecuación (7) para *definir*

$$y_n = z_0 - [\, a_1 y_{n-1} + \cdots + a_{n-1} y_1 + a_n y_0 \,]$$

Y ahora que $y_1, \ldots, y_n$ se especifican, utilice (7) para definir y_{n+1}. En general, utilice la relación de recurrencia

$$y_{n+k} = z_k - [\, a_1 y_{k+n-1} + \cdots + a_n y_k \,] \tag{8}$$

para definir y_{n+k} para $k \geq 0$. Para definir y_k para $k < 0$, utilice la relación de recurrencia

$$y_k = \frac{1}{a_n} z_k - \frac{1}{a_n} [\, y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} \,] \tag{9}$$

Esto produce una señal que satisface la ecuación (7). A la inversa, cualquier señal que satisface la ecuación (7) para toda k sin duda satisface las ecuaciones (8) y (9), por lo que la solución de la ecuación (7) es única. ■

TEOREMA 17

El conjunto H de todas las soluciones de la ecuación lineal homogénea en diferencias de n-ésimo orden

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k \tag{10}$$

es un espacio vectorial de dimensión n.

DEMOSTRACIÓN Como se indicó antes, H es un subespacio de $\mathbb{S}$ porque H es el núcleo de una transformación lineal. Para $\{y_k\}$ en H, sea $F\{y_k\}$ el vector en $\mathbb{R}^n$ dado por $(y_0, y_1, \ldots, y_{n-1})$. Es posible comprobar fácilmente que $F : H \to \mathbb{R}^n$ es una transformación lineal. Dado cualquier vector $(y_0, y_1, \ldots, y_{n-1})$ en $\mathbb{R}^n$, el teorema 16 dice que hay una única señal $\{y_k\}$ en H tal que $F\{y_k\} = (y_0, y_1, \ldots, y_{n-1})$. Esto significa que F es una transformación lineal uno a uno de H sobre $\mathbb{R}^n$, es decir, F es un isomorfismo. Así, dim H = dim $\mathbb{R}^n = n$. (Véase el ejercicio 32 en la sección 4.5). ■

EJEMPLO 5 Encuentre una base para el conjunto de todas las soluciones a la ecuación en diferencias

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k$$

SOLUCIÓN Nuestro trabajo en álgebra lineal ¡realmente está rindiendo frutos ahora! Sabemos a partir de los ejemplos 2 y 4 que 1^k, $(-2)^k$ y 3^k son soluciones linealmente independientes. En general, puede ser difícil comprobar directamente que un conjunto de las señales *genera* el espacio solución. Pero eso no es problema en este caso, debido a dos importantes teoremas: el teorema 17, que demuestra que el espacio solución es exactamente de dimensión 3, y el teorema de la base de la sección 4.5, el cual afirma que un conjunto linealmente independiente de n vectores en un espacio de dimensión n es automáticamente una base. Así, 1^k, $(-2)^k$ y 3^k forman una base para el espacio solución. ■

La solución estándar para describir la "solución general" de la ecuación en diferencias (10) es mostrar una base para el subespacio de todas las soluciones. Dicha base se suele denominar **conjunto fundamental de soluciones** de (10). En la práctica, si se encuentran n señales linealmente independientes que satisfagan (10), de manera automática generarán el espacio de soluciones de dimensión n, como se explica en el ejemplo 5.

Ecuaciones no homogéneas

La solución general de la ecuación no homogénea en diferencias

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{11}$$

se puede escribir como una solución particular de (11) más una combinación lineal arbitraria de un conjunto fundamental de soluciones de la ecuación homogénea correspondiente (10). Este hecho es análogo al resultado de la sección 1.5 que muestra que los conjuntos solución de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$ son paralelos. Ambos resultados tienen la misma explicación: el mapeo $\mathbf{x} \mapsto A\mathbf{x}$ es lineal, y el mapeo que transforma la señal $\{y_k\}$ en la señal $\{z_k\}$ en (11) es lineal. Véase el ejercicio 35.

EJEMPLO 6 Compruebe que la señal $y_k = k^2$ satisface la ecuación en diferencias

$$y_{k+2} - 4y_{k+1} + 3y_k = -4k \quad \text{para toda } k \tag{12}$$

Luego, encuentre una descripción de todas las soluciones de esta ecuación.

SOLUCIÓN Sustituya y_k por k^2 en el lado izquierdo de (12):

$$\begin{aligned}(k+2)^2 &- 4(k+1)^2 + 3k^2 \\ &= (k^2 + 4k + 4) - 4(k^2 + 2k + 1) + 3k^2 \\ &= -4k\end{aligned}$$

Así, k^2 es de hecho una solución de (12). El siguiente paso es resolver la ecuación homogénea

$$y_{k+2} - 4y_{k+1} + 3y_k = 0 \tag{13}$$

La ecuación auxiliar es

$$r^2 - 4r + 3 = (r-1)(r-3) = 0$$

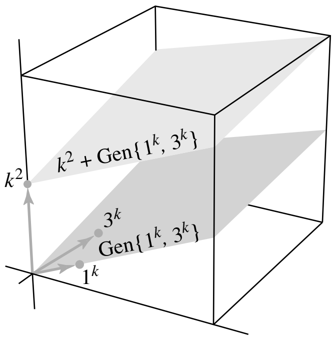


FIGURA 4
Conjuntos solución de las ecuaciones en diferencias (12) y (13).

Las raíces son $r = 1, 3$. Así que dos soluciones de la ecuación homogénea en diferencias son 1^k y 3^k. Evidentemente, no son múltiplos una de la otra, por lo que son señales linealmente independientes. De acuerdo con el teorema 17, el espacio solución es de dimensión 2, de manera que 3^k y 1^k forman una base para el conjunto de soluciones de la ecuación (13). Al trasladar ese conjunto mediante una solución particular de la ecuación no homogénea (12), se obtiene la solución general de (12):

$$k^2 + c_1 1^k + c_2 3^k \quad \text{o} \quad k^2 + c_1 + c_2 3^k$$

La figura 4 muestra una visualización geométrica de los dos conjuntos solución. Cada punto de la figura corresponde a una señal en $\mathbb{S}$. ■

Reducción a sistemas de ecuaciones de primer orden

Una forma moderna de estudiar una ecuación en diferencias homogénea de n-ésimo orden es remplazarla por un sistema equivalente de ecuaciones en diferencias de primer orden, escrito en la forma

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para toda } k$$

donde los vectores $\mathbf{x}_k$ están en $\mathbb{R}^n$ y A es una matriz de $n \times n$.

Un ejemplo sencillo de tal ecuación en diferencias (con valor vectorial) se estudió ya en la sección 1.10. Otros ejemplos se tratarán en las secciones 4.9 y 5.6.

EJEMPLO 7 Escriba la siguiente ecuación en diferencias como un sistema de primer orden:

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k$$

SOLUCIÓN Para cada k, establezca

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

La ecuación en diferencias dice que $y_{k+3} = -6y_k + 5y_{k+1} + 2y_{k+2}$, por lo que

$$\mathbf{x}_{k+1} = \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix} = \begin{bmatrix} 0 & + & y_{k+1} & + 0 \\ 0 & + 0 & + & y_{k+2} \\ -6y_k & + 5y_{k+1} & + 2y_{k+2} & \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix} \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

Es decir,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para toda } k, \quad \text{donde} \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$$ ■

En general, la ecuación

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

se puede rescribir como $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para toda k, donde

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

Lecturas adicionales

Hamming, R. W., *Digital Filters*, 3a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 1989), pp. 1-37.

Kelly, W. G. y A. C. Peterson, *Difference Equations*, 2a. ed. (San Diego: Harcourt-Academic Press, 2001).

Mickens, R. E., *Difference Equations*, 2a. ed. (Nueva York: Van Nostrand Reinhold, 1990), pp. 88-141.

Oppenheim, A. V. y A. S. Willsky, *Signals and Systems*, 2a. ed. (Upper Saddle River, NJ: Prentice-Hall, 1997), pp. 1-14, 21-30, 38-43.

PROBLEMA DE PRÁCTICA

Es posible demostrar que las señales 2^k, $3^k \operatorname{sen} \frac{k\pi}{2}$, y $3^k \cos \frac{k\pi}{2}$ son soluciones de

$$y_{k+3} - 2y_{k+2} + 9y_{k+1} - 18y_k = 0$$

Demuestre que estas señales forman una base para el conjunto de todas las soluciones de la ecuación en diferencias.

4.8 EJERCICIOS

Compruebe que las señales de los ejercicios 1 y 2 son soluciones de la ecuación en diferencias correspondiente.

1. $2^k, (-4)^k;\ y_{k+2} + 2y_{k+1} - 8y_k = 0$

2. $5^k, (-5)^k;\ y_{k+2} - 25y_k = 0$

Demuestre que las señales en los ejercicios 3 a 6 forman una base para el conjunto solución de la ecuación en diferencias correspondiente.

3. Las señales y la ecuación en el ejercicio 1

4. Las señales y la ecuación en el ejercicio 2

5. $(-2)^k, k(-2)^k;\ y_{k+2} + 4y_{k+1} + 4y_k = 0$

6. $4^k \cos\left(\frac{k\pi}{2}\right), 4^k \operatorname{sen}\left(\frac{k\pi}{2}\right);\ y_{k+2} + 16y_k = 0$

En los ejercicios 7 a 12, suponga que las señales listadas son soluciones de la ecuación en diferencias dada. ¿Las señales forman una base para el espacio solución de la ecuación? Justifique sus respuestas usando los teoremas adecuados.

7. $1^k, 2^k, (-2)^k;\ y_{k+3} - y_{k+2} - 4y_{k+1} + 4y_k = 0$

8. $(-1)^k, 2^k, 3^k;\ y_{k+3} - 4y_{k+2} + 1y_{k+1} + 6y_k = 0$

9. $2^k, 5^k \cos\left(\frac{k\pi}{2}\right), 5^k \operatorname{sen}\left(\frac{k\pi}{2}\right);$

$y_{k+3} - 2y_{k+2} + 25y_{k+1} - 50y_k = 0$

10. $(-2)^k, k(-2)^k, 3^k;\ y_{k+3} + y_{k+2} - 8y_{k+1} - 12y_k = 0$

11. $(-1)^k, 2^k;\ y_{k+3} - 3y_{k+2} + 4y_k = 0$

12. $3^k, (-2)^k;\ y_{k+4} - 13y_{k+2} + 36y_k = 0$

En los ejercicios 13 a 16, encuentre una base para el espacio solución de la ecuación en diferencias. Demuestre que las soluciones que encuentre generan al conjunto solución.

13. $y_{k+2} - y_{k+1} + \frac{2}{9}y_k = 0$

14. $y_{k+2} - 5y_{k+1} + 6y_k = 0$

15. $6y_{k+2} + y_{k+1} - 2y_k = 0$

16. $y_{k+2} - 25y_k = 0$

Los ejercicios 17 y 18 se refieren a un modelo sencillo de la economía nacional descrito por la ecuación en diferencias

$$Y_{k+2} - a(1+b)Y_{k+1} + abY_k = 1 \qquad (14)$$

Aquí Y_k es el ingreso nacional total durante el año k, a es una constante menor que 1, que se llama *propensión marginal al consumo*, y b es una *constante de ajuste* positiva que describe cómo los cambios en el gasto de consumo afectan la tasa anual de inversión privada.[1]

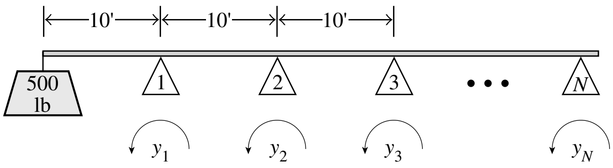
17. Determine la solución general de la ecuación (14) cuando $a = .9$ y $b = \frac{4}{9}$. ¿Qué pasa con Y_k conforme k aumenta? [*Sugerencia:* En primer lugar, encuentre una solución particular de la forma $Y_k = T$, donde T es una constante, llamada el nivel de equilibrio del ingreso nacional].

18. Determine la solución general de la ecuación (14), cuando $a = .9$ y $b = .5$.

[1] Por ejemplo, véase *Discrete Dynamical Systems*, de James T. Sandefur (Oxford: Clarendon Press, 1990), pp. 267-276. El *modelo original acelerador-multiplicador* se atribuye al economista P. A. Samuelson.

Una viga ligera en voladizo está soportada en los puntos N espaciados por una distancia de 10 ft, y se coloca un peso de 500 lb en el extremo de la viga, a 10 ft del primer soporte, como se muestra en la figura. Sea y_k el momento flector en el k-ésimo soporte. Luego, $y_1 = 5000$ ft-lb. Suponga que la viga está rígidamente sujeta al N-ésimo soporte y el momento flector ahí es cero. En medio, los momentos satisfacen la *ecuación de tres momentos*

$$y_{k+2} + 4y_{k+1} + y_k = 0 \quad \text{para } k = 1, 2, \ldots, N-2 \tag{15}$$



Momentos flectores en una viga en voladizo.

19. Determine la solución general de la ecuación en diferencias (15). Justifique su respuesta.

20. Encuentre la solución particular de la ecuación (15) que satisface las *condiciones frontera* $y_1 = 5000$ y $y_N = 0$. (La respuesta implica a N).

21. Cuando se produce una señal a partir de una secuencia de mediciones efectuadas en un proceso (una reacción química, un flujo de calor a través de un tubo, un brazo robot en movimiento, etcétera), la señal, por lo general, contiene *ruido* aleatorio producido por errores de medición. Un método estándar de procesamiento previo de los datos para reducir el ruido es suavizar o filtrar los datos. Un sencillo filtro es un *promedio móvil* que sustituye cada y_k por su promedio con los dos valores adyacentes:

$$\tfrac{1}{3}y_{k+1} + \tfrac{1}{3}y_k + \tfrac{1}{3}y_{k-1} = z_k \quad \text{para } k = 1, 2, \ldots$$

Suponga que una señal y_k, para $k = 0, \ldots, 14$, es

9, 5, 7, 3, 2, 4, 6, 5, 7, 6, 8, 10, 9, 5, 7

Utilice el filtro para calcular $z_1, \ldots, z_{13}$. Trace una gráfica de línea discontinua que se superponga a la señal original y a la señal suavizada.

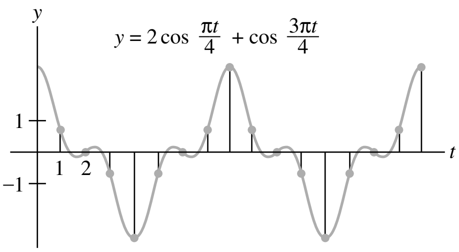
22. Sea $\{y_k\}$ la secuencia producida por la toma de muestras de señal continua $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$ en $t = 0, 1, 2, \ldots$, como se indica en la figura. Los valores de y_k, empezando con $k = 0$, son

$3, .7, 0, -.7, -3, -.7, 0, .7, 3, .7, 0, \ldots$

donde .7 es una abreviatura de $\sqrt{2}/2$.

a) Calcule la señal de salida $\{z_k\}$ cuando $\{y_k\}$ alimenta al filtro del ejemplo 3.

b) Explique cómo y por qué la salida en el inciso *a*) está relacionada con los cálculos del ejemplo 3.



Datos de la muestra de $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$.

Los ejercicios 23 y 24 se refieren a una ecuación en diferencias de la forma $y_{k+1} - ay_k = b$, para constantes adecuadas a y b.

23. Un préstamo de \$10,000 tiene una tasa de interés del 1% mensual y requiere un pago mensual de \$450. El préstamo se realiza en el mes $k = 0$, y el primer pago se realiza un mes después, en $k = 1$. Para $k = 0, 1, 2, \ldots$, sea y_k el saldo insoluto del préstamo justo después del k-ésimo pago mensual. Así,

$$\underset{\substack{\text{Nuevo}\\ \text{saldo}}}{y_1} = \underset{\substack{\text{Saldo}\\ \text{adeudado}}}{10{,}000} + \underset{\substack{\text{Intereses}\\ \text{agregados}}}{(.01)10{,}000} - \underset{\text{Pago}}{450}$$

a) Escriba una ecuación en diferencias satisfecha por $\{y_k\}$.

b) **[M]** Cree una tabla que muestre k, y el saldo y_k al mes k. Liste el programa o los tecleos (las instrucciones) para crear la tabla.

c) **[M]** ¿Cuál es el valor de k cuando se efectúa el último pago? ¿De cuánto será el último pago? ¿Cuánto dinero en total pagó el deudor?

24. En el tiempo $k = 0$, con una inversión inicial de \$1000 se abre una cuenta de ahorros que paga el 6% de interés anual con capitalización mensual. (La tasa de interés mensual es de .005). Cada mes después de la inversión inicial, se agregan \$200 a la cuenta. Para $k = 0, 1, 2, \ldots$, sea y_k la cantidad que hay en la cuenta al momento k, justo después de que se hace un depósito.

a) Escriba una ecuación en diferencias que satisfaga $\{y_k\}$.

b) **[M]** Cree una tabla que muestre k, y la cantidad total en la cuenta de ahorros en el mes k, para $k = 0$ a 60. Liste su programa o los tecleos (las instrucciones) que utilizó para crear la tabla.

c) **[M]** ¿Cuánto habrá en la cuenta después de dos años (es decir, 24 meses), cuatro años, y cinco años? ¿Qué parte del total a los cinco años corresponderá a los intereses?

En los ejercicios 25 a 28, demuestre que la señal dada es una solución de la ecuación en diferencias. Luego, determine la solución general de esa ecuación en diferencias.

25. $y_k = k^2;\ y_{k+2} + 3y_{k+1} - 4y_k = 7 + 10k$

26. $y_k = 1 + k;\ y_{k+2} - 6y_{k+1} + 5y_k = -4$

27. $y_k = k - 2;\ y_{k+2} - 4y_k = 8 - 3k$

28. $y_k = 1 + 2k;\ y_{k+2} - 25y_k = -48k - 20$

Escriba las ecuaciones en diferencias en los ejercicios 29 y 30 como sistemas de primer orden, $\mathbf{x}_{k+1} = A\mathbf{x}_k$, para toda k.

29. $y_{k+4} + 3y_{k+3} - 8y_{k+2} + 6y_{k+1} - 2y_k = 0$

30. $y_{k+3} - 5y_{k+2} + 8y_k = 0$

31. ¿La siguiente ecuación en diferencias es de orden 3? Explique su respuesta.

$$y_{k+3} + 5y_{k+2} + 6y_{k+1} = 0$$

32. ¿Cuál es el orden de la siguiente ecuación en diferencias? Explique su respuesta.

$$y_{k+3} + a_1 y_{k+2} + a_2 y_{k+1} + a_3 y_k = 0$$

33. Sea $y_k = k^2$ y $z_k = 2k|k|$. ¿Las señales $\{y_k\}$ y $\{z_k\}$ son linealmente independientes? Evalúe la matriz de Casorati $C(k)$ asociada para $k = 0$, $k = -1$ y $k = -2$, y analice sus resultados.

34. Sean f, g y h, funciones linealmente independientes definidas para todos los números reales, y construya tres señales por muestreo de los valores de las funciones de los números enteros:

$$u_k = f(k), \qquad v_k = g(k), \qquad w_k = h(k)$$

¿Es necesario que las señales sean linealmente independientes en $\mathbb{S}$? Analícelo.

35. Sean a y b números diferentes de cero. Demuestre que el mapeo T definido por $T\{y_k\} = \{w_k\}$, donde

$$w_k = y_{k+2} + ay_{k+1} + by_k$$

es una transformación lineal de $\mathbb{S}$ en $\mathbb{S}$.

36. Sea V un espacio vectorial, y sea $T : V \to V$ una transformación lineal. Dada $\mathbf{z}$ en V, suponga que $\mathbf{x}_p$ en V satisface $T(\mathbf{x}_p) = \mathbf{z}$, y sea $\mathbf{u}$ cualquier vector en el núcleo de T. Demuestre que $\mathbf{u} + \mathbf{x}_p$ satisface la ecuación no homogénea $T(\mathbf{x}) = \mathbf{z}$.

37. Sea $\mathbb{S}_0$ el espacio vectorial de todas las secuencias de la forma $(y_0, y_1, y_2, \ldots)$, y defina las transformaciones lineales T y D de $\mathbb{S}_0$ dentro de $\mathbb{S}_0$ por

$$T(y_0, y_1, y_2, \ldots) = (y_1, y_2, y_3, \ldots)$$
$$D(y_0, y_1, y_2, \ldots) = (0, y_0, y_1, y_2, \ldots)$$

Demuestre que $TD = I$ (la transformación identidad en $\mathbb{S}_0$) y que, sin embargo, $DT \neq I$.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Examine la matriz de Casorati:

$$C(k) = \begin{bmatrix} 2^k & 3^k \operatorname{sen} \frac{k\pi}{2} & 3^k \cos \frac{k\pi}{2} \\ 2^{k+1} & 3^{k+1} \operatorname{sen} \frac{(k+1)\pi}{2} & 3^{k+1} \cos \frac{(k+1)\pi}{2} \\ 2^{k+2} & 3^{k+2} \operatorname{sen} \frac{(k+2)\pi}{2} & 3^{k+2} \cos \frac{(k+2)\pi}{2} \end{bmatrix}$$

Sea $k = 0$ y reduzca por filas la matriz para comprobar que tiene tres posiciones pivote y que, por lo tanto, es invertible:

$$C(0) = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 0 \\ 4 & 0 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 0 & 0 & -13 \end{bmatrix}$$

La matriz de Casorati es invertible en $k = 0$, por lo que las señales son linealmente independientes. Puesto que hay tres señales, y el espacio solución H de la ecuación en diferencias es de dimensión 3 (teorema 17), las señales forman una base para H, de acuerdo con el teorema de la base.

4.9 APLICACIONES A CADENAS DE MARKOV

Las cadenas de Markov descritas en esta sección se utilizan como modelos matemáticos de una amplia variedad de situaciones en biología, negocios, química, ingeniería, física y otros campos. En cada caso, el modelo se utiliza para describir un experimento o una medición que se realiza muchas veces de la misma manera, cuando el resultado de cada ensayo del experimento será uno de varios posibles resultados especificados, y depende solo del ensayo inmediato anterior.

Por ejemplo, si la población de una ciudad y sus suburbios se midieran cada año, entonces un vector tal como

$$\mathbf{x}_0 = \begin{bmatrix} .60 \\ .40 \end{bmatrix} \tag{1}$$

podría indicar que el 60% de la población vive en la ciudad y el 40% en los suburbios. Los decimales en $\mathbf{x}_0$ suman 1, ya que dan cuenta de toda la población de la región. Los porcentajes son más convenientes para nuestros propósitos en este momento que los totales de población.

Un vector con entradas no negativas que suman 1 se llama **vector de probabilidad**. Una **matriz estocástica** es una matriz cuadrada cuyas columnas son vectores de probabilidad. Una **cadena de Markov** es una secuencia de vectores de probabilidad de $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$, junto con una matriz estocástica P, tal que

$$\mathbf{x}_1 = P\mathbf{x}_0, \quad \mathbf{x}_2 = P\mathbf{x}_1, \quad \mathbf{x}_3 = P\mathbf{x}_2, \quad \ldots$$

Así, la cadena de Markov se describe mediante la ecuación en diferencias de primer orden

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

Cuando una cadena de Markov de vectores en $\mathbb{R}^n$ describe un sistema o una serie de experimentos, las entradas en $\mathbf{x}_k$ listan, respectivamente, las probabilidades de que el sistema esté en cada uno de los n posibles estados, o las probabilidades de que el resultado del experimento sea uno de los n posibles resultados. Por esta razón, $\mathbf{x}_k$ con frecuencia se llama un **vector de estado**.

EJEMPLO 1 En la sección 1.10 se examinó un modelo para los movimientos de población entre una ciudad y sus suburbios. Véase la figura 1. La migración anual entre estas dos partes de la zona metropolitana se regía por la *matriz de migración M*:

$$\begin{array}{ccl} & \begin{array}{cc} \multicolumn{2}{c}{\text{De:}} \\ \text{Ciudad} & \text{Suburbios} \end{array} & \text{De:} \\ M = & \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} & \begin{array}{l} \text{Ciudad} \\ \text{Suburbios} \end{array} \end{array}$$

Es decir, cada año el 5% de la población de la ciudad se muda a los suburbios, y el 3% de la población de los suburbios se muda a la ciudad. Las columnas de M son vectores de probabilidad, por lo que M es una matriz estocástica. Suponga la población de la región en el año 2000 era de 600,000 en la ciudad y de 400,000 en los suburbios. Entonces, la distribución inicial de la población en la región está dada por $\mathbf{x}_0$ en la ecuación (1) anterior. ¿Cuál es la distribución de la población en el año 2001? ¿Y en 2002?

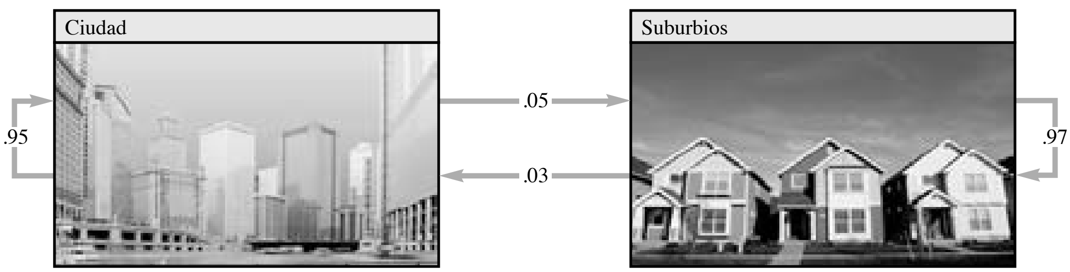


FIGURA 1 Migración porcentual anual entre la ciudad y los suburbios.

SOLUCIÓN En el ejemplo 3 de la sección 1.10, vimos que después de un año, el vector de población $\begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$ cambió a

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

Si dividimos ambos lados de esta ecuación entre la población total de un millón y consideramos el hecho de que $kM\mathbf{x} = M(k\mathbf{x})$, se encuentra que

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .600 \\ .400 \end{bmatrix} = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$$

El vector $\mathbf{x}_1 = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$ da la distribución de la población en 2001. Es decir, el 58.2% de la región vivía en la ciudad y el 41.8% vivía en los suburbios. Del mismo modo, la distribución de la población en el año 2002 se describe por un vector $\mathbf{x}_2$, donde

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .582 \\ .418 \end{bmatrix} = \begin{bmatrix} .565 \\ .435 \end{bmatrix}$$ ■

EJEMPLO 2 Suponga que el resultado de la votación durante las elecciones de un congreso en un distrito de votación está representado por un vector $\mathbf{x}$ en $\mathbb{R}^3$:

$$\mathbf{x} = \begin{bmatrix} \text{\% de votación por los demócratas (D)} \\ \text{\% de votación por los republicanos (R)} \\ \text{\% de votación por los liberales (L)} \end{bmatrix}$$

Suponga que se registra el resultado de las elecciones al congreso cada dos años mediante un vector de este tipo y que el resultado de una elección solo depende de los resultados de la elección anterior. Entonces, la secuencia de los vectores que describen los votos cada dos años puede ser una cadena de Markov. Como un ejemplo de una matriz estocástica P para esta cadena, tomamos

$$\begin{array}{c} \text{De:} \\ \begin{array}{ccc} \text{D} & \text{R} & \text{L} \end{array} \\ P = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \end{array} \begin{array}{l} \text{A:} \\ \text{D} \\ \text{R} \\ \text{L} \end{array}$$

Las entradas en la primera columna, denotada como D, describen qué van a hacer en las próximas elecciones las personas que votaron por los demócratas en una elección. Aquí supusimos que el 70% votará de nuevo por los demócratas en las próximas elecciones, el 20% votará por los republicanos, y el 10% dará su voto a los liberales. Similares interpretaciones valen para las otras columnas de P. Un diagrama de esta matriz se presenta en la figura 2.

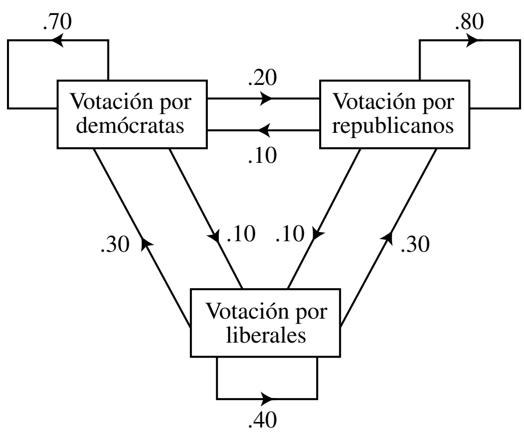


FIGURA 2 Cambios en la votación de una elección a la siguiente.

Si los porcentajes de "transición" se mantienen constantes a lo largo de muchos años de una elección a otra, entonces la secuencia de los vectores que dan los resultados de la votación constituye una cadena de Markov. Suponga que el resultado de una elección está dado por

$$\mathbf{x}_0 = \begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix}$$

Determine el resultado probable de la siguiente elección y el resultado probable de la elección posterior.

SOLUCIÓN El resultado de la siguiente elección se describe mediante el vector de estado $\mathbf{x}_1$ y el de la elección posterior por $\mathbf{x}_2$, donde

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix}\begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix} = \begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} \quad \begin{matrix} \text{El 44\% votará por D} \\ \text{El 44.5\% votará por R} \\ \text{El 11.5\% votará por L} \end{matrix}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix}\begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} = \begin{bmatrix} .3870 \\ .4785 \\ .1345 \end{bmatrix} \quad \begin{matrix} \text{El 38.7\% votará por D} \\ \text{El 47.8\% votará por R} \\ \text{El 13.5\% votará por L} \end{matrix}$$

Para entender por qué $\mathbf{x}_1$, efectivamente, da el resultado de las próximas elecciones, suponga que 1000 personas votaron en la "primera" elección, con 550 votos a favor de D, 400 votos a favor de R y 50 votos a favor de L. (Véase los porcentajes en $\mathbf{x}_0$). En la siguiente elección, el 70% de los 550 votarán de nuevo por D, el 10% de los 400 cambiará su voto de R a D, y el 30% de los 50 cambiará de L a D. De esta forma, el total de votos para D será

$$.70(550) + .10(400) + .30(50) = 385 + 40 + 15 = 440 \tag{2}$$

Así, el 44% de la próxima votación será a favor del candidato demócrata. El cálculo en (2) es, en esencia, el mismo que se utilizó para calcular la primera entrada en $\mathbf{x}_1$. Se podrían hacer cálculos análogos para las otras entradas en $\mathbf{x}_1$, para las entradas en $\mathbf{x}_2$, y así sucesivamente. ■

Predicciones del futuro lejano

El aspecto más interesante de las cadenas de Markov es el estudio del comportamiento a largo plazo de una cadena. Por ejemplo, ¿qué se puede decir en el ejemplo 2 de la votación después de que han pasado muchas elecciones (suponiendo que la matriz estocástica dada continúe describiendo los porcentajes de transición de una elección a la siguiente)? O, ¿qué sucede con la distribución de la población en el ejemplo 1 "a largo plazo"? Antes de responder estas preguntas, veamos un ejemplo numérico.

EJEMPLO 3 Sea $P = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Considere un sistema cuyo estado está descrito por la cadena de Markov $\mathbf{x}_{k+1} = P\mathbf{x}_k$, para $k = 0, 1, \ldots$ ¿Qué sucede con el sistema a medida que transcurre el tiempo? Para averiguarlo, calcule los vectores de estado de $\mathbf{x}_1, \ldots, \mathbf{x}_{15}$.

SOLUCIÓN

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix} = \begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix}$$

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix} = \begin{bmatrix} .329 \\ .525 \\ .146 \end{bmatrix}$$

Los resultados de los cálculos posteriores se muestran a continuación, con las entradas redondeadas a cuatro o cinco cifras significativas.

$$\mathbf{x}_4 = \begin{bmatrix} .3133 \\ .5625 \\ .1242 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} .3064 \\ .5813 \\ .1123 \end{bmatrix}, \quad \mathbf{x}_6 = \begin{bmatrix} .3032 \\ .5906 \\ .1062 \end{bmatrix}, \quad \mathbf{x}_7 = \begin{bmatrix} .3016 \\ .5953 \\ .1031 \end{bmatrix}$$

$$\mathbf{x}_8 = \begin{bmatrix} .3008 \\ .5977 \\ .1016 \end{bmatrix}, \quad \mathbf{x}_9 = \begin{bmatrix} .3004 \\ .5988 \\ .1008 \end{bmatrix}, \quad \mathbf{x}_{10} = \begin{bmatrix} .3002 \\ .5994 \\ .1004 \end{bmatrix}, \quad \mathbf{x}_{11} = \begin{bmatrix} .3001 \\ .5997 \\ .1002 \end{bmatrix}$$

$$\mathbf{x}_{12} = \begin{bmatrix} .30005 \\ .59985 \\ .10010 \end{bmatrix}, \quad \mathbf{x}_{13} = \begin{bmatrix} .30002 \\ .59993 \\ .10005 \end{bmatrix}, \quad \mathbf{x}_{14} = \begin{bmatrix} .30001 \\ .59996 \\ .10002 \end{bmatrix}, \quad \mathbf{x}_{15} = \begin{bmatrix} .30001 \\ .59998 \\ .10001 \end{bmatrix}$$

Estos vectores parecen acercarse a $\mathbf{q} = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$. Las probabilidades son difíciles de cambiar de un valor de k al siguiente. Observe que el siguiente cálculo es exacto (sin errores de redondeo):

$$P\mathbf{q} = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix} \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix} = \begin{bmatrix} .15 + .12 + .03 \\ .09 + .48 + .03 \\ .06 + 0 + .04 \end{bmatrix} = \begin{bmatrix} .30 \\ .60 \\ .10 \end{bmatrix} = \mathbf{q}$$

Cuando el sistema está en estado $\mathbf{q}$, no hay ningún cambio en el sistema de una medición a la siguiente. ■

Vectores de estado estable

Si P es una matriz estocástica, entonces un **vector de estado estable** (o **vector de equilibrio**) de P es un vector de probabilidad $\mathbf{q}$ tal que

$$P\mathbf{q} = \mathbf{q}$$

Es posible demostrar que toda matriz estocástica tiene un vector de estado estable. En el ejemplo 3, $\mathbf{q}$ es un vector de estado estable de P.

EJEMPLO 4 El vector de probabilidad $\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$ es un vector de estado estable para la matriz de migración de población M en el ejemplo 1, porque

$$M\mathbf{q} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} .375 \\ .625 \end{bmatrix} = \begin{bmatrix} .35625 + .01875 \\ .01875 + .60625 \end{bmatrix} = \begin{bmatrix} .375 \\ .625 \end{bmatrix} = \mathbf{q}$$ ■

Si el total de la población de la región metropolitana en el ejemplo 1 es de 1 millón, entonces **q** del ejemplo 4 correspondería a tener 375,000 personas en la ciudad y 625,000 en los suburbios. Al final de un año, la *emigración* de la ciudad sería (.05)(375,000) = 18,750 personas, y la *inmigración* a la ciudad proveniente de los suburbios sería (.03)(625,000) = 18,750 personas. Como resultado, la población de la ciudad seguiría siendo la misma. Asimismo, la población de los suburbios se mantendría estable.

El siguiente ejemplo muestra cómo *encontrar* un vector de estado estable.

EJEMPLO 5 Sea $P = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$. Encuentre un vector de estado estable para P.

SOLUCIÓN En primer lugar, se resuelve la ecuación $P\mathbf{x} = \mathbf{x}$.

$$\begin{aligned} P\mathbf{x} - \mathbf{x} &= \mathbf{0} \\ P\mathbf{x} - I\mathbf{x} &= \mathbf{0} \qquad \text{Recuerde de la sección 1.4 que } I\mathbf{x} = \mathbf{x}. \\ (P - I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

Para la P anterior,

$$P - I = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -.4 & .3 \\ .4 & -.3 \end{bmatrix}$$

Para encontrar todas las soluciones de $(P - I)\mathbf{x} = 0$, se reduce por filas la matriz aumentada:

$$\begin{bmatrix} -.4 & .3 & 0 \\ .4 & -.3 & 0 \end{bmatrix} \sim \begin{bmatrix} -.4 & .3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3/4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Entonces $x_1 = \frac{3}{4}x_2$ y x_2 es libre. La solución general es $x_2 \begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$.

A continuación, se elige una base sencilla para el espacio solución. Una elección evidente es $\begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$, pero una mejor opción, sin fracciones, es $\mathbf{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ (correspondiente a $x_2 = 4$).

Por último, encuentre un vector de probabilidad en el conjunto de todas las soluciones de $P\mathbf{x} = \mathbf{x}$. Este proceso es fácil, ya que cada solución es un múltiplo de la solución **w** anterior. Divida **w** entre la suma de sus entradas para obtener

$$\mathbf{q} = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$$

Para comprobar, calcule

$$P\mathbf{q} = \begin{bmatrix} 6/10 & 3/10 \\ 4/10 & 7/10 \end{bmatrix} \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix} = \begin{bmatrix} 18/70 + 12/70 \\ 12/70 + 28/70 \end{bmatrix} = \begin{bmatrix} 30/70 \\ 40/70 \end{bmatrix} = \mathbf{q} \qquad \blacksquare$$

El siguiente teorema demuestra que lo que sucedió en el ejemplo 3 es característico de muchas matrices estocásticas. Decimos que una matriz estocástica es **regular** si alguna potencia de la matriz P^k contiene solo entradas estrictamente positivas. Para P en el ejemplo 3,

$$P^2 = \begin{bmatrix} .37 & .26 & .33 \\ .45 & .70 & .45 \\ .18 & .04 & .22 \end{bmatrix}$$

Ya que cada entrada en P^2 es estrictamente positiva, P es una matriz estocástica regular.

Además, se dice que una secuencia de vectores $\{\mathbf{x}_k\colon k = 1, 2, \dots\}$ **converge** a un vector **q** conforme $k \to \infty$ si las entradas en $\mathbf{x}_k$ se pueden acercar tanto como se desee a las entradas correspondientes en **q** al hacer a k suficientemente grande.

TEOREMA 18

> Si P es una matriz de $n \times n$ estocástica regular, entonces P tiene un único vector de estado estable $\mathbf{q}$. Además, si $\mathbf{x}_0$ es cualquier estado inicial y una $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, 2, \ldots$, entonces la cadena de Markov $\{\mathbf{x}_k\}$ converge a $\mathbf{q}$ conforme $k \to \infty$.

Este teorema se demuestra en los libros que analizan cadenas de Markov. La parte sorprendente del teorema es que el estado inicial no tiene ningún efecto sobre el comportamiento a largo plazo de la cadena de Markov. Se verá más adelante (en la sección 5.2) por qué esto es cierto para varias matrices estocásticas estudiadas aquí.

EJEMPLO 6 En el ejemplo 2, ¿qué porcentaje de los electores tienden a votar por el candidato republicano en alguna elección muchos años después a partir de ahora, suponiendo que el resultado de las elecciones forman una cadena de Markov?

SOLUCIÓN Para los cálculos a mano, el enfoque *equivocado* es elegir algún vector inicial $\mathbf{x}_0$ y calcular $\mathbf{x}_1, \ldots, \mathbf{x}_k$ para algún valor grande de k. No hay manera de saber cuántos vectores habrá que calcular, y no se puede estar seguro de los valores límite de las entradas en $\mathbf{x}_k$.

El enfoque correcto es calcular el vector de estado estable y, luego, recurrir al teorema 18. Dado P como en el ejemplo 2, se forma $P - I$ restando 1 de cada entrada diagonal en P. Después se reduce por filas la matriz aumentada:

$$[\,(P - I) \quad \mathbf{0}\,] = \begin{bmatrix} -.3 & .1 & .3 & 0 \\ .2 & -.2 & .3 & 0 \\ .1 & .1 & -.6 & 0 \end{bmatrix}$$

Recuerde, de trabajos anteriores con decimales, que el cálculo se simplifica al multiplicar cada fila por 10.[1]

$$\begin{bmatrix} -3 & 1 & 3 & 0 \\ 2 & -2 & 3 & 0 \\ 1 & 1 & -6 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -9/4 & 0 \\ 0 & 1 & -15/4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general de $(P - I)\mathbf{x} = \mathbf{0}$ es $x_1 = \frac{9}{4}x_3$, $x_2 = \frac{15}{4}x_3$ y x_3 es libre. Eligiendo $x_3 = 4$, se obtiene una base para el espacio solución cuyas entradas son enteros, y a partir de ello se encuentra fácilmente el vector de estado estable cuyas entradas suman 1:

$$\mathbf{w} = \begin{bmatrix} 9 \\ 15 \\ 4 \end{bmatrix} \quad \text{y} \quad \mathbf{q} = \begin{bmatrix} 9/28 \\ 15/28 \\ 4/28 \end{bmatrix} \approx \begin{bmatrix} .32 \\ .54 \\ .14 \end{bmatrix}$$

Las entradas en $\mathbf{q}$ describen la distribución de los votos en una elección que se celebrará dentro de muchos años (suponiendo que la matriz estocástica continúa describiendo los cambios de una elección a otra). Así, finalmente, alrededor del 54% de los votos serán para el candidato republicano. ■

[1] *Advertencia:* No solo multiplique P por 10. En vez de ello, multiplique la matriz aumentada para la ecuación $(P - I)\mathbf{x} = 0$ por 10.

NOTA NUMÉRICA

Quizás haya notado que si $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, \ldots$ entonces

$$\mathbf{x}_2 = P\mathbf{x}_1 = P(P\mathbf{x}_0) = P^2\mathbf{x}_0,$$

y, en general,

$$\mathbf{x}_k = P^k\mathbf{x}_0 \quad \text{para } k = 0, 1, \ldots$$

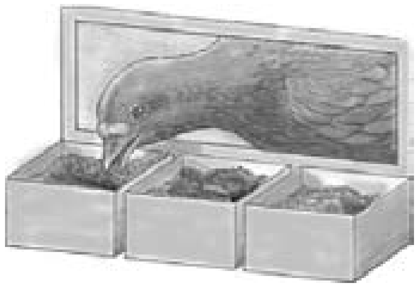
Para calcular un vector específico, tal como $\mathbf{x}_3$, se necesitan menos operaciones aritméticas para calcular $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$, más que P^3 y $P^3\mathbf{x}_0$. Sin embargo, si P es pequeña, por ejemplo, 30×30, el tiempo de cálculo de la máquina es insignificante con ambos métodos, y tal vez se prefiera un comando para calcular $P^3\mathbf{x}_0$, ya que requiere teclear menos.

PROBLEMAS DE PRÁCTICA

1. Suponga que los habitantes de una región metropolitana se desplazan de acuerdo con las probabilidades de la matriz de migración M del ejemplo 1, y que se elige un residente “al azar”. Entonces, un vector de estado para un año determinado se puede interpretar como un indicador de las probabilidades de que la persona sea residente de la ciudad o de los suburbios en ese momento.
   a) Suponga que la persona elegida es un residente de la ciudad ahora, de manera que $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. ¿Cuál es la probabilidad de que esa persona viva en los suburbios el año próximo?
   b) ¿Cuál es la probabilidad de que esa persona viva en los suburbios en dos años?
2. Sea $P = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix}$ y $\mathbf{q} = \begin{bmatrix} .3 \\ .7 \end{bmatrix}$. ¿Es $\mathbf{q}$ un vector de estado estable para P?
3. ¿Qué porcentaje de la población en el ejemplo 1 vivirá en los suburbios después de muchos años?

4.9 EJERCICIOS

1. Un pueblo remoto y pequeño recibe transmisiones de radio de dos estaciones radiofónicas: una de noticias y una emisora de música. De los oyentes que sintonizan la emisora de noticias, el 70% seguirá escuchando las noticias después del corte de estación que ocurre cada media hora, mientras que el 30% sintonizará la estación de música durante el corte de estación. De los oyentes que sintonizan la estación de música, el 60% cambiará a la estación de noticias en el corte de estación, mientras que el 40% permanecerá escuchando la música. Suponga que todos los habitantes están escuchando las noticias a las 8:15 A.M.
   a) Dé la matriz estocástica que describe cómo los oyentes de radio tienden a cambiar las estaciones en cada corte de estación. Rotule las filas y las columnas.
   b) Dé el vector de estado inicial.
   c) ¿Qué porcentaje de los oyentes estará escuchando la emisora de música a las 9:25 A.M. (después de los cortes de estación de las 8:30 y de las 9:00 A.M.)?
2. Un animal de laboratorio puede comer cualquiera de tres alimentos cada día. Los registros de laboratorio indican que si el animal elige un alimento en un ensayo, elegirá el mismo alimento en el siguiente ensayo con una probabilidad del 60%, y elegirá cualquiera de los otros alimentos en el siguiente ensayo con iguales probabilidades del 20%.
   a) ¿Cuál es la matriz estocástica de esta situación?
   b) Si el animal elige el alimento #1 en un ensayo inicial, ¿cuál es la probabilidad de que elija el alimento #2 en el segundo ensayo después del inicial?



3. En un día cualquiera, un estudiante está sano o enfermo. De los estudiantes que están sanos hoy, el 95% estarán sanos mañana.

De los estudiantes que están enfermos hoy, el 55% seguirán enfermos mañana.

a) ¿Cuál es la matriz estocástica de esta situación?

b) Suponga que el 20% de los alumnos están enfermos el lunes. ¿Qué fracción o porcentaje de los estudiantes es probable que estén enfermos el martes? ¿Y el miércoles?

c) Si un estudiante está sano hoy, ¿cuál es la probabilidad de que esté sano dentro de dos días?

4. El clima en Columbus es bueno, regular o malo en un día determinado. Si el clima es bueno hoy, hay un 40% de probabilidad de que sea bueno mañana, un 30% de probabilidad de que sea regular, y un 30% de que sea malo. Si el clima es regular hoy, existe un 50% de probabilidad de que sea bueno mañana, y un 20% de probabilidad de que sea regular. Por último, si el clima es malo hoy, existe un 30% de probabilidad de que sea bueno mañana y un 40% de que sea regular.

a) ¿Cuál es la matriz estocástica de esta situación?

b) Suponga que hay una probabilidad del 50% de buen clima hoy, y una probabilidad del 50% de clima regular. ¿Cuáles son las probabilidades de que el clima sea malo mañana?

c) Suponga que, de acuerdo con los pronósticos para el lunes, hay un 60% de probabilidad de que el clima sea regular y un 40% de que sea malo. ¿Cuáles son las probabilidades de tener buen clima el miércoles?

En los ejercicios 5 a 8, encuentre el vector de estado estable.

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

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

7. $\begin{bmatrix} .7 & .1 & .1 \\ .2 & .8 & .2 \\ .1 & .1 & .7 \end{bmatrix}$

8. $\begin{bmatrix} .4 & .5 & .8 \\ 0 & .5 & .1 \\ .6 & 0 & .1 \end{bmatrix}$

9. Determine si $P = \begin{bmatrix} .2 & 1 \\ .8 & 0 \end{bmatrix}$ es una matriz estocástica regular.

10. Determine si $P = \begin{bmatrix} 1 & .3 \\ 0 & .7 \end{bmatrix}$ es una matriz estocástica regular.

11. *a*) Encuentre el vector de estado estable para la cadena de Markov en el ejercicio 1.

b) En algún momento al final del día, ¿qué fracción de los oyentes escuchará las noticias?

12. Consulte el ejercicio 2. ¿Qué alimento prefiere el animal después de muchos ensayos?

13. *a*) Encuentre el vector de estado estable para la cadena de Markov en el ejercicio 3.

b) ¿Cuál es la probabilidad de que, después de muchos días, un estudiante en particular esté enfermo? ¿Importa si esa persona está enferma hoy?

14. Consulte el ejercicio 4. En el largo plazo, ¿qué tan probable es que el clima en Columbus sea bueno en un día determinado?

15. [**M**] La Unidad de Investigación Demográfica del Departamento de Finanzas del Estado de California proporcionó los siguientes datos para la matriz de migración, la cual describe el movimiento de la población de Estados Unidos durante 1989. En ese año, alrededor del 11.7% de la población total vivía en California. ¿Qué porcentaje de la población total con el tiempo vivirá en California si las probabilidades de migración mencionadas se mantuvieran constantes durante muchos años?

$$\begin{array}{cc} \text{De:} & \\ \begin{array}{cc} \text{CA} & \text{Resto de E.U.A.} \end{array} & \text{A:} \\ \begin{bmatrix} .9821 & .0029 \\ .0179 & .9971 \end{bmatrix} & \begin{array}{l} \text{California} \\ \text{Resto de E.U.A.} \end{array} \end{array}$$

16. [**M**] En Detroit, Hertz Rent A Car cuenta con una flota de cerca de 2000 automóviles. El patrón de puntos de alquiler y devolución de las unidades está descrito por las fracciones en la siguiente tabla. En un día típico, ¿cuántos autos estarán listos para rentarse en la sucursal ubicada en el centro de la ciudad?

$$\begin{array}{cc} \text{Autos rentados en:} & \\ \begin{array}{ccc} \text{Aeropuerto de la cd.} & \text{Centro} & \text{Apto. fuera de cd.} \end{array} & \text{Devueltos en:} \\ \begin{bmatrix} .90 & .01 & .09 \\ .01 & .90 & .01 \\ .09 & .09 & .90 \end{bmatrix} & \begin{array}{l} \text{Aeropuerto de la cd.} \\ \text{Centro} \\ \text{Apto. fuera de cd.} \end{array} \end{array}$$

17. Sea P una matriz estocástica de $n \times n$. El siguiente argumento indica que la ecuación $P\mathbf{x} = \mathbf{x}$ tiene una solución no trivial. (De hecho, una solución de estado estable existe con entradas no negativas. En algunos libros de texto avanzados se da una demostración). Justifique cada una de las siguientes afirmaciones. (Mencione un teorema cuando sea pertinente).

a) Si todas las otras filas de $P - I$ se suman a la fila de abajo, el resultado es una fila de ceros.

b) Las filas de $P - I$ son linealmente dependientes.

c) La dimensión del espacio fila de $P - I$ es menor que n.

d) $P - I$ tiene un espacio nulo no trivial.

18. Demuestre que toda matriz estocástica de 2×2 tiene al menos un vector de estado estable. Cualquier matriz de este tipo se puede representar en la forma $P = \begin{bmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{bmatrix}$, donde α y β son constantes entre 0 y 1. (Hay dos vectores de estado estable, linealmente independientes, si $\alpha = \beta = 0$. De lo contrario, solo hay uno).

19. Sea S una matriz fila de $1 \times n$ con un 1 en cada columna, $S = [1 \quad 1 \quad \cdots \quad 1]$

a) Explique por qué un vector $\mathbf{x}$ en $\mathbb{R}^n$ es un vector de probabilidad si y solo si sus entradas son no negativas y $S\mathbf{x} = 1$. (Una matriz de 1×1 como el producto $S\mathbf{x}$ generalmente se representa sin los símbolos de corchete de la matriz).

b) Sea P una matriz estocástica de $n \times n$. Explique por qué $SP = S$.

c) Sea P una matriz estocástica de $n \times n$, y sea $\mathbf{x}$ un vector de probabilidad. Demuestre que $P\mathbf{x}$ es también un vector de probabilidad.

20. Con base en el ejercicio 19, demuestre que si P es una matriz estocástica de $n \times n$, entonces también lo es P^2.

21. [**M**] Examine las potencias de una matriz estocástica regular.

a) Calcule P^k para $k = 2, 3, 4, 5$, cuando

$$P = \begin{bmatrix} .3355 & .3682 & .3067 & .0389 \\ .2663 & .2723 & .3277 & .5451 \\ .1935 & .1502 & .1589 & .2395 \\ .2047 & .2093 & .2067 & .1765 \end{bmatrix}$$

Presente los cálculos con cuatro decimales. ¿Qué ocurre a las columnas de P^k conforme k aumenta? Calcule el vector de estado estable para P.

b) Calcule Q^k para $k = 10, 20, \ldots, 80$, cuando

$$Q = \begin{bmatrix} .97 & .05 & .10 \\ 0 & .90 & .05 \\ .03 & .05 & .85 \end{bmatrix}$$

(La estabilidad de Q^k con cuatro decimales puede requerir $k = 116$ o más). Calcule el vector de estado estable para Q. Haga una conjetura sobre lo que podría ser cierto para cualquier matriz estocástica regular.

c) Con base en el teorema 18, explique lo que encontró en los incisos *a*) y *b*).

22. [**M**] Compare dos métodos para encontrar el vector de estado estable $\mathbf{q}$ de una matriz estocástica regular P: **1.** calculando $\mathbf{q}$ como en el ejemplo 5, o **2.** calculando P^k para un valor grande de k y utilizando una de las columnas de P^k como una aproximación para $\mathbf{q}$. [La *Guía de estudio* describe un programa de *base nula* que casi automatiza el método 1].

Experimente con las matrices aleatorias estocásticas más grandes que su programa de matrices le permita, y utilice $k = 100$ o algún otro valor grande. Para cada método, describa el tiempo que *usted* necesita para teclear y ejecutar su programa. (Algunas versiones de MATLAB tienen los comandos `flops` y `tic` …`toc` que registran el número de operaciones de punto flotante y el tiempo total transcurrido que utiliza MATLAB). Compare las ventajas de cada método y determine cuál prefiere.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *a*) Como el 5% de los residentes de la ciudad se mudará a los suburbios en menos de un año, hay una probabilidad del 5% de elegir a esa persona. Sin profundizar en el conocimiento acerca del individuo, se dice que hay un 5% de probabilidad de que este se mude a los suburbios. Este hecho se encuentra en la segunda entrada del vector de estado $\mathbf{x}_1$, donde

$$\mathbf{x}_1 = M\mathbf{x}_0 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .95 \\ .05 \end{bmatrix}$$

b) La probabilidad de que la persona viva en los suburbios después de dos años es del 9.6%, porque

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .95 \\ .05 \end{bmatrix} = \begin{bmatrix} .904 \\ .096 \end{bmatrix}$$

2. El vector de estado estable satisface $P\mathbf{x} = \mathbf{x}$. Como

$$P\mathbf{q} = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix}\begin{bmatrix} .3 \\ .7 \end{bmatrix} = \begin{bmatrix} .32 \\ .68 \end{bmatrix} \neq \mathbf{q}$$

llegamos a la conclusión de que $\mathbf{q}$ *no* es el vector de estado estable para P.

3. M en el ejemplo 1 es una matriz estocástica regular, ya que sus entradas son estrictamente positivas. Así que podemos utilizar el teorema 18. Ya conocemos el vector de estado estable del ejemplo 4. Por lo tanto, los vectores de distribución de la población $\mathbf{x}_k$ convergen en

$$\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$$

WEB

Al final, el 62.5% de la población vivirá en los suburbios.

CAPÍTULO 4 EJERCICIOS COMPLEMENTARIOS

1. Marque cada enunciado como verdadero o falso. Justifique sus respuestas. (Si es verdadero, cite hechos o teoremas pertinentes. Si es falso, explique por qué o dé un contraejemplo que muestre por qué el enunciado no es cierto en todos los casos). En los incisos *a*) a *f*), $\mathbf{v}_1, \ldots, \mathbf{v}_p$ son vectores en un espacio vectorial V con dimensión finita diferente de cero, y $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

a) El conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ es un espacio vectorial.

b) Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}\}$ genera a V, entonces S genera a V.

c) Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}\}$ es linealmente independiente, también S lo es.

d) Si S es linealmente independiente, entonces S es una base para V.

e) Si Gen $S = V$, entonces algún subconjunto de S es una base para V.

f) Si dim $V = p$ y Gen $S = V$, entonces S no puede ser linealmente dependiente.

g) Un plano en $\mathbb{R}^3$ es un subespacio de dimensión 2.

h) Las columnas de una matriz que no son pivote siempre son linealmente dependientes.

i) Las operaciones de fila en una matriz A pueden cambiar las relaciones de dependencia lineal entre las filas de A.

j) Las operaciones de fila en una matriz pueden cambiar el espacio nulo.

k) El rango de una matriz es igual al número de filas diferentes de cero.

l) Si una matriz A de $m \times n$ es equivalente por filas a una matriz escalonada U, y si U tiene k filas diferentes de cero, entonces la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$ es $m - k$.

m) Si B se obtiene a partir de una matriz A con varias operaciones elementales de fila, entonces rango B = rango A.

n) Las filas diferentes de cero de una matriz A forman una base para Fila A.

o) Si las matrices A y B tienen la misma forma escalonada reducida, entonces Fila A = Fila B.

p) Si H es un subespacio de $\mathbb{R}^3$, entonces hay una matriz A de 3×3 tal que H = Col A.

q) Si A es $m \times n$ y rango $A = m$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno.

r) Si A es de $m \times n$ y la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es sobre, entonces rango $A = m$.

s) Una matriz de cambio de coordenadas siempre es invertible.

t) Si $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ son bases para un espacio vectorial V, entonces la j-ésima columna de la matriz de cambio de coordenadas de la matriz $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ es el vector de coordenadas $[\mathbf{c}_j]_{\mathcal{B}}$.

2. Encuentre una base para el conjunto de todos los vectores de la forma

$$\begin{bmatrix} a - 2b + 5c \\ 2a + 5b - 8c \\ -a - 4b + 7c \\ 3a + b + c \end{bmatrix}. \text{ (Tenga cuidado).}$$

3. Sean $\mathbf{u}_1 = \begin{bmatrix} -2 \\ 4 \\ -6 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ y W = Gen $\{\mathbf{u}_1, \mathbf{u}_2\}$. Encuentre una descripción *implícita* de W, es decir, encuentre un conjunto de una o más ecuaciones homogéneas que caracterizan a los puntos de W. [*Sugerencia:* Pregúntese cuándo está $\mathbf{b}$ en W].

4. Explique lo que está mal en el siguiente análisis: Sea $\mathbf{f}(t) = 3 + t$ y $\mathbf{g}(t) = 3t + t^2$, y note que $\mathbf{g}(t) = t\mathbf{f}(t)$. Entonces, $\{\mathbf{f}, \mathbf{g}\}$ es linealmente dependiente, porque $\mathbf{g}$ es un múltiplo de $\mathbf{f}$.

5. Considere los polinomios $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, $\mathbf{p}_3(t) = 4$, $\mathbf{p}_4(t) = t + t^2$, y $\mathbf{p}_5(t) = 1 + 2t + t^2$, y sea H el subespacio de $\mathbb{P}_5$ generado por el conjunto $S = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5\}$. Utilice el método descrito en la demostración del teorema del conjunto generador (sección 4.3) con la finalidad de obtener una base para H. (Explique cómo seleccionar los miembros adecuados de S).

6. Suponga que $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$ son polinomios específicos que generan un subespacio H de dimensión 2 de $\mathbb{P}_5$. Describa cómo se puede encontrar una base para H mediante el examen de los cuatro polinomios y sin hacer casi ningún cálculo.

7. ¿Qué tendría que saber acerca del conjunto solución de un sistema homogéneo de 18 ecuaciones lineales con 20 variables con la finalidad de determinar que cada ecuación no homogénea asociada tiene una solución? Analice.

8. Sea H un subespacio de dimensión n de un espacio vectorial V de dimensión n. Explique por qué $H = V$.

9. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal.

a) ¿Cuál es la dimensión del rango de T si T es un mapeo uno a uno? Explique su respuesta.

b) ¿Cuál es la dimensión del núcleo de T (véase la sección 4.2) si T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$? Explique su respuesta.

10. Sea S un subconjunto maximal linealmente independiente de un espacio vectorial V. Es decir, S tiene la propiedad de que si un vector que no está en S se adjunta a S, entonces el nuevo conjunto ya no será linealmente independiente. Demuestre que S debe ser una base para V. [*Sugerencia:* Pregúntese qué pasa si S es linealmente independiente, pero no una base de V].

11. Sea S un conjunto generador mínimo finito de un espacio vectorial V. Es decir, S tiene la propiedad de que si un vector se elimina de S, entonces el nuevo conjunto ya no genera a V. Demuestre que S debe ser una base para V.

Los ejercicios 12 a 17 desarrollan las propiedades del rango que a veces son necesarias en las aplicaciones. Suponga que la matriz A es de $m \times n$.

12. A partir de los incisos *a*) y *b*) demuestre que el rango AB no puede exceder al rango de A o al rango de B. (En general, el rango de un producto de matrices no puede exceder el rango de cualquier otro factor en el producto).

a) Demuestre que si B es de $n \times p$, entonces rango $AB \leq$ rango A. [*Sugerencia:* Explique por qué cada vector en el espacio columna de AB está en el espacio columna de A].

b) Demuestre que si B es $n \times p$, entonces rango $AB \leq$ rango B. [*Sugerencia:* Tome en cuenta el inciso *a*) para estudiar el rango de $(AB)^T$].

13. Demuestre que si P es una matriz invertible de $m \times m$, entonces el rango PA = rango A. [*Sugerencia:* Aplique el ejercicio 12 a PA y $P^{-1}(PA)$].

14. Demuestre que si Q es invertible, entonces rango AQ = rango A. [*Sugerencia:* Tome en cuenta el ejercicio 13 para estudiar el rango de $(AQ)^T$].

15. Sea A una matriz de $m \times n$, y B una matriz de $n \times p$ tales que $AB = 0$. Demuestre que rango A + rango $B \leq n$. [*Sugerencia:* Considere que uno de los cuatro subespacios Nul A, Col A, Nul B y Col B se encuentra en uno de los otros tres subespacios].

16. Si A es una matriz de $m \times n$ de rango r, entonces la *factorización de rango* de A es una ecuación de la forma $A = CR$, donde C es una matriz de $m \times r$ de rango r, y R es una matriz de $r \times n$

de rango r. Tal factorización siempre existe (ejercicio 38 en la sección 4.6). Dadas dos matrices A y B de $m \times n$, utilice factorizaciones de rango de A y B para demostrar que

rango $(A + B) \leq$ rango A + rango B

[*Sugerencia:* Escriba $A + B$ como el producto de dos matrices particionadas].

17. Una **submatriz** de una matriz A es cualquier matriz que resulta de la eliminación de algunas filas y/o columnas (o de ninguna) de A. Es posible demostrar que A tiene rango r si y solo si A contiene una submatriz invertible de $r \times r$ y ninguna submatriz cuadrada mayor es invertible. Demuestre parte de este enunciado explicando: *a*) por qué una matriz A de $m \times n$ y rango r tiene una submatriz A_1 de $m \times r$ y rango r, y *b*) por qué A_1 tiene una submatriz invertible A_2 de $r \times r$.

El concepto de rango desempeña un papel importante en el diseño de sistemas de control de ingeniería, como el sistema del transbordador espacial mencionado en el ejemplo introductorio de este capítulo. Un *modelo de espacios de estado* de un sistema de control incluye una ecuación en diferencias de la forma

$$\mathbf{x}_{k+1} = A\mathbf{x} + B\mathbf{u}_k \quad \text{para } k = 0, 1, \ldots \tag{1}$$

donde A es de $n \times n$, B es de $n \times m$, $\{\mathbf{x}_k\}$ es una secuencia de "vectores de estado" en $\mathbb{R}^n$ que describe el estado del sistema en tiempos discretos, y $\{\boldsymbol{u}_k\}$ es una secuencia de *control* o de *entrada*. Se dice que el par (A, B) es **controlable** si

$$\text{rango}\,[B \quad AB \quad A^2B \quad \cdots \quad A^{n-1}B] = n \tag{2}$$

La matriz que aparece en la ecuación (2) se llama **matriz de controlabilidad** del sistema. Si (A, B) es controlable, entonces el sistema se puede controlar, o conducir del estado **0** a cualquier estado especificado $\mathbf{v}$ (en $\mathbb{R}^n$) en un máximo de n pasos, simplemente eligiendo una secuencia de control adecuada en $\mathbb{R}^m$. Este hecho se ilustra en el ejercicio 18 para $n = 4$ y $m = 2$. Para un análisis adicional de la capacidad de control, visite el sitio Web de este libro (estudio de caso del capítulo 4).

WEB

18. Suponga que A es una matriz de 4×4 y B es una matriz de 4×2, y sea que $\mathbf{u}_0, \ldots, \mathbf{u}_3$ representen una secuencia de vectores de entrada en $\mathbb{R}^2$.

a) Establezca $\mathbf{x}_0 = \mathbf{0}$, calcule $\mathbf{x}_1, \ldots, \mathbf{x}_4$ a partir de la ecuación (1), y escriba una fórmula para $\mathbf{x}_4$ que implique a la matriz de controlabilidad M que aparece en la ecuación (2). (*Nota:* La matriz M se construye como una matriz particionada; su tamaño total aquí es 4×8).

b) Suponga que (A, B) es controlable y $\mathbf{v}$ es cualquier vector en $\mathbb{R}^4$. Explique por qué existe una secuencia de control $\mathbf{u}_0, \ldots, \mathbf{u}_3$ en $\mathbb{R}^2$ tal que $\mathbf{x}_4 = \mathbf{v}$.

