10

Matrix Algebra and Its Applications

10.1 Geometric Vectors

In the early 1800s, physicists found that most physical quantities could be categorized in one of two ways: those that have only size, such as length, time, or mass, and those that have both size and direction, such as force, velocity, or acceleration. Quantities that have only size, or *magnitude,* are called **scalars;** those that have both magnitude and direction are called **vectors.** We use lightface, italic lowercase letters, such as a, m, or x to denote scalars; we use boldface, roman lowercase letters such as **b, v,** or **x** to denote vectors.

Although vectors may represent physical (as well as other) quantities, we will think of them geometrically in this section. In two dimensions, we visualize a vector as an arrow connecting two points, as shown in Figure 10.1. The length of the arrow represents the magnitude of the vector. The slope of the line through any two points on the arrow, along with the arrowhead, gives the direction of the vector **v**. In three dimensions, we likewise visualize a vector as an arrow connecting two points, as shown in Figure 10.2.

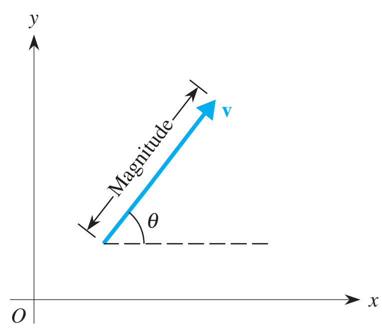


FIGURE 10.1

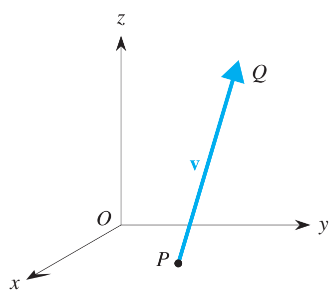


FIGURE 10.2

Any vector starting at the origin is known as a *position vector* because it gives the position of the arrowhead with respect to the origin. A vector connecting two points P and Q, sometimes written $\mathbf{v} = PQ$, is called a *displacement vector;* it indicates how to get from P to Q by moving a given distance from P in the desired direction. Obviously, a position vector is also a displacement vector, indicating how to move from the origin to point Q. But a displacement vector is not a position vector if it starts at any point P other than the origin.

If we know the coordinates of the initial point and the final point of the arrow, we can write the vector simply. First, we consider the position vector $\mathbf{v}$ from the origin to the point $(3, 4)$ shown in Figure 10.3. It involves moving 3 units to the right and 4 units upward from the origin, and we write the vector $\mathbf{v}$ as either a *row vector* $\mathbf{v} = \begin{bmatrix} 3 & 4 \end{bmatrix}$ or as a *column vector* $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. The numerical entries 3 and 4 in the vector are called its *components*. The decision to write the vector as a row vector or as a column vector is usually a matter of choice, so long as you are consistent. We cover several specific cases later in this chapter in which the choice of column vectors is essential; in this section we primarily use row vectors for convenience.

The magnitude of the vector $\mathbf{v} = \begin{bmatrix} 3 & 4 \end{bmatrix}$ is the length of the arrow, as shown in Figure 10.3. It is the distance from the origin to the point $(3, 4)$, and so is 5, using the Pythagorean theorem.

Next, we consider the displacement vector $\mathbf{w}$ from the point $(6, 15)$ to the point $(11, 3)$, as shown in Figure 10.4. It involves a move of $5\,(=11 - 6)$ to the right and a move of $-12\,(=3 - 15)$ vertically. We therefore write this vector either as the row vector $\mathbf{w} = \begin{bmatrix} 5 & -12 \end{bmatrix}$ or as the column vector $\mathbf{w} = \begin{bmatrix} 5 \\ -12 \end{bmatrix}$. Note that, in a displacement vector, the components are the differences in the coordinates of the points defining the vector. The magnitude of the vector $\mathbf{w} = \begin{bmatrix} 5 & -12 \end{bmatrix}$ equals the distance from one point to the other, or

$$\begin{aligned} \text{Magnitude} &= \sqrt{(11 - 6)^2 + (3 - 15)^2} = \sqrt{5^2 + (-12)^2} \\ &= \sqrt{25 + 144} = \sqrt{169} = 13. \end{aligned}$$

FIGURE 10.3

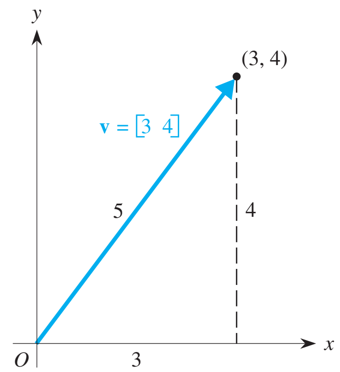
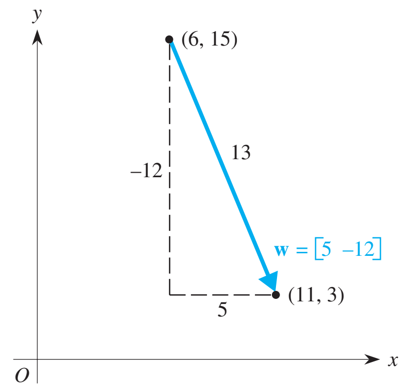


FIGURE 10.4



We write the magnitude of a vector $\mathbf{v}$ as $\|\mathbf{v}\|$. In general, in two dimensions, we have the following.

> If $\mathbf{v} = \begin{bmatrix} a & b \end{bmatrix}$, then
> $$\|\mathbf{v}\| = \sqrt{a^2 + b^2}.$$

For example, if $\mathbf{v} = [7 \quad -4]$, then $\|\mathbf{v}\| = \sqrt{7^2 + (-4)^2} = \sqrt{65}$.

Similarly, in three dimensions we can write a vector $\mathbf{v}$ in terms of three components a, b, and c as $\mathbf{v} = [a \quad b \quad c]$. The magnitude of such a vector is defined analogously with the Pythagorean theorem.

If $\mathbf{v} = [a \quad b \quad c]$, then

$$\|\mathbf{v}\| = \sqrt{a^2 + b^2 + c^2}.$$

However, specifying the direction of a vector in space is considerably harder than in the plane, and we don't go into it here.

EXAMPLE 1

Find the magnitude of the vector from the point $(1, 2, 4)$ to the point $(3, -3, 8)$.

Solution We first write this vector in terms of its components, which are the differences in each of the three coordinates. Therefore $\mathbf{v} = [3 - 1 \quad -3 - 2 \quad 8 - 4] = [2 \quad -5 \quad 4]$, and its magnitude is

$$\|\mathbf{v}\| = \sqrt{2^2 + (-5)^2 + 4^2} = \sqrt{4 + 25 + 16} = \sqrt{45}.$$

We say that two vectors $\mathbf{v}$ and $\mathbf{w}$ are equal, written $\mathbf{v} = \mathbf{w}$, if all their corresponding components are equal. For instance, $[3 \quad 4 \quad 7] = [3 \quad \sqrt{16} \quad 7]$, but

$$\begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

Geometrically, two vectors are equal if they have the same magnitude and the same direction. Figure 10.5 shows that $\mathbf{v} = \mathbf{y}$ (they have the same magnitude and the same direction); but $\mathbf{v} \neq \mathbf{x}$ (they are parallel and have the same direction, but have different magnitudes) and $\mathbf{v} \neq \mathbf{w}$ (they have the same magnitude, but do not have the same direction because they are not parallel).

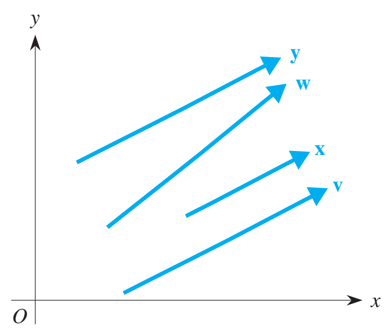


FIGURE 10.5

A Constant Multiple of a Vector

If the vector $\mathbf{v} = [3 \quad 4]$, it seems reasonable to assume that two times $\mathbf{v}$ is just

$$2\mathbf{v} = 2 \cdot [3 \quad 4] = [6 \quad 8].$$

Does this make sense geometrically? Figure 10.6 shows $\mathbf{v}$ as the displacement vector from an arbitrary point (x, y) to the point $(x + 3, y + 4)$. Note that the magnitude of $\mathbf{v}$ is 5. We also show the vector $\mathbf{w} = [6 \quad 8]$ starting from the same point (x, y); it extends 6 units to the right and 8 units up, so it ends at the point $(x + 6, y + 8)$. From the Pythagorean theorem, the magnitude of $\mathbf{w}$ is

$$\|\mathbf{w}\| = \sqrt{6^2 + 8^2} = \sqrt{100} = 10.$$

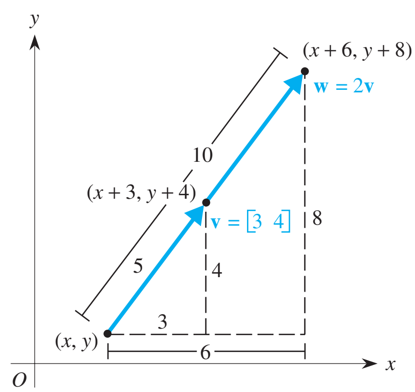


FIGURE 10.6

Thus the magnitude of $\mathbf{w}$ is twice the magnitude of $\mathbf{v}$. So multiplying $\mathbf{v}$ by 2 produces a vector $\mathbf{w} = 2\mathbf{v}$ that is twice the length of $\mathbf{v}$, but in the same direction.

In general, for any vector $\mathbf{v} = [a \quad b]$ and any scalar multiple m,

$$m \cdot \mathbf{v} = [ma \quad mb]$$

is a vector m times as long as $\mathbf{v}$ (shorter if $0 < m < 1$) that points in the same direction if $m > 0$. If the multiple $m < 0$, the resulting vector is parallel to $\mathbf{v}$, but it points in the opposite direction.

Moreover, this definition of the multiple of a vector suggests the following important and useful fact.

> Two vectors $\mathbf{v}$ and $\mathbf{w}$ are parallel if and only if one is a multiple of the other.

Unit Vectors

EXAMPLE 2

Find a vector $\mathbf{u}$ of length 1 that is in the same direction as the vector $\mathbf{v} = [-6 \quad 8]$.

Solution Because the vector $\mathbf{u}$ we want to find is in the same direction as $\mathbf{v}$, it will be parallel to $\mathbf{v}$ and so must be some multiple of $\mathbf{v}$, $\mathbf{u} = m \cdot \mathbf{v}$, as shown in Figure 10.7. The problem is to find the appropriate multiple m. The magnitude of $\mathbf{v}$ is

$$\|\mathbf{v}\| = \sqrt{(-6)^2 + 8^2} = \sqrt{36 + 64} = \sqrt{100} = 10.$$

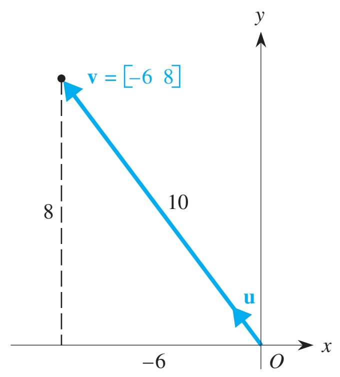


FIGURE 10.7

The vector **u** we seek is to have length 1, so it must be one-tenth of **v**. That is,

$$\mathbf{u} = \left(\frac{1}{10}\right)\mathbf{v} = \begin{bmatrix} -\frac{6}{10} & \frac{8}{10} \end{bmatrix}.$$

Any vector whose length is 1 is called a **unit vector.** In general, if **v** is any nonzero vector, a unit vector **u** in the same direction as **v** is

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

In two dimensions, the two most important unit vectors are the *coordinate vectors* along the horizontal and vertical axes. The unit coordinate vector pointing to the right is denoted by $\mathbf{i} = [1 \quad 0]$, and the unit coordinate vector pointing upward is denoted by $\mathbf{j} = [0 \quad 1]$.

The Sum of Two Vectors

We add vectors—whether they are two row vectors or two column vectors—by adding the corresponding components. For instance, if $\mathbf{v} = [4 \quad -9]$ and $\mathbf{w} = [7 \quad 3]$ are row vectors, their sum is the row vector

$$\mathbf{v} + \mathbf{w} = [4 \quad -9] + [7 \quad 3] = [4 + 7 \quad -9 + 3] = [11 \quad -6].$$

In general, if $\mathbf{v} = [v_1 \quad v_2]$ and $\mathbf{w} = [w_1 \quad w_2]$ are any two row vectors,

$$\mathbf{v} + \mathbf{w} = [v_1 + w_1 \quad v_2 + w_2].$$

Geometrically, adding vectors involves “adding” the arrows, which can be thought of in two ways. First, in Figure 10.8 vector **w** is “moved” so that it starts at the end of vector **v** and still points in the same direction. (Equivalently, vector **w** is replaced by an equal vector that starts at the end of vector **v**.) Then $\mathbf{v} + \mathbf{w}$ is the vector from the start of **v** to the end of **w**. The sum of the two vectors is the third side of the triangle formed by the vector **v** and the shifted vector **w**.

Alternatively, in Figure 10.9, **v** and **w** form adjacent sides of a parallelogram. The upper side has the same length as **v** and is parallel to **v**; therefore it equals **v**. Similarly, the right side of the parallelogram has the same magnitude as **w** and is parallel to **w**, so it equals **w**. The sum $\mathbf{v} + \mathbf{w}$ then is the long diagonal in the parallelogram.

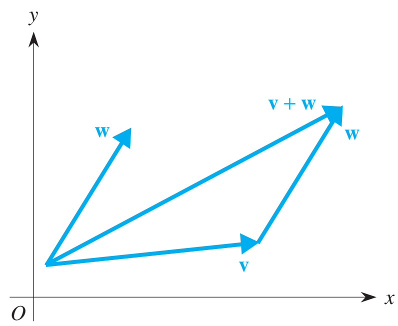


FIGURE 10.8

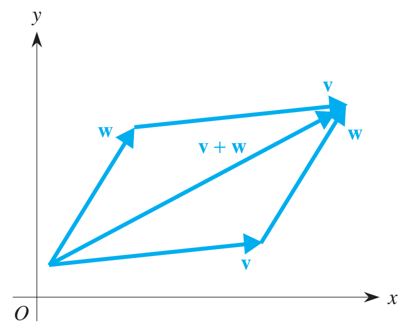


FIGURE 10.9

The Coordinate Vectors i and j

One of the advantages of the coordinate vectors $\mathbf{i} = [1 \quad 0]$ and $\mathbf{j} = [0 \quad 1]$ is that any vector in the plane $\mathbf{v} = [a \quad b]$ can be written in terms of **i** and **j**. In particular, as shown in Figure 10.10, the vector **v** can be thought of as the sum of a horizontal vector with magnitude a and a vertical vector with magnitude b. We write the horizontal vector with magnitude a as $a\mathbf{i}$ and the vertical vector with magnitude b as $b\mathbf{j}$. Consequently, $\mathbf{v} = [a \quad b] = a\mathbf{i} + b\mathbf{j}$.

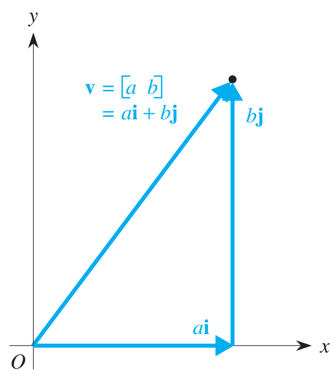


FIGURE 10.10

All operations with vectors, such as addition, can be done in terms of **i** and **j**.

EXAMPLE 3

Given the vectors $\mathbf{v} = 3\mathbf{i} + 5\mathbf{j}$ and $\mathbf{w} = 7\mathbf{i} - 4\mathbf{j}$, find **(a)** their sum and **(b)** 4 times the first vector.

Solution

a. The sum of the two vectors is $\mathbf{v} + \mathbf{w} = (3\mathbf{i} + 5\mathbf{j}) + (7\mathbf{i} - 4\mathbf{j}) = 10\mathbf{i} + \mathbf{j}$.

b. $4\mathbf{v} = 4(3\mathbf{i} + 5\mathbf{j}) = 12\mathbf{i} + 20\mathbf{j}$.

Applications of Vectors

We next look at several examples involving physical situations that use vector addition.

EXAMPLE 4

Tom is trying to open a window that is stuck. He exerts a force of 30 lb at an angle of 20° with the wall. What is the effective vertical force that he exerts upward against the window?

Solution We start by drawing a sketch of the situation, as shown in Figure 10.11. The force that Tom exerts is a vector **F** whose magnitude is 30 and whose direction is at a 20° angle with the wall. The force actually consists of two components—one vertically upward ($\mathbf{F}_y$), which represents the effective force that he exerts to raise the window, and the other ($\mathbf{F}_x$) perpendicular to the window, which doesn't have any effect on moving the window vertically. Thus the total force **F** is just the sum of the two vectors $\mathbf{F}_x$ and $\mathbf{F}_y$, and what we seek is the vertical vector $\mathbf{F}_y$.

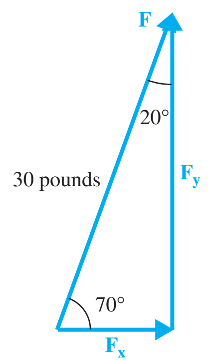


FIGURE 10.11

The angle at the upper vertex of the triangle is 20°, so the angle at the starting point of vector **F** is 70°. The length of the hypotenuse of the triangle is just the magnitude of the force vector, or $\|\mathbf{F}\| = 30$. Using trigonometry, we have

$$\sin 70° = \frac{\text{opposite}}{\text{hypotenuse}},$$

where the length of the hypotenuse is $\|\mathbf{F}\| = 30$. Therefore the length of the vertical side of the triangle is $\|\mathbf{F}\|\sin 70° = 30 \sin 70° \approx 28.19$, and the corresponding vertical vector is $\mathbf{F}_y = [0 \quad 28.19]$. Consequently, the effective force that Tom exerts upward to move the window actually is $\|\mathbf{F}_y\| = 30 \sin 70° \approx 28.19$ pounds.

EXAMPLE 5

A flock of Canadian geese is trying to fly due south for the winter with a constant velocity of 12 mph. A stiff wind is blowing at a constant rate of 20 mph from a direction 35° west of north. Find the actual direction that the geese end up flying and their actual speed with respect to the ground.

Solution We begin with a sketch of the situation, as shown in Figure 10.12. Each goose is trying to fly due south, so there is one velocity vector, **g**, for the goose having magnitude 12 and pointing vertically downward. In addition, each goose is pushed by the wind, which is coming from a northwesterly direction. The wind is represented by a second velocity vector, **w**, having magnitude 20 and pointing from a direction 35° west of north.

Gord.3896.10.pgs 4/28/03 3:29 PM Page 656

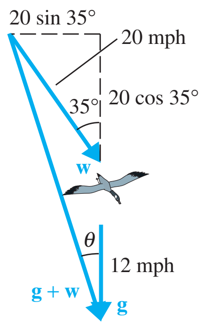


FIGURE 10.12

The actual velocity vector for the goose is the sum of these two vectors. To find the direction that a goose actually flies—and then the speed at which it flies—we need to find the components of **g** + **w**. Because the goose is trying to fly due south, **g** has only a vertical component of -12, so $\mathbf{g} = [0 \quad -12]$. The velocity vector for the wind has both a horizontal and a vertical component. Using trigonometry, we see that the horizontal component of **w** is $20 \sin 35° \approx 11.47$. Similarly, the vertical component of **w** is $-20 \cos 35° \approx -16.38$ (it is negative because it is directed downward). Consequently, the wind vector **w** is

$$\mathbf{w} = [11.47 \quad -16.38].$$

Thus the sum of the two vectors is

$$\mathbf{g} + \mathbf{w} = [0 \quad -12] + [11.47 \quad -16.38] = [11.47 \quad -28.38].$$

The actual speed with which the goose flies is the magnitude of this vector **g** + **w**, which is

$$\text{Speed} = \sqrt{(11.47)^2 + (-28.38)^2} = \sqrt{131.56 + 805.42} = \sqrt{936.90} \approx 30.61.$$

Next, to find the direction in which the goose flies, we need to find the angle θ that the vector **g** + **w** makes with the vertical. From the large right trangle in Figure 10.12, we find

$$\begin{aligned} \tan\theta &= \frac{20 \sin 35°}{20 \cos 35° + 12} \\ &= \frac{11.47}{28.38} \approx 0.404 \end{aligned}$$

so that

$$\theta = \arctan 0.404 \approx 22°.$$

Thus the geese actually end up flying in a direction 22° east of south instead of due south.

The Difference of Two Vectors

We define the difference of two vectors $\mathbf{v} = [v_1 \quad v_2]$ and $\mathbf{w} = [w_1 \quad w_2]$ to be

$$\mathbf{v} - \mathbf{w} = [v_1 - w_1 \quad v_2 - w_2];$$

that is, we simply take the difference of corresponding components.

For instance, if $\mathbf{v} = [14 \quad 3]$ and $\mathbf{w} = [6 \quad 10]$, then $\mathbf{v} - \mathbf{w} = [8 \quad -7]$. Alternatively, in terms of the unit vectors $\mathbf{i}$ and $\mathbf{j}$, $\mathbf{v} = 14\mathbf{i} + 3\mathbf{j}$ and $\mathbf{w} = 6\mathbf{i} + 10\mathbf{j}$, so that

$$\mathbf{v} - \mathbf{w} = (14\mathbf{i} + 3\mathbf{j}) - (6\mathbf{i} + 10\mathbf{j}) = 8\mathbf{i} - 7\mathbf{j}.$$

Now let's interpret the difference of two vectors geometrically. Consider the vector $\mathbf{x}$ in Figure 10.13, which connects the end of $\mathbf{v}$ to the end of $\mathbf{w}$. We know from the sum of two vectors that

$$\mathbf{v} + \mathbf{x} = \mathbf{w} \quad \text{so that} \quad \mathbf{x} = \mathbf{w} - \mathbf{v}.$$

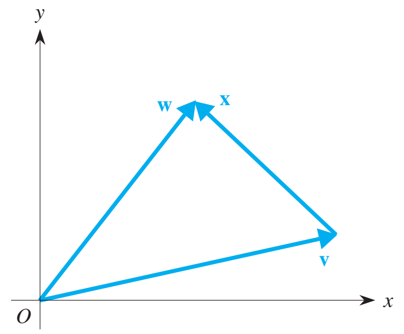


FIGURE 10.13

In general, the difference of two vectors always connects the end of the second vector to the end of the first. The only question is: Which way does the difference vector point? The easiest way to decide that is to draw a sketch such as Figure 10.13.

Another way of looking at the difference $\mathbf{v} - \mathbf{w}$ is to think of it as $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w})$, where

$$-\mathbf{w} = (-1)[w_1 \quad w_2] = [-w_1 \quad -w_2],$$

a vector with the same length as $\mathbf{w}$ but pointing in the opposite direction.

Problems

1. Plot each position vector as an arrow in the xy-plane from the origin to the point having the appropriate coordinates.

   a. $\mathbf{r} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$ **b.** $\mathbf{s} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

   c. $\mathbf{t} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$ **d.** $\mathbf{u} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$

   e. $\mathbf{v} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$

2. Using the vectors in Problem 1, plot the result of:

   a. Adding vector $\mathbf{u}$ to vector $\mathbf{r}$.
   b. Adding vector $\mathbf{r}$ to $\mathbf{u}$. Compare your answer to the vector obtained in part (a).
   c. Adding $\mathbf{t}$ to $\mathbf{s}$.
   d. Adding $\mathbf{u}$ to $\mathbf{v}$.

3. Using the vectors in Problem 1, plot each vector in a–c on the same graph.

   a. The result of adding one-half of $\mathbf{r}$ to one-half of $\mathbf{s}$.
   b. The result of adding one-quarter of $\mathbf{r}$ to three-quarters of $\mathbf{s}$.
   c. The result of adding three-quarters of $\mathbf{r}$ to one-quarter of $\mathbf{s}$.
   d. Plot $\mathbf{r}$ and $\mathbf{s}$. Draw a straight line joining these two vectors. Which of the vectors drawn in parts (a), (b), and (c) lie on this line?

4. Determine the magnitude of the position vectors from the origin to the following points.

a. $(3, 4)$ **b.** $(12, 5)$
c. $(3, -2)$ **d.** $(-7, -3)$
e. $(1, 2, 2)$ **f.** $(2, -3, 4)$

5. Determine the magnitude of the displacement vector from point A to point B for each pair of points.

a. $A = (1, 2), B = (5, 5)$
b. $A = (-2, -1), B = (4, 1)$
c. $A = (1, -3), B = (-3, -4)$
d. $A = (1, 2, 3), B = (4, 5, 3)$
e. $A = (-1, 6, 3), B = (3, -1, 2)$

6. Determine the vector that is the given multiple of the vector $[1 \quad 2]$.

a. 2 **b.** 7
c. -2 **d.** 0

7. Determine the vector that is the given multiple of the vector $[3 \quad -1 \quad 2]$.

a. 3 **b.** 10
c. -7 **d.** $\frac{1}{2}$

8. Find the unit vector that points in the same direction as the given vector.

a. $[3 \quad 4]$ **b.** $[1 \quad 1]$
c. $[0 \quad 5]$ **d.** $[1 \quad -1]$
e. $[1 \quad 2 \quad 2]$ **f.** $[1 \quad -1 \quad 1]$

9. Express each vector as a sum of multiples of the coordinate vectors $\mathbf{i} = [1 \quad 0]$ and $\mathbf{j} = [0 \quad 1]$.

a. $[2 \quad 1]$ **b.** $[-1 \quad 3]$
c. $[\frac{1}{2} \quad \frac{1}{2}]$ **d.** $[3 \quad 0]$

10. Express each vector as a sum of multiples of the three coordinate vectors $\mathbf{i} = [1 \quad 0 \quad 0]$, $\mathbf{j} = [0 \quad 1 \quad 0]$, and $\mathbf{k} = [0 \quad 0 \quad 1]$.

a. $[1 \quad 2 \quad 3]$ **b.** $[2 \quad 0 \quad 3]$
c. $[-1 \quad -3 \quad 1]$ **d.** $[0 \quad \frac{1}{2} \quad 0]$

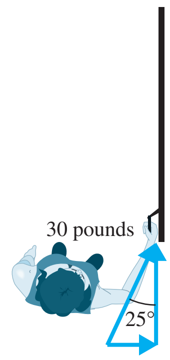
11. Refer to Example 4 in the text. What is the upward force on the window if Tom exerts

a. a force of 30 lb at an angle of 25°?
b. a force of 20 lb at an angle of 30°?
c. a force of 40 lb at an angle of 15°?

