CHAPTER

FOUR

LINEAR MODELS AND MATRIX ALGEBRA

For the one-commodity model (3.1), the solutions $\bar{P}$ and $\bar{Q}$ as expressed in (3.4) and (3.5) are relatively simple, even though a number of parameters are involved. As more and more commodities are incorporated into the model, such solution formulas quickly become cumbersome and unwieldy. That was why we had to resort to a little shorthand, even for the two-commodity case—in order that the solutions (3.14) and (3.15) can still be written in a relatively concise fashion. We did not attempt to tackle any three- or four-commodity models, even in the linear version, primarily because we did not yet have at our disposal a method suitable for handling a large system of simultaneous equations. Such a method is found in *matrix algebra*, the subject of this chapter and the next.

Matrix algebra can enable us to do many things. In the first place, it provides a compact way of writing an equation system, even an extremely large one. Second, it leads to a way of testing the existence of a solution by evaluation of a *determinant*—a concept closely related to that of a matrix. Third, it gives a method of finding that solution (if it exists). Since equation systems are encountered not only in static analysis but also in comparative-static and dynamic analyses and in optimization problems, you will find ample application of matrix algebra in almost every chapter that is to follow.

However, one slight "catch" should be mentioned at the outset. Matrix algebra is applicable only to *linear*-equation systems. How realistically linear equations can describe actual economic relationships depends, of course, on the nature of the relationships in question. In many cases, even if some sacrifice of realism is entailed by the assumption of linearity, an assumed linear relationship can produce a sufficiently close approximation to an actual nonlinear relationship to warrant its use. In other cases, the closeness of approximation may also be

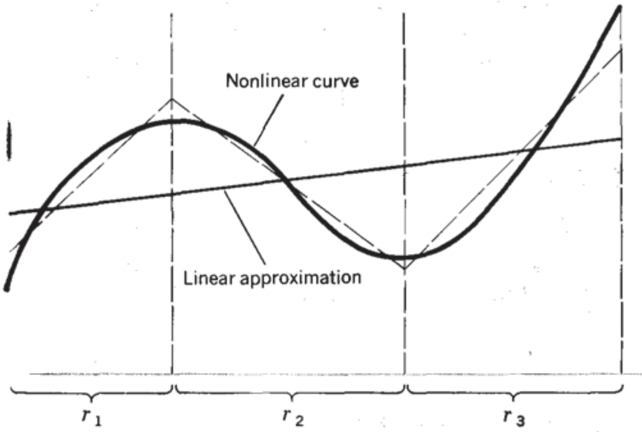

Figure 4.1

improved by having a separate linear approximation for each segment of a nonlinear relationship, as is illustrated in Fig. 4.1. If the solid curve is taken as the actual nonlinear relationship, a single linear approximation might take the form of the solid straight line, which shows substantial deviation from the curve at certain points. But if the domain is divided into three regions r_1, r_2, and r_3, we can have a much closer linear approximation (broken straight line) in each region.

In yet other cases, while preserving the nonlinearity in the model, we can effect a transformation of variables so as to obtain a linear relation to work with. For example, the nonlinear function

$$y = ax^b$$

can be readily transformed, by taking the logarithm on both sides, into the function

$$\log y = \log a + b \log x$$

which is linear in the two variables ($\log y$) and ($\log x$). (Logarithms will be discussed in detail in Chap. 10.)

In short, the linearity assumption frequently adopted in economics may in certain cases be quite reasonable and justified. On this note, then, let us proceed to the study of matrix algebra.

4.1 MATRICES AND VECTORS

The two-commodity market model (3.12) can be written—after eliminating the quantity variables—as a system of two linear equations, as in (3.13′),

$$c_1P_1 + c_2P_2 = -c_0$$
$$\gamma_1P_1 + \gamma_2P_2 = -\gamma_0$$

where the parameters c_0 and γ_0 appear to the right of the equals sign. In general, a system of m linear equations in n variables $(x_1, x_2, \ldots, x_n)$ can also be arranged into such a format:

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = d_1 \\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = d_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = d_m
\end{aligned}
\tag{4.1}
$$

In (4.1), the variable x_1 appears only within the leftmost column, and in general the variable x_j appears only in the jth column on the left side of the equals sign. The double-subscripted parameter symbol a_{ij} represents the coefficient appearing in the ith equation and attached to the jth variable. For example, a_{21} is the coefficient in the second equation, attached to the variable x_1. The parameter d_i which is unattached to any variable, on the other hand, represents the constant term in the ith equation. For instance, d_1 is the constant term in the first equation. All subscripts are therefore keyed to the specific locations of the variables and parameters in (4.1).

Matrices as Arrays

There are essentially three types of ingredients in the equation system (4.1). The first is the set of coefficients a_{ij}; the second is the set of variables $x_1, \ldots, x_n$; and the last is the set of constant terms $d_1, \ldots, d_m$. If we arrange the three sets as three rectangular arrays and label them, respectively, as A, x, and d (without subscripts), then we have

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}
\qquad
x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
\qquad
d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}
\tag{4.2}
$$

As a simple example, given the linear-equation system

$$
\begin{aligned}
6x_1 + 3x_2 + x_3 &= 22 \\
x_1 + 4x_2 - 2x_3 &= 12 \\
4x_1 - x_2 + 5x_3 &= 10
\end{aligned}
\tag{4.3}
$$

we can write

$$
A = \begin{bmatrix} 6 & 3 & 1 \\ 1 & 4 & -2 \\ 4 & -1 & 5 \end{bmatrix}
\qquad
x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}
\qquad
d = \begin{bmatrix} 22 \\ 12 \\ 10 \end{bmatrix}
\tag{4.4}
$$

Each of the three arrays in (4.2) or (4.4) constitutes a *matrix*.

A matrix is defined as a rectangular array of numbers, parameters, or variables. The members of the array, referred to as the *elements* of the matrix, are usually enclosed in brackets, as in (4.2), or sometimes in parentheses or with double vertical lines: $\| \ \|$. Note that in matrix A (the *coefficient matrix* of the equation system), the elements are separated not by commas but by blank spaces only. As a shorthand device, the array in matrix A can be written more simply as

$$A = [a_{ij}] \qquad \begin{pmatrix} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{pmatrix}$$

Inasmuch as the location of each element in a matrix is unequivocally fixed by the subscript, every matrix is an ordered set.

Vectors as Special Matrices

The number of rows and the number of columns in a matrix together define the *dimension* of the matrix. Since matrix A in (4.2) contains m rows and n columns, it is said to be of dimension $m \times n$ (read: "m by n"). It is important to remember that the row number always precedes the column number; this is in line with the way the two subscripts in a_{ij} are ordered. In the special case where $m = n$, the matrix is called a *square matrix*; thus the matrix A in (4.4) is a 3×3 square matrix.

Some matrices may contain only one column, such as x and d in (4.2) or (4.4). Such matrices are given the special name *column vectors*. In (4.2), the dimension of x is $n \times 1$, and that of d is $m \times 1$; in (4.4) both x and d are 3×1. If we arranged the variables x_j in a horizontal array, though, there would result a $1 \times n$ matrix, which is called a *row vector*. For notation purposes, a row vector is often distinguished from a column vector by the use of a primed symbol:

$$x' = [x_1 \quad x_2 \quad \cdots \quad x_n]$$

You may observe that a vector (whether row or column) is merely an ordered n-tuple, and as such it may be interpreted as a point in an n-dimensional space. In turn, the $m \times n$ matrix A can be interpreted as an ordered set of m row vectors or as an ordered set of n column vectors. These ideas will be followed up later.

An issue of more immediate interest is how the matrix notation can enable us, as promised, to express an equation system in a compact way. With the matrices defined in (4.4), we can express the equation system (4.3) simply as

$$Ax = d$$

In fact, if A, x, and d are given the meanings in (4.2), then even the general-equation system in (4.1) can be written as $Ax = d$. The compactness of this notation is thus unmistakable.

However, the equation $Ax = d$ prompts at least two questions. How do we multiply two matrices A and x? What is meant by the equality of Ax and d? Since matrices involve whole blocks of numbers, the familiar algebraic operations defined for single numbers are not directly applicable, and there is need for a new set of operational rules.

EXERCISE 4.1

1 Rewrite the equation system (3.1) in the format of (4.1), and show that, if the three variables are arranged in the order Q_d, Q_s, and P, the coefficient matrix will be

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & b \\ 0 & 1 & -d \end{bmatrix}$$

How would you write the vector of constants?

2 Rewrite the equation system (3.12) in the format of (4.1) with the variables arranged in the following order: $Q_{d1}, Q_{s1}, Q_{d2}, Q_{s2}, P_1, P_2$. Write out the coefficient matrix, the variable vector, and the constant vector.

4.2 MATRIX OPERATIONS

As a preliminary, let us first define the word *equality*. Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are said to be *equal* if and only if they have the same dimension and have identical elements in the corresponding locations in the array. In other words, $A = B$ if and only if $a_{ij} = b_{ij}$ for all values of i and j. Thus, for example, we find

$$\begin{bmatrix} 4 & 3 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 2 & 0 \end{bmatrix} \neq \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$$

As another example, if $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$, this will mean that $x = 7$ and $y = 4$.

Addition and Subtraction of Matrices

Two matrices can be added if and only if they have the same dimension. When this dimensional requirement is met, the matrices are said to be *conformable for addition.* In that case, the addition of $A = [a_{ij}]$ and $B = [b_{ij}]$ is defined as the addition of each pair of corresponding elements.

Example 1

$$\begin{bmatrix} 4 & 9 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} 4+2 & 9+0 \\ 2+0 & 1+7 \end{bmatrix} = \begin{bmatrix} 6 & 9 \\ 2 & 8 \end{bmatrix}$$

Example 2

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \\ a_{31}+b_{31} & a_{32}+b_{32} \end{bmatrix}$$

In general, we may state the rule thus:

$$[a_{ij}] + [b_{ij}] = [c_{ij}] \qquad \text{where } c_{ij} = a_{ij} + b_{ij}$$

Note that the sum matrix $[c_{ij}]$ must have the same dimension as the component matrices $[a_{ij}]$ and $[b_{ij}]$.

The subtraction operation $A - B$ can be similarly defined if and only if A and B have the same dimension. The operation entails the result

$$[a_{ij}] - [b_{ij}] = [d_{ij}] \qquad \text{where } d_{ij} = a_{ij} - b_{ij}$$

Example 3

$$\begin{bmatrix} 19 & 3 \\ 2 & 0 \end{bmatrix} - \begin{bmatrix} 6 & 8 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 19-6 & 3-8 \\ 2-1 & 0-3 \end{bmatrix} = \begin{bmatrix} 13 & -5 \\ 1 & -3 \end{bmatrix}$$

The subtraction operation $A - B$ may be considered alternatively as an addition operation involving a matrix A and another matrix $(-1)B$. This, however, raises the question of what is meant by the multiplication of a matrix by a single number (here, -1).

Scalar Multiplication

To multiply a matrix by a number—or in matrix-algebra terminology, by a *scalar* —is to multiply *every* element of that matrix by the given scalar.

Example 4

$$7\begin{bmatrix} 3 & -1 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 21 & -7 \\ 0 & 35 \end{bmatrix}$$

Example 5

$$\frac{1}{2}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}a_{11} & \frac{1}{2}a_{12} \\ \frac{1}{2}a_{21} & \frac{1}{2}a_{22} \end{bmatrix}$$

From these examples, the rationale of the name scalar should become clear, for it "scales up (or down)" the matrix by a certain multiple. The scalar can, of course, be a negative number as well.

Example 6

$$-1\begin{bmatrix} a_{11} & a_{12} & d_1 \\ a_{21} & a_{22} & d_2 \end{bmatrix} = \begin{bmatrix} -a_{11} & -a_{12} & -d_1 \\ -a_{21} & -a_{22} & -d_2 \end{bmatrix}$$

Note that if the matrix on the left represents the coefficients *and* the constant terms in the simultaneous equations

$$a_{11}x_1 + a_{12}x_2 = d_1$$
$$a_{21}x_1 + a_{22}x_2 = d_2$$

then multiplication by the scalar -1 will amount to multiplying both sides of both equations by -1, thereby changing the sign of every term in the system.

Multiplication of Matrices

Whereas a scalar can be used to multiply a matrix of any dimension, the multiplication of two matrices is contingent upon the satisfaction of a different dimensional requirement.

Suppose that, given two matrices A and B, we want to find the product AB. The conformability condition for multiplication is that the *column* dimension of A (the "lead" matrix in the expression AB) must be equal to the row dimension of B (the "lag" matrix). For instance, if

$$\text{(4.5)} \qquad \underset{(1\times 2)}{A} = \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \qquad \underset{(2\times 3)}{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$$

the product AB then *is* defined, since A has *two columns* and B has *two rows*—precisely the same number.* This can be checked at a glance by comparing the *second* number in the dimension indicator for A, which is (1×2), with the *first* number in the dimension indicator for B, (2×3). On the other hand, the reverse product BA is *not* defined in this case, because B (now the lead matrix) has *three* columns while A (the lag matrix) has only *one* row; hence the conformability condition is violated.

In general, if A is of dimension $m \times n$ and B is of dimension $p \times q$, the matrix product AB will be defined if and only if $n = p$. If defined, moreover, the product matrix AB will have the dimension $m \times q$—the same number of *rows* as the lead matrix A and the same number of *columns* as the lag matrix B. For the matrices given in (4.5), AB will be 1×3.

It remains to define the exact procedure of multiplication. For this purpose, let us take the matrices A and B in (4.5) for illustration. Since the product AB is defined and is expected to be of dimension 1×3, we may write in general (using the symbol C rather than c' for the row vector) that

$$AB = C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \end{bmatrix}$$

Each element in the product matrix C, denoted by c_{ij}, is defined as a sum of products, to be computed from the elements in the ith *row* of the lead matrix A, and those in the jth *column* of the lag matrix B. To find c_{11}, for instance, we should take the *first row* in A (since $i = 1$) and the *first column* in B (since $j = 1$)

* The matrix A, being a row vector, would normally be denoted by a'. We use the symbol A here to stress the fact that the multiplication rule being explained applies to matrices in general, not only to the product of one vector and one matrix.