Determine si los pares de matrices de los ejercicios 19 a 22 son controlables.

19. $A = \begin{bmatrix} .9 & 1 & 0 \\ 0 & -.9 & 0 \\ 0 & 0 & .5 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

20. $A = \begin{bmatrix} .8 & -.3 & 0 \\ .2 & .5 & 1 \\ 0 & 0 & -.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

21. [**M**] $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & -4.2 & -4.8 & -3.6 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

22. [**M**] $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -13 & -12.2 & -1.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

5

Valores propios y vectores propios

EJEMPLO INTRODUCTORIO

Sistemas dinámicos y búhos manchados

En 1990 el búho manchado norteño se convirtió en el centro de una controversia nacional sobre el uso y el mal uso de los majestuosos bosques en el Pacífico Noroccidental de Estados Unidos. Los ambientalistas convencieron al gobierno federal de que el búho estaría en riesgo de extinción si continuaba la tala en los antiguos bosques (con árboles de más de 200 años de edad), donde el búho prefiere vivir. La industria maderera, anticipando la pérdida de 30,000 a 100,000 puestos de trabajo como resultado de las nuevas restricciones gubernamentales sobre la tala, argumentó que el búho no debería estar clasificado como "especie amenazada" y citó diferentes reportes científicos publicados para apoyar su caso.[1]

En medio del fuego cruzado entre los grupos antagónicos, los ecologistas matemáticos intensificaron su afán de comprender la dinámica poblacional del búho. El ciclo de vida de un búho manchado se divide naturalmente en tres etapas: juvenil (hasta un año de edad), subadulto (de 1 a 2 años) y adulto (más de 2 años). Los búhos se aparean de por vida en las etapas de subadulto y adulto, y empiezan a criar cuando son adultos. Viven alrededor de 20 años y cada pareja requiere aproximadamente 1000 hectáreas (4 millas cuadradas) como su propio territorio. Un tiempo fundamental en el ciclo de vida es cuando los búhos juveniles dejan el nido. Para sobrevivir y convertirse en un subadulto, un búho juvenil debe tener éxito para encontrar un nuevo territorio para vivir (y en general, una pareja).

Un primer paso en el estudio de la dinámica poblacional consiste en modelar la población a intervalos anuales, en tiempos que se denotan con $k = 0, 1, 2,...$ En general, se supone que hay una relación 1:1 de machos a hembras en cada etapa del ciclo de vida y se cuentan solamente las hembras. La población en cada año k se describe con un vector $\boldsymbol{x}_k = (j_k, s_k, a_k)$, donde j_k, s_k y a_k son los números de hembras en las etapas juvenil, subadulto y adulto, respectivamente.

Utilizando datos reales de estudios demográficos, R. Lamberson y sus colaboradores consideraron el siguiente *modelo matricial por etapas*:[2]

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

Aquí, el número de nuevas hembras juveniles en el año $k + 1$ es .33 veces el número de hembras adultas en el año k (según la tasa promedio de nacimientos por pareja de búhos). Asimismo, 18% de los búhos juveniles sobreviven para convertirse en subadultos, en tanto que 71% de los subadultos y 94% de los adultos sobreviven para contarse como adultos.

El modelo matricial por etapas es una ecuación en diferencias de la forma $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$. A esta ecuación con frecuencia se le llama **sistema dinámico** (o **sistema**

[1] "The Great Spotted Owl War", *Reader's Digest*, noviembre de 1992, pp. 91-95.

[2] R. H. Lamberson, R. McKelvey, B. R. Noon y C. Voss, "A Dynamic Analysis of the Viability of the Northern Spotted Owl in a Fragmented Forest Environment", *Conservation Biology* **6** (1992), pp. 505-512. También, una comunicación privada del profesor Lamberson, 1993.

dinámico lineal discreto), ya que describe los cambios en un sistema conforme transcurre el tiempo.

La tasa de sobrevivencia del 18% de búhos juveniles en la matriz por etapas de Lamberson es la entrada más afectada por la cantidad de bosque antiguo disponible. En realidad, el 60% de los búhos juveniles normalmente sobreviven para dejar el nido, pero en la región de Willow Creek, California, estudiada por Lamberson y sus colaboradores, tan solo al 30% de los búhos juveniles que dejaron el nido les fue posible encontrar un nuevo territorio para vivir. Los restantes perecieron durante la búsqueda.

Un motivo significativo para el fracaso de los búhos en encontrar nuevas áreas de distribución es el aumento en la fragmentación de las áreas con árboles antiguos, debido a la tala total de áreas dispersas en terrenos con árboles antiguos. Cuando un búho sale de la región protectora del bosque y cruza un área talada, aumenta drásticamente el riesgo de sufrir un ataque por parte de los depredadores. En la sección 5.6 se mostrará que el modelo descrito predice la eventual extinción de búhos, pero que si el 50% de los búhos juveniles que sobreviven al dejar el nido también encuentran nuevas áreas de distribución, entonces se salvaría su población.

WEB

El objetivo de este capítulo es descomponer la acción de una transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ en elementos que sean de fácil visualización. Excepto por un breve paréntesis en la sección 5.4, en este capítulo todas las matrices son cuadradas. Las principales aplicaciones descritas aquí son para sistemas dinámicos discretos, incluyendo el ya mencionado asunto de los búhos manchados. Sin embargo, los conceptos básicos, vectores propios y valores propios, son útiles en matemáticas puras y aplicadas, y se presentan en situaciones más generales que las que aquí se consideran. Los valores propios también sirven para estudiar ecuaciones diferenciales y sistemas dinámicos *continuos*, brindan información crítica en diseños de ingeniería, y se originan de manera natural en campos de la física y la química.

5.1 VECTORES PROPIOS Y VALORES PROPIOS

No obstante que una transformación $\mathbf{x} \mapsto A\mathbf{x}$ puede mover vectores en diversas direcciones, a menudo sucede que hay vectores especiales sobre los cuales la acción de A es bastante simple.

EJEMPLO 1 Sean $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. En la figura 1 se muestran las imágenes de $\mathbf{u}$ y $\mathbf{v}$ bajo la multiplicación por A. En efecto, $A\mathbf{v}$ es precisamente $2\mathbf{v}$. Así, A tan solo "estira" o dilata a $\mathbf{v}$. ■

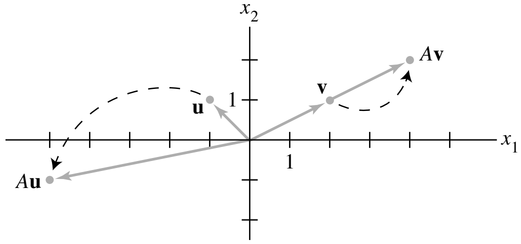


FIGURA 1 Efectos de la multiplicación por A.

Como otro ejemplo, los lectores de la sección 4.9 recordarán que si A es una matriz estocástica, entonces el vector de estado estable $\mathbf{q}$ para A satisface la ecuación $A\mathbf{x} = \mathbf{x}$. Es decir, $A\mathbf{q} = 1 \cdot \mathbf{q}$.

Esta sección estudia ecuaciones tales como

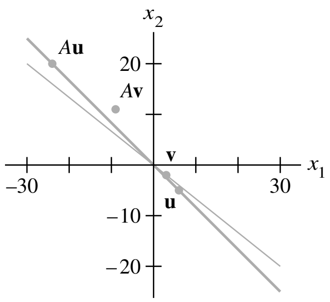
$$A\mathbf{x} = 2\mathbf{x} \quad \text{o bien} \quad A\mathbf{x} = -4\mathbf{x}$$

donde los vectores especiales son transformados por A en múltiplos escalares de sí mismos.

DEFINICIÓN

Un **vector propio** de una matriz A de $n \times n$ es un vector distinto de cero $\mathbf{x}$ tal que $A\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ. Un escalar λ es un **valor propio** de A si existe una solución no trivial $\mathbf{x}$ de $A\mathbf{x} = \lambda\mathbf{x}$; tal $\mathbf{x}$ es el *vector propio correspondiente a* λ.[1]

Es fácil determinar si un vector dado es un vector propio de una matriz. También es sencillo decidir si cierto escalar es un valor propio.



$A\mathbf{u} = -4\mathbf{u}$, pero $A\mathbf{v} \neq \lambda\mathbf{v}$.

EJEMPLO 2 Sean $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. ¿Son $\mathbf{u}$ y $\mathbf{v}$ vectores propios de A?

SOLUCIÓN

$$A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4\begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\mathbf{u}$$

$$A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda\begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Entonces, $\mathbf{u}$ es un vector propio correspondiente a un valor propio (-4), pero $\mathbf{v}$ no es un vector propio de A porque $A\mathbf{v}$ no es múltiplo de $\mathbf{v}$. ■

EJEMPLO 3 Demuestre que 7 es un valor propio de la matriz A del ejemplo 2, y determine los vectores propios correspondientes.

SOLUCIÓN El escalar 7 es un valor propio de A, si y solo si la ecuación

$$A\mathbf{x} = 7\mathbf{x} \tag{1}$$

tiene una solución no trivial. No obstante, la ecuación (1) es equivalente a $A\mathbf{x} - 7\mathbf{x} = \mathbf{0}$, o bien,

$$(A - 7I)\mathbf{x} = \mathbf{0} \tag{2}$$

Para resolver esta ecuación homogénea, forme la matriz

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

Como es evidente, las columnas de $A - 7I$ son linealmente dependientes, de manera que la ecuación (2) tiene soluciones no triviales. Por lo tanto, 7 *es* un valor propio de A. Para encontrar los vectores propios correspondientes, utilice operaciones de fila:

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

La solución general tiene la forma $x_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Cada vector de esta forma con $x_2 \neq 0$ es un vector propio correspondiente a $\lambda = 7$. ■

[1] Advierta que, por definición, un vector propio debe ser *distinto* de cero, pero un valor propio puede ser cero. Después del ejemplo 5, se analizará el caso donde el número 0 es un valor propio.

Advertencia: No obstante que en el ejemplo 3 se utilizó reducción por filas para encontrar *vectores* propios, este método no se puede emplear para determinar *valores* propios. En general, una forma escalonada de una matriz A *no* presenta los valores propios de A.

Evidentemente, la equivalencia de las ecuaciones (1) y (2) es válida para cualquier λ en vez de $\lambda = 7$. Entonces, λ es un valor propio de la matriz A de $n \times n$ si y solo si la ecuación

$$(A - \lambda I)\mathbf{x} = \mathbf{0} \tag{3}$$

tiene una solución no trivial. El conjunto de *todas* las soluciones de (3) es justo el espacio nulo de la matriz de $A - \lambda I$. Así que este conjunto es un *subespacio* de $\mathbb{R}^n$ y se denomina el **espacio propio** de A correspondiente a λ. El espacio propio consiste en el vector cero y de todos los vectores propios correspondientes a λ.

El ejemplo 3 demuestra que para la matriz A del ejemplo 2, el espacio propio correspondiente a $\lambda = 7$ consiste en *todos* los múltiplos de (1, 1), que es la recta que pasa tanto por (1, 1) como por el origen. Con el ejemplo 2, se comprueba que el espacio propio asociado con $\lambda = -4$ es la recta que pasa por $(6, -5)$. En la figura 2, se presentan esos espacios propios, junto con los vectores propios (1, 1) y $(3/2, -5/4)$ y la acción geométrica de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ sobre cada espacio propio.

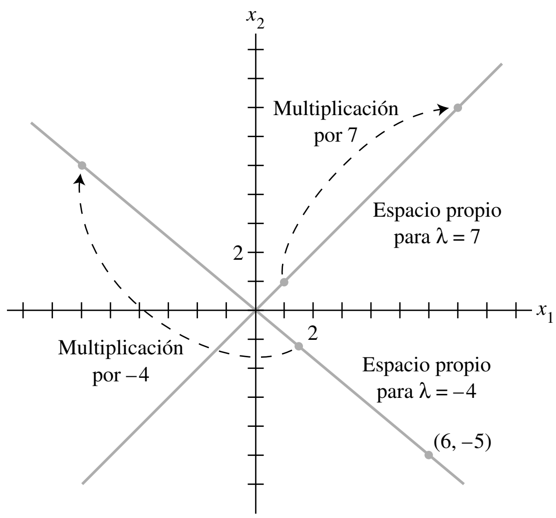


FIGURA 2 Espacios propios para $\lambda = -4$ y $\lambda = 7$.

EJEMPLO 4 Sea $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. Un valor propio de A es 2. Determine una base para el espacio propio correspondiente.

SOLUCIÓN Forme

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

y reduzca por filas la matriz aumentada para $(A - 2I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

En este punto, resulta claro que 2 es un valor propio de A porque la ecuación $(A - 2I)\mathbf{x} = \mathbf{0}$ tiene variables libres. La solución general es:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad x_2 \text{ y } x_3 \text{ son libres}$$

El espacio propio, que se muestra en la figura 3, es un subespacio bidimensional de $\mathbb{R}^3$. Una base es

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

■

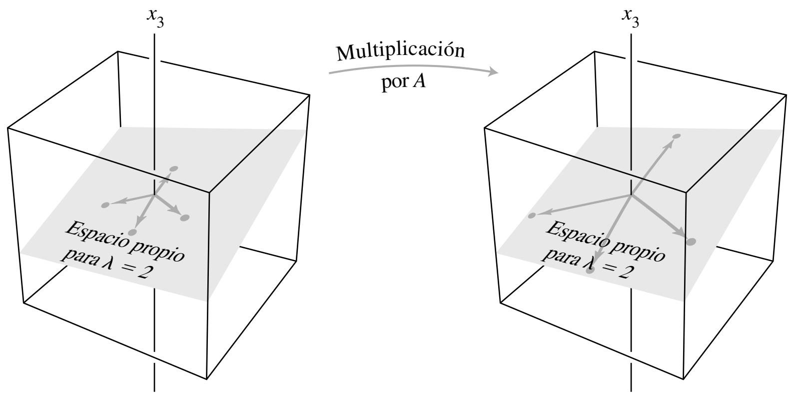


FIGURA 3 A actúa como una dilatación sobre el espacio propio.

NOTA NUMÉRICA

El ejemplo 4 es un buen método para el cálculo manual de vectores propios en casos simples donde se conoce un valor propio. En general, se recomienda el uso de un programa matricial y la reducción por filas para encontrar un espacio propio (para un valor propio dado), aunque no es totalmente confiable. Ocasionalmente, el error por redondeo lleva a una forma escalonada reducida con el número de pivotes equivocado. Los mejores programas computacionales calculan simultáneamente aproximaciones para los valores propios y los vectores propios, con el grado de exactitud que se requiera, en matrices que no son muy grandes. El tamaño de las matrices que se logra analizar se incrementa cada año conforme mejora la potencia computacional y la eficiencia del software.

El siguiente teorema describe uno de los pocos casos especiales donde los valores propios se pueden determinar con precisión. En la sección 5.2 también se analizará el cálculo de valores propios.

TEOREMA 1

Los valores propios de una matriz triangular son las entradas sobre su diagonal principal.

DEMOSTRACIÓN Por sencillez, considere el caso 3×3. Si A es triangular superior, entonces, $A - \lambda I$ tiene la forma

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \\ &= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix} \end{aligned}$$

El escalar λ es un valor propio de A si y solo si la ecuación $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial, es decir, si y solo si la ecuación tiene una variable libre. Gracias a las entradas cero en $A - \lambda I$, es fácil ver que $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene una variable libre si y solo si al menos una de las entradas sobre la diagonal de $A - \lambda I$ es cero. Esto ocurre si y solo si λ es igual a una de las entradas a_{11}, a_{22}, a_{33} en A. Véase el ejercicio 28 para el caso donde A es triangular inferior. ■

EJEMPLO 5 Sean $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$. Los valores propios de A son 3, 0 y 2. Los valores propios de B son 4 y 1. ■

¿Qué significa para una matriz A tener un valor propio de 0, tal como en el ejemplo 5? Esto ocurre si y solo si la ecuación

$$A\mathbf{x} = 0\mathbf{x} \tag{4}$$

tiene una solución no trivial. Pero la ecuación (4) es equivalente a $A\mathbf{x} = \mathbf{0}$, que tiene una solución no trivial si y solo si A no es invertible. Entonces, 0 *es un valor propio de A si y solo si A no es invertible.* En la sección 5.2 este hecho se agregará al teorema de la matriz invertible.

El siguiente teorema importante se necesitará más adelante. Su demostración muestra un cálculo típico con vectores propios.

TEOREMA 2

Si $\mathbf{v}_1, \ldots, \mathbf{v}_r$ son vectores propios que corresponden a distintos valores propios $\lambda_1, \ldots, \lambda_r$ de una matriz A de $n \times n$, entonces el conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ es linealmente independiente.

DEMOSTRACIÓN Suponga que $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ es linealmente dependiente. Puesto que $\mathbf{v}_1$ es diferente de cero, el teorema 7 de la sección 1.7 indica que uno de los vectores en el conjunto es una combinación lineal de los vectores precedentes. Sea p el índice mínimo tal que $\mathbf{v}_{p+1}$ sea una combinación lineal de los vectores precedentes (linealmente independientes). Entonces, existen escalares $c_1, \ldots, c_p$ tales que

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{v}_{p+1} \tag{5}$$

Multiplicando por A ambos miembros de la ecuación (5) y considerando el hecho de que $A\mathbf{v}_k = \lambda_k\mathbf{v}_k$ para cada k,

$$\begin{aligned} c_1A\mathbf{v}_1 + \cdots + c_pA\mathbf{v}_p &= A\mathbf{v}_{p+1} \\ c_1\lambda_1\mathbf{v}_1 + \cdots + c_p\lambda_p\mathbf{v}_p &= \lambda_{p+1}\mathbf{v}_{p+1} \end{aligned} \tag{6}$$

Multiplicando por λ_{p+1} ambos lados de la ecuación (5) y restando el resultado de (6),

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0} \tag{7}$$

Puesto que $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es linealmente independiente, todos los pesos en la ecuación (7) son iguales a cero. Sin embargo, ninguno de los factores $\lambda_i - \lambda_{p+1}$ son cero, ya que los valores propios son distintos. Por lo tanto, $c_i = 0$ para $i = 1, \ldots, p$. No obstante, la ecuación (5) indica que $\mathbf{v}_{p+1} = \mathbf{0}$, lo cual es imposible. En consecuencia, $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ no puede ser linealmente dependiente y, por consiguiente, debe ser linealmente independiente. ■

Vectores propios y ecuaciones en diferencias

Esta sección concluye con la demostración cómo construir soluciones de la ecuación en diferencias de primer orden analizada en el ejemplo introductorio del capítulo:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad (k = 0, 1, 2, \ldots) \tag{8}$$

Si A es una matriz de $n \times n$, entonces la ecuación (8) es una descripción *recursiva* de una secuencia $\{\mathbf{x}_k\}$ en $\mathbb{R}^n$. Una **solución** de la ecuación (8) es una descripción explícita de $\{\mathbf{x}_k\}$, cuya fórmula para cada $\mathbf{x}_k$ no depende directamente de A ni de los términos precedentes en la secuencia, excepto del término inicial $\mathbf{x}_0$.

La forma más sencilla de construir una solución de la ecuación (8) es tomar un vector propio $\mathbf{x}_0$ y su correspondiente valor propio λ, y hacer

$$\mathbf{x}_k = \lambda^k \mathbf{x}_0 \quad (k = 1, 2, \ldots) \tag{9}$$

Esta secuencia es una solución, ya que

$$A\mathbf{x}_k = A(\lambda^k \mathbf{x}_0) = \lambda^k (A\mathbf{x}_0) = \lambda^k (\lambda \mathbf{x}_0) = \lambda^{k+1}\mathbf{x}_0 = \mathbf{x}_{k+1}$$

Combinaciones lineales de soluciones en la forma de la ecuación (9) ¡también son soluciones! Véase el ejercicio 33.

PROBLEMAS DE PRÁCTICA

1. ¿Es 5 un valor propio de $A = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?
2. Si $\mathbf{x}$ es un vector propio de A correspondiente a λ, ¿qué ocurre con $A^3\mathbf{x}$?
3. Suponga que $\mathbf{b}_1$ y $\mathbf{b}_2$ son vectores propios correspondientes a distintos valores propios λ_1 y λ_2, respectivamente, y también que $\mathbf{b}_3$ y $\mathbf{b}_4$ son vectores propios linealmente independientes asociados a un tercer valor propio λ_3 distinto. ¿Necesariamente se deduce que $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4\}$ es un conjunto linealmente independiente? [*Sugerencia:* Considere la ecuación $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$].

5.1 EJERCICIOS

1. ¿Es $\lambda = 2$ un valor propio de $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$? ¿Por qué?
2. ¿Es $\lambda = -3$ un valor propio de $\begin{bmatrix} -1 & 4 \\ 6 & 9 \end{bmatrix}$? ¿Por qué?
3. ¿Es $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ un vector propio de $\begin{bmatrix} 1 & -1 \\ 6 & -4 \end{bmatrix}$? Si lo es, encuentre el valor propio.
4. ¿Es $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ un vector propio de $\begin{bmatrix} 5 & 2 \\ 3 & 6 \end{bmatrix}$? Si así es, determine el valor propio.
5. ¿Es $\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$ un vector propio de $\begin{bmatrix} -4 & 3 & 3 \\ 2 & -3 & -2 \\ -1 & 0 & -2 \end{bmatrix}$? En caso afirmativo, determine el valor propio.
6. ¿Es $\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$ un vector propio de $\begin{bmatrix} 3 & 6 & 7 \\ 3 & 2 & 7 \\ 5 & 6 & 4 \end{bmatrix}$? Si lo es, encuentre el valor propio.
7. ¿Es $\lambda = 4$ un valor propio de $\begin{bmatrix} 3 & 0 & -1 \\ 2 & 3 & 1 \\ -3 & 4 & 5 \end{bmatrix}$? Si es así, determine un vector propio correspondiente.
8. ¿Es $\lambda = 1$ un valor propio de $\begin{bmatrix} 4 & -2 & 3 \\ 0 & -1 & 3 \\ -1 & 2 & -2 \end{bmatrix}$? Si es así, determine un vector propio correspondiente.

En los ejercicios 9 a 16, determine una base para el espacio propio asociado con cada valor propio indicado.

9. $A = \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}$, $\lambda = 1, 3$

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

11. $A = \begin{bmatrix} 1 & -3 \\ -4 & 5 \end{bmatrix}, \lambda = -1, 7$

12. $A = \begin{bmatrix} 4 & 1 \\ 3 & 6 \end{bmatrix}, \lambda = 3, 7$

13. $A = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \lambda = 1, 2, 3$

14. $A = \begin{bmatrix} 4 & 0 & -1 \\ 3 & 0 & 3 \\ 2 & -2 & 5 \end{bmatrix}, \lambda = 3$

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

16. $A = \begin{bmatrix} 5 & 0 & -1 & 0 \\ 1 & 3 & 0 & 0 \\ 2 & -1 & 3 & 0 \\ 4 & -2 & -2 & 4 \end{bmatrix}, \lambda = 4$

En los ejercicios 17 y 18, determine los valores propios de las matrices.

17. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & -2 \end{bmatrix}$ **18.** $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 3 \end{bmatrix}$

19. Para $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, encuentre un valor propio, sin realizar cálculos. Justifique su respuesta.

20. Sin hacer cálculos, obtenga un valor propio y dos vectores propios linealmente independientes de $A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$. Justifique su respuesta.

En los ejercicios 21 y 22, A es una matriz de $n \times n$. Marque cada enunciado como verdadero o falso. Justifique su respuesta.

21. *a*) Si $A\mathbf{x} = \lambda\mathbf{x}$ para algún vector $\mathbf{x}$, entonces λ es un valor propio de A.

b) Una matriz A no es invertible si y solo si 0 es un valor propio de A.

c) Un número c es un valor propio de A si y solo si la ecuación $(A - cI)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial.

d) Quizá sea difícil obtener un vector propio de A, pero es sencillo comprobar si un vector dado es, de hecho, un vector propio.

e) Para determinar los valores propios de A, se reduce A a una forma escalonada.

22. *a*) Si $A\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ, entonces $\mathbf{x}$ es un vector propio de A.

b) Si $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios linealmente independientes, entonces corresponden a distintos valores propios.

c) Un vector de estado estable para una matriz estocástica es realmente un vector propio.

d) Los valores propios de una matriz están sobre su diagonal principal.

e) Un espacio propio de A es el espacio nulo de cierta matriz.

23. Explique por qué una matriz de 2×2 puede tener, cuando mucho, dos valores propios distintos. También indique por qué una matriz de $n \times n$ puede tener, cuando mucho, n valores propios diferentes.

24. Construya un ejemplo de una matriz de 2×2 con tan solo un valor propio distinto.

25. Sea λ un valor propio de una matriz A invertible. Demuestre que λ^{-1} es un valor propio de A^{-1}. [*Sugerencia:* Suponga que $\mathbf{x}$ distinto de cero satisface $A\mathbf{x} = \lambda\mathbf{x}$].

26. Demuestre que si A^2 es la matriz cero, entonces el único valor propio de A es 0.

27. Demuestre que λ es un valor propio de A si y solo si λ es un valor propio de A^T. [*Sugerencia:* Determine cómo se relaciona $A - \lambda I$ con $A^T - \lambda I$].

28. Utilice el ejercicio 27 para completar la demostración del teorema 1 para el caso en que A es triangular inferior.

29. Considere una matriz A de $n \times n$ con la propiedad de que la suma de las filas sea igual al mismo número s. Demuestre que s es un valor propio de A. [*Sugerencia:* Encuentre un vector propio].

30. Considere una matriz A de $n \times n$ tal que las sumas de columnas sean iguales al mismo número s. Demuestre que s es un valor propio de A. [*Sugerencia:* Utilice los ejercicios 27 y 29].

En los ejercicios 31 y 32, sea A la matriz de la transformación lineal T. Sin escribir A, encuentre un valor propio de A y describa el espacio propio.

31. T es la transformación sobre $\mathbb{R}^2$ que refleja los puntos con respecto a una recta que pasa por el origen.

32. T es la transformación en $\mathbb{R}^3$ que gira los puntos alrededor de alguna recta que pasa por el origen.

33. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores propios de una matriz A, con valores propios correspondientes λ y μ, y sean c_1 y c_2 escalares. Defina:

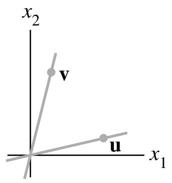
$$\mathbf{x}_k = c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v} \quad (k = 0, 1, 2, \ldots)$$

a) ¿Qué es $\mathbf{x}_{k+1}$ por definición?

b) Calcule $A\mathbf{x}_k$ de la fórmula para $\mathbf{x}_k$ y demuestre que $A\mathbf{x}_k = \mathbf{x}_{k+1}$. Este cálculo probará que la secuencia $\{\mathbf{x}_k\}$ que se acaba de definir satisface la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$.

34. Describa cómo intentaría construir una solución de una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0,1,2,\ldots)$, si le dieran $\mathbf{x}_0$ inicial y resultara que este no fuera un vector propio de A. [*Sugerencia:* ¿Cómo podría relacionar $\mathbf{x}_0$ con los vectores propios de A?]

35. Sean $\mathbf{u}$ y $\mathbf{v}$ los vectores que se muestran en la figura, y suponga que $\mathbf{u}$ y $\mathbf{v}$ son vectores propios de una matriz A de 2×2 correspondientes a los valores propios 2 y 3, respectivamente. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal dada por $T(\mathbf{x}) = A\mathbf{x}$ para cada $\mathbf{x}$ en $\mathbb{R}^2$, y sea $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Haga una copia de la figura, y sobre el mismo sistema de coordenadas grafique cuidadosamente los vectores $T(\mathbf{u})$, $T(\mathbf{v})$ y $T(\mathbf{w})$.



36. Repita el ejercicio 35, suponiendo que $\mathbf{u}$ y $\mathbf{v}$ son vectores propios de A correspondientes a los valores propios -1 y 3, respectivamente.

[**M**] En los ejercicios 37 a 40, utilice un programa matricial para determinar los valores propios de la matriz. Después, aplique el método del ejemplo 4 con una rutina de reducción por filas para producir una base para cada espacio propio.

37. $\begin{bmatrix} 12 & 1 & 4 \\ 2 & 11 & 4 \\ 1 & 3 & 7 \end{bmatrix}$

38. $\begin{bmatrix} 5 & -2 & 2 & -4 \\ 7 & -4 & 2 & -4 \\ 4 & -4 & 2 & 0 \\ 3 & -1 & 1 & -3 \end{bmatrix}$

39. $\begin{bmatrix} 12 & -90 & 30 & 30 & 30 \\ 8 & -49 & 15 & 15 & 15 \\ 16 & -52 & 12 & 0 & 20 \\ 0 & -30 & 10 & 22 & 10 \\ 8 & -41 & 15 & 15 & 7 \end{bmatrix}$

40. $\begin{bmatrix} -23 & 57 & -9 & -15 & -59 \\ -10 & 12 & -10 & 2 & -22 \\ 11 & 5 & -3 & -19 & -15 \\ -27 & 31 & -27 & 25 & -37 \\ -5 & -15 & -5 & 1 & 31 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. El número 5 es un valor propio de A si y solo si la ecuación $(A - 5I)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. Forme

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

y reduzca por filas la matriz aumentada:

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

En este punto, es evidente que el sistema homogéneo no tiene variables libres. Entonces, $A - 5I$ es una matriz invertible, lo cual significa que 5 *no* es un valor propio de A.

2. Si $\mathbf{x}$ es un vector propio de A correspondiente a λ, entonces $A\mathbf{x} = \lambda\mathbf{x}$ y

$$A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

Otra vez, $A^3\mathbf{x} = A(A^2\mathbf{x}) = A(\lambda^2\mathbf{x}) = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$. El patrón general, $A^k\mathbf{x} = \lambda^k\mathbf{x}$, se demuestra por inducción.

3. Sí. Suponga que $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + c_4\mathbf{b}_4 = \mathbf{0}$. Puesto que cualquier combinación lineal de vectores propios del mismo valor propio es otra vez un vector propio para ese valor propio, $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ es un vector propio para λ_3. De acuerdo con el teorema 2, los vectores $\mathbf{b}_1$, $\mathbf{b}_2$ y $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ son linealmente independientes, de manera que

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$$

implica que $c_1 = c_2 = 0$. Pero, c_3 y c_4 también deben ser cero porque $\mathbf{b}_3$ y $\mathbf{b}_4$ son linealmente independientes. Por lo tanto, todos los coeficientes en la ecuación original deben ser iguales a cero, y los vectores $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ y $\mathbf{b}_4$ son linealmente independientes.

5.2 LA ECUACIÓN CARACTERÍSTICA

Información útil acerca de los valores propios de una matriz cuadrada A está codificada en una ecuación escalar especial llamada ecuación característica de A. Un simple ejemplo nos llevará al caso general.

EJEMPLO 1 Determine los valores propios de $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

SOLUCIÓN Deben obtenerse todos los escalares λ tales que la ecuación matricial

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

tenga una solución no trivial. De acuerdo con el teorema de la matriz invertible de la sección 2.3, este problema es equivalente a encontrar todas las λ tales que la matriz de $A - \lambda I$ *no* sea invertible, donde

$$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$$

De acuerdo con el teorema 4 de la sección 2.2, esta matriz no es invertible precisamente cuando su determinante es cero. De manera que los valores propios de A son las soluciones de la ecuación

$$\det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = 0$$

Recuerde que

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Entonces,

$$\begin{aligned} \det(A - \lambda I) &= (2-\lambda)(-6-\lambda) - (3)(3) \\ &= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 \\ &= \lambda^2 + 4\lambda - 21 \\ &= (\lambda - 3)(\lambda + 7) \end{aligned}$$

Si $\det(A - \lambda I) = 0$, entonces $\lambda = 3$, o bien, $\lambda = -7$. Por lo tanto, los valores propios de A son 3 y -7. ■

El determinante en el ejemplo 1 transformó la ecuación matricial $(A - \lambda I)\mathbf{x} = \mathbf{0}$, que implica *dos* incógnitas (λ y $\mathbf{x}$), en la ecuación escalar $\lambda^2 + 4\lambda - 21 = 0$, que tan solo implica una incógnita. La misma idea funciona para matrices de $n \times n$. Sin embargo, antes de pasar a matrices más grandes, se resumen las propiedades necesarias de los determinantes para estudiar valores propios.

Determinantes

Sean A una matriz de $n \times n$, y U una forma escalonada de A obtenida mediante remplazos e intercambios de fila (sin escalamiento), y sea r el número de tales intercambios de fila. Entonces, el **determinante** de A, que se escribe $\det A$, es $(-1)^r$ veces el producto de las entradas diagonales $u_{11}, \ldots, u_{nn}$ en U. Si A es invertible, todas la entradas $u_{11}, \ldots, u_{nn}$ son *pivotes* (porque $A \sim I_n$ y las u_{ii} no se escalaron a 1). Es decir, al menos u_{nn} es cero y el producto $u_{11} \cdots u_{nn}$ es cero. Así,[1]

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix} \text{producto de} \\ \text{pivotes en } U \end{pmatrix}, & \text{cuando } A \text{ es invertible} \\ 0, & \text{cuando } A \text{ no es invertible} \end{cases} \tag{1}$$

[1] La fórmula (1) se dedujo en la sección 3.2. Los lectores que no hayan estudiado el capítulo 3 pueden emplear esta fórmula como la definición de det A. Es un hecho notable y no trivial que cualquier forma escalonada U obtenida de A sin escalar dé el mismo valor para det A.

EJEMPLO 2 Calcule det A para $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

SOLUCIÓN La siguiente reducción por filas utiliza un intercambio de filas:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} = U_1$$

De manera que det $A = (-1)^1(1)(-2)(-1) = -2$. La siguiente reducción por filas alternativa evita el intercambio de filas y produce una forma escalonada diferente. El último paso suma $-1/3$ por la fila 2 a la fila 3:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & 0 & 1/3 \end{bmatrix} = U_2$$

Esta vez, det A es $(-1)^0(1)(-6)(1/3) = -2$, igual que antes. ■

La fórmula (1) para el determinante muestra que A es invertible si y solo si det A es distinto de cero. Este hecho, y la caracterización de invertibilidad establecida en la sección 5.1, se agregan al teorema de la matriz invertible.

TEOREMA

El teorema de la matriz invertible (continuación)

Sea A una matriz de $n \times n$. Entonces, A es invertible si y solo si:

s) El número 0 *no es* un valor propio de A.

t) El determinante de A *no es* cero.

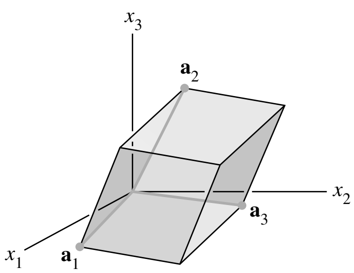


FIGURA 1

Cuando A es una matriz de 3×3, $|\det A|$ resulta ser el *volumen* del paralelepípedo definido por las columnas $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ de A, como se muestra en la figura 1. (Para más detalles, véase la sección 3.3). Este volumen es *distinto de cero* si y solo si los vectores $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ son linealmente independientes, en cuyo caso la matriz A es invertible. (Si los vectores son distintos de cero y linealmente dependientes, entonces se encuentran sobre un plano o sobre una recta).

El siguiente teorema lista los resultados, que se necesitarán, de las secciones 3.1 y 3.2. Aquí se incluye el inciso *a*) como una referencia conveniente.

TEOREMA 3

Propiedades de los determinantes

Sean A y B matrices de $n \times n$.

a) A es invertible si y solo si det $A \neq 0$.

b) det $(AB) = (\det A)(\det B)$.

c) det $A^T = \det A$.

d) Si A es triangular, entonces det A es el producto de las entradas sobre la diagonal principal de A.

e) La operación de remplazo de filas sobre A no cambia el determinante. Un intercambio de filas cambia el signo del determinante. Un escalamiento de filas también escala al determinante por el mismo factor de escala.

La ecuación característica

El teorema 3*a*) demuestra cómo determinar cuando una matriz de la forma $A - \lambda I$ *no* es invertible. La ecuación escalar $\det(A - \lambda I) = 0$ es la **ecuación característica** de A, y el argumento en el ejemplo 1 justifica el siguiente hecho.

> Un escalar λ es un valor propio de una matriz A de $n \times n$ si y solo si λ satisface la ecuación característica
>
> $$\det(A - \lambda I) = 0$$

EJEMPLO 3 Encuentre la ecuación característica de

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

SOLUCIÓN Forme $A - \lambda I$ y utilice el teorema 3*d*):

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} \\ &= (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda) \end{aligned}$$

La ecuación característica es

$$(5-\lambda)^2(3-\lambda)(1-\lambda) = 0$$

o bien,

$$(\lambda-5)^2(\lambda-3)(\lambda-1) = 0$$

Desarrollando el producto, se escribe

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0$$ ■

En los ejemplos 1 y 3, $\det(A - \lambda I)$ es un polinomio en λ. Se puede demostrar que si A es una matriz de $n \times n$, entonces $\det(A - \lambda I)$ es un polinomio de grado n llamado **polinomio característico** de A.

El valor propio 5, en el ejemplo 3, tiene *multiplicidad* 2 porque $(\lambda - 5)$ ocurre dos veces como factor del polinomio característico. En general, la **multiplicidad** (**algebraica**) de un valor propio λ es su multiplicidad como una raíz de la ecuación característica.

EJEMPLO 4 El polinomio característico de una matriz de 6×6 es $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Determine los valores propios y sus multiplicidades.

SOLUCIÓN Factorice el polinomio

$$\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda^2 - 4\lambda - 12) = \lambda^4(\lambda - 6)(\lambda + 2)$$

Los valores propios son 0 (multiplicidad 4), 6 (multiplicidad 1) y -2 (multiplicidad 1). ■

Se podrían listar los valores propios del ejemplo 4 como 0, 0, 0, 0, 6 y -2, de modo que los valores propios están repetidos de acuerdo con sus multiplicidades.

Puesto que la ecuación característica para una matriz de $n \times n$ implica un polinomio de grado n, la ecuación tiene exactamente n raíces, contando las multiplicidades, si se permiten raíces complejas. En la sección 5.5 se analizarán dichas raíces complejas, llamadas *valores propios complejos.* Hasta entonces, solo se considerarán valores propios reales, y los escalares continuarán siendo números reales.

La ecuación característica es importante para fines teóricos. Sin embargo, en el trabajo práctico, los valores propios de una matriz más grande que 2×2 deberían encontrarse usando una computadora, a menos que la matriz sea triangular o tenga otras propiedades especiales. No obstante que un polinomio característico 3×3 es fácil de calcular a mano, quizá sea difícil factorizarlo (a menos que la matriz sea cuidadosamente seleccionada). Véase las "Notas numéricas" del final de esta sección.

Similitud

El siguiente teorema ilustra un uso del polinomio característico, y brinda un fundamento para varios métodos iterativos que *aproximan* valores propios. Si A y B son matrices de $n \times n$, entonces A **es similar a** B si hay una matriz invertible P tal que $P^{-1}AP = B$, o bien, de manera equivalente, $A = PBP^{-1}$. Escribiendo Q para P^{-1}, se tiene $Q^{-1}BQ = A$. Así, B es también similar a A, y se dice simplemente que A y B **son similares**. Cambiar A por $P^{-1}AP$ es una **transformación de similitud**.

TEOREMA 4 Si las matrices A y B de $n \times n$ son similares, entonces tienen el mismo polinomio característico y, por lo tanto, los mismos valores propios (con las mismas multiplicidades).

DEMOSTRACIÓN Si $B = P^{-1}AP$, entonces,

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

Utilizando la propiedad multiplicativa *b*) del teorema 3, se calcula

$$\begin{aligned} \det(B - \lambda I) &= \det[P^{-1}(A - \lambda I)P] \\ &= \det(P^{-1}) \cdot \det(A - \lambda I) \cdot \det(P) \end{aligned} \tag{2}$$

Puesto que $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det I = 1$, de la ecuación (2) se observa que $\det(B - \lambda I) = \det(A - \lambda I)$. ■

ADVERTENCIAS:

1. Las matrices

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

no son similares a pesar de tener los mismos valores propios.

2. Similitud no es lo mismo que equivalencia por filas. (Si A es equivalente por filas a B, entonces $B = EA$ para alguna matriz E invertible). En general, las operaciones con filas sobre una matriz cambian sus valores propios.

Aplicación a sistemas dinámicos

Los valores propios y vectores propios tienen la clave para la evolución discreta de un sistema dinámico, como se mencionó en la introducción al capítulo.

EJEMPLO 5 Sea $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analice el comportamiento a largo plazo del sistema dinámico definido por $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2, \ldots$), con $\mathbf{x}_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$.

SOLUCIÓN El primer paso consiste en determinar los valores propios de A y una base para cada espacio propio. La ecuación característica de A es

$$\begin{aligned} 0 &= \det \begin{bmatrix} .95-\lambda & .03 \\ .05 & .97-\lambda \end{bmatrix} = (.95-\lambda)(.97-\lambda) - (.03)(.05) \\ &= \lambda^2 - 1.92\lambda + .92 \end{aligned}$$

Por la fórmula cuadrática,

$$\begin{aligned} \lambda &= \frac{1.92 \pm \sqrt{(1.92)^2 - 4(.92)}}{2} = \frac{1.92 \pm \sqrt{.0064}}{2} \\ &= \frac{1.92 \pm .08}{2} = 1 \quad \text{o bien} \quad .92 \end{aligned}$$

Es rápido comprobar que los vectores propios correspondientes a $\lambda = 1$ y $\lambda = 0.92$ son múltiplos de

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

respectivamente.

El siguiente paso es escribir $\mathbf{x}_0$ dada, en términos de $\mathbf{v}_1$ y $\mathbf{v}_2$. Esto se puede hacer porque $\{\mathbf{v}_1, \mathbf{v}_2\}$ evidentemente es una base para $\mathbb{R}^2$. (¿Por qué?) De manera que existen pesos c_1 y c_2 tales que

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = [\,\mathbf{v}_1 \quad \mathbf{v}_2\,]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{3}$$

De hecho,

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= [\,\mathbf{v}_1 \quad \mathbf{v}_2\,]^{-1}\mathbf{x}_0 = \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1} \begin{bmatrix} .60 \\ .40 \end{bmatrix} \\ &= \frac{1}{-8}\begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix}\begin{bmatrix} .60 \\ .40 \end{bmatrix} = \begin{bmatrix} .125 \\ .225 \end{bmatrix} \end{aligned} \tag{4}$$

Cada $\mathbf{x}_k$ es fácil de calcular, ya que $\mathbf{v}_1$ y $\mathbf{v}_2$ en la ecuación (3) son vectores propios de A, con $A\mathbf{v}_1 = \mathbf{v}_1$ y $A\mathbf{v}_2 = .92\mathbf{v}_2$:

$$\begin{aligned} \mathbf{x}_1 = A\mathbf{x}_0 &= c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 && \text{Utilizando linealidad de } \mathbf{x} \mapsto A\mathbf{x}, \\ &= c_1\mathbf{v}_1 + c_2(.92)\mathbf{v}_2 && \mathbf{v}_1 \text{ y } \mathbf{v}_2 \text{ son vectores propios.} \\ \mathbf{x}_2 = A\mathbf{x}_1 &= c_1A\mathbf{v}_1 + c_2(.92)A\mathbf{v}_2 \\ &= c_1\mathbf{v}_1 + c_2(.92)^2\mathbf{v}_2 \end{aligned}$$

y así sucesivamente. En general,

$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2(.92)^k\mathbf{v}_2 \quad (k = 0, 1, 2, \ldots)$$

Utilizando c_1 y c_2 de la ecuación (4),

$$\mathbf{x}_k = .125\begin{bmatrix} 3 \\ 5 \end{bmatrix} + .225(.92)^k\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (k = 0, 1, 2, \ldots) \tag{5}$$

Esta fórmula explícita para $\mathbf{x}_k$ da la solución de la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Conforme $k \to \infty$, $(.92)^k$ tiende a cero y $\mathbf{x}_k$ tiende a $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\mathbf{v}_1$. ■

Los cálculos del ejemplo 5 tienen una aplicación interesante a las cadenas de Markov analizadas en la sección 4.9. Quienes hayan leído esa sección reconocerán que la matriz A del ejemplo 5 anterior es la misma que la matriz M de migración de la sección 4.9, que $\mathbf{x}_0$ es la distribución de población inicial entre la ciudad y los suburbios, y que $\mathbf{x}_k$ representa la distribución de población después de k años.

El teorema 18 de la sección 4.9 estableció que para una matriz tal como A, la secuencia $\mathbf{x}_k$ tiende a un vector de estado estable. Ahora se sabe *por qué* las $\mathbf{x}_k$ se comportan de esa forma, al menos para la matriz de migración. El vector de estado estable es $.125\mathbf{v}_1$, un múltiplo del vector propio $\mathbf{v}_1$, y la fórmula (5) para $\mathbf{x}_k$ muestra precisamente por qué $\mathbf{x}_k \to .125\mathbf{v}_1$.

NOTAS NUMÉRICAS

1. El software computacional como Mathematica y Maple utilizan cálculos simbólicos para encontrar el polinomio característico de una matriz de tamaño moderado. Pero no hay una fórmula o un algoritmo finito para resolver la ecuación característica de una matriz general de $n \times n$ para $n \geq 5$.
2. Los mejores métodos numéricos para encontrar valores propios evitan totalmente el polinomio característico. En efecto, MATLAB determina el polinomio característico de una matriz A calculando primero los valores propios $\lambda_1, \ldots, \lambda_n$ de A y, luego, desarrollando el producto $(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$.
3. Varios algoritmos comunes para estimar los valores propios de una matriz A se basan en el teorema 4. En los ejercicios se analiza el poderoso *algoritmo QR*. Otra técnica, llamada *método de Jacobi*, funciona cuando $A = A^T$ y calcula una secuencia de matrices de la forma

$$A_1 = A \quad \text{y} \quad A_{k+1} = P_k^{-1} A_k P_k \quad (k = 1, 2, \ldots)$$

   Cada matriz en la secuencia es similar a A y así tiene los mismos valores propios que A. Las entradas no diagonales de A_{k+1} tienden a cero conforme aumenta k, y las entradas diagonales se van aproximando a los valores propios de A.
4. En la sección 5.8 se analizan otros métodos para calcular valores propios.

PROBLEMA DE PRÁCTICA

Encuentre la ecuación característica y los valores propios de $A = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

5.2 EJERCICIOS

En los ejercicios 1 a 8, encuentre el polinomio característico y los valores propios reales de las matrices dadas.

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

Los ejercicios 9 a 14 requieren las técnicas de la sección 3.1. Determine el polinomio característico de cada matriz, utilizando un desarrollo por cofactores o la fórmula especial para determinantes de 3×3 descrita antes de los ejercicios 15 a 18 de la sección 3.1. [*Nota:* Mediante operaciones de filas no es sencillo encontrar el polinomio característico de una matriz de 3×3, porque está implicada la variable λ].

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

11. $\begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 4 \\ 1 & 0 & 4 \end{bmatrix}$ **12.** $\begin{bmatrix} -1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

13. $\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 4 & 0 & -1 \\ -1 & 0 & 4 \\ 0 & 2 & 3 \end{bmatrix}$

Para las matrices de los ejercicios 15 a 17, liste los valores propios reales, repetidos de acuerdo con su multiplicidad.

15. $\begin{bmatrix} 5 & 5 & 0 & 2 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}$ **16.** $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 2 & 0 & 0 \\ 0 & 3 & 6 & 0 \\ 2 & 3 & 3 & -5 \end{bmatrix}$

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

18. Se puede demostrar que la multiplicidad algebraica de un valor propio λ es siempre mayor o igual que la dimensión del espacio propio correspondiente a λ. Determine h en la matriz A de abajo, tal que el espacio propio para $\lambda = 4$ sea bidimensional:

$$A = \begin{bmatrix} 4 & 2 & 3 & 3 \\ 0 & 2 & h & 3 \\ 0 & 0 & 4 & 14 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

19. Sea A una matriz de $n \times n$ y suponga que A tiene n valores propios reales, $\lambda_1, \ldots, \lambda_n$, repetidos de acuerdo con sus multiplicidades, tal que

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

Explique por qué $\det A$ es el producto de los n valores propios de A. (Este resultado es válido para cualquier matriz cuadrada cuando se consideran valores propios complejos).

20. Utilice una propiedad de determinantes para demostrar que A y A^T tienen el mismo polinomio característico.

En los ejercicios 21 y 22, A y B son matrices de $n \times n$. Indique si cada enunciado es verdadero o falso. Justifique cada respuesta.

21. *a*) El determinante de A es el producto de las entradas diagonales en A.

b) Una operación de filas elemental sobre A no cambia el valor del determinante.

c) $(\det A)(\det B) = \det AB$

d) Si $\lambda + 5$ es un factor del polinomio característico de A, entonces 5 es un valor propio de A.

22. *a*) Si A es 3×3, con columnas $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, entonces $\det A$ es igual al volumen del paralelepípedo formado por $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

b) $\det A^T = (-1) \det A$.

c) La multiplicidad de una raíz r de la ecuación característica de A se llama multiplicidad algebraica de r como un valor propio de A.

d) Una operación de remplazo por filas sobre A no altera sus valores propios.

Un método ampliamente utilizado para estimar los valores propios de una matriz general A es el *algoritmo QR*. En condiciones adecuadas, este algoritmo produce una secuencia de matrices, todas similares a A, casi triangulares superiores, con entradas diagonales que aproximan los valores propios de A. La idea principal es factorizar A (u otra matriz similar a A) en la forma $A = Q_1R_1$, donde $Q_1{}^T = Q_1{}^{-1}$ y R_1 es triangular superior. Los factores se intercambian para formar $A_1 = R_1Q_1$, que se factoriza de nuevo como $A_1 = Q_2R_2$; entonces, se forma $A_2 = R_2Q_2$, y así sucesivamente. La similitud de $A, A_1, \ldots$ se deduce del resultado más general del ejercicio 23.

23. Demuestre que si $A = QR$ con Q invertible, entonces A es similar a $A_1 = RQ$.

24. Demuestre que si A y B son similares, entonces $\det A = \det B$.

25. Sean $A = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. [*Nota:* A es la matriz estocástica estudiada en el ejemplo 5 de la sección 4.9].

a) Encuentre una base para $\mathbb{R}^2$ que consiste en $\mathbf{v}_1$ y otro vector propio $\mathbf{v}_2$ de A.

b) Compruebe que $\mathbf{x}_0$ se escribe en la forma $\mathbf{x}_0 = \mathbf{v}_1 + c\mathbf{v}_2$.

c) Para $k = 1,2,\ldots$, defina $\mathbf{x}_k = A^k\mathbf{x}_0$. Calcule $\mathbf{x}_1$ y $\mathbf{x}_2$, y escriba una fórmula para $\mathbf{x}_k$. Después, demuestre que $\mathbf{x}_k \to \mathbf{v}_1$ conforme k aumenta.

26. Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Utilice la fórmula (1) para un determinante (dada antes del ejemplo 2) para demostrar que $\det A = ad - bc$. Considere dos casos: $a \neq 0$ y $a = 0$.

27. Sean $A = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$,

$\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

a) Demuestre que $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son vectores propios de A. [*Nota:* A es la matriz estocástica estudiada en el ejemplo 3 de la sección 4.9].

b) Sea $\mathbf{x}_0$ cualquier vector en $\mathbb{R}^3$ con entradas no negativas cuya suma sea 1. (En la sección 4.9, a $\mathbf{x}_0$ se llamó vector de probabilidad). Explique por qué existen constantes c_1, c_2, c_3 tales que $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Calcule $\mathbf{w}^T\mathbf{x}_0$, y deduzca que $c_1 = 1$.

c) Para $k = 1, 2,\ldots$, defina $\mathbf{x}_k = A^k\mathbf{x}_0$, con $\mathbf{x}_0$ como en el inciso *b*). Demuestre que $\mathbf{x}_k \to \mathbf{v}_1$ cuando k aumenta.

28. [**M**] Construya una matriz A de 4×4, valuada en los enteros, y compruebe que A y A^T tienen el mismo polinomio característico (los mismos valores propios con las mismas multiplicidades). ¿A y A^T tienen los mismos vectores propios? Realice el mismo análisis con una matriz de 5×5. Escriba las matrices y sus conclusiones.

29. [**M**] Construya una matriz A de 4×4 valuada en los enteros.

a) Reduzca A a una forma escalonada U sin emplear escalamiento por filas, y utilice U en la fórmula (1) (que está antes del ejemplo 2) para calcular $\det A$. (Si A resulta ser singular, entonces inicie con una nueva matriz aleatoria).

b) Determine los valores propios de A y el producto de esos valores propios (tan exactamente como sea posible).

c) Escriba la matriz A y, con cuatro lugares decimales, liste los pivotes en U y los valores propios de A. Calcule $\det A$ con su programa matricial, y compárelo con los productos encontrados en los incisos *a*) y *b*).

30. [**M**] Sea $A = \begin{bmatrix} -6 & 28 & 21 \\ 4 & -15 & -12 \\ -8 & a & 25 \end{bmatrix}$. Para cada valor de a en el conjunto {32, 31.9, 31.8, 32.1, 32.2}, calcule el polinomio característico de A y los valores propios. En cada caso, trace una gráfica del polinomio característico $p(t) = \det(A - tI)$ para $0 \le t \le 3$. Si es posible, elabore todas las gráficas sobre un mismo sistema de coordenadas. Describa cómo las gráficas revelan los cambios en los valores propios conforme a cambia.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

La ecuación característica es

$$\begin{aligned} 0 = \det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & -4 \\ 4 & 2-\lambda \end{bmatrix} \\ &= (1-\lambda)(2-\lambda) - (-4)(4) = \lambda^2 - 3\lambda + 18 \end{aligned}$$

De la fórmula cuadrática,

$$\lambda = \frac{3 \pm \sqrt{(-3)^2 - 4(18)}}{2} = \frac{3 \pm \sqrt{-63}}{2}$$

Es evidente que la ecuación característica no tiene soluciones reales, así que A no tiene valores propios reales. La matriz A está actuando sobre el espacio vectorial real $\mathbb{R}^2$, y ahí no existe un vector $\mathbf{v}$ distinto de cero, tal que $A\mathbf{v} = \lambda\mathbf{v}$ para algún escalar λ.

5.3 DIAGONALIZACIÓN

En muchos casos, la información valor propio-vector propio contenida en una matriz A se presenta en una factorización útil de la forma $A = PDP^{-1}$ donde D es una matriz diagonal. En esta sección, la factorización permite calcular rápidamente las potencias A^k para valores grandes de k: una idea fundamental en varias aplicaciones del álgebra lineal. Posteriormente, en las secciones 5.6 y 5.7, se utilizará la factorización para analizar (y *desacoplar*) sistemas dinámicos.

El siguiente ejemplo muestra que las potencias de una matriz diagonal son fáciles de calcular.

EJEMPLO 1 Si $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, entonces, $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$

y

$$D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

En general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \quad \text{para } k \ge 1$$

■

Si $A = PDP^{-1}$ para alguna P invertible con D diagonal, entonces A^k es también fácil de calcular, como lo demuestra el siguiente ejemplo.

EJEMPLO 2 Sea $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Encuentre una fórmula para A^k, dado que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

SOLUCIÓN La fórmula estándar para la inversa de una matriz de 2×2 da

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

Entonces, por asociatividad de la multiplicación matricial,

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD\underbrace{(P^{-1}P)}_{I}DP^{-1} = PDDP^{-1}$$

$$= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Nuevamente,

$$A^3 = (PDP^{-1})A^2 = (PD\underbrace{P^{-1})P}_{I}D^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1}$$

En general, para $k \geq 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix} \qquad \blacksquare$$

Una matriz cuadrada A es **diagonalizable** si A es similar a una matriz diagonal, es decir, si $A = PDP^{-1}$ para alguna matriz P invertible y alguna matriz D diagonal. El siguiente teorema da una caracterización de matrices diagonalizables e indica cómo construir una factorización conveniente.

TEOREMA 5 **Teorema de diagonalización**

Una matriz A de $n \times n$ es diagonalizable, si y solo si A tiene n vectores propios linealmente independientes.

En efecto, $A = PDP^{-1}$, con D como matriz diagonal, si y solo si las columnas de P son n vectores propios linealmente independientes de A. En este caso, las entradas diagonales de D son valores propios de A que corresponden, respectivamente, a los vectores propios en P.

En otras palabras, A es diagonalizable si y solo si hay suficientes vectores propios para formar una base de $\mathbb{R}^n$. Tal base es una **base de vectores propios** de $\mathbb{R}^n$.

DEMOSTRACIÓN Primero, observe que si P es cualquier matriz de $n \times n$ con columnas $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y si D es cualquier matriz diagonal con entradas diagonales $\lambda_1, \ldots, \lambda_n$, entonces,

$$AP = A[\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n\,] = [\,A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n\,] \tag{1}$$

mientras que

$$PD = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\,\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n\,] \tag{2}$$

Ahora suponga que A es diagonalizable y $A = PDP^{-1}$. Por lo tanto, esta relación se multiplica por la derecha por P y se obtiene $AP = PD$. En este caso, las ecuaciones (1) y (2) implican que

$$[\, A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n \,] = [\, \lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n \,] \tag{3}$$

Igualando columnas,

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \quad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \quad \ldots, \quad A\mathbf{v}_n = \lambda_n\mathbf{v}_n \tag{4}$$

Puesto que P es invertible, sus columnas $\mathbf{v}_1,\ldots, \mathbf{v}_n$ deben ser linealmente independientes. Asimismo, como esas columnas no son cero, entonces las ecuaciones en (4) muestran que $\lambda_1,\ldots, \lambda_n$ son valores propios y $\mathbf{v}_1,\ldots, \mathbf{v}_n$ son los vectores propios correspondientes. Este argumento demuestra las partes "solo si" del primer y segundo enunciados, junto con el tercer enunciado, del teorema.

Por último, dados cualesquiera n vectores propios $\mathbf{v}_1,\ldots, \mathbf{v}_n$, úselos para construir las columnas de P y utilice los valores propios correspondientes $\lambda_1,\ldots, \lambda_n$ para formar D. Por las ecuaciones (1) a (3), $AP = PD$. Esto es válido sin condiciones sobre los vectores propios. Si, en efecto, los vectores propios son linealmente independientes, entonces P es invertible (según el teorema de la matriz invertible), y $AP = PD$ implica que $A = PDP^{-1}$. ■

Diagonalización de matrices

EJEMPLO 3 Diagonalice la siguiente matriz, si es posible.

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

Es decir, encuentre una matriz P invertible y una matriz D diagonal tales que $A = PDP^{-1}$.

SOLUCIÓN Hay cuatro pasos para implementar la descripción en el teorema 5.

Paso 1. Determine los valores propios de A. Como se mencionó en la sección 5.2, la mecánica de este paso es adecuada para una computadora, cuando la matriz es mayor que 2×2. Para eliminar distracciones innecesarias, el texto generalmente dará información adicional de utilidad para este paso. En el presente caso, la ecuación característica implica un polinomio cúbico que se factoriza como:

$$\begin{aligned} 0 = \det(A - \lambda I) &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

Los valores propios son $\lambda = 1$ y $\lambda = -2$.

Paso 2. Encuentre tres vectores propios de A linealmente independientes. Se necesitan *tres* vectores porque A es una matriz de 3×3. Este es el paso crítico. Si falla, entonces el teorema 5 indica que A no se puede diagonalizar. El método de la sección 5.1 produce una base para cada espacio propio:

$$\text{Base para } \lambda = 1: \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -2: \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Puede comprobar que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto linealmente independiente.

Paso 3. Construya P con los vectores del paso 2. No es importante el orden de los vectores. Utilizando el orden seleccionado en el paso 2, forme

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Paso 4. Construya D con los valores propios correspondientes. En este paso, resulta esencial que el orden de los valores propios coincida con el orden elegido para las columnas de P. Utilice dos veces el valor propio $\lambda = -2$, uno para cada vector propio correspondiente a $\lambda = -2$:

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

Es buena idea comprobar que P y D realmente funcionen. Para evitar el cálculo de P^{-1}, simplemente compruebe que $AP = PD$. Esto es equivalente a $A = PDP^{-1}$ cuando P es invertible. (Sin embargo, ¡asegúrese que P sea invertible!) Calcule

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

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

EJEMPLO 4 Si es posible, diagonalice la siguiente matriz.

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

SOLUCIÓN La ecuación característica de A resulta ser exactamente la misma que en el ejemplo 3:

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$$

Los valores propios son $\lambda = 1$ y $\lambda = -2$. No obstante, es fácil comprobar que cada espacio propio es tan solo unidimensional:

$$\text{Base para } \lambda = 1: \qquad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -2: \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

No existen más valores propios, en tanto que cada vector propio de A es un múltiplo de $\mathbf{v}_1$ o de $\mathbf{v}_2$. Por consiguiente, es imposible construir una base de $\mathbb{R}^3$ utilizando vectores propios de A. De acuerdo con el teorema 5, A *no es* diagonalizable. ■

El siguiente teorema proporciona una condición *suficiente* para que una matriz sea diagonalizable.

TEOREMA 6 Una matriz de $n \times n$ con n valores propios distintos es diagonalizable.

DEMOSTRACIÓN Sean $\mathbf{v}_1, \ldots, \mathbf{v}_n$ los vectores propios correspondientes a los n valores propios diferentes de una matriz A. Entonces, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es linealmente independiente, de acuerdo con el teorema 2 de la sección 5.1. Por lo tanto, A es diagonalizable, según el teorema 5. ■

No es *necesario* que una matriz de $n \times n$ tenga n valores propios distintos para ser diagonalizable. La matriz de 3×3 del ejemplo 3 es diagonalizable aun cuando tenga solamente dos valores propios diferentes.

EJEMPLO 5 Determine si la siguiente matriz es diagonalizable.

$$A = \begin{bmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{bmatrix}$$

SOLUCIÓN ¡Esto es fácil! Como la matriz es triangular, sus valores propios son evidentemente 5, 0 y -2. Al ser A una matriz de 3×3 con tres valores propios distintos, entonces A es diagonalizable. ■

Matrices cuyos valores propios no son diferentes

Si una matriz A de $n \times n$ tiene n valores propios distintos, con los vectores propios correspondientes $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y si $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, entonces P es automáticamente invertible porque sus columnas son linealmente independientes, según el teorema 2. Cuando A es diagonalizable, pero tiene menos de n valores propios diferentes, sigue siendo posible construir P en una forma que la hace automáticamente invertible, como lo demuestra el siguiente teorema.[1]

TEOREMA 7

Sea A una matriz de $n \times n$ cuyos distintos valores propios son $\lambda_1, \ldots, \lambda_p$.

a) Para $1 \le k \le p$, la dimensión del espacio propio para λ_k es menor o igual que la multiplicidad del valor propio λ_k.

b) La matriz A es diagonalizable si y solo si la suma de las dimensiones de los espacios propios es igual a n, y esto ocurre si y solo si: **i.** se factoriza completamente el polinomio característico en factores lineales y **ii.** la dimensión del espacio propio para cada λ_k es igual a la multiplicidad de λ_k.

c) Si A es diagonalizable y $\mathcal{B}_k$ es una base para el espacio propio asociado con λ_k para cada k, entonces la colección de vectores total en los conjuntos $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forma una base de vectores propios para $\mathbb{R}^n$.

EJEMPLO 6 Diagonalice la siguiente matriz, si es posible.

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

[1] La demostración del teorema 7 es larga, pero no difícil. Por ejemplo, véase S. Friedberg, A. Insel, y L. Spence, *Linear Algebra*, 4a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 2002), sección 5.2.

SOLUCIÓN Puesto que A es una matriz triangular, los valores propios son 5 y -3, cada uno con multiplicidad 2. Utilizando el método de la sección 5.1, se encuentra una base para cada espacio propio.

$$\text{Base para } \lambda = 5: \quad \mathbf{v}_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -3: \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

El conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es linealmente independiente, según el teorema 7. Así, la matriz $P = [\mathbf{v}_1 \cdots \mathbf{v}_4]$ es invertible, en tanto que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$ ■

PROBLEMAS DE PRÁCTICA

1. Calcule A^8, donde $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.
2. Sean $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suponga que le han dicho que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios de A. Utilice esta información para diagonalizar A.
3. Sea A una matriz de 4×4 con valores propios 5, 3 y -2, y suponga que se sabe que el espacio propio para $\lambda = 3$ es bidimensional. ¿Se dispone de suficiente información para determinar si A es diagonalizable?