12. a. A sliding door is difficult to open. If Claire exerts a horizontal force of 30 lb at an angle of 25° to the sliding door, what is the effective force on the door in the direction in which the door slides?

b. Repeat part (a) with a force of 25 lb at an angle of 40°.

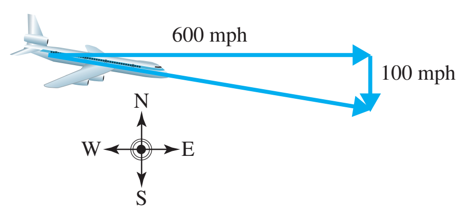
c. Repeat part (a) with a force of 40 lb at an angle of 20°.



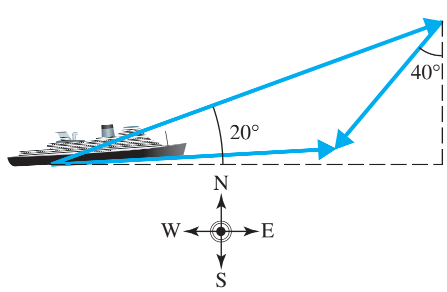
13. a. If a jet plane is flying on a heading of due east at 600 mph and the wind is blowing due south at 100 mph, what are the actual direction and speed of the plane?

b. Repeat part (a) if the plane is flying due east at 600 mph and the wind is blowing in a direction that is 40° south of east at 100 mph.

c. Repeat part (a) if the plane is flying southwest (45° south of west) at 300 mph and the wind is blowing in a direction 50° south of west at 100 mph.



14. Suppose that a boat is moving at 10 mph in the direction of 20° north of east across a bay and the tide is moving the water in the bay at 4 mph in the direction of 40° west of south. What are the actual direction and speed of the boat? (*Hint:* Express both the boat's vector and the tide's vector in terms of the two coordinate vectors **i** and **j**).



10.2 Linear Models

The real-world problems to which people apply mathematical models often involve large and very complex situations. For instance, one might want to analyze the effect that imposing a 50¢ per gallon tax on gasoline would have on the national economy with its thousands of interdependent businesses and industries. An airline must have a reservations system that takes into account all its aircraft, the cities it serves, flight schedules, dates, and different fare structures in effect. A company may need a battery of tests that can predict how well applicants will perform at a given job. The data in such problems usually come in the form of a rectangular array of numbers, called a **matrix.** For example, if four students, Ann, Bob, Carol, and Dan, take exams in French, mathematics, and sociology, the set of exam results could be displayed in the matrix

$$\begin{array}{l} \\ \text{French} \\ \text{Mathematics} \\ \text{Sociology} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Ann} & \text{Bob} & \text{Carol} & \text{Dan} \end{array} \\ \begin{bmatrix} 84 & 73 & 82 & 85 \\ 88 & 78 & 94 & 92 \\ 76 & 81 & 83 & 78 \end{bmatrix} \end{array}$$

Thus, for instance, Bob received an 81 in sociology and Carol a 94 in math.

Matrix algebra provides a systematic way of working with such arrays of numbers. In this chapter, we develop the basic language and methods of matrix algebra that will allow you to use matrices to solve various problems involving systems of linear equations.

We refer to a rectangular array of numbers as an $m \times n$ matrix when it has m rows horizontally and it has n columns vertically. A 3×4 matrix thus has 3 horizontal rows and 4 vertical columns, as in the preceding matrix of exam scores. We use boldface capital letters, such as **A**, to denote matrices in print. (When writing matrices by hand, you may find it convenient to use a wavy line under the letter, as in $\utilde{A}$.)

Two more examples of matrices are

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & -1 \\ 1 & 7 & 2 \\ 6 & 5 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 7 & 4 \end{bmatrix}.$$

Here **A** is a 3×3 matrix and **N** is a 2×3 matrix because it has two rows across and three columns vertically.

We denote the entry in row i and column j of matrix **A** by a_{ij}. Thus in matrix **A**, $a_{13} = -1$ because -1 is the entry in the first row and the third column, whereas $a_{31} = 6$ because 6 is the entry in the third row and first column. Similarly, in matrix **N**, $n_{23} = 4$ because 4 is the entry in the second row and the third column.

When a matrix has only one row or only one column, we call it a **row vector** or a **column vector,** respectively, or simply a vector. A vector having three numbers is called a **3-vector,** whereas a vector consisting of n numbers is called an ***n*-vector.** As noted in Section 10.1, we use boldface lowercase letters, such as **b** or **x**, to denote vectors. Note that any vector is also a matrix. Some examples of vectors are

$$\mathbf{b} = [1 \quad 3 \quad 0 \quad 5] \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}.$$

Here **b** is a row 4-vector and **x** is a column 3-vector. We write b_1 for the first entry in vector **b,** b_2 for the second entry in **b,** and b_i for the ith entry in **b.** Thus for the vectors **b** and **x,** we have $b_2 = 3$ and $x_2 = 4$.

Recall from Section 10.1 that two vectors **v** and **w** are *equal,* written $\mathbf{v} = \mathbf{w}$, if all their corresponding entries, or *components,* are equal. Similarly, two matrices are *equal* if all their corresponding components are the same.

Any list of numbers can be thought of as a column vector or a row vector. Whether we choose a column or row format if only vectors are involved usually doesn't matter, but when a vector and a matrix are multiplied, it is important to distinguish clearly whether the vector is a row vector or a column vector. For reasons that will be clear shortly, we *usually treat most vectors as column vectors.* Note that a column n-vector is an $n \times 1$ matrix and a row n-vector is a $1 \times n$ matrix.

An $m \times n$ matrix **A** can be thought of as a set of n column m-vectors or as a set of m row n-vectors. In the case of students and their test results, each column vector of the matrix gives the scores for one student in all these courses, whereas each row vector gives the scores in one course for all these students.

We use the following notation to refer to rows and columns in a matrix:

$\mathbf{a}_j$ denotes the jth column vector in **A**; and

$\mathbf{a}_i'$ denotes the ith row vector in **A**.

For instance, in the matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & -1 \\ 1 & 7 & 2 \\ 6 & 5 & 0 \end{bmatrix}, \qquad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 7 \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{a}_1' = \begin{bmatrix} 5 & 1 & -1 \end{bmatrix}.$$

A Geometric View of Vectors

In Section 10.1, we presented vectors geometrically as positions and displacements in coordinate space. As we pointed out there, vectors can be used to represent points in space. In two-dimensional space, we use a 2-vector; in three-dimensional space, we use a 3-vector. The point $(5, 2)$ in the plane can be thought of as the 2-vector $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$. Similarly, the point $(3, 2, 7)$ in three-dimensional space with coordinates $x = 3$, $y = 2$, $z = 7$, or equivalently, $x_1 = 3$, $x_2 = 2$, $x_3 = 7$, can be written as the 3-vector

$$\begin{bmatrix} 3 \\ 2 \\ 7 \end{bmatrix}.$$

Thus the coordinates of a point become the components of a position vector.

We next consider how matrices in general and vectors in particular occur in applied problems from many different fields.

A Clothes Production Model

A textile company runs three clothing factories. Each factory produces three types of women's clothing: vests, pants, and coats. For simplicity, we assume that one size fits all. Suppose that the first factory produces 20 vests, 10 pants, and 5 coats from each roll of cloth. The second and third factories produce different amounts of these three products, as described in matrix **A**.

$$\begin{array}{r} \\ \text{Vests} \\ \text{Pants} \\ \text{Coats} \end{array}
\begin{array}{c} \begin{array}{ccc} \text{Factory 1} & \text{Factory 2} & \text{Factory 3} \end{array} \\ \begin{bmatrix} 20 & 4 & 4 \\ 10 & 14 & 5 \\ 5 & 5 & 12 \end{bmatrix} \end{array} = \mathbf{A}$$

Each column of **A** is a vector of clothing produced by a factory from one roll of cloth. For instance, Factory 3 has the output vector

$$\mathbf{a}_3 = \begin{bmatrix} 4 \\ 5 \\ 12 \end{bmatrix},$$

which indicates that it makes 4 vests, 5 pants, and 12 coats from each roll. Each row of **A** is a vector of factory production of one particular type of clothing from one roll of cloth. The row vector for coats is $\mathbf{a}_3' = [5 \quad 5 \quad 12]$, which indicates that Factory 1 produces 5 coats, Factory 2 produces 5 coats, and Factory 3 produces 12 coats from each roll.

Let x_1 denote the number of rolls of cloth used by the first factory; similarly, x_2 and x_3 denote the numbers of rolls used by the second and third factories, respectively. Suppose that the company gets an order for 500 vests, 850 pants, and 1000 coats. This triple of numbers is called the *demand,* which we write as a column vector

$$\begin{bmatrix} 500 \\ 850 \\ 1000 \end{bmatrix}.$$

Then x_1, x_2, and x_3 need to satisfy the system of linear equations

$$\begin{aligned}
\text{vests:}\quad 20x_1 + 4x_2 + 4x_3 &= 500 \\
\text{pants:}\quad 10x_1 + 14x_2 + 5x_3 &= 850 \\
\text{coats:}\quad 5x_1 + 5x_2 + 12x_3 &= 1000.
\end{aligned}$$

In words, the vests equation says: The number of vests produced by Factory 1, $20x_1$ (this expression is 20 vests per roll times the x_1 rolls used by Factory 1), plus the number of vests produced by Factory 2, which is $4x_2$, plus the number of vests produced by Factory 3, which is $4x_3$, must equal the demand of 500 vests.

As we demonstrate in Section 10.3, we can write this system of linear equations as the *matrix–vector equation*

$$\begin{bmatrix} 20 & 4 & 4 \\ 10 & 14 & 5 \\ 5 & 5 & 12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 500 \\ 850 \\ 1000 \end{bmatrix}.$$

We can also write this as a single *vector equation* in the column vectors of the matrix as

$$x_1 \begin{bmatrix} 20 \\ 10 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 14 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ 5 \\ 12 \end{bmatrix} = \begin{bmatrix} 500 \\ 850 \\ 1000 \end{bmatrix}.$$

If **b** is the demand vector on the right side of the matrix–vector equation and **x** is a (column) vector of the x_i's, matrix algebra gives us a way to write the system of linear equations concisely in terms of **A**, **b**, and **x** as $\mathbf{Ax} = \mathbf{b}$. We discuss how to do this in Section 10.3. In Section 10.5, we extend these ideas to solve any system of three equations in three variables by using matrix algebra techniques.

The expression $2x_1 + 4x_2 - \frac{1}{2}x_3$ is a *linear expression* in three variables; it involves only the first power of the variables x_1, x_2, and x_3. Other cases are the expressions on the left side of the preceding system of linear equations. More formally, a linear expression is one that involves a sum of terms made up of constants multiplying individual variables that are raised only to the first power. In contrast, a nonlinear expression involves one or more variables that are raised to various powers (different from 1), or exponential, logarithmic, trigonometric, or other more complex expressions. (This terminology is similar to that used to describe linear difference equations.) The term *linear* is used to indicate that a "line-like" graph is associated with each variable in the expression. For example, the vest expression $20x_1 + 4x_2 + 4x_3$ is a linear expression; if x_2 and x_3 are fixed—say, $x_2 = x_3 = 3$—with only x_1 remaining as a free variable, the resulting expression $20x_1 + 4(3) + 4(3)$, or $20x_1 + 24$, defines a function $y = 20x_1 + 24$ whose graph is a line.

The clothing production equations form what is called a **linear model** because the equations involve only linear expressions (linear equations). In the following sections, we will return to this and other models introduced here as we develop the mathematical methods needed to analyze linear models.

A Markov Chain Model for the Stock Market

We next develop a linear model for the behavior of the stock market. Here we show how matrix methods can be used to represent a situation in which the values of a number of variables at one stage of a process are related to their values at the preceding stage.

Each business day, the stock market goes up, goes down, or stays the same. Suppose that historical studies show that if the market goes up one day—say today—the probability is $\frac{1}{4}$ that it will go up tomorrow, the probability is $\frac{1}{2}$ that it will go down tomorrow, and the probability is $\frac{1}{4}$ that it will stay the same tomorrow. If the market goes down today, there are three other observed probabilities for tomorrow's market performance. Similarly, if the market stays the same today, there is a third set of three probabilities for what will happen tomorrow. We can conveniently display all nine of these probabilities in a matrix **A:**

$$\begin{array}{cc} & \textbf{Market Today} \\ & \begin{array}{ccc} \text{Up} & \text{Down} & \text{Same} \end{array} \\ \textbf{Market Tomorrow} \begin{array}{r} \text{Up} \\ \text{Down} \\ \text{Same} \end{array} & \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} = \mathbf{A}. \end{array}$$

The probabilities in this matrix are called *transition probabilities* because they give us information about how to relate one stage of a process to the next. The matrix **A** is called a **transition matrix.** Each column corresponds to a type of market movement today, and each row corresponds to a type of market movement tomorrow. The matrix entry a_{23}, which equals $\frac{1}{2}$, is in the "down tomorrow" row and in the "same today" column. The value $\frac{1}{2}$ represents the probability that the market will go down tomorrow given that it stays the same today.

Note that the probabilities in each column of the transition matrix *must add to 1* because they include all possible outcomes for tomorrow given a particular type of market behavior today. A mathematical model such as this with given transition probabilities is known as a **Markov process,** or **Markov chain** (named after Russian mathematician Andrei Markov, who first developed these ideas). A convenient way to display the information in a Markov chain is with a *transition diagram,* such as the one shown in Figure 10.14. In this diagram, there are three *nodes,* one for each type of market movement: go up (U), go down (D), or stay the same (S). These are the possible *states* for the system. Note also that we indicate each transition probability with an arrow.

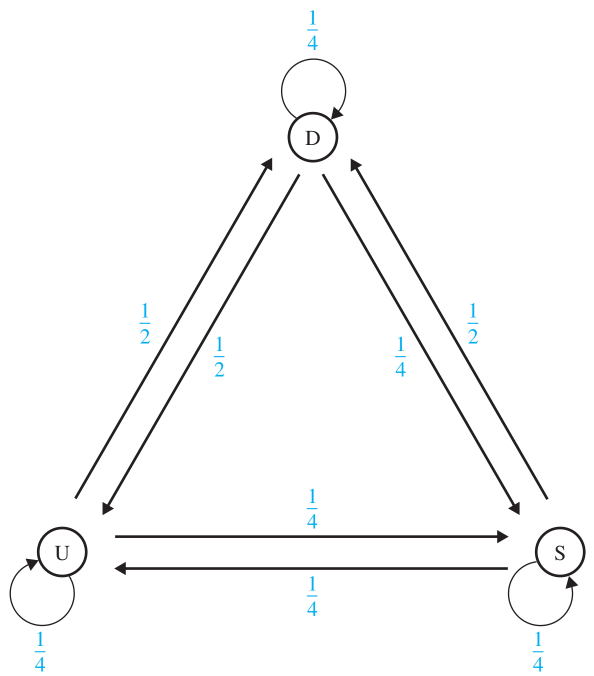


FIGURE 10.14

EXAMPLE 1

Suppose that, before the stock market opens today, we believe that there is a 50–50 chance of the market going down or staying the same, but no chance of its going up. Use the values in the preceding transition matrix **A** to compute the probabilities of the market being in any of the three states tomorrow—up, down, or the same—based on the probabilities of the market being up, down, or the same today.

Solution Let p_1, p_2, and p_3 denote today's probabilities of the market being up, down, and the same, respectively, and let p_1^+, p_2^+, and p_3^+ denote tomorrow's probabilities of being up, down, and the same, respectively. Let's see how to compute p_1^+. First, to compute the probability of two successive events—such as (i) being in State 1 today (probability p_1) followed by (ii) switching from State 1 today to State 1 tomorrow (probability $\frac{1}{4}$)—we multiply the probabilities of the two events and get $\frac{1}{4}p_1$. Similarly, the probability of (i) being in State 2 today (probability p_2) followed by (ii) switching from State 2 to State 1 tomorrow (probability $\frac{1}{2}$) is the product $\frac{1}{2}p_2$. Also, the probability of (i) being in State 3 today (probability p_3) followed by (ii) switching from State 3 to State 1 tomorrow (probability $\frac{1}{4}$) is $\frac{1}{4}p_3$.

To get the total probability p_1^+ that the stock market will go up tomorrow, we add these three values to get

$$p_1^+ = \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3.$$

In the same way, we calculate the probabilities p_2^+ and p_3^+ that the stock market goes down or stays the same tomorrow and so obtain a set of three equations based on the transition matrix **A**:

$$\begin{aligned} p_1^+ &= \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3 \\ p_2^+ &= \frac{1}{2}p_1 + \frac{1}{4}p_2 + \frac{1}{2}p_3 \\ p_3^+ &= \frac{1}{4}p_1 + \frac{1}{4}p_2 + \frac{1}{4}p_3. \end{aligned} \tag{1}$$

Note that the coefficients in this system of linear equations come directly from the entries in the transition matrix **A.**

Believing that $p_1 = 0$, $p_2 = \frac{1}{2}$, and $p_3 = \frac{1}{2}$ are today's probabilities of the market being up, down, and the same, respectively, we calculate the probability p_1^+ that the market goes up tomorrow, using the first of Equations (1), as follows:

$$p_1^+ = \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3 = \frac{1}{4}\cdot 0 + \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2} = \frac{3}{8}.$$

In the same way, using Equations (1), we obtain the probabilities for the other two market outcomes tomorrow. Thus tomorrow's probabilities p_1^+, p_2^+, and p_3^+ are

$$\begin{aligned} p_1^+ &= \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3 = \frac{1}{4}\cdot 0 + \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2} = \frac{3}{8}, \\ p_2^+ &= \frac{1}{2}p_1 + \frac{1}{4}p_2 + \frac{1}{2}p_3 = \frac{1}{4}\cdot 0 + \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{3}{8}, \\ p_3^+ &= \frac{1}{4}p_1 + \frac{1}{4}p_2 + \frac{1}{4}p_3 = \frac{1}{4}\cdot 0 + \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2} = \frac{1}{4}. \end{aligned}$$

EXAMPLE 2

Use the equations for p_1^+, p_2^+, and p_3^+ to predict the market probabilities p_1^{++}, p_2^{++}, and p_3^{++} two days ahead. Then predict the market probabilities farther into the future.

Solution We repeat the process in Example 1, using $p_1^+ = \frac{3}{8}$, $p_2^+ = \frac{3}{8}$, and $p_3^+ = \frac{1}{4}$ to obtain

$$\begin{aligned} p_1^{++} &= \frac{1}{4}p_1^+ + \frac{1}{2}p_2^+ + \frac{1}{4}p_3^+ = \frac{1}{4}\cdot\frac{3}{8} + \frac{1}{2}\cdot\frac{3}{8} + \frac{1}{4}\cdot\frac{1}{4} = \frac{11}{32}, \\ p_2^{++} &= \frac{1}{2}p_1^+ + \frac{1}{4}p_2^+ + \frac{1}{2}p_3^+ = \frac{1}{2}\cdot\frac{3}{8} + \frac{1}{4}\cdot\frac{3}{8} + \frac{1}{2}\cdot\frac{1}{4} = \frac{13}{32}, \\ p_3^{++} &= \frac{1}{4}p_1^+ + \frac{1}{4}p_2^+ + \frac{1}{4}p_3^+ = \frac{1}{4}\cdot\frac{3}{8} + \frac{1}{4}\cdot\frac{3}{8} + \frac{1}{4}\cdot\frac{1}{4} = \frac{8}{32} = \frac{1}{4}. \end{aligned}$$

From these probabilities for 2 days hence, we can predict the market 3 days ahead, and so on indefinitely, so long as the probabilities of the market going up, going down, or staying the same continue to hold. In Section 10.4, we introduce a far simpler way to perform these calculations based on matrix algebra. For now, we simply indicate the results in the following table, assuming that today's market probabilities are 0 for going up, $\frac{1}{2}$ for going down, and $\frac{1}{2}$ for staying the same.

	Up	**Down**	**Same**
Today	0	$\frac{1}{2}$	$\frac{1}{2}$
Tomorrow	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{4}$
2 days ahead	$\frac{11}{32}$	$\frac{13}{32}$	$\frac{1}{4}$
3 days ahead	$\frac{45}{128}$	$\frac{51}{128}$	$\frac{1}{4}$
5 days ahead	≈ 0.35	≈ 0.40	0.25
10 days ahead	≈ 0.35	≈ 0.40	0.25
100 days ahead	0.35	0.40	0.25

Each triple of probabilities for any day—say p_1, p_2, p_3 for the first day— can be thought of as the components of a 3-vector of probabilities for that day. Also, note that on any given day, the sum of the probabilities is always 1 because one of the possibilities (up, down, or the same) must occur.

The sequence of all the successive vectors $\mathbf{p}, \mathbf{p}^+, \mathbf{p}^{++}, \mathbf{p}^{+++}, \ldots$ associated with any transition matrix $\mathbf{A}$ is called a *Markov chain* because the successive vectors are linked by the matrix $\mathbf{A}$. Eventually, the successive probabilities in this Markov chain stabilize at 0.35 for the market going up, 0.40 for the market going down, and 0.25 for the market staying the same. That is, the probabilities converge over time to these limiting values. This behavior occurs regardless of the initial values we used for today's probabilities. Later, we formulate a system of three linear equations in three variables and solve it to determine these stable probabilities directly.

A Population Growth Model

The following model relates populations of hares and wolves from one week to the next. To make the numbers work out conveniently, we measure the hare population in groups of 10 hares and the wolf population in single wolves. Suppose that the number of groups of hares H grows by 20% per week when no wolves are present, so the population of hares next week H^+ would be $1.2H$. But W wolves *are* present and the wolves eat the hares at the rate of each wolf eating 3 hares each week, which is $30\% = 0.3$ of a group of 10 hares. Thus the hare population is reduced by $0.3W$ per week, giving $H^+ = 1.2H - 0.3W$ as next week's hare population.

Next, without hares present, the wolf population decreases at a rate of 30% per week, so the wolf population next week W^+ would be $0.7W$. But the wolf population grows each week when hares are present at the rate of one wolf for every 50 hares or every 5 groups of 10 hares, which is $\frac{1}{5} = 0.2$ wolf per group of 10 hares. Hence $W^+ = 0.2H + 0.7W$.

Together, these two equations are our model for the hare and wolf population over time:

$$H^+ = 1.2H - 0.3W$$
$$W^+ = 0.2H + 0.7W.$$

EXAMPLE 3

Suppose that we start with 1000 groups of hares and 800 wolves. Use the preceding expressions for H^+ and W^+ to calculate the populations of hares and wolves over time.

Solution If we start with $H = 1000$ and $W = 800$, the hare–wolf model predicts that

$$H^+ = 1.2(1000) - 0.3(800) = 960$$
$$W^+ = 0.2(1000) + 0.7(800) = 760$$

as the populations after the first week. We now use these values to predict the two populations after the second week:

$$H^{++} = 1.2H^+ - 0.3W^+ = 1.2(960) - 0.3(760) = 924$$
$$W^{++} = 0.2H^+ + 0.7W^+ = 0.2(960) + 0.7(760) = 724.$$

Extending these values from one week to the next, we obtain the following table for the sizes of the hare and wolf populations over time.

Weeks	Groups of hares	Wolves
0	1000	800
1	960	760
2	924	724
3	892	692
10	739	539
20	649	449
50	602	402
100	600	400

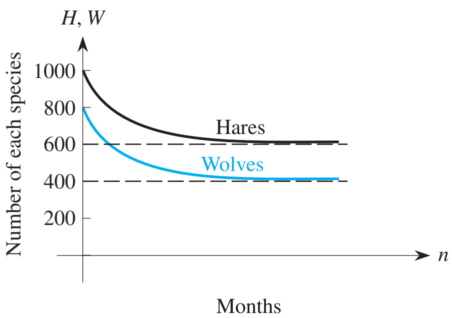


FIGURE 10.15

Note that over time the populations converge to 600 groups of 10 hares and 400 wolves. Figure 10.15 shows the graphs of both populations as functions of time.

We can visualize this situation another way: We can think one population depends on the other. That is, the number of wolves W can be viewed as a function of the number of hares H. If we plot the number of wolves versus groups of hares, we find that they fall in a straight line. In particular, the linear function that fits the points in the preceding table is $W = H - 200$. The graph in Figure 10.16 shows several *trajectories* for the populations (in hundreds) of groups of 10 hares and wolves. One trajectory shown starts from the initial point $(10, 8)$ and leads to the point $(6, 4)$.

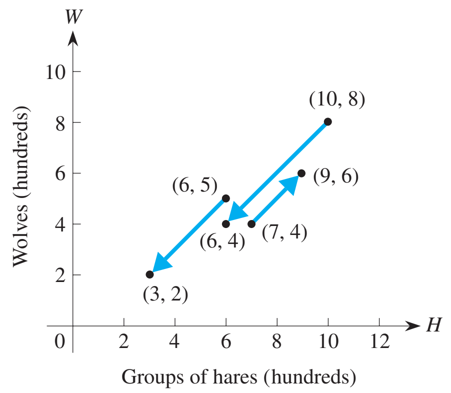


FIGURE 10.16

Suppose we start with a different set of initial values for the two populations—say, $H_0 = 700$ and $W_0 = 400$. Then

$$H^+ = 1.2(700) - 0.3(400) = 720$$
$$W^+ = 0.2(700) + 0.7(400) = 420$$

and so on. The resulting points all lie on a line starting at the point $(700, 400)$ and converge to the point $(900, 600)$. Similarly, if we start with $H_0 = 600$ and $W_0 = 500$, the resulting points all lie on a line and converge to $(300, 200)$, as also shown in Figure 10.16. In each case, the limiting values for H and W satisfy $W = \frac{2}{3}H$. That is, the limiting points lie on the line $W = \frac{2}{3}H$, as shown in Figure 10.17. In fact, for *any* initial pair of population values (H_0, W_0), the points of successive pairs (H_n, W_n) all lie on some line, and in each case the successive points are converging (as indicated by the arrows) toward a limiting point on the line $W = \frac{2}{3}H$, as illustrated in Figure 10.18. Thus under this model, all populations, regardless of the initial values, converge over time to populations in which the number of wolves is two-thirds the number of groups of 10 hares.