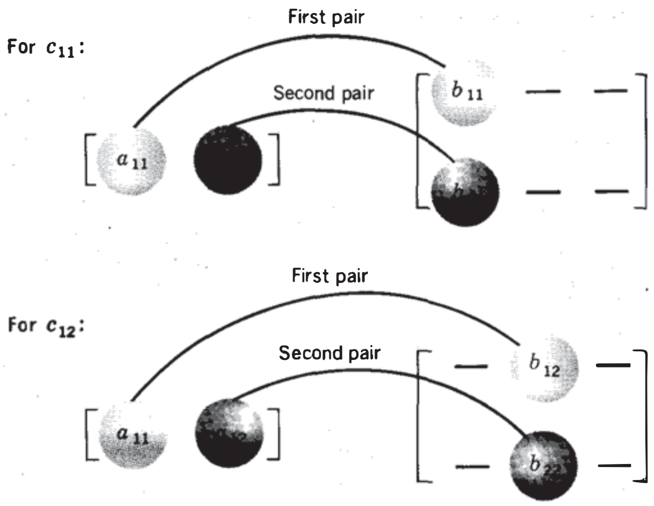

Figure 4.2

—as shown in the top panel of Fig. 4.2—and then pair the elements together sequentially, multiply out each pair, and take the sum of the resulting products, to get

(4.6) $$c_{11} = a_{11}b_{11} + a_{12}b_{21}$$

Similarly, for c_{12}, we take the *first row* in A (since $i = 1$) and the *second column* in B (since $j = 2$), and calculate the indicated sum of products—in accordance with the lower panel of Fig. 4.2—as follows:

(4.6′) $$c_{12} = a_{11}b_{12} + a_{12}b_{22}$$

By the same token, we should also have

(4.6″) $$c_{13} = a_{11}b_{13} + a_{12}b_{23}$$

It is the particular pairing requirement in this process which necessitates the matching of the column dimension of the lead matrix and the row dimension of the lag matrix before multiplication can be performed.

The multiplication procedure illustrated in Fig. 4.2 can also be described by using the concept of the *inner product* of two vectors. Given two vectors u and v with n elements each, say, $(u_1, u_2, \ldots, u_n)$ and $(v_1, v_2, \ldots, v_n)$, arranged *either* as two rows *or* as two columns *or* as one row and one column, their inner product, written as $u \cdot v$, is defined as

$$u \cdot v = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

This is a sum of products of corresponding elements, and hence the inner product of two vectors is a scalar. If, for instance, we prepare after a shopping trip a vector of quantities purchased of n goods and a vector of their prices (listed in the corresponding order), then their inner product will give the total purchase cost.

Note that the inner-product concept is exempted from the conformability condition, since the arrangement of the two vectors in rows or columns is immaterial.

Using this concept, we can describe the element c_{ij} in the product matrix $C = AB$ simply as the inner product of the ith row of the lead matrix A and the jth column of the lag matrix B. By examining Fig. 4.2, we can easily verify the validity of this description.

The rule of multiplication outlined above applies with equal validity when the dimensions of A and B are other than those illustrated above; the only prerequisite is that the conformability condition be met.

Example 7 Given

$$\underset{(2\times 2)}{A} = \begin{bmatrix} 3 & 5 \\ 4 & 6 \end{bmatrix} \quad \text{and} \quad \underset{(2\times 2)}{B} = \begin{bmatrix} -1 & 0 \\ 4 & 7 \end{bmatrix}$$

find AB. The product AB is obviously defined, and will be 2×2:

$$AB = \begin{bmatrix} 3(-1) + 5(4) & 3(0) + 5(7) \\ 4(-1) + 6(4) & 4(0) + 6(7) \end{bmatrix} = \begin{bmatrix} 17 & 35 \\ 20 & 42 \end{bmatrix}$$

Example 8 Given

$$\underset{(3\times 2)}{A} = \begin{bmatrix} 1 & 3 \\ 2 & 8 \\ 4 & 0 \end{bmatrix} \quad \text{and} \quad \underset{(2\times 1)}{b} = \begin{bmatrix} 5 \\ 9 \end{bmatrix}$$

find Ab. This time the product matrix should be 3×1, that is, a column vector:

$$Ab = \begin{bmatrix} 1(5) + 3(9) \\ 2(5) + 8(9) \\ 4(5) + 0(9) \end{bmatrix} = \begin{bmatrix} 32 \\ 82 \\ 20 \end{bmatrix}$$

Example 9 Given

$$\underset{(3\times 3)}{A} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 0 & 3 \\ 4 & 0 & 2 \end{bmatrix} \quad \text{and} \quad \underset{(3\times 3)}{B} = \begin{bmatrix} 0 & -\frac{1}{5} & \frac{3}{10} \\ -1 & \frac{1}{5} & \frac{7}{10} \\ 0 & \frac{2}{5} & -\frac{1}{10} \end{bmatrix}$$

find AB. The same rule of multiplication now yields a very special product matrix:

$$AB = \begin{bmatrix} 0+1+0 & -\frac{3}{5} - \frac{1}{5} + \frac{4}{5} & \frac{9}{10} - \frac{7}{10} - \frac{2}{10} \\ 0+0+0 & -\frac{1}{5} + 0 + \frac{6}{5} & \frac{3}{10} + 0 - \frac{3}{10} \\ 0+0+0 & -\frac{4}{5} + 0 + \frac{4}{5} & \frac{12}{10} + 0 - \frac{2}{10} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This last matrix—a square matrix with 1s in its *principal diagonal* (the diagonal running from northwest to southeast) and 0s everywhere else—exemplifies the important type of matrix known as *identity matrix*. This will be further discussed below.

Example 10 Let us now take the matrix A and the vector x as defined in (4.4) and find Ax. The product matrix is a 3×1 column vector:

$$\underset{(3\times 3)}{Ax = \begin{bmatrix} 6 & 3 & 1 \\ 1 & 4 & -2 \\ 4 & -1 & 5 \end{bmatrix}} \underset{(3\times 1)}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}} = \underset{(3\times 1)}{\begin{bmatrix} 6x_1 + 3x_2 + x_3 \\ x_1 + 4x_2 - 2x_3 \\ 4x_1 - x_2 + 5x_3 \end{bmatrix}}$$

Repeat: the product on the right is a *column* vector, its corpulent appearance notwithstanding! When we write $Ax = d$, therefore, we have

$$\begin{bmatrix} 6x_1 + 3x_2 + x_3 \\ x_1 + 4x_2 - 2x_3 \\ 4x_1 - x_2 + 5x_3 \end{bmatrix} = \begin{bmatrix} 22 \\ 12 \\ 10 \end{bmatrix}$$

which, according to the definition of matrix equality, is equivalent to the statement of the entire equation system in (4.3).

Note that, to use the matrix notation $Ax = d$, it is necessary, because of the conformability condition, to arrange the variables x_j into a *column* vector, even though these variables are listed in a horizontal order in the original equation system.

Example 11 The simple national-income model in two endogenous variables Y and C,

$$Y = C + I_0 + G_0$$
$$C = a + bY$$

can be rearranged into the standard format of (4.1) as follows:

$$Y - C = I_0 + G_0$$
$$-bY + C = a$$

Hence the coefficient matrix A, the vector of variables x, and the vector of constants d are:

$$\underset{(2\times 2)}{A} = \begin{bmatrix} 1 & -1 \\ -b & 1 \end{bmatrix} \qquad \underset{(2\times 1)}{x} = \begin{bmatrix} Y \\ C \end{bmatrix} \qquad \underset{(2\times 1)}{d} = \begin{bmatrix} I_0 + G_0 \\ a \end{bmatrix}$$

Let us verify that this given system can be expressed by the equation $Ax = d$. By the rule of matrix multiplication, we have

$$Ax = \begin{bmatrix} 1 & -1 \\ -b & 1 \end{bmatrix} \begin{bmatrix} Y \\ C \end{bmatrix} = \begin{bmatrix} 1(Y) + (-1)(C) \\ -b(Y) + 1(C) \end{bmatrix} = \begin{bmatrix} Y - C \\ -bY + C \end{bmatrix}$$

Thus the matrix equation $Ax = d$ would give us

$$\begin{bmatrix} Y - C \\ -bY + C \end{bmatrix} = \begin{bmatrix} I_0 + G_0 \\ a \end{bmatrix}$$

Since matrix equality means the equality between corresponding elements, it is clear that the equation $Ax = d$ does precisely represent the original equation system, as expressed in the (4.1) format above.

The Question of Division

While matrices, like numbers, can undergo the operations of addition, subtraction, and multiplication—subject to the conformability conditions—it is not possible to divide one matrix by another. That is, we cannot write A/B.

For two numbers a and b, the quotient a/b (with $b \neq 0$) can be written alternatively as ab^{-1} or $b^{-1}a$, where b^{-1} represents the *inverse* or *reciprocal* of b. Since $ab^{-1} = b^{-1}a$, the quotient expression a/b can be used to represent both ab^{-1} and $b^{-1}a$. The case of matrices is different. Applying the concept of inverses to matrices, we may in certain cases (discussed below) define a matrix B^{-1} that is the inverse of matrix B. But from the discussion of conformability condition it follows that, if AB^{-1} is defined, there can be no assurance that $B^{-1}A$ is also defined. Even if AB^{-1} and $B^{-1}A$ are indeed both defined, they still may not represent the same product. Hence the expression A/B cannot be used without ambiguity, and it must be avoided. Instead, you must specify whether you are referring to AB^{-1} or $B^{-1}A$—provided that the inverse B^{-1} does exist and that the matrix product in question is defined. Inverse matrices will be further discussed below.

Digression on Σ Notation

The use of subscripted symbols not only helps in designating the locations of parameters and variables but also lends itself to a flexible shorthand for denoting sums of terms, such as those which arose during the process of matrix multiplication.

The summation shorthand makes use of the Greek letter Σ (sigma, for "sum"). To express the sum of x_1, x_2, and x_3, for instance, we may write

$$x_1 + x_2 + x_3 = \sum_{j=1}^{3} x_j$$

which is read: "the sum of x_j as j ranges from 1 to 3." The symbol j, called the *summation index*, takes only integer values. The expression x_j represents the *summand* (that which is to be summed), and it is in effect a function of j. Aside from the letter j, summation indices are also commonly denoted by i or k, such as

$$\sum_{i=3}^{7} x_i = x_3 + x_4 + x_5 + x_6 + x_7$$

$$\sum_{k=0}^{n} x_k = x_0 + x_1 + \cdots + x_n$$

The application of Σ notation can be readily extended to cases in which the x term is prefixed with a coefficient or in which each term in the sum is raised to some integer power. For instance, we may write:

$$\sum_{j=1}^{3} ax_j = ax_1 + ax_2 + ax_3 = a(x_1 + x_2 + x_3) = a\sum_{j=1}^{3} x_j$$

$$\sum_{j=1}^{3} a_j x_j = a_1x_1 + a_2x_2 + a_3x_3$$

$$\sum_{i=0}^{n} a_i x^i = a_0x^0 + a_1x^1 + a_2x^2 + \cdots + a_nx^n$$

$$= a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

The last example, in particular, shows that the expression $\sum_{i=0}^{n} a_i x^i$ can in fact be used as a shorthand form of the general polynomial function of (2.4).

It may be mentioned in passing that, whenever the context of the discussion leaves no ambiguity as to the range of summation, the symbol Σ can be used alone, without an index attached (such as Σx_i), or with only the index letter underneath (such as $\sum_i x_i$).

Let us apply the Σ shorthand to matrix multiplication. In (4.6), (4.6′), and (4.6″), each element of the product matrix $C = AB$ is defined as a sum of terms, which may now be rewritten as follows:

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} = \sum_{k=1}^{2} a_{1k}b_{k1}$$

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} = \sum_{k=1}^{2} a_{1k}b_{k2}$$

$$c_{13} = a_{11}b_{13} + a_{12}b_{23} = \sum_{k=1}^{2} a_{1k}b_{k3}$$

In each case, the first subscript of c_{1j} is reflected in the first subscript of a_{1k}, and the second subscript of c_{1j} is reflected in the second subscript of b_{kj} in the Σ expression. The index k, on the other hand, is a "dummy" subscript; it serves to indicate which particular pair of elements is being multiplied, but it does not show up in the symbol c_{1j}.

Extending this to the multiplication of an $m \times n$ matrix $A = [a_{ik}]$ and an $n \times p$ matrix $B = [b_{kj}]$, we may now write the elements of the $m \times p$ product matrix $AB = C = [c_{ij}]$ as

$$c_{11} = \sum_{k=1}^{n} a_{1k}b_{k1} \qquad c_{12} = \sum_{k=1}^{n} a_{1k}b_{k2} \qquad \cdots$$

or more generally,

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad \begin{pmatrix} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, p \end{pmatrix}$$

This last equation represents yet another way of stating the rule of multiplication for the matrices defined above.

EXERCISE 4.2

1 Given $A = \begin{bmatrix} 4 & -1 \\ 6 & 9 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 3 \\ 3 & -2 \end{bmatrix}$, and $C = \begin{bmatrix} 8 & 3 \\ 6 & 1 \end{bmatrix}$, find:

(a) $A + B$ (b) $C - A$ (c) $3A$ (d) $4B + 2C$

2 Given $A = \begin{bmatrix} 2 & 8 \\ 3 & 0 \\ 5 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 3 & 8 \end{bmatrix}$, and $C = \begin{bmatrix} 7 & 2 \\ 6 & 3 \end{bmatrix}$:

(a) Is AB defined? Calculate AB. Can you calculate BA? Why?

(b) Is BC defined? Calculate BC. Is CB defined? If so, calculate CB. Is it true that $BC = CB$?

3 On the basis of the matrices given in Example 9, is the product BA defined? If so, calculate the product. In this case do we have $AB = BA$?

