V

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

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 2 \\ 4 \end{pmatrix}.$$

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

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

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

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

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 and only if corresponding components of the vector are equal. Thus for the vectors

$$u = (1, 2), \quad v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad w = (1, 2), \quad x = (2, 1),$$

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 component-wise 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

$$\begin{aligned} u + v &= (u_1, u_2, u_3) + (v_1, v_2, v_3) \\ &= (u_1 + v_1, u_2 + v_2, u_3 + v_3). \end{aligned}$$

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 component-wise 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 is immaterial as far as the answer goes, 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,

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

and

$$\begin{aligned} (1, 0, 0) + (0, 2, 0) + (0, 0, 3) &= (1, 2, 0) + (0, 0, 3) = (1, 2, 3) \\ &= (1, 0, 0) + (0, 2, 3) = (1, 2, 3). \end{aligned}$$

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, etc.

The multiplication of a number a times a vector v is defined by component-wise 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 component-wise multiplication.

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, i.e., 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

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

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 one can denote a whole collection of numbers by a single letter such as u, v, . . . , and treat such a collection as if it were a single quantity. By using the vector notation it is possible to state very complicated relationships in a simple manner. The student will see many examples of this in the remainder of the present chapter and the two succeeding chapters.

EXERCISES

1. Compute the quantities below for the vectors

$$u = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix}, \quad w = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

(a) $2u$. [*Ans.* $\begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix}$.]

(b) $-v$.
(c) $2u - v$.

(d) $v + w$. [*Ans.* $\begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$.]

(e) $u + v - w$.
(f) $2u - 3v - w$.

(g) $3u - v + 2w$. [*Ans.* $\begin{pmatrix} 9 \\ -2 \\ 8 \end{pmatrix}$.]

2. Compute (a) through (g) of Exercise 1 if the vectors u, v, and w are

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

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 u and v are two row or two column vectors having the same number of components, prove that $u + 0v = u$ and $0u + v = v$.

5. If $2u - v = 0$, what is the relationship between the components of u and those of v? [*Ans.* $v_i = 2u_i$.]

6. Answer the question in Exercise 5 for the equation $-3u + 5v + u - 7v = 0$. Do the same for the equation $20v - 3u + 5v + 8u = 0$.

7. When possible, compute the following sums; when not possible, give reasons.

(a) $\begin{pmatrix} -1 \\ 3 \end{pmatrix} + \begin{pmatrix} 6 \\ -2 \\ 5 \\ -4 \end{pmatrix} = ?$

(b) $(2, -1, -1) + 0(4, 7, -2) = ?$

(c) $(5, 6) + 7 - 21 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = ?$

(d) $1\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = ?$

8. If $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, find u_1, u_2, and u_3. [*Ans.* 0; -2; -2.]

9. If $2\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$, find the components of v.

10. 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?

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

12. 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?

13. 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 12.

14. 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 12, 13, and 14.

SUPPLEMENTARY EXERCISES

15. 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?
(b) What are the possibilities if he has a 3.0 average?
(c) What are the possibilities if he has a 2.0 average?

16. Consider the vectors

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

Show that the vector

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

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

17. (a) Show that the vector equation

$$x\begin{pmatrix}3\\-4\end{pmatrix}+y\begin{pmatrix}-4\\5\end{pmatrix}=\begin{pmatrix}2\\-3\end{pmatrix}$$

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

(b) Solve these equations for x and y and substitute into the above vector equation to check your work.

18. Write the following simultaneous linear equations in vector form

$$\begin{aligned}ax+by&=e\\cx+dy&=f.\end{aligned}$$

[*Hint:* Follow the form given in Exercise 17.]

19. Let $x=\begin{pmatrix}x_1\\x_2\end{pmatrix}$. Define $x\geq 0$ to be the conjunction of the statements $x_1\geq 0$ *and* $x_2\geq 0$. Define $x\leq 0$ analogously. Now prove that if $x\geq 0$, then $-x\leq 0$.

20. Using the definition in Exercise 19, define $x\geq y$ to mean $x-y\geq 0$, where x and y are vectors of the same shape. Consider the following four vectors:

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

(a) Show that $x\geq y$.
(b) Show that $u\geq y$.
(c) Is there any relationship between x and u?
(d) Show that $v\geq x$, $v\geq y$, and $v\geq u$.

21. If $x\geq y$ and $y\geq u$, prove that $x\geq u$.

22. If $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ is a set of n vectors, show how to find a vector u such that $u\geq x^{(i)}$ for all i. Also show how to find a vector v such that $v\leq x^{(i)}$ for all i.

2. THE PRODUCT OF VECTORS

The reader may have wondered why it was necessary to introduce both column and row vectors when their properties are so similar. This question can be answered in several different ways. In the first place, in many applications there are two kinds of quantities which 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 eggs 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:

$$\begin{aligned} x &= [6\text{ (apples)}, 12\text{ (eggs)}, 3\text{ (lemons)}, 12\text{ (oranges)}, 6\text{ (pears)}] \\ &= (6, 12, 3, 12, 6). \end{aligned}$$

Suppose that apples are 4 cents each, eggs are 6 cents each, lemons are 9 cents each, oranges are 5 cents each, and pears are 7 cents each. We can then represent the prices of these items as a column vector,

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

The obvious question to ask now is, what is the total amount that Smith must pay for his purchases? What we would like to do is 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:

$$\begin{aligned} x \cdot y &= (6, 12, 3, 12, 6) \begin{pmatrix} 4 \\ 6 \\ 9 \\ 5 \\ 7 \end{pmatrix} \\ &= 6 \cdot 4 + 12 \cdot 6 + 3 \cdot 9 + 12 \cdot 5 + 6 \cdot 7 \\ &= 24 + 72 + 27 + 60 + 42 \\ &= 225 \text{ cents or } \$2.25. \end{aligned}$$

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 product $u \cdot v$ 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 given below.

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

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

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

Example 2. Consider an oversimplified economy which 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

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

and the demand vectors of each of the industries might be

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

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

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

Hence the profit of the steel industry is \$120 − \$40 = \$80. In the exercises below 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. The latter complication will be introduced in Chapter VII.

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 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.

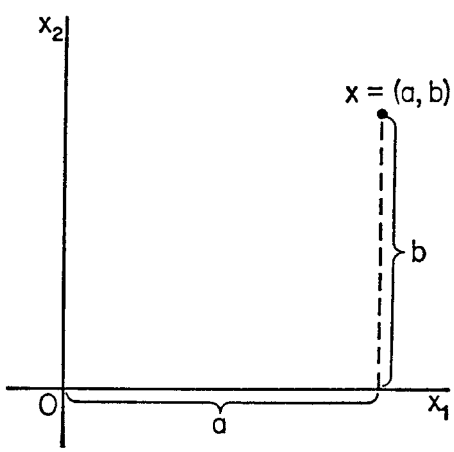


Figure 1

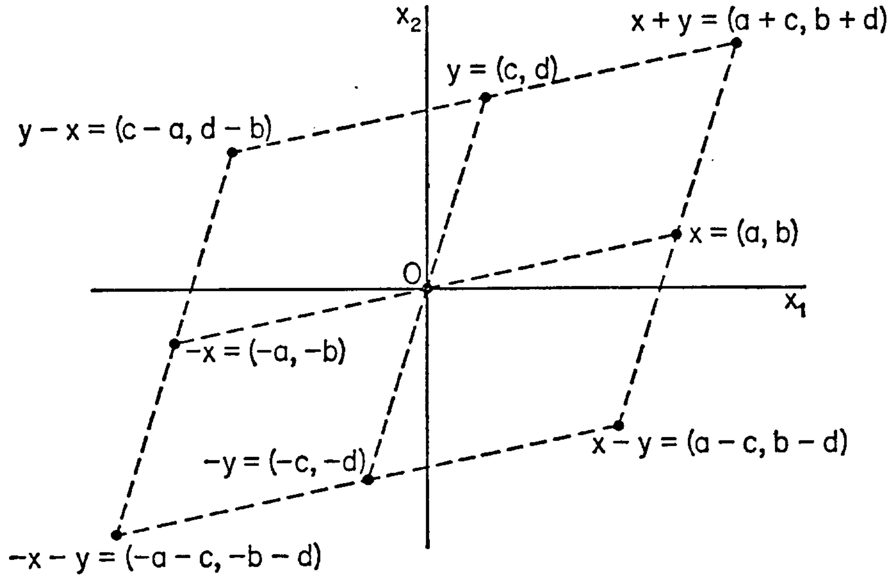


Figure 2

The idea of multiplying a row vector by a number can also be given a geometric meaning, see Figure 3. There we have plotted the points 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 which has geometrical significance is the vector $z = ax + (1 - a)y$, where a is any number between 0 and 1. Observe in Figure 4 that the points z all lie on the line segment between

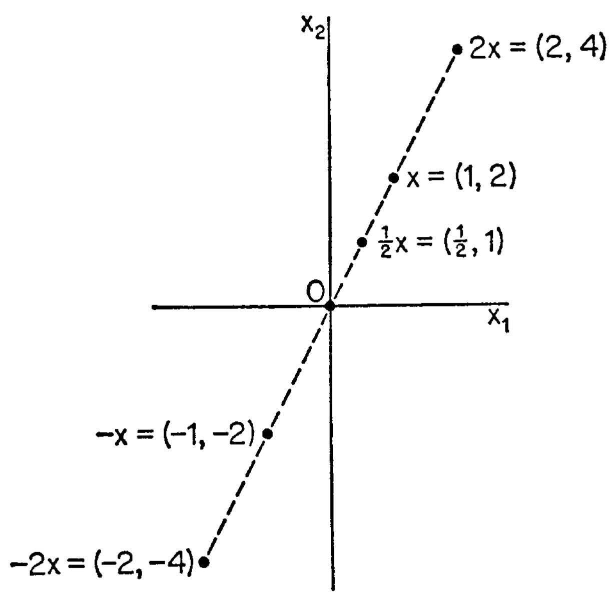


Figure 3

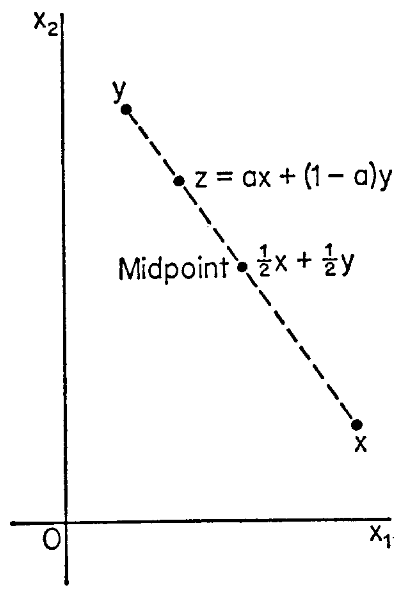


Figure 4

the points x and y. If $a = \frac{1}{2}$, the corresponding point on the line segment is the mid-point of the segment. Thus, if $x = (a, b)$ and $y = (c, d)$, then the point

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

is the mid-point of the line segment between x and y.

EXERCISES

1. Compute the quantities below for the following vectors:

$$u = (1, -1, 4), \qquad x = (0, 1, 2),$$

$$v = \begin{pmatrix}5\\0\\1\end{pmatrix}, \qquad y = \begin{pmatrix}-1\\-1\\2\end{pmatrix}.$$

(a) $u \cdot v + x \cdot y = ?$ [*Ans.* 12.]
(b) $(-u + 5x) \cdot (3v - 2y) = ?$
(c) $5u \cdot v + 10[x \cdot (2v - y)] = ?$ [*Ans.* 55.]
(d) $2[(u - x) \cdot (v + y)] = ?$

2. Plot the points corresponding to the row vectors $x = (3, 4)$ and $y = (-2, 7)$. Then compute and plot the following vectors.
(a) $\frac{1}{2}x + \frac{1}{2}y$.
(b) $x + y$.
(c) $x - 2y$.
(d) $\frac{7}{8}x + \frac{1}{8}y$.
(e) $3x - 2y$.
(f) $4y - 3x$.

3. If $x = (1, -1, 2)$ and $y = (0, 1, 3)$ are points in space, what is the mid-point of the line segment joining x to y? [*Ans.* $(\frac{1}{2}, 0, \frac{5}{2})$.]