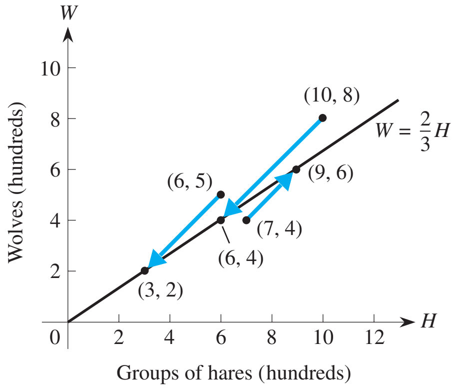


FIGURE 10.17

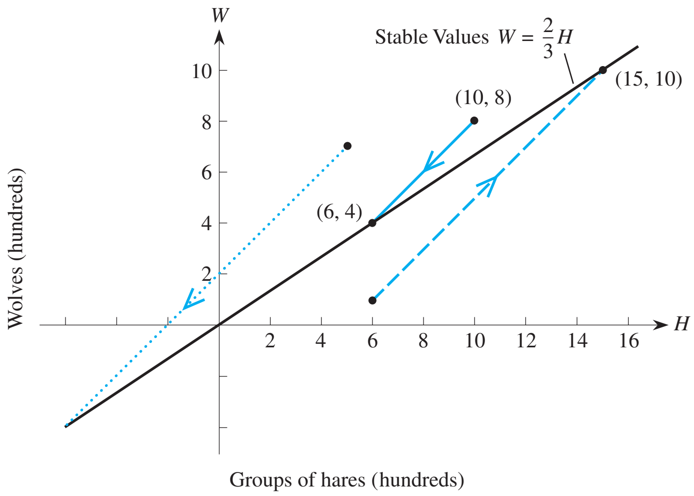


FIGURE 10.18

Note that the pair of equations we used to define this model can be rewritten as a pair of difference equations:

$$H_{n+1} = 1.2H_n - 0.3W_n$$
$$W_{n+1} = 0.2H_n + 0.7W_n.$$

We present a more detailed analysis of population models based on systems of difference equations, including a more sophisticated predator–prey model, in supplementary Section 12.6.

Problems

1. Ted, Carol, and Alice took tests in German, physics, theater, and politics. Ted's test scores in these subjects were 64, 73, 86, 85; Carol's scores were 82, 69, 77, 91; Alice's were 82, 84, 81, 83. Construct a matrix of these test results. Label the columns and rows.

2. For the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 5 & 3 \\ 6 & 1 & 7 \\ 6 & 9 & 5 \\ 0 & 2 & 8 \end{bmatrix},$$

write the following row and column vectors and entries.

a. $\mathbf{a}_1'$
b. $\mathbf{a}_3$
c. $\mathbf{a}_4$
d. a_{12}
e. a_{23}
f. a_{41}

3. In the matrix of letters

$$\mathbf{A} = \begin{bmatrix} H & R & B & I \\ N & S & O & A \\ E & T & Y & L \\ M & G & D & I \end{bmatrix},$$

spell the words represented by the following sequences of entries.

a. $a_{11}\ a_{23}\ a_{12}\ a_{21}$
b. $a_{41}\ a_{23}\ a_{32}\ a_{11}\ a_{31}\ a_{12}$
c. $a_{31}\ a_{24}\ a_{12}\ a_{34}\ a_{33}$
d. $a_{33}\ a_{31}\ a_{22}\ a_{32}\ a_{31}\ a_{12}\ a_{43}\ a_{24}\ a_{33}$

4. A clothing company's three factories (1, 2, and 3) produce the following numbers of vests, pants, and coats from each roll of cloth.

$$\begin{array}{l} \\ \text{Vests} \\ \text{Pants} \\ \text{Coats} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Factory 1} & \text{Factory 2} & \text{Factory 3} \end{array} \\ \begin{bmatrix} 6 & 4 & 2 \\ 4 & 8 & 4 \\ 3 & 2 & 8 \end{bmatrix} \end{array}$$

Suppose that the company has a demand for 400 vests, 800 pants, and 500 coats. Write a system of equations whose solution would determine production levels to yield the desired numbers of vests, pants, and coats. As in the clothes production model, let x_i be the number of rolls of cloth processed by the ith factory.

5. Three oil refineries (1, 2, and 3) produce the following amounts, in thousands of gallons, of heating oil, diesel oil, and gasoline from each shipment of crude petroleum.

$$\begin{array}{l} \\ \text{Heating Oil} \\ \text{Diesel Oil} \\ \text{Gasoline} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Refinery 1} & \text{Refinery 2} & \text{Refinery 3} \end{array} \\ \begin{bmatrix} 8 & 5 & 3 \\ 2 & 5 & 5 \\ 3 & 7 & 6 \end{bmatrix} \end{array}$$

Suppose that demand is for 6200 thousand gallons of heating oil, 4000 thousand gallons of diesel oil, and 4700 thousand gallons of gasoline. Write a system of equations whose solution would determine production levels to yield the desired amounts of heating oil, diesel oil, and gasoline. Let x_i be the number of shipments processed by the ith refinery.

6. The staff dietitian at the California Institute of Trigonometry has to make up a meal with 600 calories, 20 grams of protein, and 200 mg of vitamin C. The three food types that the dietitian can choose from are gelatin, fish sticks, and mystery meat. They have the following nutritional content per unit.

$$\begin{array}{l} \\ \text{Calories} \\ \text{Protein} \\ \text{Vitamin C} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Gelatin} & \text{Fish Sticks} & \text{Mystery Meat} \end{array} \\ \begin{bmatrix} 10 & 50 & 200 \\ 1 & 3 & 0.2 \\ 30 & 10 & 0 \end{bmatrix} \end{array}$$

Construct a mathematical model for this situation, based on a system of three linear equations.

7. A company has a budget of \$280, 000 for computing equipment. The types of equipment available are microcomputers at \$2000 each, terminals at \$500 each, and workstations at \$5000 each. There should be five times as many terminals as microcomputers and twice as many microcomputers as workstations. Write a system of three linear equations to describe this situation.

8. In the clothes production model in the text, suppose that Factory 1 processes 15 rolls of cloth, Factory 2 processes 20 rolls, and Factory 3 processes 60 rolls. For which product, vests, pants, or coats, does production deviate the most from the demand for 600, 800, 1000?

9. Refer to the stock market Markov chain in Example 1. Determine the set of probabilities for tomorrow's market for each set of probabilities that the market will be up, down, or the same today.

   a. $p_1 = 1, p_2 = 0, p_3 = 0$
   b. $p_1 = 0, p_2 = \frac{1}{2}, p_3 = \frac{1}{2}$
   c. $p_1 = \frac{1}{2}, p_2 = 0, p_3 = \frac{1}{2}$
   d. $p_1 = \frac{1}{4}, p_2 = \frac{1}{2}, p_3 = \frac{1}{4}$
   e. $p_1 = 0.35, p_2 = 0.40, p_3 = 0.25$

10. The copy machine at the student union breaks down according to the following pattern. If it is working today, it has a 70% chance of working tomorrow (and a 30% chance of breaking down). If the copy machine is broken today, it has a 50% chance of working tomorrow (and a 50% chance of being broken again).

    a. Construct a Markov chain for this situation; give the matrix of transition probabilities and draw the transition diagram.
    b. If there is a 50–50 chance of the copy machine's working today, what is the chance of its working tomorrow?
    c. Based on the situation in part (b), what is the chance that the copy machine is working the day after tomorrow?
    d. If the copy machine is working today, what is the chance that it is working the day after tomorrow?

11. The Pins, a bowling team, plays in a bowling league each week. If they win this week's game, they have a $\frac{2}{3}$ chance of winning next week's game. If they lose this week's game, they have a $\frac{1}{2}$ chance of winning next week's game.

    a. Construct a Markov chain for this situation; give the matrix of transition probabilities and draw the transition diagram.
    b. If there is a 50–50 chance of the Pins' winning this week's game, what is their chance of winning next week's game?
    c. If they won this week, what is their chance of winning the game 2 weeks from now?

12. Consider a weather Markov chain having two states: sunny and cloudy. If today is sunny, there is a $\frac{3}{4}$ probability that tomorrow will be sunny and a $\frac{1}{4}$ probability that tomorrow will be cloudy. If today is cloudy, there is a $\frac{1}{4}$ probability that tomorrow will be sunny and a $\frac{3}{4}$ probability that tomorrow will be cloudy.

    a. Write the transition matrix for this Markov chain and draw its transition diagram.
    b. In this weather Markov chain, starting with the vector of probabilities $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (a sunny day), compute and plot the vectors of probabilities for four successive days.
    c. Repeat the process starting with the probability vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (a cloudy day). Can you guess the values of the equilibrium state to which your probability vectors are converging?

13. The following model for learning a concept over a set of lessons identifies four states of learning: I = ignorance, E = exploratory thinking, S = superficial understanding, and M = mastery. If you are now in state I, after one lesson you have a probability of $\frac{1}{2}$ of still being in I and a probability of $\frac{1}{2}$ of being in E. If you are now in state E, after one lesson you have a probability of $\frac{1}{4}$ of being in I, $\frac{1}{2}$ in E, and $\frac{1}{4}$ in S. If you are now in state S, after one lesson you have a probability of $\frac{1}{4}$ of being in E, $\frac{1}{2}$ in S, and $\frac{1}{4}$ in M. If you are in M, you always stay in M (with a probability of 1).

    a. Construct a Markov chain for this learning model.
    b. If you start in state I, what is your probability vector after two lessons? After three lessons?

14. **a.** In Example 3, if initially there were 800 groups of 10 hares and 300 wolves, how many hares and wolves would there be after 1 month, 3 months,

5 months, and 10 months? What do the limiting values for the two populations appear to be?

b. Repeat part (a) with an initial population of 600 groups of 10 hares and 800 wolves.

c. Repeat part (a) with an initial population of 900 groups of 10 hares and 600 wolves.

15. Consider the following cattle–sheep models in which the two species compete for common grazing land. In each case, compute the populations after 1 month, 2 months, and 3 months if the initial populations are 50 cattle and 100 sheep.

a. $C^+ = 1.2C - 0.3S$
$S^+ = -0.2C + 1.2S$

b. $C^+ = 1.2C - 0.1S$
$S^+ = 0.5C + 1.4S$

16. Consider the rabbit–fox model

$$R^+ = 1.1R - 0.2F$$
$$F^+ = 0.2R + 0.6F.$$

Plot the following on the same graph.

a. The trajectory of populations starting from (10, 15).

b. The trajectory of populations starting from (10, 30).

c. The trajectory of populations starting from (20, 10).

17. In Example 3 we found that the solution to the system

$$H^+ = 1.2H - 0.3W$$
$$W^+ = 0.2H + 0.7W$$

converged to a point on the line $W = \frac{2}{3}H$ for any starting values H_0 and W_0.

a. To find the equation of this limiting line, assume that $H^+ = H$ and $W^+ = W$ in the two equations defining the system and solve these two equations in two unknowns.

b. For any given starting populations—say, $H_0 = 1000$ and $W_0 = 800$—calculate the next point (H^+, W^+) on the trajectory. What is the equation of the line through (H_0, W_0) and (H^+, W^+)?

c. You now have the equation of the limiting line and the equation of the trajectory. Describe how you would use them to find the final population values for H and W. What are the values for initial population values of $H_0 = 1000$ and $W_0 = 800$?

18. a. The population models in Example 3 and Problems 15–17 all involve linear expressions in H and W. Suppose that the equations for H^+ and W^+ contained nonlinear expressions in H and W. How might such expressions affect the trajectory?

b. We develop a more sophisticated mathematical model for two species, known as the predator–prey model, in supplementary Section 12.6. It is based on equations such as

$$H^+ = 1.2H - 0.003HW$$
$$W^+ = 0.2W + 0.005HW.$$

If the initial population values are $H_0 = 200$ and $W_0 = 100$, calculate and plot the population values over the first 3 months. What do you observe about the trajectory?

10.3 Scalar Products

In this section, we explain how to multiply two vectors in what is called a scalar product. For instance, suppose that a family normally eats three vegetables—asparagus, beans, and corn. Suppose that the costs per pound for these vegetables are \$0.80 for asparagus, \$1.00 for beans, and \$0.60 for corn. We can then form a vector $\mathbf{p} = [0.80 \quad 1.00 \quad 0.60]$ for the prices of these three vegetables. Suppose further that the family consumes 2 lb of asparagus, 5 lb of beans, and 3 lb of corn each week, so we can also form a vector $\mathbf{d} = [2 \quad 5 \quad 3]$ of the family's weekly demand for these vegetables. The total cost of the family's weekly demand for the vegetables is

$$2 \times 0.80 + 5 \times 1.00 + 3 \times 0.60 = 1.60 + 5.00 + 1.80 = \$8.40$$

because the cost for asparagus is $2 \times \$0.80$, the cost for beans is $5 \times \$1.00$, and the cost for corn is $3 \times \$0.60$. This result suggests a natural way to define the product of the two vectors **d** and **p**. We write this product as $\mathbf{d} \cdot \mathbf{p}$ and call it the **scalar product** of **d** and **p**. Here the scalar product $\mathbf{d} \cdot \mathbf{p}$ is

$$\begin{aligned}\mathbf{d} \cdot \mathbf{p} &= [2 \quad 5 \quad 3] \cdot [0.80 \quad 1.00 \quad 0.60] \\ &= 2(0.80) + 5(1.00) + 3(0.60) \\ &= 1.60 + 5.00 + 1.80 = 8.40,\end{aligned}$$

or \$8.40. Note that the scalar product involves multiplying the corresponding entries in each position of the vectors and adding the results. Each vector in a scalar product can be either a row or a column vector, but multiplication makes sense only when the two vectors have the same size (that is, they have the same number of entries). The scalar product of two vectors of different sizes can't be formed because there would be terms that do not match. For instance, the scalar product of $[1 \quad 2 \quad 3 \quad 4 \quad 5]$ and $[20 \quad 40 \quad 60]$ can't be formed.

Recall that the word *scalar* means a single number, as opposed to a vector or matrix. The scalar product of two vectors is so named because its result is a single number—a scalar. The scalar product also is known as the *dot product* and the *inner product,* but we will not use either of these terms.

The product of two vectors can also be defined in a different way, known as the *vector product,* which produces a vector instead of a scalar, or number, as the result. However, we don't consider it here.

More formally we have the following definition of the scalar product.

Scalar Product

Let

$$\mathbf{a} = [a_1 \quad a_2 \quad \dots \quad a_n] \quad \text{and} \quad \mathbf{b} = [b_1 \quad b_2 \quad \dots \quad b_n]$$

be vectors of the same size n. Then the *scalar product* $\mathbf{a} \cdot \mathbf{b}$ of **a** and **b** is the single number (a scalar) equal to the sum of the products,

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Figure 10.19 will help you visualize how to calculate the scalar product of two vectors **a** and **b**.

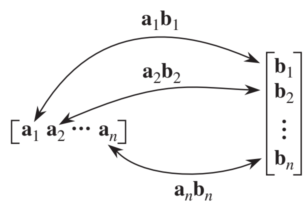


FIGURE 10.19

For instance, the scalar product of the vectors $\mathbf{a} = [3 \quad 2 \quad -5 \quad 4]$ and $\mathbf{b} = [6 \quad -3 \quad 2 \quad 0]$ is

$$\mathbf{a} \cdot \mathbf{b} = 3(6) + 2(-3) + (-5)(2) + 4(0) = 18 - 6 - 10 + 0 = 2.$$

We now see how the scalar product arises in a variety of situations.

EXAMPLE 1

Suppose that peaches cost 30¢ each, pears cost 20¢ each, apples cost 35¢ each, and grapefruits cost 50¢ each. Amy wants to get 5 peaches, 3 pears, 2 apples, and 2 grapefruits, and Bill wants to get 3 peaches, 4 pears, 3 apples, and 3 grapefruits.

a. Write vectors to represent the prices of the fruits and the amount of each fruit that Amy and Bill will purchase.

b. Write the total costs of their fruit purchases, using vector methods.

Solution

a. We form the price vector $\mathbf{p} = [0.30 \quad 0.20 \quad 0.35 \quad 0.50]$ for the prices of the respective fruits and the two demand vectors, $\mathbf{a} = [5 \quad 3 \quad 2 \quad 2]$ for Amy and $\mathbf{b} = [3 \quad 4 \quad 3 \quad 3]$ for Bill.

b. The scalar products $\mathbf{a} \cdot \mathbf{p}$ and $\mathbf{b} \cdot \mathbf{p}$ give the total costs of Amy's and Bill's purchases:

$$\begin{aligned}
\mathbf{a} \cdot \mathbf{p} &= [5 \quad 3 \quad 2 \quad 2] \cdot [0.30 \quad 0.20 \quad 0.35 \quad 0.50] \\
&= 5(0.30) + 3(0.20) + 2(0.35) + 2(0.50) \\
&= 1.50 + 0.60 + 0.70 + 1.00 = 3.80; \\
\mathbf{b} \cdot \mathbf{p} &= [3 \quad 4 \quad 3 \quad 3] \cdot [0.30 \quad 0.20 \quad 0.35 \quad 0.50] \\
&= 3(0.30) + 4(0.20) + 3(0.35) + 3(0.50) \\
&= 0.90 + 0.80 + 1.05 + 1.50 = 4.25.
\end{aligned}$$

Thus the cost of fruit was \$3.80 for Amy and \$4.25 for Bill.

A Geometric View of Scalar Products

An interesting special case of the scalar product involves the use of coordinate vectors, which we introduced in Section 10.1. In two dimensions, they are the unit vectors $[1 \quad 0]$ along the horizontal axis and $[0 \quad 1]$ along the vertical axis. See Figure 10.20. In three dimensions, the coordinate vectors are $[1 \quad 0 \quad 0]$, $[0 \quad 1 \quad 0]$, and $[0 \quad 0 \quad 1]$. They lie along three mutually perpendicular axes, as illustrated in Figure 10.21.

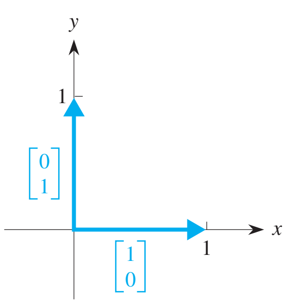


FIGURE 10.20

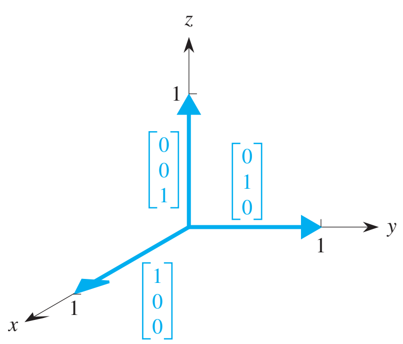


FIGURE 10.21

Any vector can be built from its components. For instance, for the 3-vector $[2 \quad 5 \quad 3]$, we use the components 2, 5, and 3 and the coordinate vectors to write

$$[2 \quad 5 \quad 3] = 2[1 \quad 0 \quad 0] + 5[0 \quad 1 \quad 0] + 3[0 \quad 0 \quad 1].$$

In general, for any 3-vector $\mathbf{a} = [a_1 \quad a_2 \quad a_3]$, we can write

$$\mathbf{a} = a_1[1 \quad 0 \quad 0] + a_2[0 \quad 1 \quad 0] + a_3[0 \quad 0 \quad 1].$$

In many geometric uses of vectors, we need to know whether two vectors "point" in the same general direction or in opposite directions. When two vectors point in approximately the same general direction, as shown by the arrows in Figure 10.22, their scalar product will be positive. When two vectors point in approximately opposite directions, as shown in Figure 10.23, their scalar product will be negative. Most interestingly, when two vectors form a right angle, as shown in Figure 10.24, their scalar product will be zero.

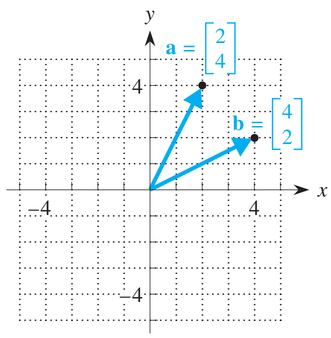


FIGURE 10.22

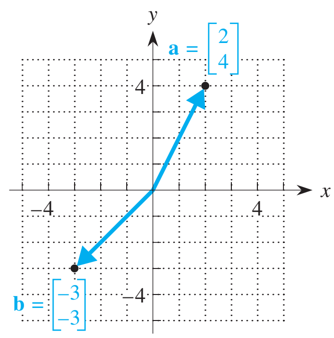


FIGURE 10.23

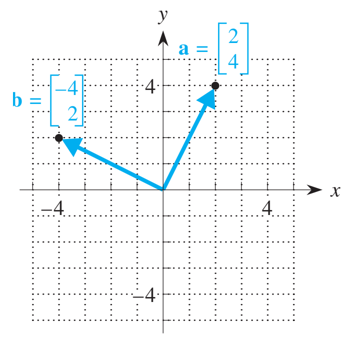


FIGURE 10.24

These assertions follow from the fact that, if θ is the angle between the vectors **a** and **b**, as shown in Figure 10.25, we have the following relationship.

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{\sqrt{\mathbf{a}\cdot\mathbf{a}}\sqrt{\mathbf{b}\cdot\mathbf{b}}}$$

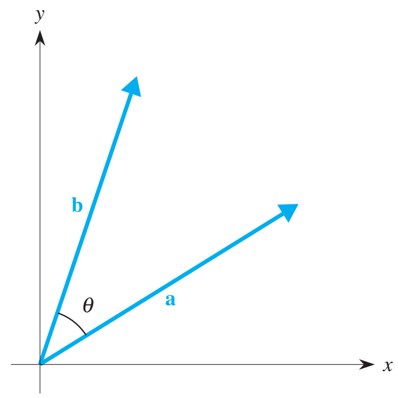


FIGURE 10.25

(This formula is based on the Law of Cosines introduced in Section 6.5; we ask you to derive this formula in a problem at the end of the section.) Note that, if $\mathbf{a} = [a_1 \quad a_2]$ is a 2-vector, $\mathbf{a}\cdot\mathbf{a} = a_1^2 + a_2^2$. We interpret $\mathbf{a}\cdot\mathbf{a}$ geometrically by using the Pythagorean theorem to represent the square of the hypotenuse in a right triangle with sides a_1 and a_2, as shown in Figure 10.26. Thus, in the formula for $\cos\theta$, $\sqrt{\mathbf{a}\cdot\mathbf{a}}$ is the length of the vector **a**. Similarly, if **a** is a 3-vector, $\mathbf{a} = [a_1 \quad a_2 \quad a_3]$, then

$$\mathbf{a}\cdot\mathbf{a} = a_1^2 + a_2^2 + a_3^2$$

and so $\sqrt{\mathbf{a}\cdot\mathbf{a}}$ is the length of **a**. Similarly, $\sqrt{\mathbf{b}\cdot\mathbf{b}}$ is the length of **b**. Therefore we can rewrite the formula for $\cos\theta$ as

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{(\text{length of }\mathbf{a})(\text{length of }\mathbf{b})}$$

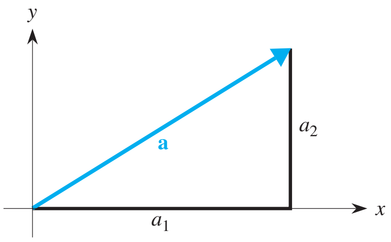


FIGURE 10.26

Clearly the length of **a** and the length of **b** are both positive. Thus, when the vectors point in roughly the same direction and the angle θ is between 0° and 90°, $\cos\theta$ is positive and so is $\mathbf{a}\cdot\mathbf{b}$. When the vectors point in roughly opposite directions and θ is between 90° and 180°, $\cos\theta$ is negative and so is $\mathbf{a}\cdot\mathbf{b}$. When θ is 90°, so that **a** and **b** are perpendicular, $\cos\theta$ is zero and so $\mathbf{a}\cdot\mathbf{b} = 0$.

EXAMPLE 2

Find the angle between the lines $y = 2x$ and $y = 3x$.

Solution We know that both lines pass through the origin, as illustrated in Figure 10.27. To find the angle θ between the lines at the origin, we need to find vectors along each line. Suppose that we arbitrarily choose $x = 1$. On the first line $y = 2x$, the corresponding value of y is 2, so the point $(1, 2)$ is on that line and the vector from the origin to that

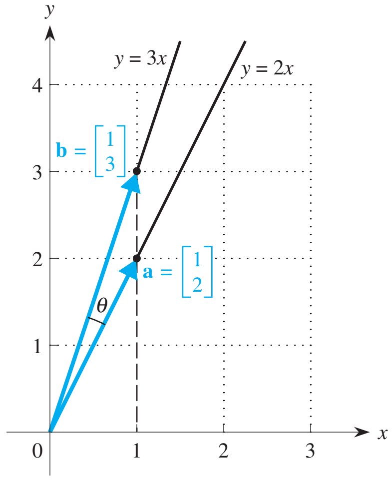


FIGURE 10.27

point is $\mathbf{a} = [1 \quad 2]$. Similarly, using $x = 1$, we find the corresponding point $(1, 3)$ on the second line $y = 3x$, so the vector from the origin to that point is $\mathbf{b} = [1 \quad 3]$. The angle θ between the lines is the same as the angle between the vectors, so

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{\sqrt{\mathbf{a}\cdot\mathbf{a}}\sqrt{\mathbf{b}\cdot\mathbf{b}}} = \frac{(1)(1) + (2)(3)}{\sqrt{(1)(1) + (2)(2)}\sqrt{(1)(1) + (3)(3)}}$$
$$= \frac{7}{\sqrt{5}\sqrt{10}} \approx 0.98995.$$

Consequently, the angle between the two lines is

$$\theta = \arccos(0.98995) \approx 8.13° \approx 0.1419 \text{ radian.}$$

The Clothes Production Model Using Scalar Products

In Section 10.2 we introduced a mathematical model for production in three clothing factories to meet a demand for 500 vests, 850 pants, and 1000 coats. If x_i denotes the number of rolls of cloth used by the ith factory ($i = 1, 2, 3$), the x_i's must satisfy the system of linear equations

$$\begin{aligned} \text{vests: } & 20x_1 + 4x_2 + 4x_3 = 500 \\ \text{pants: } & 10x_1 + 14x_2 + 5x_3 = 850 \\ \text{coats: } & 5x_1 + 5x_2 + 12x_3 = 1000. \end{aligned}$$

A key property of scalar products is that $\mathbf{a}\cdot\mathbf{b}$ is a linear combination of the entries in each vector. (Similarly, you can think of a polynomial as being a linear combination of power functions.) Conversely, any linear combination of variables or numbers always can be interpreted as a scalar product of two vectors.