4 Find the product matrices in the following (in each case, append beneath every matrix a dimension indicator):

(a) $\begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 4 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 1 \\ 3 & 5 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 2 & 0 \\ 4 & 2 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

(b) $\begin{bmatrix} 6 & 5 & 1 \\ 3 & 0 & 4 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 5 & 2 \\ 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & 2 \\ 1 & 4 \end{bmatrix}$

5 Expand the following summation expressions:

(a) $\sum_{i=2}^{5} x_i$ (c) $\sum_{i=1}^{4} bx_i$ (e) $\sum_{i=0}^{3} (x + i)^2$

(b) $\sum_{i=5}^{8} a_i x_i$ (d) $\sum_{i=1}^{n} a_i x^{i-1}$

$\sum_{i=1}^{3} a_i x_i (x_i - 1)$

6 Rewrite the following in Σ notation:

(a) $x_1(x_1 - 1) + 2x_2(x_2 - 1) + 3x_3(x_3 - 1)$

(b) $a_2(x_3 + 2) + a_3(x_4 + 3) + a_4(x_5 + 4)$

(c) $\dfrac{1}{x} + \dfrac{1}{x^2} + \cdots + \dfrac{1}{x^n} \quad (x \neq 0)$

(d) $1 + \dfrac{1}{x} + \dfrac{1}{x^2} + \cdots + \dfrac{1}{x^n} \quad (x \neq 0)$

7 Show that the following are true:

(*a*) $\left(\sum_{i=0}^{n} x_i \right) + x_{n+1} = \sum_{i=0}^{n+1} x_i$

(*b*) $\sum_{j=1}^{n} ab_j y_j = a \sum_{j=1}^{n} b_j y_j$

(*c*) $\sum_{j=1}^{n} (x_j + y_j) = \sum_{j=1}^{n} x_j + \sum_{j=1}^{n} y_j$

4.3 NOTES ON VECTOR OPERATIONS

In the above, vectors are considered as special types of matrix. As such, they qualify for the application of all the algebraic operations discussed. Owing to their dimensional peculiarities, however, some additional comments on vector operations are useful.

Multiplication of Vectors

An $m \times 1$ column vector u, and a $1 \times n$ row vector v', yield a product matrix uv' of dimension $m \times n$.

Example 1 Given $u = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $v' = [1 \quad 4 \quad 5]$, we can get

$$uv' = \begin{bmatrix} 3(1) & 3(4) & 3(5) \\ 2(1) & 2(4) & 2(5) \end{bmatrix} = \begin{bmatrix} 3 & 12 & 15 \\ 2 & 8 & 10 \end{bmatrix}$$

Since each row in u consists of one element only, as does each column in v', each element of uv' turns out to be a single product instead of a sum of products. The product uv' is a 2×3 matrix, even though we started out only with two vectors.

On the other hand, given a $1 \times n$ row vector u' and an $n \times 1$ column vector v, the product $u'v$ will be of dimension 1×1.

Example 2 Given $u' = [3 \quad 4]$ and $v = \begin{bmatrix} 9 \\ 7 \end{bmatrix}$, we have

$$u'v = [3(9) + 4(7)] = [55]$$

As written, $u'v$ is a matrix, despite the fact that only a single element is present. However, 1×1 matrices behave exactly like scalars with respect to addition and multiplication: $[4] + [8] = [12]$, just as $4 + 8 = 12$; and $[3]\ [7] = [21]$, just as $3(7) = 21$. Moreover, 1×1 matrices possess no major properties that scalars do not have. In fact, there is a one-to-one correspondence between the set of all scalars and the set of all 1×1 matrices whose elements are scalars. For this reason, we may redefine $u'v$ to be the *scalar* corresponding to the 1×1 product matrix. For the above example, we can accordingly write $u'v = 55$. Such a product is called a *scalar product*.* Remember, however, that while a 1×1 matrix can be treated as a scalar, a scalar cannot be replaced by a 1×1 matrix at will if further calculation is to be carried out, unless conformability conditions are fulfilled.

Example 3 Given a row vector $u' = [3 \quad 6 \quad 9]$, find $u'u$. Since u is merely the column vector with the elements of u' arranged vertically, we have

$$u'u = [3 \quad 6 \quad 9]\begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} = (3)^2 + (6)^2 + (9)^2$$

where we have omitted the brackets from the 1×1 product matrix on the right. Note that the product $u'u$ gives the sum of squares of the elements of u.

In general, if $u' = [u_1 \quad u_2 \quad \cdots \quad u_n]$, then $u'u$ will be the sum of squares (a scalar) of the elements u_j:

$$u'u = u_1^2 + u_2^2 + \cdots + u_n^2 = \sum_{j=1}^{n} u_j^2$$

Had we calculated the inner product $u \cdot u$ (or $u' \cdot u'$), we would have, of course, obtained exactly the same result.

To conclude, it is important to distinguish between the meanings of uv' (a matrix larger than 1×1) and $u'v$ (a 1×1 matrix, or a scalar). Observe, in particular, that a scalar product must have a *row* vector as the lead matrix and a *column* vector as the lag matrix; otherwise the product cannot be 1×1.

Geometric Interpretation of Vector Operations

It was mentioned earlier that a column or row vector with n elements (referred to hereafter as an *n-vector*) can be viewed as an n-tuple, and hence as a point in an n-dimensional space (referred to hereafter as an n-space). Let us elaborate on this idea. In Fig. 4.3a, a point (3, 2) is plotted in a 2-space and is labeled u. This is the geometric counterpart of the vector $u = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ or the vector $u' = [3 \quad 2]$, both of which indicate in this context one and the same ordered pair. If an arrow (a directed-line segment) is drawn from the point of origin (0, 0) to the point u, it will specify the unique straight route by which to reach the destination point u from the point of origin. Since a unique arrow exists for each point, we can regard the vector u as graphically represented *either* by the point (3, 2), *or* by the

* The concept of scalar product is thus akin to the concept of inner product of two vectors with the same number of elements in each, which also yields a scalar. Recall, however, that the inner product is exempted from the conformability condition for multiplication, so that we may write it as $u \cdot v$. In the case of scalar product (denoted without a dot between the two vector symbols), on the other hand, we can express it only as a row vector multiplied by a column vector, with the row vector in the lead.

corresponding arrow. Such an arrow, which emanates from the origin (0, 0) like the hand of a clock, with a definite length and a definite direction, is called a *radius vector*.

Following this new interpretation of a vector, it becomes possible to give geometric meanings to (*a*) the scalar multiplication of a vector, (*b*) the addition and subtraction of vectors, and more generally, (*c*) the so-called "linear combination" of vectors.

First, if we plot the vector $\begin{bmatrix} 6 \\ 4 \end{bmatrix} = 2u$ in Fig. 4.3*a*, the resulting arrow will overlap the old one but will be twice as long. In fact, the multiplication of vector u by any scalar k will produce an overlapping arrow, but the arrowhead will be

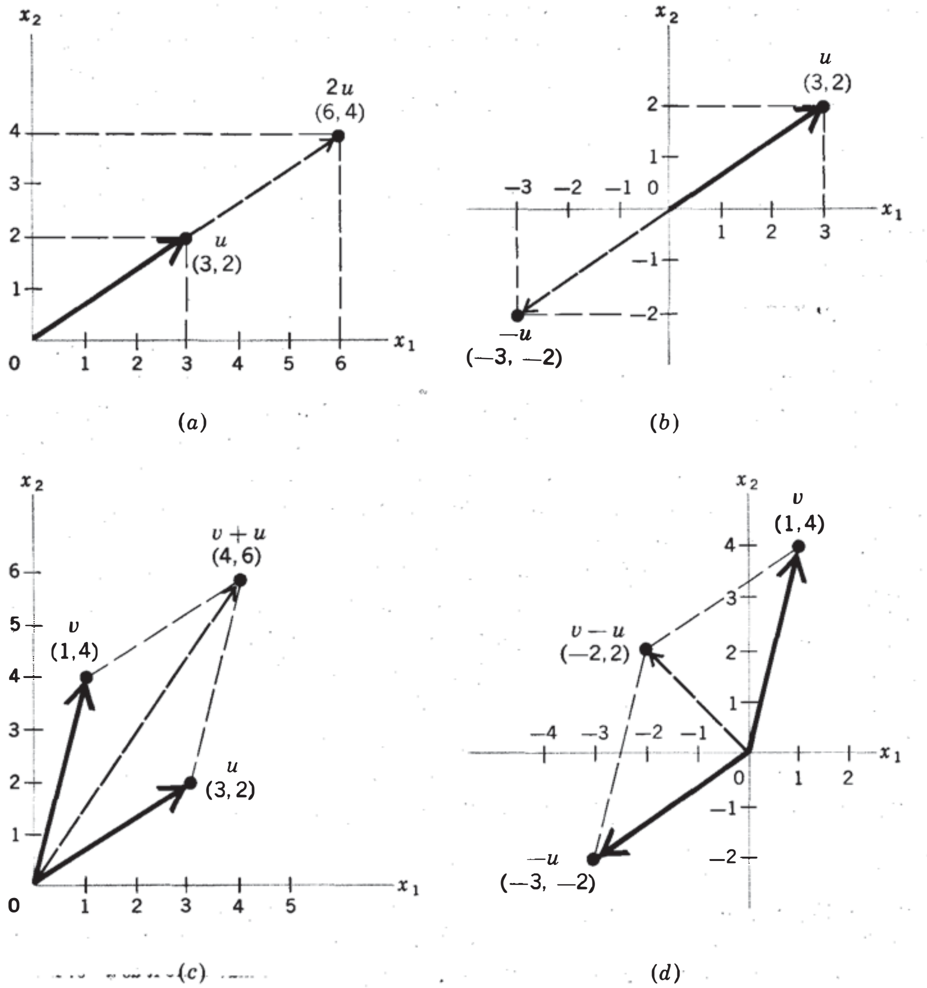

Figure 4.3

relocated, unless $k = 1$. If the scalar multiplier is $k > 1$, the arrow will be extended out (scaled up); if $0 < k < 1$, the arrow will be shortened (scaled down); if $k = 0$, the arrow will shrink into the point of origin—which represents a *null vector*, $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. A negative scalar multiplier will even reverse the direction of the arrow. If the vector u is multiplied by -1, for instance, we get $-u = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$, and this plots in Fig. 4.3*b* as an arrow of the same length as u but diametrically opposite in direction.

Next, consider the addition of two vectors, $v = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and $u = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. The sum $v + u = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ can be directly plotted as the broken arrow in Fig. 4.3*c*. If we construct a parallelogram with the two vectors u and v (solid arrows) as two of its sides, however, the diagonal of the parallelogram will turn out exactly to be the arrow representing the vector sum $v + u$. In general, a vector sum can be obtained geometrically from a parallelogram. Moreover, this method can also give us the *vector difference* $v - u$, since the latter is equivalent to the *sum* of v and $(-1)u$. In Fig. 4.3*d*, we first reproduce the vector v and the negative vector $-u$ from diagrams *c* and *b*, respectively, and then construct a parallelogram. The resulting diagonal represents the vector difference $v - u$.

It takes only a simple extension of the above results to interpret geometrically a linear combination (i.e., a linear sum or difference) of vectors. Consider the simple case of

$$3v + 2u = 3\begin{bmatrix} 1 \\ 4 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 16 \end{bmatrix}$$

The scalar multiplication aspect of this operation involves the relocation of the respective arrowheads of the two vectors v and u, and the addition aspect calls for the construction of a parallelogram. Beyond these two basic graphical operations, there is nothing new in a linear combination of vectors. This is true even if there are more terms in the linear combination, as in

$$\sum_{i=1}^{n} k_i v_i = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$$

where k_i are a set of scalars but the subscripted symbols v_i now denote a set of vectors. To form this sum, the first two terms may be added first, and then the resulting sum is added to the third, and so forth, till all terms are included.

Linear Dependence

A set of vectors $v_1, \ldots, v_n$ is said to be *linearly dependent* if (and only if) any one of them can be expressed as a linear combination of the remaining vectors; otherwise they are *linearly independent*.

Example 4 The three vectors $v_1 = \begin{bmatrix} 2 \\ 7 \end{bmatrix}$, $v_2 = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$, and $v_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ are linearly dependent because v_3 is a linear combination of v_1 and v_2:

$$3v_1 - 2v_2 = \begin{bmatrix} 6 \\ 21 \end{bmatrix} - \begin{bmatrix} 2 \\ 16 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} = v_3$$

Note that this last equation is alternatively expressible as

$$3v_1 - 2v_2 - v_3 = 0$$

where $0 \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ represents a null vector (also called *zero vector*).

Example 5 The two row vectors $v'_1 = [5 \quad 12]$ and $v'_2 = [10 \quad 24]$ are linearly dependent because

$$2v'_1 = 2[5 \quad 12] = [10 \quad 24] = v'_2$$

The fact that one vector is a multiple of another vector illustrates the simplest case of linear combination. Note again that this last equation may be written equivalently as

$$2v'_1 - v'_2 = 0'$$

where $0'$ represents the null row vector $[0 \quad 0]$.

With the introduction of null vectors, linear dependence may be redefined as follows. A set of m-vectors $v_1, \ldots, v_n$ is *linearly dependent* if and only if there exists a set of scalars $k_1, \ldots, k_n$ (not all zero) such that

$$\sum_{i=1}^{n} k_i v_i = \underset{(m \times 1)}{0}$$

If this equation can be satisfied *only when* $k_i = 0$ for all i, on the other hand, these vectors are linearly independent.

The concept of linear dependence admits of an easy geometric interpretation also. Two vectors u and $2u$—one being a multiple of the other—are obviously dependent. Geometrically, in Fig. 4.3*a*, their arrows lie on a single straight line. The same is true of the two dependent vectors u and $-u$ in Fig. 4.3*b*. In contrast, the two vectors u and v of Fig. 4.3*c* are linearly *independent*, because it is impossible to express one as a multiple of the other. Geometrically, their arrows do not lie on a single straight line.

When more than two vectors in the 2-space are considered, there emerges this significant conclusion: once we have found two linearly *independent* vectors in the 2-space (say, u and v), all the other vectors in that space will be expressible as a linear combination of these (u and v). In Fig. 4.3*c* and *d*, it has already been illustrated how the two simple linear combinations $v + u$ and $v - u$ can be found. Furthermore, by extending, shortening, and reversing the given vectors u and v and then combining these into various parallelograms, we can generate an infinite number of new vectors, which will exhaust the set of all 2-vectors. Because of this, any set of three or more 2-vectors (three or more vectors in a 2-space) must be linearly dependent. Two of them can be independent, but then the third must be a linear combination of the first two.

Vector Space

The totality of the 2-vectors generated by the various linear combinations of two independent vectors u and v constitutes the two-dimensional *vector space*. Since we are dealing only with vectors with real-valued elements, this vector space is none other than R^2, the 2-space we have been referring to all along. The 2-space cannot be generated by a single 2-vector, because "linear combinations" of the latter can only give rise to the set of vectors lying on a single straight line. Nor does the generation of the 2-space require more than two linearly independent 2-vectors—at any rate, it would be impossible to find more than two.