4. If u is a three-component row vector and v is a three-component column vector, and a is a number, prove that $a(u \cdot v) = (au) \cdot v = u \cdot (av)$.

5. Suppose that Brown, Jones, and Smith go to the grocery store and purchase the following items:
Brown: two apples, six lemons, and five pears;
Jones: two dozen eggs, two lemons, and two dozen oranges;
Smith: ten apples, one dozen eggs, two dozen oranges, and a half dozen pears.

(a) How many different kinds of items did they purchase? [*Ans.* 5.]

(b) Write each of their purchases as row vectors with as many components as the answer found in (a).

(c) Using the price vector given in Example 1, compute each man's grocery bill. [*Ans.* \$0.97, \$2.82, \$2.74.]

(d) By means of vector addition, find the total amount of their purchases as a row vector.

(e) Compute in two different ways the total amount spent by the three men at the grocery store. [*Ans.* \$6.53.]

6. Prove that vector multiplication satisfies the following property:

$$u \cdot (v + w) = u \cdot v + u \cdot w,$$

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

7. The production of a book involves several steps: first it must be set in type, then it must be printed, and finally it must be supplied with covers and bound. Suppose that the typesetter charges \$6 an hour, paper costs $\frac{1}{4}$ cent per sheet, that the printer charges 11 cents for each minute that his press runs, that the cover costs 28 cents, and that 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 five minutes of press time per book.

(a) Write a five-component row vector which gives the requirements for the first book. Write another row vector which 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.* \$1,801.53.]

(c) Using vector addition and multiplication, find the cost of printing a first edition run of 5000 copies. [*Ans.* \$9,450.]

(d) Assuming that the type plates from the first edition are used again, find the cost of printing a second edition of 5000 copies. [*Ans.* \$7,650.]

8. 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.

9. A building contractor has accepted orders for five ranch style houses, seven Cape Cod houses, and twelve Colonial style houses. Write a three-component row vector x whose components give the numbers of each type of house to be built. Suppose that he knows that a ranch style house requires 20 units of wood, a Cape Cod 18 units, and a Colonial style 25 units of wood. Write a column vector u whose components give the various quantities of wood needed for each type of house. Find the total amount of wood needed by computing the matrix product xu. [*Ans.* 526.]

10. Let $x = (x_1, x_2)$ and let a and b be the vectors

$$a = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \qquad b = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

If $x \cdot a = -1$ and $x \cdot b = 7$, determine x_1 and x_2. [*Ans.* $x_1 = -31$; $x_2 = 23$.]

11. Let $x = (x_1, x_2)$ and let a and b be the vectors

$$a = \begin{pmatrix} 4 \\ 8 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

If $x \cdot a = x_1$ and $x \cdot b = x_2$, determine x_1 and x_2.

SUPPLEMENTARY EXERCISES

12. Consider the vectors

$$x = (5, 8), \qquad y = (3, 7), \qquad f = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(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. [*Ans.* 6.5, 5.]

(b) Compute $\frac{1}{4}(x + y)f$ and give an interpretation for this number. [*Partial Ans.* 5.75.]

13. 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 $\frac{1}{n}xf$ and $\frac{1}{n}yf$ and interpret the result.

(b) Compute $\frac{1}{2n}(x + y)f$ and interpret the result.

[*Hint:* Exercise 12 is a special case.]

14. Consider an experiment in which there are two outcomes; we get \$2 with probability $\frac{1}{3}$ and \$3 with probability $\frac{2}{3}$. Let

$$a = (2, 3) \quad \text{and} \quad p = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}.$$

Show that the expected outcome of the experiment is ap.

15. 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

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

Show that the expected outcome is ap.

16. Consider the vectors

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

and a number c. Show that the equation $ax = c$ is a single equation in two variables.

17. Consider the vectors

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

and two numbers c_1 and c_2. Show that the equations

$$\begin{aligned} ax &= c_1 \\ bx &= c_2 \end{aligned}$$

represent two simultaneous equations in two unknowns.

18. Show that every set of two simultaneous equations in two unknowns can be written as in Exercise 17.

3. *MATRICES AND THEIR COMBINATION WITH VECTORS*

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

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

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.

$$(1, 2, 3), \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix},$$

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

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 IV, Section 13, 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.)

$$\begin{array}{l c} & \begin{array}{ccccc} \text{Steel} & \text{Wood} & \text{Glass} & \text{Paint} & \text{Labor} \end{array} \\ \begin{array}{l} \text{Ranch:} \\ \text{Cape Cod:} \\ \text{Colonial:} \end{array} & \begin{pmatrix} 5 & 20 & 16 & 7 & 17 \\ 7 & 18 & 12 & 9 & 21 \\ 6 & 25 & 8 & 5 & 13 \end{pmatrix} = R \end{array}$$

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:

$$
\begin{aligned}
xR = (5, 7, 12)\begin{pmatrix} 5 & 20 & 16 & 7 & 17 \\ 7 & 18 & 12 & 9 & 21 \\ 6 & 25 & 8 & 5 & 13 \end{pmatrix}
&= (5\cdot 5 + 7\cdot 7 + 12\cdot 6, \quad 5\cdot 20 + 7\cdot 18 + 12\cdot 25, \\
&\qquad 5\cdot 16 + 7\cdot 12 + 12\cdot 8, \quad 5\cdot 7 + 7\cdot 9 + 12\cdot 5, \\
&\qquad 5\cdot 17 + 7\cdot 21 + 12\cdot 13) \\
&= (146, 526, 260, 158, 388).
\end{aligned}
$$

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:

$$xA = (x_1, x_2, \ldots, x_m) \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

$$= (x_1a_{11} + x_2a_{21} + \ldots + x_ma_{m1},\quad x_1a_{12} + x_2a_{22} + \ldots + x_ma_{m2},$$
$$\ldots,\quad x_1a_{1n} + x_2a_{2n} + \ldots + x_ma_{mn});$$

$$Au = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{pmatrix} = \begin{pmatrix} a_{11}u_1 + a_{12}u_2 + \dots + a_{1n}u_n \\ a_{21}u_1 + a_{22}u_2 + \dots + a_{2n}u_n \\ \dots \\ a_{m1}u_1 + a_{m2}u_2 + \dots + a_{mn}u_n \end{pmatrix}.$$

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. In Exercise 6 of Chapter IV, Section 13, we considered a Markov chain with transition matrix

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

The initial state was chosen by a random device that selected 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

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

Using the methods of Chapter IV, 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

$$y = \begin{pmatrix} 5 \\ 5 \\ 10 \\ 4 \\ 6 \end{pmatrix} \begin{array}{l} \text{cents per apple} \\ \text{cents per egg} \\ \text{cents per lemon} \\ \text{cents per orange} \\ \text{cents per pear.} \end{array}$$

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:

$$\begin{array}{cc} & \begin{array}{ccc} \text{Prices,} & \text{Prices,} & \text{Minimum} \\ \text{Store 1} & \text{Store 2} & \text{Price} \end{array} \\ P = & \begin{pmatrix} 4 & 5 & 4 \\ 6 & 5 & 5 \\ 9 & 10 & 9 \\ 5 & 4 & 4 \\ 7 & 6 & 6 \end{pmatrix}. \end{array}$$

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

Example 4. 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:

$$M = \begin{array}{c c} & \begin{array}{ccc} \text{Ticket A} & \text{Ticket B} & \text{Ticket C} \end{array} \\ \begin{array}{r} \text{Red} \\ \text{Green} \\ \text{White} \end{array} & \begin{pmatrix} 1 & 3 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 16 \end{pmatrix} \end{array}.$$

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

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

From this we have

$$\begin{aligned} pM &= (\tfrac{5}{9}, \tfrac{3}{9}, \tfrac{1}{9})\begin{pmatrix} 1 & 3 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 16 \end{pmatrix} \\ &= (1\cdot\tfrac{5}{9} + 4\cdot\tfrac{3}{9} + 0\cdot\tfrac{1}{9},\quad 3\cdot\tfrac{5}{9} + 1\cdot\tfrac{3}{9} + 0\cdot\tfrac{1}{9},\quad 0\cdot\tfrac{5}{9} + 0\cdot\tfrac{3}{9} + 16\cdot\tfrac{1}{9}) \\ &= (\tfrac{17}{9},\ \tfrac{18}{9},\ \tfrac{16}{9}). \end{aligned}$$

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) $\begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} 7 \\ 2 \end{pmatrix} = ?$

(b) $(3, -4)\begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} = ?$ [*Ans.* (11, −11).]

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

(d) $(2, 2)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = ?$ [*Ans.* (0, 0).]

(e) $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 5 \end{pmatrix} = ?$

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

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

(h) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = ?$

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

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

2. What number does the matrix in parts (i) and (j) above resemble?

3. Notice that in Exercise 1(d) above the product of a row vector, none of whose components is zero, times a matrix, none of whose components is zero, yields the zero row vector. Find another example which is similar to this one. Answer the analogous question for Exercise 1(e).

4. When possible, solve for the indicated quantities.

(a) $(x_1, x_2)\begin{pmatrix} 0 & -1 \\ 7 & 3 \end{pmatrix} = (7, 0)$. Find the vector x. [*Ans.* $(3, 1)$.]

(b) $(2, -1)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (6, 3)$. Find the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. In this case can you find more than one solution?

(c) $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Find the vector u.

(d) $\begin{pmatrix} -1 & 4 \\ 2 & -8 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -6 \end{pmatrix}$. Find u.

How many solutions can you find?

[*Ans.* $u = \begin{pmatrix} 4k - 3 \\ k \end{pmatrix}$, for any number k.]

5. 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.* $u = \begin{pmatrix} k \\ 2k \end{pmatrix}$ for any number k.]

(c) $\begin{pmatrix} \frac{5}{8} & \frac{1}{8} \\ \frac{3}{8} & \frac{7}{8} \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$; find u. How many answers are there?

6. In Exercise 5 of the preceding section construct the 3×5 matrix whose rows give the various purchases of Brown, Jones, and Smith. Multiply on the right by the five-component price (column) vector to find the three-component column vector whose entries give each person's grocery bill. Multiply on the left by the row vector $x = (1, 1, 1)$ and on the right by the price vector to find the total amount that they spent in the store.

7. In Example 1 of this section, assume that the contractor is to build seven ranch style, three Cape Cod, and five 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 of this section, assume that the initial probability vector is $p^{(0)} = (\frac{1}{3}, \frac{2}{3})$. Find the vector $p^{(1)}$. [*Ans.* $(\frac{4}{9}, \frac{5}{9})$.]

9. For the Markov chain whose transition matrix is

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

assume the initial probability vector is $p^{(0)} = (\frac{1}{2}, \frac{1}{6}, \frac{1}{3})$. Draw the tree of the process and find the tree measures. Compute $p^{(1)}$ by means of the tree measure and also from the formula $p^{(1)} = p^{(0)}P$ and show that the two answers agree.

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. 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. In Example 2 use tree measures to show that $p^{(2)} = p^{(1)}P$.

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

$$\begin{array}{ll} & \begin{array}{cccc} \text{A} & \text{B} & \text{C} & \text{D} \end{array} \\ \begin{array}{l} \text{Vitamin:} \end{array} & \\ \begin{array}{l} \text{Food I:} \\ \text{Food II:} \\ \text{Food III:} \end{array} & \begin{pmatrix} .5 & .5 & 0 & 0 \\ .3 & 0 & .2 & .1 \\ .1 & .1 & .2 & .5 \end{pmatrix} \end{array}$$

If we eat five units of food I, ten units of food II, and eight 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 eaten.

[*Ans.* $(6.3, 3.3, 3.6, 5.0)$, $\begin{pmatrix} 15 \\ 13 \\ 33 \end{pmatrix}$, \$4.69.]

13. In Example 3, by how much would store 1 have to reduce the price of oranges in order to make Smith's purchases less expensive at store 1 than at store 2?

14. In Example 3, find the store at which the total cost to Smith is less when he wishes to purchase

(a) $x = (4, 1, 2, 0, 1)$. [*Ans.* Store 1, cost 47 cents.]

(b) $x = (2, 1, 3, 1, 0)$.

(c) $x = (2, 1, 1, 2, 0)$.

15. In Example 4, let us assume that an individual chooses ticket A with probability r_1, ticket B with probability r_2, and ticket C with probability r_3. Let $r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$. Give an interpretation for pMr. Compute this for the case that $r_1 = r_2 = r_3$. [*Ans.* $pMr = \frac{17}{9}$, which is the expected return.]

