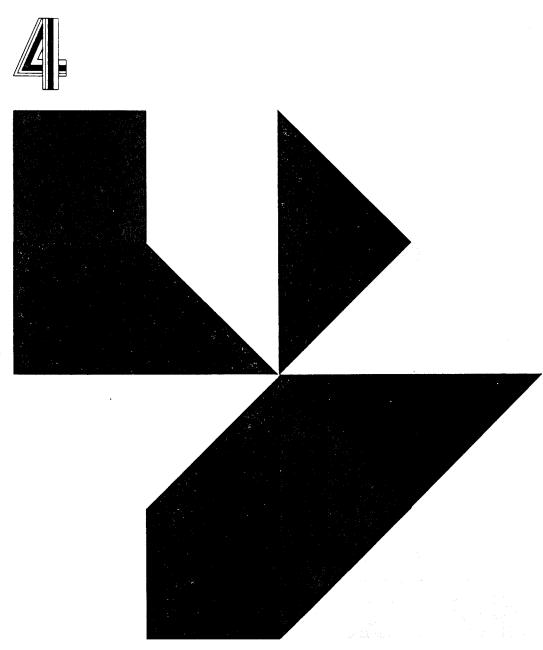
Vectors and Matrices





1 COLUMN AND ROW VECTORS

A column vector is an ordered collection of numbers written in a column. Examples of such vectors are

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} .6 \\ .4 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \\ 4 \end{pmatrix}.$$

The individual numbers in these vectors are called *components*, and the number of components a vector has is one of its distinguishing characteristics. Thus the first two vectors above have two components; the next two have three components; and the last has four components. When talking more generally about *n*-component column vectors we shall write

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Analogously, a *row vector* is an ordered collection of numbers written in a row. Examples of row vectors are

$$(1,0), (-2,1), (2,-3,4,0), (-1,2,-3,4,-5).$$

Each number appearing in the vector is again called a *component* of the vector, and the number of components a row vector has is again one of its important characteristics. Thus the first two examples are two-component, the third a four-component, and the fourth a five-component vector. The vector $v = (v_1, v_2, \ldots, v_n)$ is an *n*-component row vector.

Two row vectors, or two column vectors, are said to be equal if and only if corresponding components of the vector are equal. Thus for the vectors

$$u = (1, 2),$$
 $v = {1 \choose 2},$ $w = (1, 2),$ $x = (2, 1)$

we see that u = w but $u \neq v$, and $u \neq x$.

If u and v are three-component column vectors, we shall define their sum u + v by componentwise addition as follows:

$$u + v = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{pmatrix}.$$

Similarly, if u and v are three-component row vectors, their sum is defined to be

$$u + v = (u_1, u_2, u_3) + (v_1, v_2, v_3)$$

= $(u_1 + v_1, u_2 + v_2, u_3 + v_3).$

Note that the sum of two three-component vectors yields another three-component vector. For example,

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

and

$$(4, -7, 12) + (3, 14, -14) = (7, 7, -2).$$

The sum of two *n*-component vectors (either row or column) is defined by componentwise addition in an analogous manner, and yields another *n*-component vector. Observe that we do not define the addition of vectors unless they are both row or both column vectors having the same number of components.

Because the order in which two numbers are added does not affect the answer, it is also true that the order in which vectors are added does not matter; that is,

$$u + v = v + u$$

where u and v are both row or both column vectors. This is the so-called commutative law of addition. A numerical example is

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

Once we have the definition of the addition of two vectors, we can easily see how to add three or more vectors by grouping them in pairs as in the addition of numbers. For example,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

and

$$(1,0,0) + (0,2,0) + (0,0,3) = (1,2,0) + (0,0,3) = (1,2,3)$$

= $(1,0,0) + (0,2,3) = (1,2,3)$.

In general, the sum of any number of vectors (row or column), each having the same number of components, is the vector whose first component is the sum of the first components of the vectors, whose second component is the sum of the second components, and so on.

The multiplication of a number a times a vector v is defined by componentwise multiplication of a times the components of v. For the three-component case we have

$$au = a \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} au_1 \\ au_2 \\ au_3 \end{pmatrix}$$

for column vectors and

$$av = a(v_1, v_2, v_3) = (av_1, av_2, av_3)$$

for row vectors. If u is an n-component vector (row or column), then au is defined similarly by componentwise multiplication. This operation is sometimes called *scalar multiplication* of a vector, where *scalar* is another name for a number.

If u is any vector, we define its negative -u to be the vector -u = (-1)u. Thus in the three-component case for row vectors we have

$$-u = (-1)(u_1, u_2, u_3) = (-u_1, -u_2, -u_3).$$

Once we have the negative of a vector it is easy to see how to subtract vectors: we simply add "algebraically." For the three-component column-vector case we have

$$u - v = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} - \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1 - v_1 \\ u_2 - v_2 \\ u_3 - v_3 \end{pmatrix}.$$

Specific examples of subtraction of vectors occur in the exercises at the end of this section.

An important vector is the zero vector, all of whose components are zero. For example, three-component zero vectors are

$$0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
 and $0 = (0, 0, 0)$.

When there is no danger of confusion we shall use the symbol 0, as above, to denote the zero (row or column) vector. The meaning will be clear from the context. The zero vector has the important property that, if u is any vector, then u + 0 = u. A proof for the three-component column-vector case is as follows:

$$u + 0 = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} u_1 + 0 \\ u_2 + 0 \\ u_3 + 0 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = u.$$

One of the chief advantages of the vector notation is that we can denote a whole collection of numbers by a single letter such as u, v, \ldots , and treat such a collection as if it were a single quantity. By using the vector notation we can state very complicated relationships in a simple manner. The student will see many examples of this in the remainder of the present chapter and in the three succeeding chapters.

EXERCISES

1. Compute the quantities below for the vectors

$$u = \begin{pmatrix} 2 \\ 1 \\ 8 \\ 6 \end{pmatrix}, \qquad v = \begin{pmatrix} 3 \\ -5 \\ 2 \\ 0 \end{pmatrix}, \qquad w = \begin{pmatrix} 6 \\ 6 \\ -6 \\ -6 \end{pmatrix}.$$

(a)
$$u + w$$
.
$$[Ans. \begin{pmatrix} 8 \\ 7 \\ 2 \\ 0 \end{pmatrix}]$$

- **(b)** 5w.
- (c) v-u.

(d)
$$3u + 7v - 2w$$
. [Ans. $\begin{pmatrix} 15 \\ -44 \\ 50 \\ 30 \end{pmatrix}$.]

- (e) $\frac{1}{2}w + \frac{3}{4}u$.
- (f) u-w-v.
- (g) -2u + 3v 100w.
- 2. Compute (a) through (g) of Exercise 1 if the vectors u, v, and w are

$$u = (7, 0, -3),$$
 $v = (2, 1, -5),$ $w = (1, -1, 0).$

- 3. (a) Show that the zero vector is not changed when multiplied by any number.
 - **(b)** If u is any vector, show that 0 + u = u.

- 4. If 2u v = 0, what is the relationship between the components of u and those of v?

 [Ans. $v_i = 2u_i$.]
- 5. When possible, compute the following sums; when not possible, give reasons.

(a)
$$(2,3) + 2 \binom{1}{5} = ?$$

(b)
$$0(5, 1, 7) + 3(6, 2, 6) = ?$$

(c)
$$\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + 5 + \begin{pmatrix} 6 \\ 2 \\ 6 \end{pmatrix} = ?$$
 [Ans. Not possible.]

(d)
$$\begin{pmatrix} 21\\22\\23\\24 \end{pmatrix} + 32 \begin{pmatrix} 1\\2\\0\\1 \end{pmatrix} = ?$$
 [Ans. $\begin{pmatrix} 53\\86\\23\\56 \end{pmatrix}$.]

6. If
$$\begin{pmatrix} 6 \\ 6 \\ 0 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -5 \\ 5 \end{pmatrix}$$
, find u_1 , u_2 , and u_3 . [Ans. -1 , -11 , 5 .]

7. If
$$8 \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 7 \\ -16 \\ 0 \end{pmatrix}$$
, find the components of v .

8. Find three vectors u, v, and w such that w = 3u, v = 2u, and $2u + 3v + 4w = \begin{pmatrix} 20 \\ 10 \\ -25 \end{pmatrix}.$

9. If
$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
, what can be said concerning the components u_1 , u_2 , u_3 ?

10. If $0 \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, what can be said concerning the components u_1 , $u_2 u_3$?

11. If
$$(u_1 + u_2 + u_3 + u_4) \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix}$$
, what are u_1, u_2, u_3 , and u_4 ?

12. (a) Show that the vector equation

$$x \binom{5}{7} + y \binom{3}{-10} = \binom{-16}{77}$$

represents two simultaneous linear equations for the two variables x and y.

- (b) Solve the equations for x and y from (a) and substitute into the vector equation above to check your work.
- 13. Write the following simultaneous linear equations in vector form:

$$ax + by = e$$
$$cx + dy = f.$$

[Hint: Follow the form given in Exercise 12.]

- 14. Suppose that we associate with each person a three-component row vector having the following entries: age, height, and weight. Would it make sense to add together the vectors associated with two different persons? Would it make sense to multiply one of these vectors by a constant?
- 15. Suppose that we associate with each person leaving a supermarket a row vector whose components give the quantities of each available item that he has purchased. Answer the same questions as those in Exercise 14.
- 16. Let us associate with each supermarket a column vector whose entries give the prices of each item in the store. Would it make sense to add together the vectors associated with two different supermarkets? Would it make sense to multiply one of these vectors by a constant? Discuss the differences in the situations given in Exercises 14, 15, and 16.
- 17. Consider the vectors

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Show that the vector

$$\frac{1}{2}(x+y)$$

has components that are the averages of the components of x and y. Generalize this result to the case of n vectors.

- 18. Would the concept of averages, as discussed in Exercise 17, be applicable to the vectors mentioned in Exercises 14, 15, and 16? How would the averages be interpreted?
- 19. In a certain school students take four courses each semester. At the end of the semester the registrar records the grades of each student as a row vector. He then gives the student 4 points for each A, 3 points for each B, 2 points for each C, 1 point for each D, and 0 for each F. The sum of these numbers, divided by 4, is the student's grade point average.
 - (a) If a student has a 4.0 average, what are the logical possibilities for his grade vector? [Hint: Each grade vector will have five components.]
 - (b) What are the possibilities if he has a 3.0 average?
 - (c) What are the possibilities if he has a 2.0 average?

- **20.** Let $x = \binom{x_1}{x_2}$. Define $x \ge 0$ to be the conjunction of the statements $x_1 \ge 0$ and $x_2 \ge 0$. Define $x \le 0$ analogously. If $(x_1 + x_2) x \ge 0$, what must be true of x?
- 21. Using the definition in Exercise 20, define $x \ge y$ to mean $x y \ge 0$, where x and y are vectors of the same shape. Consider the following four vectors:

$$x = \begin{pmatrix} 3 \\ 5 \\ -1 \end{pmatrix}, \quad y = \begin{pmatrix} 6 \\ 5 \\ 6 \end{pmatrix}, \quad u = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}, \quad v = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}.$$

- (a) Show that x > u.
- **(b)** Show that v > u.
- (c) Is there any relationship between x and v?
- (d) Show that $y \ge x$, $y \ge u$, and $y \ge v$.
- 22. (a) If $x^{(1)}$, $x^{(2)}$, ..., $x^{(n)}$ is a set of *n* vectors, show how to find a vector *u* such that $u \ge x^{(i)}$ for all *i*. Also show how to find a vector *v* such that $v \le x^{(i)}$ for all *i*.
 - (b) Apply the results of part (a) to the vectors in Exercise 21.

(c) Let
$$u = \begin{pmatrix} 630 \\ 520 \\ 310 \end{pmatrix}$$
, $v = \begin{pmatrix} 960 \\ 200 \\ 400 \end{pmatrix}$, and $w = \begin{pmatrix} 600 \\ 750 \\ 490 \end{pmatrix}$. If $x = \begin{pmatrix} -4 \\ -5 \\ -2 \end{pmatrix}$, find

the largest number n such that $nx \ge u$, $nx \ge v$, and $nx \ge w$. [Ans. -235.]

2 THE PRODUCT OF VECTORS; EXAMPLES

The reader may wonder why it is necessary to introduce both column and row vectors when their properties are so similar. This question can be answered in several different ways. First, in many applications two kinds of quantities are studied simultaneously, and it is convenient to represent one of them as a row vector and the other as a column vector. Second, there is a way of combining row and column vectors that is very useful for certain types of calculations. To bring out these points let us look at the following simple economic example.

EXAMPLE 1 Suppose a man named Smith goes into a grocery store to buy a dozen each of peaches and oranges, a half-dozen each of apples and pears, and three lemons. Let us represent his purchases by means of the following row vector:

$$x = [6 \text{ (apples)}, 12 \text{ (peaches)}, 3 \text{ (lemons)}, 12 \text{ (oranges)}, 6 \text{ (pears)}]$$

= $(6, 12, 3, 12, 6)$.

Suppose that apples are 4 cents each, peaches 6 cents, lemons 9 cents, oranges 5 cents, and pears 7 cents. We can then represent the prices of

these items as a column vector:

$$y = \begin{pmatrix} 4 \\ 6 \\ 9 \\ 5 \\ 7 \end{pmatrix}$$
 cents per apple cents per peach cents per lemon cents per orange cents per pear.

The obvious question is: What is the total amount that Smith must pay for his purchases? We would like to multiply the quantity vector x by the price vector y, and we would like the result to be Smith's bill. We see that our multiplication should have the following form:

$$x \cdot y = (6, 12, 3, 12, 6) \begin{pmatrix} 4 \\ 6 \\ 9 \\ 5 \\ 7 \end{pmatrix}$$

$$= 6 \cdot 4 + 12 \cdot 6 + 3 \cdot 9 + 12 \cdot 5 + 6 \cdot 7$$

$$= 24 + 72 + 27 + 60 + 42$$

$$= 225 \text{ cents or } \$2.25.$$

This is, of course, the computation that the cashier performs in figuring Smith's bill.

We shall adopt in general the above definition of multiplication of row times column vectors.

Definition Let u be a row vector and v a column vector each having the same number n of components; then we shall define the produce $u \cdot v$ to be

$$u \cdot v = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n.$$

Notice that we always write the row vector first and the column vector second, and this is the only kind of vector multiplication that we consider. Some examples of vector multiplication are

$$(2, 1, -1) \cdot \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} = 2 \cdot 3 + 1 \cdot (-1) + (-1) \cdot 4 = 1,$$

$$(1, 0) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \cdot 0 + 0 \cdot 1 = 0 + 0 = 0.$$

Note that the result of vector multiplication is always a number.

EXAMPLE 2 Consider an oversimplified economy that has three industries, which we call coal, electricity, and steel, and three consumers 1, 2, and 3. Suppose that each consumer uses some of the output of each industry and also that each

industry uses some of the output of each other industry. We assume that the amounts used are positive or zero, since using a negative quantity has no immediate interpretation. We can represent the needs of each consumer and industry by a three-component demand (row) vector, the first component measuring the amount of coal needed by the consumer or industry; the second component the amount of electricity needed; and the third component the amount of steel needed, in some convenient units. For example, the demand vectors of the three consumers might be

$$d_1 = (3, 2, 5),$$
 $d_2 = (0, 17, 1),$ $d_3 = (4, 6, 12)$

and the demand vectors of each of the industries might be

$$d_C = (0, 1, 4),$$
 $d_E = (20, 0, 8),$ $d_S = (30, 5, 0),$

where the subscript C stands for coal, the subscript E for electricity, and the subscript S for steel. Then the total demand for these goods by the consumers is given by the sum

$$d_1 + d_2 + d_3 = (3, 2, 5) + (0, 17, 1) + (4, 6, 12) = (7, 25, 18).$$

Also, the total industrial demand for these goods is given by the sum

$$d_C + d_E + d_S = (0, 1, 4) + (20, 0, 8) + (30, 5, 0) = (50, 6, 12).$$

Therefore the total overall demand is given by the sum

$$(7, 25, 18) + (50, 6, 12) = (57, 31, 30).$$

Suppose now that the price of coal is \$1 per unit, the price of electricity is \$2 per unit, and the price of steel is \$4 per unit. Then these prices can be represented by the column vector

$$p = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

Consider the steel industry: it sells a total of 30 units of steel at \$4 per unit, so that its total income is \$120. Its bill for the various goods is given by the vector product

$$d_{S} \cdot p = (30, 5, 0) \cdot \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = 30 + 10 + 0 = $40.$$

Hence the profit of the steel industry is \$120 - \$40 = \$80. (In the exercises the profits of the other industries will be found.)

This model of an economy is unrealistic in two senses. First, we have not chosen realistic numbers for the various quantities involved. Second, and more important, we have neglected the fact that the more an industry produces the more inputs it requires.

EXAMPLE 3 Consider the rectangular coordinate system in the plane shown in Figure 1. A two-component row vector x = (a, b) can be regarded as a point in

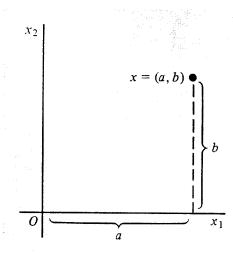


Figure 1

the plane located by means of the coordinate axes as shown. The point x can be found by starting at the origin of coordinates O and moving a distance a along the x_1 axis; then moving a distance b along a line parallel to the x_2 axis. If we have two such points, say x = (a, b) and y = (c, d), then the points x + y, -x, -y, x - y, y - x, -x - y have the geometric significance shown in Figure 2.