5.3 EJERCICIOS

En los ejercicios 1 y 2, sea $A = PDP^{-1}$ y calcule A^4.

1. $P = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}, D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

2. $P = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

En los ejercicios 3 y 4, use la factorización $A = PDP^{-1}$ para calcular A^k, donde k representa un entero positivo arbitrario.

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

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

En los ejercicios 5 y 6, la matriz A se factoriza en la forma PDP^{-1}. Utilice el teorema de diagonalización para encontrar los valores propios de A y una base para cada espacio propio.

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

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

6. $A = \begin{bmatrix} 3 & 0 & 0 \\ -3 & 4 & 9 \\ 0 & 0 & 3 \end{bmatrix}$

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

En los ejercicios 7 a 20, diagonalice las matrices, si es posible. Para los ejercicios 11 a 16 y 18, se incluyen los valores propios reales debajo de la matriz.

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

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

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

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

11. $\begin{bmatrix} 0 & 1 & 1 \\ 2 & 1 & 2 \\ 3 & 3 & 2 \end{bmatrix}$
$\lambda = -1,\ 5$

12. $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$
$\lambda = 2,\ 5$

13. $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$
$\lambda = 1,\ 5$

14. $\begin{bmatrix} 2 & 0 & -2 \\ 1 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$
$\lambda = 2,\ 3$

15. $\begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$
$\lambda = 0,\ 1$

16. $\begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -2 \\ 1 & 3 & 1 \end{bmatrix}$
$\lambda = 0$

17. $\begin{bmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix}$

18. $\begin{bmatrix} 2 & -2 & -2 \\ 3 & -3 & -2 \\ 2 & -2 & -2 \end{bmatrix}$
$\lambda = -2, -1, 0$

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

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

En los ejercicios 21 y 22, A, B, P y D son matrices de $n \times n$. Indique si cada enunciado es verdadero o falso. Justifique cada respuesta. (Antes de intentar resolver estos ejercicios, estudie con cuidado los teoremas 5 y 6, y los ejemplos de esta sección).

21. *a*) A es diagonalizable si $A = PDP^{-1}$ para alguna matriz D y alguna matriz P invertible.

b) Si $\mathbb{R}^n$ tiene una base de vectores propios de A, entonces A es diagonalizable.

c) A es diagonalizable si y solo si A tiene n valores propios, contando multiplicidades.

d) Si A es diagonalizable, entonces A es invertible.

22. *a*) A es diagonalizable si A tiene n vectores propios.

b) Si A es diagonalizable, entonces A tiene n valores propios distintos.

c) Si $AP = PD$, con D diagonal, entonces las columnas distintas de cero de P deben ser vectores propios de A.

d) Si A es invertible, entonces A es diagonalizable.

23. A es una matriz de 5×5 con dos valores propios. Un espacio propio es tridimensional y el otro espacio propio es bidimensional. ¿A es diagonalizable? ¿Por qué?

24. A es una matriz de 3×3 con dos valores propios. Cada espacio propio es unidimensional. ¿A es diagonalizable? ¿Por qué?

25. A es una matriz de 4×4 con tres valores propios. Un espacio propio es unidimensional, y uno de los otros espacios propios es bidimensional. ¿Es posible que A *no* sea diagonalizable? Justifique su respuesta.

26. A es una matriz de 7×7 con tres valores propios. Un espacio propio es bidimensional, y uno de los otros espacios propios es tridimensional. ¿Es posible que A *no* sea diagonalizable? Justifique su respuesta.

27. Demuestre que si A es diagonalizable e invertible, entonces también lo es A^{-1}.

28. Demuestre que si A tiene n vectores propios linealmente independientes, entonces ocurre lo mismo con A^T. [*Sugerencia:* Use el teorema de diagonalización].

29. Una factorización $A = PDP^{-1}$ no es única. Demuestre esto para la matriz A del ejemplo 2. Con $D_1 = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$, utilice la información del ejemplo 2 para encontrar una matriz P_1 tal que $A = P_1D_1P_1{}^{-1}$.

30. Con A y D como en el ejemplo 2, encuentre una P_2 invertible, distinta de la P del ejemplo 2, tal que $A = P_2DP_2{}^{-1}$.

31. Construya una matriz de 2×2 distinta de cero que sea invertible, pero no diagonalizable.

32. Construya una matriz de 2×2 no diagonal que sea diagonalizable, pero no invertible.

[**M**] En los ejercicios 33 a 36, diagonalice las matrices. Utilice la instrucción de valores propios de su programa matricial para obtener los valores propios y, luego, calcule bases para los espacios propios como en la sección 5.1.

33. $\begin{bmatrix} 9 & -4 & -2 & -4 \\ -56 & 32 & -28 & 44 \\ -14 & -14 & 6 & -14 \\ 42 & -33 & 21 & -45 \end{bmatrix}$

34. $\begin{bmatrix} 4 & -9 & -7 & 8 & 2 \\ -7 & -9 & 0 & 7 & 14 \\ 5 & 10 & 5 & -5 & -10 \\ -2 & 3 & 7 & 0 & 4 \\ -3 & -13 & -7 & 10 & 11 \end{bmatrix}$

35. $\begin{bmatrix} 13 & -12 & 9 & -15 & 9 \\ 6 & -5 & 9 & -15 & 9 \\ 6 & -12 & -5 & 6 & 9 \\ 6 & -12 & 9 & -8 & 9 \\ -6 & 12 & 12 & -6 & -2 \end{bmatrix}$

36. $\begin{bmatrix} 24 & -6 & 2 & 6 & 2 \\ 72 & 51 & 9 & -99 & 9 \\ 0 & -63 & 15 & 63 & 63 \\ 72 & 15 & 9 & -63 & 9 \\ 0 & 63 & 21 & -63 & -27 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. det $(A - \lambda I) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$. Los valores propios son 2 y 1, y los vectores propios correspondientes son $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Ahora, forme

$$P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

Como $A = PDP^{-1}$,

$$\begin{aligned} A^8 = PD^8P^{-1} &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix} \end{aligned}$$

2. Calcule $A\mathbf{v}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{v}_1$, y

$$A\mathbf{v}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \cdot \mathbf{v}_2$$

De manera que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios para los valores propios 1 y 3, respectivamente. Entonces,

$$A = PDP^{-1} \quad \text{donde,} \quad P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

3. Sí, A es diagonalizable. Existe una base $\{\mathbf{v}_1, \mathbf{v}_2\}$ para el espacio propio asociado a $\lambda = 3$. Asimismo, habrá al menos un vector propio para $\lambda = 5$ y uno para $\lambda = -2$, denotados mediante $\mathbf{v}_3$ y $\mathbf{v}_4$. Entonces, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es linealmente independiente de acuerdo con el teorema 2 y el problema de práctica 3 de la sección 5.1. No pueden existir más vectores propios que sean linealmente independientes de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, porque los vectores están en $\mathbb{R}^4$. Por consiguiente, los espacios propios para $\lambda = 5$ y $\lambda = -2$ son ambos unidimensionales. De acuerdo con el teorema 7*b*) se tiene que A es diagonalizable.

5.4 VECTORES PROPIOS Y TRANSFORMACIONES LINEALES

El objetivo de esta sección es entender la factorización matricial $A = PDP^{-1}$ como un enunciado sobre transformaciones lineales. Se verá que la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es esencialmente lo mismo que el muy simple mapeo $\mathbf{u} \mapsto D\mathbf{u}$, cuando se estudia desde un enfoque adecuado. Una interpretación similar se aplicará a D y A, aun cuando D no sea una matriz diagonal.

De la sección 1.9, recuerde que cualquier transformación lineal T de $\mathbb{R}^n$ a $\mathbb{R}^m$ se implementa mediante la multiplicación por la izquierda por una matriz A, llamada *matriz estándar* de T. Ahora se necesita el mismo tipo de representación para cualquier transformación lineal entre dos espacios vectoriales de dimensión finita.

La matriz de una transformación lineal

Sean V un espacio vectorial n-dimensional, W un espacio vectorial m-dimensional y T cualquier transformación lineal de V a W. Para asociar una matriz con T, se eligen bases (ordenadas) $\mathcal{B}$ y $\mathcal{C}$ para V y W, respectivamente.

Dada cualquier $\mathbf{x}$ en V, el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ está en $\mathbb{R}^n$ y el vector de coordenadas de su imagen, $[T(\mathbf{x})]_{\mathcal{C}}$, está en $\mathbb{R}^m$, como se indica en la figura 1.

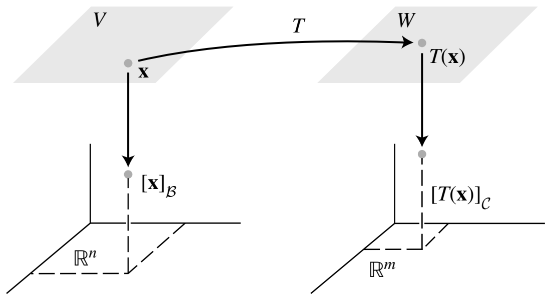


FIGURA 1 Una transformación lineal de V a W.

Es fácil encontrar la conexión entre $[\mathbf{x}]_{\mathcal{B}}$ y $[T(\mathbf{x})]_{\mathcal{C}}$. Sea $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ la base $\mathcal{B}$ para V. Si $\mathbf{x} = r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n$, entonces,

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

y

$$T(\mathbf{x}) = T(r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n) = r_1T(\mathbf{b}_1) + \cdots + r_nT(\mathbf{b}_n) \tag{1}$$

porque T es lineal. Ahora, como el mapeo coordenado de W a $\mathbb{R}^m$ es lineal (teorema 8 de la sección 4.4), la ecuación (1) conduce a

$$[\,T(\mathbf{x})\,]_{\mathcal{C}} = r_1[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} + \cdots + r_n[\,T(\mathbf{b}_n)\,]_{\mathcal{C}} \tag{2}$$

Puesto que los vectores de coordenadas C están en $\mathbb{R}^m$, así la ecuación vectorial (2) se escribe como una ecuación matricial; a saber,

$$[T(\mathbf{x})]_{\mathcal{C}} = M[\mathbf{x}]_{\mathcal{B}} \tag{3}$$

donde

$$M = \Big[[T(\mathbf{b}_1)]_{\mathcal{C}} \quad [T(\mathbf{b}_2)]_{\mathcal{C}} \quad \cdots \quad [T(\mathbf{b}_n)]_{\mathcal{C}}\Big] \tag{4}$$

$$\begin{array}{ccc} \mathbf{x} & \xrightarrow{\quad T \quad} & T(\mathbf{x}) \\ \downarrow & & \downarrow \\ [\mathbf{x}]_{\mathcal{B}} & \xrightarrow[\text{por } M]{\text{Multiplicación}} & [T(\mathbf{x})]_{\mathcal{C}} \end{array}$$

FIGURA 2

La matriz M es una representación matricial de T, llamada **matriz para T respecto a las bases $\mathcal{B}$ y $\mathcal{C}$**. Véase la figura 2.

La ecuación (3) indica que, en lo concerniente a los vectores de coordenadas, la acción de T sobre $\mathbf{x}$ se puede considerar una multiplicación por la izquierda por M.

EJEMPLO 1 Suponga que $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ es una base para V y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ es una base para W. Sea $T : V \to W$ una transformación lineal con la propiedad

$$T(\mathbf{b}_1) = 3\mathbf{c}_1 - 2\mathbf{c}_2 + 5\mathbf{c}_3 \quad \text{y} \quad T(\mathbf{b}_2) = 4\mathbf{c}_1 + 7\mathbf{c}_2 - \mathbf{c}_3$$

Determine la matriz M para T respecto a $\mathcal{B}$ y $\mathcal{C}$.

SOLUCIÓN Los vectores de coordenadas $\mathcal{C}$ de las *imágenes* de $\mathbf{b}_1$ y $\mathbf{b}_2$ son

$$[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \quad \text{y} \quad [\,T(\mathbf{b}_2)\,]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

Por lo tanto,

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$ ■

Si $\mathcal{B}$ y $\mathcal{C}$ son bases para el mismo espacio V y si T es la transformación identidad $T(\mathbf{x}) = \mathbf{x}$ para $\mathbf{x}$ en V, entonces la matriz M en la ecuación (4) es justo una matriz de cambio de coordenadas (véase la sección 4.7).

Transformaciones lineales de V en V

$$\begin{tikzcd} \mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\ {[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplicación}", "\text{por } [T]_{\mathcal{B}}"'] & {[T(\mathbf{x})]_{\mathcal{B}}} \end{tikzcd}$$

FIGURA 3

En el caso común donde W es igual a V y la base $\mathcal{C}$ coincide con $\mathcal{B}$, la matriz M en (4) se denomina **matriz para *T* respecto a $\mathcal{B}$**, o simplemente **$\mathcal{B}$-matriz para *T***, y se denota con $[T]_{\mathcal{B}}$. Véase la figura 3.

La $\mathcal{B}$-matriz para $T: V \to V$ satisface

$$[\,T(\mathbf{x})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}[\,\mathbf{x}\,]_{\mathcal{B}} \quad \text{para toda } \mathbf{x} \text{ en } V \tag{5}$$

EJEMPLO 2 El mapeo $T: \mathbb{P}_2 \to \mathbb{P}_2$ definido por

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

es una transformación lineal. (Los estudiantes de cálculo reconocerán T como el operador derivada).

a) Determine la $\mathcal{B}$-matriz para T, cuando $\mathcal{B}$ es la base $\{1, t, t^2\}$.

b) Compruebe que $[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}[\,\mathbf{p}\,]_{\mathcal{B}}$ para toda $\mathbf{p}$ en $\mathbb{P}_2$.

SOLUCIÓN

a) Determine las imágenes de los vectores básicos:

$T(1) = 0$ El polinomio cero

$T(t) = 1$ El polinomio cuyo valor siempre es 1

$T(t^2) = 2t$

Después, escriba los vectores de $\mathcal{B}$-coordenadas de $T(1)$, $T(t)$ y $T(t^2)$ (que en este ejemplo se localizan por inspección) y júntelos en la $\mathcal{B}$-matriz para T:

$$[\,T(1)\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [\,T(t)\,]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [\,T(t^2)\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

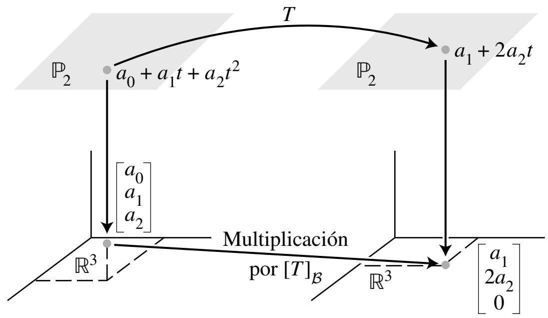
$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

b) Para un general $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$,

$$[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,a_1 + 2a_2 t\,]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [\,T\,]_{\mathcal{B}}[\,\mathbf{p}\,]_{\mathcal{B}}$$

Véase la figura 4. ■



WEB

FIGURA 4 Representación matricial de una transformación lineal.

Transformaciones lineales sobre $\mathbb{R}^n$

En un problema aplicado que implica a $\mathbb{R}^n$, en general una transformación lineal T aparece primero como una transformación matricial, $\mathbf{x} \mapsto A\mathbf{x}$. Si A es diagonalizable, entonces hay una base $\mathcal{B}$ para $\mathbb{R}^n$ que consiste en vectores propios de A. A continuación, el teorema 8 demuestra que, en este caso, la $\mathcal{B}$-matriz para T es diagonal. Diagonalizar A equivale a encontrar una representación matricial diagonal $\mathbf{x} \mapsto A\mathbf{x}$.

TEOREMA 8

Representación matricial diagonal

Suponga que $A = PDP^{-1}$, donde D es una matriz de $n \times n$ diagonal. Si $\mathcal{B}$ es la base para $\mathbb{R}^n$ formada con las columnas de P, entonces D es la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$.

DEMOSTRACIÓN Denote las columnas de P con $\mathbf{b}_1, \ldots, \mathbf{b}_n$, de manera que $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $P = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$. En este caso, P es la matriz de cambio de coordenadas $P_{\mathcal{B}}$ analizada en la sección 4.4, donde

$$P[\,\mathbf{x}\,]_{\mathcal{B}} = \mathbf{x} \quad \text{y} \quad [\,\mathbf{x}\,]_{\mathcal{B}} = P^{-1}\mathbf{x}$$

Si $T(\mathbf{x}) = A\mathbf{x}$ para $\mathbf{x}$ en $\mathbb{R}^n$, entonces,

$$\begin{aligned}
[\,T\,]_{\mathcal{B}} &= \big[\,[\,T(\mathbf{b}_1)\,]_{\mathcal{B}} \quad \cdots \quad [\,T(\mathbf{b}_n)\,]_{\mathcal{B}}\,\big] && \text{Definición de } [T]_{\mathcal{B}} \\
&= \big[\,[\,A\mathbf{b}_1\,]_{\mathcal{B}} \quad \cdots \quad [\,A\mathbf{b}_n\,]_{\mathcal{B}}\,\big] && \text{Porque } T(\mathbf{x}) = A\mathbf{x} \\
&= [\,P^{-1}A\mathbf{b}_1 \quad \cdots \quad P^{-1}A\mathbf{b}_n\,] && \text{Cambio de coordenadas} \\
&= P^{-1}A[\,\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n\,] && \text{Multiplicación matricial} \\
&= P^{-1}AP && (6)
\end{aligned}$$

Puesto que $A = PDP^{-1}$, entonces se tiene $[\,T\,]_{\mathcal{B}} = P^{-1}AP = D$. ■

EJEMPLO 3 Defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ por $T(\mathbf{x}) = A\mathbf{x}$, donde $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Encuentre una base $\mathcal{B}$ para $\mathbb{R}^2$ con la propiedad de que la $\mathcal{B}$-matriz para T es una matriz diagonal.

SOLUCIÓN Del ejemplo 2 de la sección 5.3, se conoce que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Las columnas de P, denotadas mediante $\mathbf{b}_1$ y $\mathbf{b}_2$, son vectores propios de A. Según el teorema 8, D es la $\mathcal{B}$-matriz para T cuando $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Los mapeos $\mathbf{x} \mapsto A\mathbf{x}$ y $\mathbf{u} \mapsto D\mathbf{u}$ describen la misma transformación lineal respecto a diferentes bases. ■

Similitud de representaciones matriciales

La demostración del teorema 8 no utilizó la información de que D era diagonal. Por lo tanto, si A es similar a una matriz C, con $A = PCP^{-1}$, entonces C es la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ cuando la base $\mathcal{B}$ está formada por las columnas de P. La figura 5 ilustra la factorización $A = PCP^{-1}$.

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "\text{Multiplicación}", "\text{por } A"'] \arrow[d, "\text{Multiplicación}"', "\text{por } P^{-1}"'] & A\mathbf{x} \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplicación}", "\text{por } C"'] & {[A\mathbf{x}]_{\mathcal{B}}} \arrow[u, "\text{Multiplicación}"', "\text{por } P"]
\end{tikzcd}$$

FIGURA 5 Similitud de dos representaciones matriciales: $A = PCP^{-1}$.

Inversamente, si $T : \mathbb{R}^n \to \mathbb{R}^n$ está definida por $T(\mathbf{x}) = A\mathbf{x}$, y si $\mathcal{B}$ es cualquier base para $\mathbb{R}^n$, entonces la $\mathcal{B}$-matriz para T es similar a A. En efecto, los cálculos en la demostración del teorema 8 indican que si P es la matriz cuyas columnas provienen de los vectores en $\mathcal{B}$, entonces, $[T]_{\mathcal{B}} = P^{-1}AP$. Así, el conjunto de todas las matrices similares a una matriz A coincide con el conjunto de todas las representaciones matriciales de la transformación $\mathbf{x} \mapsto A\mathbf{x}$.

EJEMPLO 4 Sean $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. El polinomio característico de A es $(\lambda + 2)^2$, sin embargo, el espacio propio para el valor propio -2 es unidimensional; de manera que A no es diagonalizable. Sin embargo, la base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ tiene la propiedad de que la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es una matriz triangular llamada *forma de Jordan* de A.[1] Determine esta $\mathcal{B}$-matriz.

SOLUCIÓN Si $P = [\mathbf{b}_1 \quad \mathbf{b}_2]$, entonces la $\mathcal{B}$-matriz es $P^{-1}AP$. Calcule

$$AP = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix}$$

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

Observe que el valor propio de A está sobre la diagonal. ■

[1] Cada matriz cuadrada A es similar a una matriz en forma de Jordan. La base utilizada para generar una forma de Jordan consiste en vectores propios y los llamados "vectores propios generalizados" de A. Véase el capítulo 9 de *Applied Linear Algebra*, 3a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 1988), de B. Noble y J. W. Daniel.

NOTA NUMÉRICA

Una forma eficiente de calcular una $\mathcal{B}$-matriz de $P^{-1}AP$ es determinando AP y, luego, reduciendo por filas la matriz aumentada $[P \quad AP]$ a $[I \quad P^{-1}AP]$. Es innecesario un cálculo por separado de P^{-1}. Véase el ejercicio 15 de la sección 2.2.

PROBLEMAS DE PRÁCTICA

1. Determine $T(a_0 + a_1t + a_2t^2)$, si T es la transformación lineal de $\mathbb{P}_2$ a $\mathbb{P}_2$ cuya matriz respecto a $\mathcal{B} = \{1, t, t^2\}$ es

$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

2. Sean A, B y C matrices de $n \times n$. En el texto se ha demostrado que si A es similar a B, entonces B es similar a A. Esta propiedad, junto con los enunciados que se muestran debajo, indica que “similar a” es una *relación de equivalencia*. (La equivalencia por fila es otro ejemplo de una relación de equivalencia). Compruebe los incisos *a*) y *b*).

   a) A es similar a A.

   b) Si A es similar a B, y B es similar a C, entonces A es similar a C.

5.4 EJERCICIOS

1. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ las bases de los espacios vectoriales V y W, respectivamente. Sea $T : V \to W$ una transformación lineal con la propiedad

$$T(\mathbf{b}_1) = 3\mathbf{d}_1 - 5\mathbf{d}_2, \quad T(\mathbf{b}_2) = -\mathbf{d}_1 + 6\mathbf{d}_2, \quad T(\mathbf{b}_3) = 4\mathbf{d}_2$$

   Determine la matriz para T respecto a $\mathcal{B}$ y $\mathcal{D}$.

2. Sean $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ las bases para los espacios vectoriales V y W, respectivamente. Sea $T : V \to W$ una transformación lineal con la propiedad

$$T(\mathbf{d}_1) = 3\mathbf{b}_1 - 3\mathbf{b}_2, \quad T(\mathbf{d}_2) = -2\mathbf{b}_1 + 5\mathbf{b}_2$$

   Determine la matriz para T respecto a $\mathcal{D}$ y $\mathcal{B}$.

3. Sean $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base estándar para $\mathbb{R}^3$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V, y $T : \mathbb{R}^3 \to V$ una transformación lineal con la propiedad

$$T(x_1, x_2, x_3) = (2x_3 - x_2)\mathbf{b}_1 - (2x_2)\mathbf{b}_2 + (x_1 + 3x_3)\mathbf{b}_3$$

   a) Calcule $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$ y $T(\mathbf{e}_3)$.

   b) Obtenga $[T(\mathbf{e}_1)]_{\mathcal{B}}$, $[T(\mathbf{e}_2)]_{\mathcal{B}}$ y $[T(\mathbf{e}_3)]_{\mathcal{B}}$.

   c) Determine la matriz para T respecto a $\mathcal{E}$ y $\mathcal{B}$.

4. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V y $T : V \to \mathbb{R}^2$ una transformación lineal con la propiedad

$$T(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3) = \begin{bmatrix} 2x_1 - 3x_2 + x_3 \\ -2x_1 + 5x_3 \end{bmatrix}$$

   Encuentre la matriz para T respecto a $\mathcal{B}$ y la base estándar para $\mathbb{R}^2$.

5. Sea $T : \mathbb{P}_2 \to \mathbb{P}_3$ la transformación que mapea un polinomio $\mathbf{p}(t)$ en el polinomio $(t + 3)\mathbf{p}(t)$.

   a) Encuentre la imagen de $\mathbf{p}(t) = 3 - 2t + t^2$.

   b) Demuestre que T es una transformación lineal.

   c) Obtenga la matriz para T respecto a las bases $\{1, t, t^2\}$ y $\{1, t, t^2, t^3\}$.

6. Sea $T : \mathbb{P}_2 \to \mathbb{P}_4$ la transformación que mapea un polinomio $\mathbf{p}(t)$ en el polinomio $\mathbf{p}(t) + 2t^2\mathbf{p}(t)$.

   a) Determine la imagen de $\mathbf{p}(t) = 3 - 2t + t^2$.

   b) Demuestre que T es una transformación lineal.

   c) Encuentre la matriz para T respecto a las bases $\{1, t, t^2\}$ y $\{1, t, t^2, t^3, t^4\}$.

7. Suponga que el mapeo $T : \mathbb{P}_2 \to \mathbb{P}_2$ definido por

$$T(a_0 + a_1t + a_2t^2) = 3a_0 + (5a_0 - 2a_1)t + (4a_1 + a_2)t^2$$

   es lineal. Obtenga la representación matricial de T respecto a la base $\mathcal{B} = \{1, t, t^2\}$.

8. Sea $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V. Encuentre $T(4\mathbf{b}_1 - 3\mathbf{b}_2)$ cuando T es una transformación lineal de V a V cuya matriz respecto a $\mathcal{B}$ es

$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & -2 \\ 1 & 3 & 1 \end{bmatrix}$$

9. Defina $T : \mathbb{P}_2 \to \mathbb{R}^3$ por $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-1) \\ \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$.

a) Encuentre la imagen bajo T de $\mathbf{p}(t) = 5 + 3t$.

b) Pruebe que T es una transformación lineal.

c) Obtenga la matriz para T respecto a la base $\{1, t, t^2\}$ para $\mathbb{P}_2$ y la base estándar para $\mathbb{R}^3$.

10. Defina $T : \mathbb{P}_3 \to \mathbb{R}^4$ por $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-2) \\ \mathbf{p}(3) \\ \mathbf{p}(1) \\ \mathbf{p}(0) \end{bmatrix}$.

a) Demuestre que T es una transformación lineal.

b) Encuentre la matriz para T respecto a la base $\{1, t, t^2, t^3\}$ para $\mathbb{P}_3$ y la base estándar para $\mathbb{R}^4$.

En los ejercicios 11 y 12, determine la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ donde $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

11. $A = \begin{bmatrix} -4 & -1 \\ 6 & 1 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

12. $A = \begin{bmatrix} -6 & -2 \\ 4 & 0 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

En los ejercicios 13 a 16, defina $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ por $T(\mathbf{x}) = A\mathbf{x}$. Determina una base $\mathcal{B}$ para $\mathbb{R}^2$ con la propiedad de que $[T]_{\mathcal{B}}$ es diagonal.

13. $A = \begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}$

14. $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$

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

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

17. Sean $A = \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ para $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(\mathbf{x}) = A\mathbf{x}$.

a) Compruebe que $\mathbf{b}_1$ es un vector propio de A, pero que A no es diagonalizable.

b) Encuentre la $\mathcal{B}$-matriz para T.

18. Defina $T : \mathbb{R}^3 \to \mathbb{R}^3$ mediante $T(\mathbf{x}) = A\mathbf{x}$, donde A es una matriz de 3×3 con valores propios de 5, 5 y -2. ¿Existe una base $\mathcal{B}$ para $\mathbb{R}^3$ tal que la $\mathcal{B}$-matriz para T sea una matriz diagonal? Analice.

En los ejercicios 19 a 24, compruebe los enunciados. Las matrices son cuadradas.

19. Si A es invertible y similar a B, entonces B es invertible y A^{-1} es similar a B^{-1}. [*Sugerencia:* $P^{-1}AP = B$ para alguna P invertible. Explique por qué B es invertible. Luego, determine una Q invertible tal que $Q^{-1}A^{-1}Q = B^{-1}$].

20. Si A es similar a B, entonces A^2 es similar a B^2.

21. Si B es similar a A y C es similar a A, entonces B es similar a C.

22. Si A es diagonalizable y B es similar a A, entonces B también es diagonalizable.

23. Si $B = P^{-1}AP$ y $\mathbf{x}$ es un vector propio de A correspondiente a un valor propio λ, entonces $P^{-1}\mathbf{x}$ es un vector propio de B también asociado con λ.

24. Si A y B son similares, entonces tienen el mismo rango (dimensión del espacio imagen). [*Sugerencia:* Consulte los ejercicios complementarios 13 y 14 del capítulo 4].

25. La *traza* de una matriz cuadrada A es la suma de las entradas diagonales en A y se denota con tr A. Se puede comprobar que $\mathrm{tr}(FG) = \mathrm{tr}(GF)$ para cualesquiera dos matrices F y G de $n \times n$. Demuestre que si A y B son similares, entonces $\mathrm{tr}\,A = \mathrm{tr}\,B$.

26. Se puede demostrar que la traza de una matriz A es igual a la suma de los valores propios de A. Compruebe este enunciado para el caso cuando A es diagonalizable.

27. Sean $V = \mathbb{R}^n$ con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, y $W = \mathbb{R}^n$ con la base estándar que se denota mediante $\mathcal{E}$; y considere la transformación identidad $I : \mathbb{R}^n \to \mathbb{R}^n$, donde $I(\mathbf{x}) = \mathbf{x}$. Obtenga la matriz para I respecto a $\mathcal{B}$ y $\mathcal{E}$. ¿Cómo se llamó a esta matriz en la sección 4.4?

28. Sean V un espacio vectorial con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y W el mismo espacio V con una base $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$; I es la transformación identidad $I : V \to W$. Encuentre la matriz para I respecto a $\mathcal{B}$ y $\mathcal{C}$. ¿Cómo se llamó a esta matriz en la sección 4.7?

29. Sea V un espacio vectorial con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Determine la $\mathcal{B}$-matriz para la transformación identidad $I : V \to V$.

[**M**] En los ejercicios 30 y 31, determine la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ donde $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

30. $A = \begin{bmatrix} 6 & -2 & -2 \\ 3 & 1 & -2 \\ 2 & -2 & 2 \end{bmatrix}$,

$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$

31. $A = \begin{bmatrix} -7 & -48 & -16 \\ 1 & 14 & 6 \\ -3 & -45 & -19 \end{bmatrix}$,

$\mathbf{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$

32. [**M**] Sea T la transformación cuya matriz estándar se presenta a continuación. Determine una base para $\mathbb{R}^4$ con la propiedad de que $[T]_{\mathcal{B}}$ sea diagonal.

$$A = \begin{bmatrix} -6 & 4 & 0 & 9 \\ -3 & 0 & 1 & 6 \\ -1 & -2 & 1 & 0 \\ -4 & 4 & 0 & 7 \end{bmatrix}$$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sea $\mathbf{p}(t) = a_0 + a_1t + a_2t^2$, y calcule

$$[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}[\,\mathbf{p}\,]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

Así, $T(\mathbf{p}) = (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

2. *a*) $A = (I)^{-1}AI$, entonces A es similar a A.

b) Por hipótesis, existen matrices invertibles P y Q con la propiedad de que $B = P^{-1}AP$ y $C = Q^{-1}BQ$. Sustituya la fórmula para B en la expresión para C, y utilice un resultado acerca de la inversa de un producto:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$$

Esta ecuación tiene la forma adecuada para demostrar que A es similar a C.

5.5 VALORES PROPIOS COMPLEJOS

Puesto que la ecuación característica de una matriz de $n \times n$ implica un polinomio de grado n, la ecuación siempre tiene exactamente n raíces, contando las multiplicidades, *siempre y cuando se admitan raíces complejas.* Esta sección demuestra que si la ecuación característica de una matriz real A tiene algunas raíces complejas, entonces esas raíces aportan información importante acerca de A. La clave es dejar que A actúe sobre el espacio $\mathbb{C}^n$ de n-adas de números complejos.[1]

Nuestro interés en $\mathbb{C}^n$ no se origina por el deseo de "generalizar" los resultados de los capítulos anteriores, no obstante que ello abriría nuevas aplicaciones significativas en álgebra lineal.[2] Más bien, este estudio de valores propios complejos es esencial para revelar información "oculta" sobre ciertas matrices con entradas reales que se presentan en gran variedad de problemas de la vida cotidiana. Tales problemas incluyen muchos sistemas dinámicos reales que implican movimiento periódico, vibración o algún tipo de rotación en el espacio.

La teoría del valor propio-vector propio matricial ya desarrollada para $\mathbb{R}^n$ se aplica bien a $\mathbb{C}^n$. Así que un escalar complejo λ satisface $\det(A - \lambda I) = 0$ si y solo si existe un vector distinto de cero $\mathbf{x}$ en $\mathbb{C}^n$ tal que $A\mathbf{x} = \lambda\mathbf{x}$. Decimos que λ es un **valor propio (complejo)** y $\mathbf{x}$ su correspondiente **vector propio (complejo)**.

EJEMPLO 1 Si $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ on $\mathbb{R}^2$ gira el plano en sentido antihorario, un ángulo de 90º. La acción de A es periódica, ya que, después de cuatro de tales aplicaciones, un vector regresa a su posición original. Evidentemente, ningún vector distinto de cero se mapea sobre un múltiplo de sí mismo, de manera que A no tiene vectores propios en $\mathbb{R}^2$ y, por lo tanto, carece de vectores propios reales. En efecto, la ecuación característica de A es

$$\lambda^2 + 1 = 0$$

[1] Consulte el apéndice B para un breve análisis de números complejos. El álgebra matricial y los conceptos sobre espacios vectoriales reales se podrían ampliar al caso de entradas y escalares complejos. En particular, $A(c\mathbf{x} + d\mathbf{y}) = cA\mathbf{x} + dA\mathbf{y}$, para A de $m \times n$, con entradas complejas, $\mathbf{x}$, $\mathbf{y}$ en $\mathbb{C}^n$, y c, d en $\mathbb{C}$.

[2] Con frecuencia, en un segundo curso en álgebra lineal se analizan dichos temas, que son de mucha importancia en ingeniería eléctrica.

Las únicas raíces son complejas: $\lambda = i$ y $\lambda = -i$. Sin embargo, si se permite que A actúe sobre $\mathbb{C}^2$, entonces,

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

Así, i y $-i$ son valores propios, con $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ y $\begin{bmatrix} 1 \\ i \end{bmatrix}$ como sus vectores propios correspondientes. (En el ejemplo 2 se analiza un método para *determinar* vectores propios complejos). ■

El principal interés de esta sección será la matriz del siguiente ejemplo.

EJEMPLO 2 Sea $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$. Determine los valores propios de A, y encuentre una base para cada espacio propio.

SOLUCIÓN La ecuación característica de A es

$$\begin{aligned} 0 = \det\begin{bmatrix} .5-\lambda & -.6 \\ .75 & 1.1-\lambda \end{bmatrix} &= (.5-\lambda)(1.1-\lambda) - (-.6)(.75) \\ &= \lambda^2 - 1.6\lambda + 1 \end{aligned}$$

De la fórmula cuadrática, $\lambda = \frac{1}{2}[1.6 \pm \sqrt{(-1.6)^2 - 4}] = .8 \pm .6i$. Para el valor propio $\lambda = 0.8 \pm .6i$, construya

$$\begin{aligned} A - (.8 - .6i)I &= \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} - \begin{bmatrix} .8-.6i & 0 \\ 0 & .8-.6i \end{bmatrix} \\ &= \begin{bmatrix} -.3+.6i & -.6 \\ .75 & .3+.6i \end{bmatrix} \end{aligned} \tag{1}$$

La reducción por filas de la matriz aumentada usual es muy engorrosa a mano debido a su aritmética compleja. Sin embargo, a continuación se presenta una agradable observación que realmente simplifica el asunto: Puesto que $.8 - .6i$ es un valor propio, el sistema

$$\begin{aligned} (-.3+.6i)x_1 - \quad\quad .6x_2 &= 0 \\ .75x_1 + (.3+.6i)x_2 &= 0 \end{aligned} \tag{2}$$

tiene una solución no trivial (con x_1 y x_2 posiblemente números complejos). Por lo tanto, *ambas ecuaciones en la ecuación (2) determinan la relación entre x_1 y x_2, y cualquier ecuación sirve para expresar una variable en términos de la otra.*[3]

La segunda ecuación en (2) conduce a

$$\begin{aligned} .75x_1 &= (-.3-.6i)x_2 \\ x_1 &= (-.4-.8i)x_2 \end{aligned}$$

Se elige $x_2 = 5$ para así eliminar los decimales, y se obtiene $x_1 = -2 - 4i$. Una base para el espacio propio correspondiente a $\lambda = .8 - .6i$ es

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$$

[3] Otra forma de ver esto es considerando que la matriz en la ecuación (1) no es invertible, de manera que sus filas son linealmente dependientes (como vectores en $\mathbb{C}^2$) y, por consiguiente, una fila es múltiplo (complejo) de la otra.

Cálculos semejantes para $\lambda = .8 + .6i$ producen el vector propio

$$\mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix}$$

Como una comprobación, calcule

$$A\mathbf{v}_2 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4+2i \\ 4+3i \end{bmatrix} = (.8+.6i)\mathbf{v}_2 \qquad \blacksquare$$

Sorprendentemente, la matriz A del ejemplo 2 determina una transformación $\mathbf{x} \mapsto A\mathbf{x}$ que en esencia es una rotación. Este hecho se hace evidente cuando se grafican los puntos adecuados.

EJEMPLO 3 Una manera de ver cómo la multiplicación por la matriz A del ejemplo 2 afecta los puntos consiste en graficar un punto inicial arbitrario, por ejemplo, $\mathbf{x}_0 = (2, 0)$, y después graficar imágenes sucesivas de este punto con multiplicaciones repetidas por A. Es decir, grafique

$$\begin{aligned} \mathbf{x}_1 = A\mathbf{x}_0 &= \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} \\ \mathbf{x}_2 = A\mathbf{x}_1 &= \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -.4 \\ 2.4 \end{bmatrix} \\ \mathbf{x}_3 = A\mathbf{x}_2, &\ldots \end{aligned}$$

La figura 1 muestra $\mathbf{x}_0, \ldots, \mathbf{x}_8$ como los puntos más grandes. Los puntos más pequeños son las ubicaciones de $\mathbf{x}_9, \ldots, \mathbf{x}_{100}$. La secuencia se encuentra sobre una órbita elíptica. $\blacksquare$

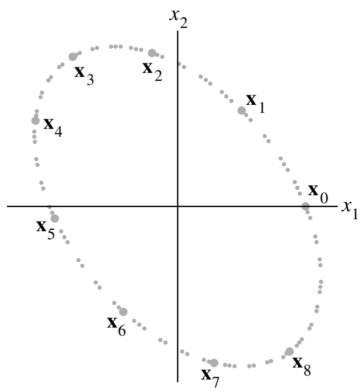


FIGURA 1 Iteración de un punto $\mathbf{x}_0$ por la acción de una matriz con un valor propio complejo.

Desde luego, la figura 1 no explica *por qué* ocurre la rotación. El secreto de la rotación está oculto en las partes real e imaginaria de un vector propio complejo.

Partes real e imaginaria de vectores

El complejo conjugado de un vector complejo $\mathbf{x}$ en $\mathbb{C}^n$ es el vector $\bar{\mathbf{x}}$ en $\mathbb{C}^n$ cuyas entradas están los complejos conjugados de las entradas en $\mathbf{x}$. Las **partes real** e **imaginaria** de un vector complejo $\mathbf{x}$ son los vectores Re $\mathbf{x}$ e Im $\mathbf{x}$ en $\mathbb{R}^n$ formados con las partes real e imaginaria de las entradas de $\mathbf{x}$.

EJEMPLO 4 Si $\mathbf{x} = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i\begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$; entonces

$$\operatorname{Re}\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \quad \operatorname{Im}\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} \quad \text{y} \quad \overline{\mathbf{x}} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i\begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix} \qquad \blacksquare$$

Si B es una matriz de $m \times n$ con posibles entradas complejas, entonces $\overline{B}$ denota la matriz cuyas entradas son los complejos conjugados de las entradas en B. Las propiedades de conjugados de números complejos son válidas también en álgebra matricial compleja:

$$\overline{r\mathbf{x}} = \overline{r}\,\overline{\mathbf{x}}, \quad \overline{B\mathbf{x}} = \overline{B}\,\overline{\mathbf{x}}, \quad \overline{BC} = \overline{B}\,\overline{C} \quad \text{y} \quad \overline{rB} = \overline{r}\,\overline{B}$$

Valores propios y vectores propios de una matriz real que actúa sobre $\mathbb{C}^n$

Sea A una matriz de $n \times n$ cuyas entradas son reales. Entonces, $\overline{A\mathbf{x}} = \overline{A}\overline{\mathbf{x}} = A\overline{\mathbf{x}}$. Si λ es un valor propio de A y $\mathbf{x}$ su vector propio correspondiente en $\mathbb{C}^n$, entonces,

$$A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$$

Por lo tanto, $\bar{\lambda}$ también es un valor propio de A, con $\bar{\mathbf{x}}$ su vector propio correspondiente. Esto demuestra que *cuando A es real, sus valores propios complejos se presentan en pares conjugados*. (Aquí y en todas partes, se utiliza el término *valor propio complejo* para referirse a un valor propio $\lambda = a + bi$, con $b \neq 0$).

EJEMPLO 5 Los valores propios de la matriz real del ejemplo 2 son complejos conjugados; a saber, $.8 - .6i$ y $.8 + .6i$. Los vectores propios correspondientes que se encontraron en el ejemplo 2 también son conjugados:

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \overline{\mathbf{v}}_1 \qquad \blacksquare$$

El siguiente ejemplo proporciona el “bloque básico” para todas las matrices reales de 2×2 con valores propios complejos.

EJEMPLO 6 Si $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, donde a y b son reales, pero no ambos cero, entonces los valores propios de C son $\lambda = a \pm bi$. (Véase el problema de práctica al final de esta sección). También, si $r = |\lambda| = \sqrt{a^2 + b^2}$, entonces,

$$C = r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$$

donde φ es el ángulo entre el eje x positivo y el rayo que pasa por el origen y por (a, b). Véase la figura 2 y el apéndice B. El ángulo φ se llama el *argumento* de $\lambda = a + bi$. Por consiguiente, la transformación $\mathbf{x} \mapsto C\mathbf{x}$ se puede ver como la composición de una rotación por un ángulo φ y un escalamiento por $|\lambda|$ (véase la figura 3). $\blacksquare$

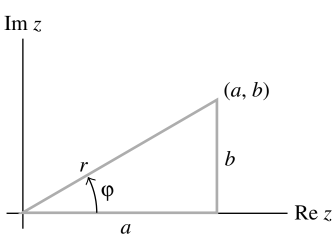


FIGURA 2

Finalmente, se está listo para descubrir la rotación que está oculta dentro de una matriz real que tiene un valor propio complejo.

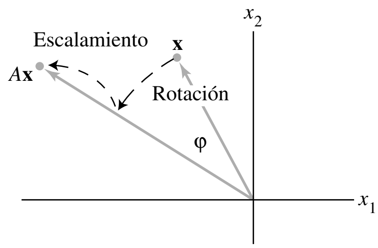


FIGURA 3 Una rotación seguida por un escalamiento.

EJEMPLO 7 Sea $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$, $\lambda = .8 - .6i$ y $\mathbf{v}_1 = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$, como en el ejemplo 2. También, sea P la matriz real de 2×2

$$P = \begin{bmatrix} \operatorname{Re} \mathbf{v}_1 & \operatorname{Im} \mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$$

y sea

$$C = P^{-1}AP = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} .8 & -.6 \\ .6 & 8 \end{bmatrix}$$

Por el ejemplo 6, C es una rotación pura porque $|\lambda|^2 = (.8)^2 + (.6)^2 = 1$. De $C = P^{-1}AP$, se obtiene

$$A = PCP^{-1} = P\begin{bmatrix} .8 & -.6 \\ .6 & 8 \end{bmatrix}P^{-1}$$

¡Aquí está la rotación "dentro" de A! La matriz P proporciona un cambio de variable, por ejemplo, $\mathbf{x} = P\mathbf{u}$. La acción de A equivale a un cambio de variable de $\mathbf{x}$ a $\mathbf{u}$, seguido de una rotación y, después, de regreso a la variable original. Véase la figura 4. La rotación produce una elipse, como en la figura 1, en vez de un círculo, ya que el sistema de coordenadas determinado por las columnas de P no es rectangular ni tiene longitudes unitarias idénticas sobre los dos ejes. ■

$$\begin{tikzcd}
\mathbf{x} \arrow[rr, "A"] \arrow[d, "P^{-1}", "\text{Cambio de variable}"'] & & A\mathbf{x} \\
\mathbf{u} \arrow[rr, "C", "\text{Rotación}"'] & & C\mathbf{u} \arrow[u, "P"', "\text{Cambio de variable}"]
\end{tikzcd}$$

FIGURA 4 Rotación debida a un valor propio complejo.

El siguiente teorema demuestra que los cálculos del ejemplo 7 se adaptan a cualquier matriz real de 2×2 que tenga un valor propio complejo λ. La demostración utiliza el hecho de que si las entradas de A son reales, entonces $A(\operatorname{Re} \mathbf{x}) = \operatorname{Re} A\mathbf{x}$ y $A(\operatorname{Im} \mathbf{x}) = \operatorname{Im} A\mathbf{x}$, y si $\mathbf{x}$ es un vector propio para un valor propio complejo, entonces $\operatorname{Re} \mathbf{x}$ e $\operatorname{Im} \mathbf{x}$ son linealmente independientes en $\mathbb{R}^2$. (Véase los ejercicios 25 y 26). Se omiten los detalles.

TEOREMA 9

Sea A una matriz real de 2×2 con un valor propio complejo $\lambda = a - bi$ ($b \neq 0$) y un vector propio asociado $\mathbf{v}$ en $\mathbb{C}^2$. Entonces,

$$A = PCP^{-1} \quad \text{donde} \quad P = [\operatorname{Re} \mathbf{v} \quad \operatorname{Im} \mathbf{v}] \quad \text{y} \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

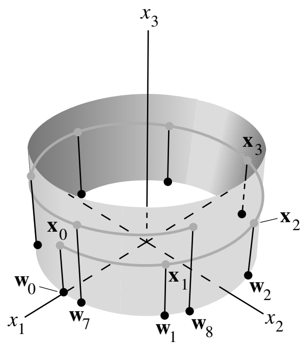


FIGURA 5
Iteraciones de dos puntos bajo la acción de una matriz de 3×3 con un valor propio complejo.

El fenómeno que se ilustra en el ejemplo 7 persiste en dimensiones más altas. Por ejemplo, si A es una matriz de 3×3 con un valor propio complejo, entonces existe un plano en $\mathbb{R}^3$ sobre el cual A actúa como una rotación (posiblemente combinada con escalamiento). Cada vector en ese plano se gira en otro punto sobre el mismo plano. Se dice que el plano es **invariante** bajo A.

EJEMPLO 8 La matriz $A = \begin{bmatrix} .8 & -.6 & 0 \\ .6 & .8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$ tiene valores propios $.8 \pm .6i$ y 1.07. Cualquier vector $\mathbf{w}_0$ en el plano x_1x_2 (con la tercera coordenada cero) es girado por A en otro punto del plano. Cualquier vector $\mathbf{x}_0$ que no esté en el plano tiene su coordenada x_3 multiplicada por 1.07. En la figura 5 se presentan las iteraciones de los puntos $\mathbf{w}_0 = (2, 0, 0)$ y $\mathbf{x}_0 = (2, 0, 1)$ con multiplicación por A. ■

PROBLEMA DE PRÁCTICA

Demuestre que si a y b son reales, entonces los valores propios de $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ son $a \pm bi$, con los vectores propios correspondientes $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ y $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

5.5 EJERCICIOS

En los ejercicios 1 a 6, cada matriz actúa sobre $\mathbb{C}^2$. Determine los valores propios y una base para cada espacio propio en $\mathbb{C}^2$.

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

3. $\begin{bmatrix} 5 & 1 \\ -8 & 1 \end{bmatrix}$ **4.** $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$

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

En los ejercicios 7 a 12, utilice el ejemplo 6 para listar los valores propios de A. En cada caso, la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es la composición de una rotación y de un escalamiento. Determine el ángulo φ de la rotación, donde $-\pi < \varphi \le \pi$, así como el factor de escala de r.

7. $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$ **8.** $\begin{bmatrix} 3 & 3\sqrt{3} \\ -3\sqrt{3} & 3 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$ **10.** $\begin{bmatrix} 0 & .5 \\ -.5 & 0 \end{bmatrix}$

11. $\begin{bmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{bmatrix}$ **12.** $\begin{bmatrix} 3 & -\sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$

En los ejercicios 13 a 20, encuentre una matriz P invertible y una matriz C de la forma $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ tal que la matriz dada tenga la forma $A = PCP^{-1}$.

13. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 5 \\ -2 & 2 \end{bmatrix}$ **16.** $\begin{bmatrix} 4 & -2 \\ 1 & 6 \end{bmatrix}$

17. $\begin{bmatrix} -11 & -4 \\ 20 & 5 \end{bmatrix}$ **18.** $\begin{bmatrix} 3 & -5 \\ 2 & 5 \end{bmatrix}$

19. $\begin{bmatrix} 1.52 & -.7 \\ .56 & .4 \end{bmatrix}$ **20.** $\begin{bmatrix} -3 & -8 \\ 4 & 5 \end{bmatrix}$

21. En el ejemplo 2, despeje x_2 en la primera ecuación en (2) en términos de x_1, para de ahí producir el vector propio $\mathbf{y} = \begin{bmatrix} 2 \\ -1 + 2i \end{bmatrix}$ para la matriz A. Demuestre que este $\mathbf{y}$ es un múltiplo (complejo) del vector $\mathbf{v}_1$ utilizado en el ejemplo 2.

22. Sean A una matriz de $n \times n$ compleja (o real), y $\mathbf{x}$ en $\mathbb{C}^n$ un vector propio correspondiente a un valor propio λ en $\mathbb{C}$. Demuestre que para cada escalar complejo μ distinto de cero, el vector $\mu\mathbf{x}$ es un vector propio de A.

El capítulo 7 se enfocará en matrices con la propiedad $A^T = A$. Los ejercicios 23 y 24 demuestran que cada valor propio de dichas matrices es necesariamente real.

23. Sean A una matriz real de $n \times n$ tal que $A^T = A$, $\mathbf{x}$ cualquier vector en $\mathbb{C}^n$ y $q = \bar{\mathbf{x}}^T A\mathbf{x}$. Las igualdades que se presentan a continuación demuestran que q es un número real comprobando que $\bar{q} = q$. Dé una razón para cada paso.

$$\bar{q} = \overline{\bar{\mathbf{x}}^T A\mathbf{x}} \underset{a)}{=} \mathbf{x}^T \overline{A\mathbf{x}} \underset{b)}{=} \mathbf{x}^T A\bar{\mathbf{x}} \underset{c)}{=} (\mathbf{x}^T A\bar{\mathbf{x}})^T \underset{d)}{=} \bar{\mathbf{x}}^T A^T \mathbf{x} \underset{e)}{=} q$$

24. Sea A una matriz real de $n \times n$ con la propiedad $A^T = A$. Demuestre que si $A\mathbf{x} = \lambda\mathbf{x}$ para algún vector $\mathbf{x}$ distinto de cero en $\mathbb{C}^n$, entonces, en efecto, λ es real y la parte real de $\mathbf{x}$ es un vector propio de A. [*Sugerencia:* Calcule $\bar{\mathbf{x}}^T A\mathbf{x}$, y utilice el ejercicio 23. Asimismo, examine las partes real e imaginaria de $A\mathbf{x}$].

25. Sean A una matriz real de $n \times n$, y $\mathbf{x}$ un vector en $\mathbb{C}^n$. Demuestre que $\operatorname{Re}(A\mathbf{x}) = A(\operatorname{Re}\mathbf{x})$ y $\operatorname{Im}(A\mathbf{x}) = A(\operatorname{Im}\mathbf{x})$.

26. Sea A una matriz real de 2×2 con un valor propio complejo $\lambda = a - bi$ $(b \neq 0)$ y un vector propio asociado $\mathbf{v}$ en $\mathbb{C}^2$.

a) Demuestre que $A(\operatorname{Re}\mathbf{v}) = a\operatorname{Re}\mathbf{v} + b\operatorname{Im}\mathbf{v}$ y $A(\operatorname{Im}\mathbf{v}) = -b\operatorname{Re}\mathbf{v} + a\operatorname{Im}\mathbf{v}$. [*Sugerencia:* Escriba $\mathbf{v} = \operatorname{Re}\mathbf{v} + i\operatorname{Im}\mathbf{v}$, y calcule $A\mathbf{v}$].

b) Compruebe que si P y C están dadas como en el teorema 9, entonces $AP = PC$.

[**M**] En los ejercicios 27 y 28, encuentre una factorización de la matriz A dada en la forma $A = PCP^{-1}$, donde C sea una matriz diagonal a bloques de 2×2 como se indica en el ejemplo 6. (Para cada par conjugado de valores propios, utilice las partes real e imaginaria de un vector propio en $\mathbb{C}^4$ para crear dos columnas de P).

27. $A = \begin{bmatrix} 26 & 33 & 23 & 20 \\ -6 & -8 & -1 & -13 \\ -14 & -19 & -16 & 3 \\ -20 & -20 & -20 & -14 \end{bmatrix}$

28. $A = \begin{bmatrix} 7 & 11 & 20 & 17 \\ -20 & -40 & -86 & -74 \\ 0 & -5 & -10 & -10 \\ 10 & 28 & 60 & 53 \end{bmatrix}$

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Recuerde que es fácil comprobar si un vector es un vector propio. No hay necesidad de examinar la ecuación característica. Calcule

$$A\mathbf{x} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} a+bi \\ b-ai \end{bmatrix} = (a+bi)\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Así $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ es un vector propio correspondiente a $\lambda = a + bi$. Del análisis en esta sección, resulta que $\begin{bmatrix} 1 \\ i \end{bmatrix}$ debe ser un vector propio asociado con $\overline{\lambda} = a - bi$.

5.6 SISTEMAS DINÁMICOS DISCRETOS

Los valores propios y vectores propios ofrecen la clave para entender el comportamiento a largo plazo, o *evolución*, de un sistema dinámico descrito por una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Dicha ecuación se utilizó en la sección 1.10 para modelar el movimiento de población; en la sección 4.9, varias cadenas de Markov; y en el ejemplo introductorio de este capítulo, la población de búhos manchados. Los vectores $\mathbf{x}_k$ dan información acerca del sistema conforme transcurre el tiempo (que se denota con k). En el caso de los búhos manchados, por ejemplo, $\mathbf{x}_k$ listó el número de búhos en tres clases de edades al tiempo k.

En esta sección las aplicaciones se enfocan sobre problemas ecológicos, ya que son más fáciles de plantear y de explicar que, por ejemplo, problemas de física o de ingeniería. Sin embargo, los sistemas dinámicos se originan en muchos campos científicos. En cursos estándar de licenciatura de sistemas de control se estudian diversos aspectos de sistemas dinámicos. En tales cursos el método moderno de diseño del *espacio de estados* se basa significativamente en álgebra de matrices.[1] La respuesta de *estado estable* de un sistema de control es el equivalente en ingeniería de lo que aquí se denomina “comportamiento a largo plazo” del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

[1] Véase G. F. Franklin, J. D. Powell y A. Emami-Naeimi, *Feedback Control of Dynamic Systems*, 5a. ed. (Upper Saddle River, NJ: Prentice-Hall, 2006). Este texto tiene una excelente introducción a modelos dinámicos (cap.2). Los capítulos 7 y 8 tratan el diseño del espacio de estados.

Hasta el ejemplo 6, se supone que A es diagonalizable con n vectores propios linealmente independientes, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y sus valores propios correspondientes $\lambda_1, \ldots, \lambda_n$. Por conveniencia, se supone que los vectores propios están arreglados de tal manera que $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. Puesto que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base para $\mathbb{R}^n$, entonces cualquier vector inicial $\mathbf{x}_0$ se puede escribir unívocamente en la forma

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \tag{1}$$

Esta *descomposición de vectores propios* de $\mathbf{x}_0$ determina lo que sucede a la secuencia $\{\mathbf{x}_k\}$. El siguiente cálculo generaliza el caso simple examinado en el ejemplo 5 de la sección 5.2. Como los $\mathbf{v}_i$ son vectores propios

$$\begin{aligned}\mathbf{x}_1 = A\mathbf{x}_0 &= c_1A\mathbf{v}_1 + \cdots + c_nA\mathbf{v}_n \\ &= c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n\end{aligned}$$

En general,

$$\mathbf{x}_k = c_1(\lambda_1)^k\mathbf{v}_1 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n \qquad (k = 0, 1, 2, \ldots) \tag{2}$$

Los siguientes ejemplos ilustran qué sucede en (2) conforme $k \to \infty$.

Un sistema depredador-presa

En lo profundo de los bosques de secuoyas en California, las ratas de bosque de patas oscuras suministran el 80% de la dieta de los búhos manchados, el principal depredador de esos roedores nocturnos. El ejemplo 1 utiliza un sistema dinámico lineal para modelar el sistema físico de los búhos y las ratas. (Hay que admitir que aunque el modelo no es realista en varios aspectos, puede dar el punto de partida para el estudio de modelos no lineales más complicados utilizados por científicos ambientales).

EJEMPLO 1 Denote las poblaciones de búhos y ratas de bosque al tiempo k por el vector $\mathbf{x}_k = \begin{bmatrix} O_k \\ R_k \end{bmatrix}$, donde k es el tiempo en meses, O_k es el número de búhos en la región bajo estudio y R_k es el número de ratas (medido en miles). Suponga que

$$\begin{aligned} O_{k+1} &= (.5)O_k + (.4)R_k \\ R_{k+1} &= -p \cdot O_k + (1.1)R_k \end{aligned} \tag{3}$$

donde p es un parámetro positivo que se debe especificar. $(.5)O_k$ en la primera ecuación indica que si no hay ratas para alimentarse, entonces cada mes sobreviviría tan solo la mitad de los búhos; mientras que $(1.1)R_k$ en la segunda ecuación implica que sin los búhos como depredadores, la población de ratas crecería 10% cada mes. Si las ratas son abundantes, $(.4)R_k$ hará que crezca la población de búhos, en tanto que el término negativo $-p \cdot O_k$ mide las muertes de ratas por la depredación de los búhos. (En efecto, $1000p$ es el número promedio de ratas devoradas por un búho en un mes). Determine la evolución de este sistema, cuando el parámetro de depredación p es .104.

SOLUCIÓN Cuando $p = .104$, los valores propios de la matriz de coeficientes A para las ecuaciones (3) resultan ser $\lambda_1 = 1.02$ y $\lambda_2 = .58$. Los vectores propios correspondientes son

$$\mathbf{v}_1 = \begin{bmatrix} 10 \\ 13 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

Un vector inicial $\mathbf{x}_0$ se escribe como $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Entonces, para $k \geq 0$,

$$\begin{aligned}\mathbf{x}_k &= c_1(1.02)^k\mathbf{v}_1 + c_2(.58)^k\mathbf{v}_2 \\ &= c_1(1.02)^k\begin{bmatrix} 10 \\ 13 \end{bmatrix} + c_2(.58)^k\begin{bmatrix} 5 \\ 1 \end{bmatrix}\end{aligned}$$

Conforme $k \to \infty$, entonces $(.58)^k$ rápidamente se aproxima a cero. Suponga que $c_1 > 0$. Por lo tanto, para toda k suficientemente grande, $\mathbf{x}_k$ es aproximadamente igual a $c_1(1.02)^k\mathbf{v}_1$, y se escribe

$$\mathbf{x}_k \approx c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \tag{4}$$

La aproximación en (4) mejora conforme k se incrementa, y así para k grandes,

$$\mathbf{x}_{k+1} \approx c_1(1.02)^{k+1} \begin{bmatrix} 10 \\ 13 \end{bmatrix} = (1.02)c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \approx 1.02\mathbf{x}_k \tag{5}$$

La aproximación en la ecuación (5) indica que a final de cuentas ambas entradas de $\mathbf{x}_k$ (los números de búhos y ratas) crecerán por un factor de casi 1.02 cada mes, un 2% de tasa de crecimiento mensual. Por (4), $\mathbf{x}_k$ es aproximadamente un múltiplo de (10, 13), de manera que las entradas en $\mathbf{x}_k$ están casi en la misma razón de 10 a 13. Es decir, por cada 10 búhos hay cerca de 13 mil ratas. ∎

El ejemplo 1 muestra dos hechos generales acerca de un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ donde A es $n \times n$, sus valores propios satisfacen $|\lambda_1| \geq 1$ y $1 > |\lambda_1|$ para $j = 2, \ldots, n$ y $\mathbf{v}_1$ es un vector propio correspondiente a λ_1. Si $\mathbf{x}_0$ está dado por la ecuación (1), con $c_1 \neq 0$, entonces para todas las k suficientemente grandes,

$$\mathbf{x}_{k+1} \approx \lambda_1 \mathbf{x}_k \tag{6}$$

y

$$\mathbf{x}_k \approx c_1(\lambda_1)^k \mathbf{v}_1 \tag{7}$$

Las aproximaciones en (6) y (7) se pueden volver tan cercanas como se desee tomando una k muy grande. Por (6), $\mathbf{x}_k$ finalmente crecerá casi por un factor de λ_1 cada vez, de manera que λ_1 determina la eventual tasa de crecimiento del sistema. Asimismo, por (7), la razón de cualesquiera dos entradas en $\mathbf{x}_k$ (para k grande) es casi la misma que la razón de las entradas correspondientes en $\mathbf{v}_1$. En el ejemplo 5 de la sección 5.2, se ilustra el caso donde $\lambda_1 = 1$.

Descripción gráfica de soluciones

Cuando A es una matriz de 2×2, los cálculos algebraicos se complementan con una descripción geométrica de la evolución del sistema. La ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ se puede considerar una descripción de qué sucede al punto inicial $\mathbf{x}_0$ en $\mathbb{R}^2$, conforme se transforma repetidamente por el mapeo $\mathbf{x} \mapsto A\mathbf{x}$. La gráfica de $\mathbf{x}_0, \mathbf{x}_1, \ldots$ es la **trayectoria** del sistema dinámico.

EJEMPLO 2 Grafique varias trayectorias del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, cuando

$$A = \begin{bmatrix} .80 & 0 \\ 0 & .64 \end{bmatrix}$$

SOLUCIÓN Los valores propios de A son .8 y .64, con vectores propios $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Si $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, entonces,

$$\mathbf{x}_k = c_1(.8)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Desde luego, $\mathbf{x}_k$ tiende a $\mathbf{0}$ porque $(.8)^k$ y $(.64)^k$ se aproximan a 0 conforme $k \to \infty$. No obstante, resulta interesante *la manera* en que $\mathbf{x}_k$ va hacia $\mathbf{0}$. La figura 1 (en la página 304) muestra los primeros términos escasos de varias trayectorias que empiezan en puntos sobre la frontera de la caja con vértices en $(\pm 3, \pm 3)$. Los puntos sobre cada trayectoria están conectados por una curva delgada, para que así la trayectoria sea más fácil de observar. ∎

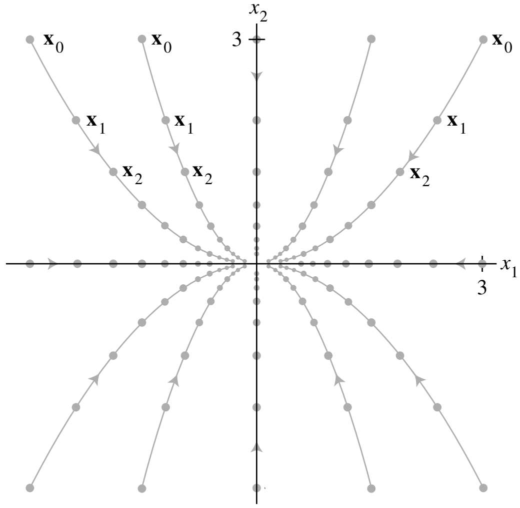


FIGURA 1 El origen como un atractor.

En el ejemplo 2, al origen se le denomina **atractor** del sistema dinámico porque todas las trayectorias tienden hacia **0**. Esto ocurre siempre que ambos valores propios sean menores que 1 en magnitud. La dirección de la atracción más grande está a lo largo de la recta que pasa por **0** y el vector propio $\mathbf{v}_2$ para el valor propio de menor magnitud.