The two linearly independent vectors u and v are said to *span* the 2-space. They are also said to constitute a *basis* for the 2-space. Note that we said *a* basis, not *the* basis, because any pair of 2-vectors can serve in that capacity as long as they are linearly independent. In particular, consider the two vectors $[1 \quad 0]$ and $[0 \quad 1]$, which are called *unit vectors*. The first one plots as an arrow lying along the horizontal axis, and the second, an arrow lying along the vertical axis. Because they are linearly independent, they can serve as a basis for the 2-space, and we do in fact ordinarily think of the 2-space as spanned by its two axes, which are nothing but the extended versions of the two unit vectors.

By analogy, the three-dimensional vector space is the totality of 3-vectors, and it must be spanned by exactly three linearly independent 3-vectors. As an illustration, consider the set of three unit vectors

$$(4.7) \qquad e_1 \equiv \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad e_2 \equiv \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad e_3 \equiv \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

where each e_i is a vector with 1 as its ith element and with zeros elsewhere. These three vectors are obviously linearly independent; in fact, their arrows lie on the three axes of the 3-space in Fig. 4.4. Thus they span the 3-space, which implies that the entire 3-space (R^3, in our framework) can be generated from these unit vectors. For example, the vector $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ can be considered as the linear combination $e_1 + 2e_2 + 2e_3$. Geometrically, we can first add the vectors e_1 and $2e_2$ in Fig. 4.4 by the parallelogram method, in order to get the vector represented by the point $(1, 2, 0)$ in the x_1x_2 plane, and then add the latter vector to $2e_3$—via the parallelogram constructed in the shaded vertical plane—to obtain the desired final result, at the point $(1, 2, 2)$.

The further extension to n-space should be obvious. The n-space can be defined as the totality of n-vectors. Though nongraphable, we can still think of the n-space as being spanned by a total of n (n-element) unit vectors that are all

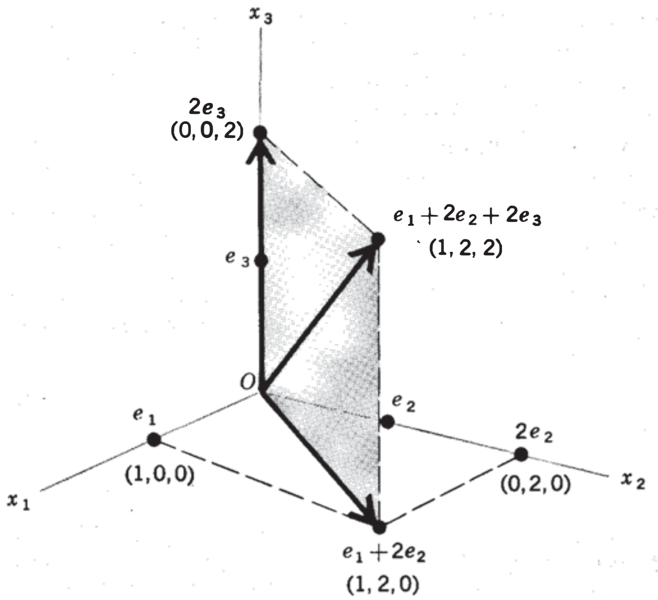

Figure 4.4

linearly independent. Each n-vector, being an ordered n-tuple, represents a *point* in the n-space, or an arrow extending from the point of origin (i.e., the n-element null vector) to the said point. And any given set of n linearly independent n-vectors is, in fact, capable of generating the entire n-space. Since, in our discussion, each element of the n-vector is restricted to be a real number, this n-space is in fact R^n.

The n-space referred to above is sometimes more specifically called the *euclidean n-space* (named after Euclid). To explain this latter concept, we must first comment briefly on the concept of *distance* between two vector points. For any pair of vector points u and v in a given space, the distance from u to v is some real-valued function

$$d = d(u, v)$$

with the following properties: (1) when u and v coincide, the distance is zero; (2) when the two points are distinct, the distance from u to v and the distance from v to u are represented by an identical positive real number; and (3) the distance between u and v is never longer than the distance from u to w (a point distinct from u and v) plus the distance from w to v. Expressed symbolically,

$$d(u, v) = 0 \qquad (\text{for } u = v)$$

$$d(u, v) = d(v, u) > 0 \qquad (\text{for } u \neq v)$$

$$d(u, v) \leq d(u, w) + d(w, v) \qquad (\text{for } w \neq u, v)$$

The last property is known as the *triangular inequality*, because the three points u, v, and w together will usually define a triangle.

When a vector space has a distance function defined that fulfills the above three properties it is called a *metric space*. However, note that the distance $d(u, v)$ has been discussed above only in general terms. Depending on the specific form assigned to the d function, there may result a variety of metric spaces. The so-called "euclidean space" is one specific type of metric space, with a distance function defined as follows. Let point u be the n-tuple $(a_1, a_2, \dots, a_n)$ and point v be the n-tuple $(b_1, b_2, \dots, b_n)$; then the euclidean distance function is

$$d(u, v) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

where the square root is taken to be positive. As can be easily verified, this specific distance function satisfies all three properties enumerated above. Applied to the two-dimensional space in Fig. 4.3a, the distance between the two points (6, 4) and (3, 2) is found to be

$$\sqrt{(6-3)^2 + (4-2)^2} = \sqrt{3^2 + 2^2} = \sqrt{13}$$

This result is seen to be consistent with *Pythagoras' theorem*, which states that the length of the hypotenuse of a right-angled triangle is equal to the (positive) square root of the sum of the squares of the lengths of the other two sides. For if we take (6,4) and (3, 2) to be u and v, and plot a new point w at (6, 2), we shall indeed have a right-angled triangle with the lengths of its horizontal and vertical sides equal to 3 and 2, respectively, and the length of the hypotenuse (the distance between u and v) equal to $\sqrt{3^2 + 2^2} = \sqrt{13}$.

The euclidean distance function can also be expressed in terms of the square root of a scalar product of two vectors. Since u and v denote the two n-tuples $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$, we can write a column vector $u - v$, with elements $a_1 - b_1, a_2 - b_2, \dots, a_n - b_n$. What goes under the square-root sign in the euclidean distance function is, of course, simply the sum of squares of these n elements, which, in view of Example 3 above, can be written as the scalar product $(u - v)'(u - v)$. Hence we have

$$d(u, v) = \sqrt{(u - v)'(u - v)}$$

EXERCISE 4.3

1 Given $u' = [5 \quad 2 \quad 3]$, $v' = [3 \quad 1 \quad 9]$, $w' = [7 \quad 5 \quad 8]$, and $x' = [x_1 \quad x_2 \quad x_3]$, write out the column vectors, u, v, w, and x, and find

(a) uv' (c) xx' (e) $u'v$ (g) $u'u$
(b) uw' (d) $v'u$ (f) $w'x$ (h) $x'x$

2 Given $w = \begin{bmatrix} 3 \\ 2 \\ 16 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, and $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$:

(*a*) Which of the following are defined: $w'x$, $x'y'$, xy', $y'y$, zz', yw', $x \cdot y$?
(*b*) Find all the products that are defined.

3 Having bought n items of merchandise at quantities $Q_1, \ldots, Q_n$ and prices $P_1, \ldots, P_n$, how would you express the total cost of purchase in (*a*) Σ notation and (*b*) vector notation?

4 Given two nonzero vectors w_1 and w_2, the angle θ ($0° \le \theta \le 180°$) they form is related to the scalar product $w_1'w_2$ ($= w_2'w_1$) as follows:

$$\theta \text{ is a(n) } \left\{ \begin{matrix} \text{acute} \\ \text{right} \\ \text{obtuse} \end{matrix} \right\} \text{ angle if and only if } w_1'w_2 \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0$$

Verify this by computing the scalar product for each of the following pair of vectors (see Figs. 4.3 and 4.4):

(*a*) $w_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $w_2 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ (*d*) $w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $w_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$

(*b*) $w_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $w_2 = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$ (*e*) $w_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, $w_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

(*c*) $w_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $w_2 = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$

5 Given $u = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, find the following graphically:

(*a*) $2v$ (*c*) $u - v$ (*e*) $2u + 3v$
(*b*) $u + v$ (*d*) $v - u$ (*f*) $4u - 2v$

6 Since the 3-space is spanned by the three unit vectors defined in (4.7), any other 3-vector should be expressible as a linear combination of e_1, e_2, and e_3. Show that the following 3-vectors can be so expressed:

(*a*) $\begin{bmatrix} 4 \\ 7 \\ 0 \end{bmatrix}$ (*b*) $\begin{bmatrix} 15 \\ -2 \\ 1 \end{bmatrix}$ (*c*) $\begin{bmatrix} -1 \\ 3 \\ 9 \end{bmatrix}$ (*d*) $\begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix}$

7 In the three-dimensional euclidean space, what is the distance between the following points?

(*a*) $(3, 2, 8)$ and $(0, -1, 5)$ (*b*) $(9, 0, 4)$ and $(2, 0, -4)$

8 The triangular inequality is written with the *weak* inequality sign $\le$, rather than the strict inequality sign $<$. Under what circumstances would the "$=$" part of the inequality apply?

9 Express the length of a radius vector v in the euclidean n-space (i.e., the distance from the origin to point v) in terms of:

(*a*) scalars (*b*) a scalar product (*c*) an inner product

4.4 COMMUTATIVE, ASSOCIATIVE, AND DISTRIBUTIVE LAWS

In ordinary scalar algebra, the additive and multiplicative operations obey the commutative, associative, and distributive laws as follows:

Commutative law of addition:	$a + b = b + a$
Commutative law of multiplication:	$ab = ba$
Associative law of addition:	$(a + b) + c = a + (b + c)$
Associative law of multiplication:	$(ab)c = a(bc)$
Distributive law:	$a(b + c) = ab + ac$

These have been referred to during the discussion of the similarly named laws applicable to the union and intersection of sets. Most, but not all, of these laws also apply to matrix operations—the significant exception being the commutative law of multiplication.

Matrix Addition

Matrix addition is commutative as well as associative. This follows from the fact that matrix addition calls only for the addition of the corresponding elements of two matrices, and that the order in which each pair of corresponding elements is added is immaterial. In this context, incidentally, the subtraction operation $A - B$ can simply be regarded as the addition operation $A + (-B)$, and thus no separate discussion is necessary.

The commutative and associative laws can be stated as follows:

Commutative law $A + B = B + A$

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

Example 1 Given $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 6 & 2 \\ 3 & 4 \end{bmatrix}$, we find that

$$A + B = B + A = \begin{bmatrix} 9 & 3 \\ 3 & 6 \end{bmatrix}$$

Associative law $(A + B) + C = A + (B + C)$

PROOF
$$\begin{aligned}(A + B) + C &= [a_{ij} + b_{ij}] + [c_{ij}] = [a_{ij} + b_{ij} + c_{ij}] \\ &= [a_{ij}] + [b_{ij} + c_{ij}] = A + (B + C)\end{aligned}$$

Example 2 Given $v_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, $v_2 = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$, and $v_3 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, we find that

$$(v_1 + v_2) - v_3 = \begin{bmatrix} 12 \\ 5 \end{bmatrix} - \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$$

which is equal to

$$v_1 + (v_2 - v_3) = \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 7 \\ -4 \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$$

Applied to the linear combination of vectors $k_1 v_1 + \cdots + k_n v_n$, this law permits us to select any pair of terms for addition (or subtraction) first, instead of having to follow the sequence in which the n terms are listed.

Matrix Multiplication

Matrix multiplication is *not* commutative, that is,

$$AB \neq BA$$

As explained previously, even when AB is defined, BA may not be; but even if both products are defined, the general rule is still $AB \neq BA$.

Example 3 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & -1 \\ 6 & 7 \end{bmatrix}$; then

$$AB = \begin{bmatrix} 1(0) + 2(6) & 1(-1) + 2(7) \\ 3(0) + 4(6) & 3(-1) + 4(7) \end{bmatrix} = \begin{bmatrix} 12 & 13 \\ 24 & 25 \end{bmatrix}$$

but

$$BA = \begin{bmatrix} 0(1) - 1(3) & 0(2) - 1(4) \\ 6(1) + 7(3) & 6(2) + 7(4) \end{bmatrix} = \begin{bmatrix} -3 & -4 \\ 27 & 40 \end{bmatrix}$$

Example 4 Let u' be 1×3 (a row vector); then the corresponding column vector u must be 3×1. The product $u'u$ will be 1×1, but the product uu' will be 3×3. Thus, obviously, $u'u \neq uu'$.

In view of the general rule $AB \neq BA$, the terms *premultiply* and *postmultiply* are often used to specify the order of multiplication. In the product AB, the matrix B is said to be *pre*multiplied by A, and A to be *post*multiplied by B.

There do exist interesting exceptions to the rule $AB \neq BA$, however. One such case is when A is a square matrix and B is an identity matrix. Another is when A is the inverse of B, that is, when $A = B^{-1}$. Both of these will be taken up again later. It should also be remarked here that the scalar multiplication of a matrix *does* obey the commutative law; thus

$$kA = Ak$$

if k is a scalar.

Although it is not in general commutative, matrix multiplication *is* associative.

Associative law $\quad (AB)C = A(BC) = ABC$

In forming the product ABC, the conformability condition must naturally be satisfied by each *adjacent* pair of matrices. If A is $m \times n$ and if C is $p \times q$, then conformability requires that B be $n \times p$:

$$\underset{(m\times n)}{A}\ \underset{(n\times p)}{B}\ \underset{(p\times q)}{C}$$

Note the dual appearance of n and p in the dimension indicators. If the conformability condition is met, the associative law states that any *adjacent* pair of matrices may be multiplied out first, provided that the product is duly inserted in the exact place of the original pair.

Example 5 If $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}$, then

$$x'Ax = x'(Ax) = [x_1 \quad x_2]\begin{bmatrix} a_{11}x_1 \\ a_{22}x_2 \end{bmatrix} = a_{11}x_1^2 + a_{22}x_2^2$$

which is a "weighted" sum of squares, in contrast to the simple sum of squares given by $x'x$. Exactly the same result comes from

$$(x'A)x = [a_{11}x_1 \quad a_{22}x_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a_{11}x_1^2 + a_{22}x_2^2$$

Matrix multiplication is also distributive.