SUPPLEMENTARY EXERCISES

16. 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

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

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.]

17. 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.

18. Consider the matrices

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

Show that $Pf = f$.

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. If $Ax = 0$ and $Ay = 0$, show that $A(x + y) = 0$.

21. If $Ax = b$ and $Ay = 0$, show that $A(x + y) = b$.

4. THE ADDITION AND MULTIPLICATION OF MATRICES

Two matrices of the same shape, that is, having the same numbers 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

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

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 the following:

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

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

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

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

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

Observe that this is merely component-wise multiplication, as was the analogous concept for vectors. Some 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, of course, 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 as

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

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

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

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

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.

$$\begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 7 & 0 \\ -2 & -3 \end{pmatrix} = \begin{pmatrix} 16 & 3 \\ -6 & -9 \end{pmatrix};$$

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

$$\begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 5 \end{pmatrix}\begin{pmatrix} 1 & 3 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 10 & 4 & 4 \\ 2 & 6 & 5 & 5 \end{pmatrix}.$$

One obvious question that now arises is that of multiplying 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, i.e.,

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

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 operations.

(a) $2\begin{pmatrix} 6 & 1 \\ 0 & -3 \\ -1 & 2 \end{pmatrix} - 3\begin{pmatrix} 4 & 2 \\ 0 & 1 \\ -5 & -1 \end{pmatrix} = ?$ [*Ans.* $\begin{pmatrix} 0 & -4 \\ 0 & -9 \\ 13 & 7 \end{pmatrix}$.]

(b) $\begin{pmatrix} 6 & 1 & -1 \\ 1 & -3 & 2 \end{pmatrix} - 5\begin{pmatrix} 4 & 0 & -5 \\ 2 & 1 & -1 \end{pmatrix} = ?$

(c) $\begin{pmatrix} 6 & 1 \\ 0 & -3 \end{pmatrix}\begin{pmatrix} 4 & 0 & -4 \\ 2 & 1 & -1 \end{pmatrix} = ?$

(d) $\begin{pmatrix} 6 & 0 & -1 \\ 1 & -3 & 2 \end{pmatrix}\begin{pmatrix} 4 & 2 \\ 0 & 1 \\ -5 & -1 \end{pmatrix} = ?$ [*Ans.* $\begin{pmatrix} 29 & 13 \\ -6 & -3 \end{pmatrix}$.]

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

(f) $\begin{pmatrix} 4 & 1 & 4 \\ -1 & -2 & -1 \\ 2 & -1 & -2 \end{pmatrix}\begin{pmatrix} 3 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = ?$ [*Ans.* $\begin{pmatrix} 11 & 2 & 12 \\ -1 & -4 & -3 \\ 7 & -2 & -2 \end{pmatrix}$.]

(g) $\begin{pmatrix} 1 & -2 \\ 0 & 0 \\ 7 & 5 \\ -4 & 8 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} -7 & 9 & -5 & 6 & 0 \\ -1 & 0 & 3 & -4 & 1 \end{pmatrix} = ?$

2. 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.

3. 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.

4. 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.

5. Verify the associative law for the special case when

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

6. Consider the matrices

$$A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 17 & 57 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix},$$

$$D = \begin{pmatrix} -1 & -1 \\ 2 & 2 \\ 1 & 1 \end{pmatrix}.$$

The shapes of these are 2×3, 4×3, 3×3, and 3×2, respectively. What is the shape of

(a) AC.
(b) DA.
(c) AD.
(d) BC.
(e) CB.
(f) DAC.
(g) $BCDA$. [*Ans.* 4×3.]

7. In Exercise 6 find

(a) The component in the second row and second column of AC. [*Ans.* 40.]
(b) The component in the fourth row and first column of BC.
(c) The component in the last row and last column of DA. [*Ans.* 58.]
(d) The component in the first row and first column of CB.

8. If A is a square matrix, it can be multiplied by itself; hence we can define (using the associative law)

$$\begin{aligned} A^2 &= A \cdot A \\ A^3 &= A^2 \cdot A = A \cdot A \cdot A \\ &\dots \\ A^n &= A^{n-1} \cdot A = A \cdot A \cdot \ldots A \qquad (n \text{ factors}). \end{aligned}$$

These are naturally called "powers" of a matrix—the first one being called the square, the second, the cube, etc. Compute the indicated powers of the following matrices.

(a) If $A = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}$, find A^2, A^3, and A^4.

[*Ans.* $\begin{pmatrix} 1 & 0 \\ 15 & 16 \end{pmatrix}$; $\begin{pmatrix} 1 & 0 \\ 63 & 64 \end{pmatrix}$; $\begin{pmatrix} 1 & 0 \\ 255 & 256 \end{pmatrix}$.]

(b) If I and 0 are the matrices defined in Exercises 2 and 3, find I^2, I^3, I^n, 0^2, 0^3, and 0^n.

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

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

9. 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 IV, Section 13, and comment on the result.

10. Consider a two-stage 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 $P^{(3)} = P^3$.

11. Show that the fifth and all higher powers of the matrix

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

have all entries positive. Show that no smaller power has this property.

12. 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.

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

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.]

13. A college survey at an all-male school shows that dates of students are distributed as follows: a freshman dates one blonde and one brunette during the year; each sophomore dates one blonde, three brunettes, and one redhead; each junior dates three blondes, two brunettes, and two redheads; each senior dates three redheads. It is further known that each blonde brings three dresses with her, two skirts, two blouses, and one sweater; each brunette brings five dresses, four skirts, one blouse, and three sweaters; each redhead brings one dress, four skirts, and four sweaters. If each dress costs \$50, each skirt \$15, each blouse \$10, and each sweater \$5; and if there are 500 freshmen, 400 sophomores, 300 juniors, and 200 seniors,

(a) What is the total number of blondes, brunettes, and redheads dated?

(b) What is the total number of each type of clothing item in the dates' wardrobes?

(c) What is the cost of the wardrobe of a blonde? A brunette? A redhead?

(d) What is the total cost of all the wardrobes of all the dates? Calculate two ways. [*Ans.* \$1,347,500.]

SUPPLEMENTARY EXERCISES

14. Find three different 2×2 matrices A such that $A^2 = I$.

15. 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 the fact that it is true for ordinary numbers.

16. 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.

17. 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.

18. A *diagonal matrix* is square and its only nonzero entries are on the main diagonal. For instance, the matrices

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

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.

19. Consider the matrices

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

(a) Show that $A^2 = B$ and $A^3 = I$. What is A^4?

(b) Show that $B^2 = A$ and $B^3 = I$.

(c) Show that $A^3 = BA = AB = B^3 = I$, hence A and B commute.

20. For the matrix

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

what is the smallest k such that $A^k = I$?

21. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

(a) Find a matrix B such that $AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

(b) Find a matrix D such that $AD = \begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix}$.

[*Ans.* (a) $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$; (b) $\begin{pmatrix} -2 & -3 \\ 4 & 3 \end{pmatrix}$.]

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.

$$x_1 + 4x_2 + 3x_3 = 1 \tag{1}$$

$$2x_1 + 5x_2 + 4x_3 = 4 \tag{2}$$

$$x_1 - 3x_2 - 2x_3 = 5. \tag{3}$$

Equations such as these, that contain 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 their conjunction

$$\begin{aligned} (x_1 + 4x_2 + 3x_3 = 1) \wedge (2x_1 + 5x_2 + 4x_3 = 4) \\ \wedge (x_1 - 3x_2 - 2x_3 = 5) \end{aligned}$$

true. 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 equations 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

$$
\begin{aligned}
&(1') && x_1 + 4x_2 + 3x_3 = 1\\
&(2') && -3x_2 - 2x_3 = 2\\
&(3') && -7x_2 - 5x_3 = 4.
\end{aligned}
$$

Next we 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

$$
\begin{aligned}
&(1'') && x_1 + 0 + \tfrac{1}{3}x_3 = \tfrac{11}{3}\\
&(2'') && x_2 + \tfrac{2}{3}x_3 = -\tfrac{2}{3}\\
&(3'') && -\tfrac{1}{3}x_3 = -\tfrac{2}{3}.
\end{aligned}
$$

The last step is to divide 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:

$$
\begin{aligned}
&(1''') && x_1 + 0 + 0 = 3\\
&(2''') && x_2 + 0 = -2\\
&(3''') && x_3 = 2.
\end{aligned}
$$

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.

$$\begin{aligned} &(4) && x_1 - 2x_2 - 3x_3 = 2 \\ &(5) && x_1 - 4x_2 - 13x_3 = 14 \\ &(6) && -3x_1 + 5x_2 + 4x_3 = 0. \end{aligned}$$

Let us proceed as before and use equation (4) to eliminate the variable x_1 from the other two equations. We have

$$\begin{aligned} &(4') && x_1 - 2x_2 - 3x_3 = 2 \\ &(5') && -2x_2 - 10x_3 = 12 \\ &(6') && -x_2 - 5x_3 = 6. \end{aligned}$$

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′).

$$\begin{aligned} &(4'') && x_1 + 0 + 7x_3 = -10 \\ &(5'') && x_2 + 5x_3 = -6 \\ &(6'') && 0 = 0. \end{aligned}$$

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

$$\begin{aligned} &(4''') && x_1 = -10 - 7x_3 \\ &(5''') && x_2 = -\ 6 - 5x_3. \end{aligned}$$

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.

$$\begin{aligned} x_1 &= -17, & x_2 &= -11, & x_3 &= 1 \\ x_1 &= -10, & x_2 &= -6, & x_3 &= 0 \\ x_1 &= 4, & x_2 &= 4, & x_3 &= -2. \end{aligned}$$

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

$$x_1 - 2x_2 - 3x_3 = 2 \tag{7}$$

$$x_1 - 4x_2 - 13x_3 = 14 \tag{8}$$

$$-3x_1 + 5x_2 + 4x_3 = 2. \tag{9}$$

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

$$x_1 - 2x_2 - 3x_3 = 2 \tag{7'}$$

$$-2x_2 - 10x_3 = 12 \tag{8'}$$

$$-x_2 - 5x_3 = 8. \tag{9'}$$

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

$$x_1 + 0 + 7x_3 = -10 \tag{7''}$$

$$x_2 + 5x_3 = -6 \tag{8''}$$

$$0 = 2. \tag{9''}$$

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.

$$-4x_1 + 3x_2 + 2x_3 = -2 \tag{10}$$

$$5x_1 - 4x_2 + x_3 = 3. \tag{11}$$

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

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

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

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

(13) $$5x_1 - 4x_2 = 0$$

(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′) $$x_1 - \tfrac{3}{4}x_2 = -\tfrac{1}{2}$$

(13′) $$-\tfrac{1}{4}x_2 = \tfrac{5}{2}$$

(14′) $$\tfrac{1}{2}x_2 = a + 1.$$

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 will 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