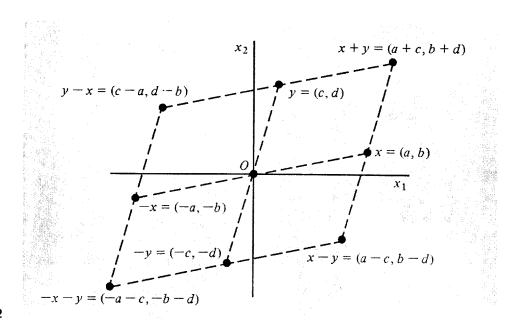


Figure 2

The idea of multiplying a row vector by a number can also be given a geometric meaning. In Figure 3 we have plotted the point corresponding to the vector x = (1, 2), and 2x, $\frac{1}{2}x$, -x, and -2x. Observe that all these points lie on a line through the origin of coordinates. Another vector quantity that has geometrical significance is the vector z = ax + (1 - a)y,

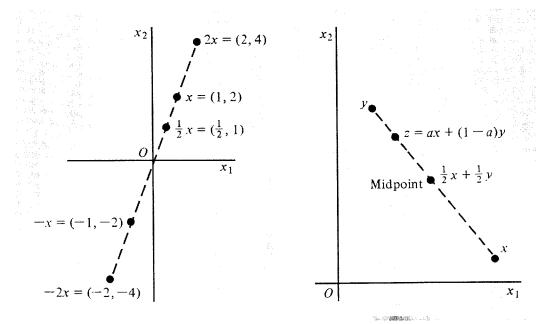


Figure 3 Figure 4

where a is any number between 0 and 1. Observe in Figure 4 that the points z all lie on the line segment between the points x and y. If $a = \frac{1}{2}$, the corresponding point on the line segment is the midpoint of the segment. Thus, if x = (a, b) and y = (c, d), then the point

$$\frac{1}{2}x + \frac{1}{2}y = \frac{1}{2}(a,b) + \frac{1}{2}(c,d)$$
$$= \left(\frac{a+c}{2}, \frac{b+d}{2}\right)$$

is the midpoint of the line segment between x and y.

EXERCISES

1. Let
$$u = (2, -1, 3), v = (5, 0, 2), x = \begin{pmatrix} -7 \\ 1 \\ 2 \end{pmatrix}$$
, and $y = \begin{pmatrix} 0 \\ 8 \\ 3 \end{pmatrix}$.

Compute the following:

(a)
$$(u + v) \cdot (x + y)$$
.

(b)
$$((3u \cdot x) \cdot v) \cdot y$$
.

(c)
$$u \cdot x - 4v \cdot y$$
.

(d)
$$u \cdot x + 3u \cdot y - v \cdot y$$
. [Ans. -12.]

(e) $(2(v+u)\cdot y) - 5uy$.

(f)
$$4u \cdot x + 6[v \cdot (3x - y)]$$
. [Ans. -630.]

- 2. If x = (5, -4, 2) and y = (1, 8, 1) are points in space, what is the midpoint of the line segment joining x to y? [Ans. $(3, 2, \frac{3}{2})$.]
- 3. Let x = (-1, 3) and y = (4, 1) be row vectors. Plot the points corresponding to x and y, and compute and plot the following vectors:

- (a) $\frac{1}{2}x + \frac{1}{2}y$.
- (d) 2x y.
- (b) x + y.
- (e) y + 4x.
- (c) $\frac{2}{3}x + \frac{1}{3}y$.
- (f) x-y.
- 4. Prove that vector multiplication satisfies the following two properties:
 - (i) $u \cdot (av) = a(u \cdot v)$,
 - (ii) $u \cdot (v + w) = u \cdot v + u \cdot w$,

where u is a three-component row vector, v and w are three-component column vectors, and a is a number.

- 5. If u is a three-component row vector, v is a three-component column vector having the same number of components, and a is a number, prove that $a(u \cdot v) = u \cdot (av)$.
- 6. A certain football stadium has three gates. After one game, the ticket-taker at gate 1 reported admitting 275 adults, 300 students, and 15 children; the ticket-taker at gate 2 admitted 200 adults, 107 students, and 40 children; and 65 adults, 250 students, and 60 children were admitted at gate 3.
 - (a) Write the numbers of people admitted through each gate as a three-component row vector.
 - (b) Use vector addition to find how many people in each category attended the game.
 - (c) Suppose that adults pay \$3.00 to attend the game, students pay \$2.00, and children pay 50¢. Assuming that each person buys his ticket at the gate, calculate the value of the tickets sold at each gate.
 - (d) Compute in two different ways the total value of tickets sold.
- 7. Perform the following calculations for Example 2.
 - (a) Compute the amount that each industry and each consumer has to pay for the goods it receives.
 - (b) Compute the profit made by each of the industries.
 - (c) Find the total amount of money that is paid out by all the industries and consumers.
 - (d) Find the proportion of the total amount of money found in (c) paid out by the industries. Find the proportion of the total money that is paid out by the consumers.
- 8. A farmer intends to plant 40 acres of corn, 25 acres of wheat, and 30 acres of rye. Write a three-component row vector whose components give the number of acres of each grain he wants to plant. Suppose an acre of corn requires an hour to plant, an acre of wheat 45 minutes, and an acre of rye one-half hour. Write a column vector whose components give the number of minutes needed to plant an acre of each crop. Use vector multiplication to find the total time required for planting.
- 9. In Exercise 8, suppose the seed for an acre of corn cost \$40, for an acre of wheat \$15, and for an acre of rye \$10. Find the total amount

the farmer must spend on seed. If an acre of corn brings the farmer a *profit* of \$19, an acre of wheat \$15, and an acre of rye \$4, find the total profit the farmer will make.

- 10. The production of a book involves several steps: first it must be set into type, then it must be printed, and finally it must be supplied with covers and bound. Suppose the typesetter charges \$6 an hour, paper costs \(\frac{1}{4}\) cent per sheet, the printer charges 11 cents for each minute that his press runs, the cover costs 28 cents, and the binder charges 15 cents to bind each book. Suppose now that a publisher wishes to print a book that requires 300 hours of work by the typesetter, 220 sheets of paper per book, and 5 minutes of press time per book.
 - (a) Write a five-component row vector that gives the requirements for the first book. Write another row vector that gives the requirements for the second, third, . . . copies of the book. Write a five-component column vector whose components give the prices of the various requirements for each book, in the same order as they are listed in the requirement vectors above.
 - (b) Using vector multiplication, find the cost of publishing one copy of a book.

 [Ans. \$1801.53.]
 - (c) Using vector addition and multiplication, find the cost of printing a first-edition run of 5000 copies. [Ans. \$9450.]
 - (d) Assuming that the printing plates from the first edition are used again, find the cost of printing a second edition of 5000 copies.

 [Ans. \$7650.]
- 11. Let $x = {x_1 \choose x_2}$, and let a and b be the vectors a = (-1, 4), b = (2, -7). If ax = 1 and bx = 2, find x_1 and x_2 .

 [Ans. $x_1 = 15, x_2 = 4$.]
- 12. Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, and let a and b be the vectors a = (3, 3), b = (-1, 4). If $ax = x_1$ and $bx = x_2$, find x_1 and x_2 .
- 13. Consider the vectors

$$a = (a_1, a_2),$$
 $b = (b_1, b_2),$ $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

and two numbers c_1 and c_2 . Show that the equations

$$ax = c_1, \quad bx = c_2$$

represent two simultaneous equations in two unknowns.

- 14. Show that every set of two simultaneous equations in two unknowns can be written as in Exercise 13.
- 15. Consider an experiment in which there are two outcomes: we win \$10 with probability $\frac{1}{5}$ and lost \$5 with probability $\frac{4}{5}$. Let a = (10, -5) and $p = \begin{pmatrix} \frac{1}{5} \\ \frac{4}{5} \end{pmatrix}$. Show that the expected outcome of the experiment is ap.

- 16. A gambling game works as follows. Two dice are rolled. If no sixes turn up, we lose \$1; if one six turns up, we win \$1; if two sixes turn up we win \$10. Set up a row vector representing the various outcomes and a column vector representing the probability of those outcomes. Use vector multiplication to find the expected outcomes. Is the game fair?
- 17. If an experiment has outcomes a_1, a_2, \ldots, a_n occurring with probabilities p_1, p_2, \ldots, p_n , define the vectors

$$a = (a_1, \dots, a_n)$$
 and $p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}$

Show that the expected outcome is ap.

- **18.** Consider the vectors x = (1, 5), y = (3, 1), and $f = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
 - (a) Compute $\frac{1}{2}xf$ and $\frac{1}{2}yf$, and show that these numbers are the averages of the components of x and y, respectively. [Partial Ans. 3, 2.]
 - (b) Compute $\frac{1}{4}(x+y)f$, and give an interpretation for this number. [Partial Ans. $2\frac{1}{2}$.]
- 19. Let x and y be two n-component row vectors, and let f be an n-component column vector all of whose entries are 1's.
 - (a) Compute (1/n)xf and (1/n)yf and interpret the result.
 - (b) Compute $(\frac{1}{2}n)(x + y)f$ and interpret the result. [Hint: Exercise 18 is a special case.]
- 20. How would the results of Exercise 19 change if we used three vectors: x, y, and z?

3 MATRICES AND THEIR COMBINATION WITH VECTORS

A matrix is a rectangular array of numbers written in the form

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

Here the letters a_{ij} stand for real numbers and m and n are integers. Observe that m is the number of rows and n is the number of columns of the matrix. For this reason we call it an $m \times n$ matrix. If m = n, the matrix is square. The following are examples of matrices:

$$\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix};
\begin{pmatrix}
1 & 7 & -8 & 9 & 10 \\
3 & -1 & 14 & 2 & -6 \\
0 & 3 & -5 & 7 & 0
\end{pmatrix}.$$

The first example is a row vector which is a 1×3 matrix; the second is a column vector which is a 3×1 matrix; the third example is a 2×2 square matrix; the fourth is a 4×4 square matrix; and the last is a 3×5 matrix.

Two matrices having the same shape (i.e., having the same number of rows and columns) are said to be equal if and only if the corresponding entries are equal.

Recall that in Chapter 3, Section 12, we found that a matrix arose naturally in the consideration of a Markov chain process. To give another example of how matrices occur in practice and are used in connection with vectors, we consider the following example.

EXAMPLE 1 Suppose that a building contractor has accepted orders for five ranch-style houses, seven Cape Cod houses, and twelve colonial-style houses. We can represent his orders by means of a row vector x = (5, 7, 12). The contractor is familiar, of course, with the kinds of "raw materials" that go into each type of house. Let us suppose that these raw materials are steel, wood, glass, paint, and labor. The numbers in the matrix below give the amounts of each raw material going into each type of house, expressed in convenient units. (The numbers are put in arbitrarily, and are not meant to be realistic.)

	Steel	Wood	Glass	Paint	Labor	
Ranch:	5	20	16	7	17	
Cape Cod:	7	18	12	9	21	= R.
Cape Cod: Colonial:	\ 6	25	8	5	13 /	

Observe that each row of the matrix is a five-component row vector which gives the amounts of each raw material needed for a given kind of house. Similarly, each column of the matrix is a three-component column vector which gives the amounts of a given raw material needed for each kind of house. Clearly, a matrix is a very succinct way of summarizing this information.

Suppose now that the contractor wishes to compute how much of each raw material to obtain in order to fulfill his contracts. Let us denote the matrix above by R; then he would like to obtain something like the product xR, and he would like the product to tell him what orders to make out. The product should have the following form:

$$xR = (5,7,12) \begin{pmatrix} 5 & 20 & 16 & 7 & 17 \\ 7 & 18 & 12 & 9 & 21 \\ 6 & 25 & 8 & 5 & 13 \end{pmatrix}$$

$$= (5 \cdot 5 + 7 \cdot 7 + 12 \cdot 6, \quad 5 \cdot 20 + 7 \cdot 18 + 12 \cdot 25, \\ 5 \cdot 16 + 7 \cdot 12 + 12 \cdot 8, \quad 5 \cdot 7 + 7 \cdot 9 + 12 \cdot 5, \\ 5 \cdot 17 + 7 \cdot 21 + 12 \cdot 13)$$

$$= (146,526,260,158,388).$$

Thus we see that the contractor should order 146 units of steel, 526 units of wood, 260 units of glass, 158 units of paint, and 388 units of labor. Observe that the answer we get is a five-component row vector and that each entry in

this vector is obtained by taking the vector product of x times the corresponding column of the matrix R.

The contractor is also interested in the prices that he will have to pay for these materials. Suppose that steel costs \$15 per unit, wood costs \$8 per unit, glass costs \$5 per unit, paint costs \$1 per unit, and labor costs \$10 per unit. Then we can write the cost as a column vector as follows:

$$y = \begin{pmatrix} 15\\8\\5\\1\\10 \end{pmatrix}.$$

Here the product Ry should give the costs of each type of house, so that the multiplication should have the form

$$Ry = \begin{pmatrix} 5 & 20 & 16 & 7 & 17 \\ 7 & 18 & 12 & 9 & 21 \\ 6 & 25 & 8 & 5 & 13 \end{pmatrix} \begin{pmatrix} 15 \\ 8 \\ 5 \\ 1 \\ 10 \end{pmatrix}$$
$$= \begin{pmatrix} 5 \cdot 15 + 20 \cdot 8 + 16 \cdot 5 + 7 \cdot 1 + 17 \cdot 10 \\ 7 \cdot 15 + 18 \cdot 8 + 12 \cdot 5 + 9 \cdot 1 + 21 \cdot 10 \\ 6 \cdot 15 + 25 \cdot 8 + 8 \cdot 5 + 5 \cdot 1 + 13 \cdot 10 \end{pmatrix}$$
$$= \begin{pmatrix} 492 \\ 528 \\ 465 \end{pmatrix}.$$

Thus the cost of materials for the ranch style house is \$492, for the Cape Cod house is \$528, and for the Colonial house \$465.

The final question which the contractor might ask is what is the total cost of raw materials for all the houses he will build. It is easy to see that this is given by the vector xRy. We can find it in two ways as shown below.

$$xRy = (xR)y = (146, 526, 260, 158, 388) \cdot \begin{pmatrix} 15\\8\\5\\1\\10 \end{pmatrix} = 11,736$$

 $xRy = x(Ry) = (5, 7, 12) \cdot \begin{pmatrix} 492\\528\\465 \end{pmatrix} = 11,736.$

The total cost is then \$11,736.

We shall adopt, in general, the above definitions for the multiplication of a matrix times a row or a column vector. **Definition** Let A be an $m \times n$ matrix, let x be an m-component row vector, and let u be an n-component column vector; then we define the products xA and Au as follows:

$$xA = (x_{1}, x_{2}, \dots, x_{m}) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

$$= (x_{1}a_{11} + x_{2}a_{21} + \dots + x_{m}a_{m1}, \quad x_{1}a_{12} + x_{2}a_{22} + \dots + x_{m}a_{m2}, \\ & \dots, \quad x_{1}a_{1n} + x_{2}a_{2n} + \dots + x_{m}a_{mn});$$

$$Au = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} u_{1} \\ u_{2} \\ \vdots \\ u_{n} \end{pmatrix} = \begin{pmatrix} a_{11}u_{1} + a_{12}u_{2} + \dots + a_{1n}u_{n} \\ a_{21}u_{1} + a_{22}u_{2} + \dots + a_{2n}u_{n} \\ & \dots & \\ a_{m1}u_{1} + a_{m2}u_{2} + \dots + a_{mn}u_{n} \end{pmatrix}.$$

The reader will find these formulas easy to work with if he observes that each entry in the products xA or Au is obtained by vector multiplication of x or u by a column or row of the matrix A. Notice that in order to multiply a row vector times a matrix, the number of rows of the matrix must equal the number of components of the vector, and the result is another row vector; similarly, to multiply a matrix times a column vector, the number of columns of the matrix must equal the number of components of the vector, and the result of such a multiplication is another column vector.

Some numerical examples of the multiplication of vectors and matrices are:

$$(1,0,-1)\begin{pmatrix} 3 & 1 \\ 2 & 3 \\ 2 & 8 \end{pmatrix} = (1 \cdot 3 + 0 \cdot 2 - 1 \cdot 2, 1 \cdot 1 + 0 \cdot 3 - 1 \cdot 8)$$

$$= (1,-7);$$

$$\begin{pmatrix} 3 & 1 & 2 \\ 2 & 3 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 - 1 + 4 \\ 2 - 3 + 16 \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \end{pmatrix};$$

$$\begin{pmatrix} 3 & 2 & -1 \\ 1 & 0 & 2 \\ 0 & 3 & 1 \\ 5 & -4 & 7 \\ -3 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \\ -2 \\ -9 \\ -1 \end{pmatrix}.$$

Observe that if x is an m-component row vector and A is $m \times n$, then xA is an n-component row vector; similarly, if u is an n-component column vector, then Au is an m-component column vector. These facts can be observed in the examples above.