EXAMPLE 3

Rewrite the three linear equations for the clothes production model using scalar products of vectors.

Solution Consider the first linear equation of the clothes production model

$$20x_1 + 4x_2 + 4x_3 = 500.$$

The left-hand side of this equation is a linear combination of the three variables. If

$$\mathbf{a} = [20 \quad 4 \quad 4] \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

(we explain shortly the reason for writing $\mathbf{a}$ as a row vector and $\mathbf{x}$ as a column vector), we can write the left-hand side of this equation as the scalar product

$$\mathbf{a}\cdot\mathbf{x} = [20 \quad 4 \quad 4]\cdot\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 20x_1 + 4x_2 + 4x_3.$$

The first equation is then $\mathbf{a} \cdot \mathbf{x} = 500$. Note that the vector **a** consists of the entries in the first row of the matrix

$$A = \begin{bmatrix} 20 & 4 & 4 \\ 10 & 14 & 5 \\ 5 & 5 & 12 \end{bmatrix}$$

that we used in the clothes production model of Section 10.2.

Using vectors **b** and **c** to represent the second and third rows of matrix **A,** we can write the other two linear equations in the same way as $\mathbf{b} \cdot \mathbf{x} = 850$ and $\mathbf{c} \cdot \mathbf{x} = 1000$.

Any system of linear equations can be written in terms of a system of scalar products. For example, the left-hand sides of the equations for the clothes production model are

$$\text{vests: } 20x_1 + 4x_2 + 4x_3 = [20 \quad 4 \quad 4] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{a}_1' \cdot \mathbf{x},$$

$$\text{pants: } 10x_1 + 14x_2 + 5x_3 = [10 \quad 14 \quad 5] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{a}_2' \cdot \mathbf{x},$$

$$\text{coats: } 5x_1 + 5x_2 + 12x_3 = [5 \quad 5 \quad 12] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{a}_3' \cdot \mathbf{x},$$

where $\mathbf{a}_i'$ is the ith row of the clothes production coefficient matrix

$$\begin{array}{l} \\ \text{Shirts} \\ \text{Pants} \\ \text{Coats} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Factory 1} & \text{Factory 2} & \text{Factory 3} \end{array} \\ \begin{bmatrix} 20 & 4 & 4 \\ 10 & 14 & 5 \\ 5 & 5 & 12 \end{bmatrix} \end{array} = \mathbf{A}.$$

The Matrix–Vector Product

Although we encounter many important uses of single scalar products in matrix algebra, their most important use is as a building block for defining the product of a matrix and a vector and, in Section 10.4, the product of two matrices. We define the matrix–vector product as follows.

Matrix–Vector Product

The product of an $m \times n$ matrix **A** and a column n-vector **c** is a column m-vector of scalar products

$$\mathbf{Ac} = \mathbf{a}_i' \cdot \mathbf{c} \quad (\text{each } \textit{row } \mathbf{a}_i' \text{ of } \mathbf{A} \text{ multiplies } \mathbf{c}).$$

The diagram in Figure 10.28 will help you visualize this definition. Think of each row in the first matrix **A** as a row vector (or think of **A** as consisting of a collection of row vectors). The scalar product of each row vector in **A** and the column vector **c** creates a new column vector. In order for this definition to make sense, the entries in each row of the matrix **A** (equivalently, the number of columns in matrix **A**) must equal the size of the vector **c**.

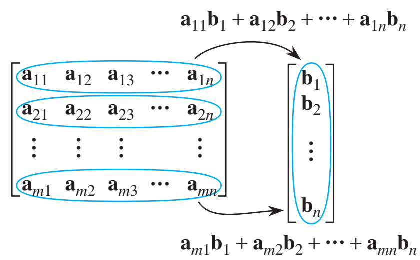


FIGURE 10.28

For instance, if **A** is a 2×3 matrix and **c** is a column 3-vector,

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix},$$

then

$$\mathbf{Ac} = \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 2(3) + 1(4) + 0(2) \\ 5(3) + 3(4) + 6(2) \end{bmatrix} = \begin{bmatrix} 10 \\ 39 \end{bmatrix}.$$

As we said previously, each row of matrix **A** is treated as if it were a row vector, and it is used to form a scalar product with the column vector **c**. However, we cannot multiply

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 6 \end{bmatrix} \quad \text{and} \quad c = \begin{bmatrix} 3 \\ 4 \\ 2 \\ 5 \end{bmatrix}$$

because the number of entries in each row of **A** does not match the number of entries in **C**.

We can recast this illustration symbolically as follows. If

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

then

$$\mathbf{Ac} = \begin{bmatrix} \mathbf{a}_1{}' \cdot \mathbf{c} \\ \mathbf{a}_2{}' \cdot \mathbf{c} \end{bmatrix} = \begin{bmatrix} a_{11}c_1 + a_{12}c_2 + a_{13}c_3 \\ a_{21}c_1 + a_{22}c_2 + a_{23}c_3 \end{bmatrix}.$$

The only other allowable way to multiply a vector and a matrix is to multiply a row vector by a matrix. In that case, the matrix must have the same number of rows as the row vector has entries. That is, we can multiply

$$\mathbf{c} = [1 \quad 4 \quad -2 \quad 3] \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} 2 & 0 \\ 6 & -2 \\ -3 & 5 \\ 1 & -5 \end{bmatrix}$$

because there are four entries in the row vector, which matches the number of rows (vertical entries per column) in the matrix **A**.

EXAMPLE 4

Consider again the set of equations discussed in Example 3 for the clothes production model:

$$\begin{aligned} \text{vests:}\quad & 20x_1 + 4x_2 + 4x_3 = 500 \\ \text{pants:}\quad & 10x_1 + 14x_2 + 5x_3 = 850 \\ \text{coats:}\quad & 5x_1 + 5x_2 + 12x_3 = 1000. \end{aligned}$$

Rewrite the system of equations as a matrix–vector equation.

Solution If we make vectors of the left-hand sides of the three equations, we have

$$\begin{bmatrix} 20x_1 + 4x_2 + 4x_3 \\ 10x_1 + 14x_2 + 5x_3 \\ 5x_1 + 5x_2 + 12x_3 \end{bmatrix} = \begin{bmatrix} [20 \quad 4 \quad 4]\cdot\mathbf{x} \\ [10 \quad 14 \quad 5]\cdot\mathbf{x} \\ [5 \quad 5 \quad 12]\cdot\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1'\cdot\mathbf{x} \\ \mathbf{a}_2'\cdot\mathbf{x} \\ \mathbf{a}_3'\cdot\mathbf{x} \end{bmatrix} = \mathbf{Ax}.$$

If we let

$$\mathbf{b} = \begin{bmatrix} 500 \\ 850 \\ 1000 \end{bmatrix}$$

represent the column vector of demands for the different products, the system of equations becomes

$$\mathbf{Ax} = \mathbf{b} \quad \text{or} \quad \begin{bmatrix} 20x_1 + 4x_2 + 4x_3 \\ 10x_1 + 14x_2 + 5x_3 \\ 5x_1 + 5x_2 + 12x_3 \end{bmatrix} = \begin{bmatrix} 500 \\ 850 \\ 1000 \end{bmatrix}.$$

As Example 4 suggests, *any* system of linear equations can be written as a matrix–vector equation. As we develop the tools for working with and solving such equations, we demonstrate that having such a formulation has significant advantages, especially for systems that are larger than three equations in three unknowns.

The Fruit Purchase Model Revisited In Example 1, we computed the scalar products for the costs of fruit purchased by Amy and Bill. Recall that Amy wanted 5 peaches, 3 pears, 2 apples, and 2 grapefruits, whereas Bill wanted 3 peaches, 4 pears, 3 apples, and 3 grapefruits. Also, we know that peaches cost 30¢ each, pears 20¢ each, apples 35¢ each, and grapefruits 50¢ each.

Gord.3896.10.pgs 4/28/03 3:29 PM Page 679

EXAMPLE 5

Write a matrix–vector equation to represent the costs of the fruit purchases by Amy and Bill.

Solution Let's make a matrix **A** of fruit purchases. The columns represent the different fruits, the first row gives Amy's fruit shopping list, and the second row gives Bill's list. We also make a (column) vector **p** of the costs, in cents.

$$\mathbf{A} = \begin{array}{c} \begin{array}{cccc} \text{Peaches} & \text{Pears} & \text{Apples} & \text{Grapefruits} \end{array} \\ \begin{bmatrix} 5 & 3 & 2 & 2 \\ 3 & 4 & 3 & 3 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Amy} \\ \text{Bill} \end{array} \qquad \mathbf{p} = \begin{bmatrix} 30 \\ 20 \\ 35 \\ 50 \end{bmatrix} \begin{array}{l} \text{Peaches} \\ \text{Pears} \\ \text{Apples} \\ \text{Grapefruits} \end{array}$$

We can now write the costs of Amy's and Bill's fruit purchases as the matrix–vector product

$$\mathbf{Ap} = \begin{bmatrix} 5 & 3 & 2 & 2 \\ 3 & 4 & 3 & 3 \end{bmatrix} \begin{bmatrix} 30 \\ 20 \\ 35 \\ 50 \end{bmatrix} = \begin{bmatrix} 5(30) + 3(20) + 2(35) + 2(50) \\ 3(30) + 4(20) + 3(35) + 3(50) \end{bmatrix}$$

$$= \begin{bmatrix} 380 \\ 425 \end{bmatrix} = \begin{bmatrix} \$3.80 \\ \$4.25 \end{bmatrix}.$$

This result is the same as we obtained in Example 1.

Markov Chain for the Stock Market Revisited Recall the Markov chain introduced in Example 1 of Section 10.2. The equations for determining the probabilities p_1^+, p_2^+, and p_3^+ that the stock market goes up, goes down, or stays the same tomorrow given the probabilities p_1, p_2, and p_3 of its going up, going down, or staying the same today were

$$p_1^+ = \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3$$

$$p_2^+ = \frac{1}{2}p_1 + \frac{1}{4}p_2 + \frac{1}{2}p_3$$

$$p_3^+ = \frac{1}{4}p_1 + \frac{1}{4}p_2 + \frac{1}{4}p_3.$$

EXAMPLE 6

Write the preceding Markov chain model for the stock market as a matrix–vector equation.

Solution Let

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \quad \text{and} \quad \mathbf{p}^+ = \begin{bmatrix} p_1^+ \\ p_2^+ \\ p_3^+ \end{bmatrix},$$

respectively, be the vectors of today's and tomorrow's probabilities and let the matrix of transition probabilities be **A**:

$$\begin{array}{cc|ccc} & & \multicolumn{3}{c}{\textbf{Market Today}} \\ & & \text{Up} & \text{Down} & \text{Same} \\ \hline \textbf{Market} & \text{Up} & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \textbf{Tomorrow} & \text{Down} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ & \text{Same} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} = \mathbf{A}.$$

Then the preceding set of probability equations can be written simply as $\mathbf{p}^+ = \mathbf{Ap}$ because

$$\mathbf{Ap} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3 \\ \frac{1}{2}p_1 + \frac{1}{4}p_2 + \frac{1}{2}p_3 \\ \frac{1}{4}p_1 + \frac{1}{4}p_2 + \frac{1}{4}p_3 \end{bmatrix}.$$

Similarly, the probabilities $\mathbf{p}^{++}$ for the day after tomorrow can be found from

$$\mathbf{p}^{++} = \mathbf{Ap}^+$$

and so on.

In Section 10.4 we present a concise way of writing these probability vectors.

A Geometric View of Matrix–Vector Products

In geometric terms, when we multiply a vector $\mathbf{v}$ by some matrix $\mathbf{A}$, the vector $\mathbf{v}$ is transformed by the multiplication into another vector $\mathbf{w} = \mathbf{Av}$. Because of the dimensions of $\mathbf{A}$ and $\mathbf{v}$, this product makes sense only in the order $\mathbf{Av}$ (rather than $\mathbf{vA}$); we thus call this operation *pre-multiplication* of $\mathbf{v}$ by $\mathbf{A}$. We can view pre-multiplication by $\mathbf{A}$ as defining a function: $\mathbf{w} = f(\mathbf{v})$, where $f(\mathbf{v}) = \mathbf{Av}$. This type of transformation is used in computer graphics, for instance, to produce creative lettering and moving images (animation) on the screen. It also provides a way to visualize the effect on a vector of multiplying it by a matrix.

Rather than using arrows, here we simply represent 2-vectors as points in the plane.

EXAMPLE 7

Consider the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad \mathbf{v}_4 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$\mathbf{v}_5 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{v}_6 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \mathbf{v}_7 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \mathbf{v}_8 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Here, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ point to the corners of a square whose sides are of length 2 and $\mathbf{v}_5$, $\mathbf{v}_6$, $\mathbf{v}_7$, and $\mathbf{v}_8$ point to the midpoints of the sides of this square, as shown in Figure 10.29(a). Describe the effects of pre-multiplying each of these eight vectors by the matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

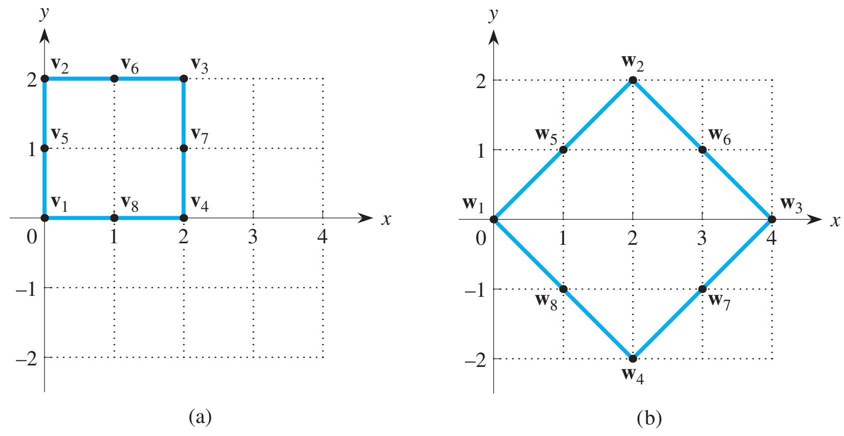


FIGURE 10.29

Solution When we pre-multiply the **v**'s by the matrix **A**, we obtain the following eight vectors $\mathbf{w}_i = \mathbf{A}\mathbf{v}_i$.

$$\mathbf{w}_1 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathbf{w}_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$\mathbf{w}_3 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \qquad \mathbf{w}_4 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$\mathbf{w}_5 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{w}_6 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{w}_7 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \qquad \mathbf{w}_8 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

When we examine these eight new vectors, shown in Figure 10.29(b), we observe that they also form a square, but it has a different orientation and size. Each of the four vertex vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ of the original square was transformed into a vertex vector $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$, and $\mathbf{w}_4$ of the new square. Each of the midpoint vectors $\mathbf{v}_5$, $\mathbf{v}_6$, $\mathbf{v}_7$, and $\mathbf{v}_8$ in the original square was transformed into a corresponding midpoint vector $\mathbf{w}_5$, $\mathbf{w}_6$, $\mathbf{w}_7$, and $\mathbf{w}_8$ in the new square. Also, in the original square, the sides were of length 2 and each diagonal was $2\sqrt{2}$, from the Pythagorean theorem. In the transformed square, each diagonal has length 4, so that the sides have length $2\sqrt{2}$. Thus pre-multiplying the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ forming the original square by **A** does three things to the square.

1. It rotates the square clockwise through 45° or $\pi/4$.
2. It increases the length of each side by a factor of $\sqrt{2}$.
3. It moves the center of the square from $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ to $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

Example 7 demonstrates that one square can be transformed into another by multiplying a set of vectors by an appropriate matrix. The same principle applies to any shape whose corners are determined by a set of vectors.

If we pre-multiplied the vectors forming the original square by other (appropriately chosen) matrices, we could get rotations of the square through any desired angle, increase or decrease the lengths of the sides by any desired multiple, and place the center in any desired location. However, the origin will always be unaffected because $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ multiplied by any matrix yields $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

For example, using $\mathbf{B} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ rotates the square counterclockwise through an angle of $\pi/2$ and transforms the center to $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. We ask you to verify this result in the Problems at the end of this section.

In general, pre-multiplying by the matrix

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

rotates the square counterclockwise through an angle of θ.

EXAMPLE 8

Show that the effect of pre-multiplying any vector $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ by the matrix $\mathbf{R}$ is to rotate $\mathbf{v}$ through an angle θ.

Solution We start with an arbitrary vector $\mathbf{v}$, as shown in Figure 10.30, that is inclined at an angle α from the horizontal, so that

$$\tan\alpha = \frac{b}{a}.$$

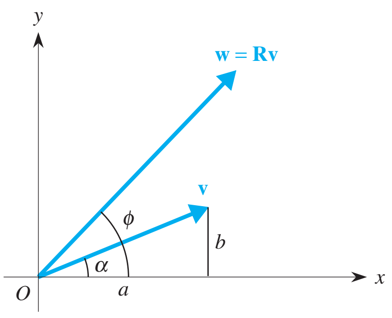


FIGURE 10.30

Also, we let the angle associated with the vector $\mathbf{w} = \mathbf{Rv}$ be ϕ. We want to show that the difference $\phi - \alpha$ in the two angles must equal θ.

When we pre-multiply vector $\mathbf{v}$ by $\mathbf{R}$ to form vector $\mathbf{w}$, we get

$$\mathbf{w} = \mathbf{Rv} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a\cos\theta - b\sin\theta \\ a\sin\theta + b\cos\theta \end{bmatrix}.$$

Consequently, the tangent of angle ϕ is

$$\tan\phi = \frac{a\sin\theta + b\cos\theta}{a\cos\theta - b\sin\theta}.$$

We factor a out of both the numerator and denominator and divide both the numerator and denominator by $\cos\theta$:

$$\tan\phi = \frac{a\cdot\left(\dfrac{\sin\theta}{\cos\theta}+\dfrac{b}{a}\right)}{a\cdot\left(1-\dfrac{b\sin\theta}{a\cos\theta}\right)} = \frac{\tan\theta+\dfrac{b}{a}}{1-\dfrac{b}{a}\tan\theta}.$$

However, we know that $\tan\alpha = b/a$, so that this expression reduces to

$$\tan\phi = \frac{\tan\theta+\tan\alpha}{1-\tan\alpha\tan\theta}.$$

When we compare this expression to the sum identity for the tangent from Problem 37 of Section 8.1, we see that

$$\tan\phi = \tan(\theta+\alpha) \quad \text{so that} \quad \phi = \theta+\alpha.$$

Therefore the effect of pre-multiplying any vector **v** by the matrix **R** is to rotate the resulting vector through an angle of θ.

What if we want to change the length of a vector by using matrix multiplication? We note that the lengths of the sides of the square in Example 8 are unchanged when we use this rotation matrix **R**. If we want to enlarge or contract the square, we must use an appropriate *scalar multiple* of the rotation matrix, so that we multiply every entry in the matrix by a constant amount. Thus, if

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \text{then} \quad 2\mathbf{A} = 2\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ -2 & 2 \end{bmatrix}.$$

In Example 7, to make each side of the square grow by a factor of 2 instead of a factor of $\sqrt{2}$, we would multiply **A** by $\sqrt{2}$ because $\sqrt{2}\cdot\sqrt{2} = 2$. Incidentally, not every matrix can be a rotation matrix; a special form is necessary.

Any transformation that takes a 2-vector **v** into the 2-vector $\mathbf{w} = \mathbf{A}\mathbf{v}$, for any 2×2 matrix **A**, always *transforms* or *maps* lines into lines. Because of this linearity property, a mapping of the form $\mathbf{v}\to\mathbf{w} = \mathbf{A}\mathbf{v}$ is called a *linear transformation.*

Suppose that we start with any figure comprising (very short) line segments. Each segment can be interpreted as a vector, and appropriate matrices can be constructed to create any kind of transformation—a shift, a stretch, or a rotation—that we desire. This method is the mathematical foundation of the computer graphics animation that appears in movies, on television, and on computer screens.

Problems

1. Let

$$\mathbf{a} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 5 \\ -1 \end{bmatrix},$$
$$\mathbf{d} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

Compute the scalar products.

a. $\mathbf{a}\cdot\mathbf{c}$ **b.** $\mathbf{b}\cdot\mathbf{c}$ **c.** $\mathbf{b}\cdot\mathbf{d}$
d. $\mathbf{a}\cdot\mathbf{d}$ **e.** $\mathbf{c}\cdot\mathbf{d}$ **f.** $\mathbf{b}\cdot\mathbf{e}$
g. $\mathbf{c}\cdot\mathbf{e}$

2. Let

$$\mathbf{a} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} -4 \\ 2 \end{bmatrix},$$
$$\mathbf{d} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \quad \text{and} \quad \mathbf{e} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Plot the pairs of vectors and compute their scalar products.

a. **a**, **b** **b.** **a**, **c** **c.** **a**, **d**
d. **b**, **d** **e.** **c**, **d** **f.** **d**, **e**

What geometric pattern do you find in pairs of vectors whose scalar products are zero? What numerical pattern do you find in the ratios of components in these pairs?

3. Let

$$\mathbf{a} = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$
$$\mathbf{c} = \begin{bmatrix} 7 \\ -2 \\ -1 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} \mathbf{-2} \\ \mathbf{3} \\ \mathbf{1} \end{bmatrix}.$$

Compute the scalar products.

a. $\mathbf{a} \cdot \mathbf{c}$ **b.** $\mathbf{b} \cdot \mathbf{c}$ **c.** $\mathbf{b} \cdot \mathbf{d}$
d. $\mathbf{a} \cdot \mathbf{d}$ **e.** $\mathbf{a} \cdot \mathbf{a}$ **f.** $\mathbf{c} \cdot \mathbf{d}$

4. Let

$$\mathbf{a} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix},$$
$$\mathbf{c} = \begin{bmatrix} 1 \\ 4 \\ 8 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}.$$

Compute the scalar products.

a. $\mathbf{a} \cdot \mathbf{c}$ **b.** $\mathbf{b} \cdot \mathbf{d}$ **c.** $\mathbf{c} \cdot \mathbf{d}$
d. $(\mathbf{a} + \mathbf{c}) \cdot \mathbf{a}$ **e.** $\mathbf{a} \cdot (\mathbf{c} + \mathbf{d})$
f. $(\mathbf{b} - \mathbf{a}) \cdot (\mathbf{a} + \mathbf{c})$

5. Let **a**, **b**, **c**, and **d** be as in Problem 1 and let

$$\mathbf{A} = \begin{bmatrix} 1 & 7 \\ 4 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 5 & 2 \\ -1 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & 1 \\ 4 & -2 \\ 3 & 1 \end{bmatrix}.$$

Compute the matrix–vector products.

a. **Ab** **b.** **Ac** **c.** **Ba**
d. **Bb** **e.** **Cc** **f.** **Ca**

6. Let **a**, **b**, **c**, and **d** be as in Problem 3 and let

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 0 & 4 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 3 \\ -1 & 5 & 1 \\ 4 & -1 & 6 \end{bmatrix},$$
$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \\ 9 & 10 & 11 \end{bmatrix}.$$

Find the matrix–vector products.

a. **Aa** **b.** **Ab** **c.** **Bc**
d. **Bd** **e.** **Cb** **f.** **Cc**

7. Let **a**, **b**, **c**, and **d** be as in Problem 3 and let

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

Compute the matrix–vector products and describe in words the effect of multiplying a vector by the particular matrix.

a. **Ac** **b.** **Ad** **c.** **Ba**
d. **Bd** **e.** **Cb** **f.** **Ca**

8. Explain why it is not possible to multiply the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}.$$

9. Write each vector–matrix equation as a system of equations. Here **x** denotes a column vector of variables $x_1, x_2, \ldots,$ where the number of variables equals the number of columns in **A**.

a. $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} = \begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$

b. $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & -3 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$$

c. $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 5 & 2 & 1 \\ 4 & 1 & 6 \\ 3 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix}$$

d. $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 5 \\ 3 & 1 & 2 \\ 5 & 1 & -3 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

10. Write each system of equations in matrix notation. Define any matrix or vector that you use.

a. $x_1 + 2x_2 = 6$
$2x_1 - 6x_2 = 4$

b. $5x_1 + 2x_2 + 4x_3 = 6$
$x_1 + 3x_2 - 2x_3 = 2$
$2x_1 - 5x_2 + 5x_3 = 5$

c. $2x_1 + 5x_2 - 2x_3 = 0$
$3x_1 + 8x_2 + 4x_3 = 0$
$x_1 + x_2 - 7x_3 = 0$

11. Write in matrix notation the systems of linear equations obtained in the following Problems in Section 10.2. Define any matrix or vector that you use.

a. Problem 4 **b.** Problem 5
c. Problem 6 **d.** Problem 7

12. Write in matrix notation the systems of linear equations obtained in the hare–wolf population model in Example 3 of Section 10.2. Define any matrix or vector that you use.

13. Suppose that you will need 10 hero sandwiches, 6 quarts of fruit punch, 3 pounds of potato salad, and 2 plates of hors d'oeuvres for a party. The matrix shows the cost per unit of these supplies from three different caterers.

	Caterer A	Caterer B	Caterer C
Hero sandwich	\$4	\$6	\$5
Fruit punch	\$2	\$1	\$0.85
Potato salad	\$1.50	\$2	\$2.50
Hors d'oeuvres	\$6	\$5	\$7

a. Express the cost of catering the party by each caterer as a matrix–vector product.
b. Determine the cost of each caterer.