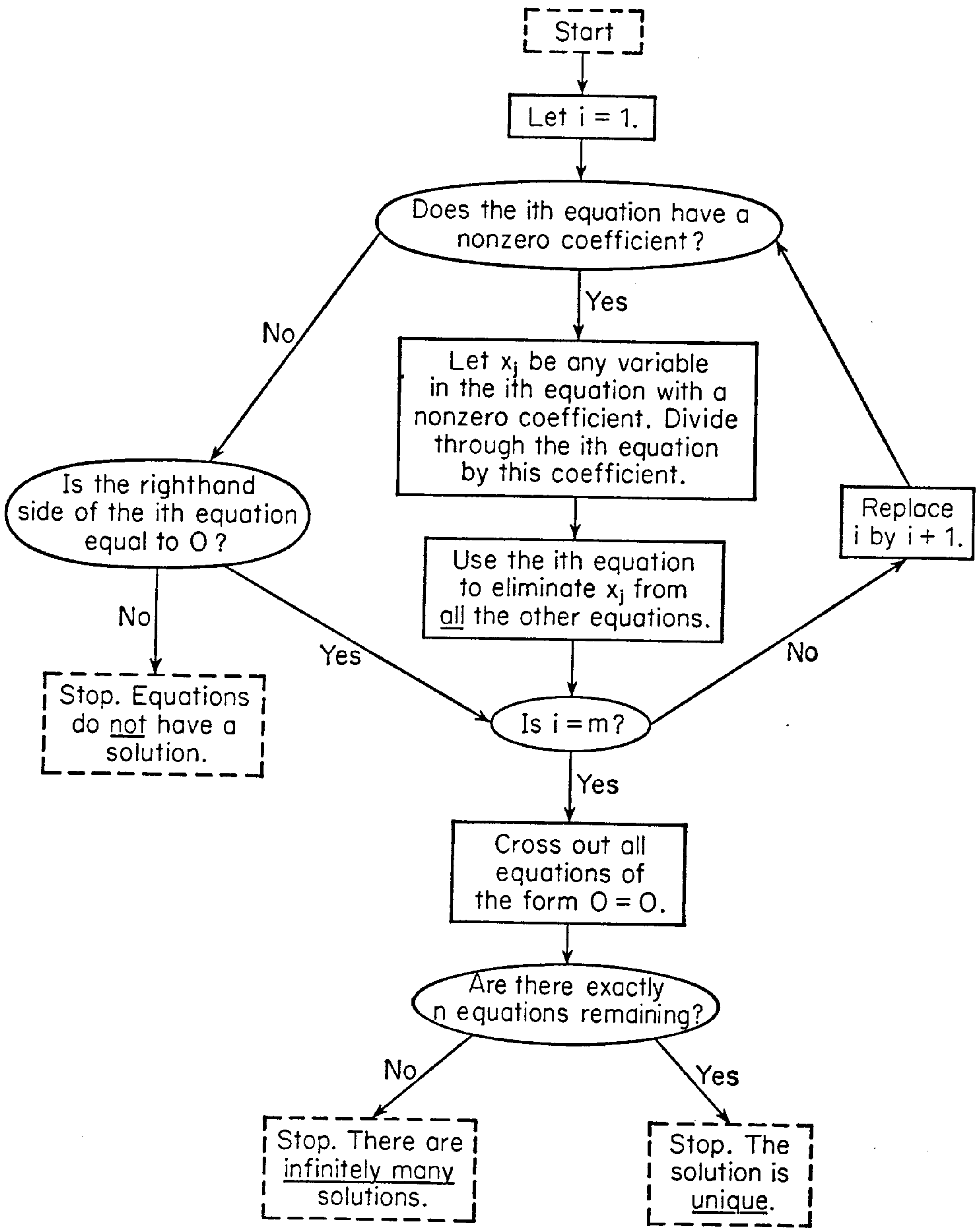


Figure 5 *Flow diagram for solving m equations in n variables.*

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 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 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*

$$(15) \qquad \left(\begin{array}{rrr|r} 1 & 4 & 3 & 1 \\ 2 & 5 & 4 & 4 \\ 1 & -3 & -2 & 5 \end{array}\right).$$

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.

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

$$\begin{pmatrix} 1 & 4 & 3 & \big| & 1 \\ 0 & -3 & -2 & \big| & 2 \\ 0 & -7 & -5 & \big| & 4 \end{pmatrix} \tag{16}$$

$$\begin{pmatrix} 1 & 0 & \frac{1}{3} & \big| & \frac{11}{3} \\ 0 & 1 & \frac{2}{3} & \big| & -\frac{2}{3} \\ 0 & 0 & -\frac{1}{3} & \big| & -\frac{2}{3} \end{pmatrix} \tag{17}$$

$$\begin{pmatrix} 1 & 0 & 0 & \big| & 3 \\ 0 & 1 & 0 & \big| & -2 \\ 0 & 0 & 1 & \big| & 2 \end{pmatrix}. \tag{18}$$

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:

(1), (2), and (3)	correspond to	(15)
(1′), (2′), and (3′)	correspond to	(16)
(1″), (2″), and (3″)	correspond to	(17)
(1‴), (2‴), and (3‴)	correspond to	(18)

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

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

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

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.

$$\begin{pmatrix} 1 & 4 & 3 & \big| & 1 & -1 \\ 2 & 5 & 4 & \big| & 4 & 0 \\ 1 & -3 & -2 & \big| & 5 & 2 \end{pmatrix} \tag{20}$$

$$\begin{pmatrix} 1 & 4 & 3 & \big| & 1 & -1 \\ 0 & -3 & -2 & \big| & 2 & 2 \\ 0 & -7 & -5 & \big| & 4 & 3 \end{pmatrix} \tag{21}$$

$$(22)\qquad \left(\begin{array}{ccc|cc} 1 & 0 & \frac{1}{3} & \frac{11}{3} & \frac{5}{3} \\ 0 & 1 & \frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} \\ 0 & 0 & -\frac{1}{3} & -\frac{2}{3} & -\frac{5}{3} \end{array}\right)$$

$$(23)\qquad \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & -2 & -4 \\ 0 & 0 & 1 & 2 & 5 \end{array}\right).$$

We find the answers

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

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. Work again Examples 2–4 using the detached coefficient tableau.

2. Find all the solutions of the following simultaneous equations.

(a)
$$\begin{aligned} 4x_1 \phantom{{}-x_2} + 5x_3 &= 6 \\ x_2 - 6x_3 &= -2 \\ 3x_1 \phantom{{}-x_2} + 4x_3 &= 3. \end{aligned}$$
[*Ans.* $x_1 = 9, x_2 = -38, x_3 = -6$.]

(b)
$$\begin{aligned} 3x_1 - x_2 - 2x_3 &= 2 \\ 2x_2 - x_3 &= -1 \\ 3x_1 - 5x_2 \phantom{{}-x_3} &= 3. \end{aligned}$$
[*Ans.* No solution.]

(c)
$$\begin{aligned} -x_1 + 2x_2 + 3x_3 &= 0 \\ x_1 - 4x_2 - 13x_3 &= 0 \\ -3x_1 + 5x_2 + 4x_3 &= 0. \end{aligned}$$
[*Ans.* $x_1 = -7x_3, x_2 = -5x_3$.]

3. Find all the solutions of the following simultaneous equations.

(a)
$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ 2x_1 + 4x_2 + 3x_3 &= 0 \\ 4x_2 + 4x_3 &= 0. \end{aligned}$$

(b)
$$\begin{aligned} x_1 + x_2 + x_3 &= -2 \\ 2x_1 + 4x_2 + 3x_3 &= 3 \\ 4x_2 + 2x_3 &= 2. \end{aligned}$$

(c)
$$\begin{aligned} 4x_1 \phantom{{}+x_2} + 4x_3 &= 8 \\ x_2 - 6x_3 &= -3 \\ 3x_1 + x_2 - 3x_3 &= 3. \end{aligned}$$

4. Find all solutions of the following equations using the detached coefficient tableau.

(a)
$$\begin{aligned} 5x_1 - 3x_2 &= -7 \\ -2x_1 + 9x_2 &= 4 \\ 2x_1 + 4x_2 &= -2. \end{aligned}$$
[*Ans.* $x_1 = -\frac{17}{13}$; $x_2 = \frac{2}{13}$.]

(b)
$$\begin{aligned} x_1 + 2x_2 &= 1 \\ -3x_1 + 2x_2 &= -2 \\ 2x_1 + 3x_2 &= 1. \end{aligned}$$
[*Ans.* No solution.]

(c)
$$\begin{aligned} 5x_1 - 3x_2 - 7x_3 + x_4 &= 10 \\ -x_1 + 2x_2 + 6x_3 - 3x_4 &= -3 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 0. \end{aligned}$$

5. Find all solutions of:
$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 10 \\ 2x_1 - x_2 + x_3 - x_4 &= 1 \\ 3x_1 + x_2 + 4x_3 + 3x_4 &= 11 \\ -2x_1 + 6x_2 + 4x_3 + 10x_4 &= 18 \end{aligned}$$
[*Ans.* $x_1 = \frac{12}{5} - x_3 - \frac{2}{5}x_4$; $x_2 = \frac{19}{5} - x_3 - \frac{9}{5}x_4$, x_3 and x_4 arbitrary.]

6. We consider buying three kinds of food. Food I has one unit of vitamin A, three units of vitamin B, and four units of vitamin C. Food II has two, three, and five units, respectively. Food III has three units each of vitamin A and vitamin C, none of vitamin B. We need to have 11 units of vitamin A, nine of vitamin B, and 20 of vitamin C.

(a) Find all possible amounts of the three foods that will provide precisely these amounts of the vitamins.

(b) If food I costs 60 cents and the others cost 10 cents each per unit, is there a solution costing exactly \$1? [*Ans.* (b) Yes; 1, 2, 2.]

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

$$\begin{array}{rcccc} & \text{(a)} & \text{(b)} & \text{(c)} & \text{(d)} \\ 4x_1 \qquad + 5x_3 = & 1 & 1 & 0 & 0 \\ x_2 - 6x_3 = & 2 & 0 & 0 & 1 \\ 3x_1 \qquad + 4x_3 = & 3 & 0 & 1 & 0. \end{array}$$
[*Ans.* (a) $x_1 = -11$, $x_2 = 56$, $x_3 = 9$.]

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

$$\begin{array}{rcccc} & \text{(a)} & \text{(b)} & \text{(c)} & \text{(d)} \\ x_1 + x_2 + x_3 = & 3 & 0 & 12 & 0 \\ x_1 - x_2 + 2x_3 = & 2 & -1 & 7 & 0 \\ 2x_1 + x_2 - x_3 = & 2 & 3 & 11 & 0. \end{array}$$

9. Solve the following three sets of simultaneous equations.

$$\begin{array}{rcrrr} & & \text{(a)} & \text{(b)} & \text{(c)} \\ x_1 + x_2 + x_3 & = & 1 & 2 & 0 \\ x_1 - x_2 + 2x_3 & = & -2 & 2 & 0 \\ 3x_1 - x_2 + 5x_3 & = & -3 & 2 & 0. \end{array}$$

10. Show that the equations

$$\begin{aligned} -4x_1 + 3x_2 + ax_3 &= c \\ 5x_1 - 4x_2 + bx_3 &= d \end{aligned}$$

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

$$\begin{aligned} -4x_1 + 3x_2 &= a \\ 5x_1 - 4x_2 &= b \\ -3x_1 + 2x_2 &= c \end{aligned}$$

have a solution. [*Ans.* $2a + b = c$.]

12. (a) Let $x = (x_1, x_2)$ and let A be the matrix

$$A = \begin{pmatrix} 3 & -4 \\ 2 & -6 \end{pmatrix}.$$

Find all solutions of the equation $xA = x$. [*Ans.* $x = (0, 0)$.]

(b) Let $x = (x_1, x_2)$ and let A be the matrix

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

Find all solutions of the equation $xA = x$.

[*Ans.* $x = (k, k)$ for any number k.]

13. Let $x = (x_1, x_2)$ and let P be the matrix

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

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

14. If $x = (x_1, x_2, x_3)$ and A is the matrix

$$A = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 5 & 4 \\ 0 & -6 & -4 \end{pmatrix},$$

find all solutions of the equation $xA = x$.

[*Ans.* $x = (-k/2, 5k/4, k)$ for any number k.]

15. If $x = (x_1, x_2, x_3)$ and P is the matrix

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

find all solutions of the equation $xP = x$. Select the unique solution for which $x_1 + x_2 + x_3 = 1$.

16. (a) Show that the simultaneous linear equations

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

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

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

(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.

17. (a) Show that the equations of Exercise 16(a) can be interpreted as a row-vector-times-matrix equation of the form

$$(x_1 \quad x_2 \quad x_3) \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = (1 \quad 0).$$

(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.

18. (a) Show that the simultaneous linear equations of Exercise 16(a) can be interpreted as asking for all possible ways of expressing the column vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in terms of the column vectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, and $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

(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.

SUPPLEMENTARY EXERCISES

19. 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?

$$\begin{aligned} 2x \quad\quad\;\; + 4z &= 6 \\ 3x + y + \;\; z &= -1 \\ 2y - \;\; z &= -2 \\ x - y + kz &= -5. \end{aligned}$$

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

20. Consider the following set of simultaneous equations.

$$\begin{array}{lllll} x_1 & & y_1 + & & = a \\ x_1 & & & + y_2 & = b \\ & x_2 + y_1 & & & = c \\ & x_2 & & + y_2 & = d. \end{array}$$

(a) For what conditions on a, b, c, and d will these equations have a solution?

(b) Give a set of values 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.

21. 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.]

22. Let

$$A = (a, b, c), \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

and let d be any number. Consider the open statement $Ax = d$.