EXAMPLE 2 Consider a Markov chain with transition matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Choose the initial state by a random device that selects states a_1 and a_2 each with probability $\frac{1}{2}$. Let us indicate the choice of initial state by the vector $p^{(0)} = (\frac{1}{2}, \frac{1}{2})$ where the first component gives the probability of choosing state a_1 and the second the probability of choosing state a_2 . Let us compute the product $p^{(0)}P$. We have

$$p^{(0)}P = \begin{pmatrix} \frac{1}{2}, \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{6} + \frac{1}{4}, \frac{1}{3} + \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{5}{12}, \frac{7}{12} \end{pmatrix}.$$

Using the methods of Chapter 3, one can show that after one step there is probability $\frac{5}{12}$ that the process will be in state a_1 and probability $\frac{7}{12}$ that it will be in state a_2 . Let $p^{(1)}$ be the vector whose first component gives the probability of the process being in state a_1 after one step and whose second component gives the probability of it being in state a_2 after one step. In our example we have $p^{(1)} = (\frac{5}{12}, \frac{7}{12}) = p^{(0)}P$.

In general, the formula $p^{(1)} = p^{(0)}P$ holds for any Markov process with transition matrix P and initial probability vector $p^{(0)}$.

EXAMPLE 3 In Example 1 of Section 2 assume that Smith has two stores at which he can make his purchases, and let us assume that the prices charged at these two stores are slightly different. Let the price vector at the second store be

$$y = \begin{pmatrix} 5 \\ 5 \\ 10 \\ 4 \end{pmatrix}$$
 cents per apple cents per peach cents per lemon cents per orange cents per pear.

Smith now has the option of buying all his purchases at store 1, all at store 2, or buying just the lower-priced items at the store charging the lower price. To help him decide, we form a price matrix as follows:

	Prices,		Prices,		Minimum	
	Store 1		Store 2		Price	
	/ 4		5		4	١
	6		5		5	- 1
P =	9	`	10		9	
	5		4		4	- 1
	\ 7		, 6		6	

The first column lists the prices of store 1, the second column lists the price of store 2, and the third column lists the lower of these two prices. To compute Smith's bill under the three possible ways he can make his pur-

chases, we compute the produce xP, as follows:

$$xP = (6, 12, 3, 12, 6) \begin{pmatrix} 4 & 5 & 4 \\ 6 & 5 & 5 \\ 9 & 10 & 9 \\ 5 & 4 & 4 \\ 7 & 6 & 6 \end{pmatrix} = (225, 204, 195).$$

We thus see that if Smith buys only in store 1, his bill will be \$2.25; if he buys only in store 2, his bill will be \$2.04; but if he buys each item in the cheaper of the two stores (apples and lemons in store 1, and the rest in store 2), his bill will be \$1.95.

Exactly what Smith will, or should, do depends upon circumstances. If both stores are equally close to him, he will probably split his purchases and obtain the smallest bill. If store 1 is close and store 2 is very far away, he may buy everything at store 1. If store 2 is closer and store 1 is far enough away so that the 9 cents he would save by splitting his purchases is not worth the travel effort, he may buy everything at store 2.

The problem just cited is an example of a decision problem. In such problems it is necessary to choose one of several courses of action, or strategies. For each such course of action or strategy, it is possible to compute the cost or worth of such a strategy. The decision maker will choose a strategy with maximum worth.

Sometimes the worth of an outcome must be measured in psychological units and we then say that we measure the *utility* of an outcome. For the purposes of this book we shall always assume that the utility of an outcome is measured in monetary units, so that we can compare the worths of two different outcomes to the decision maker.

As a second example of a decision problem, consider the following. An urn contains five red, three green, and one white ball. One ball will be drawn at random, and then payments will be made to holders of three kinds of lottery tickets, A, B, and C, according to the following schedule:

$$M = \begin{array}{ccc} \text{Red} & \text{Ticket A} & \text{Ticket B} & \text{Ticket C} \\ M = \begin{array}{cccc} \text{Red} & 1 & 3 & 0 \\ 4 & 1 & 0 \\ \text{White} & 0 & 0 & 16 \end{array} \right).$$

Thus, if a red ball is selected, holders of ticket A will get \$1, holders of ticket B will get \$3, and holders of ticket C will get nothing. If green is chosen, the payments are 4, 1, and 0, respectively. If white is chosen, holders of ticket C get \$16, and the others nothing. Which ticket would we prefer to have?

Our decision will depend upon the concept of expected value discussed in the preceding chapter. The statements "draw a red ball," "draw a green ball," and "draw a white ball" have probabilities $\frac{5}{9}$, $\frac{3}{9}$, and $\frac{1}{9}$, respectively.

From these probabilities we can calculate the expected value of holding each of the lottery tickets as described in the last chapter. However, a compact way of performing all these calculations is to compute the product pM, where p is the probability vector

$$p = (\frac{5}{9}, \frac{3}{9}, \frac{1}{9}).$$

From this we have

$$pM = \begin{pmatrix} \frac{5}{9}, \frac{3}{9}, \frac{1}{9} \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 16 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \cdot \frac{5}{9} + 4 \cdot \frac{3}{9} + 0 \cdot \frac{1}{9}, & 3 \cdot \frac{5}{9} + 1 \cdot \frac{3}{9} + 0 \cdot \frac{1}{9}, & 0 \cdot \frac{5}{9} + 0 \cdot \frac{3}{9} + 16 \cdot \frac{1}{9} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{17}{9}, \frac{18}{9}, \frac{16}{9} \end{pmatrix}.$$

It is easy to see that the three components of pM give the expected values of holding lottery tickets A, B, and C, respectively. From these numbers we can see that ticket B is the best, A is the next best, and C is third best.

If we have to pay for the tickets, then the cost of the tickets will determine which is the best buy. If each ticket costs \$3 we would be better off by not buying any ticket, since we would then expect to lose money. If each ticket costs \$1 then we should buy ticket B, since it would give us a net expected gain of 2 - 1 = 1. If the first two tickets cost 2.10, and the third cost 1.50, we should buy ticket C since it is the only one for which we would have a positive net expectation.

EXERCISES

1. Perform the following multiplications:

(a)
$$\binom{7}{-3} \binom{6}{2} \binom{1}{-3} = ?$$
 [Ans. $\binom{-11}{-9}$.]
(b) $(2, -2) \binom{-3}{-1} \binom{2}{0} = ?$

(c)
$$\begin{pmatrix} 0 & 3 & 1 \\ 3 & -1 & 7 \\ -5 & 14 & -8 \\ 7 & 2 & 9 \\ 0 & -6 & 10 \end{pmatrix} \begin{pmatrix} 6 \\ 1 \\ -1 \end{pmatrix} = ?$$

(d)
$$(5,5)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = ?$$
 [Ans. $(0,0)$.]

(e)
$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 12 \\ 12 \end{pmatrix} = ?$$

(f)
$$(0, 1, -5)$$
 $\begin{pmatrix} 3 & 1 & 2 & 0 & -8 \\ 6 & 8 & 2 & 1 & 14 \\ 2 & 15 & 2 & 0 & -5 \end{pmatrix} = ?$

(g)
$$(x_1, x_2) \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ?$$
 [Ans. $ax_1 + cx_2, bx_1 + dx_2$.]

(h)
$$\binom{w}{y} \binom{x}{z} \binom{p_1}{p_2} = ?$$

(i)
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = ?$$

(j)
$$(x_1, x_2, x_3, x_4)$$
 $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = ?$

- 2. What number does the matrix in parts (i) and (j) of Exercise 1 resemble?
- 3. Notice that in Exercise 1(d), above, the product of a row vector none of whose components is zero and a matrix none of whose components is zero is the zero row vector. Find a second example, this time using a 3 × 3 matrix, which is similar to this one. Answer the analogous question for Exercise 1(e).
- 4. Consider the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

- (a) Show that the equation Ax = b represents two simultaneous equations in two unknowns.
- (b) Show that every set of two simultaneous equations in two unknowns can be written in this form for the proper choice of A and b.
- 5. When possible, solve for the indicated quantities.

(a)
$$(x_1, x_2) \begin{pmatrix} 6 & 8 \\ -9 & 0 \end{pmatrix} = (-45, 72)$$
. Find the vector x.

(b)
$$(6,9) \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (12, -15)$$
. Find the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

In this case can you find more than one solution?

(c)
$$\begin{pmatrix} 5 & -5 \\ -5 & 5 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}$$
. Find the vector u .

(d)
$$\binom{3}{-12} \binom{-7}{28} \binom{u_1}{u_2} = \binom{-1}{4}$$
. Find u . How many solutions can you find?

Ans. Infinitely many solutions, all of the form
$$u = \begin{pmatrix} k \\ 3k+1 \\ 7 \end{pmatrix}$$
.

6. Solve for the indicated quantities below and give an interpretation for each.

(a)
$$(1,-1)\begin{pmatrix} 0 & 2 \\ -2 & 4 \end{pmatrix} = a(1,-1)$$
; find a. [Ans. $a = 2$.]

(b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 5 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$; find u. How many answers can you find?

[Ans.
$$u = \binom{k}{2k}$$
 for any number k .]

(c)
$$\begin{pmatrix} 3 & 3 \\ 6 & 10 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$
; find u . How many answers are there?

- 7. In Example 1 of this section, assume that the contractor is to build eight ranch-style, four Cape Cod, and four colonial-type houses. Recompute, using matrix multiplication, the total cost of raw materials, in two different ways, as in the example.
- 8. In Example 2 use tree measures to show that $p^{(2)} = p^{(1)}P$.
- 9. In Example 2 of this section, assume that the initial probability vector is $p^{(0)} = (\frac{1}{6}, \frac{5}{6})$. Find the vector $p^{(1)}$. [Ans. $(\frac{17}{36}, \frac{19}{36})$.]
- 10. Consider the Markov chain with two states whose transition matrix is

$$P = \begin{pmatrix} a & 1 - a \\ 1 - b & b \end{pmatrix},$$

where a and b are nonnegative numbers less than 1. Suppose the initial probability vector for the process is $p^{(0)} = (p_1^{(0)}, p_2^{(0)})$, where $p_1^{(0)}$ is the initial probability of choosing state 1 and $p_2^{(0)}$ is the initial probability of choosing state 2. Derive the formulas for the components of the vector $p^{(1)}$.

[Ans. $p^{(1)} = ap_1^{(0)} + (1-b)p_2^{(0)}$, $(1-a)p_1^{(0)} + bp_2^{(0)}$.]

11. Suppose that $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ is the transition matrix of a three-state

Markov chain. Find
$$(x, y, z) \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
. [Ans. $x + y + z$.]

12. The following matrix gives the vitamin contents of three food items, in conveniently chosen units:

If we eat 11 units of food I, 6 units of food II, and 4 units of food III, how much of each type of vitamin have we consumed? If we pay only for the vitamin content of each food, paying 10 cents, 20 cents, 25 cents, and 50 cents, respectively, for units of the four vitamins, how much does a unit of each type of food cost? Compute in two ways the total cost of the food we ate.

[Partial Ans. \$3.75.]

- 13. In Example 3, by how much would store 1 have to reduce the price of apples to make Smith's total purchases less expensive there than at store 2?
- 14. In Example 3, find the store at which the total cost to Smith is the least when he wishes to purchase
 - (a) x = (4, 1, 2, 0, 1).

[Ans. Store 1; cost 47 cents.]

- **(b)** x = (1, 3, 2, 4, 0).
- (c) x = (2, 2, 2, 0, 2).
- 15. In Example 4, let us assume that an individual chooses ticket 1 with probability r_1 , ticket 2 with probability r_2 , and ticket 3 with

probability r_3 . Let $r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$. Give an interpretation for pMr. Compute this for $r_1 = \frac{1}{4}$, $r_2 = \frac{1}{4}$, and $r_3 = \frac{1}{2}$.

[Ans. $pMr = \frac{67}{36}$, which is the expected return.]

16. A game room contains three pinball machines. Either a game on one of these machines terminates normally, or else the machine refunds enough money for one or two games. A game also ends if the machine is tilted. The probability of each of these events is given by the following matrix:

	Normal	Refunds 1 Game	Refunds 2	Tilt
Machine 1	8.	.09	.01	.1 \
M = Machine 2	.75	.045	.005	.2
Machine 3	9.	.04	.01	.05 /

Assume that n = (25, 20, 30) represents the number of games played on machines 1, 2, and 3, respectively. Compute and interpret nM.

- In Exercise 16, Suppose it costs 10¢ to play one game.
 - (a) Construct a column vector with entries being the profit per game made by the owner of the machines for each of the different outcomes.
 - (b) What is the expected profit made by the owner if 100 people play machine 1, 80 play machine 2, and 120 play machine 3.
 - Due to space limitations the owner must sell one of the machines. Which one should he sell?
- (a) Consider the matrices 18.

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} \quad \text{and} \quad f = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Show that Pf = f. The vector f is called a fixed vector on the right of P.

- **(b)** Let $w = (\frac{9}{17}, \frac{8}{17})$ and let P be the matrix in part (a). Show that wP = w. For this reason w is called a fixed vector on the left of Р.
- 19. Let P be the matrix of transition probabilities for a Markov chain having n states, and let f be a column matrix all of whose entries

are 1's. Show that Pf = f. [Hint: Exercise 18 provides a special case.]

- 20. Let A, B, and C be matrices of the same shape, and let h and k be numbers. Use the ordinary rules for numbers plus the definitions of this section to show that the following laws hold:
 - (a1) A + B = B + A (commutative law of addition).
 - (a2) A + (B + C) = (A + B) + C (associative law of addition).
 - (a3) If 0 is the zero matrix of the same shape, then A + 0 = A (additive identity law).
 - (a4) Define -A = (-1)A; then A + (-A) = 0 (additive inverse law).
 - (s1) h(kA) = (hk)A (mixed associative law).
 - (s2) 1A = A for all A (unity law).
 - (s3) h(A + B) = hA + hB (first distributive law).
 - (s4) (h + k)A = hA + kA (second distributive law).
- 21. A company is considering which of three methods of production it should use in producing three goods, A, B, and C. The amount of each good produced by each method is shown in the matrix:

$$R = \begin{pmatrix} A & B & C \\ 2 & 3 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 1 \end{pmatrix} \begin{array}{l} \text{Method 1} \\ \text{Method 2} \\ \text{Method 3.} \end{array}$$

Let p be a vector whose components represent the profit per unit for each of the goods. What does the vector Rp represent? Find three different vectors p such that under each of these profit vectors a different method would be most profitable.

[Partial Ans. For
$$p = \begin{pmatrix} 10 \\ 8 \\ 7 \end{pmatrix}$$
 method 3 is most profitable.]

4 THE ADDITION AND MULTIPLICATION OF MATRICES

Two matrices of the same shape—that is, having the same number of rows and columns—can be added together by adding corresponding components. For example, if A and B are two 2×3 matrices, we have

$$A + B = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \end{pmatrix}.$$

Observe that the addition of vectors (row or column) is simply a special case of the addition of matrices. Numerical examples of the addition of matrices are

$$(1,0,-2) + (0,5,0) = (1,5,-2);$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

$$\begin{pmatrix} 7 & 0 & 0 \\ -3 & 1 & -6 \\ 4 & 0 & 7 \\ 0 & -2 & -2 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} -8 & 0 & 1 \\ 4 & 5 & -1 \\ 0 & 3 & 0 \\ -1 & 1 & -1 \\ 0 & -4 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 6 & -7 \\ 4 & 3 & 7 \\ -1 & -1 & -3 \\ 1 & -3 & 3 \end{pmatrix}.$$

Other examples occur in the exercises. The reader should observe that we do not add matrices of different shapes.

If A is a matrix and k is any number, we define the matrix kA as

$$kA = k \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}.$$

Observe that this is merely entrywise multiplication, as was the analogous concept for vectors. Examples of multiplication of matrices by constants are

$$-2\begin{pmatrix} 7 & -2 & 8 \\ 0 & 5 & -1 \end{pmatrix} = \begin{pmatrix} -14 & 4 & -16 \\ 0 & -10 & 2 \end{pmatrix};$$
$$6\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 6 \\ 18 & -24 \end{pmatrix}.$$

The multiplication of a vector by a number is a special case of the multiplication of a matrix by a number.

Under certain conditions two matrices can be multiplied together to give a new matrix. As an example, let A be a 2×3 matrix and B be a 3×2 matrix. Then the product AB is found to be

$$AB = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}.$$

Observe that the product is a 2×2 matrix. Also notice that each entry in the new matrix is the product of one of the rows of A times one of the columns of B; for example, the entry in the second row and first column is found as the product

$$(a_{21} \quad a_{22} \quad a_{23}) \begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}.$$

The following definition holds for the general case of matrix multiplication:

Definition Let A be an $m \times k$ matrix and B be a $k \times n$ matrix; then the product matrix C = AB is an $m \times n$ matrix whose components are

$$c_{ij} = (a_{i1} \ a_{i2} \ \cdots \ a_{ik}) \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{kj} \end{pmatrix}$$
$$= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}.$$

The important things to remember about this definition are: first, in order to be able to multiply matrix A times matrix B, the number of columns of A must be equal to the number of rows of B; second, the product matrix C = AB has the same number of rows as A and the same number of columns as B; finally, to get the entry in the ith row and jth column of AB we multiply the ith row of A times the jth column of B. Notice that the product of a vector times a matrix is a special case of matrix multiplication.

Below are several examples of matrix multiplication:

We next ask how we multiply more than two matrices together. Let A be an $m \times h$ matrix, let B be an $h \times k$ matrix, and let C be a $k \times n$ matrix. Then we can certainly define the products (AB)C and A(BC). It turns out that these two products are equal, and we define the product ABC to be their common value; that is,

$$ABC = A(BC) = (AB)C.$$

The rule expressed in the above equation is called the associative law for multiplication. We shall not prove the associative law here, although the student will be asked to check an example of it in Exercise 5.