14. Plot a square with corners determined by the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, and $\mathbf{v}_4$ in (a)and (b). Then, for the matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, plot the transformed corners $\mathbf{w}_1 = \mathbf{Av}_1$, $\mathbf{w}_2 = \mathbf{Av}_2$, $\mathbf{w}_3 = \mathbf{Av}_3$, and $\mathbf{w}_4 = \mathbf{Av}_4$. Confirm that the midpoints of the sides of the original square are mapped to the midpoints of the sides of the transformed square.

a. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$,
$\mathbf{v}_4 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

b. $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$,
$\mathbf{v}_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

15. Repeat Problem 14 with the matrix $\mathbf{B} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

In Example 8, we proved that the mapping of $\mathbf{v}$ to $\mathbf{w} = \mathbf{Bv}$ acts to rotate a square counterclockwise through an angle of $\pi/2$. Do your results confirm this outcome?

16. Transform each square in Problem 14 by using the matrix $\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Does the mapping of $\mathbf{v}$ to $\mathbf{w} = \mathbf{Cv}$ transform the square into a square?

17. Consider the rotation matrix $\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Show that, for any vector $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$, the vector $\mathbf{w} = \mathbf{Rv}$ has the same length as $\mathbf{v}$.

18. Use the matrix $\mathbf{R}$ from Problem 17 to determine the effect that the matrix $2\mathbf{R}$ has on the vector $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$.

19. The rotation matrix $\mathbf{R}$ in Problem 17 acts to rotate any nonzero vector counterclockwise through an angle θ.

a. Modify matrix $\mathbf{R}$ to produce a matrix that rotates any nonzero vector clockwise through an angle θ.
b. Write a matrix that will rotate any nonzero vector counterclockwise through an angle of 30°.
c. Plot the position vector $\mathbf{v} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, pre-multiply it by the matrix you created in part (b), and then plot the resulting vector $\mathbf{w}$ on the same graph.
d. Pre-multiply the vector $\mathbf{w}$ by the matrix you created in part (a) with $\theta = 30°$. What is the result of this operation?

20. The five points A, B, C, D, and E at $(3, 2)$, $(7, -1)$, $(9, 1)$, $(10, 3)$, and $(6, 6)$ determine a five sided figure

Gord.3896.10.pgs 11/21/03 11:05 AM Page 686

having two right angles. Use vectors to determine which of the five angles in the figure are the right angles.

21. Find the acute angle between each pair of vectors.
   a. $[3 \quad 5]$ and $[-2 \quad 4]$
   b. $[1 \quad 4]$ and $[-2 \quad 5]$
   c. $[1 \quad 4 \quad 5]$ and $[2 \quad 3 \quad -2]$
   d. $[6 \quad 4 \quad -1]$ and $[5 \quad -3 \quad 2]$

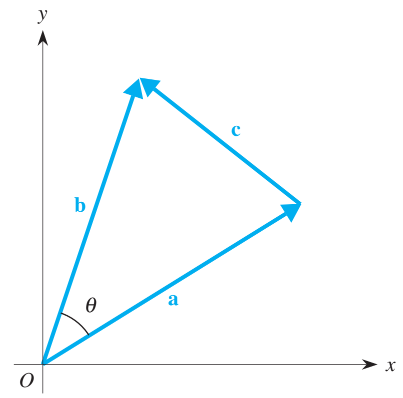
22. Rewrite each polynomial as the scalar product of a vector of numbers and a vector of power function terms of the form $\mathbf{x} = [1 \quad x \quad x^2]$ or $\mathbf{x}' = [1 \quad x \quad x^2 \quad x^3]$.
   a. $5 + 3x + 2x^2$
   b. $8 - 7x + 3x^2$
   c. $4 - 2x - 6x^2 + 5x^3$

23. When Susan was applying to college, she was turned down by her top choice, Ivy Tech, but she was accepted by State Tech and the Hawaii Institute of Technology. To decide on which school to attend, she rated each school (on a scale of 0 to 10) on five important criteria and then selected the one that was closer to Ivy Tech. She used her knowledge of vectors to decide how to interpret *closer*—the two vectors with the smallest angle between them. Which college did she choose?

	Ivy	State	Hawaii
Location	7	6	10
Size	6	3	7
Campus	8	4	7
Faculty	9	6	4
Programs	5	7	6

24. (Continuation of Problem 23) Suggest some other ways that Susan could have decided which school was closer to Ivy Tech? Would you necessarily make the same decision as to which school to choose? Explain.

25. (Derivation of the formula for cos θ.) In the triangle shown, let $\mathbf{a} = [a_1 \quad a_2]$ and $\mathbf{b} = [b_1 \quad b_2]$. The lengths of $\mathbf{a}$ and $\mathbf{b}$ can be written $\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2}$ and $\|\mathbf{b}\| = \sqrt{b_1^2 + b_2^2}$, so that $\|\mathbf{a}\|^2 = \mathbf{a} \cdot \mathbf{a}$ and $\|\mathbf{b}\|^2 = \mathbf{b} \cdot \mathbf{b}$.



   a. Write the vector $\mathbf{c}$ in terms of the vectors $\mathbf{a}$ and $\mathbf{b}$.
   b. Use the expression from part (a) to form the scalar product $\mathbf{c} \cdot \mathbf{c}$.
   c. Use the law of cosines to write an equation giving $\|\mathbf{c}\|^2$.
   d. Compare the expressions from part (b) and part (c) and use the facts that $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$, $\mathbf{b} \cdot \mathbf{b} = \|\mathbf{b}\|^2$, and $\mathbf{c} \cdot \mathbf{c} = \|\mathbf{c}\|^2$ to show that

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|}.$$

10.4 Matrix Multiplication

In Section 10.3, we introduced the concept of the scalar product $\mathbf{a} \cdot \mathbf{b}$ of two vectors $\mathbf{a}$ and $\mathbf{b}$. For example, if $\mathbf{a} = [2 \quad 1]$ and $\mathbf{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$, then $\mathbf{a} \cdot \mathbf{b} = 2(4) + 1(6) = 14$.

We used this scalar product of vectors to define the matrix–vector product $\mathbf{Ab}$ of a matrix $\mathbf{A}$ times a vector $\mathbf{b}$. For instance, if

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{and again} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix},$$

then

$$\mathbf{Ab} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 2(4) + 1(6) \\ 0(4) + 3(6) \end{bmatrix} = \begin{bmatrix} 14 \\ 18 \end{bmatrix}.$$

Thus **Ab** is a column vector consisting of the scalar products of each row of **A** with **b**.

In this section, we extend this process to define the product of two matrices **A** and **B.** In particular, in the product **AB** we think of the second matrix **B** as consisting of a series of column vectors and pre-multiply each of them by the matrix **A**. That is, we multiply the first column of **B** by **A**, then multiply the second column of **B** by **A**, and so on, as illustrated in Figure 10.31. The result of each product is a column vector, so the product **AB** will be a matrix having the same number of columns as there are in **B**. We can think of the product **AB** as a matrix whose columns are a sequence of matrix–vector products $[\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \ldots \quad \mathbf{Ab}_n]$.

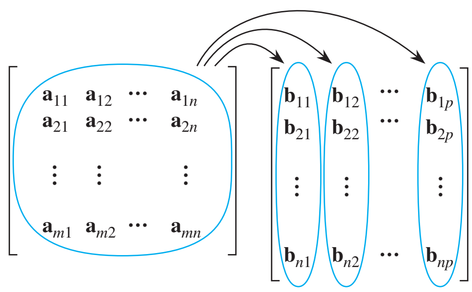


FIGURE 10.31

Think About This

In the matrix product **AB**, you can also think of the first matrix **A** as consisting of a series of row vectors and the second matrix **B** as consisting of a series of column vectors and then take the scalar product of each of the row vectors making up **A** with each of the column vectors making up **B**. Draw a sketch comparable to Figure 10.31 to illustrate this interpretation.

To demonstrate the product of two matrices, consider again the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{and let} \quad \mathbf{B} = \begin{bmatrix} 4 & 7 & -1 \\ 6 & 5 & 9 \end{bmatrix}.$$

We think of each of the three columns of matrix **B** as a column vector. Note that the first column $\mathbf{b}_1 = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ of **B** is precisely the column vector $\mathbf{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ we used above. The corresponding first column in the product matrix **AB** is then $\mathbf{Ab}_1$, which we computed above to be $\begin{bmatrix} 14 \\ 18 \end{bmatrix}$.

Similarly, we take the matrix–vector product of the matrix **A** with the second column of **B** thinking of it as the vector $\mathbf{b}_2 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$ to produce the second column of the product **AB.** That gives

$$\mathbf{Ab}_2 = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 7 \\ 5 \end{bmatrix} = \begin{bmatrix} 2(7) + 1(5) \\ 0(7) + 3(5) \end{bmatrix} = \begin{bmatrix} 19 \\ 15 \end{bmatrix}.$$

Finally, we take the matrix–vector product of **A** with the third column of **B**, which is the vector $\mathbf{b}_3 = \begin{bmatrix} -1 \\ 9 \end{bmatrix}$, to produce the third column of the product **AB**. The complete product **AB** is therefore

$$\mathbf{AB} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 4 & 7 & -1 \\ 6 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 2(4)+1(6) & 2(7)+1(5) & 2(-1)+1(9) \\ 0(4)+3(6) & 0(7)+3(5) & 0(-1)+3(9) \end{bmatrix}$$
$$= \begin{bmatrix} 14 & 19 & 7 \\ 18 & 15 & 27 \end{bmatrix}.$$

In general, we have the following definition for the product of two matrices.

Matrix Multiplication

Let **A** be an $m \times r$ matrix and **B** be an $r \times n$ matrix. The number of columns in **A** must equal the number of rows in **B**.
The matrix product **AB** is the $m \times n$ matrix obtained by forming the scalar product of each row $\mathbf{a}_i'$ in **A** with each column $\mathbf{b}_j$ in **B**. That is, the (i, j)th entry in **AB** is $\mathbf{a}_i'\mathbf{b}_j$. Thus

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1'\mathbf{b}_1 & \mathbf{a}_1'\mathbf{b}_2 & \cdots & \mathbf{a}_1'\mathbf{b}_n \\ \mathbf{a}_2'\mathbf{b}_1 & \mathbf{a}_2'\mathbf{b}_2 & \cdots & \mathbf{a}_2'\mathbf{b}_n \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_m'\mathbf{b}_1 & \mathbf{a}_m'\mathbf{b}_2 & \cdots & \mathbf{a}_m'\mathbf{b}_n \end{bmatrix}.$$

In summary, there are three ways to interpret matrix multiplication.

1. We can think of **AB** as the scalar product of each row of **A** with each column of **B**. Figure 10.32 illustrates this interpretation. The product requires that the number of entries in each row of **A** must equal the number of entries in each column of **B**.

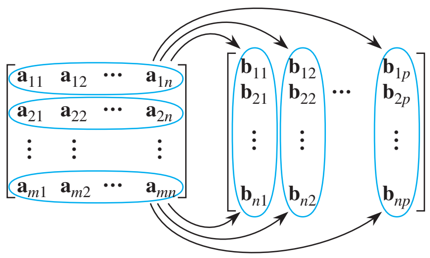


FIGURE 10.32

2. We can think of **AB** as a sequence of matrix–vector products—that is, the product of **A** with each of the column vectors making up **B**, so that

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \ldots \quad \mathbf{b}_n] = [\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \ldots \quad \mathbf{Ab}_n].$$

For example, for the matrices **A** and **B** above, we verify that the first column of **AB** is the matrix–vector product $\mathbf{Ab}_1$. Then

$$\mathbf{Ab}_1 = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 2(4) + 1(6) \\ 0(4) + 3(6) \end{bmatrix} = \begin{bmatrix} 14 \\ 18 \end{bmatrix}.$$

3. We can view **AB** as a set of vector-matrix products $\mathbf{a}_i{}'\mathbf{B}$ (defined analogously to matrix–vector products), where each row of **A**, considered as a row vector, pre-multiplies the matrix **B**. Thus

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1{}' \\ \mathbf{a}_2{}' \\ \vdots \\ \mathbf{a}_m{}' \end{bmatrix}\mathbf{B} = \begin{bmatrix} \mathbf{a}_1{}'\mathbf{B} \\ \mathbf{a}_2{}'\mathbf{B} \\ \vdots \\ \mathbf{a}_m{}'\mathbf{B} \end{bmatrix}.$$

For instance, for the matrices **A** and **B**, we verify that the first row of **AB** is the vector–matrix product $\mathbf{a}_1{}'\mathbf{B}$:

$$\mathbf{a}_1{}'\mathbf{B} = \begin{bmatrix} 2 & 1 \end{bmatrix}\begin{bmatrix} 4 & 7 & -1 \\ 6 & 5 & 9 \end{bmatrix}$$
$$= \begin{bmatrix} 2(4) + 1(6) & 2(7) + 1(5) & 2(-1) + 1(9) \end{bmatrix} = \begin{bmatrix} 14 & 19 & 7 \end{bmatrix}.$$

For the matrix product **AB** to make sense, the number of columns in **A** must equal the number of rows in **B.** Thus, if **A** is $m \times n$ (m rows and n columns) and **B** is $n \times k$ (n rows and k columns), the product **AB** is $m \times k$. For instance, the product of a 3×5 matrix and a 5×8 matrix will be a 3×8 matrix. But, the product of a 5×8 matrix and a 3×5 matrix is not defined—the numbers of columns and rows do not match, so it is not possible to perform the multiplication.

Note that our definition of a matrix–vector product in Section 10.3 is a special case of matrix multiplication because, in the matrix–vector product **Ab,** the column vector **b** can be interpreted as an $n \times 1$ matrix. Then **Ab** is the matrix product of an $m \times n$ matrix **A** and an $n \times 1$ matrix **b.** The result **Ab** is an $m \times 1$ matrix (a column m-vector).

In general, except in rather unusual circumstances, matrix multiplication is not commutative; that is, $\mathbf{AB} \neq \mathbf{BA}$. In fact, unless **A** and **B** are both square matrices with the same size, only one at most of **AB** and **BA** is defined. For instance, if **A** is 3×5 and **B** is 5×2, we can form **AB**, but not **BA.**

EXAMPLE 1

Given

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

find **AB** and **BA** and decide whether matrix multiplication is commutative.

Solution Using the given matrices, we find that

$$\mathbf{AB} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 2(3) + 1(1) & 2(-1) + 1(4) \\ 0(3) + 3(1) & 0(-1) + 3(4) \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 3 & 12 \end{bmatrix},$$

whereas

$$\mathbf{BA} = \begin{bmatrix} 3 & -1 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 3(2) + (-1)(0) & 3(1) + (-1)(3) \\ 1(2) + 4(0) & 1(1) + 4(3) \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 2 & 13 \end{bmatrix}.$$

Thus, for these two matrices, $\mathbf{AB} \neq \mathbf{BA}$, so matrix multiplication is not, in general, a commutative operation.

We also need a way to add two matrices. Recall how we added vectors in Example 3 of Section 10.1: We simply added the entries in the corresponding positions. For instance, the sum of the row vectors $\mathbf{c} = [2 \quad 1 \quad 5]$ and $\mathbf{d} = [4 \quad 3 \quad 0]$ is

$$\mathbf{c} + \mathbf{d} = [2 \quad 1 \quad 5] + [4 \quad 3 \quad 0] = [2+4 \quad 1+3 \quad 5+0] = [6 \quad 4 \quad 5].$$

We add matrices in the same way by simply adding all corresponding entries. However, for addition to make sense, the two matrices or the two vectors must be of the same size.

EXAMPLE 2

Find the sum of the two matrices

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

Solution Matrix addition involves adding the entries in the corresponding positions, so we have

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 3 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 2+3 & 1+(-1) \\ 0+1 & 3+4 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 1 & 7 \end{bmatrix}.$$

EXAMPLE 3

Find the sum of the 3×3 matrices

$$\mathbf{E} = \begin{bmatrix} 4 & 2 & 3 \\ 2 & 1 & 3 \\ 0 & 5 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{F} = \begin{bmatrix} 2 & 6 & -1 \\ 4 & 0 & 2 \\ 7 & 3 & 8 \end{bmatrix}.$$

Solution Again, by adding the entries in the corresponding positions, we find that

$$\mathbf{E} + \mathbf{F} = \begin{bmatrix} 4 & 2 & 3 \\ 2 & 1 & 3 \\ 0 & 5 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 6 & -1 \\ 4 & 0 & 2 \\ 7 & 3 & 8 \end{bmatrix} = \begin{bmatrix} 4+2 & 2+6 & 3+(-1) \\ 2+4 & 1+0 & 3+2 \\ 0+7 & 5+3 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 & 2 \\ 6 & 1 & 5 \\ 7 & 8 & 12 \end{bmatrix}.$$

Subtraction of matrices is defined in the analogous way—we simply take the difference of the corresponding entries in each position. However, the quotient of two matrices can't be defined.

We now summarize the laws of matrix algebra for matrix addition and multiplication. In each case we assume that the matrices have the appropriate sizes so that the operations make sense.

Basic Laws of Matrix Algebra

Associative Law

Matrix addition and matrix multiplication are associative:

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \quad \text{and} \quad (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}).$$

Commutative Law

Matrix addition is commutative:

$$A + B = B + A.$$

Matrix multiplication is *not* commutative (except in special cases):

$$AB \neq BA.$$

Distributive Law

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \quad \textit{and} \quad (\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$$

Law of Scalar Factoring

$$k(\mathbf{AB}) = (k\mathbf{A})\mathbf{B} = \mathbf{A}(k\mathbf{B}),$$

where k is a scalar constant.

With the exception that matrix multiplication is not commutative, these laws are basically the same as the laws used in algebra for working with real numbers. However, because the objects now are arrays and the operation of matrix multiplication is much more complicated than real-number multiplication, it is not at all obvious that these matrix laws should be true. Some effort is required to verify them (but this is beyond the scope of this chapter).

Because a vector is just a $1 \times n$ matrix or an $n \times 1$ matrix, these laws also apply to vectors. However, scalar products of vectors are commutative.

Matrix calculations are standard features on most calculators and in many software packages. Typically, you have to give a name for a matrix, such as **A**, then specify the *dimensions* of the matrix (the number of rows by the number of columns)—say, 3×4—and then enter the values in the appropriate positions. Once you have entered the matrices, you can use the calculator to perform any of the allowable operations such as sums, differences, and products. See your instruction manual for details.

The Fruit Purchase Model Revisited

In Example 1 of Section 10.3, we computed the scalar products of fruit costs and quantities of fruit to be purchased by Amy and Bill. Recall that Amy wanted 5 peaches, 3 pears, 2 apples, and 2 grapefruits, whereas Bill wanted 3 peaches, 4 pears, 3 apples, and 3 grapefruits. Also, peaches cost 30¢ each, pears 20¢ each, apples 35¢ each, and grapefruits 50¢ each. In Example 6 of Section 10.2, we constructed a matrix **A** of the fruit shopping lists of Amy and Bill and made a column vector **p** of the fruit costs (in cents).

$$\begin{array}{cc} & \begin{array}{cccc}\text{Peaches} & \text{Pears} & \text{Apples} & \text{Grapefruits}\end{array} \\ \mathbf{A} = & \begin{bmatrix} 5 & 3 & 2 & 2 \\ 3 & 4 & 3 & 3 \end{bmatrix} \begin{array}{l}\text{Amy}\\ \text{Bill}\end{array}\end{array} \qquad \mathbf{p} = \begin{bmatrix} 30 \\ 20 \\ 35 \\ 50 \end{bmatrix} \begin{array}{l}\text{Peaches}\\ \text{Pears}\\ \text{Apples}\\ \text{Grapefruits}\end{array}$$

The matrix-vector product $\mathbf{Ap} = \begin{bmatrix} 380 \\ 425 \end{bmatrix} = \begin{bmatrix} \$3.80 \\ \$4.25 \end{bmatrix}$ gave the cost of the fruit purchases of Amy and Bill.

EXAMPLE 4

Suppose now that there are two other stores at which Amy and Bill can shop for fruit. Instead of a vector of fruit prices, we now have a matrix **P** of fruit prices (whose first column is the original store's set of prices).

$$\begin{array}{cc} & \begin{array}{ccc}\text{Store 1} & \text{Store 2} & \text{Store 3}\end{array} \\ \mathbf{P} = & \begin{bmatrix} 30 & 25 & 30 \\ 20 & 25 & 25 \\ 35 & 40 & 30 \\ 50 & 60 & 45 \end{bmatrix} \begin{array}{l}\text{Peaches}\\ \text{Pears}\\ \text{Apples}\\ \text{Grapefruits}\end{array}\end{array}$$

Construct a matrix giving the cost to Amy and Bill of their fruit purchases at each store.

Solution We need to compute the matrix product **AP**, which is well defined becasue **A** is a 2×4 matrix and **P** is a 4×3 matrix:

$$\mathbf{AP} = \begin{bmatrix} 5 & 3 & 2 & 2 \\ 3 & 4 & 3 & 3 \end{bmatrix} \begin{bmatrix} 30 & 25 & 30 \\ 20 & 25 & 25 \\ 35 & 40 & 30 \\ 50 & 60 & 45 \end{bmatrix}$$

$$= \begin{bmatrix} 5(30) + 3(20) + 2(35) + 2(50) & 5(25) + 3(25) + 2(40) + 2(60) & 5(30) + 3(25) + 2(30) + 2(45) \\ 3(30) + 4(20) + 3(35) + 3(50) & 3(25) + 4(25) + 3(40) + 3(60) & 3(30) + 4(25) + 3(30) + 3(45) \end{bmatrix}$$

$$= \begin{bmatrix} 380 & 400 & 375 \\ 425 & 475 & 415 \end{bmatrix}.$$

Alternatively, had we entered the cost of fruit in the form \$0.30 instead of 30¢, say, we would get

$$\mathbf{AP} = \begin{bmatrix} \$3.80 & \$4.00 & \$3.75 \\ \$4.25 & \$4.75 & \$4.15 \end{bmatrix}.$$

Note that we would have gotten the same results using a calculator to multiply the two matrices.

Powers of Markov Chain Transition Matrices

Next, we consider a more substantial use of matrix multiplication—one that greatly expands the power of the Markov chain model for the stock market introduced in Section 10.2.

In the Markov chain model, the equations for determining the probabilities p_1^+, p_2^+, and p_3^+ of the market going up, going down, or staying the same tomorrow given the probabilities p_1, p_2, and p_3 of the market going up, going down, or staying the same today were

$$p_1^+ = \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3$$
$$p_2^+ = \frac{1}{2}p_1 + \frac{1}{4}p_2 + \frac{1}{2}p_3$$
$$p_3^+ = \frac{1}{4}p_1 + \frac{1}{4}p_2 + \frac{1}{4}p_3.$$

If

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \quad \text{and} \quad \mathbf{p}^+ = \begin{bmatrix} p_1^+ \\ p_2^+ \\ p_3^+ \end{bmatrix}$$

are the vector of today's probabilities and the vector of tomorrow's probabilities, respectively, and the matrix **A** of transition probabilities is

$$\begin{array}{cc} & \textbf{Market Today} \\ & \begin{array}{ccc} \text{Up} & \text{Down} & \text{Same} \end{array} \\ \begin{array}{c} \textbf{Market} \\ \textbf{Tomorrow} \end{array} \begin{array}{c} \text{Up} \\ \text{Down} \\ \text{Same} \end{array} & \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}, \end{array}$$

then the given system of transition probabilities can be written simply as $\mathbf{p}^+ = \mathbf{Ap}$, or, equivalently, as

$$\begin{bmatrix} p_1^+ \\ p_2^+ \\ p_3^+ \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}p_1 + \frac{1}{2}p_2 + \frac{1}{4}p_3 \\ \frac{1}{2}p_1 + \frac{1}{4}p_2 + \frac{1}{2}p_3 \\ \frac{1}{4}p_1 + \frac{1}{4}p_2 + \frac{1}{4}p_3 \end{bmatrix}.$$

EXAMPLE 4

Use matrices to find the probabilities of the stock market going up, going down, or staying the same tomorrow if

$$\mathbf{p} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

is the vector of probabilities of the market going up, going down, or staying the same today, as in Example 1 of Section 10.2.