Distributive law $A(B+C) = AB + AC$ [premultiplication by A]
$(B+C)A = BA + CA$ [postmultiplication by A]

In each case, the conformability conditions for addition as well as for multiplication must, of course, be observed.

EXERCISE 4.4

1 Given $A = \begin{bmatrix} 3 & 6 \\ 2 & 4 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 7 \\ 8 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} 3 & 4 \\ 1 & 9 \end{bmatrix}$, verify that
(a) $(A + B) + C = A + (B + C)$
(b) $(A + B) - C = A + (B - C)$

2 The subtraction of a matrix B may be considered as the addition of the matrix $(-1)B$. Does the commutative law of addition permit us to state that $A - B = B - A$? If not, how would you correct the statement?

3 Test the associative law of multiplication with the following matrices:

$$A = \begin{bmatrix} 5 & 3 \\ 0 & 5 \end{bmatrix} \qquad B = \begin{bmatrix} -8 & 0 & 7 \\ 1 & 3 & 2 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 3 \\ 7 & 1 \end{bmatrix}$$

4 Prove that for any two scalars g and k
(a) $k(A + B) = kA + kB$
(b) $(g + k)A = gA + kA$

5 Prove that $(A + B)(C + D) = AC + AD + BC + BD$.

6 If the matrix A in Example 5 had all its four elements nonzero, would $x'Ax$ still give a weighted sum of squares? Would the associative law still apply?

4.5 IDENTITY MATRICES AND NULL MATRICES

Identity Matrices

Reference has been made earlier to the term *identity matrix*. Such a matrix is defined as a *square* (repeat: square) matrix with 1s in its principal diagonal and 0s everywhere else. It is denoted by the symbol I, or I_n, in which the subscript n serves to indicate its row (as well as column) dimension. Thus,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

But both of these can also be denoted by I.

The importance of this special type of matrix lies in the fact that it plays a role similar to that of the number 1 in scalar algebra. For any number a, we have $1(a) = a(1) = a$. Similarly, for any matrix A, we have

$$(4.8) \qquad IA = AI = A$$

Example 1 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 3 \end{bmatrix}$, then

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

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

Because A is 2×3, premultiplication and postmultiplication of A by I would call for identity matrices of different dimensions, namely, I_2 and I_3, respectively. But in case A is $n \times n$, then the same identity matrix I_n can be used, so that (4.8) becomes $I_nA = AI_n$, thus illustrating an exception to the rule that matrix multiplication is not commutative.

The special nature of identity matrices makes it possible, during the multiplication process, to *insert* or *delete* an identity matrix without affecting the matrix product. This follows directly from (4.8). Recalling the associative law, we have, for instance,

$$\underset{(m\times n)}{A}\ \underset{(n\times n)}{I}\ \underset{(n\times p)}{B} = (AI)B = \underset{(m\times n)}{A}\ \underset{(n\times p)}{B}$$

which shows that the presence or absence of I does not affect the product.

Observe that dimension conformability is preserved whether or not I appears in the product.

An interesting case of (4.8) occurs when $A = I_n$, for then we have

$$AI_n = (I_n)^2 = I_n$$

which states that an identity matrix squared is equal to itself. A generalization of this result is that

$$(I_n)^k = I_n \quad (k = 1, 2, \ldots)$$

An identity matrix remains unchanged when it is multiplied by itself any number of times. Any matrix with such a property (namely, $AA = A$) is referred to as an *idempotent matrix*.

Null Matrices

Just as an identity matrix I plays the role of the number 1, a *null matrix*—or *zero matrix*—denoted by 0, plays the role of the number 0. A null matrix is simply a matrix whose elements are all zero. Unlike I, the zero matrix is not restricted to being square. Thus it is possible to write

$$\underset{(2\times 2)}{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \underset{(2\times 3)}{0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and so forth. A square null matrix is idempotent, but a nonsquare one is not. (Why?)

As the counterpart of the number 0, null matrices obey the following rules of operation (subject to conformability) with regard to addition and multiplication:

$$\underset{(m\times n)}{A} + \underset{(m\times n)}{0} = \underset{(m\times n)}{0} + \underset{(m\times n)}{A} = \underset{(m\times n)}{A}$$

$$\underset{(m\times n)}{A}\ \underset{(n\times p)}{0} = \underset{(m\times p)}{0} \quad \text{and} \quad \underset{(q\times m)}{0}\ \underset{(m\times n)}{A} = \underset{(q\times n)}{0}$$

Note that, in multiplication, the null matrix to the left of the equals sign and the one to the right may be of different dimensions.

Example 2

$$A + 0 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = A$$

Example 3

$$\underset{(2\times 3)}{A}\ \underset{(3\times 1)}{0} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \underset{(2\times 1)}{0}$$

To the left, the null matrix is a 3×1 null vector; to the right, it is a 2×1 null vector.

Idiosyncracies of Matrix Algebra

Despite the apparent similarities between matrix algebra and scalar algebra, the case of matrices does display certain idiosyncracies that serve to warn us not to "borrow" from scalar algebra too unquestioningly. We have already seen that, in general, $AB \neq BA$ in matrix algebra. Let us look at two more such idiosyncracies of matrix algebra.

For one thing, in the case of scalars, the equation $ab = 0$ always implies that either a or b is zero, but this is not so in matrix multiplication. Thus, we have

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

although neither A nor B is itself a zero matrix.

As another illustration, for scalars, the equation $cd = ce$ (with $c \neq 0$) implies that $d = e$. The same does not hold for matrices. Thus, given

$$C = \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad E = \begin{bmatrix} -2 & 1 \\ 3 & 2 \end{bmatrix}$$

we find that

$$CD = CE = \begin{bmatrix} 5 & 8 \\ 15 & 24 \end{bmatrix}$$

even though $D \neq E$.

These strange results actually pertain only to the special class of matrices known as *singular matrices*, of which the matrices A, B, and C are examples. (Roughly, these matrices contain a row which is a multiple of another row.) Nevertheless, such examples do reveal the pitfalls of unwarranted extension of algebraic theorems to matrix operations.

EXERCISE 4.5

Given $A = \begin{bmatrix} -1 & 8 & 7 \\ 0 & -2 & 4 \end{bmatrix}$, $b = \begin{bmatrix} 9 \\ 6 \\ 0 \end{bmatrix}$, and $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$:

1 Calculate: (*a*) AI (*b*) IA (*c*) Ix (*d*) $x'I$
Indicate the dimension of the identity matrix used in each case.

2 Calculate: (*a*) Ab (*b*) AIb (*c*) $x'IA$ (*d*) $x'A$
Does the insertion of I in (*b*) affect the result in (*a*)? Does the deletion of I in (*d*) affect the result in (*c*)?

3 What is the dimension of the null matrix resulting from each of the following?
(*a*) Premultiply A by a 4×2 null matrix.
(*b*) Postmultiply A by a 3×6 null matrix.
(*c*) Premultiply b by a 4×3 null matrix.
(*d*) Postmultiply x by a 1×5 null matrix.

4 Show that a *diagonal matrix*, i.e., a matrix of the form

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

can be idempotent only if each diagonal element is either 1 or 0. How many different numerical idempotent diagonal matrices of dimension $n \times n$ can be constructed altogether from the matrix above?

4.6 TRANSPOSES AND INVERSES

When the rows and columns of a matrix A are interchanged—so that its first row becomes the first column, and vice versa—we obtain the *transpose* of A, which is denoted by A' or A^T. The prime symbol is by no means new to us; it was used earlier to distinguish a row vector from a column vector. In the newly introduced terminology, a row vector x' constitutes the transpose of the column vector x. The superscript T in the alternative symbol is obviously shorthand for the word transpose.

Example 1 Given $\underset{(2\times 3)}{A} = \begin{bmatrix} 3 & 8 & -9 \\ 1 & 0 & 4 \end{bmatrix}$ and $\underset{(2\times 2)}{B} = \begin{bmatrix} 3 & 4 \\ 1 & 7 \end{bmatrix}$, we can interchange the rows and columns and write

$$\underset{(3\times 2)}{A'} = \begin{bmatrix} 3 & 1 \\ 8 & 0 \\ -9 & 4 \end{bmatrix} \qquad \text{and} \qquad \underset{(2\times 2)}{B'} = \begin{bmatrix} 3 & 1 \\ 4 & 7 \end{bmatrix}$$

By definition, if a matrix A is $m \times n$, then its transpose A' must be $n \times m$. An $n \times n$ square matrix, however, possesses a transpose with the same dimension.

Example 2 If $C = \begin{bmatrix} 9 & -1 \\ 2 & 0 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 3 & 7 \\ 4 & 7 & 2 \end{bmatrix}$, then

$$C' = \begin{bmatrix} 9 & 2 \\ -1 & 0 \end{bmatrix} \qquad \text{and} \qquad D' = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 3 & 7 \\ 4 & 7 & 2 \end{bmatrix}$$

Here, the dimension of each transpose is identical with that of the original matrix.

In D', we also note the remarkable result that D' inherits not only the dimension of D but also the original array of elements! The fact that $D' = D$ is the result of the symmetry of the elements with reference to the principal diagonal. Considering the principal diagonal in D as a mirror, the elements located to its northeast are exact images of the elements to its southwest; hence the first row reads identically with the first column, and so forth. The matrix D exemplifies the special class of square matrices known as *symmetric matrices*. Another example of such a matrix is the identity matrix I, which, as a symmetric matrix, has the transpose $I' = I$.

Properties of Transposes

The following properties characterize transposes:

(4.9) $(A')' = A$

(4.10) $(A + B)' = A' + B'$

(4.11) $(AB)' = B'A'$

The first says that the transpose of the transpose is the original matrix—a rather self-evident conclusion.

The second property may be verbally stated thus: the transpose of a sum is the sum of the transposes.

Example 3 If $A = \begin{bmatrix} 4 & 1 \\ 9 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 7 & 1 \end{bmatrix}$, then

$$(A + B)' = \begin{bmatrix} 6 & 1 \\ 16 & 1 \end{bmatrix}' = \begin{bmatrix} 6 & 16 \\ 1 & 1 \end{bmatrix}$$

and $$A' + B' = \begin{bmatrix} 4 & 9 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 7 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 16 \\ 1 & 1 \end{bmatrix}$$

The third property is that the transpose of a product is the product of the transposes *in reverse order*. To appreciate the necessity for the reversed order, let us examine the dimension conformability of the two products on the two sides of (4.11). If we let A be $m \times n$ and B be $n \times p$, then AB will be $m \times p$, and $(AB)'$ will be $p \times m$. For equality to hold, it is necessary that the right-hand expression $B'A'$ be of the identical dimension. Since B' is $p \times n$ and A' is $n \times m$, the product $B'A'$ is indeed $p \times m$, as required. The dimension of $B'A'$ thus works out. Note that, on the other hand, the product $A'B'$ is not even defined unless $m = p$.

Example 4 Given $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & -1 \\ 6 & 7 \end{bmatrix}$, we have

$$(AB)' = \begin{bmatrix} 12 & 13 \\ 24 & 25 \end{bmatrix}' = \begin{bmatrix} 12 & 24 \\ 13 & 25 \end{bmatrix}$$

and $$B'A' = \begin{bmatrix} 0 & 6 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 12 & 24 \\ 13 & 25 \end{bmatrix}$$

This verifies the property.

Inverses and Their Properties

For a given matrix A, the transpose A' is always derivable. On the other hand, its *inverse* matrix—another type of "derived" matrix—may or may not exist. The inverse of matrix A, denoted by A^{-1}, is defined only if A is a square matrix, in which case the inverse is the matrix that satisfies the condition

$$AA^{-1} = A^{-1}A = I \tag{4.12}$$

That is, whether A is pre- or postmultiplied by A^{-1}, the product will be the same identity matrix. This is another exception to the rule that matrix multiplication is not commutative.

The following points are worth noting:

1. Not every square matrix has an inverse—squareness is a *necessary* condition, but *not* a *sufficient* condition, for the existence of an inverse. If a square matrix A has an inverse, A is said to be *nonsingular*; if A possesses no inverse, it is called a *singular* matrix.
2. If A^{-1} does exist, then the matrix A can be regarded as the inverse of A^{-1}, just as A^{-1} is the inverse of A. In short, A and A^{-1} are inverses of each other.
3. If A is $n \times n$, then A^{-1} must also be $n \times n$; otherwise it cannot be conformable for *both* pre- and postmultiplication. The identity matrix produced by the multiplication will also be $n \times n$.
4. If an inverse exists, then it is unique. To prove its uniqueness, let us suppose that B has been found to be an inverse for A, so that

   $$AB = BA = I$$

   Now assume that there is another matrix C such that $AC = CA = I$. By premultiplying both sides of $AB = I$ by C, we find that

   $$CAB = CI(= C) \qquad \text{[by (4.8)]}$$

   Since $CA = I$ by assumption, the preceding equation is reducible to

   $$IB = C \qquad \text{or} \qquad B = C$$

   That is, B and C must be one and the same inverse matrix. For this reason, we can speak of *the* (as against *an*) inverse of A.
5. The two parts of condition (4.12)—namely, $AA^{-1} = I$ and $A^{-1}A = I$—actually imply each other, so that satisfying either equation is sufficient to establish the inverse relationship between A and A^{-1}. To prove this, we should show that if $AA^{-1} = I$, and if there is a matrix B such that $BA = I$, then $B = A^{-1}$ (so that $BA = I$ must in effect be the equation $A^{-1}A = I$). Let us postmultiply both sides of the given equation $BA = I$ by A^{-1}; then

   $$(BA)A^{-1} = IA^{-1}$$

   $$B(AA^{-1}) = IA^{-1} \qquad \text{[associative law]}$$

   $$BI = IA^{-1} \qquad [AA^{-1} = I \text{ by assumption}]$$

Therefore, as required,

$$B = A^{-1} \qquad [\text{by (4.8)}]$$

Analogously, it can be demonstrated that, if $A^{-1}A = I$, then the only matrix C which yields $CA^{-1} = I$ is $C = A$.