En el siguiente ejemplo, ambos valores propios de A son mayores que 1 en magnitud, y **0** es un **repulsor** del sistema dinámico. Todas las soluciones de $\mathbf{x}_{k+1} = A\mathbf{x}_k$ excepto la solución cero (constante), no están acotadas y tienden a alejarse del origen.[2]

EJEMPLO 3 Grafique varias soluciones típicas de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

SOLUCIÓN Los valores propios de A son 1.44 y 1.2. Si $\mathbf{x}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, entonces,

$$\mathbf{x}_k = c_1(1.44)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(1.2)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Ambos términos aumentaron en tamaño, pero el primer término lo hizo más rápido. Así, la dirección de la mayor repulsión es la recta que pasa por **0** y por el vector propio para el valor propio de magnitud más grande. La figura 2 muestra varias trayectorias que inician en puntos muy cercanos a **0**. ■

En el siguiente ejemplo, **0** es un **punto silla** porque el origen atrae soluciones de algunas direcciones y las repele en otras direcciones. Esto ocurre siempre que un valor propio sea más grande que 1 en magnitud y el otro sea menor que 1 en magnitud. La dirección de mayor atracción está determinada por un vector propio del valor propio de menor magnitud. La dirección de la mayor repulsión se determina mediante un vector propio del valor propio de magnitud más grande.

[2] El origen es el único posible atractor o repulsor en un sistema dinámico *lineal*, aunque podrían existir múltiples atractores y repulsores en un sistema dinámico más general para el cual el mapeo $\mathbf{x}_k \to \mathbf{x}_{k+1}$ no es lineal. En tal sistema, los atractores y los repulsores están definidos en términos de los valores propios de una matriz especial (con entradas variables) llamada *matriz jacobiana* del sistema.

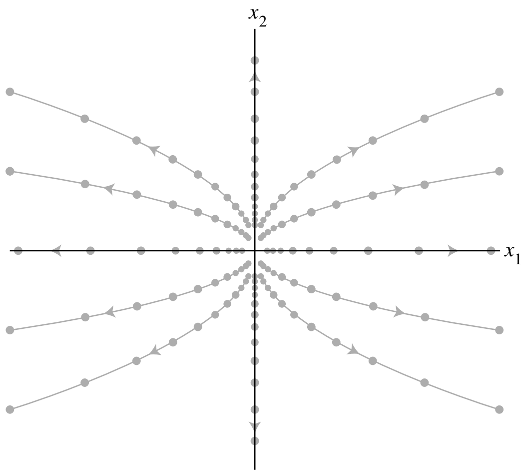


FIGURA 2 El origen como un repulsor.

EJEMPLO 4 Grafique varias soluciones típicas de la ecuación $\mathbf{y}_{k+1} = D\mathbf{y}_k$, donde

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

(Aquí se usaron D y $\mathbf{y}$ en vez de A y $\mathbf{x}$ porque este ejemplo se utilizará más adelante). Demuestre que una solución $\{\mathbf{y}_k\}$ no está acotada, cuando su punto inicial no está sobre el eje x_2.

SOLUCIÓN Los valores propios de D son 2 y 0.5. Si $\mathbf{y}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, entonces,

$$\mathbf{y}_k = c_1 2^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.5)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{8}$$

Si $\mathbf{y}_0$ está sobre el eje x_2, entonces $c_1 = 0$ y $\mathbf{y}_k \to \mathbf{0}$ conforme $k \to \infty$. Pero si $\mathbf{y}_0$ no se encuentra sobre el eje x_2, entonces el primer término en la suma para $\mathbf{y}_k$ se vuelve arbitrariamente grande y, por lo tanto, $\{\mathbf{y}_k\}$ no está acotada. La figura 3 presenta diez trayectorias que inician cerca del eje x_2 o sobre este. ■

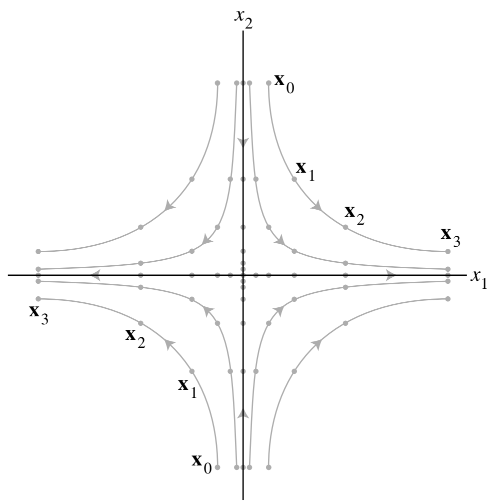


FIGURA 3 El origen como un punto silla.

Cambio de variable

Los tres ejemplos anteriores implicaron matrices diagonales. Para tratar el caso no diagonal, regrese por un momento al caso $n \times n$ donde los vectores propios de A forman una base $\{\mathbf{v}_1,\ldots, \mathbf{v}_n\}$ para $\mathbb{R}^n$. Sean $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, y D la matriz diagonal con los valores propios correspondientes sobre la diagonal. Dada una secuencia $\{\mathbf{x}_k\}$ que comprueba $\mathbf{x}_{k+1} = A\mathbf{x}_k$, defina una nueva secuencia $\{\mathbf{y}_k\}$ mediante

$$\mathbf{y}_k = P^{-1}\mathbf{x}_k \quad \text{o, de manera equivalente,} \quad \mathbf{x}_k = P\mathbf{y}_k$$

Sustituyendo esas relaciones en la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ y utilizando que $A = PDP^{-1}$, se tiene que

$$P\mathbf{y}_{k+1} = AP\mathbf{y}_k = (PDP^{-1})P\mathbf{y}_k = PD\mathbf{y}_k$$

Se multiplican ambos lados por la izquierda por P^{-1} y se obtiene

$$\mathbf{y}_{k+1} = D\mathbf{y}_k$$

Si se escribe $\mathbf{y}_k$ como $\mathbf{y}(k)$ y se denotan las entradas en $\mathbf{y}(k)$ por $y_1(k),\ldots, y_n(k)$, entonces,

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ \vdots \\ y_n(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix}$$

El cambio de variable de $\mathbf{x}_k$ a $\mathbf{y}_k$ ha *desacoplado* el sistema de ecuaciones en diferencias. La evolución de $y_1(k)$, por ejemplo, no se ve afectada por lo que suceda a $y_2(k),\ldots, y_n(k)$, porque $y_1(k + 1) = \lambda_1 \cdot y_1(k)$ para cada k.

La ecuación $\mathbf{x}_k = P\mathbf{y}_k$ indica que $\mathbf{y}_k$ es el vector de coordenadas de $\mathbf{x}_k$ con respecto a la base de vectores propios $\{\mathbf{v}_1,\ldots, \mathbf{v}_n\}$. Se puede desacoplar el sistema $\mathbf{x}_{k+1} = A\mathbf{x}_k$ efectuando cálculos en el nuevo sistema de coordenadas de vectores propios. Cuando $n = 2$, esto equivale a emplear papel para graficar con ejes en las direcciones de los dos vectores propios.

EJEMPLO 5 Demuestre que el origen es un punto silla para las soluciones de $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 1.25 & -.75 \\ -.75 & 1.25 \end{bmatrix}$$

Encuentre las direcciones de mayor atracción y de mayor repulsión.

SOLUCIÓN Utilizando técnicas estándar, se determina que los valores propios de A son 2 y .5, con los vectores propios correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectivamente. Como $|2| > 1$ y $|.5| < 1$, entonces el origen es un punto silla del sistema dinámico. Si $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$,

$$\mathbf{x}_k = c_1 2^k \mathbf{v}_1 + c_2(.5)^k \mathbf{v}_2 \tag{9}$$

Esta ecuación se ve exactamente igual que la ecuación (8) del ejemplo 4, con $\mathbf{v}_1$ y $\mathbf{v}_2$ en lugar de la base estándar.

Sobre el papel para graficar, dibuje ejes a través de $\mathbf{0}$ y de los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$. Véase la figura 4. El movimiento a lo largo de los ejes corresponde al movimiento sobre los ejes estándar de la figura 3. En la figura 4, la dirección de mayor *repulsión* es la recta que pasa por $\mathbf{0}$ y el vector propio $\mathbf{v}_1$ cuyo valor propio es mayor que 1 en magnitud. Si $\mathbf{x}_0$ está sobre esta recta, c_2 de la ecuación (9) es cero y $\mathbf{x}_k$ se aleja rápidamente de $\mathbf{0}$. La dirección de mayor *atracción* está determinada por el vector propio $\mathbf{v}_2$ cuyo valor propio es menor que 1 en magnitud.

En la figura 4 se ilustra un número de trayectorias. Cuando esta gráfica se observa en términos de los ejes de vectores propios, el esquema “parece” en esencia igual que el esquema de la figura 3. ■

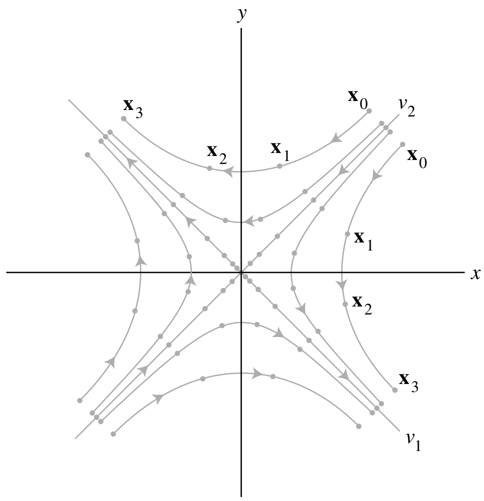


FIGURA 4 El origen como un punto silla.

Valores propios complejos

Cuando una matriz real A de 2×2 tiene valores propios complejos, entonces A no es diagonalizable (cuando actúa sobre $\mathbb{R}^2$), pero el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ es fácil de describir. El ejemplo 3 de la sección 5.5 ilustró el caso donde los valores propios tienen valor absoluto 1. La iteración de un punto $\mathbf{x}_0$ giró alrededor del origen con una trayectoria elíptica.

Si A tiene dos valores propios complejos cuyo valor absoluto es mayor que 1, entonces $\mathbf{0}$ es un repulsor y la iteración de $\mathbf{x}_0$ girará en espiral hacia fuera en torno al origen. Si los valores absolutos de los valores propios complejos son menores que 1, entonces el origen es un atractor y la iteración de $\mathbf{x}_0$ da una espiral hacia adentro acercándose al origen, como en el siguiente ejemplo.

EJEMPLO 6 Se puede comprobar que la matriz

$$A = \begin{bmatrix} .8 & .5 \\ -.1 & 1.0 \end{bmatrix}$$

tiene valores propios $.9 \pm .2i$, con vectores propios $\begin{bmatrix} 1 \mp 2i \\ 1 \end{bmatrix}$. La figura 5 (en la página 308) muestra tres trayectorias del sistema $\mathbf{x}_{k+1} = A\mathbf{x}_k$, con vectores iniciales $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ -2.5 \end{bmatrix}$. ■

Supervivencia de los búhos manchados

Recuerde del ejemplo introductorio de este capítulo que la población de búhos manchados en el área Willow Creek, California, se modeló con un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ donde las entradas en $\mathbf{x}_k = (j_k, s_k, a_k)$ listaban los números de hembras (al tiempo k) en las etapas de vida juvenil, subadulto y adulto, respectivamente, y A es la matriz de etapas

$$A = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \tag{10}$$

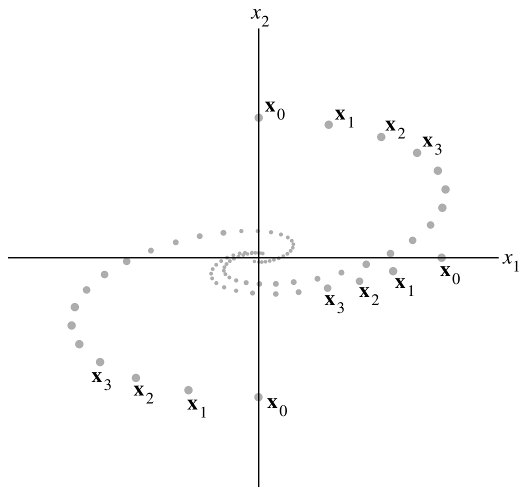


FIGURA 5 Rotación asociada con valores propios complejos.

MATLAB muestra que los valores propios de A son aproximadamente $\lambda_1 = .98$, $\lambda_2 = -.02 + .21i$ y $\lambda_3 = -.02 - .21i$. Observe que los tres valores propios son menores que 1 en magnitud, ya que $|\lambda_2|^2 = |\lambda_3|^2 = (-.02)^2 + (.21)^2 = 0.445$.

Por el momento, deje que A actúe sobre el espacio vectorial complejo $\mathbb{C}^3$. Puesto que A tiene tres valores propios distintos, entonces los tres vectores propios correspondientes son linealmente independientes y forman una base para $\mathbb{C}^3$. Denote los vectores propios con $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Así, la solución general de $\mathbf{x}_{k+1} = A\mathbf{x}_k$ (utilizando vectores en $\mathbb{C}^3$) tiene la forma

$$\mathbf{x}_k = c_1(\lambda_1)^k\mathbf{v}_1 + c_2(\lambda_2)^k\mathbf{v}_2 + c_3(\lambda_3)^k\mathbf{v}_3 \tag{11}$$

Si $\mathbf{x}_0$ es un vector inicial real, entonces $\mathbf{x}_1 = A\mathbf{x}_0$ es real porque A es real. Asimismo, la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ indica que cada $\mathbf{x}_k$ en el miembro izquierdo de la ecuación (11) es real, aun cuando se expresa como una suma de vectores complejos. Sin embargo, cada término en el miembro derecho de la ecuación (11) se aproxima al vector cero, ya que todos los valores propios son menores que 1 en magnitud. Por lo tanto, la secuencia real $\mathbf{x}_k$ también tiende al vector cero. Tristemente, este modelo predice que, a final de cuentas, perecerán todos los búhos.

¿Hay alguna esperanza para los búhos manchados? Recuerde del ejemplo introductorio que la entrada de 18% en la matriz A de la ecuación (10) es porque, no obstante que el 60% de los búhos juveniles sobreviven lo suficiente para dejar el nido y buscar nuevas áreas de distribución, tan solo el 30% de ese grupo sobrevive a dicha búsqueda y encuentra un nuevo hogar. La supervivencia a la mencionada búsqueda está significativamente influida por el número de áreas deforestadas en el bosque, lo cual vuelve más difícil y peligrosa esa búsqueda.

Algunas poblaciones de búhos habitan en zonas con pocas o ninguna área deforestada. Quizás un porcentaje aún mayor de búhos juveniles sobrevivan ahí y encuentren nuevos territorios. De hecho, el problema del búho manchado es más complicado que como se ha expuesto, no obstante que el ejemplo final da un final feliz a la historia.

EJEMPLO 7 Suponga que la tasa de supervivencia a la búsqueda, de los búhos juveniles, sea de 50%, de manera que la entrada (2, 1) de la matriz por etapas (de estados) A en (10) es .3 en vez de .18. ¿Qué predice el modelo de la matriz por etapas acerca de la población de búhos manchados?

SOLUCIÓN Ahora los valores propios de A resultan ser aproximadamente $\lambda_1 = 1.01$, $\lambda_2 = -.03 + .26i$ y $\lambda_3 = -.03 - .26i$. Un vector propio para λ_1 es aproximadamente $\mathbf{v}_1 = (10, 3, 31)$. Sean $\mathbf{v}_2$ y $\mathbf{v}_3$ vectores propios (complejos) para λ_2 y λ_3. En este caso, la ecuación (11) se convierte en

$$\mathbf{x}_k = c_1(1.01)^k\mathbf{v}_1 + c_2(-.03 + .26i)^k\mathbf{v}_2 + c_3(-.03 - .26i)^k\mathbf{v}_3$$

Cuando $k \to \infty$, el segundo de los dos vectores tiende a cero. Así, $\mathbf{x}_k$ se vuelve cada vez más semejante al vector (real) $c_1(1.01)^k\mathbf{v}_1$. Las aproximaciones aplicadas en (6) y (7) provienen del ejemplo 1. Además, se puede demostrar que la constante c_1 en la descomposición inicial de $\mathbf{x}_0$ es positiva cuando las entradas en $\mathbf{x}_0$ son no negativas. Entonces, la población de búhos crecerá lentamente, con una tasa de crecimiento asintótica de 1.01. El vector propio $\mathbf{v}_1$ describe la distribución final de los búhos en sus etapas de vida: por cada 31 adultos, habrá 10 juveniles y 3 subadultos. ■

Lecturas adicionales

Franklin G. F., J. D. Powell y M. L. Workman. *Digital Control of Dynamic Systems*, 3a. ed., Reading, MA: Addison-Wesley, 1998.

Sandefur, James T. *Discrete Dynamical Systems-Theory and Applications*. Oxford: Oxford University Press, 1990.

Tuchinsky, Phlip. *Management of a Buffalo Herd*, UMAP módulo 207. Lexington, MA: COMAP, 1980.

PROBLEMAS DE PRÁCTICA

1. La matriz A que se muestra a continuación tiene los valores propios 1, $\frac{2}{3}$ y $\frac{1}{3}$ con los vectores propios correspondientes $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$:

$$A = \frac{1}{9}\begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

Encuentre la solución general de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ si $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 11 \\ -2 \end{bmatrix}$.

2. ¿Qué ocurre en la secuencia $\{\mathbf{x}_k\}$ del problema de práctica 1, conforme $k \to \infty$?

5.6 EJERCICIOS

1. Sea A una matriz de 2×2 con valores propios 3 y $1/3$, y vectores propios correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Sea $\{\mathbf{x}_k\}$ una solución de la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$, $\mathbf{x}_0 = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$.
   a) Calcule $\mathbf{x}_1 = A\mathbf{x}_0$. [*Sugerencia:* No se necesita conocer A].
   b) Determine una fórmula para $\mathbf{x}_k$ que implique k y los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$.

2. Suponga que los valores propios de A de 3×3 son 3, $4/5$ y $3/5$, con sus vectores propios correspondientes $\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}$ y $\begin{bmatrix} -3 \\ -3 \\ 7 \end{bmatrix}$. Sea $\mathbf{x}_0 = \begin{bmatrix} -2 \\ -5 \\ 3 \end{bmatrix}$. Encuentre la solución de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $\mathbf{x}_0$ especificado, y describa qué sucede cuando $k \to \infty$.

En los ejercicios 3 a 6, suponga que cualquier vector inicial $\mathbf{x}_0$ tiene una descomposición de vectores propios tal que el coeficiente c_1, en la ecuación (1) de esta sección, es positivo.[3]

3. Determine la evolución del sistema dinámico del ejemplo 1, cuando el parámetro de depredación p es .2 en la ecuación (3). (Dé una fórmula para $\mathbf{x}_k$). ¿Crece o disminuye la población de búhos? ¿Qué se puede decir sobre la población de ratas de bosque?

4. Determine la evolución del sistema dinámico del ejemplo 1, cuando el parámetro de depredación p es .125. (Dé una fórmula para $\mathbf{x}_k$). Conforme el tiempo transcurre, ¿qué ocurre a los tamaños de las poblaciones de búhos y de ratas bosque? El sistema tiende hacia lo que algunas veces se denomina equilibrio inestable. ¿Qué piensa que pueda ocurrirle al sistema si algún aspecto del modelo (como las tasas de nacimiento o de depredación) cambiara ligeramente?

5. En los antiguos bosques de abetos en Douglas, los búhos manchados cenan básicamente ardillas voladoras. Suponga que la matriz depredador-presa para esas dos poblaciones es $A = \begin{bmatrix} .4 & .3 \\ -p & 1.2 \end{bmatrix}$. Demuestre que si el parámetro de depredación p es .325, entonces ambas poblaciones crecen. Estime la tasa de crecimiento a largo plazo y la proporción final entre búhos y ardillas voladoras.

6. Demuestre que si el parámetro de depredación p del ejercicio 5 es .5, entonces a final de cuentas desaparecerán tanto los búhos como las ardillas. Encuentre un valor de p para el cual ambas poblaciones tienden a mantener niveles constantes. En este caso, ¿cuál es la proporción entre las dos poblaciones?

7. Sea que A tiene las propiedades descritas en el ejercicio 1.

a) ¿El origen es un atractor, un repulsor, o un punto silla del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$?

b) Encuentre las direcciones de mayor atracción y/o repulsión para este sistema dinámico.

c) Haga una descripción gráfica del sistema, que indique las direcciones de mayor atracción o repulsión. Incluya un esquema de varias trayectorias típicas (sin calcular puntos específicos).

8. Determine la naturaleza del origen (atractor, repulsor o punto silla) para el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ si A tiene las propiedades descritas en el ejercicio 2. Obtenga las direcciones de mayor atracción o repulsión.

En los ejercicios 9 a 14, clasifique el origen como atractor, repulsor o punto silla del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Encuentre las direcciones de mayor atracción y/o repulsión.

9. $A = \begin{bmatrix} 1.7 & -.3 \\ -1.2 & .8 \end{bmatrix}$

10. $A = \begin{bmatrix} .3 & .4 \\ -.3 & 1.1 \end{bmatrix}$

11. $A = \begin{bmatrix} .4 & .5 \\ -.4 & 1.3 \end{bmatrix}$

12. $A = \begin{bmatrix} .5 & .6 \\ -.3 & 1.4 \end{bmatrix}$

13. $A = \begin{bmatrix} .8 & .3 \\ -.4 & 1.5 \end{bmatrix}$

14. $A = \begin{bmatrix} 1.7 & .6 \\ -.4 & .7 \end{bmatrix}$

15. Sea $A = \begin{bmatrix} .4 & 0 & .2 \\ .3 & .8 & .3 \\ .3 & .2 & .5 \end{bmatrix}$. El vector $\mathbf{v}_1 = \begin{bmatrix} .1 \\ .6 \\ .3 \end{bmatrix}$ es un vector propio de A, y dos valores propios son .5 y .2. Construya la solución del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ que satisface $\mathbf{x}_0 = (0, .3, .7)$. ¿Qué sucede a $\mathbf{x}_k$ conforme $k \to \infty$?

16. [**M**] Obtenga la solución general del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ cuando A es la matriz estocástica para el modelo Rent A Car de Hertz del ejercicio 16 de la sección 4.9.

17. Construya un modelo matricial por etapas para una especie animal que tiene dos etapas de vida: juvenil (hasta 1 año de edad) y adulto. Suponga que las hembras adultas cada año dan a luz a un promedio de 1.6 hembras juveniles. Cada año, el 30% de las hembras juveniles sobreviven para convertirse en adultos y el 80% de los adultos sobreviven. Para $k \geq 0$, sea $\mathbf{x}_k = (j_k, a_k)$, donde las entradas en $\mathbf{x}_k$ son los números de hembras juveniles y hembras adultas en el año k.

a) Construya la matriz por etapas A tal que $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $k \geq 0$.

b) Demuestre que la población está creciendo, calcule la tasa de crecimiento final de la población y dé la razón final entre adultos y juveniles.

c) [**M**] Suponga que inicialmente hay 15 hembras juveniles y 10 hembras adultas en la población. Elabore cuatro gráficas que muestren cómo cambia la población durante 8 años: *a*) el número de juveniles, *b*) el número de adultas, *c*) la población total y *d*) la razón entre adultas y juveniles (cada año). ¿Cuándo parece estabilizarse la razón en *d*)? Incluya una lista del programa o las instrucciones empleadas para realizar las gráficas en *c*) y *d*).

18. Una manada de búfalos americanos (bisontes) se podría modelar con una matriz por etapas, similar a la empleada para los búhos manchados. Las hembras se pueden dividir en terneras (hasta 1 año de edad), becerras (de 1 a 2 años) y adultas. Suponga que un promedio de 42 hembras terneras nacen cada año por cada 100 adultas. (Tan solo las adultas tienen descendencia). Cada año sobreviven cerca del 60% de terneras, 75% de becerras y 95% de adultas. Para $k \geq 0$, sea $\mathbf{x}_k = (c_k, y_k, a_k)$, donde las entradas en $\mathbf{x}_k$ son los números de hembras en cada etapa de vida en el año k.

a) Construya la matriz por etapas A para la manada de búfalos, tal que $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $k \geq 0$.

b) [**M**] Demuestre que la manada de búfalos está creciendo, determine la tasa de crecimiento esperada después de muchos años, y dé los números esperados de terneras y becerras por cada 100 adultas.

[3] Una de las limitaciones del modelo del ejemplo 1 es que ahí siempre existen vectores $\mathbf{x}_0$ de población inicial con entradas positivas, tales que el coeficiente c_1 es negativo. La aproximación (7) aún es válida, pero las entradas en $\mathbf{x}_k$ finalmente se vuelven negativas.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. El primer paso consiste en escribir $\mathbf{x}_0$ como una combinación lineal de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. La reducción por filas de $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{x}_0]$ genera los pesos $c_1 = 2$, $c_2 = 1$ y $c_3 = 3$, tal que

$$\mathbf{x}_0 = 2\mathbf{v}_1 + 1\mathbf{v}_2 + 3\mathbf{v}_3$$

Como los valores propios son 1, $\frac{2}{3}$ y $\frac{1}{3}$, la solución general es

$$\begin{aligned}\mathbf{x}_k &= 2 \cdot 1^k \mathbf{v}_1 + 1 \cdot \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3 \cdot \left(\frac{1}{3}\right)^k \mathbf{v}_3 \\ &= 2\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} + \left(\frac{2}{3}\right)^k \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + 3 \cdot \left(\frac{1}{3}\right)^k \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \end{aligned} \tag{12}$$

2. Conforme $k \to \infty$, el segundo y tercer términos en (12) tienden al vector cero, y

$$\mathbf{x}_k = 2\mathbf{v}_1 + \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3\left(\frac{1}{3}\right)^k \mathbf{v}_3 \to 2\mathbf{v}_1 = \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}$$

5.7 APLICACIONES A ECUACIONES DIFERENCIALES

Esta sección describe análogos continuos de las ecuaciones en diferencias estudiadas en la sección 5.6. En muchos problemas aplicados, algunas cantidades varían continuamente en el tiempo, y se relacionan mediante un sistema de ecuaciones diferenciales:

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

Aquí $x_1, \ldots, x_n$ son funciones diferenciables de t, con derivadas $x_1', \ldots, x_n'$, y las a_{ij} son constantes. El aspecto esencial de este sistema es que es *lineal*. Para observarlo, escriba el sistema como una ecuación diferencial matricial

$$\mathbf{x}'(t) = A\mathbf{x}(t) \tag{1}$$

donde

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Una **solución** de la ecuación (1) es una función valuada en vectores que satisface la ecuación (1) para toda t en algún intervalo de números reales, tales como $t \geq 0$.

La ecuación (1) es *lineal* porque la derivación de funciones y la multiplicación de vectores por una matriz son transformaciones lineales. Así, cuando $\mathbf{u}$ y $\mathbf{v}$ son soluciones de $\mathbf{x}' = A\mathbf{x}$, entonces $c\mathbf{u} + d\mathbf{v}$ también es una solución, porque

$$\begin{aligned} (c\mathbf{u} + d\mathbf{v})' &= c\mathbf{u}' + d\mathbf{v}' \\ &= cA\mathbf{u} + dA\mathbf{v} = A(c\mathbf{u} + d\mathbf{v}) \end{aligned}$$

(A esta propiedad los ingenieros la denominan *superposición* de soluciones). Asimismo, la función idénticamente cero es una solución (trivial) de (1). En la terminología del capítulo 4, el conjunto de todas las soluciones de (1) es un *subespacio* del conjunto de todas las funciones continuas con valores en $\mathbb{R}^n$.

Los libros estándar de ecuaciones diferenciales demuestran que siempre existe el **conjunto fundamental de soluciones** de (1). Si A es $n \times n$, entonces hay n funciones linealmente independientes en un conjunto fundamental, y cada solución de (1) es una combinación lineal única de tales n funciones. Es decir, un conjunto fundamental de soluciones es una *base* para el conjunto de todas las soluciones de (1), y el conjunto solución es un espacio vectorial n-dimensional de funciones. Si se especifica un vector $\mathbf{x}_0$, entonces el **problema con valores iniciales** consiste en construir la función (única) $\mathbf{x}$ tal que $\mathbf{x}' = A\mathbf{x}$ y $\mathbf{x}(0) = \mathbf{x}_0$.

Cuando A es una matriz diagonal, las soluciones de (1) se generan mediante cálculo elemental. Por ejemplo, considere

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{2}$$

es decir,

$$\begin{aligned} x_1'(t) &= 3x_1(t) \\ x_2'(t) &= -5x_2(t) \end{aligned} \tag{3}$$

Se dice que el sistema (2) está *desacoplado* porque cada derivada de una función tan solo depende de la propia función, y no de alguna combinación o “acoplamiento” de $x_1(t)$ y $x_2(t)$. De cálculo, las soluciones de (3) son $x_1(t) = c_1e^{3t}$ y $x_2(t) = c_2e^{-5t}$, para las constantes cualesquiera c_1 y c_2. Cada solución de la ecuación (2) se escribe en la forma

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1e^{3t} \\ c_2e^{-5t} \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix}e^{3t} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}e^{-5t}$$

Este ejemplo sugiere que para la ecuación general $\mathbf{x}' = A\mathbf{x}$, una solución podría ser una combinación lineal de funciones con la forma

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t} \tag{4}$$

para algún escalar λ y algún vector fijo $\mathbf{v}$ distinto de cero. [Si $\mathbf{v} = \mathbf{0}$, la función $\mathbf{x}(t)$ es idénticamente cero y por lo tanto satisface $\mathbf{x}' = A\mathbf{x}$]. Observe que

$$\begin{aligned} \mathbf{x}'(t) &= \lambda\mathbf{v}e^{\lambda t} && \text{Por Cálculo, ya que } \mathbf{v} \text{ es un vector constante} \\ A\mathbf{x}(t) &= A\mathbf{v}e^{\lambda t} && \text{Multiplicando ambos lados de (4) por } A \end{aligned}$$

Como $e^{\lambda t}$ nunca es cero, $\mathbf{x}'(t)$ será igual a $A\mathbf{x}(t)$ si y solo si $\lambda\mathbf{v} = A\mathbf{v}$, es decir, si y solo si λ es un valor propio de A y $\mathbf{v}$ es un vector propio correspondiente. Así, cada par valor propio-vector propio ofrece una solución (4) de $\mathbf{x}' = A\mathbf{x}$. Algunas veces dichas soluciones se conocen como *funciones propias* de la ecuación diferencial. Las funciones propias dan la clave para resolver sistemas de ecuaciones diferenciales.

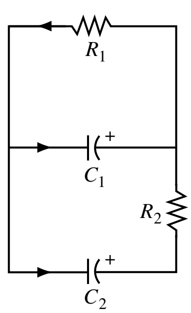


FIGURA 1

EJEMPLO 1 El circuito de la figura 1 se describe con la ecuación diferencial

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -(1/R_1 + 1/R_2)/C_1 & 1/(R_2C_1) \\ 1/(R_2C_2) & -1/(R_2C_2) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

donde $x_1(t)$ y $x_2(t)$ son los voltajes en los dos capacitores al tiempo t. Suponga que el resistor R_1 es de 1 ohm, $R_2 = 2$ ohms, capacitor $C_1 = 1$ farad, y $C_2 = .5$ farad; con cargas iniciales de 5 volts en el capacitor C_1 y de 4 volts en C_2. Encuentre las fórmulas para $x_1(t)$ y $x_2(t)$ que describan cómo los voltajes cambian en el tiempo.

SOLUCIÓN Sea que A denote la matriz mostrada arriba y sea $\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$. Para los datos dados, $A = \begin{bmatrix} -1.5 & .5 \\ 1 & -1 \end{bmatrix}$ y $\mathbf{x}(0) = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$. Los valores propios de A son $\lambda_1 = -.5$ y $\lambda_2 = -2$, con los vectores propios correspondientes

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

Ambas funciones propias $\mathbf{x}_1(t) = \mathbf{v}_1 e^{\lambda_1 t}$ y $\mathbf{x}_2(t) = \mathbf{v}_2 e^{\lambda_2 t}$ satisfacen $\mathbf{x}' = A\mathbf{x}$, y lo mismo ocurre con cualquier combinación lineal de $\mathbf{x}_1$ y $\mathbf{x}_2$. Se hace

$$\mathbf{x}(t) = c_1 \mathbf{v}_1 e^{\lambda_1 t} + c_2 \mathbf{v}_2 e^{\lambda_2 t} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

y observe que $\mathbf{x}(0) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. De manera evidente, puesto que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes y por consiguiente generan $\mathbb{R}^2$, c_1 y c_2 se determinan tal que $\mathbf{x}(0)$ sea igual a $\mathbf{x}_0$. En efecto, la ecuación

$$c_1 \underset{\substack{\uparrow \\ \mathbf{v}_1}}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}} + c_2 \underset{\substack{\uparrow \\ \mathbf{v}_2}}{\begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \underset{\substack{\uparrow \\ \mathbf{x}_0}}{\begin{bmatrix} 5 \\ 4 \end{bmatrix}}$$

conduce fácilmente a $c_1 = 3$ y $c_2 = -2$. Por lo tanto, la solución deseada de la ecuación diferencial $\mathbf{x}' = A\mathbf{x}$ es

$$\mathbf{x}(t) = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - 2\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

o bien,

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 3e^{-.5t} + 2e^{-2t} \\ 6e^{-.5t} - 2e^{-2t} \end{bmatrix}$$

La figura 2 muestra la gráfica, o *trayectoria*, de $\mathbf{x}(t)$, para $t \geq 0$, junto con trayectorias para algunos otros puntos iniciales. Las trayectorias de las dos funciones propias $\mathbf{x}_1$ y $\mathbf{x}_2$ están en los espacios propios de A.

Las funciones $\mathbf{x}_1$ y $\mathbf{x}_2$ decaen a cero cuando $t \to \infty$, pero los valores de $\mathbf{x}_2$ decrecen más rápido porque su exponente es más negativo. Las entradas en el vector propio correspondiente $\mathbf{v}_2$ muestran que los voltajes en los capacitores decaerán a cero tan rápido como sea posible, si los voltajes iniciales son iguales en magnitud pero opuestos en signo. ■

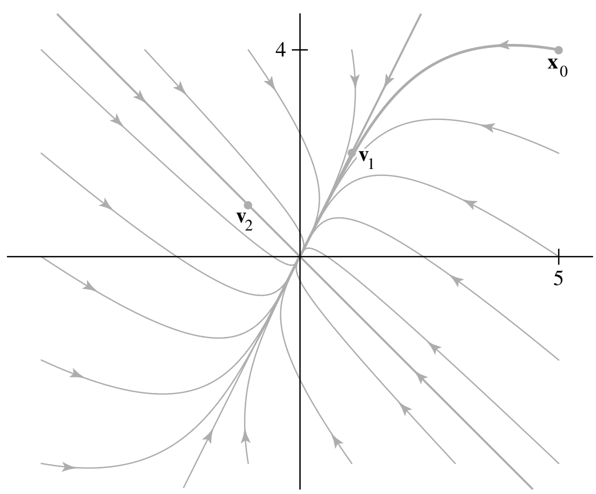


FIGURA 2 El origen como un atractor.

En la figura 2, el origen es un **atractor** o **sumidero** del sistema dinámico porque todas las trayectorias caen al origen. La dirección de mayor atracción está sobre la trayectoria de la función propia $\mathbf{x}_2$ (a lo largo de la recta que pasa por $\mathbf{0}$ y $\mathbf{v}_2$) correspondiente al valor propio más negativo, $\lambda = -2$. Las trayectorias que inician en puntos fuera de esta recta se convierten en asíntotas a la recta que pasa por $\mathbf{0}$ y $\mathbf{v}_1$, ya que sus componentes en la dirección $\mathbf{v}_2$ decaen muy rápidamente.

Si los valores propios del ejemplo 1 fueran positivos en vez de negativos, las trayectorias correspondientes serían similares en forma, aunque las trayectorias estarían *alejándose* del origen. En tal caso, el origen es un **repulsor**, o **fuente**, del sistema dinámico, y la dirección de mayor repulsión es la recta que contiene la trayectoria de la función propia asociada con el valor propio más positivo.

EJEMPLO 2 Suponga que una partícula se mueve en un campo de fuerza plano y que su vector de posición $\mathbf{x}$ satisface $\mathbf{x}' = A\mathbf{x}$ y $\mathbf{x}(0) = \mathbf{x}_0$, donde

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 1 \end{bmatrix}, \qquad \mathbf{x}_0 = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Resuelva este problema de valor inicial para $t \geq 0$, y bosqueje la trayectoria de la partícula.

SOLUCIÓN Los valores propios de A resultan ser $\lambda_1 = 6$ y $\lambda_2 = -1$, con los vectores propios correspondientes $\mathbf{v}_1 = (-5, 2)$ y $\mathbf{v}_2 = (1, 1)$. Para cualesquiera constantes c_1 y c_2, la función

$$\mathbf{x}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t} = c_1\begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + c_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

es una solución de $\mathbf{x}' = A\mathbf{x}$. Se requiere que c_1 y c_2 satisfagan $\mathbf{x}(0) = \mathbf{x}_0$, es decir,

$$c_1\begin{bmatrix} -5 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} -5 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Los cálculos demuestran que $c_1 = -3/70$ y $c_2 = 188/70$, y así la función deseada es

$$\mathbf{x}(t) = \frac{-3}{70}\begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + \frac{188}{70}\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

En la figura 3 se ilustran las trayectorias de $\mathbf{x}$ y otras soluciones. ■

En la figura 3, el origen es un **punto silla** del sistema dinámico porque algunas trayectorias primero se acercan al origen y, después, cambian de dirección y se alejan de él. Un punto silla se presenta cuando la matriz tiene valores propios tanto positivos como negativos. La dirección de mayor repulsión es la recta que pasa por $\mathbf{v}_1$ y $\mathbf{0}$, asociada con el valor propio positivo. La dirección de mayor atracción es la recta que pasa por $\mathbf{v}_2$ y $\mathbf{0}$, correspondiente al valor propio negativo.

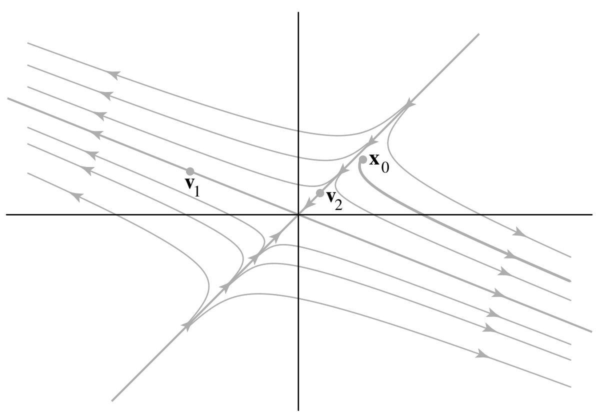


FIGURA 3 El origen como un punto silla.

Desacoplamiento de un sistema dinámico

El siguiente análisis muestra que el método de los ejemplos 1 y 2 produce un conjunto fundamental de soluciones para cualquier sistema dinámico descrito por $\mathbf{x}' = A\mathbf{x}$ cuando A es de $n \times n$ y tiene n vectores propios linealmente independientes, es decir, cuando A es diagonalizable. Suponga que las funciones propias para A son

$$\mathbf{v}_1 e^{\lambda_1 t}, \quad \ldots, \mathbf{v}_n e^{\lambda_n t}$$

con $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vectores propios linealmente independientes. Sean $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ y D la matriz diagonal con entradas $\lambda_1, \ldots, \lambda_n$ tal que $A = PDP^{-1}$. Ahora se hace un *cambio de variable*, definiendo una nueva función $\mathbf{y}$ por

$$\mathbf{y}(t) = P^{-1}\mathbf{x}(t) \quad \text{o, de manera equivalente,} \quad \mathbf{x}(t) = P\mathbf{y}(t)$$

La ecuación $\mathbf{x}(t) = P\mathbf{y}(t)$ indica que $\mathbf{y}(t)$ es el vector de coordenadas de $\mathbf{x}(t)$ respecto a la base de vectores propios. La sustitución de $P\mathbf{y}$ para $\mathbf{x}$ en la ecuación $\mathbf{x}' = A\mathbf{x}$ da

$$\frac{d}{dt}(P\mathbf{y}) = A(P\mathbf{y}) = (PDP^{-1})P\mathbf{y} = PD\mathbf{y} \tag{5}$$

Como P es una matriz constante, el lado izquierdo de (5) es $P\mathbf{y}'$. Ambos lados de (5) se multiplican por la izquierda por P^{-1} y se obtiene $\mathbf{y}' = D\mathbf{y}$, o bien,

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

El cambio de variable de $\mathbf{x}$ a $\mathbf{y}$ *desacopló* el sistema de ecuaciones diferenciales, porque la derivada de cada función escalar y_k tan solo depende de y_k. (Revise el cambio de variables análogo de la sección 5.6). Como $y_1' = \lambda_1 y_1$, se tiene que $y_1(t) = c_1 e^{\lambda_1 t}$, con fórmulas similares para $y_2, \ldots, y_n$. Así,

$$\mathbf{y}(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}, \quad \text{donde} \quad \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{y}(0) = P^{-1}\mathbf{x}(0) = P^{-1}\mathbf{x}_0$$

Para obtener la solución general $\mathbf{x}$ del sistema original, calcule

$$\begin{aligned} \mathbf{x}(t) &= P\mathbf{y}(t) = [\,\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n\,]\,\mathbf{y}(t) \\ &= c_1\mathbf{v}_1 e^{\lambda_1 t} + \cdots + c_n \mathbf{v}_n e^{\lambda_n t} \end{aligned}$$

Este es el desarrollo de funciones propias construidas como en el ejemplo 1.

Valores propios complejos

En el siguiente ejemplo, una matriz real A tiene un par de valores propios complejos λ y $\bar{\lambda}$, con vectores propios complejos asociados $\mathbf{v}$ y $\bar{\mathbf{v}}$. (Recuerde de la sección 5.5 que para una matriz real, los valores propios complejos y sus vectores propios correspondientes se presentan en pares conjugados). Entonces, dos soluciones de $\mathbf{x}' = A\mathbf{x}$ son

$$\mathbf{x}_1(t) = \mathbf{v}e^{\lambda t} \quad \text{y} \quad \mathbf{x}_2(t) = \bar{\mathbf{v}}e^{\bar{\lambda} t} \tag{6}$$

Utilizando una representación en serie de potencias de la función exponencial compleja se demuestra que $\mathbf{x}_2(t) = \overline{\mathbf{x}_1(t)}$. No obstante que las funciones propias complejas $\mathbf{x}_1$ y $\mathbf{x}_2$ son convenientes para algunos cálculos (sobre todo en ingeniería eléctrica), las funciones reales son más adecuadas para diversos fines. Afortunadamente, las partes real e imaginaria de $\mathbf{x}_1$ son soluciones (reales) de $\mathbf{x}' = A\mathbf{x}$, porque son combinaciones lineales de las soluciones de (6):

$$\operatorname{Re}(\mathbf{v}e^{\lambda t}) = \frac{1}{2}[\,\mathbf{x}_1(t) + \overline{\mathbf{x}_1(t)}\,], \qquad \operatorname{Im}(\mathbf{v}e^{\lambda t}) = \frac{1}{2i}[\,\mathbf{x}_1(t) - \overline{\mathbf{x}_1(t)}\,]$$

Para entender la naturaleza de $\operatorname{Re}(\mathbf{v}e^{\lambda t})$, recuerde que en cálculo para cualquier número x, la función exponencial e^x se determina mediante la serie de potencias:

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots$$

la cual se utiliza para definir $e^{\lambda t}$ cuando λ es complejo:

$$e^{\lambda t} = 1 + (\lambda t) + \frac{1}{2!}(\lambda t)^2 + \cdots + \frac{1}{n!}(\lambda t)^n + \cdots$$

Escribiendo $\lambda = a + bi$ (con a y b reales), y utilizando series de potencias similares para las funciones coseno y seno, se obtiene que

$$e^{(a+bi)t} = e^{at} \cdot e^{ibt} = e^{at}(\cos bt + i \operatorname{sen} bt) \tag{7}$$

Por lo tanto,

$$\begin{aligned} \mathbf{v}e^{\lambda t} &= (\operatorname{Re}\mathbf{v} + i \operatorname{Im}\mathbf{v}) \cdot e^{at}(\cos bt + i \operatorname{sen} bt) \\ &= [\,(\operatorname{Re}\mathbf{v})\cos bt - (\operatorname{Im}\mathbf{v})\operatorname{sen} bt\,]e^{at} \\ &\quad + i[\,(\operatorname{Re}\mathbf{v})\operatorname{sen} bt + (\operatorname{Im}\mathbf{v})\cos bt\,]e^{at} \end{aligned}$$

Entonces, dos soluciones reales de $\mathbf{x}' = A\mathbf{x}$ son

$$\begin{aligned} \mathbf{y}_1(t) &= \operatorname{Re}\mathbf{x}_1(t) = [(\operatorname{Re}\mathbf{v})\cos bt - (\operatorname{Im}\mathbf{v})\operatorname{sen} bt]e^{at} \\ \mathbf{y}_2(t) &= \operatorname{Im}\mathbf{x}_1(t) = [(\operatorname{Re}\mathbf{v})\operatorname{sen} bt + (\operatorname{Im}\mathbf{v})\cos bt]e^{at} \end{aligned}$$

Se puede demostrar que $\mathbf{y}_1$ y $\mathbf{y}_2$ son funciones linealmente independientes (cuando $b \neq 0$).[1]

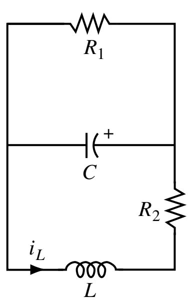


FIGURA 4

EJEMPLO 3 El circuito de la figura 4 se describe con la ecuación

$$\begin{bmatrix} i_L' \\ v_C' \end{bmatrix} = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1C) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

donde i_L es la corriente que pasa en el inductor L y v_C es la caída de voltaje en el capacitor C. Suponga que $R_1 = 5$ ohms, $R_2 = .8$ ohms, $C = .1$ farad y $L = .4$ henry. Encuentre las fórmulas para i_L y v_C, si la corriente inicial en el inductor es 3 amperes y el voltaje inicial en el capacitor es 3 volts.

SOLUCIÓN Para los datos dados, $A = \begin{bmatrix} -2 & -2.5 \\ 10 & -2 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. El método analizado en la sección 5.5 produce el valor propio $\lambda = -2 + 5i$ y el vector propio correspondiente $\mathbf{v}_1 = \begin{bmatrix} i \\ 2 \end{bmatrix}$. Las soluciones complejas de $\mathbf{x}' = A\mathbf{x}$ son combinaciones lineales complejas de

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{(-2+5i)t} \quad \text{y} \quad \mathbf{x}_2(t) = \begin{bmatrix} -i \\ 2 \end{bmatrix} e^{(-2-5i)t}$$

[1] Puesto que $\mathbf{x}_2(t)$ es el complejo conjugado de $\mathbf{x}_1(t)$, las partes real e imaginaria de $\mathbf{x}_2(t)$ son $\mathbf{y}_1(t)$ y $-\mathbf{y}_2(t)$, respectivamente. Entonces, se utiliza ya sea $\mathbf{x}_1(t)$ o $\mathbf{x}_2(t)$, pero no ambas, para producir dos soluciones reales linealmente independientes de $\mathbf{x}' = A\mathbf{x}$.

Ahora, utilice la ecuación (7) para escribir

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{-2t}(\cos 5t + i \operatorname{sen} 5t)$$

Las partes real e imaginaria de $\mathbf{x}_1$ dan soluciones reales:

$$\mathbf{y}_1(t) = \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t}, \qquad \mathbf{y}_2(t) = \begin{bmatrix} \cos 5t \\ 2\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

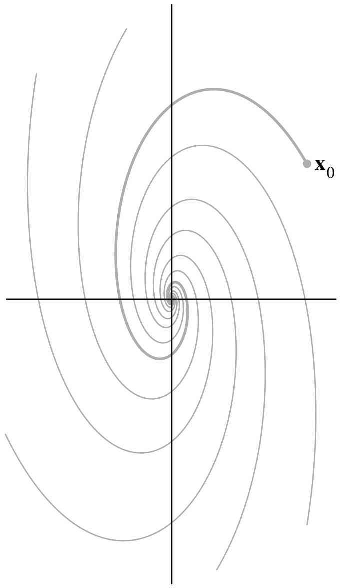


FIGURA 5
El origen como un punto espiral.

Como $\mathbf{y}_1$ y $\mathbf{y}_2$ son funciones linealmente independientes, forman una base para el espacio vectorial bidimensional real de soluciones de $\mathbf{x}' = A\mathbf{x}$. Así, la solución general es

$$\mathbf{x}(t) = c_1 \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \cos 5t \\ 2\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

Para satisfacer $\mathbf{x}(0) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, se necesita $c_1 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, lo que conduce a $c_1 = 1.5$ y $c_2 = 3$. Por lo tanto,

$$\mathbf{x}(t) = 1.5 \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + 3 \begin{bmatrix} \cos 5t \\ 2\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

o bien,

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -1.5\operatorname{sen} 5t + 3\cos 5t \\ 3\cos 5t + 6\operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

Véase la figura 5. ■

En la figura 5, el origen es un **punto espiral** del sistema dinámico. La rotación es causada por las funciones seno y coseno provenientes de un valor propio complejo. Las trayectorias en espiral entran, ya que el factor e^{-2t} tiende a cero. Recuerde que -2 es la parte real del valor propio del ejemplo 3. Cuando A tiene un valor propio complejo con parte real positiva, entonces las trayectorias en espiral son hacia fuera. Si la parte real del valor propio es cero, las trayectorias forman elipses alrededor del origen.

PROBLEMAS DE PRÁCTICA

A es una matriz real de 3×3 con valores propios $-.5$, $.2 + .3i$, y $.2 - .3i$, y vectores propios correspondientes

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_3 = \begin{bmatrix} 1-2i \\ -4i \\ 2 \end{bmatrix}$$

1. ¿A es diagonalizable como $A = PDP^{-1}$, utilizando matrices complejas?
2. Escriba la solución general de $\mathbf{x}' = A\mathbf{x}$ utilizando funciones propias complejas y, después, encuentre la solución general real.
3. Describa las formas de las trayectorias típicas.

5.7 EJERCICIOS

1. Una partícula que se mueve bajo un campo de fuerza plano tiene un vector de posición $\mathbf{x}$ que satisface $\mathbf{x}' = A\mathbf{x}$. La matriz A de 2×2 tiene valores propios 4 y 2, con sus vectores propios asociados $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Encuentre la posición de la partícula al tiempo t, suponiendo que $\mathbf{x}(0) = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$.

2. Sea A una matriz de 2×2 con valores propios -3 y -1 con vectores propios asociados $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Sea $\mathbf{x}(t)$ la posición de una partícula al tiempo t. Resuelva el problema de valor inicial $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

En los ejercicios 3 a 6, resuelva el problema de valor inicial $\mathbf{x}'(t) = A\mathbf{x}(t)$ para $t \geq 0$, con $\mathbf{x}(0) = (3, 2)$. Clasifique la naturaleza del origen como un atractor, repulsor o punto silla del sistema dinámico descrito por $\mathbf{x}' = A\mathbf{x}$. Encuentre las direcciones de mayor atracción y/o repulsión. Cuando el origen sea un punto silla, dibuje trayectorias típicas.

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

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

En los ejercicios 7 y 8, haga un cambio de variable que desacople la ecuación $\mathbf{x}' = A\mathbf{x}$. Escriba la ecuación $\mathbf{x}(t) = P\mathbf{y}(t)$ y muestre el cálculo que conduce al sistema desacoplado $\mathbf{y}' = D\mathbf{y}$, especificando P y D.

7. A como en el ejercicio 5. **8.** A como en el ejercicio 6.

En los ejercicios 9 a 18, construya la solución general de $\mathbf{x}' = A\mathbf{x}$ que implique funciones propias complejas y, después, determine la solución general real. Describa las formas de trayectorias típicas.

9. $A = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix}$ **10.** $A = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$

11. $A = \begin{bmatrix} -3 & -9 \\ 2 & 3 \end{bmatrix}$ **12.** $A = \begin{bmatrix} -7 & 10 \\ -4 & 5 \end{bmatrix}$

13. $A = \begin{bmatrix} 4 & -3 \\ 6 & -2 \end{bmatrix}$ **14.** $A = \begin{bmatrix} -2 & 1 \\ -8 & 2 \end{bmatrix}$

15. [**M**] $A = \begin{bmatrix} -8 & -12 & -6 \\ 2 & 1 & 2 \\ 7 & 12 & 5 \end{bmatrix}$

16. [**M**] $A = \begin{bmatrix} -6 & -11 & 16 \\ 2 & 5 & -4 \\ -4 & -5 & 10 \end{bmatrix}$

17. [**M**] $A = \begin{bmatrix} 30 & 64 & 23 \\ -11 & -23 & -9 \\ 6 & 15 & 4 \end{bmatrix}$

18. [**M**] $A = \begin{bmatrix} 53 & -30 & -2 \\ 90 & -52 & -3 \\ 20 & -10 & 2 \end{bmatrix}$

19. [**M**] Encuentre fórmulas para los voltajes v_1 y v_2 (como funciones del tiempo t) en el circuito del ejemplo 1, suponiendo que $R_1 = 1/5$ ohm, $R_2 = 1/3$ ohm, $C_1 = 4$ farads, $C_2 = 3$ farads, y el voltaje inicial en cada capacitor es 4 volts.

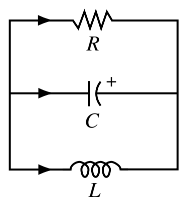
20. [**M**] Obtenga fórmulas para los voltajes v_1 y v_2 en el circuito del ejemplo 1, suponiendo que $R_1 = 1/15$ ohm, $R_2 = 1/3$ ohm, $C_1 = 9$ farads, $C_2 = 2$ farads y el voltaje inicial sobre cada capacitor es 3 volts.

21. [**M**] Encuentre fórmulas para la corriente i_L y el voltaje v_C para el circuito del ejemplo 3, suponiendo que $R_1 = 1$ ohm, $R_2 = .125$ ohm, $C = .2$ farad, $L = .125$ henry, la corriente inicial es 0 amp y el voltaje inicial es 15 volts.

22. [**M**] El circuito en la figura se describe por la ecuación

$$\begin{bmatrix} i'_L \\ v'_C \end{bmatrix} = \begin{bmatrix} 0 & 1/L \\ -1/C & -1/(RC) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

donde i_L es la corriente en el inductor L y v_C es la caída de voltaje en el capacitor C. Obtenga fórmulas para i_L y v_C cuando $R = .5$ ohm, $C = 2.5$ farads, $L = .5$ henry, la corriente inicial es 0 amp y el voltaje inicial de 12 volts.



SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sí, la matriz de 3×3 es diagonalizable ya que tiene tres valores propios distintos. El teorema 2 de la sección 5.1 y el teorema 5 de la sección 5.3 son válidos cuando se utilizan escalares complejos. (Las pruebas son esencialmente las mismas que para los escalares reales).

2. La solución general tiene la forma

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} e^{(.2+.3i)t} + c_3 \begin{bmatrix} 1-2i \\ -4i \\ 2 \end{bmatrix} e^{(.2-.3i)t}$$

Aquí los escalares c_1, c_2, c_3 pueden ser números complejos cualesquiera. El primer término en $\mathbf{x}(t)$ es real. Es posible obtener dos soluciones reales adicionales utilizando las partes real e imaginaria del segundo término en $\mathbf{x}(t)$:

$$\begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} e^{.2t}(\cos .3\mathrm{t} + i \operatorname{sen} .3t)$$

La solución general real tiene la siguiente forma, con escalares *reales* c_1, c_2, c_3:

$$c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} \cos .3t - 2 \operatorname{sen} .3t \\ -4 \operatorname{sen} .3t \\ 2\cos .3t \end{bmatrix} e^{.2t} + c_3 \begin{bmatrix} \operatorname{sen} .3t + 2\cos .3t \\ 4\cos .3t \\ 2 \operatorname{sen} .3t \end{bmatrix} e^{.2t}$$

3. Cualquier solución con $c_2 = c_3 = 0$ es atraída hacia el origen debido al factor exponencial negativo. Otras soluciones tienen componentes que crecen sin cota, en tanto que las trayectorias son espirales hacia afuera.

Sea cuidadoso para no confundir este problema con el de la sección 5.6. Ahí la condición para atracción hacia **0** fue que un valor propio era menor que 1 en magnitud, para que $|\lambda|^k \to 0$. Aquí la parte real del valor propio debe ser negativa, para que $e^{\lambda t} \to 0$.

5.8 ESTIMACIONES ITERATIVAS PARA VALORES PROPIOS

En las aplicaciones científicas del álgebra lineal, los valores propios rara vez se conocen con precisión. Por fortuna, una aproximación numérica cercana es bastante satisfactoria. En efecto, algunas aplicaciones tan solo requieren una aproximación burda al valor propio más grande. El primer algoritmo que se describe abajo sirve muy bien para este caso. Además, proporciona un fundamento para un método más poderoso que también aportaría estimaciones rápidas para los valores propios restantes.

El método de potencias

El método de potencias se aplica a una matriz A de $n \times n$, con un **valor propio estrictamente dominante** λ_1, lo cual significa que λ_1 debe ser, en valor absoluto, mucho más grande que todos los demás valores propios. En tal caso, el método de potencias produce una secuencia escalar que se aproxima a λ_1 y una secuencia vectorial que se aproxima al vector propio correspondiente. El antecedente para el método es la descomposición de vectores propios que se utilizó al inicio de la sección 5.6.

Por sencillez, suponga que A es diagonalizable y que $\mathbb{R}^n$ tiene una base de vectores propios $\mathbf{v}_1, \ldots, \mathbf{v}_n$, colocados de manera que los valores propios correspondientes $\lambda_1, \ldots, \lambda_n$ disminuyan de tamaño, con el primer valor propio estrictamente dominante. Es decir,

$$\underbrace{|\lambda_1| > |\lambda_2|}_{\text{estrictamente mayor}} \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \tag{1}$$

Como se vio en la ecuación (2) de la sección 5.6, si $\mathbf{x}$ en $\mathbb{R}^n$ se escribe en la forma $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, entonces,

$$A^k\mathbf{x} = c_1(\lambda_1)^k\mathbf{v}_1 + c_2(\lambda_2)^k\mathbf{v}_2 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n \quad (k = 1, 2, \ldots)$$

Suponga que $c_1 \neq 0$. Por lo tanto, dividiendo entre $(\lambda_1)^k$,

$$\frac{1}{(\lambda_1)^k}A^k\mathbf{x} = c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n \quad (k = 1, 2, \ldots) \tag{2}$$

De la desigualdad (1), las fracciones $\lambda_2/\lambda_1, \ldots, \lambda_n/\lambda_1$ son menores que 1 en magnitud y por ende sus potencias tienden a cero. Así,

$$(\lambda_1)^{-k}A^k\mathbf{x} \to c_1\mathbf{v}_1 \quad \text{como } k \to \infty \tag{3}$$

De manera que para k grande, un escalar múltiplo de $A^k\mathbf{x}$ casi determina la misma *dirección* como el vector propio $c_1\mathbf{v}_1$. Puesto que un múltiplo escalar positivo no cambia la dirección de un vector, entonces $A^k\mathbf{x}$ casi apunta en la misma dirección que $\mathbf{v}_1$ o $-\mathbf{v}_1$, si $c_1 \neq 0$.

EJEMPLO 1 Sean $A = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} -.5 \\ 1 \end{bmatrix}$. Entonces, A tiene los valores propios 2 y 1, y el espacio propio para $\lambda_1 = 2$ es la línea que pasa por $\mathbf{0}$ y $\mathbf{v}_1$. Para $k = 0, \ldots, 8$, calcule $A^k\mathbf{x}$ y construya la línea por $\mathbf{0}$ y $A^k\mathbf{x}$. ¿Qué sucede conforme k se incrementa?

SOLUCIÓN Los primeros tres cálculos son

$$A\mathbf{x} = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} -.5 \\ 1 \end{bmatrix} = \begin{bmatrix} -.1 \\ 1.1 \end{bmatrix}$$

$$A^2\mathbf{x} = A(A\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} -.1 \\ 1.1 \end{bmatrix} = \begin{bmatrix} .7 \\ 1.3 \end{bmatrix}$$

$$A^3\mathbf{x} = A(A^2\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} .7 \\ 1.3 \end{bmatrix} = \begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$$

Cálculos semejantes permiten elaborar la tabla 1.

TABLA 1 Iteración de un vector

k	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$A^k\mathbf{x}$	$\begin{bmatrix} -.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -.1 \\ 1.1 \end{bmatrix}$	$\begin{bmatrix} .7 \\ 1.3 \end{bmatrix}$	$\begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$	$\begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$	$\begin{bmatrix} 11.9 \\ 4.1 \end{bmatrix}$	$\begin{bmatrix} 24.7 \\ 7.3 \end{bmatrix}$	$\begin{bmatrix} 50.3 \\ 13.7 \end{bmatrix}$	$\begin{bmatrix} 101.5 \\ 26.5 \end{bmatrix}$

En la figura 1 se ilustran los vectores $\mathbf{x}, A\mathbf{x}, \ldots, A^4\mathbf{x}$. Los otros crecen demasiado y se dificulta dibujarlos. Sin embargo, se han trazado segmentos de recta para mostrar las direcciones de esos vectores. En efecto, las direcciones de los vectores es lo que realmente se desea observar, no los vectores mismos. Las rectas parecen aproximarse a la recta que representa el espacio propio generado por $\mathbf{v}_1$. Más precisamente, el ángulo entre la recta (subespacio) determinada por $A^k\mathbf{x}$ y la recta (espacio propio) definida por $\mathbf{v}_1$ tiende a cero cuando $k \to \infty$. ■

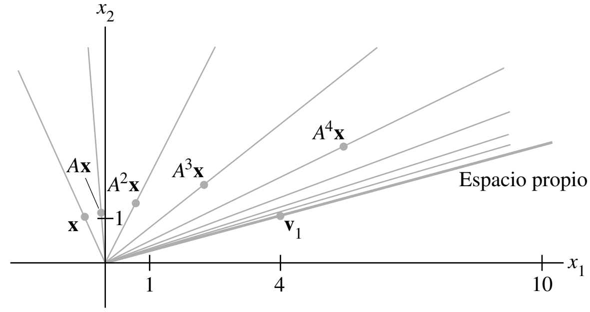


FIGURA 1 Direcciones determinadas por $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots, A^7\mathbf{x}$.

Se escalan los vectores $(\lambda_1)^{-k}A^k\mathbf{x}$ en la ecuación (3) para hacerlos converger a $c_1\mathbf{v}_1$, si $c_1 \neq 0$. No se puede escalar $A^k\mathbf{x}$ en esta forma porque se desconoce λ_1. Sin embargo, sí es posible escalar cada $A^k\mathbf{x}$ para hacer que su entrada sea 1. Entonces, resulta que la secuencia $\{\boldsymbol{x}_k\}$ convergerá a un múltiplo de $\mathbf{v}_1$ cuya mayor entrada es 1. La figura 2 muestra la secuencia escalada para el ejemplo 1. El valor propio λ_1 también se estima a partir de la secuencia $\{\mathbf{x}_k\}$. Cuando $\mathbf{x}_k$ se acerca a un vector propio para λ_1, el vector $A\mathbf{x}_k$ está cerca de $\lambda_1\mathbf{x}_k$, con cada entrada en $A\mathbf{x}_k$ aproximadamente igual a λ_1 veces la entrada correspondiente en $\mathbf{x}_k$. Como la entrada más grande en $\mathbf{x}_k$ es 1, por lo tanto, la mayor entrada en $A\mathbf{x}_k$ es casi λ_1. (Se omiten pruebas cuidadosas de estos enunciados).

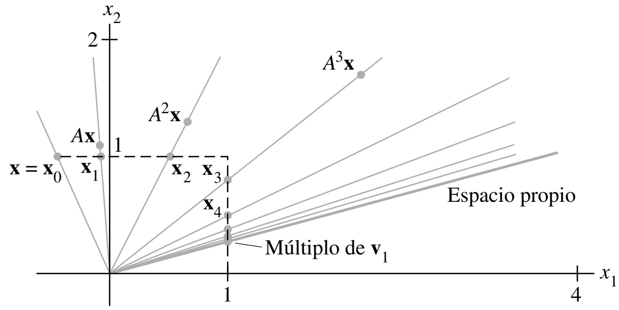


FIGURA 2 Múltiplos escalados de $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots, A^7\mathbf{x}$.