(a) If $A \neq 0$, show that the truth set of $Ax = d$ is not empty.

(b) If $A = 0$ and $d = 0$, show that $Ax = d$ is logically true.
(c) If $A = 0$ and $d \neq 0$, show that $Ax = d$ is logically false.
(d) Use (a), (b), and (c) to prove the following theorem: A single open statement $Ax = d$ is logically false if and only if $A = 0$ and $d \neq 0$.

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.

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

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

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

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

It is easy to see that a square matrix can have only one inverse. 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 non-zero number has a reciprocal, but there are matrices, not the zero matrix, which have no inverse. For example, if

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

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 it follows 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,

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

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

$$(3) \qquad \left(\begin{array}{ccc|ccc} 4 & 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 3 & 0 & 4 & 0 & 0 & 1 \end{array}\right)$$

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

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{5}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{3}{4} & 0 & 1 \end{array}\right). \tag{4}$$

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

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & 0 & -5 \\ 0 & 1 & 0 & -18 & 1 & 24 \\ 0 & 0 & 1 & -3 & 0 & 4 \end{array}\right). \tag{5}$$

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

$$(A \mid I)$$

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

$$(I \mid B).$$

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:

$$(6)\qquad \left(\begin{array}{ccc|ccc} 4 & 0 & 8 & 1 & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 2 & 0 & 4 & 0 & 0 & 1 \end{array}\right).$$

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

$$(7)\qquad \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & \frac{1}{4} & 0 & 0 \\ 0 & 1 & -6 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 1 \end{array}\right).$$

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,

$$Ax = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad Ax = \begin{pmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad \ldots, \quad Ax = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix}$$

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

$$(A \mid I)$$

can be transformed by row transformations into the tableau

$$(I \mid A^{-1}).$$

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

$$\left(\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 2 & 5 & 4 & 0 & 1 & 0 \\ 1 & -3 & -2 & 0 & 0 & 1 \end{array}\right).$$

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

$$\left(\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & -3 & -2 & -2 & 1 & 0 \\ 0 & -7 & -5 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{3} & -\frac{5}{3} & \frac{4}{3} & 0 \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} & \frac{11}{3} & -\frac{7}{3} & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & 8 & -5 & 2 \\ 0 & 0 & 1 & -11 & 7 & -3 \end{array}\right).$$

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

EXERCISES

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

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 5 \\ -2 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & 3 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 7 \end{pmatrix},$$

$$C = \begin{pmatrix} 9 & -1 & 0 & 0 \\ 0 & 8 & -2 & 0 \\ 0 & 0 & 7 & -3 \\ 0 & 0 & 0 & 6 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 4 & 0 \\ \frac{1}{2} & 3 & 2 \end{pmatrix}.$$

[*Partial Ans.* $A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -13 & 1 & -5 \\ 2 & 0 & 1 \end{pmatrix}$; $D^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{12} & \frac{1}{4} & 0 \\ -\frac{1}{8} & -\frac{3}{8} & \frac{1}{2} \end{pmatrix}$.]

2. Show that each of the following matrices fails to have an inverse.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 0 \\ 0 & 3 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 5 \\ -1 & 1 & -5 \end{pmatrix},$$

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

3. Let A, B, and D be the matrices of Exercise 1; let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{and} \quad w = (w_1, w_2, w_3);$$

let b, c, d, e, and f be the following vectors.

$$b = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}, \quad c = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}, \quad d = (3, 7, -2), \quad e = (1, 1, 1), \quad f = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Use the inverses you computed in Exercise 1 to solve the following equations.

(a) $Ax = b$. (b) $Bx = c$. (c) $wD = e$.
(d) $wB = d$. (e) $wA = e$. (f) $Dx = f$.

[*Partial Ans.* (a) $x = \begin{pmatrix} 3 \\ -40 \\ 6 \end{pmatrix}$; (e) $w = (-10, 1, -4)$; (f) $x = \begin{pmatrix} 1 \\ \frac{1}{6} \\ 0 \end{pmatrix}$.]

4. Rework Exercise 7 of Section 5 by first writing the equations in the form $Ax = b$, and finding the inverse of A.

5. Solve the following problem by first inverting the matrix. (Assume $ad \neq bc$.) If a grinding machine is supplied with 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?

6. For each of the matrices A and D in Exercise 2 find a nonzero vector whose product with the given matrix is 0.

7. Show that if A has no inverse, then neither does any of its positive powers A^k.

8. 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.

9. Expand the formula $(AB)^{-1} = B^{-1}A^{-1}$ into a two-part statement analogous to the one in the exercise above. Then prove both parts of your statement.

10. (a) Show that $(AB)^{-1} \neq A^{-1}B^{-1}$ for the matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$.

(b) Find $(AB)^{-1}$ in two different ways. [*Hint:* Use Exercise 9.]

11. Give a criterion for deciding whether the 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has an inverse. [*Ans.* $ad \neq bc$.]

12. Give a formula for $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$, when it exists.

13. 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?

SUPPLEMENTARY EXERCISES

14. Let A be the matrix $\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$.

(a) Find A^{-1}.

(b) Use the result of (a) to solve the matrix equation $A^2x = b$, where

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

[*Ans.* $x = \begin{pmatrix} 36 \\ 13 \end{pmatrix}$.]

15. Let A be a square matrix that has an inverse. Show that the inverse of A^2 is $(A^{-1})^2$. What is the inverse of A^n?

Note: The remaining exercises refer to the problem of computing $(I - Q)^{-1}$ where Q is a lower triangular matrix.

16. A matrix is *lower triangular* if it has zeros on and above its main diagonal. For instance,

$$Q = \begin{pmatrix} 0 & 0 & 0 \\ 4 & 0 & 0 \\ 10 & 5 & 0 \end{pmatrix}$$

is lower triangular.

(a) Compute Q^2, saving the result for later exercises.
(b) Show that $Q^3 = 0$, and also that $Q^k = 0$ for $k \geq 3$.

17. Consider the equation $w = wQ + d$ where Q is as in Exercise 16, and

$$w = (w_1, w_2, w_3), \qquad d = (20, 5, 3).$$

Solve symbolically for w. [*Ans.* $w = d(I - Q)^{-1}$.]

18. (a) Establish the identity

$$(I - Q)(I + Q + Q^2) = I - Q^3 = I$$

where Q is as in Exercise 16.
(b) Show from (a) that $(I - Q)^{-1} = I + Q + Q^2$.
(c) Use (b) to compute $(I - Q)^{-1}$.

[*Ans.* $(I - Q)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 30 & 5 & 1 \end{pmatrix}$.]

(d) Use (c) to solve for the w of Exercise 17. [*Ans.* $w = (130, 20, 3)$.]

19. Let Q be any $n \times n$ lower triangular matrix.
(a) Show that $Q^k = 0$ for $k \geq n$.
(b) Show that $(I - Q)(I + Q + \ldots + Q^{n-1}) = I - Q^n = I$.
(c) Show that $(I - Q)^{-1} = I + Q + \ldots + Q^{n-1}$.

20. Find $(I - Q)^{-1}$ for Q being each of the following.

(a) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 3 & 4 & 5 & 0 \end{pmatrix}$. (b) $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 & 0 \\ 0 & 1 & 8 & 0 & 0 \\ 10 & 5 & 0 & 3 & 0 \end{pmatrix}$.

[*Ans.* (a) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 7 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 41 & 4 & 5 & 1 \end{pmatrix}$.]

21. Use Exercise 19 to show that if Q is lower triangular and has non-negative *integer* entries, then so does the matrix $(I - Q)^{-1}$.

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 two-state and three-state Markov chains, but a similar procedure will work for any other Markov chain.

In Section 13 of Chapter IV, we noted that to each Markov chain there was a matrix of transition probabilities. For instance, if there are three states, a_1, a_2, and a_3, then

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

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 10 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.

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

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 above equations as

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

If we substitute values of n we get the equations $p^{(1)} = p^{(0)}P$, $p^{(2)} = p^{(1)}P = p^{(0)}P^2$, $p^{(3)} = p^{(2)}P = p^{(0)}P^3$, etc. In general, it can be seen that

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

Thus we see that, if we multiply the vector $p^{(0)}$ of initial probabilities by the nth 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 13 of Chapter IV, we considered special Markov chains that started in given fixed states. There we arrived at a matrix $P^{(n)}$ whose ith 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 nth power of P, that is, $P^{(n)} = P^n$. (Compare Exercise 10 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 a 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 (continued). Suppose that we compute powers of the matrix P in the example above. We have

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

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

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

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 a 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

(a) The powers P^n approach a matrix W.
(b) Each row of W is the same probability vector w.
(c) 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 vector w. This proves part (a).

Since the powers of P approach W, $P^{n+1} = P^nP$ 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) 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).

The following is an important consequence of this 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. Hence the Markov chain approaches equilibrium.

We noted for an independent trials process that if p is the probability of a given outcome d, 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

* For an elementary proof see Kemeny, Mirkil, Snell, and Thompson, *Finite Mathematical Structures* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1959) Chapter 6, Section 3.

approximation gets better and better as the number of experiments increases. For a regular Markov chain it is the components of the vector w that play the analogous role. That is, the fraction of times that the chain is in state a_i approaches w_i, no matter how the process is started.

Example 1 (continued). 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

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

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

$$\begin{aligned} w_1(1-a) + w_2 b &= w_1 \\ w_1 a + w_2(1-b) &= w_2. \end{aligned}$$

Each of these equations reduces to the single equation $w_1 a = w_2 b$. 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 .he unique fixed-point probability vector of S.

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, etc., 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

$$\begin{array}{c c} & \begin{array}{cc} \text{yes} & \text{no} \end{array} \\ \begin{array}{c} \text{yes} \\ \text{no} \end{array} & \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \end{array}.$$

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 section, we find that $w = (\frac{1}{2}, \frac{1}{2})$. Hence the probabilities for the nth man 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

$$\begin{array}{c c} & \begin{array}{cc} \text{yes} & \text{no} \end{array} \\ \begin{array}{c} \text{yes} \\ \text{no} \end{array} & \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \end{array}.$$

In this case $w = [b/(a + b), a/(a + b)]$. Thus there is a probability of approximately $b/(a + b)$ that the nth 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 IV, Section 13. 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

$$\begin{array}{c c}
 & \begin{array}{cccc} RR & DR & RD & DD \end{array} \\
\begin{array}{c} RR \\ DR \\ RD \\ DD \end{array} & \begin{pmatrix} 1-a & 0 & a & 0 \\ b & 0 & 1-b & 0 \\ 0 & 1-c & 0 & c \\ 0 & d & 0 & 1-d \end{pmatrix}.
\end{array}$$

If a, b, c, and d are all different from 0 or 1, then the square of the matrix has no zeros, and hence the matrix is regular. The fixed probability vector is found in the usual way (see Exercise 18) and is

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

Note that the probability of being in state DR after a large number of steps is equal to the probability of being in state RD. 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

1. 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

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

is a regular transition matrix if and only if either

(i) $0 < a \leq 1$ and $0 < b < 1$; or
(ii) $0 < a < 1$ and $0 < b \leq 1$.

3. 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.]

4. Find the fixed point w for each of the following regular matrices.

(a) $\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$. [*Ans.* $w = (\frac{2}{3}, \frac{1}{3})$.]

(b) $\begin{pmatrix} .9 & .1 \\ .1 & .9 \end{pmatrix}$.

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

5. Let $p^0 = (\frac{1}{2}, \frac{1}{2})$ and compute $p^{(1)}$, $p^{(2)}$, and $p^{(3)}$ for the matrices in Exercises 4(a) and 4(b). Do they approach the fixed points of these matrices?

6. Give a probability theory interpretation to the condition of regularity.

7. Consider the two-state Markov chain with transition matrix

$$P = \begin{array}{c c} & \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}.$$

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?

8. 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.

9. Compute the first five powers of the matrix

$$P = \begin{pmatrix} .8 & .2 \\ .2 & .8 \end{pmatrix}.$$

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

10. Show that all transition matrices of the form

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

where $0 < a < 1$, have the same unique fixed point. [*Ans.* $w = (\frac{1}{2}, \frac{1}{2})$.]