Example 5 Let $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ and $B = \frac{1}{6}\begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}$; then, since the scalar multiplier $(\frac{1}{6})$ in B can be moved to the rear (commutative law), we can write

$$AB = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}\frac{1}{6} = \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}\frac{1}{6} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This establishes B as the inverse of A, and vice versa. The reverse multiplication, as expected, also yields the same identity matrix:

$$BA = \frac{1}{6}\begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The following three properties of inverse matrices are of interest. If A and B are nonsingular matrices with dimension $n \times n$, then:

(4.13) $\qquad (A^{-1})^{-1} = A$

(4.14) $\qquad (AB)^{-1} = B^{-1}A^{-1}$

(4.15) $\qquad (A')^{-1} = (A^{-1})'$

The first says that the inverse of an inverse is the original matrix. The second states that the inverse of a product is the product of the inverses *in reverse order*. And the last one means that the inverse of the transpose is the transpose of the inverse. Note that in these statements the existence of the inverses and the satisfaction of the conformability condition are presupposed.

The validity of (4.13) is fairly obvious, but let us prove (4.14) and (4.15). Given the product AB, let us find its inverse—call it C. From (4.12) we know that $CAB = I$; thus, postmultiplication of both sides by $B^{-1}A^{-1}$ will yield

(4.16) $\qquad CABB^{-1}A^{-1} = IB^{-1}A^{-1} \; (= B^{-1}A^{-1})$

But the left side is reducible to

$$CA(BB^{-1})A^{-1} = CAIA^{-1} \qquad [\text{by (4.12)}]$$
$$= CAA^{-1} = CI = C \qquad [\text{by (4.12) and (4.8)}]$$

Substitution of this into (4.16) then tells us that $C = B^{-1}A^{-1}$ or, in other words, that the inverse of AB is equal to $B^{-1}A^{-1}$, as alleged. In this proof, the equation $AA^{-1} = A^{-1}A = I$ was utilized twice. Note that the application of this equation is permissible if and only if a matrix and its inverse are strictly adjacent to each other in a product. We may write $AA^{-1}B = IB = B$, but *never* $ABA^{-1} = B$.

The proof of (4.15) is as follows. Given A', let us find its inverse—call it D. By definition, we then have $DA' = I$. But we know that

$$(AA^{-1})' = I' = I$$

produces the same identity matrix. Thus we may write

$$\begin{aligned} DA' &= (AA^{-1})' \\ &= (A^{-1})'A' \qquad \text{[by (4.11)]} \end{aligned}$$

Postmultiplying both sides by $(A')^{-1}$, we obtain

$$DA'(A')^{-1} = (A^{-1})'A'(A')^{-1}$$

or

$$D = (A^{-1})' \qquad \text{[by (4.12)]}$$

Thus, the inverse of A' is equal to $(A^{-1})'$, as alleged.

In the proofs just presented, mathematical operations were performed on whole blocks of numbers. If those blocks of numbers had not been treated as mathematical entities (matrices), the same operations would have been much more lengthy and involved. The beauty of matrix algebra lies precisely in its simplification of such operations.

Inverse Matrix and Solution of Linear-Equation System

The application of the concept of inverse matrix to the solution of a simultaneous-equation system is immediate and direct. Referring to the equation system in (4.3), we pointed out earlier that it can be written in matrix notation as

$$\underset{(3\times 3)}{A}\ \underset{(3\times 1)}{x} = \underset{(3\times 1)}{d} \tag{4.17}$$

where A, x, and d are as defined in (4.4). Now if the inverse matrix A^{-1} exists, the premultiplication of both sides of the equation (4.17) by A^{-1} will yield

$$A^{-1}Ax = A^{-1}d$$

or

$$\underset{(3\times 1)}{x} = \underset{(3\times 3)}{A^{-1}}\ \underset{(3\times 1)}{d} \tag{4.18}$$

The left side of (4.18) is a column vector of variables, whereas the right-hand product is a column vector of certain known numbers. Thus, by definition of the equality of matrices or vectors, (4.18) shows the set of values of the variables that satisfy the equation system, i.e., the solution values. Furthermore, since A^{-1} is unique if it exists, $A^{-1}d$ must be a unique vector of solution values. We shall therefore write the x vector in (4.18) as $\bar{x}$, to indicate its status as a (unique) solution.

Methods of testing the existence of the inverse and of its calculation will be discussed in the next chapter. It may be stated here, however, that the inverse of the matrix A in (4.4) is

$$A^{-1} = \frac{1}{52}\begin{bmatrix} 18 & -16 & -10 \\ -13 & 26 & 13 \\ -17 & 18 & 21 \end{bmatrix}$$

Thus (4.18) will turn out to be

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \frac{1}{52}\begin{bmatrix} 18 & -16 & -10 \\ -13 & 26 & 13 \\ -17 & 18 & 21 \end{bmatrix}\begin{bmatrix} 22 \\ 12 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

which gives the solution: $\bar{x}_1 = 2$, $\bar{x}_2 = 3$, and $\bar{x}_3 = 1$.

The upshot is that, as one way of finding the solution of a linear-equation system $Ax = d$, where the coefficient matrix A is nonsingular, we can first find the inverse A^{-1}, and then postmultiply A^{-1} by the constant vector d. The product $A^{-1}d$ will then give the solution values of the variables.

EXERCISE 4.6

1 Given $A = \begin{bmatrix} 2 & 4 \\ -1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 8 \\ 0 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 1 & 0 & 9 \\ 6 & 1 & 1 \end{bmatrix}$, find A', B', and C'.

2 Use the matrices given in the preceding problem to verify that

(a) $(A + B)' = A' + B'$ (b) $(AC)' = C'A'$

3 Generalize the result (4.11) to the case of a product of three matrices by proving that, for any conformable matrices A, B, and C, the equation $(ABC)' = C'B'A'$ holds.

4 Given the following four matrices, test whether any one of them is the inverse of another:

$$D = \begin{bmatrix} 1 & 12 \\ 0 & 3 \end{bmatrix} \quad E = \begin{bmatrix} 1 & 1 \\ 6 & 8 \end{bmatrix} \quad F = \begin{bmatrix} 1 & -4 \\ 0 & \frac{1}{3} \end{bmatrix} \quad G = \begin{bmatrix} 4 & -\frac{1}{2} \\ -3 & \frac{1}{2} \end{bmatrix}$$

5 Generalize the result (4.14) by proving that, for any conformable nonsingular matrices A, B, and C, the equation $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

6 Let $A = I - X(X'X)^{-1}X'$.

(a) Must A be square? Must $(X'X)$ be square? Must X be square?

(b) Show that matrix A is idempotent. [*Note:* If X' and X are not square, it is inappropriate to apply (4.14).]

CHAPTER

FIVE

LINEAR MODELS AND MATRIX ALGEBRA (CONTINUED)

In Chap. 4, it was shown that a linear-equation system, however large, may be written in a compact matrix notation. Furthermore, such an equation system can be solved by finding the inverse of the coefficient matrix, provided the inverse exists. Now we must address ourselves to the questions of how to test for the existence of the inverse and how to find that inverse. Only after we have answered these questions will it be possible to apply matrix algebra meaningfully to economic models.

5.1 CONDITIONS FOR NONSINGULARITY OF A MATRIX

A given coefficient matrix A can have an inverse (i.e., can be "nonsingular") only if it is square. As was pointed out earlier, however, the squareness condition is necessary but not sufficient for the existence of the inverse A^{-1}. A matrix can be square, but singular (without an inverse) nonetheless.

Necessary versus Sufficient Conditions

The concepts of "necessary condition" and "sufficient condition" are used frequently in economics. It is important that we understand their precise meanings before proceeding further.

A necessary condition is in the nature of a prerequisite: suppose that a statement p is true *only if* another statement q is true; then q constitutes a necessary condition of p. Symbolically, we express this as follows:

$$(5.1) \qquad p \Rightarrow q$$

which is read: "p *only if* q," or alternatively, "if p, then q." It is also logically correct to interpret (5.1) to mean "p *implies* q." It may happen, of course, that we also have $p \Rightarrow w$ at the same time. Then both q and w are necessary conditions for p.

Example 1 If we let p be the statement "a person is a father" and q be the statement "a person is male," then the logical statement $p \Rightarrow q$ applies. A person is a father *only if* he is male, and to be male is a necessary condition for fatherhood. Note, however, that the converse is not true: fatherhood is not a necessary condition for maleness.

A different type of situation is that in which a statement p is true if q is true, but p can also be true when q is not true. In this case, q is said to be a sufficient condition for p. The truth of q suffices for the establishment of the truth of p, but it is not a necessary condition for p. This case is expressed symbolically by

$$(5.2) \qquad p \Leftarrow q$$

which is read: "p *if* q" (without the word "only")—or alternatively, "if q, then p," as if reading (5.2) backwards. It can also be interpreted to mean "q implies p."

Example 2 If we let p be the statement "one can get to Europe" and q be the statement "one takes a plane to Europe," then $p \Leftarrow q$. Flying can serve to get one to Europe, but since ocean transportation is also feasible, flying is not a prerequisite. We can write $p \Leftarrow q$, but not $p \Rightarrow q$.

In a third possible situation, q is *both* necessary and sufficient for p. In such an event, we write

$$(5.3) \qquad p \Leftrightarrow q$$

which is read: "p *if and only if* q" (also written as "p iff q"). The double-headed arrow is really a combination of the two types of arrow in (5.1) and (5.2); hence the joint use of the two terms "if" and "only if." Note that (5.3) states not only that p implies q but also that q implies p.

Example 3 If we let p be the statement "there are less than 30 days in the month" and q be the statement "it is the month of February," then $p \Leftrightarrow q$. To have less than 30 days in the month, it is necessary that it be February. Conversely, the specification of February is sufficient to establish that there are less than 30 days in the month. Thus q is a necessary-and-sufficient condition for p.

In order to prove $p \Rightarrow q$, it needs to be shown that q follows logically from p. Similarly, to prove $p \Leftarrow q$ requires a demonstration that p follows logically from q. But to prove $p \Leftrightarrow q$ necessitates a demonstration that p and q follow from each other.

Conditions for Nonsingularity

When the squareness condition is already met, a sufficient condition for the nonsingularity of a matrix is that its rows be linearly independent (or, what amounts to the same thing, that its *columns* be linearly independent). When the dual conditions of squareness and linear independence are taken together, they constitute the necessary-and-sufficient condition for nonsingularity (nonsingularity $\Leftrightarrow$ squareness *and* linear independence).

An $n \times n$ coefficient matrix A can be considered as an ordered set of row vectors, i.e., as a column vector whose elements are themselves row vectors:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{bmatrix}$$

where $v'_i = [a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}]$, $i = 1, 2, \ldots, n$. For the rows (row vectors) to be linearly independent, none must be a linear combination of the rest. More formally, as was mentioned in Sec. 4.3, linear row independence requires that the only set of scalars k_i which can satisfy the vector equation

$$\text{(5.4)} \qquad \sum_{i=1}^{n} k_i v'_i = \underset{(1 \times n)}{0}$$

be $k_i = 0$ for all i.

Example 4 If the coefficient matrix is

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 6 & 8 & 10 \end{bmatrix} = \begin{bmatrix} v'_1 \\ v'_2 \\ v'_3 \end{bmatrix}$$

then, since $[6 \quad 8 \quad 10] = 2[3 \quad 4 \quad 5]$, we have $v'_3 = 2v'_1 = 2v'_1 + 0v'_2$. Thus the third row is expressible as a linear combination of the first two, and the rows are *not* linearly independent. Alternatively, we may write the above equation as

$$2v'_1 + 0v'_2 - v'_3 = [6 \quad 8 \quad 10] + [0 \quad 0 \quad 0] - [6 \quad 8 \quad 10] = [0 \quad 0 \quad 0]$$

Inasmuch as the set of scalars that led to the zero vector of (5.4) is not $k_i = 0$ for all i, it follows that the rows are linearly dependent.

Unlike the squareness condition, the linear-independence condition cannot normally be ascertained at a glance. Thus a method of testing linear independence among rows (or columns) needs to be developed. Before we concern ourselves with that task, however, it would strengthen our motivation first to have an intuitive understanding of why the linear-independence condition is heaped together with the squareness condition at all. From the discussion of counting equations and unknowns in Sec. 3.4, we recall the general conclusion that, for a system of equations to possess a unique solution, it is not sufficient to have the same number of equations as unknowns. In addition, the *equations* must be consistent with and functionally independent (meaning, in the present context of linear systems, *linearly* independent) of one another. There is a fairly obvious tie-in between the "same number of equations as unknowns" criterion and the *squareness* (same number of rows and columns) of the coefficient matrix. What the "linear independence among the rows" requirement does is to preclude the inconsistency and the linear dependence *among the equations* as well. Taken together, therefore, the dual requirement of squareness and row independence in the coefficient matrix is tantamount to the conditions for the existence of a unique solution enunciated in Sec. 3.4.

Let us illustrate how the linear dependence *among the rows* of the coefficient matrix can cause inconsistency or linear dependence *among the equations* themselves. Let the equation system $Ax = d$ take the form

$$\begin{bmatrix} 10 & 4 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

where the coefficient matrix A contains linearly dependent rows: $v_1' = 2v_2'$. (Note that its columns are also dependent, the first being $\frac{5}{2}$ of the second.) We have not specified the values of the constant terms d_1 and d_2, but there are only *two* distinct possibilities regarding their relative values: (1) $d_1 = 2d_2$ and (2) $d_1 \neq 2d_2$. Under the first—with, say, $d_1 = 12$ and $d_2 = 6$—the two equations are consistent but *linearly dependent* (just as the two rows of matrix A are), for the first equation is merely the second equation times 2. One equation is then redundant, and the system reduces in effect to a single equation, $5x_1 + 2x_2 = 6$, with an infinite number of solutions. For the second possibility—with, say, $d_1 = 12$ but $d_2 = 0$—the two equations are *inconsistent*, because if the first equation ($10x_1 + 4x_2 = 12$) is true, then, by halving each term, we can deduce that $5x_1 + 2x_2 = 6$; consequently the second equation ($5x_1 + 2x_2 = 0$) cannot possibly be true also. Thus no solution exists.

The upshot is that no unique solution will be available (under either possibility) so long as the rows in the coefficient matrix A are linearly dependent. In fact, the only way to have a unique solution is to have linearly independent rows (or columns) in the coefficient matrix. In that case, matrix A will be nonsingular, which means that the inverse A^{-1} does exist, and that a unique solution $\bar{x} = A^{-1}d$ can be found.

Rank of a Matrix

Even though the concept of row independence has been discussed only with regard to square matrices, it is equally applicable to any $m \times n$ rectangular matrix. If the maximum number of linearly independent rows that can be found in such a matrix is r, the matrix is said to be of *rank* r. (The rank also tells us the maximum number of linearly independent *columns* in the said matrix.) The rank of an $m \times n$ matrix can be at most m or n, whichever is smaller.