MÉTODO DE POTENCIAS PARA ESTIMAR UN VALOR PROPIO ESTRICTAMENTE DOMINANTE

1. Seleccione un vector inicial $\mathbf{x}_0$ cuya entrada más grande sea 1.
2. Para $k = 0, 1, \ldots,$
   a) Calcule $A\mathbf{x}_k$.
   b) Sea μ_k una entrada en $A\mathbf{x}_k$ cuyo valor absoluto es tan grande como sea posible.
   c) Determine $\mathbf{x}_{k+1} = (1/\mu_k)A\mathbf{x}_k$.
3. Para casi todas las elecciones de $\mathbf{x}_0$, la secuencia $\{\mu_k\}$ se aproxima al valor propio dominante, y la secuencia $\{\mathbf{x}_k\}$ tiende a un vector propio correspondiente.

EJEMPLO 2 Aplique el método de potencias a $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ con $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Pare cuando $k = 5$, y estime el valor propio dominante y un vector propio correspondiente de A.

SOLUCIÓN Los cálculos en este ejemplo y en el siguiente fueron realizados con MATLAB, que calcula con 16 dígitos de exactitud, no obstante que aquí tan solo se muestran pocos decimales. Para comenzar, calcule $A\mathbf{x}_0$ e identifique la entrada más grande μ_0 en $A\mathbf{x}_0$:

$$A\mathbf{x}_0 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \quad \mu_0 = 5$$

Escale $A\mathbf{x}_0$ por $1/\mu_0$ para obtener $\mathbf{x}_1$, calcule $A\mathbf{x}_1$ e identifique la mayor entrada en $A\mathbf{x}_1$:

$$\mathbf{x}_1 = \frac{1}{\mu_0}A\mathbf{x}_0 = \frac{1}{5}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ .4 \end{bmatrix}$$

$$A\mathbf{x}_1 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix}, \quad \mu_1 = 8$$

Escale $A\mathbf{x}_1$ por $1/\mu_1$ para obtener $\mathbf{x}_2$, determine $A\mathbf{x}_2$ e identifique la mayor entrada en $A\mathbf{x}_2$:

$$\mathbf{x}_2 = \frac{1}{\mu_1}A\mathbf{x}_1 = \frac{1}{8}\begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ .225 \end{bmatrix}$$

$$A\mathbf{x}_2 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}, \quad \mu_2 = 7.125$$

Escale $A\mathbf{x}_2$ por $1/\mu_2$ para obtener $\mathbf{x}_3$, y así sucesivamente. Los resultados de los cálculos con MATLAB para las primeras cinco iteraciones se dan en la tabla 2.

TABLA 2 Método de potencias para el ejemplo 2

k	**0**	**1**	**2**	**3**	**4**	**5**
$\mathbf{x}_k$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .20007 \end{bmatrix}$
$A\mathbf{x}_k$	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

La evidencia de la tabla 2 sugiere significativamente que $\{\mathbf{x}_k\}$ se aproxima a (1, .2) y que $\{\mu_k\}$ se aproxima a 7. Si es así, entonces (1, .2) es un vector propio y 7 es el valor propio dominante. Esto se comprueba fácilmente calculando

$$A\begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1.4 \end{bmatrix} = 7\begin{bmatrix} 1 \\ .2 \end{bmatrix} \qquad \blacksquare$$

La secuencia $\{\mu_k\}$ del ejemplo 2 tuvo convergencia rápida a $\lambda_1 = 7$ porque el segundo valor propio de A fue mucho más pequeño. (En efecto, $\lambda_2 = 1$). En general, la rapidez de convergencia depende de la razón $|\lambda_2/\lambda_1|$, porque en la ecuación (2) el vector $c_2(\lambda_2/\lambda_1)^k\mathbf{v}_2$ es la principal fuente de error, cuando se utiliza una versión escalada de $A^k\mathbf{x}$ como un estimador de $c_1\mathbf{v}_1$. (Las otras fracciones λ_j/λ_1 son mucho más pequeñas). Si $|\lambda_2/\lambda_1|$ es cercana a 1, entonces $\{\mu_k\}$ y $\{\mathbf{x}_k\}$ pueden converger muy lentamente y se preferirían otros métodos de aproximación.

Con el método de potencias, hay una leve posibilidad de que el vector inicial elegido $\mathbf{x}$ no tenga componente en la dirección $\mathbf{v}_1$ (cuando $c_1 = 0$). Pero el error de redondeo computacional durante los cálculos de $\mathbf{x}_k$ podría crear un vector con al menos una pequeña componente en la dirección de $\mathbf{v}_1$. Si eso ocurre, entonces $\mathbf{x}_k$ empezará a converger a un múltiplo de $\mathbf{v}_1$.

Método de potencias inverso

Este método brinda una aproximación para *cualquier* valor propio, si se tiene una buena estimación inicial α del valor propio λ. En tal caso, sea $B = (A - \alpha I)^{-1}$ y se le aplica el método de potencias a B. Se puede demostrar que si los valores propios de A son $\lambda_1, \ldots, \lambda_n$, entonces los valores propios de B son

$$\frac{1}{\lambda_1 - \alpha}, \quad \frac{1}{\lambda_2 - \alpha}, \quad \ldots, \quad \frac{1}{\lambda_n - \alpha}$$

y los vectores propios correspondientes coinciden con los de A. (Véase los ejercicios 15 y 16).

Suponga, por ejemplo, que α está más cerca de λ_2 que de los otros valores propios de A. Entonces, $1/(\lambda_2 - \alpha)$ será un valor propio estrictamente dominante de B. Si α está realmente cerca de λ_2, entonces $1/(\lambda_2 - \alpha)$ será *mucho* mayor que los demás valores propios de B, y el método de potencias inverso produce una aproximación rápida a λ_2 para casi todas las elecciones de $\mathbf{x}_0$. El siguiente algoritmo presenta los detalles.

MÉTODO DE POTENCIAS INVERSO PARA ESTIMAR UN VALOR PROPIO λ DE A

1. Seleccione una estimación inicial λ lo suficientemente cerca a λ.
2. Elija un vector inicial $\mathbf{x}_0$ cuya entrada más grande sea 1.
3. Para $k = 0, 1, \ldots,$
   a) Resuelva $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ para $\mathbf{y}_k$.
   b) Sea μ_k una entrada en $\mathbf{y}_k$ cuyo valor absoluto es tan grande como sea posible.
   c) Calcule $\nu_k = \alpha + (1/\mu_k)$.
   d) Determine $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$.
4. Para casi todas las elecciones de $\mathbf{x}_0$, la secuencia $\{\nu_k\}$ se aproxima al valor propio λ de A, y la secuencia $\{\mathbf{x}_k\}$ tiende a un vector propio correspondiente.

Observe que B, o bien $(A - \alpha I)^{-1}$, no se encuentra en el algoritmo. En vez de calcular $(A - \alpha I)^{-1}\mathbf{x}_k$ para obtener el siguiente vector en la secuencia, es mucho mejor *resolver* la ecuación $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ para $\mathbf{y}_k$ (y después escalar $\mathbf{y}_k$ para producir $\mathbf{x}_{k+1}$). Como esta ecuación para $\mathbf{y}_k$ se debe resolver para cada k, entonces una factorización LU de $A - \alpha I$ acelerará el proceso.

EJEMPLO 3 En algunas aplicaciones no es poco común necesitar conocer el valor propio más pequeño de una matriz A y tener a mano estimaciones burdas de los valores propios. Suponga que 21, 3.3 y 1.9 son estimaciones para los valores propios de la matriz A que se muestra a continuación. Encuentre el valor propio más pequeño, exacto a seis decimales.

$$A = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

SOLUCIÓN Los dos valores propios menores parecen estar cercanos entre sí, de modo que se emplea el método de potencias inverso para $A - 1.9I$. La tabla 3 muestra los resultados de los cálculos obtenidos con MATLAB. Aquí $\mathbf{x}_0$ se eligió arbitrariamente, $\mathbf{y}_k = (A - 1.9I)^{-1}\mathbf{x}_k$, μ_k es la entrada más grande en $\mathbf{y}_k$, $\mathbf{v}_k = 1.9 + 1/\mu_k$ y $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$. Resulta que la estimación inicial del valor propio fue muy buena y la secuencia de potencias inversa convergió de manera rápida. El valor propio más pequeño es exactamente 2. ■

TABLA 3 Método de potencias inverso

k	**0**	**1**	**2**	**3**	**4**
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5736 \\ .0646 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5054 \\ .0045 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5004 \\ .0003 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .50003 \\ .00002 \\ 1 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 4.45 \\ .50 \\ 7.76 \end{bmatrix}$	$\begin{bmatrix} 5.0131 \\ .0442 \\ 9.9197 \end{bmatrix}$	$\begin{bmatrix} 5.0012 \\ .0031 \\ 9.9949 \end{bmatrix}$	$\begin{bmatrix} 5.0001 \\ .0002 \\ 9.9996 \end{bmatrix}$	$\begin{bmatrix} 5.000006 \\ .000015 \\ 9.999975 \end{bmatrix}$
μ_k	7.76	9.9197	9.9949	9.9996	9.999975
ν_k	2.03	2.0008	2.00005	2.000004	2.0000002

Si no se dispone de una estimación para el valor propio más pequeño de una matriz, se puede simplemente tomar $\alpha = 0$ en el método de potencias inverso. Esta elección de α funciona razonablemente bien, cuando el valor propio más pequeño está mucho más cerca de cero que los demás valores propios.

Los dos algoritmos presentados en esta sección son herramientas prácticas para muchas situaciones simples, y brindan una introducción al problema de la estimación de valores propios. Un método iterativo más robusto y ampliamente utilizado es el algoritmo QR. Por ejemplo, este es el corazón del comando `eig(A)` de MATLAB, que rápidamente calcula valores propios y vectores propios de A. En los ejercicios de la sección 5.2 se presenta una breve descripción del algoritmo QR. En casi todos los libros modernos de análisis numérico se presentan más detalles.

PROBLEMA DE PRÁCTICA

¿Cómo se podría decidir si un vector **x** dado es una buena aproximación a un vector propio de una matriz A? Si lo es, ¿cómo se estimaría el valor propio correspondiente? Experimente con

$$A = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} 1.0 \\ -4.3 \\ 8.1 \end{bmatrix}$$

5.8 EJERCICIOS

En los ejercicios 1 a 4, la matriz A va seguida por una secuencia $\{\mathbf{x}_k\}$ generada por el método de potencias. Utilice esos datos para estimar el mayor valor propio de A, y dé un vector propio correspondiente.

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

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .25 \end{bmatrix}, \begin{bmatrix} 1 \\ .3158 \end{bmatrix}, \begin{bmatrix} 1 \\ .3298 \end{bmatrix}, \begin{bmatrix} 1 \\ .3326 \end{bmatrix}$$

2. $A = \begin{bmatrix} 1.8 & -.8 \\ -3.2 & 4.2 \end{bmatrix}$;

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -.5625 \\ 1 \end{bmatrix}, \begin{bmatrix} -.3021 \\ 1 \end{bmatrix}, \begin{bmatrix} -.2601 \\ 1 \end{bmatrix}, \begin{bmatrix} -.2520 \\ 1 \end{bmatrix}$$

3. $A = \begin{bmatrix} .5 & .2 \\ .4 & .7 \end{bmatrix}$;

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .8 \end{bmatrix}, \begin{bmatrix} .6875 \\ 1 \end{bmatrix}, \begin{bmatrix} .5577 \\ 1 \end{bmatrix}, \begin{bmatrix} .5188 \\ 1 \end{bmatrix}$$

4. $A = \begin{bmatrix} 4.1 & -6 \\ 3 & -4.4 \end{bmatrix}$;

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .7368 \end{bmatrix}, \begin{bmatrix} 1 \\ .7541 \end{bmatrix}, \begin{bmatrix} 1 \\ .7490 \end{bmatrix}, \begin{bmatrix} 1 \\ .7502 \end{bmatrix}$$

5. Sea $A = \begin{bmatrix} 15 & 16 \\ -20 & -21 \end{bmatrix}$. Los vectores $\mathbf{x}, \ldots, A^5\mathbf{x}$ son $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 31 \\ -41 \end{bmatrix}, \begin{bmatrix} -191 \\ 241 \end{bmatrix}, \begin{bmatrix} 991 \\ -1241 \end{bmatrix}, \begin{bmatrix} -4991 \\ 6241 \end{bmatrix}, \begin{bmatrix} 24991 \\ -31241 \end{bmatrix}.$$

Encuentre un vector con un 1 en la segunda entrada que sea cercano a un vector propio de A. Use cuatro lugares decimales. Compruebe su respuesta y dé una estimación para el valor propio dominante de A.

6. Sea $A = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix}$. Repita el ejercicio 5, utilizando la siguiente secuencia $\mathbf{x}, A\mathbf{x}, \ldots, A^5\mathbf{x}$.

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 13 \end{bmatrix}, \begin{bmatrix} -29 \\ 61 \end{bmatrix}, \begin{bmatrix} -125 \\ 253 \end{bmatrix}, \begin{bmatrix} -509 \\ 1021 \end{bmatrix}, \begin{bmatrix} -2045 \\ 4093 \end{bmatrix}$$

[**M**] Los ejercicios 7 a 12 requieren MATLAB u otra herramienta computacional. En los ejercicios 7 y 8, utilice el método de potencias con $\mathbf{x}_0$ dado. Liste $\{\mathbf{x}_k\}$ y $\{\mu_k\}$ para $k = 1, \ldots, 5$. En los ejercicios 9 y 10, indique μ_5 y μ_6.

7. $A = \begin{bmatrix} 6 & 7 \\ 8 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

8. $A = \begin{bmatrix} 2 & 1 \\ 4 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

9. $A = \begin{bmatrix} 8 & 0 & 12 \\ 1 & -2 & 1 \\ 0 & 3 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 9 \\ 0 & 1 & 9 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Otra estimación se puede realizar para un valor propio cuando está disponible un vector propio aproximado. Observe que si $A\mathbf{x} = \lambda\mathbf{x}$, entonces $\mathbf{x}^TA\mathbf{x} = \mathbf{x}^T(\lambda\mathbf{x}) = \lambda(\mathbf{x}^T\mathbf{x})$, y el **cociente de Rayleigh**

$$R(\mathbf{x}) = \frac{\mathbf{x}^TA\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

es igual a λ. Si **x** es cercano a un vector propio para λ, entonces este cociente es cercano a λ. Cuando A es una matriz simétrica ($A^T = A$), el cociente de Rayleigh $R(\mathbf{x}_k) = (\mathbf{x}_k^TA\mathbf{x}_k)/(\mathbf{x}_k^T\mathbf{x}_k)$ tendrá aproximadamente el doble de dígitos de exactitud que el factor de escalamiento μ_k en el método de potencias. En los ejercicios 11 y 12 compruebe el aumento de precisión mediante el cálculo de μ_k y $R(\mathbf{x}_k)$ para $k = 1, \ldots, 4$.

11. $A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

12. $A = \begin{bmatrix} -3 & 2 \\ 2 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Los ejercicios 13 y 14 se aplican a una matriz A de 3×3 cuyos valores propios se estiman como 4, -4 y 3.

13. Si se conoce que los valores propios cercanos a 4 y -4 tienen diferentes valores absolutos, ¿funcionará el método de potencias? ¿El método continúa siendo útil?

14. Suponga que se conoce que los valores propios cercanos a 4 y -4 tienen exactamente el mismo valor absoluto. Describa cómo se podría obtener una secuencia que estime el valor propio cercano a 4.

15. Suponga que $A\mathbf{x} = \lambda\mathbf{x}$ con $\mathbf{x} \neq \mathbf{0}$. Sean α un escalar distinto de los valores propios de A, y $B = (A - \alpha I)^{-1}$. Reste $\alpha\mathbf{x}$ en ambos miembros de la ecuación $A\mathbf{x} = \lambda\mathbf{x}$, y utilice álgebra para demostrar que $1/(\lambda - \alpha)$ es un valor propio de B, con $\mathbf{x}$ como un vector propio correspondiente.

16. Suponga que μ es un valor propio de B en el ejercicio 15, y que $\mathbf{x}$ es un vector propio correspondiente, de manera que $(A - \alpha I)^{-1}\mathbf{x} = \mu\mathbf{x}$. Utilice esta ecuación para encontrar un valor propio de A en términos de μ y α. [*Nota:* $\mu \neq 0$ porque B es invertible].

17. [**M**] Utilice el método de potencias inverso para estimar el valor propio intermedio de la A del ejemplo 3, con una exactitud a cuatro lugares decimales. Haga $\mathbf{x}_0 = (1, 0, 0)$.

18. [**M**] Sea A como en el ejercicio 9. Aplique el método de potencias inverso con $\mathbf{x}_0 = (1, 0, 0)$ para estimar el valor propio de A cercano a $\alpha = -1.4$, con una exactitud a cuatro decimales.

[**M**] En los ejercicios 19 y 20, determine: *a*) el valor propio más grande y *b*) el valor propio más cercano a cero. En cada caso, $\mathbf{x}_0 = (1, 0, 0, 0)$ y realice las aproximaciones hasta que la secuencia de acercamiento parezca exacta con cuatro decimales. Incluya el vector propio aproximado.

19. $A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}$

20. $A = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 12 & 13 & 11 \\ -2 & 3 & 0 & 2 \\ 4 & 5 & 7 & 2 \end{bmatrix}$

21. Un error común es que si A tiene un valor propio estrictamente dominante, entonces, para cualquier k suficientemente grande, el vector $A^k\mathbf{x}$ es aproximadamente igual a un vector propio de A. Para las tres matrices que se muestran a continuación, estudie qué sucede a $A^k\mathbf{x}$ cuando $\mathbf{x} = (.5, .5)$, e intente obtener conclusiones generales (para una matriz de 2×2).

a) $A = \begin{bmatrix} .8 & 0 \\ 0 & .2 \end{bmatrix}$ *b*) $A = \begin{bmatrix} 1 & 0 \\ 0 & .8 \end{bmatrix}$

c) $A = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Para A y $\mathbf{x}$ dados,

$$A\mathbf{x} = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1.00 \\ -4.30 \\ 8.10 \end{bmatrix} = \begin{bmatrix} 3.00 \\ -13.00 \\ 24.50 \end{bmatrix}$$

Si $A\mathbf{x}$ es cercanamente un múltiplo de $\mathbf{x}$, entonces las razones de las entradas correspondientes en los dos vectores deberían ser casi constantes. Entonces, calcule:

$$\begin{array}{ccccc} \{\text{entrada en } A\mathbf{x}\} & \div & \{\text{entrada en } \mathbf{x}\} & = & \{\text{razón}\} \\ 3.00 & & 1.00 & & 3.000 \\ -13.00 & & -4.30 & & 3.023 \\ 24.50 & & 8.10 & & 3.025 \end{array}$$

Cada entrada en $A\mathbf{x}$ es alrededor de tres veces la entrada correspondiente en $\mathbf{x}$, de manera que $\mathbf{x}$ es cercano a un vector propio de A. Cualesquiera de las razones anteriores es una estimación del valor propio. (Con cinco decimales, el valor propio es 3.02409).

WEB

CAPÍTULO 5 EJERCICIOS COMPLEMENTARIOS

En estos ejercicios suplementarios, A y B representan matrices cuadradas de tamaño adecuado.

1. Marque cada enunciado como verdadero o falso. Justifique cada respuesta.

a) Si A es invertible y 1 es uno de sus valores propios, entonces 1 también es valor propio de A^{-1}.

b) Si A es equivalente por filas a la matriz identidad I, entonces A es diagonalizable.

c) Si A contiene una fila o una columna de ceros, entonces 0 es un valor propio de A.

d) Cada valor propio de A también lo es de A^2.

e) Cada vector propio de A también lo es de A^2.

f) Cada vector propio de una matriz invertible A también lo es de A^{-1}.

g) Los valores propios deben ser escalares diferentes de cero.

h) Los vectores propios deben ser vectores distintos de cero.

i) Dos vectores propios correspondientes al mismo valor propio siempre son linealmente dependientes.

j) Las matrices similares siempre tienen exactamente los mismos valores propios.

k) Las matrices similares siempre tienen exactamente los mismos vectores propios.

l) La suma de dos vectores propios de una matriz A también es vector propio de esta.

m) Los valores propios de una matriz triangular A superior son exactamente las entradas distintas de cero sobre la diagonal de A.

n) Las matrices A y A^T tienen los mismos valores propios, contando sus multiplicidades.

o) Si una matriz A de 5×5 tiene menos de 5 valores propios distintos, entonces A no es diagonalizable.

p) Existe una matriz de 2×2 que no tiene vectores propios en $\mathbb{R}^2$.

q) Si A es diagonalizable, entonces las columnas de A son linealmente independientes.

r) Un vector distinto de cero no puede corresponder a dos valores propios diferentes de A.

s) Una matriz (cuadrada) A es invertible si y solo si existe un sistema de coordenadas donde la transformación $\mathbf{x} \mapsto A\mathbf{x}$ esté representada por una matriz diagonal.

t) Si cada vector $\mathbf{e}_j$ en la base estándar para $\mathbb{R}^n$ es un vector propio de A, entonces A es una matriz diagonal.

u) Si A es similar a una matriz diagonalizable B, entonces A también es diagonalizable.

v) Si A y B son matrices de $n \times n$ invertibles, entonces AB es similar a BA.

w) Una matriz de $n \times n$ con n vectores propios linealmente independientes es invertible.

x) Si A es una matriz de $n \times n$ diagonalizable, entonces cada vector en $\mathbb{R}^n$ se escribe como una combinación lineal de vectores propios de A.

2. Demuestre que si $\mathbf{x}$ es un vector propio de la matriz producto AB y $B\mathbf{x} \neq \mathbf{0}$, entonces $B\mathbf{x}$ es un vector propio de BA.

3. Suponga que $\mathbf{x}$ es un vector propio de A correspondiente a un valor propio λ.

a) Pruebe que $\mathbf{x}$ es un vector propio de $5I - A$. ¿Cuál es el valor propio asociado?

b) Demuestre que $\mathbf{x}$ es un vector propio de $5I - 3A + A^2$. ¿Cuál es el valor propio asociado?

4. Utilice inducción matemática para demostrar que si λ es un valor propio de una matriz A de $n \times n$, con $\mathbf{x}$ como un vector propio correspondiente, entonces, para cada entero positivo m, λ^m es un valor propio de A^m con $\mathbf{x}$ como un vector propio correspondiente.

5. Si $p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n$, defina $p(A)$ como la matriz formada al remplazar cada potencia de t en $p(t)$ por la potencia correspondiente de A (con $A^0 = I$). Es decir,

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots + c_n A^n$$

Demuestre que si λ es un valor propio de A, entonces un valor propio de $p(A)$ es $p(\lambda)$.

6. Suponga que $A = PDP^{-1}$, donde P es 2×2 y $D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.

a) Sea $B = 5I - 3A + A^2$. Demuestre que B es diagonalizable encontrando una factorización conveniente de B.

b) Dados $p(t)$ y $p(A)$ como en el ejercicio 5, demuestre que $p(A)$ es diagonalizable.

7. Suponga que A es diagonalizable y que $p(t)$ es el polinomio característico de A. Defina $p(A)$ como en el ejercicio 5, y pruebe que $p(A)$ es la matriz cero. Este hecho, el cual es válido para *cualquier* matriz cuadrada, se conoce como *teorema de Cayley-Hamilton*.

8. *a*) Sea A una matriz de $n \times n$ diagonalizable. Demuestre que si la multiplicidad de un valor propio λ es n, entonces $A = \lambda I$.

b) Utilice el inciso *a*) para demostrar que la matriz $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ no es diagonalizable.

9. Demuestre que $I - A$ es invertible cuando todos los valores propios de A son menores que 1 en magnitud. [*Sugerencia:* ¿Qué sería cierto si $I - A$ no fuera invertible?]

10. Demuestre que si A es diagonalizable, con todos los valores propios menores que 1 en magnitud, entonces A^k tiende a la matriz cero conforme $k \to \infty$. [*Sugerencia:* Considere $A^k\mathbf{x}$ donde $\mathbf{x}$ representa cualquiera de las columnas de I].

11. Sean $\mathbf{u}$ un vector propio de A correspondiente a un valor propio λ, y H la recta en $\mathbb{R}^n$ que pasa por $\mathbf{u}$ y por el origen.

a) Explique por qué H es invariante bajo A en el sentido de que $A\mathbf{x}$ está en H siempre que $\mathbf{x}$ esté en H.

b) Sea K el subespacio unidimensional de $\mathbb{R}^n$ que es invariante bajo A. Explique por qué K contiene un vector propio de A.

12. Sea $G = \begin{bmatrix} A & X \\ 0 & B \end{bmatrix}$. Utilice la fórmula (1) para el determinante dado en la sección 5.2 para explicar por qué $\det G = (\det A)(\det B)$. De esto, deduzca que el polinomio característico de G es el producto de los polinomios característicos de A y B.

Utilice el ejercicio 12 para encontrar los valores propios de las matrices en los ejercicios 13 y 14.

13. $A = \begin{bmatrix} 3 & -2 & 8 \\ 0 & 5 & -2 \\ 0 & -4 & 3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 5 & -6 & -7 \\ 2 & 4 & 5 & 2 \\ 0 & 0 & -7 & -4 \\ 0 & 0 & 3 & 1 \end{bmatrix}$

15. Sea J la matriz de $n \times n$ con todas las entradas igual a 1, y considere $A = (a - b)I + bJ$; es decir,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Con los resultados del ejercicio 16 de los ejercicios complementarios del capítulo 3 demuestre que los valores propios de A son $a - b$ y $a + (n - 1)b$. ¿Cuáles son las multiplicidades de esos valores propios?

16. Aplique el resultado del ejercicio 15 para obtener los valores propios de las matrices $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$

y $\begin{bmatrix} 7 & 3 & 3 & 3 & 3 \\ 3 & 7 & 3 & 3 & 3 \\ 3 & 3 & 7 & 3 & 3 \\ 3 & 3 & 3 & 7 & 3 \\ 3 & 3 & 3 & 3 & 7 \end{bmatrix}$.

17. Sea $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Recuerde del ejercicio 25 de la sección 5.4 que tr A (la traza de A) es la suma de las entradas diagonales en A. Demuestre que el polinomio característico de A es

$\lambda^2 - (\operatorname{tr} A)\lambda + \det A$

Entonces, demuestre que los valores propios de una matriz A de 2×2 son reales si y solo si $\det A \leq \left(\frac{\operatorname{tr} A}{2}\right)^2$.

18. Sea $A = \begin{bmatrix} .4 & -.3 \\ .4 & 1.2 \end{bmatrix}$. Explique por qué A^k se aproxima a $\begin{bmatrix} -.5 & -.75 \\ 1.0 & 1.50 \end{bmatrix}$ cuando $k \to \infty$.

Los ejercicios 19 a 23 conciernen al polinomio

$p(t) = a_0 + a_1 t + \cdots + a_{n-1}t^{n-1} + t^n$

y a una matriz C_p de $n \times n$ llamada **matriz compañera** de p:

$$C_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}$$

19. Escriba la matriz compañera C_p para $p(t) = 6 - 5t + t^2$ y, luego, determine el polinomio característico de C_p.

20. Sea $p(t) = (t - 2)(t - 3)(t - 4) = -24 + 26t - 9t^2 + t^3$. Escriba la matriz compañera para $p(t)$, y utilice las técnicas del capítulo 3 para obtener su polinomio característico.

21. Utilice inducción matemática y pruebe que para $n \geq 2$,

$$\begin{aligned} \det(C_p - \lambda I) &= (-1)^n(a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n) \\ &= (-1)^n p(\lambda) \end{aligned}$$

[*Sugerencia:* Desarrolle por cofactores sobre la primera columna, para demostrar que det $(C_p - \lambda I)$ tiene la forma $(-\lambda)B + (-1)^n a_0$, donde B es cierto polinomio (por la suposición de inducción)].

22. Sean $p(t) = a_0 + a_1 t + a_2 t^2 + t^3$, y λ un cero de p.

a) Escriba la matriz compañera para p.

b) Explique por qué $\lambda^3 = -a_0 - a_1\lambda - a_2\lambda^2$, y demuestre que $(1, \lambda, \lambda^2)$ es un vector propio de la matriz compañera de p.

23. Sea p el polinomio del ejercicio 22, y suponga que la ecuación $p(t) = 0$ tiene las raíces distintas $\lambda_1, \lambda_2, \lambda_3$. Sea V la matriz de Vandermonde

$$V = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix}$$

(La transpuesta de V se consideró en el ejercicio complementario 11 del capítulo 2). Utilice el ejercicio 22 y un teorema de este capítulo para deducir que V es invertible (pero no calcule V^{-1}). Después, explique por qué $V^{-1}C_pV$ es una matriz diagonal.

24. [**M**] El comando `roots(p)` de MATLAB calcula las raíces de la ecuación polinomial $p(t) = 0$. Lea un manual de MATLAB y, luego, describa la idea básica que sustenta el algoritmo para el comando `roots`.

25. [**M**] Aplique un programa matricial para diagonalizar

$$A = \begin{bmatrix} -3 & -2 & 0 \\ 14 & 7 & -1 \\ -6 & -3 & 1 \end{bmatrix}$$

si es posible. Utilice el comando de valores propios para construir la matriz diagonal D. Si el programa tiene un comando que genere vectores propios, úselo para crear una matriz invertible P. Después, determine $AP - PD$ y PDP^{-1}. Analice sus resultados.

26. [**M**] Repita el ejercicio 25 para $A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$.

6

Ortogonalidad y mínimos cuadrados

EJEMPLO INTRODUCTORIO

Base de datos geográficos de Norteamérica y sistema de navegación GPS

Imagine que inicia un enorme proyecto que, según las estimaciones, durará 10 años y requiere los esfuerzos de muchas personas para construir y resolver un sistema de 1,800,000 por 900,000 ecuaciones lineales. Esto es exactamente lo que se hizo en 1974 con la encuesta geodésica nacional de Estados Unidos, cuando se actualizó el North American Datum (NAD), una red de 268,000 puntos de referencia localizados de manera precisa que generan el territorio completo de América del Norte, junto con Groenlandia, Hawai, las Islas Vírgenes, Puerto Rico y otras islas del Caribe.

Las latitudes y longitudes registradas en el NAD se deben determinar con una precisión dentro del rango de unos cuantos centímetros, ya que constituyen la base de encuestas, mapas, límites legales de propiedad y planos de proyectos de ingeniería civil, como los de carreteras e instalaciones de servicios públicos. Sin embargo, se tuvieron que agregar más de 200,000 nuevos puntos a los ya existentes desde la última actualización en 1927, y se acumularon errores con el paso de los años debido a mediciones imprecisas y a corrimientos de la corteza terrestre. En 1983 concluyó la recolección de datos para el reajuste del NAD.

El sistema de ecuaciones del NAD no tenía solución de la manera habitual, pero sí una *solución de mínimos cuadrados*, que asignó latitudes y longitudes a los puntos de referencia en la forma que mejor correspondía a las 1.8 millones de observaciones. En 1986 se encontró la solución de mínimos cuadrados al resolver un sistema relacionado de las llamadas *ecuaciones normales*, lo que implicó 928,735 ecuaciones con 928,735 variables.[1]

Más recientemente, el conocimiento de puntos de referencia en el suelo se ha vuelto crucial para determinar la ubicación exacta de satélites con el *Sistema de posicionamiento global* (*Global Positioning System*, GPS). Un satélite GPS calcula su posición en relación con la Tierra midiendo el tiempo que tardan las señales en llegar desde tres transmisores terrestres. Para hacer esto, los satélites utilizan relojes atómicos precisos que se han sincronizado con estaciones terrestres (cuyas ubicaciones se conocen con exactitud gracias al NAD).

El *Sistema de posicionamiento global* se utiliza para determinar las ubicaciones de los nuevos puntos de referencia terrestres y también para encontrar la ubicación del usuario sobre el suelo tomando como base los mapas ya existentes. Cuando el conductor de un automóvil (o un alpinista) enciende un receptor GPS, este último mide los tiempos de llegada de señales provenientes de al menos tres satélites. Esta información, junto con los datos transmitidos sobre las ubicaciones de los satélites y los tiempos del mensaje, se utiliza para ajustar el tiempo del receptor GPS y así determinar su ubicación aproximada sobre la Tierra. Con la información proveniente de un cuarto satélite, el receptor GPS puede incluso establecer su altura aproximada.

[1] Se presenta un análisis matemático de la estrategia de solución (junto con detalles de todo el proyecto NAD) en *North American Datum of 1983,* Charles R. Schwarz (ed.), National Geodetic Survey, National Oceanic and Atmospheric Administration (NOAA) Professional Paper NOS 2, 1989.

Los problemas del NAD y el GPS se resuelven encontrando un vector que "satisface aproximadamente" un sistema de ecuaciones inconsistente. Una cuidadosa explicación de esta aparente contradicción requerirá de las ideas desarrolladas en las primeras cinco secciones de este capítulo.

WEB

Para obtener una solución aproximada de un sistema de ecuaciones inconsistente que carece de solución, se necesita una idea bien definida de cercanía. La sección 6.1 introduce los conceptos de distancia y ortogonalidad en un espacio vectorial. Las secciones 6.2 y 6.3 muestran cómo se puede emplear la ortogonalidad para identificar aquel punto dentro de un subespacio W que es el más cercano al punto $\mathbf{y}$ externo a W. Considerando a W como el espacio columna de una matriz, la sección 6.5 desarrolla un método para obtener soluciones aproximadas ("mínimos cuadrados") de sistemas lineales inconsistentes, tal como el sistema resuelto para el informe NAD.

La sección 6.4 brinda otra oportunidad para ver proyecciones ortogonales en acción, creando una factorización matricial ampliamente utilizada en álgebra lineal numérica. Las secciones restantes examinan algunos de los muchos problemas de mínimos cuadrados que surgen en las aplicaciones, incluyendo aquellas en espacios más generales que $\mathbb{R}^n$.

6.1 PRODUCTO INTERIOR, LONGITUD Y ORTOGONALIDAD

Aquí se definen para $\mathbb{R}^n$ los conceptos geométricos de longitud, distancia y perpendicularidad, que son bien conocidos para $\mathbb{R}^2$ y $\mathbb{R}^3$. Esos conceptos ofrecen poderosas herramientas geométricas para resolver muchos problemas aplicados, incluyendo los problemas ya mencionados de mínimos cuadrados. Los tres conceptos se definen en términos del producto interior de dos vectores.

Producto interior

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$, entonces $\mathbf{u}$ y $\mathbf{v}$ se consideran matrices de $n \times 1$. La transpuesta $\mathbf{u}^T$ es una matriz de $1 \times n$, y el producto matricial $\mathbf{u}^T\mathbf{v}$ es una matriz de 1×1, que se representa como un solo número real (un escalar) sin corchetes. El número $\mathbf{u}^T\mathbf{v}$ se llama el **producto interior** de $\mathbf{u}$ y $\mathbf{v}$, y con frecuencia se representa como $\mathbf{u} \cdot \mathbf{v}$. Este producto interior, que se mencionó en los ejercicios de la sección 2.1, también se llama **producto punto**. 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 interior de $\mathbf{u}$ y $\mathbf{v}$ es

$$[u_1 \quad u_2 \quad \cdots \quad u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

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

SOLUCIÓN

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = [\,2 \quad -5 \quad -1\,]\begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v}\cdot\mathbf{u} = \mathbf{v}^T\mathbf{u} = [\,3 \quad 2 \quad -3\,]\begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = (3)(2) + (2)(-5) + (-3)(-1) = -1 \qquad \blacksquare$$

A partir de los cálculos en el ejemplo 1 resulta claro por qué $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$. La conmutatividad del producto interior es válida en general. Las siguientes propiedades del producto interior se deducen fácilmente de las características de la operación de transponer estudiada en la sección 2.1. (Véase los ejercicios 21 y 22 al final de esta sección).

TEOREMA 1

Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en $\mathbb{R}^n$, y c un escalar. Entonces,

a) $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$

b) $(\mathbf{u} + \mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{w}$

c) $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot(c\mathbf{v})$

d) $\mathbf{u}\cdot\mathbf{u} \geq 0$, y $\mathbf{u}\cdot\mathbf{u} = 0$ si y solo si $\mathbf{u} = \mathbf{0}$

Las propiedades *b*) y *c*) se pueden combinar varias veces para obtener la siguiente regla útil:

$$(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{w} = c_1(\mathbf{u}_1\cdot\mathbf{w}) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{w})$$

Longitud de un vector

Si $\mathbf{v}$ está en $\mathbb{R}^n$, con entradas $v_1,\ldots, v_n$, entonces la raíz cuadrada de $\mathbf{v}\cdot\mathbf{v}$ está definida, ya que $\mathbf{v}\cdot\mathbf{v}$ es no negativo.

DEFINICIÓN

La **longitud** (o **norma**) de $\mathbf{v}$ es el escalar no negativo $\|\mathbf{v}\|$ definido por

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \quad \text{y} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$

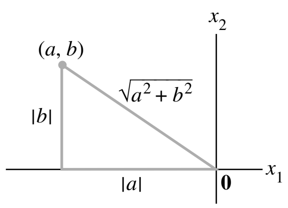


FIGURA 1
Interpretación de $\|\mathbf{v}\|$ como longitud.

Suponga que $\mathbf{v}$ está en $\mathbb{R}^2$, es decir, $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Como es usual, si se identifica a $\mathbf{v}$ con un punto geométrico en el plano, entonces $\|\mathbf{v}\|$ coincide con el concepto estándar de la longitud del segmento de recta que va del origen a $\mathbf{v}$. Esto se deduce a partir del teorema de Pitágoras aplicado a un triángulo como el de la figura 1.

Un cálculo similar con la diagonal de una caja rectangular demuestra que la definición de longitud de un vector $\mathbf{v}$ en $\mathbb{R}^3$ coincide con el concepto habitual de longitud.

Para cualquier escalar c, la longitud de $c\mathbf{v}$ es $|c|$ veces la longitud de $\mathbf{v}$. Es decir,

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$$

(Para ver esto, calcule $\|c\mathbf{v}\|^2 = (c\mathbf{v})\cdot(c\mathbf{v}) = c^2\mathbf{v}\cdot\mathbf{v} = c^2\|\mathbf{v}\|^2$ y obtenga raíces cuadradas).

Un **vector unitario** es un vector cuya longitud es 1. Si un vector **v** distinto de cero se *divide* entre su longitud —es decir, se multiplica por $1/\|\mathbf{v}\|$—, se obtiene un vector unitario **u** ya que la longitud de **u** es $(1/\|\mathbf{v}\|)\|\mathbf{v}\|$. El proceso de crear **u** a partir de **v** en ocasiones se llama **normalización de v**, y se dice que **u** está *en la misma dirección* que **v**.

Los ejemplos que se presentan a continuación emplean notación de vectores (columnas), para ahorrar espacio.

EJEMPLO 2 Sea $\mathbf{v} = (1, -2, 2, 0)$. Encuentre un vector unitario **u** en la misma dirección que **v**.

SOLUCIÓN Primero, calcule la longitud de **v**:

$$\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9$$
$$\|\mathbf{v}\| = \sqrt{9} = 3$$

Después, multiplique **v** por $1/\|\mathbf{v}\|$ para obtener

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{3}\mathbf{v} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

Para comprobar que $\|\mathbf{u}\| = 1$ es suficiente probar que $\|\mathbf{u}\|^2 = 1$.

$$\|\mathbf{u}\|^2 = \mathbf{u}\cdot\mathbf{u} = \left(\tfrac{1}{3}\right)^2 + \left(-\tfrac{2}{3}\right)^2 + \left(\tfrac{2}{3}\right)^2 + (0)^2$$
$$= \tfrac{1}{9} + \tfrac{4}{9} + \tfrac{4}{9} + 0 = 1$$

■

EJEMPLO 3 Sea W el subespacio de $\mathbb{R}^2$ generado por $\mathbf{x} = (\frac{2}{3}, 1)$. Obtenga un vector unitario **z** que sea una base para W.

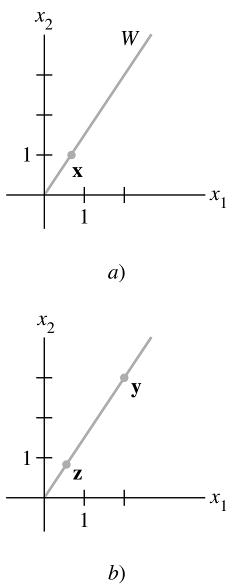


FIGURA 2
Normalización de un vector para obtener un vector unitario.

SOLUCIÓN W consiste en todos los múltiplos de **x**, como el de la figura 2*a*). Cualquier vector distinto de cero en W es una base para W. Para simplificar los cálculos, "escale" **x** para eliminar fracciones. Es decir, multiplique **x** por 3 para obtener

$$\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Ahora calcule $\|\mathbf{y}\|^2 = 2^2 + 3^2 = 13$, $\|\mathbf{y}\| = \sqrt{13}$, y normalice **y** para obtener

$$\mathbf{z} = \frac{1}{\sqrt{13}}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

Véase la figura 2*b*). Otro vector unitario es $(-2/\sqrt{13}, -3/\sqrt{13})$. ■

Distancia en $\mathbb{R}^n$

Ahora ya estamos listos para describir qué tan cerca está un vector de otro. Recuerde que si a y b son números reales, la distancia sobre la recta numérica entre a y b es el número $|a - b|$. En la figura 3 se ilustran dos ejemplos. Esta definición de distancia en $\mathbb{R}$ tiene una analogía directa en $\mathbb{R}^n$.

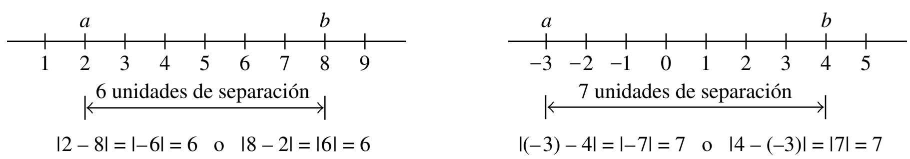


FIGURA 3 Distancias en $\mathbb{R}$.

DEFINICIÓN

Para $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, la **distancia entre u y v**, que se escribe como dist($\mathbf{u}$, $\mathbf{v}$), es la longitud del vector $\mathbf{u} - \mathbf{v}$. Es decir,

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

En $\mathbb{R}^2$ y $\mathbb{R}^3$, esta definición de distancia coincide con las fórmulas usuales de la distancia euclidiana entre dos puntos, como muestran los siguientes dos ejemplos.

EJEMPLO 4 Determine la distancia entre los vectores $\mathbf{u} = (7, 1)$ y $\mathbf{v} = (3, 2)$.

SOLUCIÓN Calcule

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

La figura 4 muestra los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} - \mathbf{v}$. Cuando el vector $\mathbf{u} - \mathbf{v}$ se suma a $\mathbf{v}$, el resultado es $\mathbf{u}$. Observe que en la figura 4 el paralelogramo revela que la distancia de $\mathbf{u}$ a $\mathbf{v}$ es igual a la distancia de $\mathbf{u} - \mathbf{v}$ a $\mathbf{0}$. ■

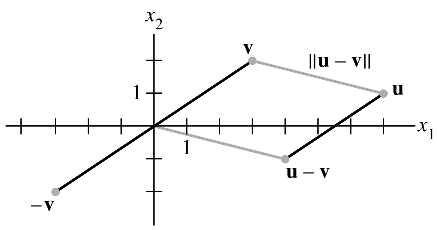


FIGURA 4 La distancia entre $\mathbf{u}$ y $\mathbf{v}$ es la longitud de $\mathbf{u} - \mathbf{v}$.

EJEMPLO 5 Si $\mathbf{u} = (u_1, u_2, u_3)$ y $\mathbf{v} = (v_1, v_2, v_3)$, entonces

$$\begin{aligned} \text{dist}(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})} \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2} \end{aligned}$$

■

Vectores ortogonales

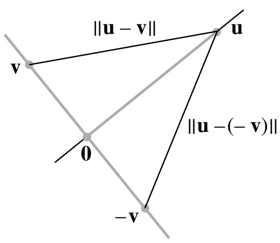


FIGURA 5

El resto de este capítulo depende del hecho de que el concepto de rectas perpendiculares de geometría euclidiana tiene un análogo en $\mathbb{R}^n$.

Considere $\mathbb{R}^2$ o $\mathbb{R}^3$ y dos rectas que pasan por el origen determinado por los vectores $\mathbf{u}$ y $\mathbf{v}$. Las dos rectas que se muestran en la figura 5 son geométricamente perpendiculares si y solo si la distancia de $\mathbf{u}$ a $\mathbf{v}$ es la misma que la distancia de $\mathbf{u}$ a $-\mathbf{v}$. Esto es lo mismo que requerir que los cuadrados de las distancias sean iguales. Ahora

$$\begin{aligned} [\,\text{dist}(\mathbf{u}, -\mathbf{v})\,]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\ &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) && \text{Teorema 1(b)} \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} && \text{Teorema 1(a), (b)} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v} && \text{Teorema 1(a)} \end{aligned} \tag{1}$$

Los mismos cálculos con $\mathbf{v}$ y $-\mathbf{v}$ intercambiados indican que

$$\begin{aligned}[\operatorname{dist}(\mathbf{u}, \mathbf{v})]^2 &= \|\mathbf{u}\|^2 + \|-\mathbf{v}\|^2 + 2\mathbf{u}\cdot(-\mathbf{v}) \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v}\end{aligned}$$

Las dos distancias al cuadrado son iguales si y solo si $2\mathbf{u}\cdot\mathbf{v} = -2\mathbf{u}\cdot\mathbf{v}$, lo que ocurre si y solo si $\mathbf{u}\cdot\mathbf{v} = 0$.

Este cálculo demuestra que cuando los vectores $\mathbf{u}$ y $\mathbf{v}$ se identifican con puntos geométricos, las rectas correspondientes que pasan por los puntos y por el origen son perpendiculares si y solo si $\mathbf{u}\cdot\mathbf{v} = 0$. La siguiente definición generaliza a $\mathbb{R}^n$ este concepto de perpendicularidad (u *ortogonalidad*, como se le llama con frecuencia en álgebra lineal).

DEFINICIÓN

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

Observe que el vector cero es ortogonal a todo vector en $\mathbb{R}^n$ porque $\mathbf{0}^T\mathbf{v} = 0$ para toda $\mathbf{v}$.

El siguiente teorema expone un dato útil en relación con los vectores ortogonales. La demostración se deduce inmediatamente de los cálculos en la ecuación (1) y de la definición de ortogonalidad. El triángulo rectángulo que se ilustra en la figura 6 permite visualizar las longitudes que aparecen en el teorema.

TEOREMA 2

Teorema de Pitágoras

Dos vectores $\mathbf{u}$ y $\mathbf{v}$ son ortogonales si y solo si $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

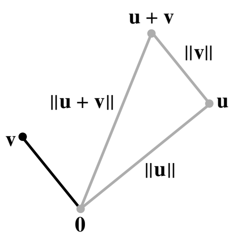


FIGURA 6

Complementos ortogonales

Para practicar el uso de productos interiores, aquí se introduce un concepto que será útil en la sección 6.3 y en todo el capítulo. Si un vector $\mathbf{z}$ es ortogonal a todo vector en un subespacio W de $\mathbb{R}^n$, entonces se dice que $\mathbf{z}$ es **ortogonal a** W. El conjunto de todos los vectores $\mathbf{z}$ que son ortogonales a W se llama **complemento ortogonal** de W y se denota con $W^\perp$ (que se lee "perpendicular a W").

EJEMPLO 6 Sean W un plano a través del origen en $\mathbb{R}^3$, y L la recta que pasa por el origen y es perpendicular a W. Si $\mathbf{z}$ y $\mathbf{w}$ son distintos de cero, $\mathbf{z}$ está sobre L y $\mathbf{w}$ está en W, entonces el segmento de recta de $\mathbf{0}$ a $\mathbf{z}$ es perpendicular al segmento de recta de $\mathbf{0}$ a $\mathbf{w}$; es decir, $\mathbf{z}\cdot\mathbf{w} = 0$. Véase la figura 7. De manera que todo vector sobre L es ortogonal a cada $\mathbf{w}$ en W. En efecto, L consiste en *todos* los vectores que son ortogonales a todos los vectores $\mathbf{w}$ en W, y W consiste en todos los vectores ortogonales a todos los vectores $\mathbf{z}$ en L. Es decir,

$$L = W^\perp \quad \text{y} \quad W = L^\perp$$ ■

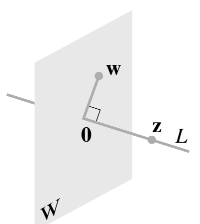


FIGURA 7
Un plano y una recta que pasan por $\mathbf{0}$ como complementos ortogonales.

Los siguientes dos hechos sobre $W^\perp$, con W como un subespacio de $\mathbb{R}^n$, se necesitarán más adelante en este capítulo. Las demostraciones se sugieren en los ejercicios 29 y 30. Los ejercicios 27 a 31 ofrecen una excelente oportunidad de practicar con las propiedades del producto interior.

1. Un vector $\mathbf{x}$ está en $W^\perp$ si y solo si $\mathbf{x}$ es ortogonal a cada vector de cualquier conjunto que genera a W.
2. $W^\perp$ es un subespacio de $\mathbb{R}^n$.

El siguiente teorema y el ejercicio 31 comprueban las afirmaciones hechas en la sección 4.6 en relación con los subespacios que se muestran en la figura 8. (Véase también el ejercicio 28 de la sección 4.6).

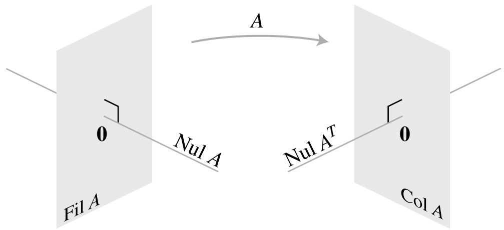


FIGURA 8 Los subespacios fundamentales determinados por una matriz A de $m \times n$.

TEOREMA 3

Sea A una matriz de $m \times n$. El complemento ortogonal del espacio fila de A es el espacio nulo de A, y el complemento ortogonal del espacio columna de A es el espacio nulo de A^T:

$$(\text{Fil } A)^{\perp} = \text{Nul } A \quad \text{y} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T$$

DEMOSTRACIÓN La regla fila-columna para calcular $A\mathbf{x}$ indica que si $\mathbf{x}$ está en Nul A, entonces $\mathbf{x}$ es ortogonal a cada fila de A (con las filas consideradas como vectores en $\mathbb{R}^n$). Como las filas de A generan el espacio fila, entonces $\mathbf{x}$ es ortogonal a Fil A. Y a la inversa, si $\mathbf{x}$ es ortogonal a Fil A, entonces $\mathbf{x}$, evidentemente, es ortogonal a cada fila de A, y por lo tanto, $A\mathbf{x} = \mathbf{0}$. Esto demuestra el primer enunciado del teorema. Ya que este enunciado es verdadero para cualquier matriz, es válido para A^T. Es decir, el complemento ortogonal del espacio fila de A^T es el espacio nulo de A^T. Esto demuestra la segunda afirmación porque Fil A^T = Col A. ■

Ángulos en $\mathbb{R}^2$ y $\mathbb{R}^3$ (opcional)

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores distintos de cero en $\mathbb{R}^2$ o $\mathbb{R}^3$, entonces existe una agradable conexión entre su producto interior y el ángulo ϑ entre los dos segmentos de recta que van del origen a los puntos identificados con $\mathbf{u}$ y $\mathbf{v}$. La fórmula es

$$\boxed{\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos \vartheta} \tag{2}$$

Si se desea comprobar esta fórmula para vectores en $\mathbb{R}^2$, considere el triángulo que se ilustra en la figura 9, con lados de longitudes $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ y $\|\mathbf{u} - \mathbf{v}\|$. De acuerdo con la ley de los cosenos,

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\| \cos \vartheta$$

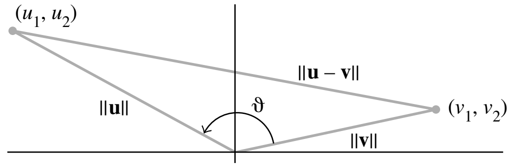


FIGURA 9 El ángulo entre dos vectores.

que se puede reordenar para obtener

$$\begin{aligned}\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta &= \frac{1}{2}\left[\|\mathbf{u}\|^2+\|\mathbf{v}\|^2-\|\mathbf{u}-\mathbf{v}\|^2\right]\\ &= \frac{1}{2}\left[u_1^2+u_2^2+v_1^2+v_2^2-(u_1-v_1)^2-(u_2-v_2)^2\right]\\ &= u_1v_1+u_2v_2\\ &= \mathbf{u}\cdot\mathbf{v}\end{aligned}$$

La comprobación es similar para $\mathbb{R}^3$. Cuando $n > 3$, la fórmula (2) se puede utilizar para *definir* el ángulo entre dos vectores en $\mathbb{R}^n$. En estadística, por ejemplo, el valor de $\cos\vartheta$ definido por la ecuación (2) para vectores convenientes $\mathbf{u}$ y $\mathbf{v}$ es lo que los especialistas en estadística llaman *coeficiente de correlación*.

PROBLEMAS DE PRÁCTICA

Sean $\mathbf{a}=\begin{bmatrix}-2\\1\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}-3\\1\end{bmatrix}$, $\mathbf{c}=\begin{bmatrix}4/3\\-1\\2/3\end{bmatrix}$ y $\mathbf{d}=\begin{bmatrix}5\\6\\-1\end{bmatrix}$.

1. Calcule $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}$ y $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a}$.
2. Encuentre un vector unitario $\mathbf{u}$ en la dirección de $\mathbf{c}$.
3. Demuestre que $\mathbf{d}$ es ortogonal a $\mathbf{c}$.
4. Utilice los resultados de los problemas de práctica 2 y 3 para explicar por qué $\mathbf{d}$ debe ser ortogonal al vector unitario $\mathbf{u}$.

6.1 EJERCICIOS

En los ejercicios 1 a 8, determine las cantidades indicadas utilizando los vectores

$$\mathbf{u}=\begin{bmatrix}-1\\2\end{bmatrix},\quad \mathbf{v}=\begin{bmatrix}4\\6\end{bmatrix},\quad \mathbf{w}=\begin{bmatrix}3\\-1\\-5\end{bmatrix},\quad \mathbf{x}=\begin{bmatrix}6\\-2\\3\end{bmatrix}$$

1. $\mathbf{u}\cdot\mathbf{u}$, $\mathbf{v}\cdot\mathbf{u}$ y $\dfrac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$

2. $\mathbf{w}\cdot\mathbf{w}$, $\mathbf{x}\cdot\mathbf{w}$ y $\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}$

3. $\dfrac{1}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}$

4. $\dfrac{1}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$

5. $\left(\dfrac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$

6. $\left(\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{x}\cdot\mathbf{x}}\right)\mathbf{x}$

7. $\|\mathbf{w}\|$

8. $\|\mathbf{x}\|$

En los ejercicios 9 a 12, obtenga un vector unitario en la dirección del vector indicado.

9. $\begin{bmatrix}-30\\40\end{bmatrix}$

10. $\begin{bmatrix}-6\\4\\-3\end{bmatrix}$

11. $\begin{bmatrix}7/4\\1/2\\1\end{bmatrix}$

12. $\begin{bmatrix}8/3\\2\end{bmatrix}$

13. Encuentre la distancia entre $\mathbf{x}=\begin{bmatrix}10\\-3\end{bmatrix}$ y $\mathbf{y}=\begin{bmatrix}-1\\-5\end{bmatrix}$.

14. Calcule la distancia entre $\mathbf{u}=\begin{bmatrix}0\\-5\\2\end{bmatrix}$ y $\mathbf{z}=\begin{bmatrix}-4\\-1\\8\end{bmatrix}$.

En los ejercicios 15 a 18, determine qué pares de vectores son ortogonales.

15. $\mathbf{a}=\begin{bmatrix}8\\-5\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}-2\\-3\end{bmatrix}$

16. $\mathbf{u}=\begin{bmatrix}12\\3\\-5\end{bmatrix}$, $\mathbf{v}=\begin{bmatrix}2\\-3\\3\end{bmatrix}$

17. $\mathbf{u}=\begin{bmatrix}3\\2\\-5\\0\end{bmatrix}$, $\mathbf{v}=\begin{bmatrix}-4\\1\\-2\\6\end{bmatrix}$

18. $\mathbf{y}=\begin{bmatrix}-3\\7\\4\\0\end{bmatrix}$, $\mathbf{z}=\begin{bmatrix}1\\-8\\15\\-7\end{bmatrix}$

En los ejercicios 19 y 20, todos los vectores están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

19. *a*) $\mathbf{v}\cdot\mathbf{v}=\|\mathbf{v}\|^2$.

b) Para cualquier escalar c, $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$.

c) Si la distancia de $\mathbf{u}$ a $\mathbf{v}$ es igual a la distancia de $\mathbf{u}$ a $-\mathbf{v}$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

d) Para una matriz cuadrada A, los vectores en Col A son ortogonales a los vectores en Nul A.

e) Si los vectores $\mathbf{v}_1,\ldots,\mathbf{v}_p$ generan un subespacio W y si $\mathbf{x}$ es ortogonal a cada $\mathbf{v}_j$ para $j = 1,\ldots, p$, entonces $\mathbf{x}$ está en $W^{\perp}$.

20. *a*) $\mathbf{u}\cdot\mathbf{v}$ -$\mathbf{v}\cdot\mathbf{u} = 0$.

b) Para cualquier escalar c, $\|c\mathbf{v}\| = c\|\mathbf{v}\|$.

c) Si $\mathbf{x}$ es ortogonal a cada vector en un subespacio W, entonces $\mathbf{x}$ está en $W^{\perp}$.

d) Si $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u} + \mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

e) Si A es una matriz de $m \times n$, los vectores en el espacio nulo de A son ortogonales a los vectores en el espacio fila de A.

21. Utilice la definición de transpuesta del producto interior para comprobar los incisos *b*) y *c*) del teorema 1. Mencione los hechos pertinentes del capítulo 2.

22. Sea $\mathbf{u} = (u_1, u_2, u_3)$. Explique por qué $\mathbf{u}\cdot\mathbf{u} \geq 0$. ¿Cuándo $\mathbf{u}\cdot\mathbf{u} = 0$?

23. Sean $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -7 \\ -4 \\ 6 \end{bmatrix}$. Calcule y compare $\mathbf{u}\cdot\mathbf{v}$, $\|\mathbf{u}\|^2$, $\|\mathbf{v}\|^2$ y $\|\mathbf{u} + \mathbf{v}\|^2$. No utilice el teorema de Pitágoras.

24. Compruebe la *ley del paralelogramo* para vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$:

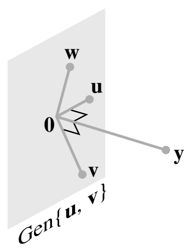
$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

25. Sea $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Describa el conjunto H de vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ que son ortogonales a $\mathbf{v}$. [*Sugerencia:* Considere $\mathbf{v} = \mathbf{0}$ y $\mathbf{v} \neq \mathbf{0}$].

26. Sean $\mathbf{u} = \begin{bmatrix} 5 \\ -6 \\ 7 \end{bmatrix}$, y W el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^3$ tales que $\mathbf{u}\cdot\mathbf{x} = 0$. ¿Qué teorema del capítulo 4 se puede utilizar para demostrar que W es un subespacio de $\mathbb{R}^3$? Describa a W en lenguaje geométrico.

27. Suponga que un vector $\mathbf{y}$ que es ortogonal a los vectores $\mathbf{u}$ y $\mathbf{v}$. Demuestre que $\mathbf{y}$ es ortogonal al vector $\mathbf{u} + \mathbf{v}$.

28. Suponga que $\mathbf{y}$ es ortogonal a $\mathbf{u}$ y $\mathbf{v}$. Demuestre que $\mathbf{y}$ es ortogonal a cada $\mathbf{w}$ en Gen $\{\mathbf{u}, \mathbf{v}\}$. [*Sugerencia:* Una $\mathbf{w}$ arbitraria en Gen $\{\mathbf{u}, \mathbf{v}\}$ tiene la forma $\mathbf{w} = c_1\mathbf{u} + c_2\mathbf{v}$. Demuestre que $\mathbf{y}$ es ortogonal a dicho vector $\mathbf{w}$].



29. Sea $W = \text{Gen}\ \{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$. Demuestre que si $\mathbf{x}$ es ortogonal a cada $\mathbf{v}_j$, para $1 \leq j \leq p$, entonces $\mathbf{x}$ es ortogonal a todo vector en W.

30. Sean W un subespacio de $\mathbb{R}^n$, y $W^{\perp}$ el conjunto de todos los vectores ortogonales a W. Demuestre que $W^{\perp}$ es un subespacio de $\mathbb{R}^n$, considerando los siguientes pasos.

a) Tome $\mathbf{z}$ en $W^{\perp}$, y sea $\mathbf{u}$ cualquier elemento de W. Entonces $\mathbf{z}\cdot\mathbf{u} = 0$. Tome cualquier escalar c y demuestre que $c\mathbf{z}$ es ortogonal a $\mathbf{u}$. (Puesto que $\mathbf{u}$ era un elemento arbitrario de W, esto demostrará que $c\mathbf{z}$ está en $W^{\perp}$).

b) Tome $\mathbf{z}_1$ y $\mathbf{z}_2$ en $W^{\perp}$, y sea $\mathbf{u}$ cualquier elemento de W. Demuestre que $\mathbf{z}_1 + \mathbf{z}_2$ es ortogonal a $\mathbf{u}$. ¿Qué se puede concluir acerca de $\mathbf{z}_1 + \mathbf{z}_2$? ¿Por qué?

c) Termine la demostración de que $W^{\perp}$ es un subespacio de $\mathbb{R}^n$.

31. Demuestre que si $\mathbf{x}$ está en W y $W^{\perp}$, entonces $\mathbf{x} = \mathbf{0}$.

32. **[M]** Construya un par $\mathbf{u}$, $\mathbf{v}$ de vectores aleatorios en $\mathbb{R}^4$, y sea

$$A = \begin{bmatrix} .5 & .5 & .5 & .5 \\ .5 & .5 & -.5 & -.5 \\ .5 & -.5 & .5 & -.5 \\ .5 & -.5 & -.5 & .5 \end{bmatrix}$$

a) Denote las columnas de A como $\mathbf{a}_1,\ldots, \mathbf{a}_4$. Obtenga la longitud de cada columna, y calcule $\mathbf{a}_1\cdot\mathbf{a}_2$, $\mathbf{a}_1\cdot\mathbf{a}_3$, $\mathbf{a}_1\cdot\mathbf{a}_4$, $\mathbf{a}_2\cdot\mathbf{a}_3$, $\mathbf{a}_2\cdot\mathbf{a}_4$ y $\mathbf{a}_3\cdot\mathbf{a}_4$.

b) Calcule y compare las longitudes de $\mathbf{u}$, $A\mathbf{u}$, $\mathbf{v}$ y $A\mathbf{v}$.

c) Utilice la ecuación (2) de esta sección para calcular el coseno del ángulo entre $\mathbf{u}$ y $\mathbf{v}$. Compare esto con el coseno del ángulo entre $A\mathbf{u}$ y $A\mathbf{v}$.

d) Repita los incisos *b*) y *c*) para otros dos pares de vectores aleatorios. ¿Qué se puede suponer del efecto de A sobre los vectores?

33. **[M]** Genere vectores aleatorios $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{v}$ en $\mathbb{R}^4$ con entradas enteras (y $\mathbf{v} \neq \mathbf{0}$), y calcule las cantidades

$$\left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \left(\frac{\mathbf{y}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \frac{(\mathbf{x}+\mathbf{y})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v},\ \frac{(10\mathbf{x})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}$$

Repita los cálculos con nuevos vectores aleatorios $\mathbf{x}$ y $\mathbf{y}$. ¿Qué se puede suponer acerca del mapeo $\mathbf{x} \mapsto T(\mathbf{x}) = \left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$ (para $\mathbf{v} \neq \mathbf{0}$)? Compruebe su suposición algebraicamente.