If A and B are square matrices of the same size, then they can be multiplied in either order. It is not true, however, that the product AB is necessarily equal to the product BA. For example, if

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

then we have

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

whereas

$$BA = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

and it is clear that $AB \neq BA$.

EXERCISES

1. Perform the following matrix operations:

(a)
$$\begin{pmatrix} 3 & 2 \\ 1 & 4 \\ 5 & 3 \end{pmatrix} - \begin{pmatrix} 0 & 7 \\ -3 & -3 \\ 8 & 1 \end{pmatrix} = ?$$
 [Ans. $\begin{pmatrix} 3 & -5 \\ 4 & 7 \\ -3 & 2 \end{pmatrix}$.]

(b)
$$3\begin{pmatrix} 1 & 0 & 4 \\ 3 & 1 & -2 \\ 2 & 1 & -1 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 2 \\ 5 & 0 & 1 \\ 3 & 5 & 2 \end{pmatrix} = ?$$

(c)
$$\begin{pmatrix} 1 & 8 \\ -3 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} = ?$$

(d)
$$\begin{pmatrix} 0 & 10 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 8 & 1 & 3 \\ -6 & -1 & 2 \end{pmatrix} = ?$$

(e)
$$\begin{pmatrix} 8 & -6 \\ 1 & -1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ 10 & 6 \end{pmatrix} = ?$$

(f)
$$\begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix} \begin{bmatrix} \begin{pmatrix} 3 & 6 \\ 4 & 4 \end{pmatrix} - 8 \begin{pmatrix} 5 & 4 \\ 0 & 3 \end{pmatrix} \end{bmatrix} = ?$$

(g)
$$\begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = ?$$

(h)
$$\begin{pmatrix} 7 & 9 & 2 \\ 4 & 9 & 6 \\ 5 & 6 & 0 \end{pmatrix} \begin{pmatrix} 3 & 3 & 5 \\ 3 & 9 & 4 \\ 4 & 5 & 7 \end{pmatrix} = ?$$
 $[Ans. \begin{pmatrix} 56 & 112 & 85 \\ 63 & 123 & 98 \\ 33 & 69 & 49 \end{pmatrix}]$

(i)
$$\left[2\begin{pmatrix}6&0&1\\1&-3&2\end{pmatrix}+\begin{pmatrix}4&0&-4\\2&1&-1\end{pmatrix}\right]\begin{pmatrix}3&0&1\\-1&2&0\\0&0&2\end{pmatrix}=?$$

$$(\mathbf{j}) \quad \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = ?$$

2. Consider the matrices
$$A = \begin{pmatrix} -19 & 2 \\ 14 & -10 \\ 5 & 0 \end{pmatrix}$$
, $B = \begin{pmatrix} 7 & 2 & 19 \\ 4 & 26 & -2 \\ 13 & 0 & 7 \end{pmatrix}$,

$$C = \begin{pmatrix} 4 & 1 & -1 \\ 6 & 3 & -5 \end{pmatrix}$$
, and $D = \begin{pmatrix} 10 & -7 \\ 4 & 15 \end{pmatrix}$. Their shapes are 3×2 , $3 \times 3 \times 10^{-1}$

3, 2×3 , and 2×2 , respectively. What is the shape of:

[Ans. 3×3 .]

(b) CB?

(c)
$$DC$$
? [Ans. 2×3 .]

- (d) ACB?
- (e) BAC?
- (f) DCB?
- (g) DCBA?
- (h) ADCB?

3. In Exercise 2, find the component:

(a) In the second row and second column of AC. [Ans.
$$-16.$$
]

- (b) In the first row and the third column of CB.
- (c) In the last row and last column of AC.
- (d) In the last row and last column of CA. [Ans. -18.]
- (e) In the second row and first column of DC.
- 4. Let A be any 3×3 matrix and let I be the matrix

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Show that AI = IA = A. The matrix I acts for the products of matrices in the same way that the number 1 acts for products of numbers. For this reason it is called the *identity matrix*.

5. Verify the associative law for the special case when

$$A = \begin{pmatrix} -1 & 3 & 3 \\ 6 & 5 & 1 \end{pmatrix}, B = \begin{pmatrix} 4 & -3 & 3 \\ 8 & 0 & 5 \\ 5 & -4 & 3 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 9 & 4 \\ 7 & -4 \\ 0 & 1 \end{pmatrix}.$$

6. The commutative law for addition is

$$A + B = B + A$$

for any two matrices A and B of the same shape. Prove that the commutative law for addition is true from the definition of matrix addition and from the fact that it is true for ordinary numbers.

- 7. Show that there is *not* a commutative law for matrix multiplication by finding two 2×2 matrices A and B different than the ones in the text such that $A \cdot B \neq B \cdot A$.
- 8. The distributive law for numbers and matrices is

$$k(A+B) = kA + kB$$

for any number k and any two matrices A and B of the same shape. Prove that this law holds from the definitions of numerical multiplication of matrices, addition of matrices, and the ordinary rules for numbers.

9. The distributive laws for matrices are

$$(A+B)C = AC + BC,$$

$$C(A+B)=CA+CB,$$

where A, B, and C are matrices of suitable shapes. Show that these laws hold from the definitions of matrix multiplication and addition, and the ordinary rules for numbers.

10. Let A be any 3×3 matrix and let 0 be the matrix

$$0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Show that A0 = 0A = 0 for any A. Also show that A + 0 = 0 + A = A for any A. The matrix 0 acts for matrices in the same way that the number 0 acts for numbers. For this reason it is called the zero matrix.

- 11. Show that for any square matrix A there is a matrix B of the same shape as A such that A + B = 0. (B is called the *additive inverse* of A.)
- 12. Find the additive inverse of each of the following matrices:

(a)
$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$$
.

(b)
$$\begin{pmatrix} -3 & 2 \\ 6 & 1 \end{pmatrix}$$
.

(c)
$$\begin{pmatrix} -3 & 2 & 4 \\ 0 & 8 & 3 \\ 5 & 4 & -7 \end{pmatrix}$$
.

- 13. If $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, show that $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Thus the product of two matrices can be the zero matrix even though neither of the matrices is itself zero. Find another example that illustrates this point.
- 14. If A is a square matrix, it can be multiplied by itself; hence we can define (using the associative law)

$$A^{2} = A \cdot A$$

$$A^{3} = A^{2} \cdot A = A \cdot A \cdot A$$

$$A^{n} = A^{n-1} \cdot A = A \cdot A \cdot \dots \cdot A \quad (n \text{ factors}).$$

These are naturally called "powers" of the matrix A, A^2 being called the square of A, A^3 the cube of A, etc.

- (a) Compute A^2 , A^3 , and A^4 for $A = \begin{pmatrix} 2 & 0 \\ 3 & -1 \end{pmatrix}$.
- (b) If I and 0 are the matrices defined in Exercises 4 and 10, find I^2 , I^3 , I^n , 0^2 , 0^3 , and 0^n .
- (c) If $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, find A^n .

(d) If
$$A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$
, find A^2 , A^3 , and A^n .

- 15. Find the matrices $P^{(2)}$ and $P^{(3)}$ for the Markov chain whose transition matrix is $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$. Compute P^2 and P^3 and compare the results.
- 16. Cube the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}.$$

Compare your answer with the matrix $P^{(3)}$ in Example 1, Chapter 3, Section 12, and comment on the result.

17. Consider a two-state Markov process whose transition matrix is

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

- (a) Assuming that the process starts in state 1, draw the tree and set up tree measures for three stages of the process. Do the same, assuming that the process starts in state 2.
- (b) Using the trees drawn in (a), compute the quantities $p_{11}^{(3)}, p_{12}^{(3)}, p_{21}^{(3)}, p_{22}^{(3)}$. Write the matrix $P^{(3)}$.
- (c) Compute the cube $P^{(3)}$ of the matrix P.
- (d) Compare the answers you found in parts (b) and (c) and show that $P^{(3)} = P^3$.
- **18.** Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$.
 - (a) Find a matrix B such that $AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Show that $BA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ as well.
 - (b) Find a matrix C such that $AC = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$. What is CA?
- 19. A diagonal matrix is square and its only nonzero entries are on the main diagonal. For instance, the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

are 2×2 diagonal matrices.

- (a) Show that A and B commute, i.e., AB = BA.
- (b) Show that any pair of diagonal matrices of the same size commute when multiplied together.
- 20. A scalar matrix is a diagonal matrix in which all the entries on

$$B = \begin{pmatrix} -7 & 0 \\ 0 & -7 \end{pmatrix}$$
 are 2 × 2 scalar matrices.

- (a) Show that A commutes with any 2×2 matrix.
- (b) Show that any scalar matrix A can be written as kI, where k is a number and I is the identity matrix of the appropriate size.
- (c) Show that any scalar matrix commutes with any other matrix of the same size.
- 21. In Example 1 of Section 3 assume that the contractor wishes to take into account the cost of transporting raw materials to the building site as well as the purchasing cost. Suppose the costs are as given in the matrix below:

Purchase Transport
$$Q = \begin{pmatrix} 15 & 4.5 \\ 8 & 2 \\ 5 & 3 \\ 1 & 0.5 \\ 10 & 0 \end{pmatrix}$$
Steel Wood Glass Paint Labor

Referring to the example:

- (a) By computing the product RQ find a 3×2 matrix whose entries give the purchase and transportation costs of the materials for each kind of house.
- (b) Find the product xRQ, which is a two-component row vector whose first component gives the total purchase price and second component gives the total transportation cost.
- (c) Let $z = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and then compute xRQz, which is a number giving the total cost of materials and transportation for all the houses being built. [Ans. 14,304.]
- 22. A candy company packages four sizes of assorted chocolates: the Sampler, Sweetheart, Matinee, and Jumbo boxes. The Sampler contains 3 almond creams, 4 chocolate nougats, 5 caramel creams, and 3 nut clusters. The Sweetheart contains 6 almond creams, 4 chocolate nougats, 8 caramel creams, and 7 nut clusters. In the Matinee box are 10 almond creams, 15 chocolate nougats, 5 caramel creams, and 5 nut clusters. The Jumbo assortment has 10 almond creams, 15 chocolate nougats, 15 caramel creams, and 10 nut clusters. The company uses as ingredients in its manufacturing process chocolate, nuts, and cream filling. An almond cream contains 1 unit of chocolate, 2 units of nuts, and 2 units of cream filling; a chocolate nougat, 2 units each of chocolate and cream filling; a caramel cream, 4 units of cream filling; and a nut cluster, 3 units of nuts and 2 units of chocolate. Suppose a unit of chocolate costs 1¢, a unit of nuts 1.6¢, and a unit of cream filling 2¢.

If the company packages 50 Samplers, 75 Sweethearts, 40 Matinees, and 100 Jumbos,

- (a) What is the total number of each type of candy produced?
- (b) What is the total number of units of each ingredient used to make all the candy?
- (c) What is the total cost of the candy in each box?
- (d) What is the total cost of all the candy in all the boxes?

[Ans. \$651.50.]

5 THE SOLUTION OF LINEAR EQUATIONS

There are many occasions when the simultaneous solutions of linear equations is important. In this section we shall develop methods for finding out whether a set of linear equations has solutions, and for finding all such solutions.

EXAMPLE 1 Consider the following example of three linear equations in three unknowns:

$$(1) x_1 + 4x_2 + 3x_3 = 1$$

$$(2) 2x_1 + 5x_2 + 4x_3 = 4$$

$$(3) x_1 - 3x_2 - 2x_3 = 5.$$

Equations such as these, containing one or more variables, are called *open statements*. Statement (1) is true for some values of the variables (for instance, when $x_1 = 1$, $x_2 = 0$, and $x_3 = 0$), and false for other values of

the variables (for instance, when $x_1 = 0$, $x_2 = 1$, and $x_3 = 0$). The truth

set of (1) is the set of all vectors $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ for which (1) is true. Similarly, the

truth set of the three simultaneous equations (1), (2), and (3) is the set of

all vectors $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ which make true their conjunction

$$(x_1 + 4x_2 + 3x_3 = 1) \land (2x_1 + 5x_2 + 4x_3 = 4) \land (x_1 - 3x_2 - 2x_3 = 5).$$

When we say that we solve a set of simultaneous equations, we mean that we determine the truth set of their conjunction.

Before we discuss the solution of these equations we note that they can be written as a single equation in matrix form as follows:

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

One of the uses of vector and matrix notation is in writing a large number of linear equations in a single simple matrix equation such as the one above. It also leads to the detached coefficient form of solving simultaneous equa-

tions that we shall discuss at the end of the present section and in the next section.

The method of solving the linear equations above is the following. First we use equation (1) to eliminate the variable x_1 from equations (2) and (3); i.e., we subtract 2 times (1) from (2) and then subtract (1) from (3), giving

$$(1') x_1 + 4x_2 + 3x_3 = 1$$

$$(2') -3x_2 - 2x_3 = 2$$

$$(3') -7x_2 - 5x_3 = 4.$$

By pivoting we shall mean the operation of using an equation to eliminate a variable from the other equations. The pivot is the coefficient of the variable being eliminated. In this case the pivot is 1. Next we pivot on -3 in (2'): divide equation (2') through by the coefficient of x_2 , namely, -3, obtaining $x_2 + \frac{2}{3}x_3 = -\frac{2}{3}$. We use this equation to eliminate x_2 from each of the other two equations. In order to do this we subtract 4 times this equation from (1') and add 7 times this equation to (3'), obtaining

$$(1'') x_1 + 0 + \frac{1}{3}x_3 = \frac{11}{3}$$

$$(2'') x_2 + \frac{2}{3}x_3 = -\frac{2}{3}$$

$$(3'') -\frac{1}{3}x_3 = -\frac{2}{3}.$$

The last step is to pivot on $-\frac{1}{3}$ by dividing through (3") by $-\frac{1}{3}$, which is the coefficient of x_3 , obtaining the equation $x_3 = 2$; we use this equation to eliminate x_3 from the first two equations as follows:

$$(1''') x_1 + 0 + 0 = 3$$

$$(2''') x_2 + 0 = -2$$

$$(3''')$$
 $x_3 = 2.$

The solution can now be read from these equations as $x_1 = 3$, $x_2 = -2$, and $x_3 = 2$. The reader should substitute these values into the original equations (1), (2), and (3) above to see that the solution has actually been obtained.

In the example just discussed we saw that there was only one solution to the set of three simultaneous equations in three variables. Example 2 will be one in which there is *more* than one solution, and Example 3 will be one in which there are *no* solutions to a set of three simultaneous equations in three variables.

EXAMPLE 2 Consider the following linear equations:

$$(4) x_1 - 2x_2 - 3x_3 = 2$$

$$(5) x_1 - 4x_2 - 13x_3 = 14$$

$$(6) -3x_1 + 5x_2 + 4x_3 = 0.$$

Let us proceed as before and use equation (4) to eliminate the variable x_1 from the other two equations. Pivoting on the 1 coefficient of x_1 in (4), we have

$$(4') x_1 - 2x_2 - 3x_3 = 2$$

$$(5') -2x_2 - 10x_3 = 12$$

$$(6') -x_2 - 5x_3 = 6.$$

Proceeding as before, we divide equation (5') by -2, obtaining the equation $x_2 + 5x_3 = -6$. We use this equation to eliminate the variable x_2 from each of the other equations—namely, we add twice this equation to (4') and then add the equation to (6'):

$$(4'') x_1 + 0 + 7x_3 = -10$$

$$(5'') x_2 + 5x_3 = -6$$

$$(6'') 0 = 0.$$

Observe that we have eliminated the last equation completely! We also see that the variable x_3 can be chosen completely arbitrarily in these equations. To emphasize this, we move the terms involving x_3 to the right-hand side, giving

$$(4''') x_1 = -10 - 7x_3$$

$$(5''') x_2 = -6 - 5x_3.$$

The reader should check, by substituting these values of x_1 and x_2 into equations (4), (5), and (6), that they are solutions regardless of the value of x_3 . Let us also substitute particular values for x_3 to obtain numerical solutions. Thus, if we let $x_3 = 1$, 0, -2, respectively, and compute the resulting numbers, using (4"') and (5"'), we obtain the following numerical solutions:

$$x_1 = -17,$$
 $x_2 = -11,$ $x_3 = 1$
 $x_1 = -10,$ $x_2 = -6,$ $x_3 = 0$
 $x_1 = 4,$ $x_2 = 4,$ $x_3 = -2.$

The reader should also substitute these numbers into (4), (5), and (6) to show that they are solutions. To summarize, our second example has an infinite number of solutions, one for each numerical value of x_3 which is substituted into equations (4''') and (5''').

EXAMPLE 3 Suppose that we modify equation (6) by changing the number on the right-hand side to 2. Then we have

$$(7) x_1 - 2x_2 - 3x_3 = 2$$

$$(8) x_1 - 4x_2 - 13x_3 = 14$$

$$(9) -3x_1 + 5x_2 + 4x_3 = 2.$$

If we carry out the same procedure as before and use (7) to eliminate x_1 from (8) and (9), we obtain

$$(7') x_1 - 2x_2 - 3x_3 = 2$$

$$(8') -2x_2 - 10x_3 = 12$$

$$(9') -x_2 - 5x_3 = 8.$$

We divide (8') by -2, the coefficient of x_2 , obtaining, as before, $x_2 + 5x_3 = -6$. Using this equation to eliminate x_2 from the other two equations, we have

$$(7'') x_1 + 0 + 7x_3 = -10$$

$$(8'') x_2 + 5x_3 = -6$$

$$(9'')$$
 0 = 2.