11. 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 one last time, he tosses a coin, and uses question two if a head comes up. If he used question two, he tosses two coins and switches to question three if both come up heads. If he used question three, he tosses three coins and switches to question one if all three come up heads. In the long run, which question does he use most often, and how frequently is it used? [*Ans.* Question two, 40 per cent of the time.]

12. A professor tries not to be late for class too often. If he is late one day, he is 90 per cent sure to be on time next time. If he is on time, then the next day there is a 30 per cent chance of his being late. In the long run, how often is he late for class?

13. 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, $\frac{2}{5}$.]

14. Let S be the matrix

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

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

15. Show that the matrix

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

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

16. 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.

17. Consider a Markov chain such that it is possible to go from any state a_i to any state a_j 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_j via a_k.]

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

SUPPLEMENTARY EXERCISES

19. Determine whether each of the following matrices is regular.

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

(b) $\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \end{pmatrix}$.

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

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

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

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

[*Ans.* (a) and (d) are regular.]

20. Consider the three-state Markov chain with transition matrix

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

(a) Show that the matrix has a unique fixed probability vector.

[*Ans.* $(\frac{1}{2}, \frac{5}{12}, \frac{1}{12})$.]

(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)?

21. Assume that it is known that of the sons of Harvard alumni, 80 per cent go to Harvard and all the rest go to Yale; of the sons of Yale men, 40 per cent go to Yale, the remainder split evenly between Harvard and Dartmouth; and of the sons of Dartmouth men, 70 per cent go to Dartmouth, 20 per cent to Harvard, and 10 per cent to Yale.

(a) Set up this process as a Markov chain.

(b) What is the probability that the grandson of a Harvard man goes to Harvard?

(c) What is the long run fraction expected in each school?

[*Ans.* (b) .7; (c) (Harvard, $\frac{5}{9}$; Yale, $\frac{2}{9}$; Dartmouth, $\frac{2}{9}$.)]

22. 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.]

23. 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.* (c) Increase, $\frac{1}{3}$; decrease, $\frac{1}{6}$; same, $\frac{1}{2}$. (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).

Example 1. A particle moves on a line; each time it moves one unit to the right with probability $\frac{1}{2}$, 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,

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

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, etc. 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.*

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

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

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

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, i.e., $Q^n \to 0$.

From the fact that $Q^n \to 0$ it can be shown that the matrix $(I - Q)^{-1}$ exists.* The matrix $(I - Q)^{-1}$ will be called the *fundamental matrix* of the absorbing chain. It has the following important interpretation.

Let n_{ij} be the expected number of times that the chain is in state a_j if it starts in state a_i, for two nonabsorbing states a_i and a_j. Let N be the matrix whose components are n_{ij}. 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

* For a proof see Kemeny, Mirkil, Snell, and Thompson, *Finite Mathematical Structures* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1959) Chapter 6, Section 3.

nonabsorbing states.) This equation may be written in matrix form:

$$N = I + QN.$$

Thus $(I - Q)N = I$, and hence $N = (I - Q)^{-1}$, as was to be shown. Thus we have found a probabilistic interpretation for our fundamental matrix: its i,jth entry is the expected number of times that the chain is in state a_j if it starts at a_i. 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 expected number of times in each nonabsorbing state for each possible nonabsorbing starting state.

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

$$\begin{array}{c} \\ 0 \\ 4 \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{ccccc} 0 & 4 & 1 & 2 & 3 \end{array} \\ \left(\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{array}\right) \end{array}.$$

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

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

Thus, starting at state 2, the expected 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 will have the expected number of times in any of the nonabsorbing states for a given starting state, that is, the expected 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 expected number of steps before being absorbed, for each possible nonabsorbing starting state.

Example 1 (continued). For Example 1 we have

$$t = Nc = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{pmatrix} \end{array} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix}.$$

Thus the expected 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_j 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. State a_j can be reached either by stepping into it on the first step, or by going to a nonabsorbing state a_k and from there eventually reaching a_j. Hence, 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

$$B = R + QB$$

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

and hence

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

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

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

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 expected number of times in each state before absorption, depending upon the starting state. The column vector $t = Nc$ gives us the expected number of steps before absorption, depending 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

1. Which of the following transition matrices are from absorbing chains?

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

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

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

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

[*Ans.* (a) and (d).]

2. Consider the two-state transition matrix

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

For what choices of a and b do we 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 $\frac{2}{3}$ and a step to the left is $\frac{1}{3}$. Find N, t, and B. Compare these with the results for probability $\frac{1}{2}$ for a step to the right and $\frac{1}{2}$ for a step to the left.

4. In the Land of Oz example (see Exercise 13, Section 7) let us change the transition matrix by making R an absorbing state. This gives

$$\begin{array}{c c} & \begin{array}{ccc} R & N & S \end{array} \\ \begin{array}{c} R \\ N \\ S \end{array} & \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}. \end{array}$$

Find the fundamental matrix N, and also t and B. What is the interpretation of these quantities?

5. 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.* $\frac{4}{7}$.]

6. 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

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

where I is the 4×4 identity matrix. What is the expected number of draws?

[*Ans.* $\frac{37}{12}$.]

7. Using the result of Exercise 6, make a conjecture for the form of the fundamental matrix if we start with integers from 1 to n. What would the expected number of draws be if we started with numbers from 1 to 10?

8. 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.

9. The following is an alternative method of finding the probability of absorption in a particular absorbing 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_j if the process starts in a_i. Use this method to find the probability of absorption in state 1 in the random walk example given in this section.

10. The following is an alternative method for finding the expected number of steps to absorption. Let t_i be the expected number of steps to absorption starting at state a_i. This must be the same as taking one more step and then adding $p_{ij}t_j$ for every nonabsorbing state a_j.

(a) Give reasons for the above claim 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.

SUPPLEMENTARY EXERCISES

11. Peter and Paul are matching pennies, and each player flips his (fair) coin before revealing it. They initially have three pennies between them and the game ends whenever one of them has all the pennies. Let the states be labelled 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?

12. Peter and Paul are matching pennies as in Exercise 11, except that whenever one of the players gets all three pennies, he returns one to his opponent, and the game continues.

(a) Set up the transition matrix.

(b) Identify the kind of Markov chain that results. [*Ans.* Regular.]

(c) Find the long run probabilities of being in each of the states.

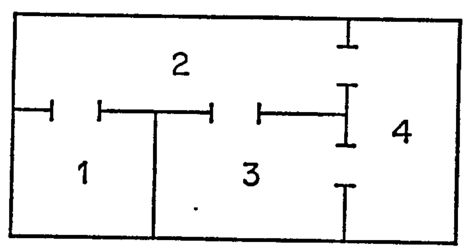
13. Peter and Paul are matching pennies as in Exercise 11, except that if Peter gets all the pennies, the game is over, while if Paul gets all the pennies, he gives one back to Peter, and the game continues.

(a) Set up the transition matrix.

(b) Identify the resulting Markov chain. [*Ans.* Absorbing.]

(c) If Peter initially has one penny, what is the probability of his winning the game? If he has two pennies?

14. 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) 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.]

(d) In part (c), if the rat starts in compartment 1, how many steps will it take him, on the average, to reach compartment 4?

[*Ans.* $4\frac{1}{3}$ steps.]

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, and there is a 10 per cent chance of his selling during a high-profit month, and a 40 per cent chance during a low-profit month. If he does not sell, with probability $\frac{2}{3}$ the profits will be the same the next month, and with probability $\frac{1}{3}$ they will change.

(a) Set up the transition matrix.

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

(b) Compute N, Nc, and NR and interpret each.

(c) Let $f = \begin{pmatrix} 5000 \\ 2000 \end{pmatrix}$ and compute the vector $g = Nf$.

$$[\textit{Ans.}\ g = \begin{pmatrix} 20{,}000 \\ 10{,}000 \end{pmatrix}.]$$

(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.

16. Suppose that P is the transition matrix of an absorbing Markov chain. Assume that each time the process is in a nonabsorbing state i, a reward f_i is received (including the starting state). Let f be the vector with components f_i and let $g = Nf$. Show that g_i is the expected winnings before absorption if the process starts in state i. [*Hint:* Exercise 15 is a specific example of this process.]

*9. *LINEAR FUNCTIONS AND TRANSFORMATIONS*

The primary use of vectors and matrices in science is the representation of several different quantities as a single one. For example, the demands on all the industries in the United States may be represented by a row vector x. We have seen examples where such a vector is multiplied by a column vector y, giving the number $x \cdot y$. The components of y could be the values of unit outputs of the various industries. Then $x \cdot y$ is the total monetary value of the demand on industries.

This illustration is typical of much that we meet in the sciences. It has two fundamental properties. If the demand increases by a given factor k, then $(kx)\cdot y = k(x\cdot y)$, and hence the value increases by the same factor. And if we have two demand vectors x and x', then $(x + x')\cdot y = (x\cdot y) + (x'\cdot y)$, and hence their values are also added.

Thus we see that y has the effect of assigning to each row vector x a number $f(x)$, and has the two very simple properties,

$$\text{(i)} \qquad f(kx) = kf(x)$$

$$\text{(ii)} \qquad f(x + x') = f(x) + f(x').$$

Such an assignment of a number to each row vector x we call a *linear function.* We have seen that each column vector with n components defines a linear function for row vectors with n components.

Linear functions represent the simplest type of dependence. Fortunately, very many problems can be represented at least approximately by linear functions. While it is not strictly true that manufacturing 100 tons of steel costs ten times as much as manufacturing ten tons, this is at least a reasonable approximation. And the same holds for necessary raw materials, for labor needed, transportation costs, etc. Linear functions are so simple to handle that we try to use them whenever this is reasonable.

Not only is it true that every column vector represents a linear function, but every linear function of row vectors can be so represented. We will prove this for linear functions of three-component row vectors.

Let us suppose that f assigns a number $f(x)$ to each three-component vector x, and that it has the properties (i) and (ii). Consider the three special vectors,

$$e_1 = (1, 0, 0), \qquad e_2 = (0, 1, 0), \qquad e_3 = (0, 0, 1).$$

Let us call $f(e_1) = y_1$; let $y_2 = f(e_2)$, $y_3 = f(e_3)$ and let $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$. If $x = (x_1, x_2, x_3)$, we can write $x = x_1e_1 + x_2e_2 + x_3e_3$. Hence, using properties (i) and (ii), we see that

$$\begin{aligned} f(x) &= f(x_1e_1 + x_2e_2 + x_3e_3) \\ &= f(x_1e_1) + f(x_2e_2) + f(x_3e_3) \\ &= x_1f(e_1) + x_2f(e_2) + x_3f(e_3) \\ &= x_1y_1 + x_2y_2 + x_3y_3 = x\cdot y. \end{aligned}$$

Hence the column vector y represents the linear function f.

Example 1. An office buys three kinds of paper, heavy bond, light bond, and a cheaper quality for intra-office use. The amounts bought (in reams) are given by the row vector $x = (20, 50, 70)$. The prices per ream of these types of paper are given (in cents) by the column vector $y = \begin{pmatrix}160\\140\\120\end{pmatrix}$. Then $f(x) = x \cdot y = \$186$ is the cost of the order. So far, y defines a linear function of x. It is customary to give a discount if 100 or more reams are ordered of one item. The new rules for computing the bill define a new function of x, different from f. Let us call the new function by the letter g. Then $g(2x) < 2g(x)$, since the office gets a discount on the light bond and on the cheaper paper. Now we have a function that is not linear. It often happens that a function in science is nearly linear for restricted values of the components, but not even roughly linear outside this range.

Sometimes we assign, not a single number to a row vector, but several numbers. Then we say that the vector is transformed into another vector. We say further that the transformation is a *linear transformation* of the vector if each component in the resulting vector is a linear function of the given vector, that is, it satisfies (i) and (ii).

Example 2. In Example 1 of Section 3 we considered a vector $x = (5, 7, 12)$, giving the number of each of three styles of houses to be built by a contractor, and a matrix

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

which gives the raw material requirements for each type of house. Suppose that $x' = (8, 2, 3)$ is another vector of house orders that are to be built in another location. Then it is easy to check that

$$\begin{aligned}(x + x')R &= (13, 9, 15)R = (218, 797, 436, 247, 605)\\ &= (146, 526, 260, 158, 388) + (72, 271, 176, 89, 217)\\ &= xR + x'R\end{aligned}$$

Similarly, if the contractor is going to produce $2x$ houses,

$$\begin{aligned}(2x)R = (10, 14, 24)R &= (292, 1052, 520, 316, 776)\\ &= 2(146, 526, 260, 158, 388).\end{aligned}$$

It can be shown in the same way that (i) and (ii) hold true in general and $f(x) = xR$ is a linear transformation of vectors x.