Solution The vector of probabilities for the stock market tomorrow is $\mathbf{p}^+ = \mathbf{Ap}$, so we get

$$\mathbf{p}^+ = \mathbf{Ap} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{3}{8} \\ \frac{3}{8} \\ \frac{2}{8} \end{bmatrix}.$$

Alternatively, using a calculator where the entries in the matrix **A** and the vector **p** are given as decimals instead of fractions, we get

$$\mathbf{p}^+ = \mathbf{Ap} = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.5 \\ 0.25 & 0.25 & 0.25 \end{bmatrix}\begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.375 \\ 0.375 \\ 0.25 \end{bmatrix}.$$

Recall that a Markov chain can be extended farther into the future. Just as tomorrow's probability vector $\mathbf{p}^+$ can be computed by the matrix expression $\mathbf{p}^+ = \mathbf{Ap}$, so too the probability vector $\mathbf{p}^{++}$ for the day after tomorrow is given by

$$\mathbf{p}^{++} = \mathbf{Ap}^+ = \mathbf{A}(\mathbf{Ap}) = (\mathbf{AA})\mathbf{p} = \mathbf{A}^2\mathbf{p},$$

where we have written $\mathbf{A}^2$ to represent **AA**, the *square* of the matrix **A.** Note that $\mathbf{A}^2 = \mathbf{AA}$ is possible provided that **A** is a square matrix. (Don't confuse "square of a matrix" and "square matrix", where the number of rows equals the number of columns.) Similarly, the vector of probabilities $\mathbf{p}^{+++}$ three days hence is given by

$$\mathbf{p}^{+++} = \mathbf{Ap}^{++} = \mathbf{A}(\mathbf{A}^2\mathbf{p}) = \mathbf{A}^3\mathbf{p}.$$

In writing these two matrix equations for $\mathbf{p}^{++}$ and $\mathbf{p}^{+++}$, we made use of the fact that matrix algebra is an associative operation, so $\mathbf{A}(\mathbf{Ap}) = (\mathbf{AA})\mathbf{p} = \mathbf{A}^2\mathbf{p}$, and $\mathbf{A}(\mathbf{A}^2\mathbf{p}) = (\mathbf{AA}^2)\mathbf{p} = \mathbf{A}^3\mathbf{p}$, where $\mathbf{A}^3$ can be thought of as **AAA.**

EXAMPLE 5

Compute $\mathbf{A}^2$ and $\mathbf{A}^3$ for the stock market Markov transition matrix **A.**

Solution We first do the calculations by hand:

$$\mathbf{A}^2 = \begin{bmatrix} \frac14 & \frac12 & \frac14 \\ \frac12 & \frac14 & \frac12 \\ \frac14 & \frac14 & \frac14 \end{bmatrix}\begin{bmatrix} \frac14 & \frac12 & \frac14 \\ \frac12 & \frac14 & \frac12 \\ \frac14 & \frac14 & \frac14 \end{bmatrix}$$

$$= \begin{bmatrix} \frac14\cdot\frac14 + \frac12\cdot\frac12 + \frac14\cdot\frac14 & \frac14\cdot\frac12 + \frac12\cdot\frac14 + \frac14\cdot\frac14 & \frac14\cdot\frac14 + \frac12\cdot\frac12 + \frac14\cdot\frac14 \\ \frac12\cdot\frac14 + \frac14\cdot\frac12 + \frac12\cdot\frac14 & \frac12\cdot\frac12 + \frac14\cdot\frac14 + \frac12\cdot\frac14 & \frac12\cdot\frac14 + \frac14\cdot\frac12 + \frac12\cdot\frac14 \\ \frac14\cdot\frac14 + \frac14\cdot\frac12 + \frac14\cdot\frac14 & \frac14\cdot\frac12 + \frac14\cdot\frac14 + \frac14\cdot\frac14 & \frac14\cdot\frac14 + \frac14\cdot\frac12 + \frac14\cdot\frac14 \end{bmatrix}$$

$$= \begin{bmatrix} \frac38 & \frac{5}{16} & \frac38 \\ \frac38 & \frac{7}{16} & \frac38 \\ \frac14 & \frac14 & \frac14 \end{bmatrix}.$$

Similarly, we compute $\mathbf{A}^3$ as

$$\mathbf{A}^3 = \mathbf{AA}^2 = \begin{bmatrix} \frac14 & \frac12 & \frac14 \\ \frac12 & \frac14 & \frac12 \\ \frac14 & \frac14 & \frac14 \end{bmatrix}\begin{bmatrix} \frac38 & \frac{5}{16} & \frac38 \\ \frac38 & \frac{7}{16} & \frac38 \\ \frac14 & \frac14 & \frac14 \end{bmatrix}$$

$$= \begin{bmatrix} \frac14\cdot\frac38 + \frac12\cdot\frac38 + \frac14\cdot\frac14 & \frac14\cdot\frac{5}{16} + \frac12\cdot\frac{7}{16} + \frac14\cdot\frac14 & \frac14\cdot\frac38 + \frac12\cdot\frac38 + \frac14\cdot\frac14 \\ \frac12\cdot\frac38 + \frac14\cdot\frac38 + \frac12\cdot\frac14 & \frac12\cdot\frac{5}{16} + \frac14\cdot\frac{7}{16} + \frac12\cdot\frac14 & \frac12\cdot\frac38 + \frac14\cdot\frac38 + \frac12\cdot\frac14 \\ \frac14\cdot\frac38 + \frac14\cdot\frac38 + \frac14\cdot\frac14 & \frac14\cdot\frac{5}{16} + \frac14\cdot\frac{7}{16} + \frac14\cdot\frac14 & \frac14\cdot\frac38 + \frac14\cdot\frac38 + \frac14\cdot\frac14 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{11}{32} & \frac{23}{64} & \frac{11}{32} \\ \frac{13}{32} & \frac{25}{64} & \frac{13}{32} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

Alternatively, we can find these matrices by using the matrix capabilities of a calculator or computer package, though likely with the entries in decimal form. In that case, we would define the 3×3 matrix $\mathbf{A}$ and then refer to it by name to form

$$\mathbf{A}^2 = \mathbf{A}\wedge 2 = \begin{bmatrix} 0.375 & 0.3125 & 0.375 \\ 0.375 & 0.4375 & 0.375 \\ 0.25 & 0.25 & 0.25 \end{bmatrix}$$

and

$$\mathbf{A}^3 = \mathbf{A}\wedge 3 = \begin{bmatrix} 0.34375 & 0.359375 & 0.34375 \\ 0.40625 & 0.390625 & 0.40625 \\ 0.25 & 0.25 & 0.25 \end{bmatrix}.$$

You can verify that these decimal entries are equivalent to the fractions we calculated by hand initially.

The entries in $\mathbf{A}^2$ are the transition probabilities for 2 days from now and the entries in $\mathbf{A}^3$ are the transition probabilities for 3 days from now. For instance, the value $\frac{3}{8} = 0.375$ in position $(2, 1)$ of $\mathbf{A}^2$ (the entry in the second row, first column) means that, if we are now in State 1 (Up—first column), the chance is $\frac{3}{8}$ that in 2 days we will be in State 2 (Down—second row). Similarly, the value of $13/32 = 0.40625$ in entry $(2, 1)$ of $\mathbf{A}^3$ indicates that, if the market is now up, the probability is $13/32$ that in 3 days it will go down.

The values we obtained in computing $\mathbf{A}^2$ and $\mathbf{A}^3$ look reasonable. In particular, the numbers in each column of $\mathbf{A}^2$ and $\mathbf{A}^3$ sum to 1 (the sum of all probabilities for the market on any given day must be 1). Note that all the entries in each of the three rows of $\mathbf{A}^3$ have roughly the same numerical value; the entries in the first row are slightly larger than $\frac{1}{3}$, those in the second row are all about 0.40, and all the entries in the last row are 0.25.

EXAMPLE 6

If the probabilities of the stock market going up, going down, or staying the same today are given by the vector

$$\mathbf{p} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix},$$

use matrices to find the probabilities $\mathbf{p}^{++}$ of the market going up, going down, or staying the same the day after tomorrow and the probabilities $\mathbf{p}^{+++}$ for the day after that.

Solution We know that $\mathbf{p}^{++} = \mathbf{A}^2\mathbf{p}$ and $\mathbf{p}^{+++} = \mathbf{A}^3\mathbf{p}$. Multiplying, we get

$$\mathbf{p}^{++} = \mathbf{A}^2\mathbf{p} = \begin{bmatrix} 0.34375 \\ 0.40625 \\ 0.25 \end{bmatrix} \quad \text{and} \quad \mathbf{p}^{+++} = \mathbf{A}^3\mathbf{p} = \begin{bmatrix} 0.3515625 \\ 0.3984375 \\ 0.25 \end{bmatrix}.$$

You can easily verify that these entries are the same values we obtained in Section 10.2.

If you refer back to the table that gave the market probabilities over many days near the end of Example 2 in Section 10.2, you will observe that the probabilities shown there are similar to the entries *in the columns* of $\mathbf{A}^3$. That is, the probabilities after 3 days are actually fairly close to the long-term probabilities, so this Markov chain converges to a limiting state rather quickly.

Example 6 illustrates how, with concise notation, matrix algebra allows us to express quite complex expressions.

We can visualize what happens in a Markov chain graphically in the case of a two-by-two transition matrix—say, $\mathbf{A} = \begin{bmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{bmatrix}$. (Again, notice that the sum of the entries in each column is 1.) Let's start with an initial vector of probabilities $\mathbf{p} = \mathbf{p}_0 = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$. Any vector, including this vector of probabilities, can be interpreted geometrically, as shown in Figure 10.33. When we pre-multiply the vector $\mathbf{p}_0$ by the transition matrix $\mathbf{A}$, we obtain another vector of probabilities $\mathbf{p}^+ = \mathbf{p}_1 = \begin{bmatrix} 0.525 \\ 0.475 \end{bmatrix}$, which we can interpret geometrically, as shown in Figure 10.33. Note that the resulting vector $\mathbf{p}_1$ points in a direction quite different from that of the initial vector $\mathbf{p}_0$. We can think of the matrix $\mathbf{A}$ as transforming the probability vector $\mathbf{p}_0$ into the vector $\mathbf{p}_1 = \mathbf{A}\mathbf{p}_0$.

When we multiply the vector $\mathbf{p}_1$ by the matrix $\mathbf{A}$ to form the next probability vector $\mathbf{p}_2 = \mathbf{A}\mathbf{p}_1 = \begin{bmatrix} 0.4575 \\ 0.5425 \end{bmatrix}$, we get a vector $\mathbf{p}_2 = \mathbf{A}^2\mathbf{p}_0$ that points in still another direction. What happens when we continue the Markov process to get $\mathbf{p}_3 = \mathbf{A}^3\mathbf{p}_0$, $\mathbf{p}_4 = \mathbf{A}^4\mathbf{p}_0, \ldots$? (We have moved from using the notation $\mathbf{p}^+$, $\mathbf{p}^{++}$, and so on, to use subscript notation because the use of multiple $+$'s quickly becomes too unwieldy.) In the following table, we show the results of continuing this process numerically for the two components p_1 and p_2 for each successive probability vector $\mathbf{p}$. As in Example 6, the probabilities seem to converge to a pair of limiting values—approximately 0.42857, which is about 3/7, and approximately 0.57143, or 4/7.

n	0	1	2	3	4	5	6	7	...	
p_1	0.75	0.525	0.4575	0.43725	0.43118	0.42935	0.42881	0.42864	...	0.42857
p_2	0.25	0.475	0.5425	0.56275	0.56883	0.57065	0.57119	0.57136	...	0.57143

Figure 10.34 shows the corresponding geometric behavior. Note how the sequence of vectors also converges, getting closer and closer to a single limiting vector.

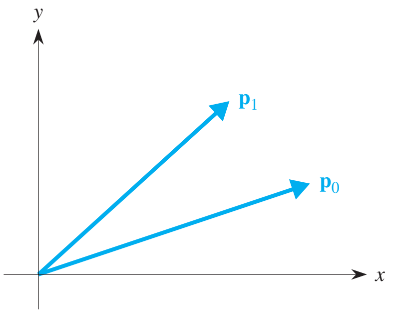


FIGURE 10.33

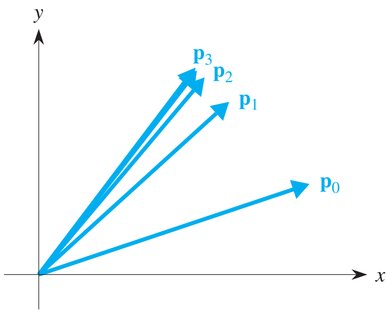


FIGURE 10.34

If we start with any other vector of probabilities, say, $\mathbf{p}_0 = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$, then the corresponding sequence of vectors will also converge to this same limiting vector $\begin{bmatrix} \frac{3}{7} \\ \frac{4}{7} \end{bmatrix}$. This special limiting vector is called an **eigenvector** of the matrix **A**. A similar type of convergence occurs with most other transition matrices, whether a two-by-two matrix or larger.

A Geometric View of Powers of a Matrix

We have shown that any matrix of the form $\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ acts as a rotation matrix to rotate a vector **v** through an angle θ. We now examine the effect of applying $\mathbf{R}^2$ to a vector.

EXAMPLE 7

Show that the matrix $\mathbf{R}^2$ is a rotation matrix that will rotate any vector through an angle 2θ.

Solution We have

$$\mathbf{R}^2 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta - \sin^2\theta & -\sin\theta\cos\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta + \sin\theta\cos\theta & -\sin^2\theta + \cos^2\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix}.$$

If we now apply the double angle identities for the sine and cosine from Chapter 8, we get $\mathbf{R}^2 = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$, so that $\mathbf{R}^2$ indeed is a rotation matrix with an associated angle 2θ.

Problems

1. Let

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

$$\mathbf{C} = \begin{bmatrix} 3 & 1 \\ -1 & -3 \end{bmatrix}.$$

Compute the matrix products.

a. **AB** **b.** **BA**
c. **CB** **d.** **BC**
e. **AC** **f.** $\mathbf{A}^2$
g. $\mathbf{B}^2$

2. For the matrices in Problem 1, compute the sums and linear combinations of matrices.

a. **A** + **B** **b.** **B** + **C**
c. **A** + **B** + **C** **d.** 2**A** + 3**C**

3. Let

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

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

Compute the products, if possible.

a. **AB** **b.** **CB** **c.** **BC**
d. **AD** **e.** **DA** **f.** **CD**
g. **DC** **h.** $\mathbf{C}^2$

4. For **A**, **B**, and **C** in Problem 1,
   - **a.** compute **AB** and then compute **(AB)C.**
   - **b.** compute **BC** and then compute **A(BC).** Does **(AB)C** equal **A(BC)?**
5. For **B**, **C**, and **D** in Problem 3,
   - **a.** compute **CB** and then compute **(CB)D.**
   - **b.** compute **BD** and then compute **C(BD).** Does **(CB)D** equal **C(BD)?**
6. Let

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

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

Compute the products, if possible. If not possible, explain why.

a. **AB** **b.** **BA** **c.** **AC**
d. **CA** **e.** **BC** **f.** **CB**
g. $\mathbf{A}^2$

7. Pre-multiply each matrix by the matrix **A** in Problem 6 and in each case describe how the columns of the product matrix compare to the columns of **A.**

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

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

8. For **A**, **B**, and **C** in Problem 6, compute the entry in the third row, third column in **(BC)A.** Explain why you do not have to multiply **BC** fully to determine this entry in **(BC)A.**
9. Perform the matrix multiplication for Amy's and Bill's fruit purchases at the three different stores if the matrix of Amy's and Bill's needs are

$$\begin{array}{c} \\ \text{Amy} \\ \text{Bill} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Peaches} & \text{Pears} & \text{Apples} & \text{Grapefruits} \end{array} \\ \begin{bmatrix} 7 & 2 & 4 & 5 \\ 2 & 5 & 1 & 8 \end{bmatrix} \end{array}.$$

10. Suppose that you have four robots: Supremo, Ultramatic, Maximus, and Gandalf, and three types of jobs: Job 1 (washing clothes), Job 2 (walking the dog), and Job 3 (doing a student's homework assignment). There are three families, the Joneses, the Smiths, and the Madonnas. Matrix **A** gives the times in hours it takes each robot to do each job. Matrix **B** tells how many jobs of each type are required by each family. Compute with **A** and **B** to find a matrix showing how long it will take each robot to do each family's set of jobs.

Matrix of Times

$$\mathbf{A} = \begin{array}{c} \begin{array}{ccc} \text{Job 1} & \text{Job 2} & \text{Job 3} \end{array} \\ \begin{bmatrix} 3 & 4 & 2 \\ 5 & 7 & 3 \\ 1 & 2 & 1 \\ 3 & 3 & 3 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Supremo} \\ \text{Ultramatic} \\ \text{Maxiumus} \\ \text{Gandalf} \end{array}$$

Matrix of Jobs

$$\mathbf{B} = \begin{array}{c} \begin{array}{ccc} \text{Jones} & \text{Smith} & \text{Madonna} \end{array} \\ \begin{bmatrix} 6 & 2 & 4 \\ 8 & 5 & 4 \\ 10 & 5 & 4 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Job 1} \\ \text{Job 2} \\ \text{Job 3} \end{array}$$

11. Suppose that you are given the following matrices involving the costs of fruit at different stores, the amounts of fruit that professors and engineers typically want, and the number of each type of person in two towns.

$$\mathbf{A} = \begin{array}{c} \begin{array}{cc} \text{Store 1} & \text{Store 2} \end{array} \\ \begin{bmatrix} 0.15 & 0.20 \\ 0.25 & 0.15 \\ 0.20 & 0.25 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Bananas} \\ \text{Peaches} \\ \text{Pears} \end{array}$$

$$\mathbf{B} = \begin{array}{c} \begin{array}{ccc} \text{Bananas} & \text{Peaches} & \text{Pears} \end{array} \\ \begin{bmatrix} 6 & 12 & 4 \\ 6 & 8 & 5 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Professors} \\ \text{Engineers} \end{array}$$

$$\mathbf{C} = \begin{array}{c} \begin{array}{cc} \text{Professors} & \text{Engineers} \end{array} \\ \begin{bmatrix} 2000 & 800 \\ 1500 & 1200 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Town 1} \\ \text{Town 2} \end{array}$$

- **a.** Compute a matrix product to find how much each type of person's fruit purchases cost at each store.
- **b.** Compute a matrix product to find how many of each fruit will be purchased in each town.
- **c.** Compute a matrix product to find how much was spent by each town at each store.

12. Consider the following population model for the numbers of goats (G) and sheep (S) from year to year.

$$G^+ = 2G + 2S$$
$$S^+ = G + 3S$$

Let $\mathbf{x}_0 = \begin{bmatrix} G_0 \\ S_0 \end{bmatrix}$ be the initial vector and $\mathbf{x}_n = \begin{bmatrix} G_n \\ S_n \end{bmatrix}$ be the vector of goats and sheep after n years. Let $\mathbf{A}$ be the matrix of coefficients in this system. (a) Write an expression for $\mathbf{x}_n$ in terms of $\mathbf{A}$ and $\mathbf{x}_0$. (b) Find $\mathbf{A}^2$ and $\mathbf{A}^3$, and use them and your formula to determine $\mathbf{x}_2$ and $\mathbf{x}_3$, if the starting population, in thousands, is $\mathbf{x}_0 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$. (Check your answer by applying the model equations for 2 and 3 years.) (c) What does the model predict for the two populations in 10 years?

13. Consider the following population model for the numbers of goats (G), sheep (S), and bears (B) from year to year.

$$G^+ = G + S + B$$
$$S^+ = G + 2S - B$$
$$B^+ = 2G - S + B$$

Let

$$\mathbf{x}_0 = \begin{bmatrix} G_0 \\ S_0 \\ B_0 \end{bmatrix}$$

be the initial vector and let $\mathbf{x}_n$ denote the vector of goats, sheep, and bears after n years. Let $\mathbf{A}$ be the matrix of coefficients in this system. (a) Write an expression for $\mathbf{x}_n$ in terms of $\mathbf{A}$ and $\mathbf{x}_0$. (b) Find $\mathbf{A}^2$ and $\mathbf{A}^3$. (c) Determine $\mathbf{x}_2$ and $\mathbf{x}_3$, if the starting population, in hundreds, is

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

(Check your answer by applying the model equations for 2 and 3 years.) (d) How long does it take for one of the populations to die out?

14. The copy machine at the student union breaks down as follows. If it is working today, it has a 70% chance of working tomorrow (and a 30% chance of breaking down). If the copy machine is broken today, it has a 50% chance of working tomorrow (and a 50% chance of being broken again). Write the Markov chain transition matrix $\mathbf{A}$ for this scenario.

   a. Compute $\mathbf{A}^2$. What probability does the entry in position $(1, 2)$ in $\mathbf{A}^2$ represent?

   b. Compute $\mathbf{A}^3$. If the copy machine is working today, what is the probability that it will be working in 3 days?

   c. If it is working today, what is the probability that it will be working a week from today?

15. Consider a weather Markov chain with 2 states, sunny and cloudy. If today is sunny, there is a $\frac{3}{4}$ probability that tomorrow will be sunny and a $\frac{1}{4}$ probability that tomorrow will be cloudy. If today is cloudy, there is a $\frac{1}{4}$ probability that tomorrow will be sunny and a $\frac{3}{4}$ probability that tomorrow will be cloudy. Write the transition matrix $\mathbf{A}$ for this Markov chain.

   a. Compute $\mathbf{A}^2$. What probability does the entry in position $(1, 2)$ in $\mathbf{A}^2$ represent?

   b. Compute $\mathbf{A}^3$. If today is cloudy, what is the probability that it will be sunny in three days?

   c. If today is cloudy, what is the probability that it will be sunny a week from today?

16. Consider the transition matrix $\mathbf{A} = \begin{bmatrix} 0.2 & 0.4 \\ 0.8 & 0.6 \end{bmatrix}$ and an initial vector $\mathbf{p} = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$.

   a. Calculate $\mathbf{p}^+$ and $\mathbf{p}^{++}$ by hand.

   b. Use the matrix features of your calculator to calculate the next four iterates $\mathbf{p}^{+++}$, $\mathbf{p}^{++++}$, and so on, corresponding to $\mathbf{A}^3$, $\mathbf{A}^4$, $\mathbf{A}^5$, and $\mathbf{A}^6$. Create a table listing the entries in the vectors that result. Do they appear to be converging?

   c. Plot the vectors you found in parts (a) and (b). Do they appear to be converging?

   d. Repeat parts (a)–(c) if the initial vector is $\mathbf{p} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$.

   e. Repeat parts (a)–(c) if the initial vector is $\mathbf{p} = \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix}$.

   f. What do you observe about the three sequences of vectors from parts (b), (d), and (e)?

17. In Problem 22 of Section 10.3, we described how to write a polynomial as a scalar product of a vector of coefficients and a vector of power functions. Suppose that you now have two cubic polynomials $P(x) = x^3 + 4x^2 - 7x + 2$ and $Q(x) = 4x^3 - 5x^2 + 8x + 17$.

   a. Express the sum of the two polynomials in terms of vectors.

b. Express the difference of the two polynomials in terms of vectors.

18. Explain why you can't have a power of a non-square matrix **A**.

19. Prove that, if **R** is the 2×2 rotation matrix associated with an angle θ, then $\mathbf{R}^3$ is a rotation matrix associated with an angle 3θ. (*Hint*: Write $\mathbf{R}^3 = \mathbf{R}\mathbf{R}^2$ and use appropriate trigonometric identities to simplify the result of the multiplication.)

20. The matrix $\mathbf{R} = \begin{bmatrix} 0.9272 & -0.3746 \\ 0.3746 & 0.9272 \end{bmatrix}$ is a rotation matrix. What is the associated angle of rotation?

21. Consider the matrix $\mathbf{R} = \begin{bmatrix} \cos 1° & -\sin 1° \\ \sin 1° & \cos 1° \end{bmatrix}$.

Describe the effect of successively applying $\mathbf{R}, \mathbf{R}^2, \mathbf{R}^3, \mathbf{R}^4, \ldots, \mathbf{R}^{90}$ to any nonzero vector **v**.

10.5 Gaussian Elimination

In this section we develop a procedure known as **Gaussian elimination** for solving any system of linear equations. The elimination method was devised by Carl Friedrich Gauss in about 1820 to solve systems of linear equations in astronomical and land-surveying computations. For our purposes here, when we speak of a system of linear equations, we assume that *the number of equations equals the number of variables.*

Suppose that we start with the system of linear equations

$$\begin{aligned} 3x - 5y + 4z &= 10 \\ 3x - 2y + 5z &= 11 \\ 6x + 2y - 2z &= 10. \end{aligned}$$

Gaussian elimination involves two stages. The first stage transforms the given system of equations, or equivalently the associated matrix of coefficients, into *upper triangular form,* with only 0's below the *main diagonal* of the matrix:

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

(We show how to do this shortly.) Once we have transformed the original system of equations to upper triangular form, solving it is quite simple. The second stage of the process uses "back substitution" to obtain values for the variables. That is, knowing from the third equation that $z = 1$, we can solve for y in the second equation:

$$3y + (1) = 1 \quad \text{so that} \quad y = 0.$$

Now, knowing y and z, we can solve for x from the first equation:

$$3x - 5(0) + 4(1) = 10 \quad \text{so that} \quad x = 2.$$

We can simplify this procedure by using vectors and matrices. Instead of using x, y, and z as the variables, we use x_1, x_2, and x_3. We then write the upper triangular form for the preceding system of equations in matrix form as $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 3 & -5 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 10 \\ 1 \\ 1 \end{bmatrix}.$$

To apply Gaussian elimination to any system of linear equations, we transform the original coefficient matrix into upper triangular form by applying the following three *elementary row operations* repeatedly.

Elementary Row Operations

- Multiply or divide any row of a matrix (or an equation) by a nonzero number.
- Add a multiple of one row (or equation) to another row (or equation).
- Interchange two rows (or two equations).

Realize that performing any of these operations on the rows of a matrix is equivalent to performing the same operation on the original equations. Becasue the operations do not materially change the equations, the equivalent operations do not materially change the matrix, so we will get the same solution.

Because the numbers, not the variables, are what matter in the equations, we work with the coefficient matrix **A** extended by a column for the right-hand side vector **b.** This matrix $[\mathbf{A}\,|\,\mathbf{b}]$ of coefficients along with the right-hand side vector is called the *augmented coefficient matrix.* For instance, if the system of equations is

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

so that

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 7 \\ 3 \end{bmatrix},$$

then the augmented matrix is

$$[\mathbf{A}\,|\,\mathbf{b}] = \left[\begin{array}{cc|c} 2 & 3 & 7 \\ 4 & -5 & 3 \end{array}\right].$$

In Example 1 we show the steps involved in applying Gaussian elimination to both the system of linear equations and the associated augmented matrix.

EXAMPLE 1

Use Gaussian elimination to solve the following system of two equations in two unknowns.

$$\begin{aligned} 2x + y &= 7 &\quad \textbf{(1)} \\ x - 2y &= -4 &\quad \textbf{(2)} \end{aligned}$$

Solution This system is equivalent to the augmented matrix

$$\left[\begin{array}{cc|c} 2 & 1 & 7 \\ 1 & -2 & -4 \end{array}\right].$$

To eliminate the x-term from Equation (2), we add $-\frac{1}{2}$ times Equation (1) to Equation (2) (the second elementary row operation). The result is a new second Equation (2′):

$$\begin{array}{lr} (2) & x - 2y = -4 \\ -\frac{1}{2}(1) & -x - 0.5y = -3.5 \\ \hline (2') = (2) - \frac{1}{2}(1) & 0 - 2.5y = -7.5 \end{array}$$

Equivalently, we perform the identical operation on the corresponding rows of the augmented matrix: Add $-\frac{1}{2}$ times the first row to the second row. Our new system of equations and the new augmented matrix become

$$2x + y = 7 \qquad \textbf{(1)}$$
$$-2.5y = -7.5, \qquad \textbf{(2')}$$

and

$$\left[\begin{array}{cc|c} 2 & 1 & 7 \\ 0 & -2.5 & -7.5 \end{array}\right].$$

Note that any solution to Equations (1) and (2) is also a solution to Equations (1) and (2′) because we can reverse the step that created Equation (2′). That is, (2′) $=$ (2) $-\frac{1}{2}$(1) implies that (2) $=$ (2′) $+\frac{1}{2}$(1). Thus Equation (2) is formed from Equation (2′) and a multiple of Equation (1), and so any solution to Equations (1) and (2′) is also a solution to Equations (1) and (2). But Equation (2′) is trivial to solve, and gives $y = 3$. Substituting $y = 3$ into Equation (1), we get

$$2x + 3 = 7, \quad \text{so} \quad x = \frac{(7-3)}{2} = 2.$$

Verify that $x = 2$ and $y = 3$ satisfy the *two* original Equations (1) and (2).