Observe that the last equation is *logically false*, that is, false for all values of x_1, x_2, x_3 . Because our elimination procedure has led to a false result we conclude that the equations (7), (8), and (9) have *no* solution. The student should always keep in mind that this possibility exists when considering simultaneous equations.

In the examples above the equations we considered had the same number of variables as equations. The next example has more variables than equations and the last has more equations than variables.

EXAMPLE 4 Consider the following two equations in three variables:

$$(10) -4x_1 + 3x_2 + 2x_3 = -2$$

$$5x_1 - 4x_2 + x_3 = 3.$$

Using the elimination method outlined above, we divide (10) by -4, and then subtract 5 times the result from (11), obtaining

$$(10') x_1 - \frac{3}{4}x_2 - \frac{1}{2}x_3 = \frac{1}{2}$$

$$(11') -\frac{1}{4}x_2 + \frac{7}{2}x_3 = \frac{1}{2}.$$

Multiplying (11') by -4 and using it to eliminate x_2 from (10'), we have

$$(10'') x_1 + 0 - 11x_3 = -1$$

$$(11'') x_2 - 14x_3 = -2.$$

We can now let x_3 take on any value whatsoever and solve these equations for x_1 and x_2 . We emphasize this fact by rewriting them as in Example 2 as

$$(10''') x_1 = 11x_3 - 1$$

$$(11''') x_2 = 14x_3 - 2.$$

The reader should check that these are solutions and also, by choosing specific values for x_3 , find numerical solutions to these equations.

EXAMPLE 5 Let us consider the other possibility suggested by Example 4, namely, the case in which we have more equations than variables. Consider the following equations:

$$(12) -4x_1 + 3x_2 = 2$$

$$5x_1 - 4x_2 = 0$$

$$(14) 2x_1 - x_2 = a,$$

where a is an arbitrary number. Using equation (12) to eliminate x_1 from the other two we obtain

$$(12') x_1 - \frac{3}{4}x_2 = -\frac{1}{2}$$

$$-\frac{1}{4}x_2 = \frac{5}{2}$$

Next we use (13') to eliminate x_2 from the other equations, obtaining

$$(12'') x_1 + 0 = -8$$

$$(13'') x_2 = -10$$

$$(14'') 0 = a + 6.$$

These equations remind us of the situation in Example 3, since we shall be led to a false result unless a = -6. We see that equations (12), (13), and (14) have the solution $x_1 = -8$ and $x_2 = -10$ only if a = -6. If $a \neq -6$, then there is no solution to these equations.

The examples above illustrate all the possibilities that can occur in the general case. There may be no solutions, exactly one solution, or an infinite number of solutions to a set of simultaneous equations.

The procedure that we have illustrated above is one that turns any set of linear equations into an equivalent set of equations from which the existence of solutions and the solutions can be easily read. A student who learned other ways of solving linear equations may wonder why we use the above procedure—one which is not always the quickest way of solving equations. The answer is that we use it because it always works, that is, it is a canonical procedure to apply to any set of linear equations. The faster methods usually work only for equations that have solutions, and even then may not find all solutions.

The computational process illustrated above is summarized in the flow diagram of Figure 5. In that diagram the instructions encircled by dotted lines are either beginning or ending instructions; those enclosed in solid rectangles are intermediate computational steps; and those enclosed in ovals ask questions, the answers to which determine which of two paths the computational process will follow.

The direction of the process is always indicated by arrows. The flow diagram of Figure 5 can easily be turned into a computer program for

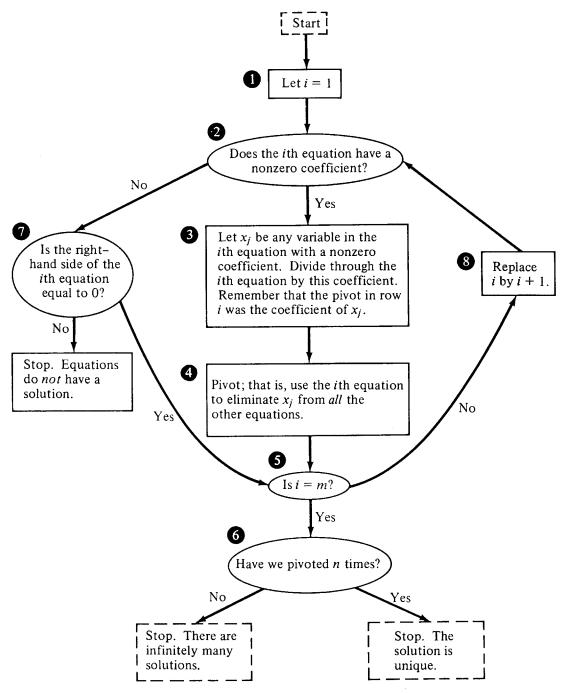


Figure 5

Flow diagram for solving m equations in n variables.

solving m linear equations in n variables. Students having access to a computer will find it a useful exercise to write such a program.

Let us return again to the equations of Example 1. Note that the variables, coefficients, and equals signs are in columns at the beginning of the solution and are always kept in the same column. It is obvious that the *location* of the coefficient is sufficient identification for it and that it is unnecessary to keep writing the variables. We can start with the format or *tableau*

$$\begin{pmatrix} 1 & 4 & 3 & 1 \\ 2 & 5 & 4 & 4 \\ 1 & -3 & -2 & 5 \end{pmatrix}.$$

Note that the coefficients of x_1 are found in the first column, the coefficients of x_2 in the second column, of x_3 in the third column, and the constants on the right-hand side of the equation all occur in the fourth column. The vertical line represents the equals signs in the equations.

The tableau of (15) will be called the detached coefficient tableau for simultaneous linear equations. We now show how to solve simultaneous equations using the detached coefficient tableau.

Starting with the tableau of (15) we carry out exactly the same calculations EXAMPLE 6 as in Example 1, which lead to the following series of tableaus:

$$\begin{pmatrix}
1 & 4 & 3 & | & 1 \\
0 & -3 & -2 & | & 2 \\
0 & -7 & -5 & | & 4
\end{pmatrix}$$

(16)
$$\begin{pmatrix} 1 & 4 & 3 & 1 \\ 0 & -3 & -2 & 2 \\ 0 & -7 & -5 & 4 \end{pmatrix}$$
(17)
$$\begin{pmatrix} 1 & 0 & \frac{1}{3} & \frac{11}{3} \\ 0 & 1 & \frac{2}{3} & -\frac{2}{3} \\ 0 & 0 & -\frac{1}{3} & -\frac{2}{3} \end{pmatrix}$$

(18)
$$\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

From the tableau of (18) we can easily read the answer $x_1 = 3$, $x_2 = -2$, and $x_3 = 2$, which is the same as before.

The correspondence between the calculations of Example 1 and of the present example is as follows:

Note that in the tableau form we are always careful to keep zero coefficients in each column when necessary.

Suppose that we have two sets of simultaneous equations to solve and that EXAMPLE 7 they differ only in their right-hand sides. For instance, suppose we want to solve

(19)
$$\begin{pmatrix} 1 & 4 & 3 \\ 2 & 5 & 4 \\ 1 & -3 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} \text{ and } = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}.$$

It is obvious that the calculations on the left-hand side will be the same regardless of the numbers appearing on the right-hand side. Therefore it is possible to solve both sets of simultaneous equations at once. We shall illustrate this in the following series of tableaus:

(20)
$$\begin{pmatrix} 1 & 4 & 3 & 1 & -1 \\ 2 & 5 & 4 & 4 & 0 \\ 1 & -3 & -2 & 5 & 2 \end{pmatrix}$$

$$\begin{pmatrix}
1 & 4 & 3 & | & 1 & -1 \\
0 & -3 & -2 & | & 2 & 2 \\
0 & -7 & -5 & | & 4 & 3
\end{pmatrix}$$

(22)
$$\begin{pmatrix} 1 & 0 & \frac{1}{3} & \frac{11}{3} & \frac{5}{3} \\ 0 & 1 & \frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} \\ 0 & 0 & -\frac{1}{3} & -\frac{2}{3} & -\frac{5}{3} \end{pmatrix}$$

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

We find the answers

$$x_1 = 3, \quad x_2 = -2, \quad x_3 = 2$$

to the first set of equations and the answers

$$x_1 = 0, \qquad x_2 = -4, \qquad x_3 = 5$$

to the second set of equations. The reader should check these answers by substituting into the original equations.

EXERCISES

1. Find all solutions to the following simultaneous equations:

(a)
$$x_1 + 2x_2 + 3x_3 = 4$$

 $4x_1 + 5x_2 + 6x_3 = -2$
 $7x_1 + 8x_2 + 27x_3 = 10$. [Ans. $x_1 = -7, x_2 = 4, x_3 = 1$.]

(b)
$$-x_1 + 3x_3 = 16$$

 $2x_1 + 7x_2 + 3x_3 = -6$
 $-x_1 + 7x_2 + 12x_3 = 41$. [Ans. No solution.]

(c)
$$-x_1 + 3x_3 = 8$$

 $2x_1 + 7x_2 + 3x_3 = -3$
 $-x_1 + 7x_2 + 12x_3 = 21$.

[Ans. $x_1 = 3x_3 - 8$, $x_2 = \frac{1}{7}(13 - 9x_3)$, $x_3 =$ any real number.]

2. Find all solutions of the following simultaneous equations:

(a)
$$x_1 + 5x_2 + 3x_3 = 2$$

 $x_1 + x_2 + x_3 = -2$
 $13x_1 + 4x_2 + 7x_3 = 4$

(b)
$$5x_1 + 3x_2 - x_3 = -2$$

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

(c)
$$2x_1 + 3x_2 - x_3 = 0$$

 $-4x_2 + x_3 = 0$
 $8x_1 - x_3 = 0$.

- 3. Rework Examples 2-4 using the detached coefficient tableau.
- **4.** Find all solutions of the following equations using all the detached coefficient tableau:

(a)
$$2x_1 + 2x_2 - x_3 = -4$$

 $-x_2 + 2x_3 = -2$
 $\frac{x_1}{2} + x_2 + x_3 = 1$.

(b)
$$5x_1 + 3x_2 + 3x_3 = 7$$
$$-2x_1 + 4x_2 + 8x_3 = -3$$
$$-x_1 + 3x_2 + 4x_3 = 2.$$

(c)
$$3x_1 + 6x_2 - x_3 = -7$$

 $x_1 + 5x_2 = 0$
 $2x_2 + x_3 = 14$.

5. Find all solutions of the following equations:

(a)
$$-x_1 + x_2 - x_3 + x_4 = 13$$
$$2x_2 - 6x_3 + 7x_4 = 0$$
$$-2x_1 + x_2 + x_3 + 5x_4 = -13$$
$$3x_1 - 2x_2 + 2x_3 + x_4 = 26.$$
(b)
$$x_1 + 2x_2 + 3x_3 + 4x_4 = 10$$
$$2x_1 - x_2 + x_3 - x_4 = 1$$
$$3x_1 + x_2 + 4x_3 + 3x_4 = 11$$
$$-2x_1 + 6x_2 + 4x_3 + 10x_4 = 18.$$

6. Solve the following four simultaneous sets whose right-hand sides are listed under (a), (b), (c), and (d) below. Use the detached coefficient tableau.

[Ans. (a)
$$x_1 = 30$$
, $x_2 = 11$, $x_3 = -40$.]

7. Solve the following four sets of simultaneous equations, which differ only in their right-hand sides:

[Ans. (d)
$$x_1 = x_2 = x_3 = 1$$
.]

8. Solve the following four sets of simultaneous equations:

- 9. A man is ordered by his doctor to take 10 units of vitamin A, 9 units of vitamin D, and 19 units of vitamin E each day. The man can choose from three brands of vitamin pills. Brand X contains two units of vitamin A, three units of vitamin D, and five units of vitamin E; brand Y has 1, 3, and 4 units, respectively; and brand Z has 1 unit of vitamin A, 1 of vitamin E, and none of vitamin D.
 - (a) Find all possible combinations of pills that will provide exactly the required amounts of vitamins.
 - (b) If brand X costs 1¢ a pill, brand Y 6¢, and brand Z 3¢, is there a solution costing exactly 15¢ a day?
 - (c) What is the least expensive solution? The most expensive?
- 10. Show that the equations

$$4x_1 - 4x_2 + ax_3 = c$$
$$3x_1 - 2x_2 + bx_3 = d$$

always have a solution for all values of a, b, c, and d.

11. Find conditions on a, b, and c in order that the equations

$$3x_1 + 2x_2 = a$$

$$-4x_1 + x_2 = b$$

$$2x_1 + 5x_2 = c$$

have a solution.

12. For what value of the constant k does the following system have a unique solution? Find the solution in this case. What is the case if k does not take on this value?

$$2x + 4z = 6$$

$$3x + y + z = -1$$

$$2y - z = -2$$

$$x - y + kz = -5$$

[Ans.
$$k = -2$$
; $x = -1$, $y = 0$, $z = 2$; no solution.]

13. Let $x = (x_1, x_2, x_3)$, let $A = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, and let d be any number. What can

you say about the truth set of the statement xA = d in the following cases:

- (a) $A \neq 0$?
- **(b)** A = 0, d = 0?
- (c) $A = 0, d \neq 0$?

What can you conclude?

- **14.** Let *P* be the matrix $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$, and let $x = (x_1, x_2)$.
 - (a) Find all solutions of the equation xP = x.
 - (b) Choose the solution(s) for which $x_1 + x_2 = 1$.

- **15.** Let P be the matrix $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$, and let $x = (x_1, x_2, x_3)$.
 - (a) Find all solutions of the equation xP = x.
 - (b) Choose the solution(s) for which $x_1 + x_2 + x_3 = 1$.
- 16. Let P be as in Exercise 15, and let $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Redo the exercise, using the equation Px = x.
- 17. Let x be as in Exercise 15, and let A be the matrix

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

- (a) Find all solutions of the equation xA = x.
- (b) Choose the solution(s) for which $x_1 + x_2 + x_3 = 1$.

[Ans.
$$x_1 = \frac{15}{11}, x_2 = -\frac{9}{11}, x_3 = \frac{5}{11}$$
.]

18. (a) Show that the simultaneous linear equations

$$3x_1 - 5x_2 + 3x_3 = 9$$
$$4x_1 + 4x_2 - 7x_3 = 0$$

can be interpreted as a single-matrix-times-column-vector equation of the form

$$\begin{pmatrix} 3 & -5 & 3 \\ 4 & 4 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 9 \\ 0 \end{pmatrix}.$$

- (b) Show that any set of simultaneous linear equations may be interpreted as a matrix equation of the form Ax = b, where A is an $m \times n$ matrix, x is an n-component column vector, and b is an m-component column vector.
- 19. (a) Show that the equations of Exercise 18(a) can be interpreted as a row-vector-times-matrix equation of the form

$$(x_1, x_2, x_3)$$
 $\begin{pmatrix} 3 & 4 \\ -5 & 4 \\ 3 & -7 \end{pmatrix} = (9, 0).$

- (b) Show that any set of simultaneous linear equations may be interpreted as a matrix equation of the form xA = b, where A is an $m \times n$ matrix, x is an m-component row vector, and b is an n-component row vector.
- 20. (a) Show that the simultaneous linear equations of Exercise 18(a) can be interpreted as asking for all possible ways of expressing the column vector $\begin{pmatrix} 9 \\ 0 \end{pmatrix}$ in terms of the column vectors $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $\begin{pmatrix} -5 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 7 \end{pmatrix}$.

- (b) Show that any set of linear equations may be interpreted as asking for all possible ways of expressing a column vector in terms of given column vectors.
- 21. Redo Exercise 20, using Exercise 19(a) and row vectors.
- 22. Consider the following set of simultaneous equations:

$$x_1 - x_2 = a$$

 $x_3 + x_4 = b$
 $x_2 - x_4 = c$
 $x_1 + x_3 = d$.

- (a) For what conditions on a, b, c, and d will these equations have a solution?
- (b) Give a set of balues for a, b, c, and d for which the equations do not have a solution.
- (c) Show that if there is one solution to these equations, then there are infinitely many solutions.
- 23. Which of the following statements are true and which false concerning the solution of m simultaneous linear equations in n unknowns written in the form Ax = b?
 - (a) If there are infinitely many solutions, then n > m.
 - (b) If the solution is unique, then n = m.
 - (c) If m = n, then the solution is unique.
 - (d) If n > m, then there cannot be a unique solution.
 - (e) If b = 0, then there is always at least one solution.
 - (f) If b = 0, then there are always infinitely many solutions.
 - (g) If b = 0, and $x^{(1)}$ and $x^{(2)}$ are solutions, then $x^{(1)} + x^{(2)}$ is also a solution. [Ans. (d), (e), and (g) are true.]

6 THE INVERSE OF A SQUARE MATRIX

If A is a square matrix and B is another square matrix of the same size having the property that BA = I (where I is the identity matrix), then we say that B is the *inverse* of A. When it exists, we shall denote the inverse of A by the symbol A^{-1} . To give a numerical example, let A and A^{-1} be the following:

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

(2)
$$A^{-1} = \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix}.$$

Then we have

$$A^{-1}A = \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

If we multiply these matrices in the other order, we also get the identity matrix; thus

$$AA^{-1} = \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

In general it can be shown that if A is a square matrix with inverse A^{-1} , then the inverse satisfies the equation

$$A^{-1}A = AA^{-1} = I$$
.