In the same manner (see Exercise 10) one can show that R is a linear transformation of five-component y (price) vectors.

Theorem. Let M be any $m \times n$ matrix; then M defines a linear transformation of m-component row vectors x, and it also defines a linear transformation of n-component column vectors y.

To prove this theorem we define $f(x) = xM$ and show, using the properties of ordinary numbers, that (i) and (ii) hold. This was done for a specific numerical example in Example 2 above. Similarly, we define $g(y) = My$ and show that (i) and (ii) hold.

It can be shown that the effect of any linear transformation can be described by a suitable matrix. This is illustrated in Example 3.

Example 3. Let us suppose that the population of the United States is divided into five groups according to income. The components of the row vector x are the number of people in each bracket. Say x_1 people have an income of \$100,000 or above, x_2 have incomes between \$40,000 and \$100,000, etc. If we know the average number of cars owned by men in a given income bracket, we can represent these five numbers as a column vector, and we get the number of privately owned cars as a linear function of x. Similarly, we could get the number of yachts, privately owned houses, or television sets. Each of these four quantities is a linear function of x (at least approximately) and each is represented by a five-component column vector whose entries are averages. Writing the four vectors together as a rectangular array, we get a 5×4 matrix. This is a linear transformation transforming x into a four-component row vector, whose components are the total number of cars, yachts, houses, and television sets, respectively.

EXERCISES

1. $x = (x_1, x_2, x_3)$. Test each of the following functions of x as to whether it has properties (i) and (ii).

(a) $f(x) = 3x_1 + x_2 - 2x_3$. [*Ans.* Linear.]

(b) $f(x) = x_1x_2x_3$.

(c) $f(x) = \sqrt{(x_1)^2 + (x_2)^2 + (x_3)^2}$. [*Ans.* Not linear.]

(d) $f(x) = x_2$.

2. $x = (x_1, x_2)$. Test each of the following transformations of x into y as to whether it is a linear transformation.

(a) $y_1 = 2x_1 + 3x_2$ and $y_2 = x_1 - x_2$. [*Ans.* Linear.]
(b) $y_1 = x_1 + 2x_2$ and $y_2 = -x_1x_2$. [*Ans.* Not linear.]
(c) $y_1 = x_2$ and $y_2 = -x_1$.

For the linear transformations above, write the matrix representing the transformation.

3. Prove that the function $f(x) = c$, where x is a two-component row vector and c is a constant, is a linear function if and only if $c = 0$.

4. Prove that the function $f(x) = ax_1 + bx_2 + c$, where x is a two-component row vector and a, b, and c are constants, is a linear function if and only if $c = 0$.

5. Prove that the transformation $T(x) = xA + C$, where x is a two-component row vector and A and C are 2×2 matrices, is a linear transformation if and only if $C = 0$.

6. Prove that $f(x) =$ (least component of x) is not a linear function.

7. Let x be a 12-component row vector. Its components are the enrollment figures in 12 mathematics courses. Give an example of

(a) A linear function of x.
[*An Ans.* The total enrollment in all mathematics courses.]
(b) A linear transformation of x.
(c) A nonlinear function of x.

8. Let the components of x be the number of fiction books, the number of nonfiction books, and the number of other publications in a library. For each of the following functions, state whether or not it is a linear function of x.

(a) The total number of publications. [*Ans.* Linear.]
(b) The total number of cards in the catalogue. (Assume that each book has two cards, each other publication has one.)

9. If in (i) and (ii), x is taken as a column vector, then the conditions define a linear function of a column vector. How can we represent such a function? How can we represent a linear transformation of column vectors?

10. Show that the matrix R defined in Example 2 can be thought of as a transformation of both row vectors and column vectors.

*10. PERMUTATION MATRICES

In Chapter III we defined a permutation of n objects to be an arrangement of these objects in a definite order. Thus the set $\{a, b, c\}$ has six permutations: *abc*, *acb*, *bac*, *bca*, *cab*, and *cba*. There is a slightly different way of thinking of a permutation. We may think of our set as given originally in a definite order, say *abc*, and then think of a permutation as a *rearrangement* of the set. Thus one permutation changes *abc* into *bac*; i.e., the first element is put into the second spot, the second into the first spot, and the third element is left unchanged. In order to arrive at the same number, $n!$, of permutations as before, we must consider the "rearrangement" that changes nothing, i.e., the permutation that "changes" *abc* into *abc*. We shall consider our n objects as components of a row vector. A permutation changes the row vector into another having the same components, but possibly in a different order.

A convenient way to describe permutations is by means of certain special matrices. For example, the rearrangement given above can be described by the product

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

In this we do not have to think of the x_i as numbers. They are objects of any sort for which multiplication by 0 and 1 and addition is defined as for numbers. The 3×3 matrix then represents our permutation. It has only 0's and 1's as components, and there is exactly one 1 in each row and in each column.

DEFINITION 1. A *permutation matrix* is a square matrix having exactly one 1 in each row and each column, and having 0's in all other places.

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

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Figure 6

Examples of permutation matrices are shown in Figure 6. Since these matrices are square matrices ($n \times n$), we can speak of the matrix as having *degree n*. Thus Figure 6 shows one matrix of degree 2, two of degree 3, and one of degree 4.

Theorem 1. Every permutation matrix of degree n represents a permutation of n objects, and every such permutation has a unique matrix representation.

Proof. Let us consider n objects $x_1, x_2, \ldots, x_n$ which by a permutation are rearranged to give $y_1, y_2, \ldots, y_n$. Here each of the y's is one of the x's, and every x is some y. If it happens that $y_j = x_i$, then the object in the ith position was changed to the jth position. In this case, define $p_{ij} = 1$ and $p_{ik} = 0$ for $k \neq i$. Doing this for every i, we obtain an $n \times n$ permutation matrix P such that

$$(x_1, x_2, \ldots, x_n)P = (y_1, y_2, \ldots, y_n). \tag{1}$$

The fact that no two elements of a single row or a single column of P are 1 (i.e., that P is a permutation matrix) follows from the fact that in a permutation each element appears once and only once in the rearrangement.

On the other hand, if we are given a permutation matrix P, then we can define a permutation by the product (1). The fact that each column of P has exactly one 1 means that each y_j is some x_i. The fact that P has only one 1 in each row means that every x_i appears as only one y_j. Hence the vector $(y_1, y_2, \ldots, y_n)$ does represent a rearrangement of the vector $(x_1, x_2, \ldots, x_n)$, completing the proof of the theorem.

We shall restrict ourselves to the case of $n = 4$ for illustrating the following discussion, but all the results we are about to establish will hold for every n. In Figure 7 we find four examples of permutation matrices of degree 4.

We want to study the product of two permutation matrices of degree 4. If $x = (x_1, x_2, x_3, x_4)$, then $xJ = (x_4, x_1, x_3, x_2)$ and $xK = (x_2, x_1, x_4, x_3)$. The former puts the first component into second place, the second component into fourth place, and the fourth component into first place; leaving the third component unchanged. The latter interchanges the first two and the last two. What happens if we perform the two permutations, one after the other? Let us first consider x_1. In the first transformation it is changed into the second component,

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$K = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Figure 7

while in the second transformation, the second component is changed into the first. Hence x_1 ends up where it started, in first place. The component x_2 is first sent into the number four slot, and then this is changed to number three by the second transformation. Hence x_2 ends up as the third component. Component x_3 is at first not changed, but later changed into component four. Component x_4 is first made into the first component, and in the second transformation it is changed into the second component. Hence, starting with x, after two transformations we end up with (x_1, x_4, x_2, x_3).

Let us now consider the product

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

The matrix JK is again a permutation matrix, and it is easy to check that it represents precisely the permutation described above.

Theorem 2. The product JK of two permutation matrices of the same degree is again such a permutation matrix. It represents the result of first performing permutation J, then permutation K.

Proof. This theorem is very easy to prove in matrix form. We wish to know what $x(JK)$ is. By the associative law (see Section 4) this is the same as $(xJ)K$. But xJ is the result of the J permutation, and $(xJ)K$ is the result of applying the K permutation to xJ. This proves the theorem.

Example. Referring to Figure 7, let us consider the products IJ and JI. We know, of course, that $IJ = JI = J$. Hence Theorem 2 tells us that performing the I permutation followed by the J permutation (or the reverse) will result simply in the J permutation. If we note that the I permutation leaves everything unchanged, this result is obvious.

Let us now consider the product JL, where again J and L are as in Figure 7. The product is equal to I; hence $L = J^{-1}$. By Theorem 2 we know that the permutation J followed by L will result in the permutation I, i.e., in no change at all. Thus we see that $L = J^{-1}$ is a permutation that undoes all changes made by J. We also note a similarity in the structure of J and L; the latter is formed from the former by turning it over its main diagonal (the diagonal slanting from the upper left-hand corner to the lower right-hand corner). In other words, L has as its i,jth component what J has as its j,ith component.

DEFINITION 2. The *transpose* A^* of a square matrix A is formed by turning it over its main diagonal; that is, the entries of A^* are given by $a_{ij}^* = a_{ji}$.

Theorem 3. If P is a permutation matrix, then P^* is its inverse; that is, P^* represents the permutation which undoes what the permutation P does.

Proof. We must show that P^* undoes what P does; the remainder will follow from the above discussion and Section 6. Let us suppose that $p_{ij}^* = 1$. Then $p_{ji} = 1$; hence the permutation P moves component x_j into position i. But then, because $p_{ij}^* = 1$, the component is moved from position i into position j. Hence x_j ends up in position j, where it started; and this holds for every component. Thus P^* undoes the work of P, which proves the theorem.

DEFINITION 3. A set of objects forms a *group* (with respect to multiplication) if

(i) The product of two elements of the set is always an element of the set.

(ii) There is in the set an element I, called the identity element, such that for every A in the set, $IA = AI = A$.

(iii) For every A in the set there is an element A^{-1} in the set such that $AA^{-1} = A^{-1}A = I$.

(iv) For every A, B, C in the set, $A(BC) = (AB)C$.

DEFINITION 4. A set of objects form a *commutative group* if, in addition to the above four properties, they also satisfy

(v) For every A and B in the set, $AB = BA$.

Theorem 4. The permutation matrices of degree n form a group (with respect to matrix multiplication), but this group is not commutative if $n > 2$.

Proof. Property (i) was shown in Theorem 2. Property (ii) follows from the more general fact that $IM = MI = M$, for every $n \times n$ matrix M. From Theorem 3 we know that A has an inverse, namely $A^{-1} = A^*$. It is easy to show that A^* is again a permutation matrix (see Exercise 1). Hence (iii) follows. And (iv) again follows from the more general theorem that all matrices obey this associative law. (See Section 4.) On the other hand it is easy to show examples, for any $n > 2$, where $AB \neq BA$. (See Exercises 2–3.) This completes the proof.

The group formed by the $n \times n$ permutation matrices is known as the *permutation group of degree n*. Since permutations are used in the study of symmetry, this group is also called the *symmetric group of degree n*.

EXERCISES

1. Prove that the transpose of a permutation matrix is a permutation matrix; i.e., that if A satisfies Definition 1, then so does A^*.

2. Write all permutation matrices of degree 1. Write all permutation matrices of degree 2. Show that these two groups are commutative.

3. For $n > 2$, we can form the matrix A which only interchanges x_1 and x_2, and the matrix B which only interchanges x_1 and x_3. What permutations are performed by AB and by BA? Are these two the same? Use this fact to show that the permutation group of order $n > 2$ is not commutative.

4. Write down the permutation matrices which change (x_1, x_2, x_3, x_4) into

(a) (x_2, x_3, x_4, x_1).
(b) (x_1, x_3, x_2, x_4).
(c) (x_2, x_3, x_1, x_4).
(d) (x_1, x_2, x_3, x_4).

[*Ans.* (a) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.]

5. For the following pairs of matrices, find the permutations they represent. In each case show that AB represents the permutation A followed by the permutation B, and that BA represents the permutation B followed by the permutation A.

(a) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$,

(b) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$,

(c) $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

[*Ans.* (a) xA is (x_3, x_1, x_2); xB is (x_1, x_3, x_2); xAB is (x_3, x_2, x_1); xBA is (x_2, x_1, x_3).]