When solving a system of linear equations, you should always check your result by substituting the values for the variables into the original equations, not the transformed equations, in case you made a mathematical error along the way.

Note also that the work we did on the system of equations exactly parallels what happens with the augmented matrix. Eventually, we will dispense with the equations altogether and work exclusively with the augmented matrix because it eliminates writing and keeping track of the variables at every step.

A Geometric Interpretation

If there are only two equations in two unknowns, you can also solve the system graphically by plotting the two lines. The first equation represents all points on one line and the second equation represents all points on the second line. Thus the solution to the system of equations corresponds to the point of intersection and you can approximate this point with your graphing calculator.

There is a comparable geometric interpretation for a system of three linear equations in three unknowns. Just as an equation of the form $ax + by = d$ represents a line in the two-dimensional coordinate plane, an equation of the form $ax + by + cz = d$ represents a plane in three-dimensional space. When you have three equations in three unknowns, you actually have three different planes in space, as suggested in Figure 10.35. Visualize, for instance, the ceiling, the wall in front of you, and the wall to your left. These three planes intersect at a single point in the upper corner of the room to your left. This point of intersection also is the solution of the system of three equations. Unfortunately, graphing calculators cannot yet use this geometric interpretation to solve such a system of equations.

Also, there can be some complications: You know that two lines can be parallel, so the resulting system of two equations in two unknowns will not have a solution. Similarly, three planes don't necessarily have a common point of intersection. Visualize

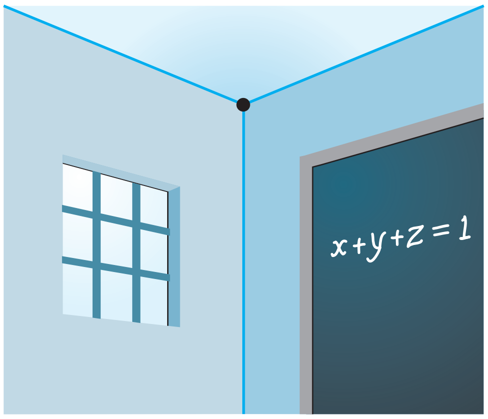


FIGURE 10.35

the ceiling, the floor, and the wall in front of you—they don't meet at a single point. Or, picture a long triangular prism made up of three flat sides that likewise have no single point of intersection. Alternatively, three planes can all pass through a common line of intersection, as with a revolving door. In that case the system of linear equations has infinitely many solutions. We consider such cases later in Example 3.

The Clothes Production Model and Gaussian Elimination

Recall the clothes production model introduced in Section 10.2 with three clothing factories whose raw material production levels had to be chosen to meet the demands for vests, pants, and coats.

$$\begin{aligned} \text{Vests:}\quad 20x_1 + 4x_2 + 4x_3 &= 500 \qquad \textbf{(3)}\\ \text{Pants:}\quad 10x_1 + 14x_2 + 5x_3 &= 850 \qquad \textbf{(4)}\\ \text{Coats:}\quad 5x_1 + 5x_2 + 12x_3 &= 1000 \qquad \textbf{(5)} \end{aligned}$$

EXAMPLE 2

Set up the corresponding augmented matrix and use Gaussian elimination to solve the system of equations.

Solution The augmented matrix is

$$\left[\begin{array}{ccc|c} 20 & 4 & 4 & 500 \\ 10 & 14 & 5 & 850 \\ 5 & 5 & 12 & 1000 \end{array}\right].$$

To solve this system, we first use multiples of Equation (3) to eliminate x_1 from Equations (4) and (5). Because $10x_1$ is half of $20x_1$ in Equation (3), we add $-\frac{1}{2}$ times Equation (3) to Equation (4) to cancel the terms $-\frac{1}{2}(20x_1)$ and $10x_1$ and so get a new second Equation $(4')$.

$$\begin{array}{lr} (4) & 10x_1 + 14x_2 + 5x_3 = 850 \\ -\frac{1}{2}(3) & -10x_1 - 2x_2 - 2x_3 = -250 \\ \hline (4') = (4) + (-\frac{1}{2})(3) & 0 + 12x_2 + 3x_3 = 600 \end{array}$$

Similarly, we add $-\frac{1}{4}$ times Equation (3) to Equation (5) to eliminate the x_1-term from Equation (5) and get a new Equation (5′).

$$\begin{array}{lrcrcrcr}
(5) & 5x_1 & + & 5x_2 & + & 12x_3 & = & 1000 \\
-\frac{1}{4}(3) & -5x_1 & - & 1x_2 & - & 1x_3 & = & -125 \\
\hline
(5') = (5) + (-\frac{1}{4})(3) & & & 4x_2 & + & 11x_3 & = & 875
\end{array}$$

Our new system of equations is now

$$\begin{aligned}
20x_1 + 4x_2 + 4x_3 &= 500 && \textbf{(3)} \\
12x_2 + 3x_3 &= 600 && \textbf{(4')} \\
4x_2 + 11x_3 &= 875, && \textbf{(5')}
\end{aligned}$$

and the corresponding augmented matrix is

$$\left[\begin{array}{rrr|r} 20 & 4 & 4 & 500 \\ 0 & 12 & 3 & 600 \\ 0 & 4 & 11 & 875 \end{array}\right].$$

Next we use Equation (4′) to eliminate the x_2-term from Equation (5′) by adding $-\frac{1}{3}$ times Equation (4′) to Equation (5′). We obtain a new Equation (5″).

$$\begin{array}{lrcrcr}
(5') & 4x_2 & + & 11x_3 & = & 875 \\
-\frac{1}{3}(4') & -4x_2 & - & 1x_3 & = & -200 \\
\hline
(5'') = (5') + (-\frac{1}{3})(4') & & & 10x_3 & = & 675
\end{array}$$

Our final system of equations in upper triangular form then is

$$\begin{aligned}
20x_1 + 4x_2 + 4x_3 &= 500 && \textbf{(3)} \\
12x_2 + 3x_3 &= 600 && \textbf{(4')} \\
10x_3 &= 675 && \textbf{(5'')}
\end{aligned}$$

with the associated augmented matrix

$$\left[\begin{array}{rrr|r} 20 & 4 & 4 & 500 \\ 0 & 12 & 3 & 600 \\ 0 & 0 & 10 & 675 \end{array}\right].$$

Any solution to the original system is also a solution to the new system. Furthermore, reversing the steps used in going from the original system to the final system (so that the original system is formed from linear combinations of the equations in the final system), we see that any solution to the new system is also a solution to the original system. The final system is in upper triangular form, so we can solve it using back substitution. From Equation (5″), we have

$$x_3 = \frac{675}{10} = 67.5.$$

Substituting this value into Equation (4′) yields

$$12x_2 + 3(67.5) = 600, \quad \text{so} \quad 12x_2 = 600 - 202.5, \quad \text{or} \quad x_2 = 33.125.$$

Next substituting $x_3 = 67.5$ and $x_2 = 33.125$ into Equation (3) gives

$$20x_1 + 4(33.125) + 4(67.5) = 500$$

so that

$$x_1 = \frac{500 - 132.5 - 270}{20} = 4.875.$$

Therefore the vector of the number of rolls of cloth needed by each of the three clothing factories is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4.875 \\ 3.125 \\ 67.5 \end{bmatrix} = \begin{bmatrix} 4\frac{7}{8} \\ 33\frac{1}{8} \\ 67\frac{1}{2} \end{bmatrix}.$$

Clearly in practice, cloths come in full rolls, so a more realistic solution would involve rounding up to the next full roll of cloth.

Incidentally, you can apply the three elementary row operations in many different ways to solve a particular system of equations. However, although the steps you use might differ from those of others solving the same problem, you should all obtain the same solution in the end, if none of you have made any mathematical errors.

Most graphing calculators have a built-in routine for applying Gaussian elimination to any set of linear equations. Typically, you would enter the coefficients, press the `SOLVE` key, and the calculator will respond with the solution or will indicate that either there is not a unique solution or no solution exists. (The specific key operations differ from one machine to another, so check your manual for details.) From one point of view, this approach is extremely simple because you get the answer instantly. However, it does have the drawback of not letting you see *how* the method works or understand what went wrong when the method fails, as we discuss below.

Systems of Linear Equations with Multiple Solutions

In our discussion regarding the geometric interpretation of systems of three equations in three unknowns, we mentioned that it is possible that three planes can have many points of intersection, as in a revolving door. We illustrate this type of situation in Example 3 where we change the number on the right-hand side of the third equation in Example 2 from 1000 to 325 and the coefficient of x_3 from 12 to 2.

EXAMPLE 3

Apply Gaussian elimination to the system of linear equations

$$20x_1 + 4x_2 + 4x_3 = 500 \quad \textbf{(6)}$$
$$10x_1 + 14x_2 + 5x_3 = 850 \quad \textbf{(7)}$$
$$5x_1 + 5x_2 + 2x_3 = 325. \quad \textbf{(8)}$$

Solution The corresponding augmented matrix is

$$\left[\begin{array}{ccc|c} 20 & 4 & 4 & 500 \\ 10 & 14 & 5 & 850 \\ 5 & 5 & 2 & 325 \end{array}\right].$$

After eliminating x_1 from Equations (7) and (8), we have

$$\begin{array}{lr} (6) & 20x_1 + 4x_2 + 4x_3 = 500 \\ (7') = (7) - \tfrac{1}{2}(6) & 12x_2 + 3x_3 = 600 \\ (8') = (8) - \tfrac{1}{4}(6) & 4x_2 + 1x_3 = 200, \end{array}$$

along with the associated augmented matrix

$$\left[\begin{array}{ccc|c} 20 & 4 & 4 & 500 \\ 0 & 12 & 3 & 600 \\ 0 & 4 & 1 & 200 \end{array}\right].$$

Next we add $-\frac{1}{3}$ times Equation (7′) to Equation (8′) to eliminate the x_2-term:

$$\begin{array}{lr} (6) & 20x_1 + 4x_2 + 4x_3 = 500 \\ (7') & 12x_2 + 3x_3 = 600 \\ (8'') = (8') - \tfrac{1}{3}(7') & 0x_3 = 0 \end{array}$$

The corresponding augmented matrix is

$$\left[\begin{array}{ccc|c} 20 & 4 & 4 & 500 \\ 0 & 12 & 3 & 600 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

But this process eliminates the x_3 term, and simultaneously the constant term on the right becomes 0. Also, note that Equation (8′) is equivalent to $\frac{1}{3}$ times Equation (7′). As a consequence, we actually have only two equations in three unknowns. This system has infinitely many solutions because we can choose *any* value for x_3 and then use back substitution to determine x_2 and x_1, based on our choice for x_3. For instance, if we choose $x_3 = 0$, Equation (7′) gives $12x_2 + 3(0) = 600$, so $x_2 = 50$, and then Equation (6) gives $20x_1 + 4(50) + 4(0) = 500$, so $x_1 = 15$. Alternatively, if we choose $x_3 = 40$, eventually Equation (7′) gives $x_2 = 40$, and so Equation (6) gives $x_1 = 9$. That is, we have obtained two very different solutions to the same system of equations. In fact, had we selected any other positive value for x_3, we would have obtained still another solution to the system (but not necessarily to the original problem). Thus, geometrically, the corresponding planes in three dimensions all pass through a common line of intersection, and every point on this line is a solution to the system.

Incidentally, if you attempted to solve the system of equations in Example 3 on a graphing calculator, say, you would get an error message of the form `SINGULAR MATRIX`. We discuss what this message means later in this section.

Systems of Linear Equations with No Solutions

We pointed out in our discussion of the geometry of systems of equations that three planes may have no common point of intersection, as with the floor, ceiling and one wall in a room. We illustrate this situation in Example 4 by making another minor change in the value on the right-hand side of the third equation in the original system in Examples 2 and 3.

EXAMPLE 4

Apply Gaussian elimination to the system of linear equations

$$\begin{aligned} 20x_1 + 4x_2 + 4x_3 &= 500 && \textbf{(9)} \\ 10x_1 + 14x_2 + 5x_3 &= 850 && \textbf{(10)} \\ 5x_1 + 5x_2 + 2x_3 &= 1000. && \textbf{(11)} \end{aligned}$$

Solution The corresponding augmented matrix is

$$\left[\begin{array}{ccc|c} 20 & 4 & 4 & 500 \\ 10 & 14 & 5 & 850 \\ 5 & 5 & 2 & 1000 \end{array}\right].$$

Eliminating x_1 as before, we get

$$\begin{aligned} 20x_1 + 4x_2 + 4x_3 &= 500 && \textbf{(9)} \\ 12x_2 + 3x_3 &= 600 && \textbf{(10')} \\ 4x_2 + 1x_3 &= 875. && \textbf{(11')} \end{aligned}$$

The associated augmented matrix is

$$\left[\begin{array}{ccc|c} 20 & 4 & 4 & 500 \\ 0 & 12 & 3 & 600 \\ 0 & 4 & 1 & 875 \end{array}\right].$$

We now use Equation $(10')$ to eliminate the x_2-term in Equation $(11')$ to get

$$\begin{aligned} &(9) && 20x_1 + 4x_2 + 4x_3 = 500 \\ &(10') && 12x_2 + 3x_3 = 600 \\ &(11'') = (11') + (-\tfrac{1}{3})(10') && 0x_3 = 675, \end{aligned}$$

along with the augmented matrix

$$\left[\begin{array}{ccc|c} 20 & 4 & 4 & 500 \\ 0 & 12 & 3 & 600 \\ 0 & 0 & 0 & 675 \end{array}\right].$$

The two Equations $(10')$ and $(11')$ are called *inconsistent equations* because they lead to the impossible Equation $(11'')$: $0 = 675$. Hence the original system has no solution. Geometrically, this outcome indicates that the corresponding planes in three dimensions do not have a common point of intersection.

If you attempted to solve this system of equations on a graphing calculator, say, you would again get an error message about a `SINGULAR MATRIX`.

In the real world, the inconsistency that occurred in Example 4 would be resolved by increasing one of the right-hand side demands (thus producing an excess amount of one type of clothing).

We now summarize the steps of Gaussian elimination.

Gaussian Elimination

1. Add multiples of the ith equation, for $i = 1, 2, \ldots, n - 1$, to the remaining equations to eliminate the ith variable from the other equations.
2. Solve the resulting upper triangular system of equations, using back substitution.

If, in the process, the coefficient of x_i in the ith equation $(i = 1, 2, \ldots, n - 1)$ is zero and the coefficient of x_i in one of the following equations is nonzero, simply interchange the two equations.

Using the Inverse Matrix

We can use a related approach with matrices to solve systems of equations. Let's begin with the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 4 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ -2 & \frac{1}{2} \end{bmatrix}.$$

Their product is

$$\mathbf{AB} = \begin{bmatrix} 1 & 0 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is called the 2×2 **identity matrix** and is denoted by $\mathbf{I}_2$. For any 2×2 matrix $\mathbf{A}$,

$$\mathbf{AI}_2 = \mathbf{I}_2\mathbf{A} = \mathbf{A},$$

so $\mathbf{I}_2$ plays the same role for 2×2 matricies as the number 1 plays in arithmetic: $a \cdot 1 = 1 \cdot a = a$. Similarly, the identity matrix for 3×3 matrices is

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and so on. The preceding matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be **inverses** of each other. We write $\mathbf{B} = \mathbf{A}^{-1}$, and the product of the matrix and its inverse is the identity matrix

$$\mathbf{AA}^{-1} = \mathbf{I}_2 = \mathbf{A}^{-1}\mathbf{A}.$$

The inverse of a matrix can be extremely useful in solving systems of linear equations.

Not every matrix has an inverse. A matrix that does not have an inverse is called a **singular matrix.** (Note that this is the same term that is used in the error message on graphing calculators when a system of equations has no solution or multiple solutions.) A square matrix $\mathbf{A}$ has an inverse if, when we use Gaussian elimination to reduce it to upper triangular form, all the diagonal elements are nonzero. However, if any of the diagonal elements are zero, no inverse exists and the matrix is singular.

Suppose now that we have the matrix-vector equation $\mathbf{Ax} = \mathbf{b}$ and that the $n \times n$ matrix $\mathbf{A}$ is not singular, so its inverse $\mathbf{A}^{-1}$ exists. Matrix $\mathbf{A}$ can be any size, so we denote the corresponding identity matrix by $\mathbf{I}$ without designating its size. Then we can left-multiply the equation by this inverse matrix and obtain

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{Ax} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{Ix} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{x} &= \mathbf{A}^{-1}\mathbf{b}, \end{aligned}$$

which automatically gives the desired solution vector $\mathbf{x}$.

Note that we haven't discussed *how* to calculate the inverse of a matrix $\mathbf{A}$. We can do so by using an extension of the Gaussian elimination process, but that is somewhat outside the scope of what we want to focus on here. Instead, note that your calculator will do this for you. Enter a square matrix $\mathbf{A}$, call it up and press the x^{-1} key, and your calculator will give you the inverse matrix $\mathbf{A}^{-1}$ if it exists. If the inverse does not exist, you will get an error message. Multiply this inverse matrix and the vector of constants $\mathbf{b}$ to get the desired solution vector $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Some of you may have seen a method called *Cramer's rule,* which uses *determinants* for solving a system of linear equations. Although fairly effective for solving systems of two or three equations, Cramer's rule is very inefficient for larger systems. Suppose, for instance, that you use a relatively slow computer that can perform only 1 million operations per second. To solve a system of 20 equations in 20 unknowns with Cramer's rule would take it about 77,000 years! The same computer could solve that system in about 0.003 second using Gaussian elimination or matrix methods.

Applications of Gaussian Elimination

We now consider a series of further applications involving systems of equations. We begin by using Gaussian elimination to investigate the long-term behavior of Markov chains.

Steady State of the Markov Chain for the Stock Market Consider again the stock market Markov chain introduced in Section 10.2 with the transition matrix

$$\begin{array}{ll} & \quad\textbf{Market Today} \\ & \begin{array}{ccc} \text{Up} & \text{Down} & \text{Same} \end{array} \\ \begin{array}{r} \textbf{Market} \\ \textbf{Tomorrow} \end{array} \begin{array}{l} \text{Up} \\ \text{Down} \\ \text{Same} \end{array} & \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} = \mathbf{A}. \end{array}$$

We noted that over many time periods, the successive probability vectors converged to 0.35 for the market going up, 0.40 for the market going down, and 0.25 for the market staying the same. We confirm the fact that

$$\mathbf{p} = \begin{bmatrix} 0.35 \\ 0.40 \\ 0.25 \end{bmatrix}$$

is the **steady state** or the **equilibrium state** of the Markov chain by showing that if the market ever reaches this state on one day, the vector $\mathbf{p}^+ = \mathbf{Ap}$ for the following day will be the same as $\mathbf{p}$. That is, at the steady state, there is no subsequent change in the values for the probabilities. Using a calculator to perform the matrix–vector product gives

$$\mathbf{p}^{+} = \mathbf{A}\mathbf{p} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 0.35 \\ 0.40 \\ 0.25 \end{bmatrix} = \begin{bmatrix} 0.35 \\ 0.40 \\ 0.25 \end{bmatrix} = \mathbf{p}.$$

In Example 5, we show how to find the equilibrium state, denoted by $\mathbf{p}^*$, for a Markov chain exactly.

EXAMPLE 5

Find the equilibrium state $\mathbf{p}^*$ for the preceding matrix $\mathbf{A}$.

Solution We want a vector $\mathbf{p}^*$ that satisfies the matrix-vector equation $\mathbf{p}^* = \mathbf{A}\mathbf{p}^*$. Let the components of $\mathbf{p}^*$ be denoted by $p_1{}^*$, $p_2{}^*$, and $p_3{}^*$. Note that $p_1{}^* + p_2{}^* + p_3{}^* = 1$ because the sum of all the probabilities associated with an event must sum to 1.

Because $\mathbf{p}^* = \mathbf{A}\mathbf{p}^*$, we have

$$\begin{bmatrix} p_1{}^* \\ p_2{}^* \\ p_3{}^* \end{bmatrix} = \begin{bmatrix} \frac{1}{4}p_1{}^* + \frac{1}{2}p_2{}^* + \frac{1}{4}p_3{}^* \\ \frac{1}{2}p_1{}^* + \frac{1}{4}p_2{}^* + \frac{1}{2}p_3{}^* \\ \frac{1}{4}p_1{}^* + \frac{1}{4}p_2{}^* + \frac{1}{4}p_3{}^* \end{bmatrix}.$$

Collecting the variables on the left, we get the following system of three equations in the three unknowns $p_1{}^*$, $p_2{}^*$, and $p_3{}^*$:

$$\begin{aligned} \tfrac{3}{4}p_1{}^* - \tfrac{1}{2}p_2{}^* - \tfrac{1}{4}p_3{}^* &= 0 \\ -\tfrac{1}{2}p_1{}^* + \tfrac{3}{4}p_2{}^* - \tfrac{1}{2}p_3{}^* &= 0 \\ -\tfrac{1}{4}p_1{}^* - \tfrac{1}{4}p_2{}^* + \tfrac{3}{4}p_3{}^* &= 0. \end{aligned}$$

Solving this system by Gaussian elimination applied to the augmented coefficient matrix, we obtain

$$\left[\begin{array}{ccc|c} \frac{3}{4} & -\frac{1}{2} & -\frac{1}{4} & 0 \\ -\frac{1}{2} & \frac{3}{4} & -\frac{1}{2} & 0 \\ -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & 0 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} \frac{3}{4} & -\frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{5}{12} & -\frac{2}{3} & 0 \\ 0 & -\frac{5}{12} & \frac{2}{3} & 0 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} \frac{3}{4} & -\frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{5}{12} & -\frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

To eliminate the fractions, we multiply the first row by 4 and the second row by 12:

$$\left[\begin{array}{ccc|c} 3 & -2 & -1 & 0 \\ 0 & 5 & -8 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

This augmented matrix is equivalent to the reduced system of equations

$$\begin{aligned} 3p_1{}^* - 2p_2{}^* - p_3{}^* &= 0 \\ 5p_2{}^* - 8p_3{}^* &= 0. \end{aligned}$$

Because there are only two equations in the three unknowns, there is no unique solution. Instead, we can solve for any two of the variables in terms of the third—say, $p_3{}^*$. The second equation then gives

$$p_2{}^* = \frac{8}{5}p_3{}^*$$

and, when we substitute it into the first equation and solve for $p_1{*}$, we have

$$p_1{*} = \frac{2p_2{*} + p_3{*}}{3} = \frac{2(8/5)p_3{*} + p_3{*}}{3} = \frac{7}{5}p_3{*}.$$

Where is our unique vector of stable probabilities? In this solution we obtained infinitely many solutions—one for each value of $p_3{*}$. We now use the fact that the sum of the three probabilities must equal 1, or $p_1{*} + p_2{*} + p_3{*} = 1$. We substitute the expressions for $p_1{*}$ and $p_2{*}$ in terms of $p_3{*}$ to obtain

$$p_1{*} + p_2{*} + p_3{*} = \frac{7}{5}p_3{*} + \frac{8}{5}p_3{*} + p_3{*} = \frac{20}{5}p_3{*} = 4p_3{*} = 1,$$

so that $p_3{*} = 0.25$. Therefore

$$p_2{*} = \frac{8}{5}(0.25) = 0.40 \quad \text{and} \quad p_1{*} = \frac{7}{5}(0.25) = 0.35.$$

These probabilities are identical to those in the equilibrium state vector of the market Markov chain given previously.

Systems of linear equations of the form $\mathbf{Ax} = \mathbf{0}$ with a zero vector, $\mathbf{0}$, on the right-hand side, such as the one in Example 5, occur frequently in matrix algebra. They are called *homogeneous systems* of linear equations. In comparison, systems of equations of the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{b} \neq \mathbf{0}$, are called *nonhomogeneous systems.*

When solving a homogeneous system, we usually are interested in a nonzero solution, as was the case here. Note, however, that $\mathbf{x} = \mathbf{0}$ is always a solution to any homogeneous system $\mathbf{Ax} = \mathbf{0}$. Thus, if we are to get a nonzero solution, we must have a case of multiple solutions.

Applications to Analytic Geometry Next, we apply Gaussian elimination to solve some systems of linear equations that arise in analytic geometry.

One of the most basic objects in geometry is a line, whose equation can be written as $y = ax + b$. Suppose that we have two points, P with coordinates $(3, 1)$ and Q with coordinates $(6, 7)$, and we want to find the equation of the line through these points. We could solve this problem by finding the slope of the line segment joining the two points and then using the point–slope formula. Here, however, we solve this problem using systems of equations to illustrate some important mathematical ideas that can be extended to answer considerably more complicated questions. The results either way are the same.

EXAMPLE 6

Use matrix methods to find the equation of the line through the points $(3, 1)$ and $(6, 7)$.