By definition, an $n \times n$ nonsingular matrix A has n linearly independent rows (or columns); consequently it must be of rank n. Conversely, an $n \times n$ matrix having rank n must be nonsingular.

EXERCISE 5.1

1 In the following paired statements, let p be the first statement and q the second. Indicate for each case whether (5.1) or (5.2) or (5.3) applies.

(a) It is a holiday; it is Thanksgiving Day.
(b) A geometric figure has four sides; it is a rectangle.
(c) Two ordered pairs (a, b) and (b, a) are equal; a is equal to b.
(d) A number is rational; it can be expressed as a ratio of two integers.
(e) A 4×4 matrix is nonsingular; the rank of the matrix is 4.
(f) The gasoline tank in my car is empty; I cannot start my car.
(g) The letter is returned to the sender for insufficient postage: the sender forgot to put a stamp on the envelope.

2 Let p be the statement "a geometric figure is a square," and let q be as follows:

(a) It has four sides.
(b) It has four equal sides.
(c) It has four equal sides each perpendicular to the adjacent one.

Which is true for each case: $p \Rightarrow q$, $p \Leftarrow q$, or $p \Leftrightarrow q$?

3 Are the rows linearly independent in each of the following?

(a) $\begin{bmatrix} 1 & 8 \\ 9 & -3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 2 \\ 3 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} -1 & 4 \\ 2 & -8 \end{bmatrix}$

4 Check whether the columns of each matrix in the preceding problem are also linearly independent. Do you get the same answer as for row independence?

5.2 TEST OF NONSINGULARITY BY USE OF DETERMINANT

To ascertain whether a square matrix is nonsingular, we can make use of the concept of determinant.

Determinants and Nonsingularity

The determinant of a square matrix A, denoted by $|A|$, is a uniquely defined scalar (number) associated with that matrix. Determinants are defined only for *square* matrices. For a 2×2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, its determinant is defined to be the sum of two terms as follows:

$$(5.5) \qquad |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} \quad [= \text{a scalar}]$$

which is obtained by multiplying the two elements in the principal diagonal of A and then subtracting the product of the two remaining elements. In view of the dimension of matrix A, $|A|$ as defined in (5.5) is called a *second-order determinant.*

Example 1 Given $A = \begin{bmatrix} 10 & 4 \\ 8 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 5 \\ 0 & -1 \end{bmatrix}$, their determinants are:

$$|A| = \begin{vmatrix} 10 & 4 \\ 8 & 5 \end{vmatrix} = 10(5) - 8(4) = 18$$

and
$$|B| = \begin{vmatrix} 3 & 5 \\ 0 & -1 \end{vmatrix} = 3(-1) - 0(5) = -3$$

While a determinant (enclosed by two vertical bars rather than brackets) is by definition a scalar, a matrix as such does not have a numerical value. In other words, a determinant is reducible to a number, but a matrix is, in contrast, a whole block of numbers. It should also be emphasized that a determinant is defined only for a square matrix, whereas a matrix as such does not have to be square.

Even at this early stage of discussion, it is possible to have an inkling of the relationship between the linear dependence of the rows in a matrix A, on the one hand, and its determinant $|A|$, on the other. The two matrices

$$C = \begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = \begin{bmatrix} 3 & 8 \\ 3 & 8 \end{bmatrix} \qquad \text{and} \qquad D = \begin{bmatrix} d_1' \\ d_2' \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 8 & 24 \end{bmatrix}$$

both have linearly dependent rows, because $c_1' = c_2'$ and $d_2' = 4d_1'$. Both of their determinants also turn out to be equal to zero:

$$|C| = \begin{vmatrix} 3 & 8 \\ 3 & 8 \end{vmatrix} = 3(8) - 3(8) = 0$$

$$|D| = \begin{vmatrix} 2 & 6 \\ 8 & 24 \end{vmatrix} = 2(24) - 8(6) = 0$$

This result strongly suggests that a "vanishing" determinant (a zero-value determinant) may have something to do with linear dependence. We shall see that this is indeed the case. Furthermore, the value of a determinant $|A|$ can serve not only as a criterion for testing the linear independence of the rows (hence the nonsingularity) of matrix A, but also as an input in the calculation of the inverse A^{-1}, if it exists.

First, however, we must widen our vista by a discussion of higher-order determinants.

Evaluating a Third-Order Determinant

A determinant of order 3 is associated with a 3×3 matrix. Given

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

its determinant has the value

$$(5.6) \qquad |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \qquad [= \text{a scalar}]$$

Looking first at the lower line of (5.6), we see the value of $|A|$ expressed as a sum of six product terms, three of which are prefixed by minus signs and three by plus signs. Complicated as this sum may appear, there is nonetheless a very easy way of "catching" all these six terms from a given third-order determinant. This is best explained diagrammatically (Fig. 5.1). In the determinant shown in Fig. 5.1,

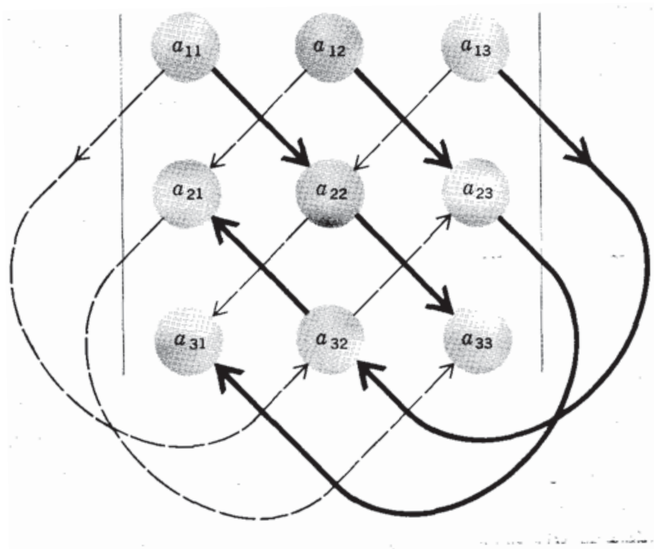

Figure 5.1

each element in the top row has been linked with two other elements via two *solid* arrows as follows: $a_{11} \to a_{22} \to a_{33}$, $a_{12} \to a_{23} \to a_{31}$, and $a_{13} \to a_{32} \to a_{21}$. Each triplet of elements so linked can be multiplied out, and their product be taken as one of the six product terms in (5.6). The solid-arrow product terms are to be prefixed with plus signs.

On the other hand, each top-row element has also been connected with two other elements via two *broken* arrows as follows: $a_{11} \to a_{32} \to a_{23}$, $a_{12} \to a_{21} \to a_{33}$, and $a_{13} \to a_{22} \to a_{31}$. Each triplet of elements so connected can also be multiplied out, and their product taken as one of the six terms in (5.6). Such products are prefixed by minus signs. The sum of all the six products will then be the value of the determinant.

Example 2

$$\begin{vmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = (2)(5)(9) + (1)(6)(7) + (3)(8)(4) - (2)(8)(6) - (1)(4)(9) - (3)(5)(7) = -9$$

Example 3

$$\begin{vmatrix} -7 & 0 & 3 \\ 9 & 1 & 4 \\ 0 & 6 & 5 \end{vmatrix} = (-7)(1)(5) + (0)(4)(0) + (3)(6)(9) - (-7)(6)(4) - (0)(9)(5) - (3)(1)(0) = 295$$

This method of cross-diagonal multiplication provides a handy way of evaluating a third-order determinant, but unfortunately it is *not* applicable to determinants of orders higher than 3. For the latter, we must resort to the so-called "Laplace expansion" of the determinant.

Evaluating an *n*th-Order Determinant by Laplace Expansion

Let us first explain the *Laplace-expansion* process for a third-order determinant. Returning to the first line of (5.6), we see that the value of $|A|$ can also be regarded as a sum of *three* terms, each of which is a product of a first-row element and a particular *second*-order determinant. This latter process of evaluating $|A|$—by means of certain lower-order determinants—illustrates the Laplace expansion of the determinant.

The three second-order determinants in (5.6) are not arbitrarily determined, but are specified by means of a definite rule. The first one, $\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$, is a *sub*determinant of $|A|$ obtained by deleting the *first* row and *first* column of $|A|$. This is called the *minor* of the element a_{11} (the element at the intersection of the deleted row and column) and is denoted by $|M_{11}|$. In general, the symbol $|M_{ij}|$ can be used to represent the minor obtained by deleting the ith row and jth column of a given determinant. Since a minor is itself a determinant, it has a value. As the reader can verify, the other two second-order determinants in (5.6) are, respectively, the minors $|M_{12}|$ and $|M_{13}|$; that is,

$$|M_{11}| \equiv \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \qquad |M_{12}| \equiv \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \qquad |M_{13}| \equiv \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

A concept closely related to the minor is that of the *cofactor*. A cofactor, denoted by $|C_{ij}|$, is a minor with a prescribed algebraic sign attached to it.* The rule of sign is as follows. If the sum of the two subscripts i and j in the minor $|M_{ij}|$ is even, then the cofactor takes the same sign as the minor; that is, $|C_{ij}| \equiv |M_{ij}|$. If it is odd, then the cofactor takes the opposite sign to the minor; that is, $|C_{ij}| \equiv -|M_{ij}|$. In short, we have

$$|C_{ij}| \equiv (-1)^{i+j}|M_{ij}|$$

where it is obvious that the expression $(-1)^{i+j}$ can be positive if and only if $(i+j)$ is even. The fact that a cofactor has a specific sign is of extreme importance and should always be borne in mind.

Example 4 In the determinant $\begin{vmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{vmatrix}$, the minor of the element 8 is

$$|M_{12}| = \begin{vmatrix} 6 & 4 \\ 3 & 1 \end{vmatrix} = -6$$

but the cofactor of the same element is

$$|C_{12}| = -|M_{12}| = 6$$

because $i + j = 1 + 2 = 3$ is odd. Similarly, the cofactor of the element 4 is

$$|C_{23}| = -|M_{23}| = -\begin{vmatrix} 9 & 8 \\ 3 & 2 \end{vmatrix} = 6$$

Using these new concepts, we can express a third-order determinant as

$$(5.7) \qquad |A| = a_{11}|M_{11}| - a_{12}|M_{12}| + a_{13}|M_{13}|$$
$$= a_{11}|C_{11}| + a_{12}|C_{12}| + a_{13}|C_{13}| = \sum_{j=1}^{3} a_{1j}|C_{1j}|$$

i.e., as a sum of three terms, each of which is the product of a first-row element and its corresponding cofactor. Note the difference in the signs of the $a_{12}|M_{12}|$ and $a_{12}|C_{12}|$ terms in (5.7). This is because 1 + 2 gives an odd number.

The Laplace expansion of a *third*-order determinant serves to reduce the evaluation problem to one of evaluating only certain *second*-order determinants.

* Many writers use the symbols M_{ij} and C_{ij} (without the vertical bars) for minors and cofactors. We add the vertical bars to give visual emphasis to the fact that minors and cofactors are in the nature of determinants and, as such, have scalar values.

A similar reduction is achieved in the Laplace expansion of higher-order determinants. In a fourth-order determinant $|B|$, for instance, the top row will contain four elements, $b_{11} \dots b_{14}$; thus, in the spirit of (5.7), we may write

$$|B| = \sum_{j=1}^{4} b_{1j}|C_{1j}|$$

where the cofactors $|C_{1j}|$ are of order 3. Each third-order cofactor can then be evaluated as in (5.6). In general, the Laplace expansion of an nth-order determinant will reduce the problem to one of evaluating n cofactors, each of which is of the $(n-1)$st order, and the repeated application of the process will methodically lead to lower and lower orders of determinants, eventually culminating in the basic second-order determinants as defined in (5.5). Then the value of the original determinant can be easily calculated.

Although the process of Laplace expansion has been couched in terms of the cofactors of the first-row elements, it is also feasible to expand a determinant by the cofactor of any row or, for that matter, of any column. For instance, if the first column of a third-order determinant $|A|$ consists of the elements a_{11}, a_{21}, and a_{31}, expansion by the cofactors of these elements will also yield the value of $|A|$:

$$|A| = a_{11}|C_{11}| + a_{21}|C_{21}| + a_{31}|C_{31}| = \sum_{i=1}^{3} a_{i1}|C_{i1}|$$

Example 5 Given $|A| = \begin{vmatrix} 5 & 6 & 1 \\ 2 & 3 & 0 \\ 7 & -3 & 0 \end{vmatrix}$, expansion by the first *row* produces the result

$$|A| = 5\begin{vmatrix} 3 & 0 \\ -3 & 0 \end{vmatrix} - 6\begin{vmatrix} 2 & 0 \\ 7 & 0 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ 7 & -3 \end{vmatrix} = 0 + 0 - 27 = -27$$

But expansion by the first *column* yields the identical answer:

$$|A| = 5\begin{vmatrix} 3 & 0 \\ -3 & 0 \end{vmatrix} - 2\begin{vmatrix} 6 & 1 \\ -3 & 0 \end{vmatrix} + 7\begin{vmatrix} 6 & 1 \\ 3 & 0 \end{vmatrix} = 0 - 6 - 21 = -27$$

Insofar as numerical calculation is concerned, this fact affords us an opportunity to choose some "easy" row or column for expansion. A row or column with the largest number of 0s or 1s is always preferable for this purpose, because a 0 times its cofactor is simply 0, so that the term will drop out, and a 1 times its cofactor is simply the cofactor itself, so that at least one multiplication step can be saved. In Example 5, the easiest way to expand the determinant is by the third column, which consists of the elements 1, 0, and 0. We could therefore have evaluated it thus:

$$|A| = 1\begin{vmatrix} 2 & 3 \\ 7 & -3 \end{vmatrix} - 0 + 0 = -27$$

To sum up, the value of a determinant $|A|$ of order n can be found by the Laplace expansion of *any row* or *any column* as follows:

$$(5.8) \qquad |A| = \sum_{j=1}^{n} a_{ij}|C_{ij}| \qquad \text{[expansion by the } i\text{th row]}$$

$$= \sum_{i=1}^{n} a_{ij}|C_{ij}| \qquad \text{[expansion by the } j\text{th column]}$$

EXERCISE 5.2

1 Evaluate the following determinants:

$(a)\ \begin{vmatrix} 8 & 1 & 3 \\ 4 & 0 & 1 \\ 6 & 0 & 3 \end{vmatrix}$ $\quad (c)\ \begin{vmatrix} 4 & 0 & 2 \\ 6 & 0 & 3 \\ 8 & 2 & 3 \end{vmatrix}$ $\quad (e)\ \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$

$(b)\ \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & 5 \\ 3 & 6 & 9 \end{vmatrix}$ $\quad (d)\ \begin{vmatrix} 1 & 1 & 4 \\ 8 & 11 & -2 \\ 0 & 4 & 7 \end{vmatrix}$ $\quad (f)\ \begin{vmatrix} x & 5 & 0 \\ 3 & y & 2 \\ 9 & -1 & 8 \end{vmatrix}$

2 Determine the signs to be attached to the relevant minors in order to get the following cofactors of a determinant: $|C_{13}|$, $|C_{23}|$, $|C_{33}|$, $|C_{41}|$, and $|C_{34}|$.