6. Prove that the set of *all* 3×3 matrices does *not* form a group (with respect to matrix multiplication).

7. Find the inverses of the six matrices in Exercise 5 by using Theorem 3. Check your answers by multiplying the matrices by their inverses.

[*Ans.* (a) $A^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$; $B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.]

8. The process of division is usually introduced by saying that b/a is the solution of the equation $ax = b$ (or of $xa = b$).

(a) Prove that in a group the equation $AX = B$ always has a unique solution.

(b) Prove that in a group the equation $XA = B$ always has a unique solution.

(c) Show by means of an example that the two equations need not have the same solution.

9. For the set of numbers $\{1, 2, 3, 4\}$ we define "multiplication" by means of the following table.

x	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

(In this table we have neglected all multiples of 5; e.g., $2 \times 4 = 8$, but we neglected the 5 and just kept the remainder 3. Again $3 \times 4 = 12$, but we ignored the 10, which is a multiple of 5, and kept the remainder 2.) Prove that this set, with multiplication so defined, forms a commutative group.

10. For the set $\{1, 2, 3, 4, 5, 6\}$ write down a multiplication table, ignoring all multiples of 7. (See Exercise 9.) Prove that the result is a commutative group.

11. For the set $\{1, 2, 3, 4, 5\}$ write down a multiplication table, ignoring multiples of 6. (See Exercises 9 and 10.) Prove that the result is *not* a group. Why do 5 and 7 give us groups, but not 6?

12. Write down all permutation matrices of degree 3, and assign letter-names to them. Write a multiplication table for this group. How, from this table alone, can we see that properties (i), (ii), and (iii) hold? How do we see that (v) does not hold?

13. Consider a group with four elements, $G = \{a, b, c, d\}$. For each x in G let x' be the vector $x' = (xa, xb, xc, xd)$; e.g., $a' = (a^2, ab, ac, ad)$. Show that x' is a permutation of (a, b, c, d). Show that the four permutations, a', b', c', d' form a permutation group having the same multiplication table as G; i.e., show $x'y' = z'$ if and only if $xy = z$.

14. Find the permutation group associated with the group in Exercise 9 by the method of Exercise 13.

*11. SUBGROUPS OF PERMUTATION GROUPS

Within a group we sometimes can find smaller groups. Here we shall study some of the subgroups of permutation groups. It will be understood that whenever we speak of a group we have a set with a finite number of elements in mind. In particular, this will be assumed for the theorems given below, since some of the theorems are not valid for groups with an infinite number of elements. The concept of a group has important applications for infinite sets, but these do not belong in this book.

Definition 1. If a given set G forms a group, and some subset H of it also forms a group, we call the subset H a *subgroup* of G. If the subset H is a proper subset of G, we speak of a *proper subgroup.*

Theorem 1. If we select any element of a group, the powers of the element form a subgroup which is commutative.

Proof. Select any element A of the given group; we must show that the powers A^n have the properties (i)–(v) given in the last section. The product of two powers is again a power, $A^jA^k = A^{j+k}$; hence (i) holds. Next we observe that the powers cannot be all different, since this would give us infinitely many elements in our group. Hence we must have an equation $A^j = A^k$, with, say, $j > k$. However, this implies that $A^{j-k} = I$. Hence I occurs among the powers of A, say $I = A^m$. Therefore (ii) holds. If $m = 1$ or 2, then A is its own inverse (see Exercise 9). On the other hand, if $m > 2$, then among the powers we find A^{m-1}, and $AA^{m-1} = A^m = I$, so that A^{m-1} is the inverse of A. This shows that property (iii) holds. The associative law (iv) follows from the fact that all matrices obey this law. Finally, we get commutativity (v) from the fact that $A^jA^k = A^{j+k} = A^{k+j} = A^kA^j$, completing the proof.

Definition 2. A group which consists of the powers of one element A is known as the *cyclic* group *generated* by A.

Thus we know that we can form a cyclic subgroup of a given group by picking any one element A and taking all its powers. The number of elements in this subgroup is called the *order* of A. In the proof above, the order of A is the smallest possible m such that $A^m = I$.

Example 1. The permutation group of degree 4 has $4! = 24$ elements. Let us consider the cyclic subgroup generated by J (see Figure 7). We find that $J^2 = L = J^{-1}$, so that $J^3 = JJ^2 = I$. Thus our cyclic subgroup consists of J, $J^2 = L$, and $J^3 = I$. If we continue to take higher powers, we get $J^4 = J$, $J^5 = L$, $J^6 = I$, etc. The elements are repeated in this fixed cycle. This is the source of the name "cyclic."

Example 2. We can get a larger cyclic subgroup by choosing the matrix M and its powers (see Figure 8). M has order 4; hence $M^{-1} = M^* = M^3$, and $M^4 = I$.

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad M^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$M^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad M^4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Figure 8

Theorem 2. If in a group we select any subset having property (i), then this subset is a subgroup.

Proof. We must show that the subset also has properties (ii)–(iv). Let A be any element of the subset. By (i), $AA = A^2$ is also in the subset, and then $AA^2 = A^3$ is in the subset, etc. Hence all powers of A are in the subset. One of these powers is I and one is A^{-1}. Hence we have properties (ii) and (iii). Property (iv) again follows from the fact that all matrices have this property, completing the proof of the theorem.

We now have a practical way of finding subgroups. We select one or more elements of the group, and form all possible products of these, using each one as many times as necessary. If we form all possible products, then the product of any two products will also be on our list, and hence property (i) holds. Then, by Theorem 2, we have a subgroup, which is called the subgroup *generated* by the elements. If we start with a single element, we obtain a cyclic subgroup. Some very interesting subgroups can be generated by two elements.

Example 3. Let us start with J (see Figure 7), and D (see Figure 6), and form the subgroup they generate. First of all we get the powers of J, namely, J and $J^2 = L$ and $J^3 = I$, as was shown in Example 1. Then we have D, and D^2, which is again I. In products formed using both J and D we need consider only J and J^2 and D, since the next higher power is I, and then the powers are repeated. Theoretically, we should consider products like $DJDJ^2$ and $JDJDJDJ$, but we can show, as follows, that such long products give nothing new. First we observe that $DJ = J^2D$, so that in a long product we may always replace DJ by J^2D, and thus put all the J's in front and all the D's at the end. (See Exercise 14.) Therefore the only new products that we need consider are of the form J^aD^b; and since J can occur only to the first or second power and D only to the first power, we arrive at JD and J^2D as the only additional products. Hence our subgroup has six elements: J, J^2, D, I, JD, and J^2D. Since $JD \neq DJ$, the subgroup is *not* commutative.

So far we have found subgroups of 3, 4, and 6 elements. Each of these numbers is a divisor of 24, the total number of elements in the group. It can be shown that the number of elements in a subgroup is always a divisor of the number of elements in the group, but we will not prove that fact here.

Example 4. Let us now form the subgroup generated by D and K. Since $D^2 = I = K^2$, both D and K will occur only to the first power in a product. Furthermore $DK = KD$; hence the subgroup will have only four elements: I, D, K, DK. This subgroup is commutative. The fact that the subgroup happens to be commutative is a consequence of the following theorem.

Theorem 3. If A and B commute (i.e., $AB = BA$), then any two products formed from A and B also commute. Hence the subgroup generated by A and B is a commutative subgroup.

Proof. Given any product formed from A and B, say $AABBBABAB$, we can make use of the fact that $AB = BA$ to move all the A's up front and all the B's to the end. Hence the product can be written A^iB^j. A second such product can be written A^kB^m. The product of these, $A^iB^jA^kB^m$, can again be rearranged so that all the A's come at the beginning. Hence $(A^iB^j)(A^kB^m) = A^{i+k}B^{j+m} = A^{k+i}B^{m+j} = (A^kB^m)(A^iB^j)$, completing the proof.

We have now found two types of commutative subgroups: (1) cyclic subgroups and (2) subgroups generated by two elements that commute. For the latter it is convenient to have a technique for finding two commuting elements. We will develop one method for finding such pairs.

DEFINITION 3. The *effective set* of a permutation matrix is the set of all those components of the row vector which are changed by the matrix.

For example, D has $\{x_1, x_2\}$ as its effective set, J has $\{x_1, x_2, x_4\}$, K has the set of all four components, and I has the empty set as its effective set. K suggests the definition:

DEFINITION 4. A permutation matrix having all the components in its effective set is called a *complete* permutation matrix.

Theorem 4. Two permutation matrices, whose effective sets are disjoint, commute.

Proof. Let A_1 have X_1 as its effective set, and A_2 have X_2, so that $X_1 \cap X_2 = \mathcal{E}$. Then A_1A_2 will make some changes on X_1 and then on X_2. The latter are not affected by the former, since X_1 and X_2 have nothing in common. Thus we get the same result if we perform A_2 followed by A_1.

We now have a simple way of getting a commutative subgroup, other than a cyclic one. Just select any two matrices (other than I) with disjoint effective sets, and form the subgroup that they generate.

EXERCISES

1. Write down the six permutation matrices of degree 3.

2. Form the cyclic subgroup for each of the six matrices in Exercise 1. Are these subgroups all different? What is the order of each matrix?

[*Ans.* Five distinct groups; one of order 1, three of order 2, two of order 3.]

3. Prove that there are no proper subgroups of the permutation group of degree 3, other than those found in Exercise 2.

4. Write the 24 permutation matrices of degree 4.

5. Form the cyclic subgroup for each of the matrices in Exercise 4. How many different ones do you get? What is the order of each matrix?

[*Ans.* 17 distinct groups; one of order 1, nine of order 2, eight of order 3, six of order 4.]

6. Show by an example that the subgroups found in Exercise 5 are not the only proper subgroups of the permutation group of degree 4.

7. Prove the following facts about orders of permutations.

(a) I has order 1.

(b) A permutation which does nothing but interchange one or more pairs of elements has order 2.

(c) Every other permutation has an order greater than 2.

8. Prove that the subgroup generated by A and B is cyclic if and only if one generator is a power of the other.

9. Prove that if a matrix has order 1 or 2, then it is its own inverse.

10. A matrix M is said to be *symmetric* if $m_{ij} = m_{ji}$ for all i and j. Prove that a permutation matrix is symmetric if and only if it has order 1 or 2.

11. Form the subgroup generated by J and K.

[*Ans.* There are 12 elements.]

12. Prove the following facts about effective sets.

(a) I has an effective set of zero elements.

(b) A matrix which simply interchanges two elements has as its effective set a set of two elements.

(c) All other matrices have an effective set of at least three elements.

(d) A matrix is complete if and only if the number of elements in its effective set equals its degree.

13. We wish to form a commutative subgroup of the permutation group of degree 4, by means of the method described above. We want to choose two matrices (other than I) with disjoint effective sets, and form the subgroup they generate.

(a) Using the results of Exercise 12, what must the number of elements be in the two effective sets? [*Ans.* 2, 2.]

(b) Choose such a pair of matrices.

(c) Form the subgroup.

14. Prove the following facts about Example 3 above.

(a) $DJ = J^2D$.

(b) From this it follows that $DJ^2 = JD$.

(c) In any product of D's and J's we can put all the J's up front.

15. If A has order m, and m is an even number, then $A^{m/2}$ is its own inverse. Prove this fact. What does this say about an element of order 2?

16. Prove that the cyclic group generated by A^2 is a subgroup of that generated by A. When will this be a proper subgroup?

SUGGESTED READING

Birkhoff, G., and S. MacLane, *A Survey of Modern Algebra*, Macmillan, third ed., 1965.

Weiss, M. J., *Higher Algebra for the Undergraduate*, Wiley, New York, 1949, Chaps. 3, 6, 7.

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

Beaumont, R. A., and R. W. Ball, *Introduction to Modern Algebra and Matrix Theory*, Holt, Rinehart and Winston, New York, 1954, Chaps. I–IV.

Thrall, R. M., and L. Tornheim, *Vector Spaces and Matrices*, Wiley, New York, 1957, Chaps. 1–3.

Hohn, F. E., *Elementary Matrix Algebra*, Macmillan, New York, 1958, Chaps. 1, 3, 5.

Hadley, G., *Linear Algebra*, Addison-Wesley, Reading, Mass., 1961, Chaps. 1–5.