Solution First, let's be clear about what we need to find. The equation of the line $y = ax + b$ is determined by two constants—the slope a and the vertical intercept b. Although a and b are constants for a particular line, we can think of them as parameters that distinguish one line from another. We need to determine appropriate values for a and b so that the line $y = ax + b$ passes through the points $(3, 1)$ and $(6, 7)$. Accordingly, a and b must satisfy the following two equations.

$$\text{For point } (3, 1): \quad 1 = a(3) + b$$
$$\text{For point } (6, 7): \quad 7 = a(6) + b$$

We rewrite these two equations in standard form to get

$$3a + b = 1$$
$$6a + b = 7.$$

Subtracting twice the first equation from the second equation yields

$$-b = 5, \quad \text{so that} \quad b = -5.$$

Substituting back into the second equation gives

$$6a + (-5) = 7, \quad \text{so} \quad 6a = 12, \quad \text{and} \quad a = 2.$$

Therefore the equation of the desired line is

$$y = 2x - 5.$$

We now extend the ideas in Example 6 to find the equation of a parabola $y = ax^2 + bx + c$. We need to determine the values of the parameters a, b, and c. To solve for these three unknowns, we need three linear equations. As with the line, we get a linear equation associated with each point the parabola goes through. For three equations, we need three points. Suppose that the three points are $(2, -1)$, $(4, 3)$ and $(-1, 8)$. Note that the three points must be noncollinear (i.e., they cannot lie on a common line). Each point determines an equation when we substitute its coordinates, x and y, into the parabola's formula $y = ax^2 + bx + c$.

$$\text{For point } (2, -1): \quad -1 = a(2)^2 + b(2) + c$$
$$\text{For point } (4, 3): \quad 3 = a(4)^2 + b(4) + c$$
$$\text{For point } (-1, 8): \quad 8 = a(-1)^2 + b(-1) + c$$

Hence we have the system of linear equations

$$4a + 2b + c = -1$$
$$16a + 4b + c = 3$$
$$a - b + c = 8.$$

In Problem 17 we ask you to solve these three equations to determine a, b, and c.

Next, we apply this curve-fitting method to a more complicated situation—namely, finding the equation of a circle. The familiar form of the equation of a circle is

$$(x - a)^2 + (y - b)^2 = c^2,$$

where (a, b) is the center of the circle and c is the radius. It turns out that, just as three noncollinear points determine a parabola, any three noncollinear points also determine a circle. We now have three parameters a, b, and c to determine, which should lead to three equations in the three unknowns. But the equations arising from giving particular values to x and y in the above equation for the circle will not be linear (there will be a^2 and b^2 terms), so we cannot use Gaussian elimination to determine a, b, and c based on the equation $(x - a)^2 + (y - b)^2 = c^2$. However, if we expand this expression algebraically, we get

$$\begin{aligned}(x - a)^2 + (y - b)^2 &= (x^2 - 2ax + a^2) + (y^2 - 2by + b^2)\\ &= x^2 + y^2 - 2ax - 2by + a^2 + b^2 = c^2.\end{aligned}$$

Therefore we can transform the equation of the circle to the form

$$x^2 + y^2 + Cx + Dy + E = 0,$$

where $C = -2a$, $D = -2b$, and $E = a^2 + b^2 - c^2$ are three other parameters. Using this form, we can set up three linear equations in these three unknowns and solve for them with Gaussian elimination. Once we have their values, we can rewrite the equation of the circle in the more familiar form and read the coordinates of the center and the radius. We illustrate these ideas in Example 7.

EXAMPLE 7

Find the equation of the circle passing through the points $(4, -2)$, $(6, 2)$ and $(5, 5)$.

Solution Using the equation

$$x^2 + y^2 + Cx + Dy + E = 0,$$

we create three linear equations that C, D, and E must satisfy.

$$\begin{aligned}&\text{For the point } (4, -2): && 4^2 + (-2)^2 + C(4) + D(-2) + E = 0\\ &\text{For the point } (6, 2): && 6^2 + 2^2 + C(6) + D(2) + E = 0\\ &\text{For the point } (5, 5): && 5^2 + 5^2 + C(5) + D(5) + E = 0\end{aligned}$$

Moving the constant terms to the right-hand side in each equation, we have

$$4C - 2D + E = -20 \qquad \textbf{(12)}$$

$$6C + 2D + E = -40 \qquad \textbf{(13)}$$

$$5C + 5D + E = -50. \qquad \textbf{(14)}$$

We now apply Gaussian elimination to this (nonhomogeneous) system of three linear equations in three unknowns. First, to eliminate C from Equations (13) and (14), we subtract $\frac{3}{2}$ of Equation (12) from Equation (13) and subtract $\frac{5}{4}$ of Equation (12) from Equation (14).

$$\begin{aligned}&(12) && 4C - 2D + E = -20\\ &(13') = (13) - \tfrac{3}{2}(12) && 5D - \frac{1}{2}E = -10\\ &(14') = (14) - \tfrac{5}{4}(12) && \frac{15}{2}D - \frac{1}{4}E = -25\end{aligned}$$

Next we eliminate D from the last equation by subtracting $\frac{3}{2}$ of Equation $(13')$ from Equation $(14')$.

$$\begin{aligned}&(12) && 4C - 2D + E = -20\\ &(13') && 5D - \frac{1}{2}E = -10\\ &(14'') = (14') - \tfrac{3}{2}(13') && \frac{1}{2}E = -10\end{aligned}$$

Gord.3896.10.pgs 4/28/03 3:30 PM Page 714

From Equation (14″), we have $E = -20$. Substituting back into Equation (13′) yields

$$5D - \frac{1}{2}(-20) = -10, \quad \text{so} \quad 5D = -20 \quad \text{and} \quad D = -4.$$

Then substituting back into Equation (12) yields

$$4C - 2(-4) + (-20) = -20, \quad \text{so} \quad 4C = -8 \quad \text{and} \quad C = -2.$$

Therefore one form of an equation for our circle is

$$x^2 + y^2 - 2x - 4y - 20 = 0.$$

We rewrite this equation in the more familiar form by completing the square on both x and y (see Appendix A8):

$$\begin{aligned}[(x^2 - 2x)] + [(y^2 - 4y)] - 20 &= 0\\ [(x^2 - 2x + 1) - 1] + [(y^2 - 4y + 4) - 4] - 20 &= 0\\ [(x-1)^2 - 1] + [(y-2)^2 - 4] - 20 &= 0\\ (x-1)^2 + (y-2)^2 &= 1 + 4 + 20 = 25.\end{aligned}$$

Thus our circle is centered at the point $(1, 2)$ and has radius 5, as shown in Figure 10.36.

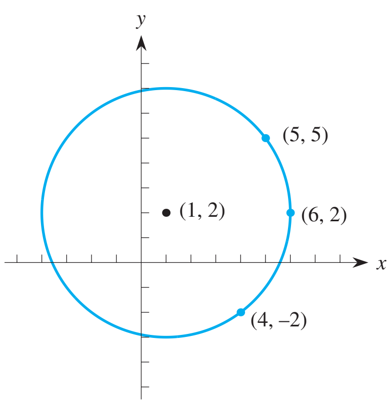


FIGURE 10.36

Think About This Substitute the coordinates of the three points in Example 7 into the usual equation for a circle, $(x - a)^2 + (y - b)^2 = c^2$, to see the types of equations involving the three parameters a, b and c that would result.

Problems

1. Solve each system of equations using Gaussian elimination.

   a. $x + y = 8$, $2x - 3y = 4$

   b. $x - 4y = 3$, $-x + 2y = 9$

   c. $2x - 3y = 4$, $-4x + y = 1$

   d. $3x - 2y = 3$, $-2x + 5y = 2$

   e. $3x - 2y = 3$, $-4x + 3y = -2$

2. Use Gaussian elimination to solve each variation on the clothes production model in Example 2. Some variations may have no solution, some may have multiple solutions (express such an infinite family of solutions in terms of x_3), and some may have an unrealistic solution involving negative values.

   a. $$\begin{aligned}20x_1 + 4x_2 + 4x_3 &= 500\\ 8x_1 + 3x_2 + 5x_3 &= 850\\ 4x_1 + 5x_2 + 11x_3 &= 2050\end{aligned}$$

   b. $$\begin{aligned}6x_1 + 5x_2 + 6x_3 &= 500\\ 10x_1 + 10x_2 &= 850\\ 2x_1 + 12x_3 &= 1000\end{aligned}$$

c. $6x_1 + 2x_2 + 2x_3 = 500$
$3x_1 + 6x_2 + 3x_3 = 300$
$3x_1 + 2x_2 + 6x_3 = 1000$

d. $8x_1 + 4x_2 + 3x_3 = 500$
$4x_1 + 8x_2 + 5x_3 = 500$
$12x_2 + 6x_3 = 500$

3. Solve each system of equations using Gaussian elimination.

a. $x_1 + 2x_2 + x_3 = 3$
$-x_1 + 2x_2 + 3x_3 = 1$
$2x_1 - x_2 + 2x_3 = 2$

b. $2x_1 - x_2 + x_3 = 2$
$-x_1 - 2x_2 + 2x_3 = -3$
$3x_1 + 2x_2 - x_3 = 0$

c. $-x_1 - 3x_2 + 2x_3 = -1$
$5x_1 + 4x_2 + 6x_3 = 12$
$2x_1 + x_2 + 3x_3 = 4$

d. $2x_1 + 4x_2 - 2x_3 = 4$
$-2x_1 + 2x_2 - 3x_3 = -4$
$x_1 - x_2 - 2x_3 = -1$

e. $x_1 + 2x_2 + 2x_3 = 3$
$x_1 + 2x_2 + 3x_3 = 11$
$2x_1 + x_2 + 2x_3 = 5$

f. $x_1 - x_2 - 2x_3 = 2$
$3x_1 + 2x_2 + 4x_3 = 2$
$3x_1 + x_2 - 2x_3 = -3$

4. In each set of three equations, show that the third equation equals some multiple r of the first equation added to the second equation: $(3) = r \cdot (1) + (2)$.

a. (1) $x + 2y = 4$
(2) $3x + y = 9$
(3) $x - 3y = 1$

b. (1) $x - y + z = 2$
(2) $x + y - z = 3$
(3) $-2x + 4y - 4z = -3$

c. (1) $2x + y - 2z = -5$
(2) $3x - y + z = 8$
(3) $6x + 0.5y - 2z = 0.5$

5. Determine whether each system of equations has a unique solution, multiple solutions, or is inconsistent (has no solution).

a. $x - 2y = -3$
$-3x + 6y = 9$

b. $x - 2y = -3$
$-3x + 6y = 4$

c. $2x - 5y = 3$
$6x - 15y = 9$

d. $x_1 + 2x_2 + 3x_3 = 10$
$2x_1 - x_2 + 4x_3 = 20$
$5x_2 + 2x_3 = 0$

e. $x_1 + 5x_2 + 4x_3 = 10$
$x_1 + 3x_2 + 6x_3 = 9$
$x_1 - x_2 + x_3 = 5$

f. $-x_1 + 2x_2 + x_3 = 0$
$2x_1 + x_2 - 3x_3 = 0$
$x_1 + x_2 + 2x_3 = 0$

6. For each augmented matrix, state whether the associated system of equations has a unique solution, multiple solutions, or no solution.

a. $\left[\begin{array}{ccc|c} 3 & 3 & 1 & 0 \\ 0 & 1 & 5 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]$ **b.** $\left[\begin{array}{ccc|c} 1 & 2 & 2 & 3 \\ 0 & 5 & 1 & 6 \\ 0 & 0 & 0 & 2 \end{array}\right]$

c. $\left[\begin{array}{ccc|c} 1 & 2 & 7 & 5 \\ 0 & 4 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$ **d.** $\left[\begin{array}{ccc|c} 4 & 2 & 3 & 3 \\ 0 & 4 & 2 & 6 \\ 0 & 2 & 1 & 3 \end{array}\right]$

e. $\left[\begin{array}{ccc|c} 2 & 1 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$

7. (Continuation of Problem 4 of Section 10.2) Consider the following clothes production model. There are three clothing factories (1, 2, and 3) and from each roll of cloth, the different factories produce the following numbers of vests, pants, and coats.

$$\begin{array}{l} \\ \text{Vests} \\ \text{Pants} \\ \text{Coats} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Factory 1} & \text{Factory 2} & \text{Factory 3} \end{array} \\ \begin{bmatrix} 6 & 4 & 2 \\ 4 & 8 & 4 \\ 3 & 2 & 8 \end{bmatrix} \end{array}$$

Suppose that the demand is for 400 vests, 800 pants, and 500 coats. Write a system of equations whose solution would determine production levels (rolls of cloth needed by each factory) to yield the desired numbers of vests, pants, and coats. Find the solution using Gaussian elimination.

8. (Continuation of Problem 5 of Section 10.2) From each shipment of crude petroleum, Refineries 1, 2, and 3 produce the following amounts (in thousands of gallons) of heating oil, diesel oil, and gasoline.

$$\begin{array}{c} \\ \text{Heating Oil} \\ \text{Diesel Oil} \\ \text{Gasoline} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Refinery 1} & \text{Refinery 2} & \text{Refinery 3} \end{array} \\ \begin{bmatrix} 8 & 5 & 3 \\ 2 & 5 & 5 \\ 3 & 7 & 6 \end{bmatrix} \end{array}$$

Suppose that the demand is for 6200 thousand gallons of heating oil, 4000 thousand gallons of diesel oil, and 4700 thousand gallons of gasoline. Write a system of equations whose solution would determine production levels to yield the desired amounts of heating oil, diesel oil, and gasoline. Find the solution using Gaussian elimination.

9. (Continuation of Problem 6 of Section 10.2) The staff dietitian at the California Institute of Trigonometry has to make up a meal with 600 calories, 20 grams of protein, and 200 mg of vitamin C. The three food types that the dietitian can choose from are gelatin, fish sticks, and mystery meat. They have the following nutritional content per ounce.

$$\begin{array}{c} \\ \text{Calories} \\ \text{Protein} \\ \text{Vitamin C} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Gelatin} & \text{Fish Sticks} & \text{Mystery Meat} \end{array} \\ \begin{bmatrix} 10 & 50 & 200 \\ 1 & 3 & 0.2 \\ 30 & 10 & 0 \end{bmatrix} \end{array}$$

Describe the dietitian's problem and create a mathematical model for it with a system of three linear equations. Find the solution using Gaussian elimination.

10. (Continuation of Problem 7 of Section 10.2) A company has a budget of \$280,000 for computing equipment. The types of equipment available are microcomputers at \$2000 each, terminals at \$500 each, and workstations at \$5000 each. There should be five times as many terminals as microcomputers and twice as many microcomputers as workstations. Set up a system of three linear equations for this situation. Find the solution using Gaussian elimination.

11. Find the equilibrium state for the Markov chains with the following transition matrices.

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

12. The copy machine at the student union breaks down with the following pattern. If it is working today, there is a 70% chance that it works tomorrow (and a 30% chance of breaking down). If it is broken today, there is a 50% chance that it works tomorrow. Construct a Markov chain for this problem and find the equilibrium state.

13. From past experience, the Pins bowling team knows that if they win this week's game, they have a $\frac{2}{3}$ chance of winning next week's game. If they lose this week's game, they have a $\frac{1}{2}$ chance of winning next week's game. Construct a Markov chain for this problem and find the stable distribution.

14. Find the equilibrium state for the Markov chains with the following transition matrices.

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

c. $\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$

15. The following model for learning a concept over a set of lessons identifies four states of learning: I = Ignorance, E = Exploratory thinking, S = superficial understanding, and M = Mastery. If you are now in State I, after one lesson you have probability $\frac{1}{2}$ of still being in I and probability $\frac{1}{2}$ of being in E. If you are now in State E, you have probability $\frac{1}{4}$ of being in I, $\frac{1}{2}$ in E, and $\frac{1}{4}$ in S. If you are now in State S, you have probability $\frac{1}{4}$ of being in E, $\frac{1}{2}$ in S, and $\frac{1}{4}$ in M. If you are in State M, you always stay in M (with probability 1). Construct a Markov chain for this problem and find the equilibrium state.

16. Find the equation of the line passing through each pairs of points, using Gaussian elimination.

a. $(1, 1), (2, 3)$ **b.** $(1, 0), (0, 2)$
c. $(2, 3), (-1, 1)$ **d.** $(2, 1), (3, 1)$
e. $(-1, 1), (2, -5)$ **f.** $(4, 1), (-1, -1)$

17. Solve the three equations for determining the parameters of the parabola $y = ax^2 + bx + c$ that passes through the three points $(2, -1)$, $(4, 3)$ and $(-1, 8)$.

18. A parabola passes through the points $(-1, 4)$, $(1, 2)$ and $(4, 7)$. Set up a system of linear equations in a, b, and c and then solve it, using Gaussian elimination, to find the equation of the parabola.

19. Find an equation of the circle passing through each triple of points, using Gaussian elimination.

a. $(1, 0), (0, 1), (1, 1)$
b. $(2, 1), (2, -3), (0, 1)$
c. $(3, 1), (2, 5), (-3, 6)$

d. $(-1, -1), (1, 3), (2, -1)$
e. $(0, 0), (4, 1), (-4, 1)$

20. Just as two distinct points determine a line uniquely, three noncollinear points determine a parabola. Moreover, four noncollinear points, all of which do not lie on a parabola, determine a cubic polynomial.

a. A parabola $y = ax^2 + bx + c$ passes through the points $(-1, 5)$, $(1, -2)$, and $(4, 6)$. Set up a system of linear equations in a, b, and c and then solve it, using Gaussian elimination, to find the equation of the parabola.

b. A cubic polynomial $y = ax^3 + bx^2 + cx + d$ passes through the four points $(-1, 4)$, $(1, -2)$, $(4, 7)$, and $(5, 2)$. Set up a system of linear equations in a, b, c, and d that could be used to find the equation of the cubic. Then solve the system using Gaussian elimination.

c. How does your answer to part (b) compare to the result you can obtain with your calculator using the routine for fitting a cubic function to a set of data?

21. Find the inverse of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 4 & 2 \end{bmatrix}$ algebraically. (*Hint*: Write the inverse as $\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and use the fact that $\mathbf{AA}^{-1} = \mathbf{I}_2$ to solve for a, b, c, and d.) How does your answer compare with what your calculator shows using its matrix features?

22. a. Consider the rotation matrix

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Based on geometric principles, predict what the inverse matrix $\mathbf{R}^{-1}$ has to be.

b. Prove algebraically that the matrix you constructed in part (a) is indeed the inverse matrix for $\mathbf{R}$.

23. The method of linear regression discussed in Chapter 3, in which the best fit line $y = ax + b$ is constructed for a set of (x, y) data, is based on solving a system of linear equations for the unknown parameters a and b. It can be shown that, if the set of n data points is $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, y_n)$, then a and b must satisfy the system of linear equations

$$\left(\sum_{i=1}^{n} x_i\right)a + nb = \left(\sum_{i=1}^{n} y_i\right)$$

$$\left(\sum_{i=1}^{n} x_i^2\right)a + \left(\sum_{i=1}^{n} x_i\right)b = \left(\sum_{i=1}^{n} x_i y_i\right).$$

Recall that the summations indicate adding up all the x-values, adding up all the y-values, adding up the squares of all the x-values, and adding up all the products of the x and y-values. Note that all the summation terms are known constants once the x's and y's have been given.

a. For the set of data $(1, 11)$, $(2, 25)$, $(3, 33)$, $(4, 45)$, $(5, 57)$, and $(6, 65)$, find the equation of the regression line using your graphing calculator or appropriate software package.

b. Calculate the sums needed in the two preceding equations by completing the entries in the table.

x	y	x^2	xy
1	11		
2	25		
3	33		
4	45		
5	57		
6	65		
$\sum x =$	$\sum y =$	$\sum x^2 =$	$\sum xy =$

c. Use the results of part (b) to write the system of linear equations in a and b that you can use to determine the values for the unknown coefficients a and b in the regression equation.

d. Solve the system of equations in part (c) using Gaussian elimination. How do your results compare to what you obtained directly in part (a)?

e. In Chapter 3 we suggested that you scale down large numbers, such as the full years 2000, 2001, 2002, . . . , in a set of data. Based on your calculations in part (b), explain why doing so is desirable. In particular, what might happen if there were many data values, if the x' s, say, consisted of full years and the y' s were also large numbers?

24. Repeat parts (a)–(d) of Problem 23 for $(0, 24)$, $(5, 21)$, $(10, 17)$, $(15, 14)$, $(20, 12)$, and $(25, 9)$.

Chapter Summary

In this chapter we introduced vectors and matrices and some of their properties and applications. In particular, we emphasized

- What a vector is geometrically, algebraically, and as an ordered pair or ordered triple of numbers.
- What a matrix is algebraically and as an array of numbers.
- How to use vectors and matrices to construct mathematical models, including Markov chains and growth models, of a wide variety of real-world problems.
- How to use vectors and matrices to rotate geometric figures.
- How to add and subtract vectors and matrices.
- How to compute the scalar product of two vectors.
- How to multiply matrices and vectors, and matrices and other matrices.
- How to use matrix multiplication to solve applied problems, including Markov chains and linear growth models.
- How to solve a system of n linear equations in n unknowns, using Gaussian elimination.
- How to apply Gaussian elimination to solve various real-world problems with a system of linear equations.
- How to find an equation of various types of curves, such as parabolas and circles, that pass through a given set of points.

Chapter Review Problems

1. Determine the magnitude of the displacement vectors from point A to point B for each pair of points.

   a. $A = (6, 6), B = (2, 3)$
   b. $A = (-1, 1), B = (4, -1)$
   c. $A = (1, 2, -3), B = (3, 3, -5)$

2. **a.** If a plane is flying due south at 200 mph and the wind is blowing *in* a direction that is 40° north of west at 50 mph, what are the actual direction and speed of the plane?
   b. Suppose instead that the wind is blowing *from* a direction that is 40° north of west at 50 mph. How do your answers to part (a) change?

3. A furniture manufacturer makes tables, chairs, and sofas. In one month, the company has available 1500 units of wood, 2300 units of labor, and 1800 units of upholstery. The manufacturer wants a production schedule for the month that uses all these resources. The different products require the following amounts of the resources.

$$\begin{array}{l} \\ \text{Wood} \\ \text{Labor} \\ \text{Upholstery} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Table} & \text{Chair} & \text{Sofa} \end{array} \\ \begin{bmatrix} 4 & 1 & 3 \\ 3 & 2 & 5 \\ 0 & 2 & 4 \end{bmatrix} \end{array}$$

   Write a system of equations whose solution would determine production levels to yield the desired numbers of tables, chairs, and sofas.

4. If the stock market goes up today, historical data show that for tomorrow it has a 60% chance of going up, a 20% chance of staying the same, and a 20% chance of going down. If the market is unchanged today, it has a 20% chance of being unchanged, a 40% chance of going up, and a 40% chance of going down tomorrow. If the market goes down today, it has a 20% chance of going up, a 20% chance of being unchanged, and a 60% chance of going down tomorrow.

   a. Construct a Markov chain for this problem. Give **A**, the matrix of transition probabilities, and draw the transition diagram.

b. If there is a 30% chance of the market going up today, a 10% chance of being unchanged, and a 60% chance of going down, what is the probability distribution for the market tomorrow?

5. Let $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$, and $\mathbf{c} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$.

Compute: **(a)** $\mathbf{a} \cdot \mathbf{b}$ **(b)** $\mathbf{b} \cdot \mathbf{c}$ **(c)** $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c})$ **(d)** $\mathbf{a} \cdot \mathbf{a}$

6. Let **a**, **b**, and **c** be as in Problem 5. Let

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

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

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

Which of the following matrix calculations are well defined (the sizes match)? If the computation makes sense, perform it. If necessary, **a, b,** or **c** may be changed to row vectors.

a. **aA** **b.** **bB** **c.** **cC**
d. **Aa** **e.** **Bb** **f.** **Cc**

7. Three different types of computers need varying amounts of four different types of integrated circuits. Matrix **A** gives the number of each circuit needed by each computer.

$$\begin{array}{cc} & \text{Circuit} \\ & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \text{Computers} \begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \end{array} & \begin{bmatrix} 2 & 3 & 2 & 1 \\ 5 & 1 & 3 & 2 \\ 3 & 2 & 2 & 2 \end{bmatrix} = \mathbf{A} \end{array}$$

Let $\mathbf{d} = [10 \quad 20 \quad 30]$ be the computer demand vector. Let

$$\mathbf{p} = \begin{bmatrix} 2 \\ 5 \\ 1 \\ 10 \end{bmatrix}$$

be the price vector for the circuits (the cost in dollars of each type of circuit).

a. Write an expression in terms of **A**, **d**, and **p** for the total cost of the circuits needed to produce the set of computers demanded; indicate where the matrix–vector product occurs and where the scalar product occurs.

b. Compute the total cost.

8. Let

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

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

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

Compute each matrix product (if possible).

a. **AB** **b.** **BA** **c.** **AC**
d. **CA** **e.** **CB**

9. Suppose that you are given the following matrices involving the costs of fruits at different stores, the amounts of fruit different types of people want, and the numbers of people of different types in different towns

$$\begin{array}{cc} & \begin{array}{cc} \text{Store 1} & \text{Store 2} \end{array} \\ \begin{array}{l} \text{Apples} \\ \text{Oranges} \\ \text{Pears} \end{array} & \begin{bmatrix} 0.10 & 0.15 \\ 0.15 & 0.20 \\ 0.10 & 0.10 \end{bmatrix} \end{array}$$

$$\begin{array}{cc} & \begin{array}{ccc} \text{Apples} & \text{Oranges} & \text{Pears} \end{array} \\ \begin{array}{l} \text{Person 1} \\ \text{Person 2} \end{array} & \begin{bmatrix} 5 & 10 & 3 \\ 4 & 5 & 5 \end{bmatrix} \end{array}$$

$$\begin{array}{cc} & \begin{array}{cc} \text{Person 1} & \text{Person 2} \end{array} \\ \begin{array}{l} \text{Town 1} \\ \text{Town 2} \end{array} & \begin{bmatrix} 1000 & 500 \\ 2000 & 1000 \end{bmatrix} \end{array}$$

a. Compute a matrix that represents the cost of each person's fruit purchases at each store.

b. Compute a matrix that represents the quantity of each fruit to be purchased in each town.

10. **a.** For the Markov chain matrix **A** in Problem 4, compute $\mathbf{A}^2$, $\mathbf{A}^3$, and $\mathbf{A}^5$.
    b. What vectors do the columns of the powers of **A** appear to be approaching?

11. Solve each system of equations, using Gaussian elimination.

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

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

12. Solve the system of equations obained for the furniture model in Problem 3.

13. Find the stable distribution for the Markov chain in Problem 4.

14. Use matrix methods to find the equation of the parabola that passes through the three points $(1, 1)$, $(2, 2)$, and $(3, 5)$.