34. **[M]** Sea $A = \begin{bmatrix} -6 & 3 & -27 & -33 & -13 \\ 6 & -5 & 25 & 28 & 14 \\ 8 & -6 & 34 & 38 & 18 \\ 12 & -10 & 50 & 41 & 23 \\ 14 & -21 & 49 & 29 & 33 \end{bmatrix}$. Construya una matriz N cuyas columnas formen una base para Nul A, y elabore una matriz R cuyas *filas* formen una base para Fil A (para detalles, véase la sección 4.6). Efectúe un cálculo matricial con N y R que muestre un hecho del teorema 3.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $\mathbf{a}\cdot\mathbf{b} = 7$, $\mathbf{a}\cdot\mathbf{a} = 5$. Así, $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}} = \dfrac{7}{5}$ y $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a} = \dfrac{7}{5}\mathbf{a} = \begin{bmatrix} -14/5 \\ 7/5 \end{bmatrix}$.

2. Escale c, multiplicando por 3 para obtener $\mathbf{y} = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$. Calcule $\|\mathbf{y}\|^2 = 29$ y $\|\mathbf{y}\| = \sqrt{29}$.

   El vector unitario en la dirección de $\mathbf{c}$ y $\mathbf{y}$ es $\mathbf{u} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y} = \begin{bmatrix} 4/\sqrt{29} \\ -3/\sqrt{29} \\ 2/\sqrt{29} \end{bmatrix}$.

3. $\mathbf{d}$ es ortogonal a $\mathbf{c}$, porque

$$\mathbf{d}\cdot\mathbf{c} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} = \frac{20}{3} - 6 - \frac{2}{3} = 0$$

4. $\mathbf{d}$ es ortogonal a $\mathbf{u}$ porque $\mathbf{u}$ tiene la forma $k\mathbf{c}$ para alguna k, y

$$\mathbf{d}\cdot\mathbf{u} = \mathbf{d}\cdot(k\mathbf{c}) = k(\mathbf{d}\cdot\mathbf{c}) = k(0) = 0$$

6.2 CONJUNTOS ORTOGONALES

Se dice que un conjunto de vectores $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ en $\mathbb{R}^n$ es un **conjunto ortogonal** si cada par de distintos vectores del conjunto es ortogonal, es decir, si $\mathbf{u}_i\cdot\mathbf{u}_j = 0$ siempre que $i \neq j$.

EJEMPLO 1 Demuestre que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal, donde

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

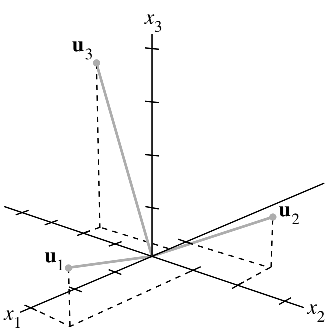


FIGURA 1

SOLUCIÓN Considere los tres posibles pares de vectores distintos, a saber, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$ y $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\begin{aligned}
\mathbf{u}_1\cdot\mathbf{u}_2 &= 3(-1) + 1(2) + 1(1) = 0 \\
\mathbf{u}_1\cdot\mathbf{u}_3 &= 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0 \\
\mathbf{u}_2\cdot\mathbf{u}_3 &= -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0
\end{aligned}$$

Cada par de vectores diferentes es ortogonal, y así $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal. Véase la figura 1; ahí los tres segmentos de recta son mutuamente perpendiculares. ■

TEOREMA 4

Si $S = \{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ es un conjunto ortogonal de vectores distintos de cero en $\mathbb{R}^n$, entonces S es linealmente independiente y, por lo tanto, es una base para el subespacio generado por S.

DEMOSTRACIÓN Si $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ para algunos escalares $c_1,\ldots,c_p$, entonces

$$\begin{aligned}
0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\
&= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\
&= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\
&= c_1(\mathbf{u}_1\cdot\mathbf{u}_1)
\end{aligned}$$

ya que $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_2,\ldots,\mathbf{u}_p$. Como $\mathbf{u}_1$ es diferente de cero, entonces $\mathbf{u}_1\cdot\mathbf{u}_1$ no es cero y así $c_1 = 0$. De manera similar, $c_2,\ldots,c_p$ deben ser cero. Por lo tanto, S es linealmente independiente. ■

DEFINICIÓN

Una **base ortogonal** para un subespacio W de $\mathbb{R}^n$ es una base para W que también es un conjunto ortogonal.

El siguiente teorema sugiere por qué una base ortogonal es mucho más agradable que otras bases. Los pesos en una combinación lineal se pueden calcular fácilmente.

TEOREMA 5

Sea $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ una base ortogonal para un subespacio W de $\mathbb{R}^n$. Para cada $\mathbf{y}$ en W, los pesos en la combinación lineal

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

están dados por

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1,\ldots,p)$$

DEMOSTRACIÓN Como en la demostración anterior, la ortogonalidad de $\{\mathbf{u}_1,\ldots,\ \mathbf{u}_p\}$ demuestra que

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

Puesto que $\mathbf{u}_1\cdot\mathbf{u}_1$ no es cero, en la ecuación anterior se puede despejar c_1. Para encontrar c_j, considerando $j = 2,\ldots,p$, calcule $\mathbf{y}\cdot\mathbf{u}_j$ y despeje c_j. ■

EJEMPLO 2 El conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ del ejemplo 1 es una base ortogonal para $\mathbb{R}^3$. Exprese el vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ como una combinación lineal de los vectores en S.

SOLUCIÓN Calcule

$$\begin{aligned} \mathbf{y}\cdot\mathbf{u}_1 &= 11, & \mathbf{y}\cdot\mathbf{u}_2 &= -12, & \mathbf{y}\cdot\mathbf{u}_3 &= -33 \\ \mathbf{u}_1\cdot\mathbf{u}_1 &= 11, & \mathbf{u}_2\cdot\mathbf{u}_2 &= 6, & \mathbf{u}_3\cdot\mathbf{u}_3 &= 33/2 \end{aligned}$$

Por el teorema 5,

$$\begin{aligned} \mathbf{y} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 \\ &= \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\ &= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3 \end{aligned}$$

■

Observe qué fácil es obtener los pesos necesarios para construir $\mathbf{y}$ a partir de una base ortogonal. Si la base no fuera ortogonal, se necesitaría resolver un sistema de ecuaciones lineales para encontrar los pesos, como en el capítulo 1.

A continuación se presenta una construcción que será un paso clave en muchos cálculos que implican ortogonalidad, y que conducirá a una interpretación geométrica del teorema 5.

Una proyección ortogonal

Dado un vector $\mathbf{u}$ distinto de cero en $\mathbb{R}^n$, considere el problema de descomponer un vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, siendo uno un múltiplo de $\mathbf{u}$ y el otro ortogonal a $\mathbf{u}$. Se desea escribir

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

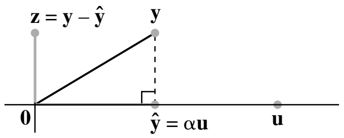


FIGURA 2
Determinación de α para hacer que $\mathbf{y} - \hat{\mathbf{y}}$ sea ortogonal a $\mathbf{u}$.

donde $\hat{\mathbf{y}} = \alpha\mathbf{u}$ para algún escalar α, y $\mathbf{z}$ es algún vector ortogonal a $\mathbf{u}$. Véase la figura 2. Dado cualquier escalar α, sea $\mathbf{z} = \mathbf{y} - \alpha\mathbf{u}$, tal que la ecuación (1) se satisface. Entonces $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a $\mathbf{u}$ si y solo si

$$0 = (\mathbf{y} - \alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - (\alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - \alpha(\mathbf{u}\cdot\mathbf{u})$$

Es decir, la ecuación (1) se satisface con $\mathbf{z}$ ortogonal a $\mathbf{u}$ si y solo si $\alpha = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$ y $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$. El vector $\hat{\mathbf{y}}$ es la **proyección ortogonal de y sobre u**, y el vector $\mathbf{z}$ es la **componente de y ortogonal a u**.

Si c es cualquier escalar distinto de cero y si $\mathbf{u}$ se remplaza por $c\mathbf{u}$ en la definición de $\hat{\mathbf{y}}$, entonces la proyección ortogonal de $\mathbf{y}$ sobre $c\mathbf{u}$ coincide exactamente con la proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$ (ejercicio 31). De ahí que esta proyección esté determinada por el *subespacio* L generado por $\mathbf{u}$ (la recta que pasa por $\mathbf{u}$ y $\mathbf{0}$). Algunas veces $\hat{\mathbf{y}}$ se denota como $\text{proy}_L\,\mathbf{y}$, y se llama la **proyección ortogonal de y sobre** L. Es decir,

$$\boxed{\hat{\mathbf{y}} = \text{proj}_L\,\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}} \tag{2}$$

EJEMPLO 3 Sean $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Encuentre la proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$. Después escriba $\mathbf{y}$ como la suma de dos vectores ortogonales, uno en Gen $\{\mathbf{u}\}$ y el otro ortogonal a $\mathbf{u}$.

SOLUCIÓN Calcule

$$\mathbf{y}\cdot\mathbf{u} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$\mathbf{u}\cdot\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$ es

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

y la componente de $\mathbf{y}$ ortogonal a $\mathbf{u}$ es

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

La suma de esos dos vectores es $\mathbf{y}$. Es decir,

$$\underset{\mathbf{y}}{\underset{\uparrow}{\begin{bmatrix} 7 \\ 6 \end{bmatrix}}} = \underset{\hat{\mathbf{y}}}{\underset{\uparrow}{\begin{bmatrix} 8 \\ 4 \end{bmatrix}}} + \underset{(\mathbf{y}-\hat{\mathbf{y}})}{\underset{\uparrow}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}}$$

En la figura 3 se muestra esta descomposición de $\mathbf{y}$. *Nota:* Si son correctos los cálculos anteriores, entonces $\{\hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}\}$ será un conjunto ortogonal. Para comprobarlo, determine

$$\hat{\mathbf{y}}\cdot(\mathbf{y} - \hat{\mathbf{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$ ■

Como el segmento de recta de la figura 3 entre $\mathbf{y}$ y $\hat{\mathbf{y}}$ es perpendicular a L, por construcción de $\hat{\mathbf{y}}$, entonces el punto identificado con $\hat{\mathbf{y}}$ es el punto más cercano de L a $\mathbf{y}$. (Esto se puede demostrar con geometría. Ahora se supondrá esto para $\mathbb{R}^2$ y en la sección 6.3 se demostrará para $\mathbb{R}^n$).

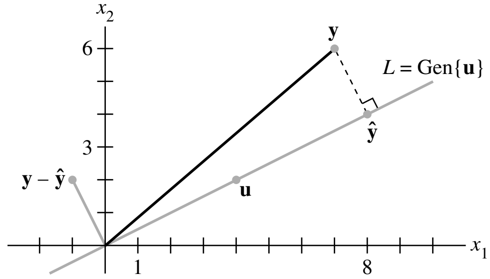


FIGURA 3 La proyección ortogonal de $\mathbf{y}$ sobre una recta L que pasa por el origen.

EJEMPLO 4 En la figura 3 encuentre la distancia de $\mathbf{y}$ a L.

SOLUCIÓN La distancia de $\mathbf{y}$ a L es la longitud del segmento de recta perpendicular de $\mathbf{y}$ a la proyección ortogonal $\hat{\mathbf{y}}$. Esta longitud es igual a la longitud de $\mathbf{y} - \hat{\mathbf{y}}$. Así, la distancia es

$$\|\mathbf{y} - \hat{\mathbf{y}}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5} \qquad \blacksquare$$

Una interpretación geométrica del teorema 5

En la ecuación (2), la fórmula para la proyección ortogonal $\hat{\mathbf{y}}$ tiene la misma apariencia que cada uno de los términos en el teorema 5. Así, el teorema 5 descompone un vector $\mathbf{y}$ en una suma de proyecciones ortogonales sobre subespacios unidimensionales.

Es fácil visualizar el caso en el cual $W = \mathbb{R}^2 = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, con $\mathbf{u}_1$ y $\mathbf{u}_2$ ortogonales. Cualquier $\mathbf{y}$ en $\mathbb{R}^2$ se puede escribir en la forma

$$\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 \tag{3}$$

El primer término en (3) es la proyección de $\mathbf{y}$ sobre el subespacio generado por $\mathbf{u}_1$ (la recta que pasa por $\mathbf{u}_1$ y por el origen), y el segundo término es la proyección de $\mathbf{y}$ sobre el subespacio generado por $\mathbf{u}_2$. Así, la ecuación (3) expresa a $\mathbf{y}$ como la suma de sus proyecciones sobre los ejes (ortogonales) determinados por $\mathbf{u}_1$ y $\mathbf{u}_2$. Véase la figura 4.

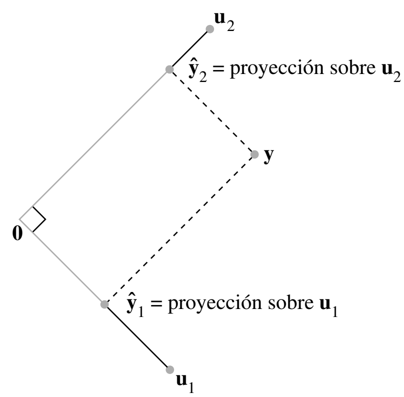


FIGURA 4 Un vector descompuesto en la suma de dos proyecciones.

El teorema 5 descompone a cada $\mathbf{y}$ en $\text{Gen}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ en la suma de p proyecciones sobre subespacios unidimensionales que son mutuamente ortogonales.

Descomposición de una fuerza en sus componentes

La descomposición que se ilustra en la figura 4 se presenta en física cuando se aplica una fuerza a un objeto. La elección de un sistema de coordenadas adecuado permite que la fuerza se represente con un vector **y** en $\mathbb{R}^2$ o $\mathbb{R}^3$. Con frecuencia el problema implica alguna dirección particular de interés, que queda representada por otro vector **u**. Por ejemplo, si el objeto se está moviendo en línea recta cuando se aplica la fuerza, entonces el vector **u** podría apuntar en la dirección de movimiento, como se muestra en la figura 5. Un paso clave en el problema es descomponer la fuerza en una componente en la dirección de **u** y en una componente ortogonal a **u**. Los cálculos serían análogos a los realizados en el ejemplo 3 anterior.

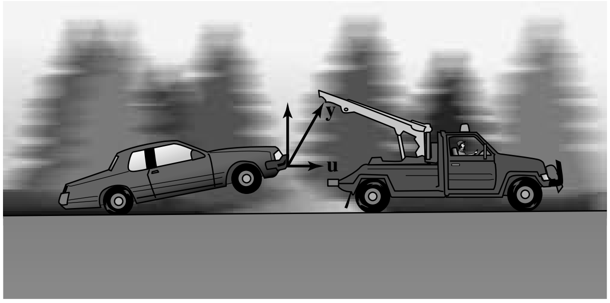


FIGURA 5

Conjuntos ortonormales

Un conjunto $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es un **conjunto ortonormal** si es un conjunto ortogonal de vectores unitarios. Si W es el subespacio generado por tal conjunto, entonces $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una **base ortonormal** para W, porque el conjunto es linealmente independiente de manera automática, de acuerdo con al teorema 4.

El ejemplo más sencillo de un conjunto ortonormal es la base estándar $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ para $\mathbb{R}^n$. Cualquier subconjunto no vacío de $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ también es ortonormal. A continuación se presenta un ejemplo más complicado.

EJEMPLO 5 Demuestre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortonormal de $\mathbb{R}^3$, donde

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

SOLUCIÓN Calcule

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_2 &= -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0 \\ \mathbf{v}_1 \cdot \mathbf{v}_3 &= -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_3 &= 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0 \end{aligned}$$

Así $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal. Además,

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_1 &= 9/11 + 1/11 + 1/11 = 1 \\ \mathbf{v}_2 \cdot \mathbf{v}_2 &= 1/6 + 4/6 + 1/6 = 1 \\ \mathbf{v}_3 \cdot \mathbf{v}_3 &= 1/66 + 16/66 + 49/66 = 1 \end{aligned}$$

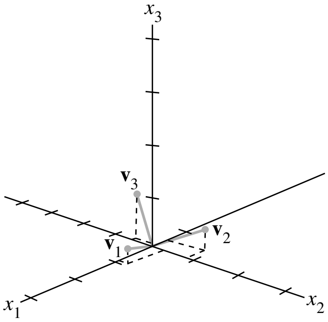


FIGURA 6

que muestra que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ son vectores unitarios. Así, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortonormal. Como el conjunto es linealmente independiente, entonces sus tres vectores forman una base para $\mathbb{R}^3$. Véase la figura 6. ■

Cuando los vectores en un conjunto ortogonal de vectores distintos de cero *se normalizan* para que tengan longitud unitaria, entonces los nuevos vectores seguirán siendo ortogonales, por lo que el nuevo conjunto será un conjunto ortonormal. Véase el ejercicio 32. En la figura 6 (ejemplo 5) es sencillo comprobar que los vectores son simplemente los vectores unitarios en las direcciones de los vectores de la figura 1 (ejemplo 1).

Las matrices cuyas columnas forman un conjunto ortonormal son importantes en aplicaciones y en algoritmos computacionales para cálculos matriciales. Sus propiedades fundamentales se exponen en los teoremas 6 y 7.

TEOREMA 6

Una matriz U de $m \times n$ tiene columnas ortonormales si y solo si $U^TU = I$.

DEMOSTRACIÓN Para simplificar la notación, suponga que U solo tiene tres columnas, cada una un vector en $\mathbb{R}^m$. En esencia, la demostración del caso general es la misma. Sea $U = [\mathbf{u}_1 \ \ \mathbf{u}_2 \ \ \mathbf{u}_3]$ y calcule

$$U^TU = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} \tag{4}$$

Las entradas en la matriz de la derecha son productos interiores, empleando la notación transpuesta. Las columnas de U son ortogonales si y solo si

$$\mathbf{u}_1^T\mathbf{u}_2 = \mathbf{u}_2^T\mathbf{u}_1 = 0, \quad \mathbf{u}_1^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_1 = 0, \quad \mathbf{u}_2^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_2 = 0 \tag{5}$$

Todas las columnas de U tienen longitud unitaria si y solo si

$$\mathbf{u}_1^T\mathbf{u}_1 = 1, \quad \mathbf{u}_2^T\mathbf{u}_2 = 1, \quad \mathbf{u}_3^T\mathbf{u}_3 = 1 \tag{6}$$

El teorema se deduce inmediatamente de las ecuaciones (4) a (6). ■

TEOREMA 7

Sea U una matriz de $m \times n$ con columnas ortonormales, y $\mathbf{x}$ y $\mathbf{y}$ están en $\mathbb{R}^n$. Entonces,

a) $\|U\mathbf{x}\| = \|\mathbf{x}\|$

b) $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$

c) $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ si y solo si $\mathbf{x}\cdot\mathbf{y} = 0$

Las propiedades *a*) y *c*) dicen que el mapeo lineal $\mathbf{x} \mapsto U\mathbf{x}$ preserva longitudes y ortogonalidad. Esas propiedades son importantes para muchos algoritmos computacionales. Véase el ejercicio 25 para la demostración del teorema 7.

EJEMPLO 6 Sean $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Observe que U tiene columnas ortonormales y

$$U^TU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Compruebe que $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

SOLUCIÓN

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|\mathbf{x}\| = \sqrt{2+9} = \sqrt{11}$$

■

Los teoremas 6 y 7 son particularmente útiles cuando se aplican a matrices *cuadradas*. Una **matriz ortogonal** es una matriz cuadrada invertible U tal que $U^{-1} = U^T$. De acuerdo con el teorema 6, tal matriz tiene columnas ortonormales.[1] Es fácil ver que cualquier matriz *cuadrada* con columnas ortonormales es una matriz ortogonal. De manera sorprendente, dicha matriz también debe tener *filas* ortonormales. Véase los ejercicios 27 y 28. En el capítulo 7 se presentarán con frecuencia matrices ortogonales.

EJEMPLO 7 La matriz

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

es una matriz ortogonal porque es cuadrada y porque sus columnas son ortonormales, de acuerdo con el ejemplo 5. Compruebe que las filas ¡también sean ortonormales! ■

PROBLEMAS DE PRÁCTICA

1. Sean $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ y $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Demuestre que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortonormal para $\mathbb{R}^2$.
2. Sean $\mathbf{y}$ y L como en el ejemplo 3 y en la figura 3. Calcule la proyección ortogonal $\hat{\mathbf{y}}$ de $\mathbf{y}$ sobre L utilizando $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ en lugar de $\mathbf{u}$ del ejemplo 3.
3. Sean U y $\mathbf{x}$ como en el ejemplo 6, y $\mathbf{y} = \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix}$. Compruebe que $U\mathbf{x}\cdot U\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$.

6.2 EJERCICIOS

En los ejercicios 1 a 6, determine qué conjuntos de vectores son ortogonales.

1. $\begin{bmatrix} -1 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -7 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -7 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

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

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

6. $\begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ -3 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 5 \\ -1 \end{bmatrix}$

En los ejercicios 7 a 10, demuestre que $\{\mathbf{u}_1, \mathbf{u}_2\}$ o $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ son una base ortogonal para $\mathbb{R}^2$ o $\mathbb{R}^3$, respectivamente. Después exprese $\mathbf{x}$ como una combinación lineal de los vectores $\mathbf{u}$.

7. $\mathbf{u}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

[1] Un mejor nombre sería *matriz ortonormal*, y este término se emplea en algunos textos de estadística. Sin embargo, *matriz ortogonal* es el término estándar en álgebra lineal.

8. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$

9. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 8 \\ -4 \\ -3 \end{bmatrix}$

10. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$

11. Calcule la proyección ortogonal de $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ sobre la recta que pasa por $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ y por el origen.

12. Calcule la proyección ortogonal de $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ sobre la recta que pasa por $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ y por el origen.

13. Sean $\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$. Escriba $\mathbf{y}$ como la suma de dos vectores ortogonales, uno en Gen $\{\mathbf{u}\}$ y el otro ortogonal a $\mathbf{u}$.

14. Sean $\mathbf{y} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Escriba $\mathbf{y}$ como la suma de un vector en Gen $\{\mathbf{u}\}$ y un vector ortogonal a $\mathbf{u}$.

15. Sean $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$. Calcule la distancia de $\mathbf{y}$ a la recta que pasa por $\mathbf{u}$ y el origen.

16. Sean $\mathbf{y} = \begin{bmatrix} -3 \\ 9 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Calcule la distancia de $\mathbf{y}$ a la recta que pasa por $\mathbf{u}$ y el origen.

En los ejercicios 17 a 22, determine cuáles conjuntos de vectores son ortonormales. Si un conjunto solamente es ortogonal, normalice los vectores para obtener un conjunto ortonormal.

17. $\begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 0 \\ 1/2 \end{bmatrix}$

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

19. $\begin{bmatrix} -.6 \\ .8 \end{bmatrix}, \begin{bmatrix} .8 \\ .6 \end{bmatrix}$

20. $\begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ 2/3 \\ 0 \end{bmatrix}$

21. $\begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{20} \\ 3/\sqrt{20} \end{bmatrix}, \begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{20} \\ -1/\sqrt{20} \end{bmatrix}, \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

22. $\begin{bmatrix} 1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

En los ejercicios 23 y 24, todos los vectores están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

23. *a*) No todo conjunto linealmente independiente en $\mathbb{R}^n$ es un conjunto ortogonal.

b) Si $\mathbf{y}$ es una combinación lineal de vectores distintos de cero de un conjunto ortogonal, entonces los pesos en la combinación lineal se pueden calcular sin operaciones de fila sobre una matriz.

c) Si se normalizan los vectores en un conjunto ortogonal de vectores distintos de cero, entonces tal vez algunos de los nuevos vectores no sean ortogonales.

d) Una matriz con columnas ortonormales es una matriz ortogonal.

e) Si L es una recta que pasa por $\mathbf{0}$ y si $\hat{\mathbf{y}}$ es la proyección ortogonal de $\mathbf{y}$ sobre L, entonces $\|\hat{\mathbf{y}}\|$ da la distancia de $\mathbf{y}$ a L.

24. *a*) No todo conjunto ortogonal en $\mathbb{R}^n$ es linealmente independiente.

b) Si un conjunto $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ tiene la propiedad de que $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ siempre que $i \neq j$, entonces S es un conjunto ortonormal.

c) Si las columnas de una matriz A de $m \times n$ son ortonormales, entonces el mapeo lineal $\mathbf{x} \mapsto A\mathbf{x}$ preserva longitudes.

d) La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{v}$ es igual a la proyección ortogonal de $\mathbf{y}$ sobre $c\mathbf{v}$ siempre que $c \neq 0$.

e) Una matriz ortogonal es invertible.

25. Demuestre el teorema 7. [*Sugerencia:* Para *a*), calcule $\|U\mathbf{x}\|^2$, o primero demuestre *b*)].

26. Suponga que W es un subespacio de $\mathbb{R}^n$ generado por n vectores ortogonales distintos de cero. Explique por qué $W = \mathbb{R}^n$.

27. Sea U una matriz cuadrada con columnas ortonormales. Explique por qué U es invertible. (Mencione los teoremas que utilice).

28. Sea U una matriz ortogonal de $n \times n$. Demuestre que las filas de U forman una base ortonormal de $\mathbb{R}^n$.

29. Sean U y V matrices ortogonales de $n \times n$. Explique por qué UV es una matriz ortogonal. [Es decir, explique por qué UV es invertible y su inversa es $(UV)^T$].

30. Sea U una matriz ortogonal; construya V intercambiando algunas de las columnas de U. Explique por qué V es una matriz ortogonal.

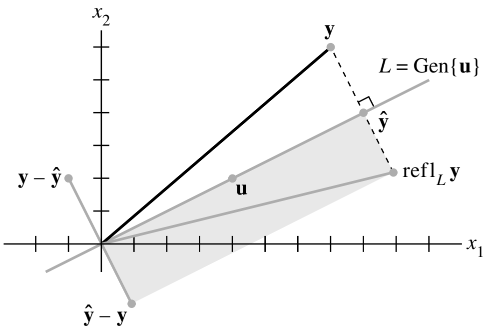
31. Demuestre que la proyección ortogonal de un vector $\mathbf{y}$ sobre una recta L que pasa por el origen en $\mathbb{R}^2$ no depende de la elección del vector $\mathbf{u}$ diferente de cero en L que se emplea en la fórmula para $\hat{\mathbf{y}}$. Para hacer esto, suponga que $\mathbf{y}$ y $\mathbf{u}$ están dados y que $\hat{\mathbf{y}}$ se calculó mediante la fórmula (2) de esta sección. Sustituya $\mathbf{u}$ en tal fórmula por $c\mathbf{u}$, donde c es un escalar distinto de cero no especificado. Demuestre que la nueva fórmula da el mismo valor para $\hat{\mathbf{y}}$.

32. Sean $\{\mathbf{v}_1, \mathbf{v}_2\}$ un conjunto ortogonal de vectores distintos de cero, y c_1, c_2 escalares diferentes de cero. Demuestre que $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2\}$ también es un conjunto ortogonal. Como la ortogonalidad de un conjunto está definida en términos de pares de vectores, esto demuestra que si se normalizan los vectores en un conjunto ortogonal, el nuevo conjunto seguirá siendo ortogonal.

33. Dado $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^n$, sea $L = \text{Gen}\{\mathbf{u}\}$. Demuestre que el mapeo $\mathbf{x} \mapsto \text{proy}_L\,\mathbf{x}$ es una transformación lineal.

34. Dado $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^n$, sea $L = \text{Gen}\{\mathbf{u}\}$. Para $\mathbf{y}$ en $\mathbb{R}^n$, la **reflexión de $\mathbf{y}$ en L** es el punto $\text{refl}_L\,\mathbf{y}$ definido por
$\text{refl}_L\,\mathbf{y} = 2\cdot\text{proy}_L\,\mathbf{y} - \mathbf{y}$.

Véase la figura, que indica que $\text{refl}_L\,\mathbf{y}$ es la suma de $\hat{\mathbf{y}} = \text{proy}_L\,\mathbf{y}$ y $\hat{\mathbf{y}} - \mathbf{y}$. Demuestre que el mapeo $\mathbf{y} \mapsto \text{refl}_L$ es una transformación lineal.



La reflexión de $\mathbf{y}$ en una recta que pasa por el origen.

35. [**M**] Demuestre que las columnas de la matriz A son ortogonales haciendo el cálculo matricial adecuado. Especifique el cálculo que realizó.

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

36. [**M**] En los incisos *a*) a *d*), sea U la matriz formada por la normalización de cada columna de la matriz A del ejercicio 35.

a) Calcule U^TU y UU^T. ¿En qué difieren?

b) Genere un vector aleatorio $\mathbf{y}$ en $\mathbb{R}^8$, y determine $\mathbf{p} = UU^T\mathbf{y}$ y $\mathbf{z} = \mathbf{y} - \mathbf{p}$. Explique por qué $\mathbf{p}$ está en Col A. Compruebe que $\mathbf{z}$ sea ortogonal a $\mathbf{p}$.

c) Compruebe que $\mathbf{z}$ es ortogonal a cada columna de U.

d) Observe que $\mathbf{y} = \mathbf{p} + \mathbf{z}$, con $\mathbf{p}$ en Col A. Explique por qué $\mathbf{z}$ está en $(\text{Col } A)^{\perp}$. (En la siguiente sección se explicará el significado de esta descomposición de $\mathbf{y}$).

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Los vectores son ortogonales porque

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = -2/5 + 2/5 = 0$$

Son unitarios porque

$$\|\mathbf{u}_1\|^2 = (-1/\sqrt{5})^2 + (2/\sqrt{5})^2 = 1/5 + 4/5 = 1$$
$$\|\mathbf{u}_2\|^2 = (2/\sqrt{5})^2 + (1/\sqrt{5})^2 = 4/5 + 1/5 = 1$$

En particular, el conjunto $\{\mathbf{u}_1, \mathbf{u}_2\}$ es linealmente independiente y, por lo tanto, es una base para $\mathbb{R}^2$ porque hay dos vectores en el conjunto.

2. Cuando $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{20}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

Esta es la misma $\hat{\mathbf{y}}$ que se encontró en el ejemplo 3. La proyección ortogonal no parece depender de la $\mathbf{u}$ seleccionada sobre la recta. Véase el ejercicio 31.

3. $$U\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}\begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix}$$

También, del ejemplo 6, $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$ y $U\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$. Por lo tanto,

$$U\mathbf{x}\cdot U\mathbf{y} = 3 + 7 + 2 = 12 \quad \text{y} \quad \mathbf{x}\cdot\mathbf{y} = -6 + 18 = 12$$

6.3 PROYECCIONES ORTOGONALES

La proyección ortogonal de un punto en $\mathbb{R}^2$ sobre una recta que pasa por el origen tiene una importante analogía en $\mathbb{R}^n$. Dado un vector $\mathbf{y}$ y un subespacio W en $\mathbb{R}^n$, existe un vector $\hat{\mathbf{y}}$ en W tal que **1.** $\hat{\mathbf{y}}$ es el único vector en W para el cual $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a W, y **2.** $\hat{\mathbf{y}}$ es el único vector en W más cercano a $\mathbf{y}$. Véase la figura 1. Estas dos propiedades de $\hat{\mathbf{y}}$ dan la clave para encontrar las soluciones de mínimos cuadrados de sistemas lineales, que se mencionaron en el ejemplo introductorio de este capítulo. En la sección 6.5 se contará la historia completa.

Como preparación para el primer teorema, observe que siempre que un vector $\mathbf{y}$ se representa como una combinación lineal de vectores $\mathbf{u}_1, \dots, \mathbf{u}_n$ en $\mathbb{R}^n$, los términos en la suma para $\mathbf{y}$ se pueden agrupar en dos partes de manera que $\mathbf{y}$ se representa como

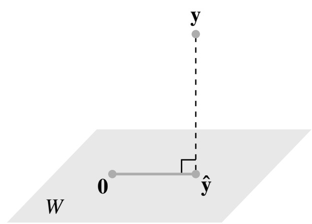


FIGURA 1

$$\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$$

donde $\mathbf{z}_1$ es una combinación lineal de algunas $\mathbf{u}_i$, y $\mathbf{z}_2$ es una combinación lineal de las $\mathbf{u}_i$ restantes. Esta idea es útil en particular cuando $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ es una base ortogonal. De la sección 6.1 recuerde que $W^\perp$ denota el conjunto de todos los vectores ortogonales a un subespacio W.

EJEMPLO 1 Sea $\{\mathbf{u}_1, \dots, \mathbf{u}_5\}$ una base ortogonal para $\mathbb{R}^5$ y

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

Considere el subespacio $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, y escriba $\mathbf{y}$ como la suma de un vector $\mathbf{z}_1$ en W y un vector $\mathbf{z}_2$ en $W^\perp$.

SOLUCIÓN Escriba

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + c_2\mathbf{u}_2}_{\mathbf{z}_1} + \underbrace{c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5}_{\mathbf{z}_2}$$

donde $\quad \mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \quad$ está en Gen $\{\mathbf{u}_1, \mathbf{u}_2\}$

y $\quad \mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5 \quad$ está en Gen $\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.

Para demostrar que $\mathbf{z}_2$ está en $W^\perp$, es suficiente probar que $\mathbf{z}_2$ es ortogonal a los vectores en la base $\{\mathbf{u}_1, \mathbf{u}_2\}$ para W. (Véase la sección 6.1). Utilizando propiedades del producto interior, calcule

$$\begin{aligned}\mathbf{z}_2 \cdot \mathbf{u}_1 &= (c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5)\cdot\mathbf{u}_1 \\ &= c_3\mathbf{u}_3 \cdot \mathbf{u}_1 + c_4\mathbf{u}_4 \cdot \mathbf{u}_1 + c_5\mathbf{u}_5 \cdot \mathbf{u}_1 \\ &= 0\end{aligned}$$

ya que $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_3$, $\mathbf{u}_4$ y $\mathbf{u}_5$. Un cálculo similar indica que $\mathbf{z}_2 \cdot \mathbf{u}_2 = 0$. Así, $\mathbf{z}_2$ está en $W^\perp$. ■

El siguiente teorema demuestra que la descomposición $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$ del ejemplo 1 se puede calcular sin tener una base ortogonal para $\mathbb{R}^n$. Es suficiente con tener una base ortogonal solo para W.

TEOREMA 8

Teorema de descomposición ortogonal

Sea W un subespacio de $\mathbb{R}^n$. Entonces toda $\mathbf{y}$ en $\mathbb{R}^n$ se puede escribir de forma única como

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

donde $\hat{\mathbf{y}}$ está en W y $\mathbf{z}$ está en $W^\perp$. De hecho, si $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ es cualquier base ortogonal de W, entonces

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \tag{2}$$

y $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

El vector $\hat{\mathbf{y}}$ en la ecuación (1) es la **proyección ortogonal de y sobre** W, y con frecuencia se escribe como $\text{proy}_W\,\mathbf{y}$. Véase la figura 2. Cuando W es un subespacio unidimensional, la fórmula para $\hat{\mathbf{y}}$ coincide con la fórmula que se presentó en la sección 6.2.

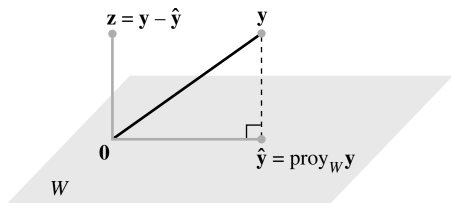


FIGURA 2 Proyección ortogonal de $\mathbf{y}$ sobre W.

DEMOSTRACIÓN Sea $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ una base ortogonal para W, y defina $\hat{\mathbf{y}}$ con la ecuación (2).[1] Entonces $\hat{\mathbf{y}}$ está en W porque $\hat{\mathbf{y}}$ es una combinación lineal de la base $\mathbf{u}_1,\ldots,\mathbf{u}_p$. Sea $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Como $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_2,\ldots,\mathbf{u}_p$, entonces a partir de la ecuación (2) se deduce que

$$\begin{aligned}\mathbf{z}\cdot\mathbf{u}_1 &= (\mathbf{y}-\hat{\mathbf{y}})\cdot\mathbf{u}_1 = \mathbf{y}\cdot\mathbf{u}_1 - \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1\cdot\mathbf{u}_1 - 0 - \cdots - 0\\ &= \mathbf{y}\cdot\mathbf{u}_1 - \mathbf{y}\cdot\mathbf{u}_1 = 0\end{aligned}$$

Por lo tanto, $\mathbf{z}$ es ortogonal a $\mathbf{u}_1$. De manera similar, $\mathbf{z}$ es ortogonal a cada $\mathbf{u}_j$ en la base para W. Por consiguiente, $\mathbf{z}$ es ortogonal para todo vector en W. Es decir, $\mathbf{z}$ está en $W^\perp$.

Para demostrar que la descomposición en la ecuación (1) es única, suponga que $\mathbf{y}$ se puede escribir como $\mathbf{y} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$ con $\hat{\mathbf{y}}_1$ en W, y $\mathbf{z}_1$ en $W^\perp$. Entonces $\hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$ (ya que ambos lados son iguales a $\mathbf{y}$), y así

$$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$$

Esta igualdad indica que el vector $\mathbf{v} = \mathbf{y} - \hat{\mathbf{y}}_1$ está en W y en $W^\perp$ (porque $\mathbf{z}_1$ y $\mathbf{z}$ están ambos en $W^\perp$, y $W^\perp$ es un subespacio). Por lo tanto, $\mathbf{v}\cdot\mathbf{v} = 0$, lo que demuestra que $\mathbf{v} = \mathbf{0}$. Esto prueba que $\mathbf{y} = \hat{\mathbf{y}}_1$ y también que $\mathbf{z}_1 = \mathbf{z}$. ■

La unicidad de la descomposición (1) demuestra que la proyección ortogonal $\hat{\mathbf{y}}$ solo depende de W y no de la base particular empleada en la ecuación (2).

[1] Se puede suponer que W no es el subespacio cero porque, de otra forma, $W^\perp = \mathbb{R}^n$ y la ecuación (1) simplemente sería $\mathbf{y} = \mathbf{0} + \mathbf{y}$. En la siguiente sección se demostrará que cualquier subespacio distinto de cero de $\mathbb{R}^n$ tiene una base ortogonal.

EJEMPLO 2 Sean $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortogonal para $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$. Escriba $\mathbf{y}$ como la suma de un vector en W y de un vector ortogonal a W.

SOLUCIÓN La proyección ortogonal de $\mathbf{y}$ sobre W es

$$\begin{aligned} \hat{\mathbf{y}} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 \\ &= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} \end{aligned}$$

Además,

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

El teorema 8 asegura que $\mathbf{y} - \hat{\mathbf{y}}$ está en $W^{\perp}$. Sin embargo, para comprobar los cálculos, es una buena idea comprobar que $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$, y por lo tanto a toda W. La descomposición deseada de $\mathbf{y}$ es

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

■

Interpretación geométrica de la proyección ortogonal

Cuando W es un subespacio unidimensional, la fórmula (2) para $\text{proy}_W\,\mathbf{y}$ solo contiene un término. Así, cuando $\dim W > 1$, cada término de la ecuación (2) es en sí mismo una proyección ortogonal de $\mathbf{y}$ sobre el subespacio unidimensional generado por uno de los vectores $\mathbf{u}$ en la base para W. La figura 3 muestra esto cuando W es un subespacio de $\mathbb{R}^3$ generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Aquí, $\hat{\mathbf{y}}_1$ y $\hat{\mathbf{y}}_2$ denotan las proyecciones de $\mathbf{y}$ sobre las rectas generadas por $\mathbf{u}_1$ y $\mathbf{u}_2$, respectivamente. La proyección ortogonal $\hat{\mathbf{y}}$ de $\mathbf{y}$ sobre W es la suma de las proyecciones de $\mathbf{y}$ sobre subespacios unidimensionales que son ortogonales entre sí. El vector $\hat{\mathbf{y}}$ en la figura 3 corresponde al vector $\mathbf{y}$ de la figura 4 de la sección 6.2, porque ahora es $\hat{\mathbf{y}}$ el que está en W.

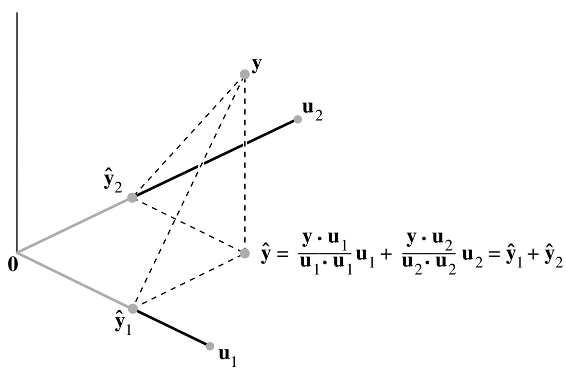


FIGURA 3 La proyección ortogonal de $\mathbf{y}$ es la suma de sus proyecciones sobre subespacios unidimensionales que son mutuamente ortogonales.

Propiedades de proyecciones ortogonales

Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una base ortogonal para W y resulta que $\mathbf{y}$ está en W, entonces la fórmula para $\text{proy}_W\, \mathbf{y}$ es exactamente la misma que la representación de $\mathbf{y}$ en el teorema 5 de la sección 6.2. En este caso, $\text{proy}_W\, \mathbf{y} = \mathbf{y}$.

Si $\mathbf{y}$ está en $W = \text{Gen}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, entonces $\text{proy}_W\, \mathbf{y} = \mathbf{y}$.

Este resultado también se deduce del siguiente teorema.

TEOREMA 9

Teorema de la mejor aproximación

W es un subespacio de $\mathbb{R}^n$, $\mathbf{y}$ es cualquier vector en $\mathbb{R}^n$, y $\hat{\mathbf{y}}$ es la proyección ortogonal de $\mathbf{y}$ sobre W. Entonces $\hat{\mathbf{y}}$ en W es el punto más cercano a $\mathbf{y}$, en el sentido de que

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \tag{3}$$

para toda $\mathbf{v}$ en W diferente de $\hat{\mathbf{y}}$.

El vector $\hat{\mathbf{y}}$ del teorema 9 se llama **la mejor aproximación a y por elementos de *W***. En secciones posteriores del libro se examinarán problemas donde una $\mathbf{y}$ dada se debe remplazar, o *aproximar*, mediante un vector $\mathbf{v}$ en algún subespacio W fijo. La distancia de $\mathbf{y}$ a $\mathbf{v}$, dada por $\|\mathbf{y} - \mathbf{v}\|$, se puede considerar como el "error" de usar $\mathbf{v}$ en lugar de $\mathbf{y}$. El teorema 9 afirma que este error se minimiza cuando $\mathbf{v} = \hat{\mathbf{y}}$.

La desigualdad (3) conduce a una nueva demostración de que $\hat{\mathbf{y}}$ no depende de la base ortogonal particular empleada para calcularlo. Si se utilizara una base ortogonal diferente para W con la finalidad de construir una proyección ortogonal de $\mathbf{y}$, entonces esta proyección también sería el punto en W más cercano a $\mathbf{y}$, a saber, $\hat{\mathbf{y}}$.

DEMOSTRACIÓN Tome $\mathbf{v}$ en W diferente de $\hat{\mathbf{y}}$. Véase la figura 4. Entonces $\hat{\mathbf{y}} - \mathbf{v}$ está en W. De acuerdo con el teorema de descomposición ortogonal, $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a W. En particular, $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a $\hat{\mathbf{y}} - \mathbf{v}$ (que está en W). Puesto que

$$\mathbf{y} - \mathbf{v} = (\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{v})$$

entonces, al utilizar el teorema de Pitágoras, se obtiene

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2$$

(Véase el triángulo rectángulo a la derecha de la figura 4. Se indica la longitud de cada lado). Ahora $\|\hat{\mathbf{y}} - \mathbf{v}\|^2 > 0$ porque $\hat{\mathbf{y}} - \mathbf{v} \neq 0$, y así se deduce inmediatamente la desigualdad (3). ∎

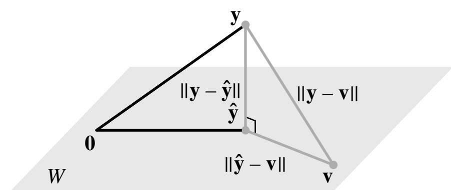


FIGURA 4 La proyección ortogonal de $\mathbf{y}$ sobre W es el punto en W más cercano a $\mathbf{y}$.

EJEMPLO 3 Si $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ y $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, como en el ejemplo 2, entonces el punto en W más cercano a $\mathbf{y}$ es

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$ ■

EJEMPLO 4 La distancia de un punto $\mathbf{y}$ en $\mathbb{R}^n$ a un subespacio W se define como la distancia de $\mathbf{y}$ al punto más cercano en W. Encuentre la distancia de $\mathbf{y}$ a $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, donde

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

SOLUCIÓN De acuerdo con el teorema de la mejor aproximación, la distancia de $\mathbf{y}$ a W es $\|\mathbf{y} - \hat{\mathbf{y}}\|$, donde $\hat{\mathbf{y}} = \text{proy}_W\,\mathbf{y}$. Ya que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortogonal para W,

$$\hat{\mathbf{y}} = \frac{15}{30}\mathbf{u}_1 + \frac{-21}{6}\mathbf{u}_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = 3^2 + 6^2 = 45$$

La distancia de $\mathbf{y}$ a W es $\sqrt{45} = 3\sqrt{5}$. ■

El teorema final en esta sección muestra cómo la fórmula (2) para $\text{proy}_W\,\mathbf{y}$ se simplifica cuando la base para W es un conjunto ortonormal.

TEOREMA 10

Si $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ es una base ortonormal para un subespacio W de $\mathbb{R}^n$, entonces

$$\text{proy}_W\,\mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p \tag{4}$$

Si $U = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_p]$, entonces

$$\text{proy}_W\,\mathbf{y} = UUT\mathbf{y} \quad \text{para toda } \mathbf{y} \text{ en } \mathbb{R}^n \tag{5}$$

DEMOSTRACIÓN La fórmula (4) se obtiene inmediatamente de la ecuación (2) del teorema 8. Además, la ecuación (4) indica que $\text{proy}_W\,\mathbf{y}$ es una combinación lineal de las columnas de U empleando los pesos $\mathbf{y}\cdot\mathbf{u}_1$, $\mathbf{y}\cdot\mathbf{u}_2,\ldots, \mathbf{y}\cdot\mathbf{u}_p$. Los pesos se pueden representar como $\mathbf{u}_1^T\mathbf{y}, \mathbf{u}_2^T\mathbf{y}, \ldots, \mathbf{u}_p^T\mathbf{y}$, probando que son las entradas en $U^T\mathbf{y}$, y justificando la ecuación (5). ■

WEB

Suponga que U es una matriz de $n \times p$ con columnas ortonormales, y sea W el espacio columna de U. Entonces,

$$U^TU\mathbf{x} = I_p\,\mathbf{x} = \mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^p \quad \text{Teorema 6}$$

$$UU^T\mathbf{y} = \text{proy}_W\,\mathbf{y} \quad \text{para toda } \mathbf{y} \text{ en } \mathbb{R}^n \quad \text{Teorema 10}$$

Si U es una matriz (cuadrada) de $n \times n$ con columnas ortonormales, entonces U es una matriz *ortogonal*, el espacio columna W es todo de $\mathbb{R}^n$, y $UU^T\mathbf{y} = I\mathbf{y} = \mathbf{y}$ para toda $\mathbf{y}$ en $\mathbb{R}^n$.

Aunque la fórmula (4) es importante para fines teóricos, en la práctica generalmente implica algunos cálculos con raíces cuadradas de números (en las entradas de $\mathbf{u}_i$). La fórmula (2) se recomienda para cálculos a mano.

PROBLEMA DE PRÁCTICA

Sean $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$ y $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$. Con base en el hecho de que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales, calcule $\text{proy}_W\, \mathbf{y}$.

6.3 EJERCICIOS

En los ejercicios 1 y 2, se puede suponer que $\{\mathbf{u}_1, \ldots, \mathbf{u}_4\}$ es una base ortogonal para $\mathbb{R}^4$.

1. $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ -4 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 3 \\ 5 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -4 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} 5 \\ -3 \\ -1 \\ 1 \end{bmatrix}$

   $\mathbf{x} = \begin{bmatrix} 10 \\ -8 \\ 2 \\ 0 \end{bmatrix}$. Escriba $\mathbf{x}$ como la suma de dos vectores, uno en $\text{Gen}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y el otro en $\text{Gen}\{\mathbf{u}_4\}$.

2. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ -1 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}$.

   $\mathbf{v} = \begin{bmatrix} 4 \\ 5 \\ -3 \\ 3 \end{bmatrix}$. Escriba $\mathbf{v}$ como la suma de dos vectores, uno en $\text{Gen}\{\mathbf{u}_1\}$ y el otro en $\text{Gen}\{\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$.

En los ejercicios 3 a 6, compruebe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es un conjunto ortogonal, y luego encuentre la proyección ortogonal de $\mathbf{y}$ sobre $\text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$.

3. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

4. $\mathbf{y} = \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$

5. $\mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 6 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$

6. $\mathbf{y} = \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} -4 \\ -1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 7 a 10, sea W el subespacio generado por los vectores $\mathbf{u}$, y escriba $\mathbf{y}$ como la suma de un vector en W y un vector ortogonal a W.

7. $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$

8. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$

9. $\mathbf{y} = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

10. $\mathbf{y} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

En los ejercicios 11 y 12, determine el punto más cercano a $\mathbf{y}$ en el subespacio W generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

11. $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

12. $\mathbf{y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$

En los ejercicios 13 y 14, calcule la mejor aproximación a $\mathbf{z}$ con vectores de la forma $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$.

13. $\mathbf{z} = \begin{bmatrix} 3 \\ -7 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ -3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$

14. $\mathbf{z} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 5 \\ -2 \\ 4 \\ 2 \end{bmatrix}$

15. Sean $\mathbf{y} = \begin{bmatrix} 5 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$. Encuentre la distancia de $\mathbf{y}$ al plano en $\mathbb{R}^3$ generado por $\mathbf{u}_1$ y $\mathbf{u}_2$.

16. Sean $\mathbf{y}$, $\mathbf{v}_1$ y $\mathbf{v}_2$ como en el ejercicio 12. Encuentre la distancia de $\mathbf{y}$ al subespacio de $\mathbb{R}^4$ generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

17. Sean $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ y $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$.

a) Sea $U = [\mathbf{u}_1 \quad \mathbf{u}_2]$. Calcule U^TU y UU^T.

b) Calcule $\text{proy}_W\, \mathbf{y}$ y $(UU^T)\mathbf{y}$.

18. Sean $\mathbf{y} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$ y $W = \text{Gen}\,\{\mathbf{u}_1\}$.

a) Sea U la matriz de 2×1 cuya única columna es $\mathbf{u}_1$. Calcule U^TU y UU^T.

b) Determine $\text{proy}_W\, \mathbf{y}$ y $(UU^T)\mathbf{y}$.

19. Sean $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$ y $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Observe que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales, pero que $\mathbf{u}_3$ no es ortogonal a $\mathbf{u}_1$ o $\mathbf{u}_2$. Es posible demostrar que $\mathbf{u}_3$ no está en el subespacio W generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Con base en este hecho, construya un vector $\mathbf{v}$ diferente de cero en $\mathbb{R}^3$ que sea ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$.

20. Sean $\mathbf{u}_1$ y $\mathbf{u}_2$ como en el ejercicio 19, y $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Es posible demostrar que $\mathbf{u}_4$ no está en el subespacio W generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Con base en este hecho, obtenga un vector $\mathbf{v}$ diferente de cero en $\mathbb{R}^3$ que sea ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$.

En los ejercicios 21 y 22, todos los vectores y subespacios están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) Si $\mathbf{z}$ es ortogonal a $\mathbf{u}_1$ y a $\mathbf{u}_2$, y si $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, entonces $\mathbf{z}$ debe estar en $W^\perp$.

b) Para cada $\mathbf{y}$ y cada subespacio W, el vector $\mathbf{y} - \text{proy}_W\, \mathbf{y}$ es ortogonal a W.

c) Algunas veces $\hat{\mathbf{y}}$, la proyección ortogonal de $\mathbf{y}$ sobre un subespacio W, puede depender de la base ortogonal de W empleada para calcular $\hat{\mathbf{y}}$.

d) Si $\mathbf{y}$ está en un subespacio W, entonces la proyección ortogonal de $\mathbf{y}$ sobre W es $\mathbf{y}$ misma.

e) Si las columnas de una matriz U de $n \times p$, son ortonormales, entonces $UU^T\mathbf{y}$ es la proyección ortogonal de $\mathbf{y}$ sobre el espacio columna de U.

22. *a*) Si W es un subespacio de $\mathbb{R}^n$ y si $\mathbf{v}$ está en W y en $W^\perp$, entonces $\mathbf{v}$ debe ser el vector cero.

b) En el teorema de descomposición ortogonal, cada término en la fórmula (2) para $\hat{\mathbf{y}}$ es en sí mismo una proyección de $\mathbf{y}$ sobre un subespacio de W.

c) Si $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$, donde $\mathbf{z}_1$ está en un subespacio W y $\mathbf{z}_2$ está en $W^\perp$, entonces $\mathbf{z}_1$ debe ser la proyección ortogonal de $\mathbf{y}$ sobre W.

d) La mejor aproximación a $\mathbf{y}$ por elementos de un subespacio W está dada por el vector $\mathbf{y} - \text{proy}_W\, \mathbf{y}$.

e) Si una matriz U de $n \times p$, tiene columnas ortonormales, entonces $UU^T\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

23. Sea A una matriz de $m \times n$. Demuestre que cada vector $\mathbf{x}$ en $\mathbb{R}^n$ se puede escribir en la forma $\mathbf{x} = \mathbf{p} + \mathbf{u}$, donde $\mathbf{p}$ está en Fil A y $\mathbf{u}$ pertenece a Nul A. Además, demuestre que si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces existe una única $\mathbf{p}$ en Fil A tal que $A\mathbf{p} = \mathbf{b}$.

24. Sea W un subespacio de $\mathbb{R}^n$ con una base ortogonal $\{\mathbf{w}_1, \ldots, \mathbf{w}_p\}$, y sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$ una base ortogonal para $W^\perp$.

a) Explique por qué $\{\mathbf{w}_1, \ldots, \mathbf{w}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$ es un conjunto ortogonal.

b) Explique por qué el conjunto del inciso *a*) genera a $\mathbb{R}^n$.

c) Demuestre que $\dim W + \dim W^\perp = n$.

25. [**M**] Sea U la matriz de 8×4 del ejercicio 36 de la sección 6.2. Encuentre el punto más cercano a $\mathbf{y} = (1, 1, 1, 1, 1, 1, 1, 1)$ en Col U. Escriba las instrucciones que utilizó para resolver este problema.

26. [**M**] Sea U la matriz del ejercicio 25. Obtenga la distancia de $\mathbf{b} = (1, 1, 1, 1, -1, -1, -1, -1)$ a Col U.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Calcule

$$\text{proy}_W\, \mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \frac{88}{66}\mathbf{u}_1 + \frac{-2}{6}\mathbf{u}_2$$

$$= \frac{4}{3}\begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix} = \mathbf{y}$$

En este caso, $\mathbf{y}$ resulta ser una combinación lineal de $\mathbf{u}_1$ y $\mathbf{u}_2$, de manera que $\mathbf{y}$ está en W. El punto en W más cerca de $\mathbf{y}$ es $\mathbf{y}$ misma.

6.4 PROCESO DE GRAM-SCHMIDT

El proceso de Gram-Schmidt es un sencillo algoritmo para obtener una base ortogonal u ortonormal para cualquier subespacio diferente de cero de $\mathbb{R}^n$. Los dos primeros ejemplos de este proceso son para realizarse a mano.

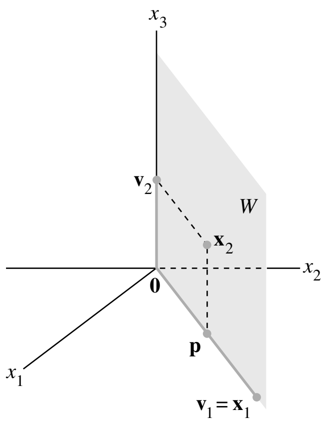


FIGURA 1
Construcción de una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$.

EJEMPLO 1 Sea $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, donde $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ y $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construya una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$ para W.

SOLUCIÓN En la figura 1 se ilustra el subespacio W, junto con $\mathbf{x}_1$, $\mathbf{x}_2$ y la proyección $\mathbf{p}$ de $\mathbf{x}_2$ sobre $\mathbf{x}_1$. La componente de $\mathbf{x}_2$ ortogonal a $\mathbf{x}_1$ es $\mathbf{x}_2 - \mathbf{p}$, que está en W porque se forma a partir de $\mathbf{x}_2$ y de un múltiplo de $\mathbf{x}_1$. Sea $\mathbf{v}_1 = \mathbf{x}_1$ y

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1}{\mathbf{x}_1 \cdot \mathbf{x}_1}\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Entonces $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto ortogonal de vectores distintos de cero en W. Como $\dim W = 2$, entonces el conjunto $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para W. ■

El siguiente ejemplo ilustra plenamente el proceso de Gram-Schmidt. Estúdielo con cuidado.

EJEMPLO 2 Sean $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Entonces, $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ es, a todas luces, linealmente independiente y, por consiguiente, es una base para un subespacio W de $\mathbb{R}^4$. Construya una base ortogonal para W.

SOLUCIÓN

Paso 1. Sean $\mathbf{v}_1 = \mathbf{x}_1$ y $W_1 = \text{Gen}\,\{\mathbf{x}_1\} = \text{Gen}\,\{\mathbf{v}_1\}$.

Paso 2. Sea $\mathbf{v}_2$ el vector producido al restar de $\mathbf{x}_2$ su proyección sobre el subespacio W_1. Es decir, sea

$$\begin{aligned} \mathbf{v}_2 &= \mathbf{x}_2 - \text{proy}_{W_1}\,\mathbf{x}_2 \\ &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \qquad \text{Ya que } \mathbf{v}_1 = \mathbf{x}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} \end{aligned}$$

Como en el ejemplo 1, $\mathbf{v}_2$ es la componente de $\mathbf{x}_2$ ortogonal a $\mathbf{x}_1$, y $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base ortogonal para el subespacio W_2 generado por $\mathbf{x}_1$ y $\mathbf{x}_2$.

Paso 2′ (opcional). Si es pertinente, escale $\mathbf{v}_2$ para simplificar cálculos posteriores. Como $\mathbf{v}_2$ tiene entradas fraccionales, entonces es conveniente escalarlo por un factor de 4 y sustituir $\{\mathbf{v}_1, \mathbf{v}_2\}$ por la base ortogonal

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Paso 3. Sea $\mathbf{v}_3$ el vector obtenido restando de $\mathbf{x}_3$ su proyección sobre el subespacio W_2. Utilice la base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2'\}$ para calcular esta proyección sobre W_2:

$$\text{proy}_{W_2}\,\mathbf{x}_3 = \underbrace{\frac{\mathbf{x}_3\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1}_{\substack{\text{Proyección}\\ \text{de } \mathbf{x}_3 \text{ sobre } \mathbf{v}_1}} + \underbrace{\frac{\mathbf{x}_3\cdot\mathbf{v}_2'}{\mathbf{v}_2'\cdot\mathbf{v}_2'}\mathbf{v}_2'}_{\substack{\text{Proyección de}\\ \mathbf{x}_3 \text{ sobre } \mathbf{v}_2'}} = \frac{2}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{2}{12}\begin{bmatrix}-3\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix}$$

Entonces, $\mathbf{v}_3$ es la componente de $\mathbf{x}_3$ ortogonal a W_2, a saber,

$$\mathbf{v}_3 = \mathbf{x}_3 - \text{proy}_{W_2}\,\mathbf{x}_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix} = \begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$$

Véase en la figura 2 un diagrama de esta construcción. Observe que $\mathbf{v}_3$ está en W porque $\mathbf{x}_3$ y $\text{proy}_{W_2}\mathbf{x}_3$ están en W. Así, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ es un conjunto ortogonal de vectores diferentes de cero y, por lo tanto, un conjunto linealmente independiente en W. Observe que W es tridimensional porque está definido con una base de tres vectores. Por lo tanto, de acuerdo con el teorema de la base de la sección 4.5, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ es una base ortogonal para W. ■

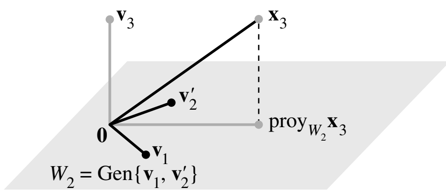


FIGURA 2 Construcción de $\mathbf{v}_3$ a partir de $\mathbf{x}_3$ y W_2.

La demostración del siguiente teorema revela que esta estrategia realmente funciona. No se menciona el escalamiento de vectores porque solo se usa para simplificar los cálculos a mano.

TEOREMA 11

Proceso de Gram-Schmidt

A partir de una base $\{\mathbf{x}_1,\ldots,\mathbf{x}_p\}$ para un subespacio W de $\mathbb{R}^n$, se define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1\\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1\\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2\\
&\vdots\\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p\cdot\mathbf{v}_{p-1}}{\mathbf{v}_{p-1}\cdot\mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Entonces, $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ es una base ortogonal para W. Además,

$$\text{Gen}\,\{\mathbf{v}_1,\ldots,\mathbf{v}_k\} = \text{Gen}\,\{\mathbf{x}_1,\ldots,\mathbf{x}_k\} \quad \text{para } 1 \le k \le p \tag{1}$$

DEMOSTRACIÓN Para $1 \leq k \leq p$, sea $W_k = \text{Gen}\,\{\mathbf{x}_1,\ldots,\mathbf{x}_k\}$. Establezca que $\mathbf{v}_1 = \mathbf{x}_1$, por lo que Gen $\{\mathbf{v}_1\}$ = Gen $\{\mathbf{x}_1\}$. Suponga que, para alguna $k < p$, se construyeron $\mathbf{v}_1,\ldots,\mathbf{v}_k$, de manera que $\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ es una base ortogonal para W_k. Se define

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - \text{proy}_{Wk}\,\mathbf{x}_{k+1} \tag{2}$$

De acuerdo con el teorema de descomposición ortogonal, $\mathbf{v}_{k+1}$ es ortogonal a W_k. Observe que $\text{proy}_{W_k}\,\mathbf{x}_{k+1}$ está en W_k y, por lo tanto, también en W_{k+1}. Como $\mathbf{x}_{k+1}$ está en W_{k+1}, entonces también $\mathbf{v}_{k+1}$ lo está (ya que W_{k+1} es un subespacio y es cerrado bajo la resta). Además, $\mathbf{v}_{k+1} \neq \mathbf{0}$ porque $\mathbf{x}_{k+1}$ no está en $W_k = \text{Gen}\,\{\mathbf{x}_1,\ldots,\mathbf{x}_k\}$. Por lo tanto, $\{\mathbf{v}_1,\ldots,\mathbf{v}_{k+1}\}$ es un conjunto ortogonal de vectores distintos de cero en el espacio $(k+1)$-dimensional W_{k+1}. De acuerdo con el teorema de la base de la sección 4.5, este conjunto es una base ortogonal para W_{k+1}. Por lo tanto, $W_{k+1} = \text{Gen}\,\{\mathbf{v}_1,\ldots,\mathbf{v}_{k+1}\}$. El proceso se detiene cuando $k + 1 = p$. ■

El teorema 11 indica que cualquier subespacio W distinto de cero de $\mathbb{R}^n$ tiene una base ortogonal, porque una base ordinaria $\{\mathbf{x}_1,\ldots,\mathbf{x}_p\}$ siempre está disponible (de acuerdo con el teorema 11 de la sección 4.5), y el proceso de Gram-Schmidt solo depende de la existencia de proyecciones ortogonales sobre subespacios de W que ya tengan bases ortogonales.

Bases ortonormales

Una base ortonormal se construye con facilidad a partir de una base ortogonal $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$: simplemente se normalizan (es decir, se "escalan") todas las $\mathbf{v}_k$. Cuando se trabajan problemas a mano, esto es más fácil que normalizar cada $\mathbf{v}_k$ conforme se vayan encontrando (porque evita la escritura innecesaria de raíces cuadradas).

EJEMPLO 3 En el ejemplo 1 se construyó la base ortogonal

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Una base ortonormal es

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{45}}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

■

Factorización QR de matrices

WEB

Si una matriz A de $m \times n$ tiene columnas linealmente independientes $\mathbf{x}_1,\ldots,\mathbf{x}_n$, la aplicación del proceso de Gram-Schmidt (con normalizaciones) a $\mathbf{x}_1,\ldots,\mathbf{x}_n$ equivale a *factorizar* A, como se describe en el siguiente teorema. Esta factorización se utiliza ampliamente en algoritmos computacionales para diversos cálculos, como la resolución de ecuaciones (que se analizó en la sección 6.5) y la determinación de valores propios (que se mencionó en los ejercicios de la sección 5.2).