Next we show that a square matrix can have only one inverse. For suppose that in addition to A^{-1} we also have a B such that

$$BA = I$$
.

Then we see that

$$B = BI = B(AA^{-1}) = (BA)A^{-1} = IA^{-1} = A^{-1}.$$

Finding the inverse of a matrix is analogous to finding the reciprocal of an ordinary number, but the analogy is not complete. Every nonzero number has a reciprocal, but there are matrices, not the zero matrix, which have no inverse. For example, if

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

then

$$AB = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

From this we shall show that neither A nor B can have an inverse. To show that A does not have an inverse, let us assume that A had an inverse A^{-1} . Then

$$B = (A^{-1}A)B = A^{-1}(AB) = A^{-1}0 = 0,$$

contradicting the fact that $B \neq 0$. The proof that B cannot have an inverse is similar.

Let us now try to calculate the inverse of the matrix A in (1). Specifically, let's try to calculate the first column of A^{-1} . Let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

be the desired entries of the first column. Then from the equation $AA^{-1} = I$ we see that we must solve

$$\begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Similarly, to find the second and third columns of A^{-1} we want to solve the additional sets of equations,

$$\begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

respectively. We thus have three sets of simultaneous equations that differ only in their right-hand sides. This is exactly the situation described in Example 7 of the previous section.

To solve them, we start with the tableau

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

and carry out the calculations as described in the last section. This gives rise to the following series of tableaus. In (3) divide the first row by 4, copy the second row, and subtract 3 times the new first row from the old third row, which yields the tableau

(4)
$$\begin{pmatrix} 1 & 0 & \frac{5}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{3}{4} & 0 & 1 \end{pmatrix}.$$

Next we multiply the third row of (4) by 4, multiply the new third row by 6 and add to the old second row, and multiply the new third row by $\frac{5}{4}$ and subtract from the old first row. We have the final tableau:

(5)
$$\begin{pmatrix} 1 & 0 & 0 & 4 & 0 & -5 \\ 0 & 1 & 0 & -18 & 1 & 24 \\ 0 & 0 & 1 & -3 & 0 & 4 \end{pmatrix}.$$

We see that the inverse A^{-1} which is given in (2) appears to the right of the vertical line in the tableau of (5).

The procedure just illustrated will find the inverse of any square matrix A, providing A has an inverse. We summarize it as follows:

Rule for Inverting a Matrix Let A be a matrix that has an inverse. To find the inverse of A start with the tableau

and change it by row transformations (as described in Section 5) into the tableau

The resulting matrix B is the inverse A^{-1} of A.

Even if A has no inverse, the procedure just outlined can be started. At some point in the procedure a tableau will be found that is not of the desired

final form and from which it is impossible to change by row transformations of the kind described.

EXAMPLE 1 Show that the matrix

$$A = \begin{pmatrix} 4 & 0 & 8 \\ 0 & 1 & -6 \\ 2 & 0 & 4 \end{pmatrix}$$

has no inverse.

We set up the initial tableau as follows:

(6)
$$\begin{pmatrix} 4 & 0 & 8 & 1 & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 2 & 0 & 4 & 0 & 0 & 1 \end{pmatrix}.$$

Carrying out one set of row transformations, we obtain the second tableau as follows:

(7)
$$\begin{pmatrix} 1 & 0 & 2 & \frac{1}{4} & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 1 \end{pmatrix}.$$

It is clear that we cannot proceed further since there is a row of zeros to the left of the equals sign on the third set of equations. Hence we conclude that A has no inverse.

Because of the form of the final tableau in (7), we see that it is impossible to solve the equations

$$\begin{pmatrix} 4 & 0 & 8 \\ 0 & 1 & -6 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

since these equations are inconsistent as is shown by the tests developed in Section 5. In other words, it is not possible to solve for the third column of the inverse matrix.

It is clear that an $n \times n$ matrix A has an inverse if and only if the following sets of simultaneous equations—

$$Ax = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad Ax = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \qquad \dots, \qquad Ax = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

—can all be uniquely solved. And these sets of simultaneous equations, since they all share the same left-hand sides, can be solved uniquely if and only if the transformation of the rule for inverting a matrix can be carried out. Hence we have proved the following theorem.

Theorem A square matrix A has an inverse if and only if the tableau

can be transformed by row transformations into the tableau

$$(I|A^{-1}).$$

EXAMPLE 2 Let us find the inverse of the matrix

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

The initial tableau is

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

Transforming it by row transformations, we obtain the following series of tableaus:

$$\begin{pmatrix}
1 & 4 & 3 & 1 & 0 & 0 \\
0 & -3 & -2 & -2 & 1 & 0 \\
0 & -7 & -5 & -1 & 0 & 1
\end{pmatrix}$$

$$\begin{pmatrix}
1 & 0 & \frac{1}{3} & -\frac{5}{3} & \frac{4}{3} & 0 \\
0 & 1 & \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\
0 & 0 & -\frac{1}{3} & \frac{11}{3} & -\frac{7}{3} & 1
\end{pmatrix}$$

$$\begin{pmatrix}
1 & 0 & 0 & 2 & -1 & 1 \\
0 & 1 & 0 & 8 & -5 & 2 \\
0 & 0 & 1 & -11 & 7 & -3
\end{pmatrix}.$$

The inverse of A is then

$$A^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ 8 & -5 & 2 \\ -11 & 7 & -3 \end{pmatrix}.$$

The reader should check that $A^{-1}A = AA^{-1} = I$.

EXAMPLE 3 A cookie recipe requires 4 cups of sugar and 2 cups of flour while a cake recipe needs 3 cups of sugar and 4 cups of flour. If we have 40 cups of sugar and 30 cups of flour on hand, how many recipes of each can we make? In order to answer this question let x be the number of cookie recipes and y the number of cake recipes to be made. Then the sugar and flour requirements give rise to the following two equations:

$$4x + 3y = 40$$
 (Sugar equation)
 $2x + 4y = 30$ (Flour equation).

Let us rewrite these equations in matrix form as

$$\begin{pmatrix} 4 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 40 \\ 30 \end{pmatrix}.$$

If we can invert the matrix we can solve the problem as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 2 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 40 \\ 30 \end{pmatrix}.$$

The initial tableau of the matrix inversion problem is

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

Pivoting on the 4 in the upper left-hand corner gives

$$\begin{pmatrix} 1 & \frac{3}{4} & \frac{1}{4} & 0 \\ 0 & \frac{5}{2} & -\frac{1}{2} & 1 \end{pmatrix}.$$

Finally, pivoting on the 5/2 term we obtain

$$\begin{pmatrix} 1 & 0 & \frac{2}{5} & -\frac{3}{10} \\ 0 & 1 & -\frac{1}{5} & \frac{2}{5} \end{pmatrix},$$

and so the inverse of the original matrix appears on the right. The solution to our problem is, then,

$$\binom{x}{y} = \binom{\frac{2}{5}}{-\frac{1}{5}} - \frac{3}{10} \binom{40}{30} = \binom{7}{4}.$$

In other words, we can make 7 batches of cookies and 4 cakes from the materials we have.

EXERCISES

1. Compute the inverse of each of the following matrices:

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} -5 & 1 & 10 \\ 9 & -2 & -17 \\ -4 & 1 & 8 \end{pmatrix},$$

$$C = \begin{pmatrix} 3 & 1 & 0 \\ 2 & 2 & 2 \\ 11 & 3 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & -7 & 3 & 2 \\ 0 & -1 & 7 & 3 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$[Partial \ ans. \ A^{-1} = \begin{pmatrix} .4 & .2 \\ .2 & .6 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ -4 & 0 & 5 \\ 1 & 1 & 1 \end{pmatrix}.]$$

2. Let B and C be the matrices of Exercise 1; let $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and $y = (y_1, y_2, y_3)$; let a, b, c, d, and e be the following vectors:

$$a = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ -1 \\ 3 \end{pmatrix}, \quad c = (1, 5, 3),$$

$$d = (1, 1, 1), \text{ and } e = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Uses the inverses computed in Exercise 1 to solve the following equations:

- (a) Bx = a.
- **(b)** yB = d.
- (c) Cx = e.
- (d) Bx = b.
- (e) yB = c.
- (f) yC = c.
- 3. Show that each of the following matrices fails to have an inverse.

$$A = \begin{pmatrix} 4 & 2 \\ 6 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 1 & 0 \\ 8 & 2 & 1 \\ 11 & 3 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} -3 & -1 & 1 \\ 4 & 3 & -2 \\ -7 & 1 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & -3 & 5 & -2 \\ 0 & 3 & -4 & 2 \\ 2 & 0 & 1 & -1 \end{pmatrix}.$$

- 4. For each of the matrices in Exercise 3 find a nonzero vector whose product with the given matrix is 0.
- 5. Solve the following four sets of simultaneous equations by first writing them in the form Ax = B, where B is a 3×4 matrix, and finding the inverse of A.

[Ans. (a)
$$x_1 = -11$$
, $x_2 = 56$, $x_3 = 9$.]

- 6. Let A be a square matrix. Show that if A has no inverse, then neither do any of its positive powers A^k . Show that if A has an inverse, then the inverse of A^2 is $(A^{-1})^2$. What is the inverse of A^3 ? Of A^n ?
- 7. The formula $(A^{-1})^{-1} = A$ states that if A has an inverse A^{-1} , then A^{-1} itself has an inverse and this inverse is A. Prove both parts of this statement.
- **8.** Expand the formula $(AB)^{-1} = B^{-1}A^{-1}$ into a two-part statement analogous to the one in Exercise 7. Then prove both parts of your statement.
- 9. Give a criterion for deciding whether the 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has an inverse. [Ans. $ad \neq bc$.]

- 10. Give a formula for $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$, when it exists.
- 11. If $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has an inverse and has integer components, what condition must it fulfill in order that $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$ have integer components?
- 12. Let A be the matrix $\begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$.
 - (a) Use Exercise 10 to find A^{-1} .
 - (b) Use the result of (a) to solve the equations Ax = b and $A^2x = c$, where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, and $c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
- 13. (a) Show that $(AB)^{-1} \neq A^{-1}B^{-1}$ for the matrices $A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}$.
 - **(b)** Find $(AB)^{-1}$ in two ways. [Hint: Use Exercises 10 and 8.]
- 14. Solve the following problems by first inverting the matrix involved.
 - (a) An automobile factory produces two models. The first requires 1 man-hour to paint and ½ man-hour to polish; the second requires 1 man-hour for each process. During each hour that the assembly line is operating, there are 100 man-hours available for painting and 80 man-hours for polishing. How many of each model can be produced each hour if all the man-hours available are to be utilized?
 - (b) Suppose each car of the first type requires 10 widgets and 14 shims, and each car of the second type requires 7 widgets and 10 shims. The factory can obtain 800 widgets and 1130 shims each hour. How many cars of each model can it produce while utilizing all the parts available?

 [Ans. 45, 50.]
- 15. Solve the following problem by first inverting the matrix. (Assume $ad \neq bc$.) If a grinding machine is supplied x pounds of meat and y pounds of scraps (meat scraps and fat) per day, then it will produce ax + by pounds of ground meat and cx + dy pounds of hamburger per day. In other words, its production vector is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$
.

What inputs are necessary in order to get 25 pounds of ground meat and 70 pounds of hamburger? In order to get 20 pounds of ground meat and 100 pounds of hamburger?

16. A square matrix is *lower-triangular* if it has zeros on and above its main diagonal. For instance, $Q = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 4 & 3 & 0 \end{pmatrix}$ is lower-triangular.

- (a) Compute Q^2 .
- (b) Compute Q^3 .
- (c) Show that $Q^k = 0$ for $k \ge 3$.
- 17. Let Q be as in Exercise 16.
 - (a) Show that $(I Q)(I + Q + Q^2) = I Q^3 = I$.
 - (b) Show that, because of (a), $I + Q + Q^2 = (I Q)^{-1}$.
 - (c) Use (b) to compute $(I Q)^{-1}$.
 - (d) Let $w = (w_1, w_2, w_3)$, d = (-1, 5, 3). Use (c) to solve the equation w = wQ + d.
- 18. (a) Show that the sum of any two lower-triangular matrices is lower-triangular.
 - (b) Show that the product of any two lower-triangular matrices is lower-triangular.
- 19. Let Q be an $n \times n$ lower-triangular matrix.
 - (a) Show that $Q^k = 0$ for $k \ge n$.
 - (b) Show that $(I Q)(I + Q + \cdots + Q^{n-1}) = I Q^n = I$.
 - (c) Show that $(I Q)^{-1} = I + Q + \cdots + Q^{n-1}$.
 - (d) Show that all entries above the main diagonal of $(I Q)^{-1}$ are 0.
 - (e) Show that if Q has nonnegative *integer* entries, then so does $(I Q)^{-1}$.
- **20.** Find $(I Q)^{-1}$ for each of the following:

(a)
$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$
.

(b)
$$\begin{pmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \\ -1 & 2 & 0 \end{pmatrix}.$$

(c)
$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 \\ 4 & -1 & 0 & 0 \\ -2 & 1 & 3 & 0 \end{pmatrix}$$

(d)
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & 5 & 4 & 0 & 0 \\ -2 & 1 & 2 & -3 & 0 \end{pmatrix} .$$

7 APPLICATIONS OF MATRIX THEORY TO MARKOV CHAINS

In this section we shall show applications of matrix theory to Markov chains. For simplicity we shall confine our discussion to three-state Markov chains, but a similar procedure will work for any other Markov chain.

In Section 12 of Chapter 3, we noted that to each Markov chain there was a matrix of transition probabilities. For example, if there are three

states, a_1 , a_2 , and a_3 , then

$$P = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_1 & p_{11} & p_{12} & p_{13} \\ a_2 & p_{21} & p_{22} & p_{23} \\ a_3 & p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Is the transition matrix for the chain. Recall that the *row sums* of P are all equal to 1. Such a matrix is called a transition matrix.

Definition A transition matrix is a square matrix with nonnegative entries such that the sum of the entries in each row is 1.

In order to obtain a Markov chain we must specify how the process starts. Suppose that the initial state is chosen by a chance device that selects state a_j with probability $p_j^{(0)}$. We can represent these initial probabilities by means of the vector $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)})$. As in Exercise 17 of Section 4, we can construct a tree measure for as many steps of the process as we wish to consider. Let $p_j^{(n)}$ be the probability that the process will be in state a_j after n steps. Let the vector of these probabilities be $p^{(n)} = (p_1^{(n)}, p_2^{(n)}, p_3^{(n)})$.

Definition A row vector p is called a probability vector if it has nonnegative components whose sum is 1.

Obviously the vectors $p^{(0)}$ and $p^{(n)}$ are probability vectors. Also each row of a transition matrix is a probability vector.

By means of the tree measure it can be shown that these probabilities satisfy the following equations:

$$\begin{split} p_1^{(n)} &= p_1^{(n-1)} p_{11} + p_2^{(n-1)} p_{21} + p_3^{(n-1)} p_{31}, \\ p_2^{(n)} &= p_1^{(n-1)} p_{12} + p_2^{(n-1)} p_{22} + p_3^{(n-1)} p_{32}, \\ p_3^{(n)} &= p_1^{(n-1)} p_{13} + p_2^{(n-1)} p_{23} + p_3^{(n-1)} p_{33}. \end{split}$$

It is not hard to give intuitive meanings to these equations. The first one, for example, expresses the fact that the probability of being in state a_1 after n steps is the sum of the probabilities of being at each of the three possible states after n-1 steps and then moving to state a_1 on the nth step. The interpretation of the other equations is similar.

If we recall the definition of the product of a vector times a matrix, we can write the equations above as

$$p^{(n)}=p^{(n-1)}P.$$

If we substitute values of n, we get the equations: $p^{(1)} = p^{(0)}P$; $p^{(2)} = p^{(1)}P = p^{(0)}P^2$; $p^{(3)} = p^{(2)}P = p^{(0)}P^3$; and so on. In general, it is evident that

$$p^{(n)} = p^{(0)}P^n.$$

Thus we see that, if we multiply the vector $p^{(0)}$ of initial probabilities by the *n*th power of the transition matrix P, we obtain the vector $p^{(n)}$, whose components give the probabilities of being in each of the states after n steps.

In particular, let us choose $p^{(0)} = (1, 0, 0)$, which is equivalent to letting the process start in state a_1 . From the equation above we see that then $p^{(n)}$ is the first row of the matrix P^n . Thus the elements of the first row of the matrix P^n give us the probabilities that after n steps the process will be in a given one of the states, under the assumption that it started in state a_1 . In the same way, if we choose $p^{(0)} = (0, 1, 0)$, we see that the second row of P^n gives the probabilities that the process will be in one of the various states after n steps, given that it started in state a_2 . Similarly the third row gives these probabilities, assuming that the process started in state a_3 .

In Section 12 of Chapter 3, we considered special Markov chains that started in given fixed states. There we arrived at a matrix $P^{(n)}$ whose *i*th row gave the probabilities of the process ending in the various states, given that it started at state a_i . By comparing the work that we did there with what we have just done, we see that the matrix $P^{(n)}$ is merely the *n*th power of P, that is, $P^{(n)} = P^n$. (Compare Exercise 17 of Section 4.) Matrix multiplication thus gives a convenient way of computing the desired probabilities.