TEOREMA 12

La factorización QR

Si A es una matriz de $m \times n$ con columnas linealmente independientes, entonces A se puede factorizar como $A = QR$, donde Q es una matriz de $m \times n$ cuyas columnas forman una base ortonormal para Col A, y R es una matriz triangular superior invertible de $n \times n$ con entradas positivas en su diagonal.

DEMOSTRACIÓN Las columnas de A forman una base $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ para Col A. Construya una base ortonormal $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ para $W = \text{Col } A$ con la propiedad (1) del teorema 11. Esta base se puede construir mediante el proceso de Gram-Schmidt o con alguna otra técnica. Sea

$$Q = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n\,]$$

Para $k = 1, \ldots, n$, $\mathbf{x}_k$ está en Gen $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \text{Gen }\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. Entonces existen constantes, $r_{1k}, \ldots, r_{kk}$, tales que

$$\mathbf{x}_k = r_{1k}\mathbf{u}_1 + \cdots + r_{kk}\mathbf{u}_k + 0 \cdot \mathbf{u}_{k+1} + \cdots + 0 \cdot \mathbf{u}_n$$

Podemos suponer que $r_{kk} \geq 0$. (Si $r_{kk} < 0$, multiplique r_{kk} y $\mathbf{u}_k$ por -1). Esto muestra que $\mathbf{x}_k$ es una combinación lineal de las columnas de Q empleando como pesos las entradas del vector

$$\mathbf{r}_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Es decir, $\mathbf{x}_k = Q\mathbf{r}_k$ para $k = 1, \ldots, n$. Sea $R = [\mathbf{r}_1 \cdots \mathbf{r}_n]$. Entonces,

$$A = [\,\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_n\,] = [\,Q\mathbf{r}_1 \quad \cdots \quad Q\mathbf{r}_n\,] = QR$$

El hecho de que R es invertible se deduce fácilmente del hecho de que las columnas de A son linealmente independientes (ejercicio 19). Como resulta evidente que R es triangular superior, sus entradas diagonales no negativas deben ser positivas. ■

EJEMPLO 4 Encuentre una factorización QR de $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

SOLUCIÓN Las columnas de A son los vectores $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ del ejemplo 2. En ese ejemplo, se encontró una base ortogonal para $\text{Col } A = \text{Gen }\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}'_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Para simplificar la aritmética que sigue, escale $\mathbf{v}_3$ dejando que $\mathbf{v}'_3 = 3\mathbf{v}_3$. Después, normalice los tres vectores para obtener $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$, y utilice esos vectores como las columnas de Q:

$$Q = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}$$

Por construcción, las primeras k columnas de Q son una base ortonormal de Gen $\{\mathbf{x}_1,\ldots,\mathbf{x}_k\}$. De la demostración del teorema 12, $A = QR$ para alguna R. Para encontrar R, observe que $Q^TQ = I$, porque las columnas de Q son ortonormales. De ahí que

$$Q^TA = Q^T(QR) = IR = R$$

y

$$R = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}$$

■

NOTAS NUMÉRICAS

1. Cuando el proceso de Gram-Schmidt se ejecuta en una computadora, el error por redondeo puede crecer conforme se van calculando los vectores $\mathbf{u}_k$, uno por uno. Cuando j y k son grandes, pero diferentes, los productos interiores $\mathbf{u}_j^T\mathbf{u}_k$ quizá no sean suficientemente cercanos a cero. Esta pérdida de ortogonalidad se puede reducir de manera sustancial reordenando los cálculos.[1] Sin embargo, en vez de este método de Gram-Schmidt modificado se prefiere el método de la factorización QR basado en computadora porque conduce a bases ortonormales más exactas, aun cuando la factorización requiere casi el doble de aritmética.
2. Para obtener una factorización QR de una matriz A, un programa de cómputo generalmente multiplica A por la izquierda por una secuencia de matrices ortogonales hasta que A se transforme en una matriz triangular superior. Esta construcción es análoga a la multiplicación por la izquierda con matrices elementales que produce una factorización LU de A.

PROBLEMA DE PRÁCTICA

Sea $W =$ Gen $\{\mathbf{x}_1, \mathbf{x}_2\}$, donde $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{x}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$. Construya una base ortonormal para W.

6.4 EJERCICIOS

En los ejercicios 1 a 6, el conjunto indicado es una base para un subespacio W. Utilice el proceso de Gram-Schmidt con la finalidad de obtener una base ortogonal para W.

1. $\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$

4. $\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -7 \\ -4 \\ 1 \end{bmatrix}$

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

[1] Véase *Fundamentals of Matrix Computations*, de David S. Watkins (Nueva York: John Wiley & Sons, 1991), pp. 167-180.

7. Encuentre una base ortonormal del subespacio generado por los vectores del ejercicio 3.

8. Obtenga una base ortonormal del subespacio generado por los vectores del ejercicio 4.

En los ejercicios 9 a 12, determine una base ortogonal para el espacio columna de cada matriz.

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

11. $\begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$ **12.** $\begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}$

En los ejercicios 13 y 14, las columnas de Q se obtuvieron aplicando el proceso de Gram-Schmidt a las columnas de A. Encuentre una matriz R triangular superior tal que $A = QR$. Compruebe su resultado.

13. $A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}, Q = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix}$

14. $A = \begin{bmatrix} -2 & 3 \\ 5 & 7 \\ 2 & -2 \\ 4 & 6 \end{bmatrix}, Q = \begin{bmatrix} -2/7 & 5/7 \\ 5/7 & 2/7 \\ 2/7 & -4/7 \\ 4/7 & 2/7 \end{bmatrix}$

15. Encuentre una factorización QR de la matriz del ejercicio 11.

16. Obtenga una factorización QR de la matriz del ejercicio 12.

En los ejercicios 17 y 18, todos los vectores y subespacios están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

17. *a*) Si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortogonal para W, entonces la multiplicación de $\mathbf{v}_3$ por un escalar c da una nueva base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2, c\mathbf{v}_3\}$.

b) El proceso de Gram-Schmidt aplicado a un conjunto linealmente independiente $\{\mathbf{x}_1,\ldots, \mathbf{x}_p\}$ produce un conjunto ortogonal $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ con la propiedad de que para cada k, los vectores $\mathbf{v}_1,\ldots, \mathbf{v}_k$ generan el mismo subespacio que se originó por $\mathbf{x}_1,\ldots, \mathbf{x}_k$.

c) Si $A = QR$, donde Q tiene columnas ortonormales, entonces $R = Q^TA$.

18. *a*) Si $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ con $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linealmente independiente, y si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal en W, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para W.

b) Si $\mathbf{x}$ no está en un subespacio W, entonces $\mathbf{x} - \text{proy}_W\, \mathbf{x}$ no es cero.

c) En una factorización QR, por ejemplo, $A = QR$ (cuando A tiene columnas linealmente independientes), las columnas de Q forman una base ortonormal para el espacio columna de A.

19. Suponga que $A = QR$, donde Q es de $m \times n$ y R es de $n \times n$. Demuestre que si las columnas de A son linealmente independientes, entonces R debe ser invertible. [*Sugerencia:* Estudie la ecuación $R\mathbf{x} = \mathbf{0}$ y considere el hecho de que $A = QR$].

20. Suponga que $A = QR$, donde R es una matriz invertible. Demuestre que A y Q tienen el mismo espacio columna. [*Sugerencia:* Dada $\mathbf{y}$ en Col A, demuestre que $\mathbf{y} = Q\mathbf{x}$ para alguna $\mathbf{x}$. Además, dada $\mathbf{y}$ en Col Q, demuestre que $\mathbf{y} = A\mathbf{x}$ para alguna $\mathbf{x}$].

21. Dada $A = QR$ como en el teorema 12, describa cómo encontrar una matriz ortogonal Q_1 de $m \times m$ y una matriz triangular superior invertible R de $n \times n$ tales que

$$A = Q_1 \begin{bmatrix} R \\ 0 \end{bmatrix}$$

La instrucción qr de **MATLAB** proporciona esta factorización QR "completa" cuando el rango de $A = n$.

22. Sean $\mathbf{u}_1,\ldots, \mathbf{u}_p$ una base ortogonal para un subespacio W de $\mathbb{R}^n$, y $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(\mathbf{x}) = \text{proy}_W\, \mathbf{x}$. Demuestre que T es una transformación lineal.

23. Suponga que $A = QR$ es una factorización QR de una matriz A de $m \times n$ (con columnas linealmente independientes). Particione A como $[A_1\ A_2]$, donde A_1 tiene p columnas. Muestre cómo obtener una factorización QR de A_1, y explique por qué su factorización tiene las características adecuadas.

24. [**M**] Utilice el proceso de Gram-Schmidt, como en el ejemplo 2, y obtenga una base ortogonal para el espacio columna de

$$A = \begin{bmatrix} -10 & 13 & 7 & -11 \\ 2 & 1 & -5 & 3 \\ -6 & 3 & 13 & -3 \\ 16 & -16 & -2 & 5 \\ 2 & 1 & -5 & -7 \end{bmatrix}$$

25. [**M**] Aplique el método expuesto en esta sección para obtener una factorización de la matriz del ejercicio 24.

26. [**M**] En un programa de matrices, el proceso de Gram-Schmidt funciona mucho mejor con vectores ortonormales. Comenzando con $\mathbf{x}_1,\ldots, \mathbf{x}_p$ como en el teorema 11, sea $A = [\mathbf{x}_1 \cdots x_p]$. Suponga que Q es una matriz de $n \times k$ cuyas columnas forman una base ortonormal para el subespacio W_k generado por las primeras k columnas de A. Entonces para $\mathbf{x}$ en $\mathbb{R}^n$, $QQ^T\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre W_k (teorema 10 de la sección 6.3). Si $\mathbf{x}_{k+1}$ es la siguiente columna de A, entonces la ecuación (2) en la demostración del teorema 11 se convierte en

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - Q(Q^T\mathbf{x}_{k+1})$$

(Los paréntesis en la ecuación anterior reducen el número de operaciones aritméticas). Sea $\mathbf{u}_{k+1} = \mathbf{v}_{k+1}/\|\mathbf{v}_{k+1}\|$. La nueva Q para el siguiente paso es $[Q\quad \mathbf{u}_{k+1}]$. Utilice este procedimiento para calcular la factorización QR de la matriz del ejercicio 24. Escriba las instrucciones que utilice.

WEB

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Sean $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \mathbf{x}_2 - \dfrac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - 0\mathbf{v}_1 = \mathbf{x}_2$. Así, $\{\mathbf{x}_1, \mathbf{x}_2\}$ ya es ortogonal. Ahora todo lo que se necesita es normalizar los vectores. Sea

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

En vez de normalizar $\mathbf{v}_2$ directamente, se normaliza $\mathbf{v}_2' = 3\mathbf{v}_2$:

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2'\|}\mathbf{v}_2' = \frac{1}{\sqrt{1^2 + 1^2 + (-2)^2}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

Entonces, $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortonormal para W.

6.5 PROBLEMAS DE MÍNIMOS CUADRADOS

El ejemplo introductorio a este capítulo describió un enorme problema del tipo $A\mathbf{x} = \mathbf{b}$ que no tuvo solución. En las aplicaciones son frecuentes los sistemas inconsistentes, aunque por lo general no aparecen matrices de coeficientes tan grandes. Cuando se pide una solución y esta no existe, lo mejor que se puede hacer es encontrar una $\mathbf{x}$ tal que $A\mathbf{x}$ esté lo más cerca posible de $\mathbf{b}$.

Piense que $A\mathbf{x}$ es una *aproximación* a $\mathbf{b}$. Cuanto menor sea la distancia entre $\mathbf{b}$ y $A\mathbf{x}$, dada por $\|\mathbf{b} - A\mathbf{x}\|$, mucho mejor será la aproximación. El **problema general de mínimos cuadrados** consiste en encontrar una $\mathbf{x}$ que haga a $\|\mathbf{b} - A\mathbf{x}\|$ tan pequeña como sea posible. El adjetivo "mínimos cuadrados" se origina en el hecho de que $\|\mathbf{b} - A\mathbf{x}\|$ es la raíz cuadrada de una suma de cuadrados.

DEFINICIÓN

Si A es de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$, una **solución de mínimos cuadrados** de $A\mathbf{x} = \mathbf{b}$ es un $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ tal que

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

para toda $\mathbf{x}$ en $\mathbb{R}^n$.

El aspecto más importante del problema de mínimos cuadrados es que sin importar cuál $\mathbf{x}$ se seleccione, el vector $A\mathbf{x}$ necesariamente estará en el espacio columna, Col A. Así que se busca un vector $\mathbf{x}$ que haga que $A\mathbf{x}$ sea el punto en Col A más cercano a $\mathbf{b}$. Véase la figura 1. (Desde luego, si $\mathbf{b}$ resulta estar en Col A, entonces $\mathbf{b}$ *es* $A\mathbf{x}$ para alguna $\mathbf{x}$, y tal $\mathbf{x}$ es una "solución de mínimos cuadrados").

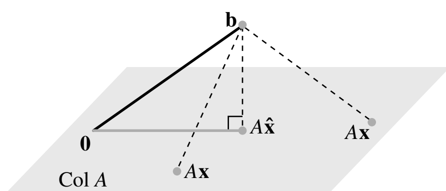


FIGURA 1 El vector $\mathbf{b}$ está más cerca de $A\hat{\mathbf{x}}$ que de $A\mathbf{x}$ para otra $\mathbf{x}$.

Solución del problema general de mínimos cuadrados

Con base en los A y $\mathbf{b}$ anteriores, aplique el teorema de la mejor aproximación que se expuso en la sección 6.3 al subespacio Col A. Sea

$$\hat{\mathbf{b}} = \text{proy}_{\text{Col } A}\, \mathbf{b}$$

Como $\hat{\mathbf{b}}$ está en el espacio columna de A, la ecuación $A\mathbf{x} = \hat{\mathbf{b}}$ *es* consistente, y existe una $\hat{\mathbf{b}}$ en $\mathbb{R}^n$ tal que

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}} \tag{1}$$

Como $\hat{\mathbf{b}}$ es el punto en Col A más cercano a $\mathbf{b}$, un vector $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ si y solo si $\hat{\mathbf{x}}$ satisface la ecuación (1). Tal $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ es un lista de pesos para construir $\hat{\mathbf{b}}$ a partir de las columnas de A. Véase la figura 2. [Existen muchas soluciones de la ecuación (1) si tiene variables libres].

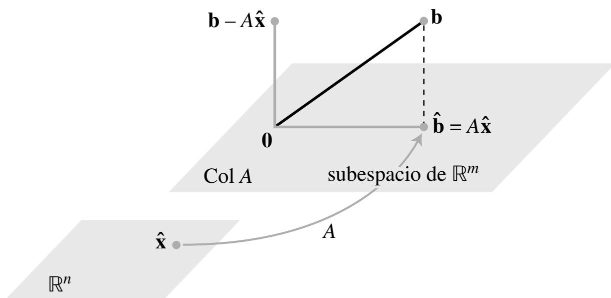


FIGURA 2 La solución de mínimos cuadrados $\hat{\mathbf{x}}$ está en $\mathbb{R}^n$.

Suponga que $\hat{\mathbf{x}}$ satisface $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Según el teorema de descomposición ortogonal de la sección 6.3, la proyección $\hat{\mathbf{b}}$ tiene la propiedad de que $\mathbf{b} - \hat{\mathbf{b}}$ es ortogonal a Col A, de manera que $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a cada columna de A. Si $\mathbf{a}_j$ es cualquier columna de A, entonces $\mathbf{a}_j \cdot (\mathbf{b} - A\hat{\mathbf{x}}) = 0$ y $\mathbf{a}_j^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0$. Como cada $\mathbf{a}_j^T$ es una fila de A^T,

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0} \tag{2}$$

(Esta ecuación también se deduce del teorema 3 de la sección 6.1). Así,

$$\begin{aligned} A^T\mathbf{b} - A^TA\hat{\mathbf{x}} &= \mathbf{0} \\ A^TA\hat{\mathbf{x}} &= A^T\mathbf{b} \end{aligned}$$

Estos cálculos indican que cada solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ satisface la ecuación

$$\boxed{A^TA\mathbf{x} = A^T\mathbf{b}} \tag{3}$$

La ecuación matricial (3) representa un sistema de ecuaciones llamado las **ecuaciones normales** para $A\mathbf{x} = \mathbf{b}$. Es frecuente que cada solución del sistema (3) se denote con $\hat{\mathbf{x}}$.

TEOREMA 13

El conjunto de soluciones de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ coincide con el conjunto no vacío de soluciones de las ecuaciones normales $A^TA\mathbf{x} = A^T\mathbf{b}$.

DEMOSTRACIÓN Como ya se mostró antes, el conjunto de soluciones de mínimos cuadrados es no vacío y cada solución de mínimos cuadrados $\hat{\mathbf{x}}$ satisface las ecuaciones normales. A la inversa, suponga que $\hat{\mathbf{x}}$ satisface $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$. Entonces, $\hat{\mathbf{x}}$ satisface la ecuación (2) anterior, lo que demuestra que $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a las filas de A^T y, por lo tanto, ortogonal a las columnas de A. Como las columnas de A generan Col A, entonces el vector $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a todo elemento de Col A. Por consiguiente, la ecuación

$$\mathbf{b} = A\hat{\mathbf{x}} + (\mathbf{b} - A\hat{\mathbf{x}})$$

es una descomposición de $\mathbf{b}$ en la suma de un vector en Col A y de un vector ortogonal a Col A. Por la singularidad de la descomposición ortogonal, $A\hat{\mathbf{x}}$ debe ser la proyección ortogonal de $\mathbf{b}$ sobre Col A. Es decir, $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, y $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados. ■

EJEMPLO 1 Encuentre una solución de mínimos cuadrados del sistema inconsistente $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

SOLUCIÓN Para emplear las ecuaciones normales (3), calcule

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Entonces, la ecuación $A^TA\mathbf{x} = A^T\mathbf{b}$ se convierte en

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Pueden utilizarse operaciones de fila para resolver este sistema, pero como A^TA es de 2×2 e invertible, tal vez sea más rápido calcular

$$(A^TA)^{-1} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

y luego resolver $A^TA\mathbf{x} = A^T\mathbf{b}$ como

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ &= \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$ ■

En muchos cálculos, A^TA es invertible, pero este no siempre es el caso. El siguiente ejemplo implica a una matriz como las que suelen presentarse en estadística, en problemas de *análisis de varianza*.

EJEMPLO 2 Encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

SOLUCIÓN Calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

La matriz aumentada para $A^TA\mathbf{x} = A^T\mathbf{b}$ es

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

La solución general es $x_1 = 3 - x_4$, $x_2 = -5 + x_4$, $x_3 = -2 + x_4$, y x_4 es libre. Así, la solución general de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ tiene la forma

$$\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$ ■

El siguiente teorema brinda útiles criterios para determinar cuándo existe solamente una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. (Desde luego, la proyección ortogonal $\hat{\mathbf{b}}$ siempre es única).

TEOREMA 14

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

a) La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución de mínimos cuadrados única para cada $\mathbf{b}$ en $\mathbb{R}^m$.

b) Las columnas de A son linealmente independientes.

c) La matriz A^TA es invertible.

Cuando estos enunciados son verdaderos, la solución $\hat{\mathbf{x}}$ de mínimos cuadrados está dada por

$$\hat{\mathbf{x}} = (A^TA)^{-1} A^T\mathbf{b} \tag{4}$$

En los ejercicios 19 a 21 se indican los elementos principales para demostrar el teorema 14; en tales ejercicios también se revisan conceptos del capítulo 4. La fórmula (4) para $\hat{\mathbf{x}}$ es útil sobre todo para fines teóricos y cálculos a mano cuando A^TA es una matriz invertible de 2×2.

Cuando se usa una solución $\hat{\mathbf{x}}$ de mínimos cuadrados para producir $A\hat{\mathbf{x}}$ como una aproximación a $\mathbf{b}$, entonces la distancia de $\mathbf{b}$ a $A\hat{\mathbf{x}}$ se llama el **error de mínimos cuadrados** de esta aproximación.

EJEMPLO 3 Dados A y $\mathbf{b}$ como en el ejemplo 1, determine el error de mínimos cuadrados en la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

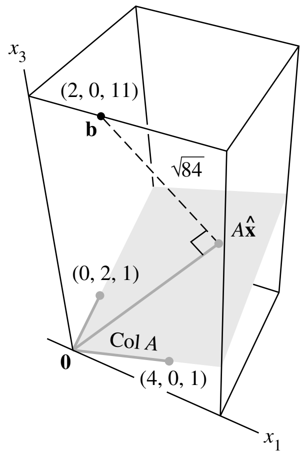


FIGURA 3

SOLUCIÓN A partir del ejemplo 1,

$$\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \quad \text{y} \quad A\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

Por lo tanto,

$$\mathbf{b} - A\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

y

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$$

El error de mínimos cuadrados es $\sqrt{84}$. Para cualquier $\mathbf{x}$ en $\mathbb{R}^2$, la distancia entre $\mathbf{b}$ y el vector $A\mathbf{x}$ es al menos $\sqrt{84}$. Véase la figura 3. Observe que la solución $\hat{\mathbf{x}}$ de mínimos cuadrados no se presenta en la figura. ■

Cálculos alternativos de soluciones de mínimos cuadrados

El siguiente ejemplo muestra cómo encontrar una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando las columnas de A son ortogonales. Con frecuencia tales matrices se presentan en problemas de regresión lineal, como los que se analizarán en la siguiente sección.

EJEMPLO 4 Encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

SOLUCIÓN Como las columnas $\mathbf{a}_1$ y $\mathbf{a}_2$ de A son ortogonales, la proyección ortogonal de $\mathbf{b}$ sobre Col A está dada por

$$\begin{aligned} \hat{\mathbf{b}} &= \frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2 = \frac{8}{4}\mathbf{a}_1 + \frac{45}{90}\mathbf{a}_2 \\ &= \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 1/2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5/2 \\ 11/2 \end{bmatrix} \end{aligned} \tag{5}$$

Ahora que se conoce $\hat{\mathbf{b}}$, se puede resolver $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Pero esto es trivial, porque ya se sabe que los pesos a colocar sobre las columnas de A producen $\hat{\mathbf{b}}$. A partir de la ecuación (5) es claro que

$$\hat{\mathbf{x}} = \begin{bmatrix} 8/4 \\ 45/90 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}$$ ■

En algunos casos, es posible que las ecuaciones normales para un problema de mínimos cuadrados estén *mal condicionadas*; es decir, en ocasiones, pequeños errores en los cálculos de las entradas de A^TA causan grandes errores en la solución $\hat{\mathbf{x}}$. Si las columnas de A son linealmente independientes, la solución de mínimos cuadrados con frecuencia se puede calcular de manera más confiable con una factorización QR de A (descrita en la sección 6.4).[1]

[1] El método QR se compara con el método de la ecuación normal estándar en G. Golub y C. Van Loan, *Matrix Computations*, 3a. ed. (Baltimore: Johns Hopkins Press, 1996), pp. 230-231.

TEOREMA 15 Dada una matriz A de $m \times n$, con columnas linealmente independientes, sea $A = QR$ una factorización QR de A como en el teorema 12. Entonces, para cada $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución de mínimos cuadrados única, dada por

$$\hat{\mathbf{x}} = R^{-1} Q^T \mathbf{b} \tag{6}$$

DEMOSTRACIÓN Sea $\hat{\mathbf{x}} = R^{-1} Q^T \mathbf{b}$. Entonces,

$$A\hat{\mathbf{x}} = QR\hat{\mathbf{x}} = QRR^{-1}Q^T\mathbf{b} = QQ^T\mathbf{b}$$

De acuerdo con el teorema 12, las columnas de Q forman una base ortonormal para Col A. Por lo tanto, según el teorema 10, $QQ^T\mathbf{b}$ es la proyección ortogonal $\hat{\mathbf{b}}$ de $\mathbf{b}$ sobre Col A. Así, $A\hat{\mathbf{x}} = \mathbf{b}$, lo que muestra que $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. La unicidad de $\hat{\mathbf{x}}$ se deduce del teorema 14. ■

NOTA NUMÉRICA

Como en el teorema 15, R es triangular superior, entonces $\hat{\mathbf{x}}$ se debería calcular como la solución exacta de la ecuación

$$R\mathbf{x} = Q^T\mathbf{b} \tag{7}$$

Es mucho más rápido resolver (7), por sustitución hacia atrás o mediante operaciones de fila, que calcular R^{-1} y utilizar la ecuación (6).

EJEMPLO 5 Encuentre la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

SOLUCIÓN La factorización QR de A se obtiene como en la sección 6.4

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

Luego,

$$Q^T\mathbf{b} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

La solución de mínimos cuadrados $\hat{\mathbf{x}}$ satisface $R\mathbf{x} = Q^T\mathbf{b}$; es decir,

$$\begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

Esta ecuación se resuelve fácilmente y conduce a $\hat{\mathbf{x}} = \begin{bmatrix} 10 \\ -6 \\ 2 \end{bmatrix}$. ■

PROBLEMAS DE PRÁCTICA

1. Sean $A = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$. Encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, y calcule el error de mínimos cuadrados asociado.
2. ¿Qué se puede decir acerca de la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando $\mathbf{b}$ es ortogonal a las columnas de A?

6.5 EJERCICIOS

En los ejercicios 1 a 4, encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ mediante *a*) la construcción de las ecuaciones normales para $\hat{\mathbf{x}}$ y *b*) el despeje de $\hat{\mathbf{x}}$.

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

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

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

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

En los ejercicios 5 y 6, describa todas las soluciones de mínimos cuadrados de la ecuación $A\mathbf{x} = \mathbf{b}$.

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

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

7. Calcule el error de mínimos cuadrados asociado con la solución de mínimos cuadrados que encontró en el ejercicio 3.

8. Determine el error de mínimos cuadrados asociado con la solución de mínimos cuadrados que encontró en el ejercicio 4.

En los ejercicios 9 a 12, encuentre *a*) la proyección ortogonal de $\mathbf{b}$ sobre Col A y *b*) una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

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

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

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

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

13. Sean $A = \begin{bmatrix} 3 & 4 \\ -2 & 1 \\ 3 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 11 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$. Calcule $A\mathbf{u}$ y $A\mathbf{v}$, y compárelos con $\mathbf{b}$. ¿Podría ser $\mathbf{u}$ una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$? (Responda sin calcular una solución de mínimos cuadrados).

14. Sean $A = \begin{bmatrix} 2 & 1 \\ -3 & -4 \\ 3 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ 4 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$. Determine $A\mathbf{u}$ y $A\mathbf{v}$, y compárelos con $\mathbf{b}$. ¿Es posible que al menos uno de los dos entre $\mathbf{u}$ o $\mathbf{v}$ sea una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$? (Responda sin obtener una solución de mínimos cuadrados).

En los ejercicios 15 y 16, utilice la factorización $A = QR$ para encontrar la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

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

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

En los ejercicios 17 y 18, A es una matriz de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

17. *a*) El problema de mínimos cuadrados general consiste en obtener una $\mathbf{x}$ que haga que $A\mathbf{x}$ esté tan cerca de $\mathbf{b}$ como sea posible.

b) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es un vector $\hat{\mathbf{x}}$ que satisface $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, donde $\hat{\mathbf{b}}$ es la proyección ortogonal de $\mathbf{b}$ sobre Col A.

c) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es un vector $\hat{\mathbf{x}}$ tal que $\|\mathbf{b} - A\mathbf{x}\| \leq \|\mathbf{b} - A\hat{\mathbf{x}}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

d) Cualquier solución de $A^TA\mathbf{x} = A^T\mathbf{b}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

e) Si las columnas de A son linealmente independientes, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene exactamente una solución de mínimos cuadrados.

18. *a*) Si $\mathbf{b}$ está en el espacio columna de A, entonces cada solución de $A\mathbf{x} = \mathbf{b}$ es una solución de mínimos cuadrados.

b) La solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es el punto en el espacio columna de A más cercano a $\mathbf{b}$.

c) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es una lista de pesos que, cuando se aplica a las columnas de A, produce la proyección ortogonal de $\mathbf{b}$ sobre Col A.

d) Si $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, entonces $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

e) Las ecuaciones normales siempre ofrecen un método confiable para calcular soluciones de mínimos cuadrados.

f) Si A tiene una factorización QR, por ejemplo, $A = QR$, entonces la mejor manera de obtener una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es calcular $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$.

19. Sea A una matriz de $m \times n$. Realice los siguientes pasos para demostrar que un vector $\mathbf{x}$ en $\mathbb{R}^n$ satisface $A\mathbf{x} = \mathbf{0}$ si y solo si $A^TA\mathbf{x} = \mathbf{0}$. Esto demostrará que Nul A = Nul A^TA.

a) Demuestre que si $A\mathbf{x} = \mathbf{0}$, entonces $A^TA\mathbf{x} = \mathbf{0}$.

b) Suponga que $A^TA\mathbf{x} = \mathbf{0}$. Explique por qué $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{0}$, y use esto para probar que $A\mathbf{x} = \mathbf{0}$.

20. Sea A una matriz de $m \times n$ tal que A^TA es invertible. Demuestre que las columnas de A son linealmente independientes. [*Precaución:* No suponga que A sea invertible; es más, tal vez ni siquiera sea cuadrada].

21. Sea A una matriz de $m \times n$ cuyas columnas son linealmente independientes. [*Precaución:* A no necesariamente es cuadrada].

a) Con base en el ejercicio 19, demuestre que A^TA es una matriz invertible.

b) Explique por qué A debe tener, al menos, tantas filas como columnas.

c) Determine el rango de A.

22. Con base en el ejercicio 19, demuestre que rango A^TA = rango A. [*Sugerencia:* ¿Cuántas columnas tiene A^TA? ¿Cómo se relaciona esto con el rango de A^TA?].

23. Suponga que A es de $m \times n$ con columnas linealmente independientes y que $\mathbf{b}$ está en $\mathbb{R}^m$. Utilice las ecuaciones normales con la finalidad de obtener una fórmula para $\hat{\mathbf{b}}$, la proyección de $\mathbf{b}$ sobre Col A. [*Sugerencia:* Primero encuentre $\hat{\mathbf{x}}$. La fórmula no requiere una base ortogonal para Col A].

24. Obtenga una fórmula para la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando las columnas de A son ortonormales.

25. Describa todas las soluciones de mínimos cuadrados del sistema

$$x + y = 2$$
$$x + y = 4$$

26. [**M**] El ejemplo 3 de la sección 4.8 mostró un filtro lineal pasa bajos que cambió la señal $\{y_k\}$ en $\{y_{k+1}\}$ y transformó la señal de alta frecuencia $\{w_k\}$ a una señal cero, donde $y_k = \cos(\pi k/4)$ y $w_k = \cos(3\pi k/4)$. Los siguientes cálculos diseñarán un filtro con aproximadamente esas propiedades. La ecuación del filtro es

$$a_0y_{k+2} + a_1y_{k+1} + a_2y_k = z_k \quad \text{para toda } k \tag{8}$$

Como las señales son periódicas, con periodo 8, basta con estudiar la ecuación (8) para $k = 0,\ldots,7$. La acción sobre las dos señales descritas se traduce en dos conjuntos de ocho ecuaciones, que se muestran a continuación:

$$\begin{array}{c} \\ k=0\\ k=1\\ \vdots\\ \\ \\ \\ \\ k=7 \end{array} \begin{array}{c} \begin{array}{ccc} y_{k+2} & y_{k+1} & y_k \end{array} \\ \begin{bmatrix} 0 & .7 & 1\\ -.7 & 0 & .7\\ -1 & -.7 & 0\\ -.7 & -1 & -.7\\ 0 & -.7 & -1\\ .7 & 0 & -.7\\ 1 & .7 & 0\\ .7 & 1 & .7 \end{bmatrix} \end{array} \begin{bmatrix} a_0\\ a_1\\ a_2 \end{bmatrix} = \begin{array}{c} y_{k+1} \\ \begin{bmatrix} .7\\ 0\\ -.7\\ -1\\ -.7\\ 0\\ .7\\ 1 \end{bmatrix} \end{array}$$

$$\begin{array}{c} \\ k=0\\ k=1\\ \vdots\\ \\ \\ \\ \\ k=7 \end{array} \begin{array}{c} \begin{array}{ccc} w_{k+2} & w_{k+1} & w_k \end{array} \\ \begin{bmatrix} 0 & -.7 & 1\\ .7 & 0 & -.7\\ -1 & .7 & 0\\ .7 & -1 & .7\\ 0 & .7 & -1\\ -.7 & 0 & .7\\ 1 & -.7 & 0\\ -.7 & 1 & -.7 \end{bmatrix} \end{array} \begin{bmatrix} a_0\\ a_1\\ a_2 \end{bmatrix} = \begin{bmatrix} 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0 \end{bmatrix}$$

Escriba una ecuación $A\mathbf{x} = \mathbf{b}$, donde A es una matriz de 16×3 formada por las dos matrices de coeficientes anteriores y donde $\mathbf{b}$ en $\mathbb{R}^{16}$ se forma a partir de los dos lados derechos de las ecuaciones. Encuentre a_0, a_1 y a_2 dadas por la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. (El .7 en los datos anteriores se empleó como una aproximación a $\sqrt{2}/2$, para ilustrar cómo se procede en un cálculo típico en un problema aplicado. Si, en vez de ello, se utilizara .707, los coeficientes del filtro resultante concordarían, al menos, en siete decimales con $\sqrt{2}/4$, $1/2$ y $\sqrt{2}/4$, los valores obtenidos mediante cálculos aritméticos exactos).

WEB

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Primero calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 0 \\ 9 & 83 & 28 \\ 0 & 28 & 14 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -3 \\ -65 \\ -28 \end{bmatrix}$$

Después, reduzca por filas la matriz aumentada para las ecuaciones normales $A^TA\mathbf{x} = A^T\mathbf{b}$:

$$\begin{bmatrix} 3 & 9 & 0 & -3 \\ 9 & 83 & 28 & -65 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 56 & 28 & -56 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -3/2 & 2 \\ 0 & 1 & 1/2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución de mínimos cuadrados general es $x_1 = 2 + \frac{3}{2}x_3$, $x_2 = -1 - \frac{1}{2}x_3$, con x_3 libre. Para una solución específica, tome $x_3 = 0$ (por ejemplo), y obtenga

$$\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

Para encontrar el error de mínimos cuadrados, calcule

$$\hat{\mathbf{b}} = A\hat{\mathbf{x}} = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$$

Resulta que $\hat{\mathbf{b}} = \mathbf{b}$, de manera que $\|\mathbf{b} - \hat{\mathbf{b}}\| = 0$. El error de mínimos cuadrados es cero porque $\mathbf{b}$ está en Col A.

2. Si $\mathbf{b}$ es ortogonal a las columnas de A, entonces la proyección de $\mathbf{b}$ sobre el espacio columna de A es $\mathbf{0}$. En este caso, una solución de mínimos cuadrados $\hat{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ satisface $A\hat{\mathbf{x}} = \mathbf{0}$.

6.6 APLICACIONES A MODELOS LINEALES

Una tarea común en ciencia e ingeniería es analizar y entender relaciones entre diferentes cantidades que varían. Esta sección describe una variedad de situaciones en las que los datos se emplean para construir o comprobar una fórmula que predice el valor de una variable como función de otras variables. En cada caso, el reto será equivalente a resolver un problema de mínimos cuadrados.

Para facilitar la aplicación del análisis a problemas reales que encontrará en sus actividades profesionales, se elige la notación que comúnmente se utiliza en el análisis estadístico de datos científicos y de ingeniería. En vez de escribir $A\mathbf{x} = \mathbf{b}$, se escribe $X\boldsymbol{\beta} = \mathbf{y}$, donde X es la **matriz de diseño**, $\boldsymbol{\beta}$ es el **vector de parámetros** y $\mathbf{y}$ es el **vector de observaciones**.

Rectas de mínimos cuadrados

La relación más sencilla entre dos variables x y y es la ecuación lineal $y = \beta_0 + \beta_1 x$.[1] Con frecuencia los datos experimentales generan los puntos $(x_1, y_1), \ldots, (x_n, y_n)$ que, cuando

[1] Es común utilizar esta notación para rectas de mínimos cuadrados en vez de $y = mx + b$.

se grafican, parecen estar cerca de una recta. Se desea determinar los parámetros β_0 y β_1 para hacer que la recta esté lo más "cerca" posible de dichos puntos.

Suponga que β_0 y β_1 están fijos, y considere la recta $y = \beta_0 + \beta_1 x$ de la figura 1. Para cada dato (x_j, y_j) existe un punto $(x_j, \beta_0 + \beta_1 x_j)$ sobre la recta con la misma coordenada x. Por otro lado, y_j es el valor *observado* de y, y $\beta_0 + \beta_1 x_j$ es el valor *predicho* para y (determinado por la recta). La diferencia entre los valores observado y predicho para y se llama *residuo*.

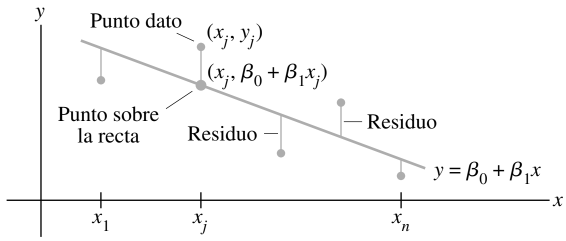


FIGURA 1 Ajuste de datos experimentales a una recta.

Existen varias maneras de medir qué tan "cerca" está la recta respecto de los datos. La elección habitual (sobre todo porque los cálculos matemáticos son sencillos) es sumar los cuadrados de los residuos. La **recta de mínimos cuadrados** es la recta $y = \beta_0 + \beta_1 x$ que minimiza la suma de los cuadrados de los residuos**.** A esta recta también se le conoce como **recta de regresión de y sobre x**, porque se supone que cualquier error en los datos solo ocurre en las coordenadas y. Los coeficientes β_0, β_1 de la recta son los **coeficientes de regresión**.[2]

Si los puntos de los datos estuvieran sobre una recta, los parámetros β_0 y β_1 satisfarían las ecuaciones

Valor predicho de y		Valor observado de y
$\beta_0 + \beta_1 x_1$	$=$	y_1
$\beta_0 + \beta_1 x_2$	$=$	y_2
$\vdots$		$\vdots$
$\beta_0 + \beta_1 x_n$	$=$	y_n

Este sistema se puede representar como

$$X\boldsymbol{\beta} = \mathbf{y} \quad \text{donde} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{1}$$

Desde luego, si los puntos de datos no están sobre una recta, entonces no existen parámetros β_0, β_1 para los cuales los valores predichos de y en $X\beta$ sean iguales a los valores observados de y en $\mathbf{y}$, y $X\beta = \mathbf{y}$ no tiene solución. Este es un problema de mínimos cuadrados, $A\mathbf{x} = \mathbf{b}$, ¡con diferente notación!

El cuadrado de la distancia entre los vectores $X\beta$ y $\mathbf{y}$ es precisamente la suma de los cuadrados de los residuos. El β que minimiza esta suma también minimiza la distancia entre $X\beta$ y $\mathbf{y}$. *Calcular la solución de mínimos cuadrados de $X\beta = \mathbf{y}$ equivale a encontrar el vector β que determina la recta de mínimos cuadrados de la figura 1.*

---

[2] Si los errores de medición estuvieran en x y no en y, simplemente se intercambiarían las coordenadas de los datos (x_j, y_j) antes de trazar la gráfica de los puntos y calcular la recta de regresión. Si ambas coordenadas están sujetas a posibles errores, entonces se podría elegir la recta que minimiza la suma de los cuadrados de las distancias *ortogonales* (perpendiculares) de los puntos a la recta. Véase los problemas de práctica de la sección 7.5.

EJEMPLO 1 Encuentre la ecuación $y = \beta_0 + \beta_1 x$ de la recta de mínimos cuadrados que mejor se ajuste a los puntos de datos (2, 1), (5, 2), (7, 3) y (8, 3).

SOLUCIÓN Utilice las coordenadas x de los datos para construir la matriz de diseño X en la ecuación (1) y las coordenadas y para construir el vector de observaciones $\mathbf{y}$:

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

Para la solución de mínimos cuadrados de $X\beta = \mathbf{y}$, obtenga las ecuaciones normales (con la nueva notación):

$$X^T X \beta = X^T \mathbf{y}$$

Es decir, calcule

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Las ecuaciones normales son

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Por lo tanto,

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

De manera que la recta de mínimos cuadrados tiene la ecuación

$$y = \frac{2}{7} + \frac{5}{14}x$$

Véase la figura 2. ■

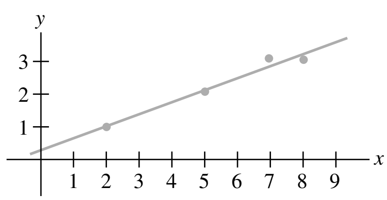


FIGURA 2 La recta de mínimos cuadrados $y = \frac{2}{7} + \frac{5}{14}x$.

Una práctica común, antes de calcular la recta de mínimos cuadrados, consiste en calcular el promedio $\bar{x}$ de los valores x originales y formar una nueva variable $x^* = x - \bar{x}$. Se dice que los nuevos datos x quedan en su **forma de desviación media**. En este caso, las dos columnas de la matriz de diseño serán ortogonales. Se simplifica la solución de las ecuaciones normales, justo como en el ejemplo 4 de la sección 6.5. Véase los ejercicios 17 y 18.

Modelo lineal general

En algunas aplicaciones, es necesario ajustar puntos de datos a algo diferente de una línea recta. En los ejemplos que siguen, la ecuación matricial continúa siendo $X\beta = \mathbf{y}$, pero la forma específica de X cambia de un problema a otro. Por lo general, los especialistas en estadística introducen un **vector residual** $\boldsymbol{\epsilon}$, definido como $\boldsymbol{\epsilon} = \mathbf{y} - X\beta$, y que se escribe

$$\mathbf{y} = X\beta + \boldsymbol{\epsilon}$$

Cualquier ecuación de esta forma es un **modelo lineal**. Una vez que X y $\mathbf{y}$ están determinadas, el objetivo es minimizar la longitud de $\boldsymbol{\epsilon}$, lo que equivale a encontrar una solución de mínimos cuadrados de $X\beta = \mathbf{y}$. En cada caso, la solución de mínimos cuadrados $\hat{\beta}$ es una solución de las ecuaciones normales

$$X^T X\beta = X^T\mathbf{y}$$

Ajuste de otras curvas con mínimos cuadrados

Cuando los puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$ en una gráfica de dispersión no se encuentran cerca de una recta, tal vez resulte pertinente postular alguna otra relación funcional entre x y y.

Los siguientes dos ejemplos muestran cómo ajustar datos mediante curvas que tienen la forma general

$$y = \beta_0 f_0(x) + \beta_1 f_1(x) + \cdots + \beta_k f_k(x) \tag{2}$$

donde $f_0, \ldots, f_k$ son funciones conocidas y $\beta_0, \ldots, \beta_k$ son parámetros que se deben determinar. Como se verá, la ecuación (2) describe un modelo lineal porque es lineal en los parámetros desconocidos.

Para un valor particular de x, la ecuación (2) da un valor predicho, o "ajustado", de y. La diferencia entre el valor observado y el valor predicho es el residuo. Los parámetros $\beta_0, \ldots, \beta_k$ se deben determinar para minimizar la suma de los cuadrados de los residuos.

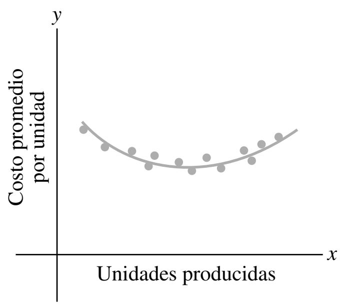


FIGURA 3
Curva de costo promedio.

EJEMPLO 2 Suponga que los puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$ parecen estar sobre una parábola y no sobre una recta. Por ejemplo, si la coordenada x denota el nivel de producción de una compañía, y y representa el costo promedio por unidad de operación en un nivel de x unidades por día, entonces una curva típica de costo promedio parece una parábola que se abre hacia arriba (figura 3). En ecología, se usa una curva parabólica que se abre hacia abajo para modelar la producción primaria neta de nutrientes en una planta, como una función del área superficial del follaje (figura 4). Suponga que deseamos aproximar los datos mediante una ecuación de la forma

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 \tag{3}$$

Describa el modelo lineal que produce un "ajuste de mínimos cuadrados" de los datos, empleando la ecuación (3).

SOLUCIÓN La ecuación (3) describe la relación ideal. Suponga que los valores reales de los parámetros son $\beta_0, \beta_1, \beta_2$. Entonces las coordenadas del primer punto de datos (x_1, y_1) satisfacen una ecuación de la forma

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \epsilon_1$$

donde ϵ_1 es el error residual entre el valor observado y_1 y el valor predicho $\beta_0 + \beta_1 x_1 + \beta_2 x_1^2$ para y. Cada punto de datos determina una ecuación similar:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \epsilon_2 \\ &\vdots \qquad\qquad \vdots \\ y_n &= \beta_0 + \beta_1 x_n + \beta_2 x_n^2 + \epsilon_n \end{aligned}$$

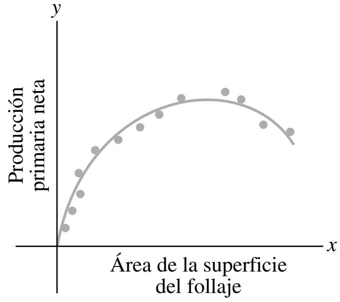


FIGURA 4
Producción de nutrientes.

Es sencillo escribir este sistema de ecuaciones en la forma $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$. Para encontrar X, inspeccione las primeras filas del sistema y observe el patrón.

$$\begin{array}{ccccccc} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} & = & \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} & \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} & + & \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \\ \mathbf{y} & = & X & \beta & + & \boldsymbol{\epsilon} \end{array}$$

■

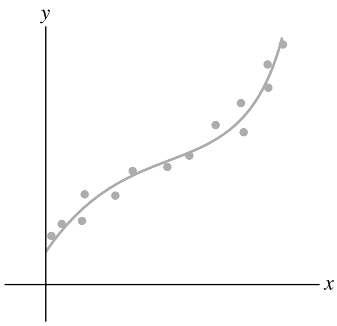


FIGURA 5
Puntos de datos a lo largo de una curva cúbica.

EJEMPLO 3 Si los puntos de datos tienden a seguir un patrón como en la figura 5, entonces un modelo adecuado podría ser una ecuación de la forma

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

Dichos datos podrían ser, por ejemplo, los costos totales de una compañía, como una función del nivel de producción. Describa el modelo lineal que da un ajuste de mínimos cuadrados de este tipo para los datos $(x_1, y_1), \dots, (x_n, y_n)$.

SOLUCIÓN Mediante un análisis similar al efectuado en el ejemplo 2, se obtiene

$$\begin{array}{cccc} \text{Vector de observaciones} & \text{Matriz de diseño} & \text{Vector de parámetros} & \text{Vector residual} \\ \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, & X = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix}, & \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, & \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \end{array}$$

■

Regresión múltiple

Suponga que en un experimento hay dos variables independientes (por ejemplo, u y v) y una variable dependiente, y. Una ecuación sencilla para predecir y a partir de u y v tiene la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v \tag{4}$$

Una ecuación predictiva más general podría tener la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v + \beta_3 u^2 + \beta_4 uv^2 + \beta_5 v^2 \tag{5}$$

Esta ecuación se utiliza en geología, por ejemplo, para modelar superficies de erosión, glaciares, el pH del suelo y otras cantidades. En tales casos, el ajuste por mínimos cuadrados se llama *superficie de tendencia*.

Tanto la ecuación (4) como la (5) conducen a un modelo lineal porque son lineales en los parámetros desconocidos (aun cuando u y v están multiplicados). En general, un modelo lineal surgirá siempre que y se prediga mediante una ecuación de la forma

$$y = \beta_0 f_0(u, v) + \beta_1 f_1(u, v) + \cdots + \beta_k f_k(u, v)$$

con las funciones conocidas $f_0, \dots, f_k$ y los pesos desconocidos $\beta_0, \dots, \beta_k$.

EJEMPLO 4 En geografía, los modelos locales del terreno se construyen mediante los datos $(u_1, v_1, y_1), \dots, (u_n, v_n, y_n)$, donde u_j, v_j y y_j son latitud, longitud y altitud, respectivamente. La solución es el *plano de mínimos cuadrados*. Véase la figura 6.

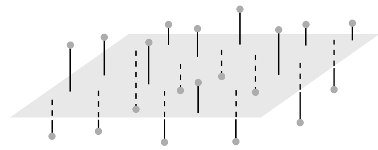


FIGURA 6 Un plano de mínimos cuadrados.

SOLUCIÓN Se espera que los datos satisfagan las siguientes ecuaciones

$$
\begin{aligned}
y_1 &= \beta_0 + \beta_1 u_1 + \beta_2 v_1 + \epsilon_1 \\
y_2 &= \beta_0 + \beta_1 u_2 + \beta_2 v_2 + \epsilon_2 \\
&\ \ \vdots \qquad\qquad\quad \vdots \\
y_n &= \beta_0 + \beta_1 u_n + \beta_2 v_n + \epsilon_n
\end{aligned}
$$

Este sistema tiene la forma matricial $\mathbf{y} = X\beta + \epsilon$, donde

$$
\begin{array}{cccc}
\text{Vector de observaciones} & \text{Matriz de diseño} & \text{Vector de parámetros} & \text{Vector residual} \\
\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, &
X = \begin{bmatrix} 1 & u_1 & v_1 \\ 1 & u_2 & v_2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{bmatrix}, &
\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}, &
\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}
\end{array}
$$

■

El ejemplo 4 revela que el modelo lineal para regresión múltiple tiene la misma forma abstracta que el modelo para regresión simple de los primeros ejemplos. El álgebra lineal nos permite entender el principio general subyacente en todos los modelos lineales. Una vez que X se define adecuadamente, las ecuaciones normales para β tienen la misma forma matricial, sin importar cuántas variables estén implicadas. Así, para cualquier modelo lineal donde X^TX sea invertible, el $\hat{\beta}$ de mínimos cuadrados está dado por $(X^TX)^{-1}X^T\mathbf{y}$.

Lecturas adicionales

Ferguson, J., *Introduction to Linear Algebra in Geology* (Nueva York: Chapman & Hall, 1994).

Krumbein, W. C. y F. A. Graybill, *An Introduction to Statistical Models in Geology* (Nueva York: McGraw-Hill, 1965).

Legendre, P. y L. Legendre, *Numerical Ecology* (Amsterdam: Elsevier, 1998).

Unwin, David J., *An Introduction to Trend Surface Analysis*, Concepts and Techniques in Modern Geography, No. 5 (Norwich, Inglaterra: Geo Books, 1975).

PROBLEMA DE PRÁCTICA

Cuando las ventas mensuales de un producto están sujetas a las fluctuaciones estacionales, entonces una curva que aproxime los datos de venta podría tener la forma

$$y = \beta_0 + \beta_1 x + \beta_2 \operatorname{sen}(2\pi x/12)$$

donde x es el tiempo en meses. El término $\beta_0 + \beta_1 x$ da la tendencia básica de ventas, y el término seno refleja los cambios estacionales en las ventas. Determine la matriz de diseño y el vector de parámetros para el modelo lineal que conduce a un ajuste de mínimos cuadrados de la ecuación anterior. Suponga que los datos son $(x_1, y_1), \ldots, (x_n, y_n)$.

6.6 EJERCICIOS

En los ejercicios 1 a 4, encuentre la ecuación $y = \beta_0 + \beta_1 x$ de la recta de mínimos cuadrados que mejor se ajuste a los puntos de datos indicados.

1. (0, 1), (1, 1), (2, 2), (3, 2)

2. (1, 0), (2, 1), (4, 2), (5, 3)

3. (-1), (0, 1), (1, 2), (2, 4)

4. (2, 3), (3, 2), (5, 1), (6, 0)

5. Sea X la matriz de diseño empleada para determinar la recta de mínimos cuadrados que se ajusta a los datos $(x_1, y_1), \ldots, (x_n, y_n)$. Utilice un teorema de la sección 6.5 para demostrar que las ecuaciones normales tienen solución única si y solo si los datos incluyen al menos dos puntos de datos con diferente coordenada x.

6. Sea X la matriz de diseño del ejemplo 2 correspondiente a un ajuste de mínimos cuadrados de los datos $(x_1, y_1), \ldots, (x_n, y_n)$ a una curva parabólica. Suponga que x_1, x_2 y x_3 son distintas. Explique por qué solo existe una parábola que mejor se ajusta a los datos, en el sentido de mínimos cuadrados. (Véase el ejercicio 5).

7. Un cierto experimento genera los datos (1, 1.8), (2, 2.7), (3, 3.4), (4, 3.8), (5, 3.9). Describa el modelo que produce un ajuste de mínimos cuadrados de esos puntos mediante una función de la forma

$$y = \beta_1 x + \beta_2 x^2$$

Se podría presentar dicha función, por ejemplo, como el ingreso derivado de la venta de x unidades de un producto, cuando la cantidad ofrecida para la venta afecta el precio del producto.

a) Determine la matriz de diseño, el vector de observaciones y el vector de parámetros desconocidos.

b) [**M**] Encuentre la curva de mínimos cuadrados asociada con los datos.

8. Una curva sencilla que con frecuencia es un buen modelo para los costos variables de una compañía, como una función del nivel x de ventas, tiene la forma $y = \beta_1 x + \beta_2 x^2 + \beta_3 x^3$. No existe término constante porque no se incluyen los costos fijos.

a) Determine la matriz de diseño y el vector de parámetros para el modelo lineal que conduce a un ajuste de mínimos cuadrados de la ecuación anterior, con los datos $(x_1, y_1), \ldots, (x_n, y_n)$.

b) [**M**] Encuentre la curva de mínimos cuadrados de la forma anterior para ajustar los datos (4, 1.58), (6, 2.08), (8, 2.5), (10, 2.8), (12, 3.1), (14, 3.4), (16, 3.8) y (18, 4.32), con valores en miles. Si es posible, realice una gráfica que muestre los puntos de datos y la curva de aproximación cúbica.

9. Un cierto experimento genera los datos (1, 7.9), (2, 5.4) y (3, $-.9$). Describa el modelo que da un ajuste de mínimos cuadrados de esos puntos mediante una función de la forma

$$y = A \cos x + B \operatorname{sen} x$$

10. Suponga que las sustancias radiactivas A y B tienen constantes de decaimiento de .02 y .07, respectivamente. Si en el momento $t = 0$ una mezcla de esas dos sustancias contiene M_A gramos de A y M_B gramos de B, entonces un modelo para la cantidad total y de la mezcla presente en el momento t es

$$y = M_A e^{-.02t} + M_B e^{-.07t}$$

Suponga que no se conocen las cantidades iniciales M_A y M_B, pero un científico logra medir las cantidades totales presentes en diferentes momentos y registra los siguientes puntos (t_i, y_i): (10, 21.34), (11, 20.68), (12, 20.05), (14, 18.87) y (15, 18.30).

a) Describa un modelo lineal que se pueda utilizar para estimar M_A y M_B.

b) [**M**] Encuentre la curva de mínimos cuadrados basada en la ecuación (6).

En 1986 fue la última aparición del cometa Halley, el cual reaparecerá en el año 2061.

11. [**M**] De acuerdo con la segunda ley de Kepler, un cometa debería tener una órbita elíptica, parabólica o hiperbólica (ignorando las atracciones gravitacionales de los planetas). En convenientes coordenadas polares, la posición (r, ϑ) de un cometa satisface una ecuación de la forma

$$r = \beta + e(r \cdot \cos \vartheta)$$

donde β es una constante y e es la *excentricidad* de la órbita, con $0 \le e < 1$ para una elipse, $e = 1$ para una parábola, y $e > 1$ para una hipérbola. Suponga que los siguientes datos corresponden a las observaciones de un cometa recién descubierto. Determine el tipo de la órbita e indique dónde estará el cometa cuando $\vartheta = 4.6$ radianes.[3]

ϑ	.88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01

12. [**M**] La presión sanguínea sistólica p (en milímetros de mercurio) de un niño saludable y su peso w (en libras) están relacionados aproximadamente mediante la ecuación

$$\beta_0 + \beta_1 \ln w = p$$

Utilice los siguientes datos experimentales para estimar la presión sanguínea sistólica de un niño saludable que pesa 100 libras.

[3] La idea básica del ajuste de mínimos cuadrados de los datos se debe a K. F. Gauss (e, independientemente, a A. Legendre), cuya fama despuntó en 1801 cuando utilizó el método para obtener la trayectoria del asteroide *Ceres*. Cuarenta días después de que el asteroide fue descubierto, desapareció detrás del Sol. Gauss predijo que Ceres aparecería 10 meses después y precisó su ubicación. La exactitud de la predicción asombró a la comunidad científica europea.

w	44	61	81	113	131
$\ln w$	3.78	4.11	4.39	4.73	4.88
p	91	98	103	110	112

13. [**M**] Para medir el desempeño de un avión durante el despegue, cada segundo se midió su posición horizontal, de $t = 0$ a $t = 12$. Las posiciones (en pies) fueron: 0, 8.8, 29.9, 62.0, 104.7, 159.1, 222.0, 294.5, 380.4, 471.1, 571.7, 686.8, y 809.2.

a) Encuentre la curva cúbica de mínimos cuadrados $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ para esos datos.

b) Con base en el resultado del inciso *a*), estime la velocidad del avión cuando $t = 4.5$ segundos.

14. Sean $\overline{x} = \frac{1}{n}(x_1 + \cdots + x_n)$ y $\overline{y} = \frac{1}{n}(y_1 + \cdots + y_n)$. Demuestre que la recta de mínimos cuadrados para los datos $(x_1, y_1), \ldots, (x_n, y_n)$ debe pasar a través de $(\bar{x}, \bar{y})$. Es decir, demuestre que $\bar{x}$ y $\bar{y}$ satisfacen la ecuación lineal $\bar{y} = \beta_0 + \beta_1\bar{x}$. [*Sugerencia:* Deduzca esta ecuación mediante la ecuación vectorial $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$. Denote la primera columna de X con $\mathbf{1}$. Con base en el hecho de que el vector residual $\boldsymbol{\epsilon}$ es ortogonal al espacio columna de X y, por lo tanto, es ortogonal a $\mathbf{1}$].

Considerando los datos para un problema de mínimos cuadrados, $(x_1, y_1), \ldots, (x_n, y_n)$, son útiles las siguientes abreviaciones:

$$\textstyle\sum x = \sum_{i=1}^n x_i, \quad \sum x^2 = \sum_{i=1}^n x_i^2,$$
$$\textstyle\sum y = \sum_{i=1}^n y_i, \quad \sum xy = \sum_{i=1}^n x_i y_i$$

Las ecuaciones normales para una recta de mínimos cuadrados $y = \hat{\beta}_0 + \hat{\beta}_1 x$ se pueden escribir en la forma

$$\begin{aligned} n\hat{\beta}_0 + \hat{\beta}_1 \textstyle\sum x &= \textstyle\sum y \\ \hat{\beta}_0 \textstyle\sum x + \hat{\beta}_1 \textstyle\sum x^2 &= \textstyle\sum xy \end{aligned} \tag{7}$$

15. Deduzca las ecuaciones normales (7) a partir de la forma matricial presentada en esta sección.

16. Utilice una matriz inversa para resolver el sistema de ecuaciones (7) y así obtener las fórmulas para $\hat{\beta}_0$ y $\hat{\beta}_1$ que aparecen en muchos libros de estadística.

17. *a*) Rescriba los datos del ejemplo 1 con nuevas coordenadas x en la forma de desviación media. Sea X la matriz de diseño asociada. ¿Por qué son ortogonales las columnas de X?

b) Escriba las ecuaciones normales para los datos del inciso *a*), y resuélvalas para encontrar la recta de mínimos cuadrados, $y = \beta_0 + \beta_1 x^*$, donde $x^* = x - 5.5$.

18. Suponga que las coordenadas x de los datos $(x_1, y_1), \ldots, (x_n, y_n)$ están en la forma de desviación media, así que $\sum x_i = 0$. Demuestre que si X es la matriz de diseño para la recta de mínimos cuadrados de este caso, entonces X^TX es una matriz diagonal.

Los ejercicios 19 y 20 implican a una matriz de diseño X con dos o más columnas y una solución de mínimos cuadrados $\hat{\beta}$ de $\mathbf{y} = X\beta$. Considere los siguientes números.

i. $\|X\hat{\beta}\|^2$, la suma de los cuadrados del "término de regresión". Denote este número con SS(R).

ii. $\|\mathbf{y} - X\hat{\beta}\|^2$, la suma de los cuadrados para el término de error. Denote este número con SS(E).

iii. $\|\mathbf{y}\|^2$, la suma "total" de los cuadrados de los valores y. Denote este número con SS(T).

Todos los libros de estadística analizan el tema de la regresión, y el modelo lineal $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$ introduce esos números, aunque la notación y la terminología tal vez varíen de un texto a otro. Para simplificar el asunto, suponga que la media de los valores y es cero. En este caso, SS(T) es proporcional a la *varianza* del conjunto de valores y.

19. Justifique la ecuación SS(T) = SS(R) + SS(E). [*Sugerencia:* Utilice un teorema, y explique por qué se satisfacen las hipótesis del teorema]. Esta ecuación es extremadamente importante en estadística, en la teoría de regresión y en el análisis de varianza.

20. Demuestre que $\|X\hat{\beta}\|^2 = \hat{\beta}^T X^T \mathbf{y}$. [*Sugerencia:* Rescriba el lado izquierdo y considere el hecho de que $\hat{\beta}$ satisface las ecuaciones normales]. Esta fórmula para SS(R) se emplea en estadística. Con base en esto y en el ejercicio 19, obtenga la fórmula estándar para SS(E):

$$\text{SS(E)} = \mathbf{y}^T\mathbf{y} - \hat{\beta}^T X^T \mathbf{y}$$

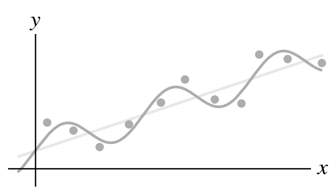
SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Construya X y β tal que la k-ésima fila de $X\beta$ sea el valor de y predicho correspondiente al punto de datos (x_k, y_k), a saber,

$$\beta_0 + \beta_1 x_k + \beta_2 \operatorname{sen}(2\pi\, x_k/12)$$

Debería ser claro que

$$X = \begin{bmatrix} 1 & x_1 & \operatorname{sen}(2\pi x_1/12) \\ \vdots & \vdots & \vdots \\ 1 & x_n & \operatorname{sen}(2\pi x_n/12) \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$



Tendencia de ventas con fluctuaciones estacionales.

6.7 ESPACIOS CON PRODUCTO INTERIOR

Los conceptos de longitud, distancia y ortogonalidad son esenciales en aplicaciones que implican un espacio vectorial. Para $\mathbb{R}^n$, esos conceptos se basaron en las propiedades del producto interior indicadas en el teorema 1 de la sección 6.1. Para otros espacios, se necesitan analogías del producto interior con las mismas propiedades. Ahora las conclusiones del teorema 1 se convierten en *axiomas* en la siguiente definición.

DEFINICIÓN

Un **producto interior** sobre un espacio vectorial V es una función que asocia un número real $\langle \mathbf{u}, \mathbf{v} \rangle$ para cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ en V, y satisface los siguientes axiomas, para toda $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ en V, y todos los escalares c:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \quad$ y $\quad \langle \mathbf{u}, \mathbf{u} \rangle = 0$ si y solo si $\mathbf{u} = \mathbf{0}$

Un espacio vectorial con un producto interior se llama **espacio con producto interior**.

El espacio vectorial $\mathbb{R}^n$ con el producto interior estándar es un espacio con producto interior, y casi todo lo analizado en este capítulo para $\mathbb{R}^n$ también es válido para los espacios con producto interior. Los ejemplos en esta sección y la siguiente establecen el fundamento para una variedad de aplicaciones que se estudian en cursos de ingeniería, física, matemáticas y estadística.

EJEMPLO 1 Elija dos números positivos cualesquiera, por ejemplo, 4 y 5; para los vectores $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$ en $\mathbb{R}^2$, establezca

$$\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 5u_2v_2 \tag{1}$$

Demuestre que la ecuación (1) define un producto interior.