Definition The probability vector w is a *fixed point* of the matrix P, if w = wP.

EXAMPLE 1 Consider the transition matrix

$$P = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} .667 & .333 \\ .500 & .500 \end{pmatrix}.$$

If w = (.6, .4), then we see that

$$wP = (.6, .4)\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (.6, .4) = w,$$

so that w is the fixed point of the matrix P.

If we had happened to choose the vector w as our initial probability vector $p^{(0)}$, we would have had $p^{(n)} = p^{(0)}P^n = wP^n = w = p^{(0)}$. In this case the probability of being at any particular state is the same at all steps of the process. Such a process is in *equilibrium*.

As seen above, in the study of Markov chains we are interested in the powers of the matrix P. To see what happens to these powers, let us further consider the example.

EXAMPLE 1 Suppose that we compute powers of the matrix P in the example above. (continued) We have

$$P^2 = \begin{pmatrix} .611 & .389 \\ .583 & .417 \end{pmatrix}, \qquad P^3 = \begin{pmatrix} .602 & .398 \\ .597 & .403 \end{pmatrix}, \quad \text{and so on.}$$

It looks as if the matrix P^n is approaching the matrix

$$W = \begin{pmatrix} .6 & .4 \\ .6 & .4 \end{pmatrix};$$

206 Vectors and Matrices Chapter 4

and, in fact, it can be shown that this is the case. (When we say that P^n approaches W we mean that each entry in the matrix P^n gets close to the corresponding entry in W.) Note that each row of W is the fixed point W of the matrix P.

Definition A transition matrix is said to be *regular* if some power of the matrix has only positive components.

Thus the matrix in the example is regular, since every entry in it is positive, so that the first power of the matrix has all positive entries. Other examples occur in the exercises.

Theorem If P is a regular transition matrix, then

- i. the powers P^n approach a matrix W;
- ii. each row of W is the same probability vector w;
- iii. the components of w are positive.

We omit the proof of this theorem;* however, we can prove the next theorem.

Theorem If P is a regular transition matrix, and W and w are given by the previous theorem, then

- a. if p is any probability vector, pP^n approaches w;
- b. the vector w is the unique fixed-point probability vector of P.

Proof First let us consider the vector pW. The first column of W has a w_1 in each row. Hence in the first component of pW each component of p is multiplied by w_1 , and therefore we have w_1 times the sum of the components of p, which is w_1 . Doing the same for the other components, we note that pW is simply w. But pP^n approaches pW; hence it approaches w. Thus if any probability vector is multiplied repeatedly by P, it approaches the fixed point w. This proves part (a).

Since the powers of P approach W, $P^{n+1} = P^n P$ approaches W, but it also approaches WP; hence WP = W. Any one row of this matrix equation states that wP = w; hence w is a fixed point (and by the previous theorem a probability vector). We must still show that it is unique. Let u be any probability-vector fixed point of P. By part a of the theorem we know that uP^n approaches w. But since u is a fixed point, $uP^n = u$. Hence u remains fixed but "approaches" w. This is possible only if u = w. Hence w is the only probability-vector fixed point. This completes the proof of part b.

^{*}For an elementary proof see John G. Kemeny and J. Laurie Snell, Finite Markov Chains, Princeton, N.J., Van Nostrand, 1960.

The following is an important consequence of the above theorem. If we take as p the vector $p^{(0)}$ of initial probabilities, then the vector $pP^n = p^{(n)}$ gives the probabilities after n steps, and this vector approaches w. Therefore no matter what the initial probabilities are, if P is regular, then after a large number of steps the probability that the process is in state a_j will be very nearly w_j .

We noted for an independent trials process that if p is the probability of a given outcome a, then this may be given an alternate interpretation by means of the law of large numbers: in a long series of experiments the fraction of outcomes in which a occurs is approximately p, and the approximation gets better and better as the number of experiments increases. For a regular Markov chain the components of the vector w play the analogous role. That is, the fraction of times that the chain is in state a_i approaches w_i , no matter how one starts.

EXAMPLE 1 Let us take $p^{(0)} = (.1, .9)$ and see how $p^{(n)}$ changes. Using *P* as in the example above, we have that $p^{(1)} = (.5167, .4833)$, $p^{(2)} = (.5861, .4139)$, and $p^{(3)} = (.5977, .4023)$. Recalling that w = (.6, .4), we see that these vectors do approach *w*.

EXAMPLE 2 As an example let us derive the formulas for the fixed point of a 2×2 transition matrix with positive components. Such a matrix is of the form

$$P = \begin{pmatrix} 1 - a & a \\ b & 1 - b \end{pmatrix},$$

where 0 < a < 1 and 0 < b < 1. Since P is regular, it has a unique probability-vector fixed point $w = (w_1, w_2)$. Its components must satisfy the equations

$$w_1(1-a) + w_2b = w_1,$$

 $w_1a + w_2(1-b) = w_2.$

Each of these equations reduces to the single equation $w_1a = w_2b$. This single equation has an infinite number of solutions. However, since w is a probability vector, we must also have $w_1 + w_2 = 1$, and the new equation gives the point [b/(a+b), a/(a+b)] as the unique fixed-point probability vector of P.

EXAMPLE 3 Suppose that the President of the United States tells person A his intention either to run or not to run in the next election. Then A relays the news to B, who in turn relays the message to C, and so on, always to some new person. Assume that there is a probability p > 0 that any one person, when he gets the message, will reverse it before passing it on to the next person. What is the probability that the nth man to hear the message will be told that the President will run? We can consider this as a two-state Markov chain, with states indicated by "yes" and "no." The process is in state "yes" at time n if the nth person to receive the message was told that the President

would run. It is in state "no" if he was told that the President would not run. The matrix P of transition probabilities is then

$$yes no yes no no p p.$$

$$no no 1-p p.$$

Then the matrix P^n gives the probabilities that the nth man is given a certain answer, assuming that the President said "yes" (first row) or assuming that the President said "no" (second row). We know that these rows approach w. From the formulas of the last example, we find that $w = (\frac{1}{2}, \frac{1}{2})$. Hence the probabilities for the nth man's being told "yes" or "no" approach $\frac{1}{2}$ independently of the initial decision of the President. For a large number of people, we can expect that approximately one-half will be told that the President will run and the other half that he will not, independently of the actual decision of the President.

Suppose now that the probability a that a person will change the news from "yes" to "no" when transmitting it to the next person is different from the probability b that he will change it from "no" to "yes." Then the matrix of transition probabilities becomes

$$yes no
 yes no
 no $(1-a a a)$
 no $(1-b)$.$$

In this case w = [b/(a+b), a/(a+b)]. Thus there is a probability of approximately b/(a+b) that the *n*th person will be told that the President will run. Assuming that *n* is large, this probability is independent of the actual decision of the president. For *n* large we can expect, in this case, that a proportion approximately equal to b/(a+b) will have been told that the President will run, and a proportion a/(a+b) will have been told that he will not run. The important thing to note is that, from the assumptions we have made, it follows that it is not the President but the people themselves who determine the probability that a person will be told "yes" or "no," and the proportion of people in the long run that are given one of these predictions.

EXAMPLE 4 For this example, we continue the study of Example 2 in Chapter 3, Section 12. The first approximation treated in that example leads to a two-state Markov chain, and the results are similar to those obtained in Example 1 above. The second approximation led to a four-state Markov chain with transition probabilities given by the matrix

[Ans. Not regular.]

$$\left(\frac{bd}{bd+2ad+ca}, \frac{ad}{bd+2ad+ca}, \frac{ad}{bd+2ad+ca}, \frac{ca}{bd+2ad+ca}\right).$$

Note that the probability of being in state RD after a large number of steps is equal to the probability of being in state DR. This shows that in equilibrium a change from R to D must have the same probability as a change from D to R.

From the fixed vector we can find the probability of being in state R in the far future. This is found by adding the probability of being in state RR and DR, giving

$$\frac{bd + ad}{bd + 2ad + ca}.$$

Notice that, to find the probability of being in state R on the election preceding some election far in the future, we should add the probabilities of being in states RR and RD. That we get the same result corresponds to the fact that predictions far in the future are essentially independent of the particular period being predicted. In other words, the process is acting as if it were in equilibrium.

EXERCISES

Which of the following matrices are regular?

which of the following matrices are regular?

(a)
$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$
.

(b) $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$. [Ans. Regular.]

(c) $\begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$.

(d) $\begin{pmatrix} \frac{1}{5} & \frac{4}{5} \\ 1 & 0 \end{pmatrix}$. [Ans. Regular.]

(e) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$.

(f) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. [Ans. Not regular.]

(g) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$.

(h) $\begin{pmatrix} \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \\ 0 & \frac{1}{5} & \frac{4}{5} \end{pmatrix}$. [Ans. Not regular.]

2. Show that the 2×2 matrix

$$P = \begin{pmatrix} 1 - a & a \\ b & 1 - b \end{pmatrix}$$

is the regular transition matrix if and only if either

- i. $0 < a \le 1$ and 0 < b < 1; or ii. 0 < a < 1 and $0 < b \le 1$.
- 3. Let P be a transition matrix in which all the entries that are not zero have been replaced by x's. Devise a method of raising such a matrix to powers in order to check for regularity. Illustrate your method by showing that

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

is regular.

4. Use the method developed in Exercise 3 to test the following matrix for regularity:

$$P = \begin{pmatrix} 0 & 0 & \frac{1}{10} & 0 & \frac{9}{10} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & \frac{3}{7} & 0 & \frac{4}{7} & 0 \end{pmatrix}.$$

- (a) Give a probability theory interpretation to the condition of regularity.
 - (b) Consider a Markov chain such that it is possible to go from any state a_i to any state a_i and such that p_{kk} is not 0 for at least one state a_k . Prove that the chain is regular. [Hint: Consider the times that it is possible to go from a_i to a_i via a_k .]
- 6. Find the fixed point for the matrix in Exercise 2 for each of the cases listed there. [Hint: Most of the cases were covered in the text above.]
- Find the fixed point w for each of the following regular matrices:

(a)
$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{5}{6} & \frac{1}{6} \end{pmatrix}$$
. [Ans. $\frac{5}{9}$, $\frac{4}{9}$.]

(b)
$$\begin{pmatrix} .37 & .63 \\ .63 & .37 \end{pmatrix}$$
.

(c)
$$\begin{pmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{3}{8} & \frac{5}{8} \end{pmatrix}$$
.

(d)
$$\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$
. [Ans. $(\frac{2}{7}, \frac{3}{7}, \frac{2}{7})$.]

(e)
$$\begin{pmatrix} \frac{5}{7} & \frac{2}{7} \\ \frac{1}{9} & \frac{8}{9} \end{pmatrix}$$
.

- **8.** Let $p^0 = (\frac{1}{2}, \frac{1}{2})$ and compute $p^{(1)}$, $p^{(2)}$, and $p^{(3)}$ for the matrices in Exercises 7(a), (b), and (c). Do they approach the fixed points of these matrices?
- 9. Consider the two-state Markov chain with transition matrix

$$P = \frac{a_1}{a_2} \begin{pmatrix} a_1 & a_2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

What is the probability that after n steps the process is in state a_1 , if it started in state a_2 ? Does this probability become independent of the initial position for large n? If not, the theorem of this section must not apply. Why? Does the matrix have a unique fixed-point probability vector?

10. Compute the first five powers of the matrix

$$P = \begin{pmatrix} .7 & .3 \\ .3 & .7 \end{pmatrix}.$$

From these, guess the fixed-point vector w. Check by computing what w is.

- 11. Prove that, if a regular 3×3 transition matrix has the property that its column sums are 1, its fixed-point probability vector is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. State a similar result for $n \times n$ transition matrices having column sums equal to 1.
- 12. The Land of Oz is blessed by many things, but not good weather. They never have two nice days in a row. If they have a nice day they are just as likely to have snow as rain the next day. If they have snow (or rain), they have an even chance of having the same the next day. If there is a change from snow or rain, only half of the time is this a change to a nice day. Set up a three-state Markov chain to describe this situation. Find the long-range probability for rain, for snow, and for a nice day. What fraction of the days does it rain in the Land of Oz?

 [Ans. The probabilities are: nice, \frac{1}{5}; rain, \frac{2}{5}; snow, \frac{2}{5}.]
- 13. A professor tries not to be late for class too often. If he is late one day, he is 95 percent sure to be on time next time. If he is on time, then the next day there is a 25 percent chance of his being late. In the long run, how often is he late for class?
- 14. Consider the three-state Markov chain with transition matrix

$$P = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{2}{5} & \frac{3}{5} & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(a) Show that the matrix has a unique fixed probability vector. [Ans. $(\frac{2}{5}, \frac{1}{2}, \frac{1}{10})$.]

- (b) Approximately what is the entry in the third column of the first row of P^{100} ?
- (c) What is the interpretation of the entry estimated in (b)?
- 15. A carnival man moves a pea among three shells, A, B, and C. Whenever the pea is under A, he moves it with equal probability to A or B. When it is under B, he is sure to move it to C. When it is under C, he is sure to put it next time under C or B, but is twice as likely to put it under C as B.

Set up a Markov chain taking as states the letters of the shells under which the pea appears after a move. Give the matrix of transition probabilities. Assume that the pea is initially under shell A. Which of the following statements are logically true?

- (a) After the first move, the pea is under A or B.
- (b) After the second move, the pea is under shell B or C.
- (c) If the pea appears under B, it will eventually appear under A again if the process goes on long enough.
- (d) If the pea appears under C, it will not appear under A again.

 [Ans. (a) and (d) are logically true.]
- 16. In Exercise 15, assume that when the pea is under C, the carnival man is sure to put it next time under C or A, but twice as likely to put it under C as A. If you arrive on the scene after he has been playing for a long time, and bet even money that next time it will turn up under a certain shell, which shell should you bet on,
 - (a) Given that you have not seen the previous play?
 - (b) Given that the last time the pea was under A? Which of the above bets would be fair?
- 17. Let P be the matrix

$$P = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Compute the unique probability-vector fixed point of P, and use your result to prove that P is not regular.

- 18. Show that the vector given in Example 4 is the fixed vector of the transition matrix.
- 19. Show that the matrix

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

has more than one probability-vector fixed point. Find the matrix that P^n approaches, and show that it is not a matrix all of whose rows are the same.

20. A businessman goes to a convention in Chicago once a year. While there, he stays at one of four hotels. Two of them, hotels 1 and 2, are expensive. The other two are very expensive. Assuming that he

goes to a given hotel one year, the hotel he goes to the next year is determined by the following matrix of probabilities:

- (a) If he stays at hotel 1 one year, what is the probability that he stays in very expensive hotels for at least two of the next three conventions?
- (b) Find the long-run probabilities for staying in each of the hotels.
- 21. For Exercise 20, compute the following:
 - (a) Given that in 1970 and 1973 he stayed in hotel 1, what is the probability that he stayed in a very expensive hotel during either 1971 or 1972?
 - (b) If in 1970 he stayed in hotel 1 and in 1972 he was in a very expensive hotel, what is the probability that in 1973 he stayed in an expensive hotel? [Ans. $\frac{7}{18}$.]
- 22. A professor has three pet questions, one of which occurs on every test he gives. The students know his habits well. He never uses the same question twice in a row. If he used question 1 last time, he tosses a coin, and uses question 2 if a head comes up. If he used question 2, he tosses two coins and switches to question 3 if at least one comes up heads. If he used question 3, he tosses three coins and switches to question 1 if at least one comes up heads. In the long run, which question does he use most often and how frequently is it used?
- 23. In some cases it makes sense to form a new Markov chain from an old one by condensing two or more states into one.
 - (a) Show that this can be done for the Land of Oz example (Exercise 12), using as the two new states nice and bad (rain or snow). Set up the matrix of transition probabilities and compute the fixed vector. [Partial ans. $(\frac{1}{5}, \frac{4}{5})$.]
 - (b) Compare the fixed vector obtained in part (a) to that obtained in Exercise 12.
 - (c) Would it make sense to condense states in the hotel example (Exercise 20), using as new states expensive (1 or 2) and very expensive (3 or 4)? Explain your answer. [Partial ans. No.]
- 24. A certain company decides each year to add a new workers to its payroll, to remove b workers from its payroll, or to leave its workforce unchanged. There is probability $\frac{3}{4}$ that the action taken in the given year will be the same as the action taken in the previous year. The president of the company has ruled that they should never fire workers the year after they added some, and that they should never hire workers the year after they fired some. Moreover, if no workers were added

or fired in the previous year, the company is twice as likely to add workers as to fire them.

- (a) Set up the problem as a Markov chain with three states.
- (b) Show that it is regular.
- (c) Find the long-run probability of each type of action.
- (d) For what values of a and b will the company tend to increase in size? To decrease? To stay the same?

[Ans. (d)
$$a > b/2$$
; $a < b/2$; $a = b/2$.]

8 ABSORBING MARKOV CHAINS

In this section we shall consider a kind of Markov chain quite different from regular chains.

Definition A state in a Markov chain is an absorbing state if it is impossible to leave it. A Markov chain is absorbing if (1) it has at least one absorbing state, and (2) from every state it is possible to go to an absorbing state (not necessarily in one step).