SOLUCIÓN Sin duda, el axioma 1 se satisface, porque $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = (v, u)$. Si $\mathbf{w} = (w_1, w_2)$, entonces,

$$\begin{aligned}\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle &= 4(u_1 + v_1)w_1 + 5(u_2 + v_2)w_2 \\ &= 4u_1w_1 + 5u_2w_2 + 4v_1w_1 + 5v_2w_2 \\ &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle\end{aligned}$$

Esto comprueba el axioma 2. Para el axioma 3, calcule

$$\langle c\mathbf{u}, \mathbf{v} \rangle = 4(cu_1)v_1 + 5(cu_2)v_2 = c(4u_1v_1 + 5u_2v_2) = c\langle \mathbf{u}, \mathbf{v} \rangle$$

Para el axioma 4, observe que $\langle \mathbf{u}, \mathbf{u} \rangle = 4u_1^2 + 5u_2^2 \geq 0$ y $4u_1^2 + 5u_2^2 = 0$ solo si $u_1 = u_2 = 0$, es decir, si $\mathbf{u} = \mathbf{0}$. Además, $\langle \mathbf{0}, \mathbf{0} \rangle = 0$. Así, (1) define un producto interior sobre $\mathbb{R}^2$. ■

Sobre $\mathbb{R}^n$ se pueden definir productos interiores similares a (1), que surgen naturalmente en problemas de "mínimos cuadrados ponderados", en los cuales los pesos se asignan a las diversas entradas en la suma para el producto interior de manera que se dé mayor importancia a las mediciones más confiables.

De ahora en adelante, cuando un espacio con producto interior implique polinomios u otras funciones, se escribirán las funciones en la forma usual, en vez de emplear negritas para los vectores. Sin duda, es importante recordar que cada función *es* un vector cuando se trata como un elemento de un espacio vectorial.

EJEMPLO 2 Sean $t_0, \ldots, t_n$ números reales distintos. Para p y q en $\mathbb{P}_n$, defina

$$\langle p, q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n) \tag{2}$$

Es sencillo comprobar los axiomas 1 a 3 del producto interior. Para el axioma 4, observe que

$$\langle p, p\rangle = [p(t_0)]^2 + [p(t_1)]^2 + \cdots + [p(t_n)]^2 \geq 0$$

Además, $\langle \mathbf{0}, \mathbf{0}\rangle = 0$. (Aquí el cero en "negritas" denota el polinomio cero, el vector cero en $\mathbb{P}_n$). Si $\langle p, p\rangle = 0$, entonces p se debe anular en $n + 1$ puntos: $t_0, \ldots, t_n$. Esto solo es posible si p es el polinomio cero, porque el grado de p es menor que $n + 1$. Así, (2) define un producto interior en $\mathbb{P}_n$. ∎

EJEMPLO 3 Considere que V está en $\mathbb{P}_2$, con el producto interior del ejemplo 2, donde $t_0 = 0$, $t_1 = \frac{1}{2}$, y $t_2 = 1$. Sean $p(t) = 12t^2$ y $q(t) = 2t - 1$. Calcule $\langle p, q\rangle$ y $\langle q, q\rangle$.

SOLUCIÓN

$$\begin{aligned}\langle p, q\rangle &= p(0)q(0) + p\left(\tfrac{1}{2}\right)q\left(\tfrac{1}{2}\right) + p(1)q(1)\\ &= (0)(-1) + (3)(0) + (12)(1) = 12\\ \langle q, q\rangle &= [q(0)]^2 + [q\left(\tfrac{1}{2}\right)]^2 + [q(1)]^2\\ &= (-1)^2 + (0)^2 + (1)^2 = 2\end{aligned}$$ ∎

Longitudes, distancias y ortogonalidad

Sea V un espacio con producto interior, con el producto interior denotado con $\langle \mathbf{u}, \mathbf{v}\rangle$. Igual que en $\mathbb{R}^n$, defina la **longitud**, o **norma**, de un vector $\mathbf{v}$ como el escalar

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$$

De manera equivalente, $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle$. (Esta definición tiene sentido porque $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$, pero la definición *no* dice que $\langle \mathbf{v}, \mathbf{v}\rangle$ sea una "suma de cuadrados" porque $\mathbf{v}$ no necesita ser un elemento de $\mathbb{R}^n$).

Un **vector unitario** es aquel cuya longitud es 1. La **distancia entre u y v** es $\|\mathbf{u} - \mathbf{v}\|$. Los vectores $\mathbf{u}$ y $\mathbf{v}$ son ortogonales si $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

EJEMPLO 4 Suponga que $\mathbb{P}_2$ tiene el producto interior (2) del ejemplo 3. Calcule las longitudes de los vectores $p(t) = 12t^2$ y $q(t) = 2t - 1$.

SOLUCIÓN

$$\begin{aligned}\|p\|^2 &= \langle p, p\rangle = [p(0)]^2 + \left[p\left(\tfrac{1}{2}\right)\right]^2 + [p(1)]^2\\ &= 0 + [3]^2 + [12]^2 = 153\\ \|p\| &= \sqrt{153}\end{aligned}$$

Del ejemplo 3, $\langle q, q\rangle = 2$. Por lo tanto, $\|q\| = \sqrt{2}$. ∎

Proceso de Gram-Schmidt

La existencia de bases ortogonales para subespacios de dimensión finita de un espacio vectorial con producto interior se puede establecer con el proceso de Gram-Schmidt, al igual que en $\mathbb{R}^n$. Mediante este proceso, es posible construir ciertas bases ortogonales que surgen a menudo en las aplicaciones.

La proyección ortogonal de un vector sobre un subespacio W con una base ortogonal se puede construir en la forma habitual. La proyección no depende de la base ortogonal seleccionada, y tiene las propiedades descritas en los teoremas de descomposición ortogonal y de la mejor aproximación.

EJEMPLO 5 Considere que V está en P_4 con el producto interior del ejemplo 2, lo que implica evaluación de polinomios en $-2, -1, 0, 1$ y 2, y que $\mathbb{P}_2$ es un subespacio de V. Obtenga una base ortogonal para $\mathbb{P}_2$ aplicando el proceso de Gram-Schmidt a los polinomios 1, t y t^2.

SOLUCIÓN El producto interior sólo depende de los valores de un polinomio en $-2,\dots, 2$, así que se listan los valores de cada polinomio como un vector en $\mathbb{R}^5$, debajo del nombre de cada polinomio:[1]

$$\begin{array}{lccc} \text{Polinomio:} & 1 & t & t^2 \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix} \end{array}$$

El producto interior de dos polinomios en V es igual al producto interior (estándar) de sus vectores correspondientes en $\mathbb{R}^5$. Observe que t es ortogonal a la función constante 1. Así, se toman $p_0(t) = 1$ y $p_1(t) = t$. Para p_2, utilice los vectores en $\mathbb{R}^5$ para calcular la proyección de t^2 sobre Gen $\{p_0, p_1\}$:

$$\begin{aligned} \langle t^2, p_0\rangle &= \langle t^2, 1\rangle = 4 + 1 + 0 + 1 + 4 = 10 \\ \langle p_0, p_0\rangle &= 5 \\ \langle t^2, p_1\rangle &= \langle t^2, t\rangle = -8 + (-1) + 0 + 1 + 8 = 0 \end{aligned}$$

La proyección ortogonal de t^2 sobre Gen $\{1, t\}$ es $\frac{10}{5}p_0 + 0p_1$. Por consiguiente,

$$p_2(t) = t^2 - 2p_0(t) = t^2 - 2$$

Una base ortogonal para el subespacio $\mathbb{P}_2$ de V es:

$$\begin{array}{lccc} \text{Polinomio:} & p_0 & p_1 & p_2 \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix} \end{array} \tag{3}$$

■

Mejor aproximación en espacios con producto interior

Un problema común en matemáticas aplicadas implica un espacio vectorial V cuyos elementos son funciones. El problema es aproximar una función f en V mediante una función g de un subespacio especificado W de V. La "cercanía" de la aproximación de f depende de la manera en que se defina $\|f - g\|$. Solo se considerará el caso en que la distancia entre f y g esté determinada por un producto interior. En tal caso, la *mejor aproximación a f mediante funciones en W* es la proyección ortogonal de f sobre el subespacio W.

EJEMPLO 6 Considere que V está en $\mathbb{P}_4$ con el producto interior del ejemplo 5; considere también que p_0, p_1 y p_2 constituyen la base ortogonal que se encontró en el ejemplo 5 para el subespacio $\mathbb{P}_2$. Determine la mejor aproximación a $p(t) = 5 - \frac{1}{2}t^4$ mediante polinomios en $\mathbb{P}_2$.

[1] Cada polinomio en $\mathbb{P}_4$ está determinado de manera unívoca por su valor en los cinco números $-2,\dots, 2$. De hecho, la correspondencia entre p y su vector de valores es un isomorfismo, es decir, un mapeo uno a uno sobre $\mathbb{R}^5$ que preserva combinaciones lineales.

SOLUCIÓN Los valores de p_0, p_1 y p_2 en los números $-2, -1, 0, 1$ y 2 están listados en los vectores de $\mathbb{R}^5$ en (3). Los valores correspondientes para p son -3, $9/2$, 5, $9/2$ y -3. Calcule

$$\begin{array}{lll} \langle p, p_0\rangle = 8, & \langle p, p_1\rangle = 0, & \langle p, p_2\rangle = -31 \\ \langle p_0, p_0\rangle = 5, & & \langle p_2, p_2\rangle = 14 \end{array}$$

Así, la mejor aproximación en V a p mediante polinomios en $\mathbb{P}_2$ es

$$\begin{aligned} \hat{p} = \text{proy}_{\mathbb{P}_2} p &= \frac{\langle p, p_0\rangle}{\langle p_0, p_0\rangle} p_0 + \frac{\langle p, p_1\rangle}{\langle p_1, p_1\rangle} p_1 + \frac{\langle p, p_2\rangle}{\langle p_2, p_2\rangle} p_2 \\ &= \tfrac{8}{5} p_0 + \tfrac{-31}{14} p_2 = \tfrac{8}{5} - \tfrac{31}{14}(t^2 - 2). \end{aligned}$$

De todos los polinomios en $\mathbb{P}_2$, este es el más cercano a p cuando la distancia entre polinomios se mide solo en $-2, -1, 0, 1$ y 2. Véase la figura 1. ■

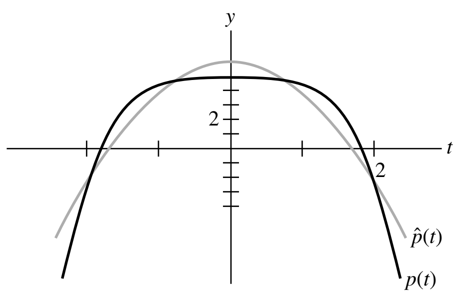


FIGURA 1

Los polinomios p_0, p_1 y p_2 de los ejemplos 5 y 6 pertenecen a una clase de polinomios que en estadística se denominan *polinomios ortogonales*.[2] La ortogonalidad se refiere al tipo de producto interior descrito en el ejemplo 2.

Dos desigualdades

Considerando un vector **v** en un espacio V con producto interior y dado un subespacio W de dimensión finita, se aplica el teorema de Pitágoras a la descomposición ortogonal de **v** con respecto a W y se obtiene

$$\|\mathbf{v}\|^2 = \|\text{proy}_W \mathbf{v}\|^2 + \|\mathbf{v} - \text{proy}_W \mathbf{v}\|^2$$

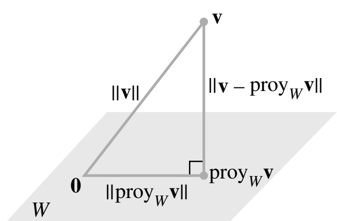


FIGURA 2
La hipotenusa es el lado más largo.

Véase la figura 2. En particular, esto muestra que la norma de la proyección de **v** sobre W no excede la propia norma de **v**. Esta sencilla observación conduce a la siguiente importante desigualdad.

TEOREMA 16

Desigualdad de Cauchy-Schwarz

Para toda **u** y **v** en V,

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\| \tag{4}$$

[2] Véase *Statistics and Experimental Design in Engineering and the Physical Sciences*, 2a. ed., de Norman L. Johnson y Fred C. Leone (Nueva York: John Wiley & Sons, 1977). En este libro las tablas listan los "polinomios ortogonales", que son simplemente los valores de los polinomios en números como $-2, -1, 0, 1$ y 2.

DEMOSTRACIÓN Si $\mathbf{u} = \mathbf{0}$, entonces ambos lados de (4) son iguales a cero, por lo tanto, la desigualdad es verdadera en este caso. (Véase el problema de práctica 1). Si $\mathbf{u} \neq \mathbf{0}$, sea W el subespacio generado por $\mathbf{u}$. Recuerde que $\|c\mathbf{u}\| = |c|\,\|\mathbf{u}\|$ para cualquier escalar c. Así,

$$\|\text{proy}_W\,\mathbf{v}\| = \left\|\frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}\right\| = \frac{|\langle \mathbf{v}, \mathbf{u}\rangle|}{|\langle \mathbf{u}, \mathbf{u}\rangle|}\|\mathbf{u}\| = \frac{|\langle \mathbf{v}, \mathbf{u}\rangle|}{\|\mathbf{u}\|^2}\|\mathbf{u}\| = \frac{|\langle \mathbf{u}, \mathbf{v}\rangle|}{\|\mathbf{u}\|}$$

Puesto que $\|\text{proy}_W\,\mathbf{v}\| \leq \|\mathbf{v}\|$, se tiene $\dfrac{|\langle \mathbf{u}, \mathbf{v}\rangle|}{\|\mathbf{u}\|} \leq \|\mathbf{v}\|$, lo que da la ecuación (4). ■

La desigualdad de Cauchy-Schwarz es útil en muchas ramas de las matemáticas. En los ejercicios se presentan unas cuantas aplicaciones. Esta desigualdad es necesaria para probar otra desigualdad fundamental que implica normas de vectores. Véase la figura 3.

TEOREMA 17

Desigualdad del triángulo

Para toda $\mathbf{u}$ y $\mathbf{v}$ en V,

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

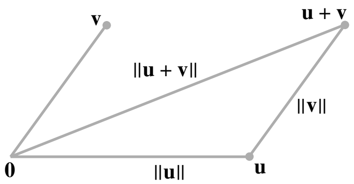


FIGURA 3
Las longitudes de los lados de un triángulo.

DEMOSTRACIÓN

$$\begin{aligned}
\|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\
&\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 \qquad \text{Cauchy–Schwarz} \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned}$$

La desigualdad del triángulo se deduce inmediatamente al sacar la raíz cuadrada en ambos lados. ■

Un producto interior para $C[a, b]$ (se requiere cálculo)

Quizás el espacio con producto interior más ampliamente utilizado en aplicaciones sea el espacio vectorial $C[a, b]$ de todas las funciones continuas en un intervalo $a \leq t \leq b$, con un producto interior que enseguida se describirá.

Se inicia considerando un polinomio p y cualquier entero n mayor o igual que el grado de p. Entonces p está en $\mathbb{P}_n$, y se puede calcular una "longitud" para p utilizando el producto interior del ejemplo 2 que implica evaluación en $n + 1$ puntos en $[a, b]$. Sin embargo, esta longitud de p solo capta el comportamiento en esos $n + 1$ puntos. Como p está en $\mathbb{P}_n$ para todas las n grandes, se podría utilizar un valor de n bastante grande, con mucho más puntos para la "evaluación" asociada con el producto interior. Véase la figura 4.

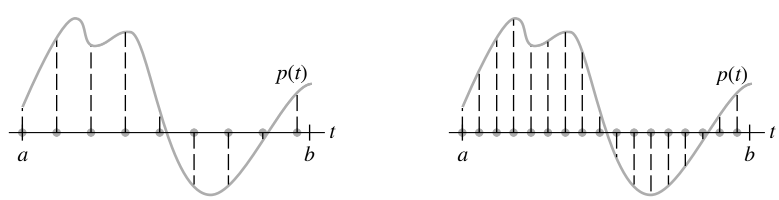
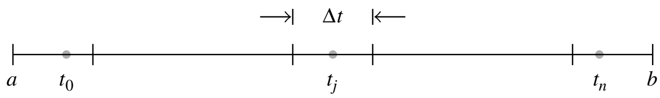


FIGURA 4 Uso de diferentes números de puntos de evaluación en $[a, b]$ para calcular $\|p\|^2$.

Se particiona $[a, b]$ en $n + 1$ subintervalos de longitud $\Delta t = (b - a)/(n + 1)$, y sean $t_0, \ldots, t_n$ puntos arbitrarios en esos subintervalos.



Si n es grande, el producto interior sobre $\mathbb{P}_n$ determinado por $t_0, \ldots, t_n$ tenderá a dar un gran valor para $\langle p, p \rangle$, así que se reduce a escala y se divide entre $n + 1$. Observe que $1/(n + 1) = \Delta t/(b - a)$, y se define

$$\langle p, q \rangle = \frac{1}{n+1} \sum_{j=0}^{n} p(t_j)q(t_j) = \frac{1}{b-a} \left[\sum_{j=0}^{n} p(t_j)q(t_j)\Delta t \right]$$

Ahora, n crece sin límite. Puesto que los polinomios p y q son funciones continuas, entonces la expresión entre corchetes es una suma de Riemann que se aproxima a una integral definida, lo que conduce a considerar el *valor promedio de* $p(t)q(t)$ sobre el intervalo $[a, b]$:

$$\frac{1}{b-a} \int_a^b p(t)q(t)\,dt$$

Esta cantidad está definida para polinomios de cualquier grado (de hecho, para todas las funciones continuas), y tiene todas las propiedades de un producto interior, como lo muestra el siguiente ejemplo. El factor de escala $1/(b - a)$ no es esencial y con frecuencia se omite con la finalidad de simplificar.

EJEMPLO 7 Para f, g en $C[a, b]$, sea

$$\langle f, g \rangle = \int_a^b f(t)g(t)\,dt \tag{5}$$

Demuestre que la ecuación (5) define un producto interior sobre $C[a, b]$.

SOLUCIÓN Los axiomas 1 a 3 del producto interior se deducen de las propiedades elementales de las integrales definidas. Para el axioma 4, observe que

$$\langle f, f \rangle = \int_a^b [f(t)]^2\,dt \geq 0$$

La función $[f(t)]^2$ es continua y no negativa en $[a, b]$. Si la integral definida de $[f(t)]^2$ es cero, entonces $[f(t)]^2$ debe ser idénticamente cero sobre $[a, b]$, de acuerdo con un teorema en cálculo avanzado; en tal caso, f es la función cero. Así $\langle f, f \rangle = 0$ implica que f es la función cero en $[a, b]$. Por lo tanto, (5) define un producto interior sobre $C[a, b]$. ■

EJEMPLO 8 Sean V el espacio $C[0, 1]$ con el producto interior del ejemplo 7, y W el subespacio generado por los polinomios $p_1(t) = 1$, $p_2(t) = 2t - 1$ y $p_3(t) = 12t^2$. Utilice el proceso de Gram-Schmidt y encuentre una base ortogonal para W.

SOLUCIÓN Sea $q_1 = p_1$, y calcule

$$\langle p_2, q_1 \rangle = \int_0^1 (2t - 1)(1)\,dt = (t^2 - t)\Big|_0^1 = 0$$

De manera que p_2 ya es ortogonal a q_1, y se puede tomar $q_2 = p_2$. Para la proyección de p_3 sobre $W_2 = \text{Gen}\{q_1, q_2\}$, se calcula

$$\begin{aligned}
\langle p_3, q_1\rangle &= \int_0^1 12t^2 \cdot 1\, dt = 4t^3\Big|_0^1 = 4 \\
\langle q_1, q_1\rangle &= \int_0^1 1 \cdot 1\, dt = t\Big|_0^1 = 1 \\
\langle p_3, q_2\rangle &= \int_0^1 12t^2(2t-1)\, dt = \int_0^1 (24t^3 - 12t^2)\, dt = 2 \\
\langle q_2, q_2\rangle &= \int_0^1 (2t-1)^2\, dt = \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}
\end{aligned}$$

Entonces

$$\text{proy}_{W_2}\, p_3 = \frac{\langle p_3, q_1\rangle}{\langle q_1, q_1\rangle} q_1 + \frac{\langle p_3, q_2\rangle}{\langle q_2, q_2\rangle} q_2 = \frac{4}{1} q_1 + \frac{2}{1/3} q_2 = 4q_1 + 6q_2$$

y

$$q_3 = p_3 - \text{proy}_{W_2}\, p_3 = p_3 - 4q_1 - 6q_2$$

Como una función, $q_3(t) = 12t^2 - 4 - 6(2t-1) = 12t^2 - 12t + 2$. La base ortogonal para el subespacio W es $\{q_1, q_2, q_3\}$. ■

PROBLEMAS DE PRÁCTICA

Aplique los axiomas del producto interior para comprobar los siguientes enunciados.

1. $\langle \mathbf{v}, \mathbf{0}\rangle = \langle \mathbf{0}, \mathbf{v}\rangle = 0$.

2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$.

6.7 EJERCICIOS

1. Suponga que $\mathbb{R}^2$ tiene el producto interior del ejemplo 1, y sean $\mathbf{x} = (1, 1)$ y $\mathbf{y} = (5, -1)$.

a) Encuentre $\|\mathbf{x}\|$, $\|\mathbf{y}\|$ y $|\langle \mathbf{x}, \mathbf{y}\rangle|^2$.

b) Describa todos los vectores (z_1, z_2) que son ortogonales a $\mathbf{y}$.

2. Considere que $\mathbb{R}^2$ tiene el producto interior del ejemplo 1. Demuestre que la desigualdad de Cauchy-Schwarz es válida para $\mathbf{x} = (3, -2)$ y $\mathbf{y} = (-2, 1)$. (*Sugerencia:* Estudie $|\langle \mathbf{x}, \mathbf{y}\rangle|^2$).

Los ejercicios 3 a 8 se refieren a $\mathbb{P}_2$ con el producto interior dado por evaluación en -1, 0 y 1. (Véase el ejemplo 2).

3. Calcule $\langle p, q\rangle$, donde $p(t) = 4 + t$, $q(t) = 5 - 4t^2$.

4. Calcule $\langle p, q\rangle$, donde $p(t) = 3t - t^2$, $q(t) = 3 + 2t^2$.

5. Obtenga $\|p\|$ y $\|q\|$, para p y q del ejercicio 3.

6. Obtenga $\|p\|$ y $\|q\|$, para p y q del ejercicio 4.

7. Determine la proyección ortogonal de q sobre el subespacio generado por p, para p y q en el ejercicio 3.

8. Obtenga la proyección ortogonal de q sobre el subespacio generado por p, para p y q en el ejercicio 4.

9. Considere que $\mathbb{P}_3$ tiene el producto interior dado por evaluación en -3, -1, 1 y 3. Sean $p_0(t) = 1$, $p_1(t) = t$ y $p_2(t) = t^2$.

a) Calcule la proyección ortogonal de p_2 sobre el subespacio generado por p_0 y p_1.

b) Encuentre un polinomio q que sea ortogonal a p_0 y p_1, tal que $\{p_0, p_1, q\}$ sea una base ortogonal para $\text{Gen}\{p_0, p_1, p_2\}$. Escale el polinomio q de modo que su vector de valores en $(-3, -1, 1, 3)$ sea $(1, -1, -1, 1)$.

10. Considere que $\mathbb{P}_3$ tiene el producto interior del ejercicio 9, con los polinomios p_0, p_1 y q ahí descritos. Determine la mejor aproximación a $p(t) = t^3$ mediante polinomios en $\text{Gen}\{p_0, p_1, q\}$.

11. Suponga que p_0, p_1 y p_2 son los polinomios ortogonales descritos en el ejemplo 5, donde el producto interior sobre $\mathbb{P}_4$ está dado por evaluación en -2, -1, 0, 1 y 2. Encuentre la proyección ortogonal de t^3 sobre $\text{Gen}\{p_0, p_1, p_2\}$.

12. Obtenga un polinomio p_3 tal que $\{p_0, p_1, p_2, p_3\}$ (véase el ejercicio 11) sea una base ortogonal para el subespacio $\mathbb{P}_3$ de $\mathbb{P}_4$. Escale el polinomio p_3 de manera que su vector de valores sea $(-1, 2, 0, -2, 1)$.

13. Sea A una matriz invertible de $n \times n$. Demuestre que para $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^n$, la fórmula $\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u}) \cdot (A\mathbf{v}) = (A\mathbf{u})^T(A\mathbf{v})$ define un producto interior sobre $\mathbb{R}^n$.

14. Sea T una transformación lineal uno a uno de un espacio vectorial V en $\mathbb{R}^n$. Demuestre que para $\mathbf{u}$, $\mathbf{v}$ en V, la fórmula $\langle \mathbf{u}, \mathbf{v} \rangle = T(\mathbf{u}) \cdot T(\mathbf{v})$ define un producto interior sobre V.

Utilice los axiomas del producto interior y otros resultados de esta sección para comprobar los enunciados de los ejercicios 15 a 18.

15. $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ para todos los escalares c.

16. Si $\{\mathbf{u}, \mathbf{v}\}$ es un conjunto ortonormal en V, entonces $\|\mathbf{u} - \mathbf{v}\| = \sqrt{2}$.

17. $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$

18. $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$

19. Dados $a \geq 0$ y $b \geq 0$, sean $\mathbf{u} = \begin{bmatrix} \sqrt{a} \\ \sqrt{b} \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} \sqrt{b} \\ \sqrt{a} \end{bmatrix}$.

Aplique la desigualdad de Cauchy-Schwarz para comparar la media geométrica $\sqrt{ab}$ con la media aritmética $(a + b)/2$.

20. Sean $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Utilice la desigualdad de Cauchy-Schwarz para demostrar que

$$\left(\frac{a+b}{2}\right)^2 \leq \frac{a^2 + b^2}{2}$$

Los ejercicios 21 a 24 se refieren a $V = C[0, 1]$, con el producto interior dado por una integral, como en el ejemplo 7.

21. Calcule $\langle f, g \rangle$, donde $f(t) = 1 - 3t^2$ y $g(t) = t - t^3$.

22. Determine $\langle f, g \rangle$, donde $f(t) = 5t - 3$ y $g(t) = t^3 - t^2$.

23. Obtenga $\|f\|$ para la f del ejercicio 21.

24. Calcule $\|g\|$ para la g del ejercicio 22.

25. Sea V el espacio $C[-1, 1]$ con el producto interior del ejemplo 7. Encuentre una base ortogonal para el subespacio generado por los polinomios 1, t y t^2. Los polinomios en esta base se llaman *polinomios de Legendre*.

26. Sea V el espacio $C[-2, 2]$ con el producto interior del ejemplo 7. Obtenga una base ortogonal para el subespacio generado por los polinomios 1, t y t^2.

27. [**M**] Considere que $\mathbb{P}_4$ tiene el producto interior como en el ejemplo 5, y sean p_0, p_1, p_2 los polinomios ortogonales de ese ejemplo. Con su programa de matrices, aplique el proceso de Gram-Schmidt al conjunto $\{p_0, p_1, p_2, t^3, t^4\}$ y cree una base ortogonal para $\mathbb{P}_4$.

28. [**M**] Sea V el espacio $C[0, 2\pi]$ con el producto interior del ejemplo 7. Aplique el proceso de Gram-Schmidt con la finalidad de crear una base ortogonal para el subespacio generado por $\{1, \cos t, \cos^2 t, \cos^3 t\}$. Utilice un programa de matrices o computacional para calcular las integrales definidas adecuadas.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Por el axioma 1, $\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle$. Entonces $\langle \mathbf{0}, \mathbf{v} \rangle = \langle 0\mathbf{v}, \mathbf{v} \rangle = 0\langle \mathbf{v}, \mathbf{v} \rangle$, por el axioma 3, de manera que $\langle \mathbf{0}, \mathbf{v} \rangle = 0$.
2. Por los axiomas 1 y 2, y después 1 otra vez, $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.

6.8 APLICACIONES DE ESPACIOS CON PRODUCTO INTERIOR

Los ejemplos de esta sección ilustran cómo surgen en problemas prácticos los espacios con producto interior definidos en la sección 6.7. El primer ejemplo se relaciona con el inmenso problema de mínimos cuadrados de actualización del North American Datum, descrito en el ejemplo introductorio de este capítulo.

Mínimos cuadrados ponderados

Sea $\mathbf{y}$ un vector de n observaciones, $y_1, \ldots, y_n$, y suponga que se desea aproximar $\mathbf{y}$ mediante un vector $\hat{\mathbf{y}}$ que pertenece a determinado subespacio de $\mathbb{R}^n$. (En la sección 6.5, $\hat{\mathbf{y}}$ se escribió como $A\mathbf{x}$, de manera que $\hat{\mathbf{y}}$ estuviera en el espacio columna de A). Denote las entradas en $\hat{\mathbf{y}}$ como $\hat{y}_1, \ldots, \hat{y}_n$. Entonces la *suma de los cuadrados de los errores*, o SS(E), al aproximar $\mathbf{y}$ por $\hat{\mathbf{y}}$ es

$$\text{SS(E)} = (y_1 - \hat{y}_1)^2 + \cdots + (y_n - \hat{y}_n)^2 \tag{1}$$

Esto es simplemente $\|\mathbf{y} - \hat{\mathbf{y}}\|^2$, utilizando la longitud estándar en $\mathbb{R}^n$.

Ahora suponga que las mediciones que generaron las entradas en **y** no son igualmente confiables. (Este fue el caso para el North American Datum, porque las mediciones se realizaron durante un periodo de 140 años). Como otro ejemplo, las entradas en **y** se podrían calcular a partir de varias muestras de medidas, con tamaños de muestra desiguales). Entonces, resulta adecuado ponderar los errores cuadráticos de la ecuación (1) de tal manera que se dé mayor importancia a las mediciones más confiables.[1] Si los pesos se denotan como $w_1^2, \ldots, w_n^2$, entonces la suma ponderada de los errores cuadráticos es

$$\text{SS(E) ponderada} = w_1^2(y_1 - \hat{y}_1)^2 + \cdots + w_n^2(y_n - \hat{y}_n)^2 \tag{2}$$

Este es el cuadrado de la longitud de $\mathbf{y} - \hat{\mathbf{y}}$, donde la longitud se deduce a partir de un producto interior análogo al del ejemplo 1 de la sección 6.7, a saber,

$$\langle \mathbf{x}, \mathbf{y} \rangle = w_1^2 x_1 y_1 + \cdots + w_n^2 x_n y_n$$

Algunas veces es conveniente transformar un problema de mínimos cuadrados ponderados en un problema equivalente de mínimos cuadrados ordinario. Sea W la matriz diagonal con $w_1, \ldots, w_n$ (positivos) en su diagonal, de manera que

$$W\mathbf{y} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & w_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} w_1 y_1 \\ w_2 y_2 \\ \vdots \\ w_n y_n \end{bmatrix}$$

con una expresión similar para $W\hat{\mathbf{y}}$. Observe que el j-ésimo término en la ecuación (2) se puede escribir como

$$w_j^2(y_j - \hat{y}_j)^2 = (w_j y_j - w_j \hat{y}_j)^2$$

Se deduce que la SS(E) ponderada en la ecuación (2) es el cuadrado de la longitud ordinaria en $\mathbb{R}^n$ de $W\mathbf{y} - W\hat{\mathbf{y}}$, que se puede representar en la forma $\|W\mathbf{y} - W\hat{\mathbf{y}}\|^2$.

Ahora suponga que el vector de aproximación $\hat{\mathbf{y}}$ se construirá mediante las columnas de A. Entonces se busca una $\hat{\mathbf{x}}$ que haga que $A\hat{\mathbf{x}} = \hat{\mathbf{y}}$ esté tan cerca de $\mathbf{y}$ como sea posible. Sin embargo, la medida de cercanía es el error ponderado,

$$\|W\mathbf{y} - W\hat{\mathbf{y}}\|^2 = \|W\mathbf{y} - WA\hat{\mathbf{x}}\|^2$$

Así, $\hat{\mathbf{x}}$ es la solución (ordinaria) de mínimos cuadrados de la ecuación

$$WA\mathbf{x} = W\mathbf{y}$$

La ecuación normal para la solución de mínimos cuadrados es

$$(WA)^T WA\mathbf{x} = (WA)^T W\mathbf{y}$$

EJEMPLO 1 Encuentre la recta de mínimos cuadrados $y = \beta_0 + \beta_1 x$ que mejor se ajuste a los datos $(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$ y $(2, 3)$. Suponga que los errores al medir los valores y de los dos últimos puntos de datos son más grandes que para los demás puntos. Pondere estos datos a la mitad en relación con los datos restantes.

[1] Nota para lectores con conocimientos de estadística: Suponga que los errores al medir y_i son variables aleatorias independientes con media igual a cero y varianzas $\sigma_1^2, \ldots, \sigma_n^2$. Entonces, los pesos adecuados en la ecuación (2) son $w_i^2 = 1/\sigma_i^2$. Cuanto mayor sea la varianza del error, menor será el peso.

SOLUCIÓN Como en la sección 6.6, escriba X para la matriz A y β para el vector $\mathbf{x}$, y obtenga

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix}$$

Para una matriz de ponderación, seleccione W con entradas diagonales 2, 2, 2, 1 y 1. Al multiplicar por la izquierda por W se escalan las filas de X y $\mathbf{y}$:

$$WX = \begin{bmatrix} 2 & -4 \\ 2 & -2 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad W\mathbf{y} = \begin{bmatrix} 6 \\ 10 \\ 10 \\ 4 \\ 3 \end{bmatrix}$$

Para la ecuación normal, calcule

$$(WX)^T WX = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \quad \text{y} \quad (WX)^T W\mathbf{y} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

y resuelva

$$\begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

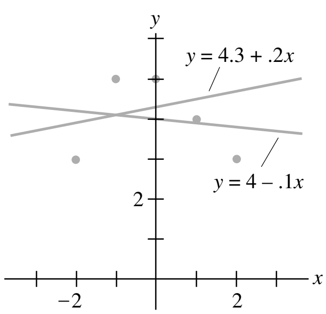


FIGURA 1
Rectas de mínimos cuadrados ordinaria y ponderada.

La solución de la ecuación normal es (a dos dígitos significativos) $\beta_0 = 4.3$ y $\beta_1 = .20$. La recta deseada es

$$y = 4.3 + .20x$$

En contraste, la recta de mínimos cuadrados ordinaria para estos datos es

$$y = 4.0 - .10x$$

En la figura 1 se representan ambas rectas. ■

Análisis de tendencia de datos

Sea f una función desconocida cuyos valores se conocen (quizá solo aproximadamente) en $t_0,\ldots, t_n$. Si existe una "tendencia lineal" en los datos $f(t_0),\ldots, f(t_n)$, entonces se espera poder aproximar los valores de f mediante una función de la forma $\beta_0 + \beta_1 t$. Si hay una "tendencia cuadrática" de los datos, podría intentarse una función con la estructura $\beta_0 + \beta_1 t + \beta_2 t^2$. Esto se analizó en la sección 6.6, pero desde otro punto de vista.

En algunos problemas estadísticos, es importante poder separar las tendencias lineal y cuadrática (y posiblemente cúbica o de mayor orden). Por ejemplo, suponga que unos ingenieros están analizando el desempeño de un nuevo automóvil, y $f(t)$ representa la distancia entre el vehículo (en el momento t) y algún punto de referencia. Si el auto viaja a velocidad constante, entonces la gráfica de $f(t)$ debería ser una recta cuya pendiente es la velocidad del auto. Si se presiona el acelerador repentinamente, entonces la gráfica de $f(t)$ cambiará para incluir un término cuadrático y posiblemente un término cúbico (debido a la aceleración). Para analizar la capacidad del auto para rebasar a otro, por ejemplo, los ingenieros tal vez quieran separar las componentes cuadrática y cúbica del término lineal.

Si la función es aproximada por una curva de la forma $y = \beta_0 + \beta_1 t + \beta_2 t^2$, es posible que el coeficiente β^2 no brinde la información deseada sobre la tendencia cuadrática en los datos, porque quizá no sea "independiente" (en un sentido estadístico) de los otros β_i.

Para efectuar el **análisis de tendencia** de los datos, se introduce un producto interior sobre el espacio $\mathbb{P}_n$ análogo al del ejemplo 2 de la sección 6.7. Para p, q en $\mathbb{P}_n$, se define

$$\langle p, q\rangle = p(t_0)q(t_0) + \cdots + p(t_n)q(t_n)$$

En la práctica, los especialistas en estadística rara vez necesitan considerar tendencias en los datos de grado mayor a tres o cuatro. Así que sean p_0, p_1, p_2, p_3 una base ortogonal del subespacio $\mathbb{P}_3$ de $\mathbb{P}_n$, obtenida al aplicar el proceso de Gram-Schmidt a los polinomios 1, t, t^2 y t^3. De acuerdo con el ejercicio complementario 11 del capítulo 2, existe un polinomio g en $\mathbb{P}_n$ cuyos valores en $t_0,\ldots, t_n$ coinciden con aquellos de la función desconocida f. Sea $\hat{g}$ la proyección ortogonal (con respecto al producto interior dado) de g sobre P$_3$, digamos,

$$\hat{g} = c_0 p_0 + c_1 p_1 + c_2 p_2 + c_3 p_3$$

Entonces $\hat{g}$ se llama una **función de tendencia** cúbica, y $c_0,\ldots, c_3$ son los **coeficientes de tendencia** de los datos. El coeficiente c_1 mide la tendencia lineal, c_2 la tendencia cuadrática, y c_3 la tendencia cúbica. Resulta que si los datos tienen ciertas propiedades, entonces esos coeficientes son estadísticamente independientes.

Como $p_0,\ldots, p_3$ son ortogonales, los coeficientes de tendencia se pueden calcular uno por uno, independientemente de los otros. (Recuerde que $c_i = \langle g, p_i\rangle/\langle p_i, p_i\rangle$). Es posible ignorar p_3 y c_3 si solo se desea la tendencia cuadrática. Y si, por ejemplo, se necesita determinar la tendencia cuártica, solo se tendría que encontrar (por medio de Gram-Schmidt) un polinomio p_4 en $\mathbb{P}_4$ ortogonal a $\mathbb{P}_3$ y calcular $\langle g, p_4\rangle/\langle p_4, p_4\rangle$.

EJEMPLO 2 El uso más sencillo y común del análisis de tendencia ocurre cuando los puntos $t_0,\ldots, t_n$ se pueden ajustar para quedar equidistantes entre sí y su suma sea cero. Ajuste una función de tendencia cuadrática a los datos $(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$ y $(2, 3)$.

SOLUCIÓN Las coordenadas t se escalan convenientemente para emplear los polinomios ortogonales encontrados en el ejemplo 5 de la sección 6.7:

$$\begin{array}{lcccc}
\text{Polinomio:} & p_0 & p_1 & p_2 & \text{Datos: } g \\
\text{Vector de valores:} & \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, & \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, & \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}, & \begin{bmatrix}3\\5\\5\\4\\3\end{bmatrix}
\end{array}$$

Los cálculos solo necesitan esos vectores, y no las fórmulas específicas para los polinomios ortogonales. La mejor aproximación a los datos mediante polinomios en $\mathbb{P}_2$ es la proyección ortogonal dada por

$$\begin{aligned}\hat{p} &= \frac{\langle g, p_0\rangle}{\langle p_0, p_0\rangle}p_0 + \frac{\langle g, p_1\rangle}{\langle p_1, p_1\rangle}p_1 + \frac{\langle g, p_2\rangle}{\langle p_2, p_2\rangle}p_2 \\ &= \tfrac{20}{5}p_0 - \tfrac{1}{10}p_1 - \tfrac{7}{14}p_2\end{aligned}$$

y

$$\hat{p}(t) = 4 - .1t - .5(t^2 - 2) \tag{3}$$

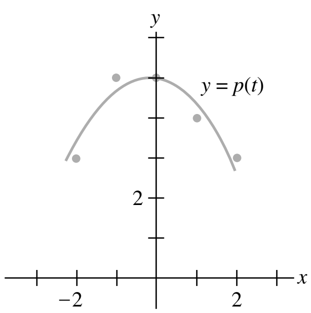


FIGURA 2
Aproximación por una función de tendencia cuadrática.

Como el coeficiente de p_2 no es extremadamente pequeño, sería razonable concluir que la tendencia es al menos cuadrática. Esto se confirma con la gráfica de la figura 2. ∎

Serie de Fourier (se requiere cálculo)

Con frecuencia, las funciones continuas se aproximan mediante combinaciones lineales de funciones seno y coseno. Por ejemplo, una función continua podría representar una onda sonora, una señal eléctrica del algún tipo, o el movimiento de un sistema mecánico vibratorio.

Para simplificar, considere funciones en $0 \le t \le 2\pi$. Resulta que cualquier función en $C[0, 2\pi]$ se puede aproximar tan cerca como se requiera con una función de la forma

$$\frac{a_0}{2} + a_1 \cos t + \cdots + a_n \cos nt + b_1 \operatorname{sen} t + \cdots + b_n \operatorname{sen} nt \tag{4}$$

para un valor de n suficientemente grande. La función (4) es un **polinomio trigonométrico**. Si a_n y b_n no son ambas cero, entonces el polinomio es de **orden *n***. La conexión entre polinomios trigonométricos y otras funciones en $C[0, 2\pi]$ depende del hecho de que para cualquier $n \ge 1$, el conjunto

$$\{1, \cos t, \cos 2t, \ldots, \cos nt, \operatorname{sen} t, \operatorname{sen} 2t, \ldots, \operatorname{sen} nt\} \tag{5}$$

sea ortogonal con respecto al producto interior

$$\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)\,dt \tag{6}$$

Esta ortogonalidad se comprueba como en el siguiente ejemplo y en los ejercicios 5 y 6.

EJEMPLO 3 Considere que $C[0, 2\pi]$ tiene el producto interior (6), y sean m y n enteros positivos diferentes. Demuestre que $\cos mt$ y $\cos nt$ son ortogonales.

SOLUCIÓN Utilice una identidad trigonométrica. Cuando $m \neq n$,

$$\begin{aligned}
\langle \cos mt, \cos nt \rangle &= \int_0^{2\pi} \cos mt \cos nt\,dt \\
&= \frac{1}{2}\int_0^{2\pi} [\cos(mt+nt) + \cos(mt-nt)]\,dt \\
&= \frac{1}{2}\left[\frac{\operatorname{sen}(mt+nt)}{m+n} + \frac{\operatorname{sen}(mt-nt)}{m-n}\right]\Bigg|_0^{2\pi} = 0
\end{aligned}$$

■

Sea W el subespacio de $C[0, 2\pi]$ generado por las funciones de la ecuación (5). Dada f en $C[0, 2\pi]$, la mejor aproximación a f mediante funciones en W es la **aproximación de Fourier de *n*-ésimo orden** sobre $[0, 2\pi]$. Puesto que las funciones de la ecuación (5) son ortogonales, la mejor aproximación está dada por la proyección ortogonal sobre W. En este caso, los coeficientes a_k y b_k de la ecuación (4) son los **coeficientes de Fourier** de f. La fórmula estándar para una proyección ortogonal indica que

$$a_k = \frac{\langle f, \cos kt \rangle}{\langle \cos kt, \cos kt \rangle}, \quad b_k = \frac{\langle f, \operatorname{sen} kt \rangle}{\langle \operatorname{sen} kt, \operatorname{sen} kt \rangle}, \quad k \ge 1$$

En el ejercicio 7 se le pide demostrar que $\langle \cos kt, \cos kt \rangle = \pi$ y $\langle \operatorname{sen} kt, \operatorname{sen} kt \rangle = \pi$. Así,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\,dt, \quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\operatorname{sen} kt\,dt \tag{7}$$

El coeficiente de la función (constante) 1 en la proyección ortogonal es

$$\frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi}\int_0^{2\pi} f(t)\cdot 1\,dt = \frac{1}{2}\left[\frac{1}{\pi}\int_0^{2\pi} f(t)\cos(0\cdot t)\,dt\right] = \frac{a_0}{2}$$

donde a_0 está definida por (7) para $k = 0$. Esto explica por qué el término constante en (4) se escribe como $a_0/2$.

EJEMPLO 4 Encuentre la aproximación de Fourier de n-ésimo orden a la función $f(t) = t$ sobre el intervalo $[0, 2\pi]$.

SOLUCIÓN Calcule

$$\frac{a_0}{2} = \frac{1}{2}\cdot\frac{1}{\pi}\int_0^{2\pi} t\,dt = \frac{1}{2\pi}\left[\frac{1}{2}t^2\Big|_0^{2\pi}\right] = \pi$$

y para $k > 0$, empleando integración por partes,

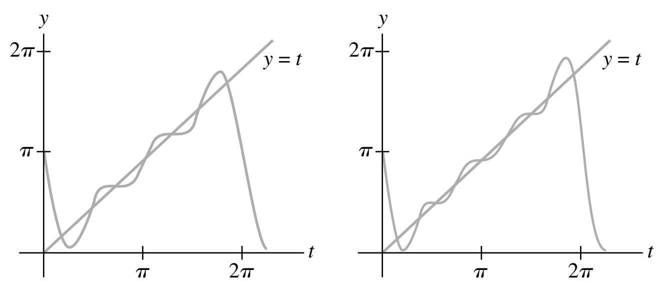
$$a_k = \frac{1}{\pi}\int_0^{2\pi} t\cos kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\cos kt + \frac{t}{k}\operatorname{sen} kt\right]_0^{2\pi} = 0$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} t\operatorname{sen} kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\operatorname{sen} kt - \frac{t}{k}\cos kt\right]_0^{2\pi} = -\frac{2}{k}$$

Así, la aproximación de Fourier de n-ésimo orden a $f(t) = t$ es

$$\pi - 2\operatorname{sen} t - \operatorname{sen} 2t - \frac{2}{3}\operatorname{sen} 3t - \cdots - \frac{2}{n}\operatorname{sen} nt$$

La figura 3 muestra las aproximaciones de Fourier de tercer y cuarto órdenes a f. ■



a) Tercer orden *b*) Cuarto orden

FIGURA 3 Aproximaciones de Fourier de la función $f(t) = t$.

La norma de la diferencia entre f y una aproximación de Fourier es el **error cuadrático medio** de la aproximación. (El término *medio* se refiere al hecho de que la norma está determinada por una integral). Es posible demostrar que el error cuadrático medio se aproxima a cero conforme se incrementa el orden de la aproximación de Fourier. Por esa razón, es común escribir

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty}(a_m\cos mt + b_m\operatorname{sen} mt)$$

Esta expresión para $f(t)$ es la **serie de Fourier** para f sobre $[0, 2\pi]$. El término $a_m\cos mt$, por ejemplo, es la proyección de f sobre un subespacio unidimensional generado por $\cos mt$.

PROBLEMAS DE PRÁCTICA

1. Sean $q_1(t) = 1$, $q_2(t) = t$ y $q_3(t) = 3t^2 - 4$. Compruebe que $\{q_1, q_2, q_3\}$ es un conjunto ortogonal en $C[-2, 2]$ con el producto interior del ejemplo 7 de la sección 6.7 (integración de -2 a 2).
2. Encuentre las aproximaciones de Fourier de primer y tercer órdenes a

$$f(t) = 3 - 2\operatorname{sen} t + 5\operatorname{sen} 2t - 6\cos 2t$$

6.8 EJERCICIOS

1. Encuentre la recta de mínimos cuadrados $y = \beta_0 + \beta_1 x$ que mejor se ajuste a los datos $(-2, 0)$, $(-1, 0)$, $(0, 2)$, $(1, 4)$ y $(2, 4)$, suponiendo que el primero y el último de los puntos de datos son menos confiables. Pondérelos a la mitad en relación con los tres restantes puntos interiores.

2. En un problema de mínimos cuadrados ponderados, suponga que 5 de 25 puntos de datos tienen una medición de y que es menos confiable que las demás, y en consecuencia se deben ponderar a la mitad en relación con los otros 20 puntos. Un método consiste en ponderar los 20 puntos por un factor de 1, y los otros 5 por un factor de $\frac{1}{2}$. Un segundo método consiste en ponderar los 20 puntos por un factor de 2, y los otros 5 por un factor de 1. ¿Estos dos métodos producen resultados diferentes? Explique su respuesta.

3. Ajuste una función de tendencia cúbica a los datos del ejemplo 2. El polinomio cúbico ortogonal es $p_3(t) = \frac{5}{6}t^3 - \frac{17}{6}t$.

4. Para realizar un análisis de tendencia de seis puntos de datos igualmente espaciados, se pueden emplear polinomios ortogonales con respecto a la evaluación en los puntos $t = -5, -3, -1, 1, 3$ y 5.

   a) Demuestre que los primeros tres polinomios ortogonales son

   $p_0(t) = 1, \quad p_1(t) = t \quad \text{y} \quad p_2(t) = \frac{3}{8}t^2 - \frac{35}{8}$

   (El polinomio p_2 se escaló para que sus valores en los puntos de evaluación fueran enteros pequeños).

   b) Ajuste una función de tendencia cuadrática a los datos

   $(-5, 1), (-3, 1), (-1, 4), (1, 4), (3, 6), (5, 8)$

En los ejercicios 5 a 14, el espacio es $C[0, 2\pi]$ con el producto interior (6).

5. Demuestre que sen mt y sen nt son ortogonales cuando $m \neq n$.

6. Demuestre que sen mt y cos nt son ortogonales para todos los enteros positivos m y n.

7. Demuestre que $\|\cos kt\|^2 = \pi$ y $\|\text{sen}\ kt\|^2$ para $k > 0$.

8. Encuentre la aproximación de Fourier de tercer orden a $f(t) = t - 1$.

9. Obtenga la aproximación de Fourier de tercer orden a $f(t) = 2\pi - t$.

10. Determine la aproximación de Fourier de tercer orden a la *función de onda cuadrada*, $f(t) = 1$ para $0 \le t < \pi$ y $f(t) = -1$ para $\pi \le t < 2\pi$.

11. Encuentre la aproximación de Fourier de tercer orden a $\text{sen}^2 t$, sin efectuar cálculos con integrales.

12. Obtenga la aproximación de Fourier de tercer orden a $\cos^3 t$, sin efectuar cálculos con integrales.

13. Explique por qué un coeficiente de Fourier de la suma de dos funciones es la suma de los coeficientes de Fourier correspondientes de las dos funciones.

14. Suponga que los primeros pocos coeficientes de Fourier de alguna función f en $C[0, 2\pi]$ son a_0, a_1, a_2, y b_1, b_2, b_3. ¿Cuál de los siguientes polinomios trigonométricos es más cercano a f? Argumente su respuesta.

$$g(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \,\text{sen}\, t$$

$$h(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \,\text{sen}\, t + b_2 \,\text{sen}\, 2t$$

15. **[M]** Consulte los datos del ejercicio 13 de la sección 6.6, concernientes al despegue de un avión. Suponga que los errores de las posibles mediciones aumentan conforme se incrementa la rapidez del avión, y sea W la matriz ponderada diagonal cuyas entradas diagonales son 1, 1, 1, .9, .9, .8, .7, .6, .5, .4, .3, .2 y .1. Encuentre la curva cúbica que se ajuste a los datos con error de mínimos cuadrados ponderados, y úsela para estimar la velocidad del avión cuando $t = 4.5$ segundos.

16. **[M]** Sean f_4 y f_5 las aproximaciones de Fourier de cuarto y quinto órdenes en $C[0, 2\pi]$ a la función de onda cuadrada del ejercicio 10. Elabore gráficas separadas de f_4 y f_5 en el intervalo $[0, 2\pi]$, y trace una gráfica de f_5 en $[-2\pi, 2\pi]$.

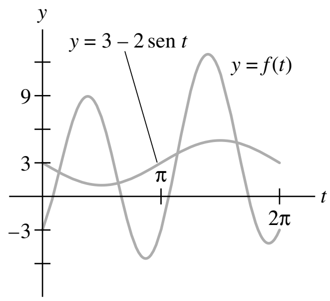
SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Calcule

$$\langle q_1, q_2 \rangle = \int_{-2}^{2} 1 \cdot t \, dt = \frac{1}{2}t^2 \Big|_{-2}^{2} = 0$$

$$\langle q_1, q_3 \rangle = \int_{-2}^{2} 1 \cdot (3t^2 - 4) \, dt = (t^3 - 4t) \Big|_{-2}^{2} = 0$$

$$\langle q_2, q_3 \rangle = \int_{-2}^{2} t \cdot (3t^2 - 4) \, dt = \left(\frac{3}{4}t^4 - 2t^2\right) \Big|_{-2}^{2} = 0$$



Aproximaciones de primer y tercer órdenes a $f(t)$.

2. La aproximación de Fourier de tercer orden a f es la mejor aproximación en $C[0, 2\varpi]$ a f mediante funciones (vectores) en el subespacio generado por 1, cos t, cos $2t$, cos $3t$, sen t, sen $2t$ y sen $3t$. Pero f está evidentemente *en* este subespacio, así que f es su propia mejor aproximación:

$$f(t) = 3 - 2 \text{ sen } t + 5 \text{ sen } 2t - 6 \cos 2t$$

Para la aproximación de primer orden, la función más cercana a f en el subespacio $W = \text{Gen }\{1, \cos t, \text{sen } t\}$ es $3 - 2$ sen t. Los otros dos términos en la fórmula para $f(t)$ son ortogonales a las funciones en W, así que no contribuyen a las integrales que dan los coeficientes de Fourier para la aproximación de primer orden.

CAPÍTULO 6 EJERCICIOS COMPLEMENTARIOS

1. Los siguientes enunciados se refieren a vectores en $\mathbb{R}^n$ (o $\mathbb{R}^m$) con el producto interior estándar. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

a) La longitud de cada vector es un número positivo.

b) Un vector $\mathbf{v}$ y su negativo $-\mathbf{v}$ tienen iguales longitudes.

c) La distancia entre $\mathbf{u}$ y $\mathbf{v}$ es $\|\mathbf{u} - \mathbf{v}\|$.

d) Si r es cualquier escalar, entonces $\|r\mathbf{v}\| = r\|\mathbf{v}\|$.

e) Si dos vectores son ortogonales, entonces son linealmente independientes.

f) Si $\mathbf{x}$ es ortogonal a $\mathbf{u}$ y $\mathbf{v}$, entonces $\mathbf{x}$ debe ser ortogonal a $\mathbf{u} - \mathbf{v}$.

g) Si $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

h) Si $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

i) La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$ es un múltiplo escalar de $\mathbf{y}$.

j) Si un vector $\mathbf{y}$ coincide con su proyección ortogonal sobre un subespacio W, entonces $\mathbf{y}$ está en W.

k) El conjunto de todos los vectores en $\mathbb{R}^n$ ortogonales a un vector fijo es un subespacio de $\mathbb{R}^n$.

l) Si W es un subespacio de $\mathbb{R}^n$, entonces W y $W^\perp$ no tienen vectores en común.

m) Si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal y si c_1, c_2 y c_3 son escalares, entonces $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, c_3\mathbf{v}_3\}$ es un conjunto ortogonal.

n) Si una matriz U tiene columnas ortonormales, entonces $UU^T = I$.

o) Una matriz cuadrada con columnas ortogonales es una matriz ortogonal.

p) Si una matriz cuadrada tiene columnas ortonormales, entonces también tiene filas ortonormales.

q) Si W es un subespacio, entonces $\|\text{proy}_W\, \mathbf{v}\|^2 + \|\mathbf{v} - \text{proy}_W\, \mathbf{v}\|^2 = \|\mathbf{v}\|^2$.

r) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es el vector $A\hat{\mathbf{x}}$ en Col A más cercano a $\mathbf{b}$, de manera que $\|\mathbf{b} - A\hat{\mathbf{x}}\| \le \|\mathbf{b} - A\hat{\mathbf{x}}\|$ para toda $\mathbf{x}$.

s) Las ecuaciones normales para una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ están dadas por $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

2. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ un conjunto ortonormal. Compruebe la siguiente igualdad mediante inducción, iniciando con $p = 2$. Si $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, entonces

$$\|\mathbf{x}\|^2 = |c_1|^2 + \cdots + |c_p|^2$$

3. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ un conjunto ortonormal en $\mathbb{R}^n$. Compruebe la siguiente *desigualdad de Bessel*, que es válida para toda $\mathbf{x}$ en $\mathbb{R}^n$:

$$\|\mathbf{x}\|^2 \ge |\mathbf{x}\cdot\mathbf{v}_1|^2 + |\mathbf{x}\cdot\mathbf{v}_2|^2 + \cdots + |\mathbf{x}\cdot\mathbf{v}_p|^2$$

4. Sea U una matriz ortogonal de $n \times n$. Demuestre que si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base ortonormal para $\mathbb{R}^n$, entonces también lo es $\{U\mathbf{v}_1, \ldots, U\mathbf{v}_n\}$.

5. Demuestre que si una matriz U de $n \times n$ satisface $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$ para todas las $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$, entonces U es una matriz ortogonal.

6. Demuestre que si U es una matriz ortogonal, entonces cualquier valor propio real de U debe ser ± 1.

7. Una *matriz de Householder*, o un *reflector elemental*, tiene la forma $Q = I - 2\mathbf{u}\mathbf{u}^T$, donde $\mathbf{u}$ es un vector unitario. (Véase el ejercicio 13 en los ejercicios complementarios del capítulo 2). Demuestre que Q es una matriz ortogonal. (Los reflectores elementales se utilizan a menudo en programas computacionales para construir una factorización QR de una matriz A. Si A tiene columnas linealmente independientes, entonces la multiplicación por la izquierda mediante una secuencia de reflectores elementales puede generar una matriz triangular superior).

8. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal que preserva longitudes; es decir, $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

a) Demuestre que T también preserva ortogonalidad; es decir, $T(\mathbf{x})\cdot T(\mathbf{y}) = 0$ siempre que $\mathbf{x}\cdot\mathbf{y} = 0$.

b) Demuestre que la matriz estándar de T es una matriz ortogonal.

9. Considere que $\mathbf{u}$ y $\mathbf{v}$ representan vectores linealmente independientes en $\mathbb{R}^n$ que *no* son ortogonales. Describa cómo encontrar la mejor aproximación a $\mathbf{z}$ en $\mathbb{R}^n$ mediante vectores de la forma $x_1\mathbf{u} + x_2\mathbf{v}$ sin construir primero una base ortogonal para Gen $\{\mathbf{u}, \mathbf{v}\}$.

10. Suponga que las columnas de A son linealmente independientes. Determine qué pasa con la solución de mínimos cuadrados $\hat{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ cuando $\mathbf{b}$ se remplaza por $c\mathbf{b}$ para algún escalar c diferente de cero.

11. Si a, b y c son números distintos, entonces el siguiente sistema es inconsistente porque las gráficas de las ecuaciones son planos paralelos. Demuestre que el conjunto de todas las soluciones de mínimos cuadrados del sistema es precisamente el plano cuya ecuación es $x - 2y + 5z = (a + b + c)/3$.

$$\begin{aligned} x - 2y + 5z &= a \\ x - 2y + 5z &= b \\ x - 2y + 5z &= c \end{aligned}$$

12. Considere el problema de encontrar un valor propio de una matriz A de $n \times n$ cuando se conoce un vector propio $\mathbf{v}$ aproximado. Puesto que $\mathbf{v}$ no es exactamente correcto, la ecuación

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

probablemente no tendrá una solución. Sin embargo, λ se puede estimar mediante una solución de mínimos cuadrados cuando la ecuación (1) se analiza adecuadamente. Piense en $\mathbf{v}$ como en una matriz V de $n \times 1$, y considere λ un vector en $\mathbb{R}^1$, y denote el vector $A\mathbf{v}$ con el símbolo $\mathbf{b}$. Luego, (1) se reduce a $\mathbf{b} = \lambda V$, que también se puede escribir como $V\lambda = \mathbf{b}$. Obtenga la solución de mínimos cuadrados de este sistema de n ecuaciones con la única incógnita λ, y escriba esta solución empleando los símbolos originales. La estimación resultante para λ se denomina *cociente de Rayleigh*. Véase los ejercicios 11 y 12 de la sección 5.8.

13. Siga los pasos descritos a continuación para probar las siguientes relaciones entre los cuatro subespacios fundamentales determinados por una matriz A de $m \times n$.

$$\text{Fil } A = (\text{Nul } A)^{\perp}, \quad \text{Col } A = (\text{Nul } A^T)^{\perp}$$

a) Demuestre que Fil A está contenida en $(\text{Nul } A)^{\perp}$. (Demuestre que si $\mathbf{x}$ está en Fil A, entonces $\mathbf{x}$ es ortogonal a cada $\mathbf{u}$ en Nul A).

b) Suponga que rango $A = r$. Encuentre dim Nul A y dim $(\text{Nul } A)^{\perp}$, y entonces deduzca del inciso *a*) que Fil $A = (\text{Nul } A)^{\perp}$. [*Sugerencia:* Estudie los ejercicios de la sección 6.3].

c) Explique por qué Col $A = (\text{Nul } A)^{\perp}$.

14. Explique por qué una ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución si y solo si $\mathbf{b}$ es ortogonal a todas las soluciones de la ecuación $A^T\mathbf{x} = \mathbf{0}$.

Los ejercicios 15 y 16 conciernen a la *factorización de Schur* (real) de una matriz A de $n \times n$ en la forma $A = URU^T$, donde U es una matriz ortogonal y R es una matriz triangular superior de $n \times n$.[1]

15. Demuestre que si A admite una factorización de Schur (real), $A = URU^T$, entonces A tiene n valores propios reales, contando multiplicidades.

16. Sea A una matriz de $n \times n$ con n valores propios reales, contando multiplicidades, denotados con $\lambda_1, \ldots, \lambda_n$. Es posible demostrar que A admite una factorización de Schur (real). Los incisos *a*) y *b*) dan las ideas clave de la demostración. El resto de la demostración equivale a repetir *a*) y *b*) para matrices sucesivamente más pequeñas, y luego unir todos los resultados.

a) Sean $\mathbf{u}_1$ un vector propio unitario correspondiente a λ_1, y $\mathbf{u}_2, \ldots, \mathbf{u}_n$ cualesquiera otros vectores tales que $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ sea una base ortonormal para $\mathbb{R}^n$, y entonces sea $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_n]$. Demuestre que la primera columna de $U^T AU$ es $\lambda_1\mathbf{e}_1$, donde $\mathbf{e}_1$ es la primera columna de la matriz identidad de $n \times n$.

b) El inciso *a*) implica que U^TAU tiene la forma que se muestra abajo. Explique por qué los valores propios de A_1 son $\lambda_2, \ldots, \lambda_n$. [*Sugerencia:* Véase los ejercicios complementarios del capítulo 5].

$$U^TAU = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{bmatrix}$$

[**M**] Cuando el lado derecho de una ecuación $A\mathbf{x} = \mathbf{b}$ se modifica ligeramente, por ejemplo, a $A\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}$ para algún vector $\Delta\mathbf{b}$, la solución cambia de $\mathbf{x}$ a $\mathbf{x} + \Delta\mathbf{x}$, donde $\Delta\mathbf{x}$ satisface $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$. El cociente $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$ es el **cambio relativo** en $\mathbf{b}$ (o el **error relativo** en $\mathbf{b}$ cuando $\Delta\mathbf{b}$ representa el error posible en las entradas de $\mathbf{b}$). El cambio relativo en la solución es $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$. Cuando A es invertible, el **número de condición** de A, representado como cond(A), aporta un límite de la magnitud del cambio relativo en $\mathbf{x}$:

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \text{cond}(A) \cdot \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \tag{2}$$

En los ejercicios 17 a 20, resuelva $A\mathbf{x} = \mathbf{b}$ y $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$, y demuestre que la desigualdad (2) es válida en cada caso. (Véase el análisis de matrices *mal condicionadas* en los ejercicios 41 a 43 de la sección 2.3).

17. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 19.249 \\ 6.843 \end{bmatrix}, \Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

18. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} .500 \\ -1.407 \end{bmatrix}, \Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

[1] Si se permiten números complejos, *cada matriz A de n × n* admite una factorización de Schur (compleja), $A = URU^{-1}$, donde R es triangular superior y U^{-1} es la transpuesta *conjugada* de U. Este hecho de gran utilidad se analiza en la obra *Matrix Analysis*, de Roger A. Horn y Charles R. Johnson (Cambridge: Cambridge University Press, 1985), pp.79-100.

19. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} .100 \\ 2.888 \\ -1.404 \\ 1.462 \end{bmatrix}$,

$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .49 \\ -1.28 \\ 5.78 \\ 8.04 \end{bmatrix}$

20. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4.230 \\ -11.043 \\ 49.991 \\ 69.536 \end{bmatrix}$,

$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .27 \\ 7.76 \\ -3.77 \\ 3.93 \end{bmatrix}$