A particle moves on a line; each time it moves one unit to the right with probability ½, or one unit to the left. We introduce barriers so that if it ever reaches one of these barriers it stays there. As a simple example, let the states be 0, 1, 2, 3, 4. States 0 and 4 are absorbing states. The transition matrix is then

$$P = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 4 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The states 1, 2, 3 are all nonabsorbing states, and from any of these it is possible to reach the absorbing states 0 and 4. Hence the chain is an absorbing chain. Such a process is usually called a *random walk*.

When a process reaches an absorbing state we shall say that it is absorbed.

Theorem In an absorbing Markov chain the probability that the process will be absorbed is 1.

We shall indicate only the basic idea of the proof of the theorem. From each nonabsorbing state a_j it is possible to reach an absorbing state. Let n_j be the minimum number of steps required to reach an absorbing state, starting from state a_j . Let p_j be the probability that, starting from state a_j , the process will *not* reach an absorbing state in n_j steps. Then $p_j < 1$. Let

n be the largest of the n_j and let p be the largest of the p_j . The probability of not being absorbed in n steps is less than p, in 2n steps is less than p^2 , and so on. Since p < 1, these probabilities tend to zero.

For an absorbing Markov chain we consider three interesting questions: (a) What is the probability that the process will end up in a given absorbing state? (b) On the average, how long will it take for the process to be absorbed? (c) On the average, how many times will the process be in each nonabsorbing state? The answer to all these questions depends, in general, on the state from which the process starts.

Consider then an arbitrary absorbing Markov chain. Let us renumber the states so that the absorbing states come first. If there are r absorbing states and s nonabsorbing states, the transition matrix will have the following canonical (or standard) form.

(1)
$$P = r \left(\begin{array}{c|c} r \text{ states} & s \text{ states} \\ \hline I & O \\ \hline R & Q \end{array} \right).$$

Here I is an r-by-r identity matrix, O is an r-by-s zero matrix, R is an s-by-r matrix, and Q is an s-by-s matrix. The first r states are absorbing and the last s states are nonabsorbing.

In Section 7 we saw that the entries of the matrix P^n gave the probabilities of being in the various states starting from the various states. It is easy to show that P^n is of the form

$$(2) P^n = \begin{pmatrix} I & O \\ * & O^n \end{pmatrix},$$

where the asterisk stands for the s-by-r matrix in the lower left-hand corner of P^n , which we do not compute here. The form of P^n shows that the entries of Q^n give the probabilities for being in each of the nonabsorbing states after n steps for each possible nonabsorbing starting state. (After zero steps the process must be in the same nonabsorbing state in which it started. Hence $Q^0 = I$.) By our first theorem, the probability of being in the nonabsorbing states after n steps approaches zero. Thus every entry of Q^n must approach zero as n approaches infinity; that is, $Q^n \to 0$.

Consider then the infinite series

$$I+Q+Q^2+Q^3+\ldots$$

Suppose that Q were a nonnegative number x instead of a nonnegative matrix. To correspond to the fact that $Q^n \to O$ we take x to be less than 1. Then

$$1 + x + x^2 + \dots = (1 - x)^{-1}$$
.

It can be proved that the matrix series behaves in exactly the same way. That is,

$$I + Q + Q^2 + \ldots = (I - Q)^{-1}$$
.

The matrix $(I - Q)^{-1}$ will be called the *fundamental matrix* for the given absorbing chain. It has the following important interpretation:

Let n_{ij} be the mean number of times that the chain is in state a_j if it starts in state a_i , for nonabsorbing states a_i and a_j . Let N be the matrix whose components are n_{ij} . We shall show that $N = (I - Q)^{-1}$. If we take into account the contribution of the original state (which is 1 if i = j and 0 otherwise), we may write the equation

$$n_{ij} = d_{ij} + (p_{i,r+1}n_{r+1,j} + p_{i,r+2}n_{r+2,j} + \ldots + p_{i,r+s}n_{r+s,j}),$$

where d_{ij} is 1 if i = j and 0 otherwise. (Note that the sum in parentheses is merely the sum of the products $p_{ik}n_{kj}$ for k running over the nonabsorbing states.) This equation may be written in matrix form:

$$N = I + QN$$
.

Then (I-Q)N=I, and hence $N=(I-Q)^{-1}$, as was to be shown. Thus we have found a probabilistic interpretation for our fundamental matrix; its i, j entry is the mean number of times that the chain is in state a_j if it starts at a_i . The fact that $N=I+Q+Q^2+\ldots$ also has a probabilistic interpretation. Since the i, j entry of Q^n is the probability of being in a_j on the nth step if we start at a_i , we have shown that the mean of the number of times in state a_j may be written as the sum of the probabilities of being there on particular steps. Thus we have answered question (c) as follows:

Theorem Let $N = (I - Q)^{-1}$ be the fundamental matrix for an absorbing chain. Then the entries of N give the mean number of times in each nonabsorbing state for each possible nonabsorbing starting state.

EXAMPLE 1 In Example 1 the transition matrix in canonical form is **(continued)**

$$\begin{array}{c|ccccc}
0 & 4 & 1 & 2 & 3 \\
0 & 1 & 0 & 0 & 0 & 0 \\
4 & 0 & 1 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\
2 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
3 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0
\end{array}$$

From this we see that the matrix Q is

$$Q = \begin{pmatrix} 0 & \frac{1}{2} & 0\\ \frac{1}{2} & 0 & \frac{1}{2}\\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

and

$$I - Q = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}.$$

Computing $(I - Q)^{-1}$, we find

$$N = (I - Q)^{-1} = \begin{cases} 1 & 2 & 3 \\ 1 & \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{cases}.$$

Thus, starting at state 2, the mean number of times in state 1 before absorption is 1, in state 2 it is 2, and in state 3 it is 1.

We next answer question (b). If we add all the entries in a row, we shall have the mean number of times in any of the nonabsorbing states for a given starting state—that is, the mean time required before being absorbed. This may be described as follows:

Theorem Consider an absorbing Markov chain with s nonabsorbing states. Let c be an s-component column vector with all entries 1. Then the vector t = Nc has as components the mean number of steps before being absorbed for each possible nonabsorbing starting state.

EXAMPLE 1 (continued)

For Example 1 we have

$$t = Nc = \frac{1}{2} \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix}.$$

Thus the mean number of steps to absorption starting at state 1 is 3, starting at state 2 it is 4, and starting at state 3 it is again 3. Since the process necessarily moves to 1 or 3 from 2, it is clear that it requires one more step starting from 2 than from 1 or 3.

We now consider question (a). That is, what is the probability that an absorbing chain will end up in a particular absorbing state? It is clear that this probability will depend upon the starting state and be interesting only for the case of a nonabsorbing starting state. We write as usual our matrix in the canonical form

$$P = \left(\begin{array}{c|c} I & O \\ \hline R & Q \end{array}\right).$$

Theorem Let b_{ij} be the probability that an absorbing chain will be absorbed in state a_i if it starts in the nonabsorbing state a_i . Let B be the matrix with

entries b_{ij} . Then

$$B = NR$$
,

where N is the fundamental matrix and R is as in the canonical form.

Proof Let a_i be a nonabsorbing state and a_j be an absorbing state. If we compute b_{ij} in terms of the possibilities on the outcome of the first step, we have the equation

$$b_{ij} = p_{ij} + \sum_{k} p_{ik} b_{kj},$$

where the summation is carried out over all nonabsorbing states a_k . Writing this in matrix form gives

$$B = R + QB$$

$$(I - Q)B = R$$

$$B = (I - Q)^{-1}R = NR.$$

and hence

EXAMPLE 1 (continued)

In the random-walk example we found that

$$N = \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix}.$$

From the canonical form we find that

$$R = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

Hence

$$B = NR = \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}.$$

Thus, for instance, starting from a_1 , there is probability $\frac{3}{4}$ of absorption in a_0 and $\frac{1}{4}$ for absorption in a_4 .

Let us summarize our results. We have shown that the answers to questions (a), (b), and (c) can all be given in terms of the fundamental matrix $N = (I - Q)^{-1}$. The matrix N itself gives us the mean number of times in each state before absorption depending upon the starting state. The column vector t = Nc gives us the mean number of steps before absorption, de-

pending upon the starting state. The matrix B = NR gives us the probability of absorption in each of the absorbing states, depending upon the starting state.

EXERCISES

Which of the following transition matrices are from absorbing chains?

$$(\mathbf{a}) \quad P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{pmatrix}.$$

(b)
$$P = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}$$
.

(c)
$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & \frac{3}{7} & 0 & \frac{4}{7} & 0 \end{pmatrix}.$$
(d)
$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

[Ans. Absorbing.]

(d)
$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(e)
$$P = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{5}{8} & 0 & \frac{1}{4} & \frac{1}{8} \end{pmatrix}.$$

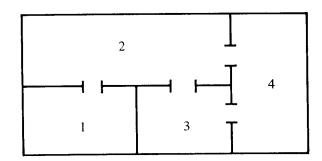
Consider the three-state transition matrix

$$P = \begin{pmatrix} 1 & 0 & 0 \\ a & \frac{1}{2} & b \\ c & 0 & d \end{pmatrix}.$$

For what choices of a, b, c, and d do we not obtain an absorbing chain?

- 3. In the random-walk example (Example 1) of the present section, assume that the probability of a step to the right is $\frac{3}{4}$ and a step to the left is $\frac{1}{4}$. Find N, t, and B. Compare these with the results for probability $\frac{1}{2}$ for a step to the right and $\frac{1}{2}$ to the left.
- In the hotel example of Exercise 20, Section 7, let us assume that the man is so impressed with the service at hotel 2 that, once he goes there, he refuses to go to any of the other hotels again. This gives

- (a) Find the fundamental matrix N, and also T and B. What is the interpretation of these quantities?
- (b) Given that in 1970 and 1973 he stayed in hotel 1, what is the probability that he stayed in a very expensive hotel during either 1971 or 1972?
- 5. A rat is put into the maze of the figure below. Each time period, it chooses at random one of the doors in the compartment it is in and moves into another compartment.



- (a) Set up the process as a Markov chain (with states being the compartments) and identify it. [Ans. Regular.]
- (b) In the long run, what fraction of his time will the rat spend in compartment 2? [Ans. \frac{3}{8}.]
- (c) What is the relation between the number of entrances to a given compartment and the fraction of the time the rat will spend in that compartment?
- (d) Make compartment 4 into an absorbing state by assuming the rat will stay in it once it reaches it. Set up the new process, and identify it as a kind of Markov chain. [Ans. Absorbing.]
- (e) In part (d), if the rat starts in compartment 2, how many steps will it take him, on the average, to reach compartment 4?
- 6. An analysis of a recent hockey game between Dartmouth and Princeton showed the following facts: If the puck was in the center (C) the probabilities that it next entered Princeton territory (P) or Dartmouth territory (D) were .4 and .6, respectively. From D it went back to C with probability .95 or into the Dartmouth goal (\overline{D}) with probability .05 (Princeton scores one point). From P it next went to C with probability .9 and to Princeton's goal (\overline{P}) with probability .1 (Dartmouth scores one point). Assuming that the puck begins in C after each point, find the transition matrix of this five-state Markov chain. Calculate the probability that Dartmouth will score. $[Ans. \frac{4}{7}]$
- 7. The following is an alternative method of finding the probability of absorption in a particular state, say a_j . Find the column vector d such that the jth component of d is 1, all other components corresponding to absorbing states are 0, and Pd = d. There is only one such vector. Component d_i is the probability of absorption in a_i if the process starts

- in a_i . Use this method to find the probability of absorption in state 0 in the random-walk example given in this section.
- 8. The following is an alternative method for finding the mean number of steps to absorption. Let t_i be the mean number of steps to absorption starting at state a_i . This must be the same as taking one more step and then adding $p_{ij}t_i$ for every nonabsorbing state a_i .
 - (a) Give reasons for the claim above that

$$t_i = 1 + \sum_j p_{ij} t_j,$$

where the summation is over the nonabsorbing states.

- (b) Solve for t for the random-walk example.
- (c) Verify that the solution agrees with that found in the text.
- 9. A man is in jail and needs \$400 for bail. Once he is out, he can recover his million-dollar loot. In jail he can make a sequence of bets in which, if he bets X dollars, he wins X dollars with probability $\frac{1}{3}$ and loses X dollars with probability $\frac{2}{3}$. He can bet any amount he wishes as long as he can pay if he loses. He has \$100. He decides to try the bold strategy of betting as much as he has each time or at least enough to get his \$400—that is, to bet \$100 if he has \$100, \$200 if he has \$200, and \$100 if he has \$300. To assess his chances of getting out now, he sets up the transition matrix:

$$P = \begin{pmatrix} 0 & 400 & 100 & 200 & 300 \\ 1 & 0 & 0 & 0 & 0 \\ 100 & \frac{2}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 200 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \end{pmatrix} = \begin{pmatrix} I & O \\ R & Q \end{pmatrix}.$$

- (a) Find the matrix $N = (I Q)^{-1}$.
- (b) Find the expected number of bets that he will make under this bold strategy.

 [Ans. \frac{4}{3}.]
- (c) Find the probability that he will get his bail money under this strategy.
- (d) Repeat parts (a), (b), and (c) assuming that he uses a more timid strategy of betting \$100 each time. Which strategy provides a longer expected game? Which strategy gives him the better chance of recovering his loot?
- 10. A number is chosen at random from the integers 1, 2, 3, 4, 5. If x is chosen, then another number is chosen from the set of integers less than or equal to x. This process is continued until the number 1 is chosen. Form a Markov chain by taking as states the largest number

that can be chosen. Show that

$$N = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 2 & 1 & 0 & 0 & 0 \\ 1 & \frac{1}{2} & 0 & 0 \\ 4 & 1 & \frac{1}{2} & \frac{1}{3} & 0 \\ 5 & 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \end{pmatrix} + I,$$

where I is the 4 \times 4 identity matrix. What is the mean number of draws? [Ans. $\frac{37}{12}$.]

- 11. Using the result of Exercise 10, make a conjecture for the form of the fundamental matrix if we start with integers from 1 to n. What would the mean number of draws be if we started with numbers from 1 to 10?
- 12. Peter and Paul are matching pennies, and each player flips his (fair) coin before revealing it. They initially have four pennies between them and the game ends whenever one of them has all the pennies. Let the states be labeled with the number of pennies that Peter has.
 - (a) Write the transition matrix.
 - **(b)** What kind of a Markov chain is it?
 - (c) If Peter initially has two pennies, what is the probability that he will win the game?
- 13. A certain college which is trying to pass several liberal measures is plagued by the problem of conservative alumni. It is determined that if an alumnus votes for a liberal policy, he will with probability $\frac{5}{6}$ vote in favor of the next policy and will with probability $\frac{1}{6}$ turn conservative. Once he turns conservative he will continue to vote against all liberal policies. Assume there are 20,000 alumni, 4000 of whom voted conservatively before the college starts trying to pass these measures. If all the alumni vote and if 12,000 opposing votes are needed to defeat a policy which the college is trying to pass, how many of its new liberal policies can the college expect to pass?
- 14. Three tanks fight a three-way duel. Tank A has probability $\frac{1}{2}$ of destroying the tank it fires at. Tank B has probability $\frac{1}{3}$ of destroying its target tank, and tank C has probability $\frac{1}{6}$ of destroying its target tank. The tanks fire together and each tank fires at the strongest opponent not yet destroyed. Form a Markov chain by taking as state the tanks which survive any one round. Find N, t, B, and interpret your results.
- 15. Consider the following model. A man buys a store. The profits of the store vary from month to month. For simplicity we assume that he earns either \$5000 or \$2000 a month ("high" or "low"). The man may sell his store at any time; there is a 10 percent chance of his selling during a high-profit month and a 40 percent chance during a low-profit month. If he does not sell, with probability \(^2_3\) the profits will be the same the next month, and with probability \(^1_3\) they will change.

(a) Set up the transition matrix.

[Ans. High
$$\begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{10} & \frac{3}{5} & \frac{3}{10} \\ \text{Low} & \begin{pmatrix} \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \end{pmatrix}$$
.]

- (b) Compute N, Nc, and NR and interpret each.
- (c) Let $f = \begin{pmatrix} 5000 \\ 2000 \end{pmatrix}$ and compute the vector g = Nf.

$$\left[Ans.\ g = \binom{20,000}{10,000}\right].$$

(d) Show that the components of g have the following interpretation: g_i is the expected amount that he will gain before selling, given that he started in state i.

SUGGESTED READING

Beaumont, R. A. Linear Algebra. New York: Harcourt, Brace and World, 1965. Birkhoff, G., and Maclane, S. A Survey of Modern Algebra. 3rd ed. New York: Macmillan, 1965.

Fraleigh, John B. First Course in Abstract Algebra. Reading, Mass.: Addison-Wesley, 1967.

Hadley, G. Linear Algebra. Reading, Mass.: Addison-Wesley, 1961. Chapters 1-5.

Hohn, Franz E. Elementary Matrix Algebra. 2nd ed. New York: Macmillan, 1964.

Johnson, R. E. First Course in Abstract Algebra. Englewood Cliffs, N.J.: Prentice-Hall, 1953.

Kemeny, John G., and Snell, J. Laurie. Finite Markov Chains. Princeton, N.J.: Van Nostrand, 1960.

Noble, Ben. Applied Linear Algebra. Englewood Cliffs, N.J.: Prentice-Hall, 1969.

Weiss, Marie J., and Dubisch, R. Higher Algebra for the Undergraduate. 2nd ed. New York: John Wiley, 1962.