3 Given $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$, find the minors and cofactors of the elements a, b, and f.

4 Evaluate the following determinants:

$(a)\ \begin{vmatrix} 1 & 2 & 0 & 9 \\ 2 & 3 & 4 & 6 \\ 1 & 6 & 0 & -1 \\ 0 & -5 & 0 & 8 \end{vmatrix}$ $\quad (b)\ \begin{vmatrix} 2 & 7 & 0 & 1 \\ 5 & 6 & 4 & 8 \\ 0 & 0 & 9 & 0 \\ 1 & -3 & 1 & 4 \end{vmatrix}$

5 In the first determinant of the preceding problem, find the value of the cofactor of the element 9.

5.3 BASIC PROPERTIES OF DETERMINANTS

We can now discuss some properties of determinants which will enable us to "discover" the connection between linear dependence among the rows of a square matrix and the vanishing of the determinant of that matrix.

Five basic properties will be discussed here. These are properties common to determinants of all orders, although we shall illustrate mostly with second-order determinants:

Property I The interchange of rows and columns does not affect the value of a determinant. In other words, the determinant of a matrix A has the same value as that of its transpose A', that is, $|A| = |A'|$.

Example 1 $\begin{vmatrix} 4 & 3 \\ 5 & 6 \end{vmatrix} = \begin{vmatrix} 4 & 5 \\ 3 & 6 \end{vmatrix} = 9$

Example 2 $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc$

Property II The interchange of any *two* rows (or any *two* columns) will alter the sign, but not the numerical value, of the determinant.

Example 3 $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$, but the interchange of the two rows yields

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - ad = -(ad - bc)$$

Example 4 $\begin{vmatrix} 0 & 1 & 3 \\ 2 & 5 & 7 \\ 3 & 0 & 1 \end{vmatrix} = -26$, but the interchange of the first and third columns yields $\begin{vmatrix} 3 & 1 & 0 \\ 7 & 5 & 2 \\ 1 & 0 & 3 \end{vmatrix} = 26.$

Property III The multiplication of any *one* row (or *one* column) by a scalar k will change the value of the determinant k-fold.

Example 5 By multiplying the top row of the determinant in Example 3 by k, we get

$$\begin{vmatrix} ka & kb \\ c & d \end{vmatrix} = kad - kbc = k(ad - bc) = k\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

It is important to distinguish between the two expressions kA and $k|A|$. In multiplying a *matrix* A by a scalar k, all the elements in A are to be multiplied by k. But, if we read the equation in the present example from right to left, it should be clear that, in multiplying a *determinant* $|A|$ by k, only a single row (or column) should be multiplied by k. This equation, therefore, in effect gives us a rule for factoring a determinant: whenever any single row or column contains a common divisor, it may be factored out of the determinant.

Example 6 Factoring the first column and the second row in turn, we have

$$\begin{vmatrix} 15a & 7b \\ 12c & 2d \end{vmatrix} = 3\begin{vmatrix} 5a & 7b \\ 4c & 2d \end{vmatrix} = 3(2)\begin{vmatrix} 5a & 7b \\ 2c & d \end{vmatrix} = 6(5ad - 14bc)$$

The direct evaluation of the original determinant will, of course, produce the same answer.

In contrast, the factoring of a *matrix* requires the presence of a common divisor for *all* its elements, as in

$$\begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix} = k\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Property IV The addition (subtraction) of a multiple of any row to (from) another row will leave the value of the determinant unaltered. The same holds true if we replace the word *row* by *column* in the above statement.

Example 7 Adding k times the top row of the determinant in Example 3 to its second row, we end up with the original determinant:

$$\begin{vmatrix} a & b \\ c + ka & d + kb \end{vmatrix} = a(d + kb) - b(c + ka) = ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Property V If one row (or column) is a multiple of another row (or column), the value of the determinant will be zero. As a special case of this, when two rows (or two columns) are *identical*, the determinant will vanish.

Example 8

$$\begin{vmatrix} 2a & 2b \\ a & b \end{vmatrix} = 2ab - 2ab = 0 \qquad \begin{vmatrix} c & c \\ d & d \end{vmatrix} = cd - cd = 0$$

Additional examples of this type of "vanishing" determinants can be found in Exercise 5.2-1.

This important property is, in fact, a logical consequence of Property IV. To understand this, let us apply Property IV to the two determinants in Example 8 and watch the outcome. For the first one, try to subtract twice the second row from the top row; for the second determinant, subtract the second column from the first column. Since these operations do not alter the values of the determinants, we can write

$$\begin{vmatrix} 2a & 2b \\ a & b \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ a & b \end{vmatrix} \qquad \begin{vmatrix} c & c \\ d & d \end{vmatrix} = \begin{vmatrix} 0 & c \\ 0 & d \end{vmatrix}$$

The new (reduced) determinants now contain, respectively, a row and a column of zeros; thus their Laplace expansion must yield a value of zero in both cases. In general, when one row (column) is a multiple of another row (column), the application of Property IV can always reduce all elements of that row (column) to zero, and Property V therefore follows.

The basic properties just discussed are useful in several ways. For one thing, they can be of great help in simplifying the task of evaluating determinants. By subtracting multiples of one row (or column) from another, for instance, the elements of the determinant may be reduced to much smaller and simpler numbers. Factoring, if feasible, can also accomplish the same. If we can indeed apply these properties to transform some row or column into a form containing mostly 0s or 1s, Laplace expansion of the determinant will become a much more manageable task.

Determinantal Criterion for Nonsingularity

Our present concern, however, is primarily to link the linear dependence of rows with the vanishing of a determinant. For this purpose, Property V can be invoked. Consider an equation system $Ax = d$:

$$\begin{bmatrix} 3 & 4 & 2 \\ 15 & 20 & 10 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

This system can have a unique solution if and only if the rows in the coefficient matrix A are linearly independent, so that A is nonsingular. But the second row is five times the first; the rows are indeed *dependent*, and hence no unique solution exists. The detection of this row dependence was by visual inspection, but by virtue of Property V we could also have discovered it through the fact that $|A| = 0$.

The row dependence in a matrix may, of course, assume a more intricate and secretive pattern. For instance, in the matrix

$$B = \begin{bmatrix} 4 & 1 & 2 \\ 5 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} v_1' \\ v_2' \\ v_3' \end{bmatrix}$$

there exists row dependence because $2v_1' - v_2' - 3v_3' = 0$; yet this fact defies visual detection. Even in this case, however, Property V will give us a vanishing determinant, $|B| = 0$, since by adding three times v_3' to v_2' and subtracting twice v_1' from it, the second row can be reduced to a zero vector. In general, *any* pattern of linear dependence among rows will be reflected in a vanishing determinant—and herein lies the beauty of Property V! Conversely, if the rows are linearly independent, the determinant must have a nonzero value.

We have, in the above, tied the nonsingularity of a matrix principally to the linear independence among *rows*. But, on occasion, we have made the claim that, for a *square* matrix A, row independence $\Leftrightarrow$ column independence. We are now equipped to prove that claim:

> According to Property I, we know that $|A| = |A'|$. Since row independence in $A \Leftrightarrow |A| \neq 0$, we may also state that row independence in $A \Leftrightarrow |A'| \neq 0$. But $|A'| \neq 0 \Leftrightarrow$ row independence in the transpose $A' \Leftrightarrow$ column independence in A (rows of A' are by definition the columns of A). Therefore, *row* independence in $A \Leftrightarrow$ *column* independence in A.

Our discussion of the test of nonsingularity can now be summarized. Given a linear-equation system $Ax = d$, where A is an $n \times n$ coefficient matrix,

$$
\begin{aligned}
|A| \neq 0 &\Leftrightarrow \text{there is row (column) independence in matrix } A \\
&\Leftrightarrow A \text{ is nonsingular} \\
&\Leftrightarrow A^{-1} \text{ exists} \\
&\Leftrightarrow \text{a unique solution } \bar{x} = A^{-1}d \text{ exists}
\end{aligned}
$$

Thus the value of the determinant of the coefficient matrix, $|A|$, provides a convenient criterion for testing the nonsingularity of matrix A and the existence of a unique solution to the equation system $Ax = d$. Note, however, that the determinantal criterion says nothing about the algebraic signs of the solution values, i.e., even though we are assured of a unique solution when $|A| \neq 0$, we may sometimes get negative solution values that are economically inadmissible.

Example 9 Does the equation system

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

possess a unique solution? The determinant $|A|$ is

$-14+(-6)+12 = -8$

$$
\begin{vmatrix} 7 & -3 & -3 \\ 2 & 4 & 1 \\ 0 & -2 & -1 \end{vmatrix} = -8 \neq 0
$$

Therefore a unique solution does exist.

Rank of a Matrix Redefined

The rank of a matrix A was earlier defined to be the maximum number of linearly independent rows in A. In view of the link between row independence and the nonvanishing of the determinant, we can redefine the rank of an $m \times n$ matrix as the maximum order of a nonvanishing determinant that can be constructed from the rows and columns of that matrix. The rank of any matrix is a unique number.

Obviously, the rank can at most be m or n, whichever is smaller, because a determinant is defined only for a square matrix, and from a matrix of dimension, say, 3×5, the largest possible determinants (vanishing or not) will be of order 3. Symbolically, this fact may be expressed as follows:

$$r(A) \leq \min\{m, n\}$$

which is read: "The rank of A is less than or equal to the minimum of the set of two numbers m and n." The rank of an $n \times n$ nonsingular matrix A must be n; in that case, we may write $r(A) = n$.

Sometimes, one may be interested in the rank of the product of two matrices. In that case, the following rule is of use:

$$r(AB) \le \min\{r(A), r(B)\}$$

EXERCISE 5.3

1 Use the determinant $\begin{vmatrix} 2 & 0 & -1 \\ 1 & 1 & 7 \\ 3 & 3 & 9 \end{vmatrix}$ to verify the first four properties of determinants.

2 Show that, when all the elements of an nth-order determinant $|A|$ are multiplied by a number k, the result will be $k^n|A|$.

3 Which properties of determinants enable us to write the following?

(*a*) $\begin{vmatrix} 9 & 18 \\ 27 & 56 \end{vmatrix} = \begin{vmatrix} 9 & 18 \\ 0 & 2 \end{vmatrix}$ (*b*) $\begin{vmatrix} 9 & 27 \\ 4 & 2 \end{vmatrix} = 18\begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix}$

4 Test whether the following matrices are nonsingular:

(*a*) $\begin{bmatrix} 4 & 0 & 1 \\ 19 & 1 & 3 \\ 5 & 4 & 7 \end{bmatrix}$ (*c*) $\begin{bmatrix} 7 & -1 & 0 \\ 1 & 1 & 4 \\ 13 & -3 & -4 \end{bmatrix}$

(*b*) $\begin{bmatrix} 4 & -2 & 1 \\ -5 & 6 & 0 \\ 7 & 0 & 3 \end{bmatrix}$ (*d*) $\begin{bmatrix} 7 & 9 & 5 \\ 3 & 0 & 1 \\ 10 & 8 & 6 \end{bmatrix}$

5 What can you conclude about the rank of each matrix in the preceding problem?

6 Can any set of 3-vectors below span the 3-space? Why or why not?

(*a*) $[1 \ \ 2 \ \ 0]$ $[2 \ \ 3 \ \ 1]$ $[3 \ \ 4 \ \ 2]$

(*b*) $[8 \ \ 1 \ \ 3]$ $[1 \ \ 2 \ \ 8]$ $[-7 \ \ 1 \ \ 5]$

7 Rewrite the simple national-income model (3.23) in the $Ax = d$ format (with Y as the first variable in the vector x), and then test whether the coefficient matrix A is nonsingular.

5.4 FINDING THE INVERSE MATRIX

If the matrix A in the linear-equation system $Ax = d$ is nonsingular, then A^{-1} exists, and the solution of the system will be $\bar{x} = A^{-1}d$. We have learned to test the nonsingularity of A by the criterion $|A| \neq 0$. The next question is: How can we find the inverse A^{-1} if A does pass that test?

Expansion of a Determinant by Alien Cofactors

Before answering this query, let us discuss another important property of determinants.

Property VI The expansion of a determinant by *alien cofactors* (the cofactors of a "wrong" row or column) always yields a value of zero.

Example 1 If we expand the determinant $\begin{vmatrix} 4 & 1 & 2 \\ 5 & 2 & 1 \\ 1 & 0 & 3 \end{vmatrix}$ by using its *first*-row elements but the cofactors of the *second*-row elements

$$|C_{21}| = -\begin{vmatrix} 1 & 2 \\ 0 & 3 \end{vmatrix} = -3 \qquad |C_{22}| = \begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix} = 10 \qquad |C_{23}| = -\begin{vmatrix} 4 & 1 \\ 1 & 0 \end{vmatrix} = 1$$

we get $a_{11}|C_{21}| + a_{12}|C_{22}| + a_{13}|C_{23}| = 4(-3) + 1(10) + 2(1) = 0$.

More generally, applying the same type of expansion by alien cofactors as described in Example 1 to the determinant $|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ will yield a zero sum of products as follows:

$$(5.9) \qquad \sum_{i=1}^{3} a_{1j}|C_{2j}| = a_{11}|C_{21}| + a_{12}|C_{22}| + a_{13}|C_{23}|$$
$$= -a_{11}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{13}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$
$$= -a_{11}a_{12}a_{33} + a_{11}a_{13}a_{32} + a_{11}a_{12}a_{33} - a_{12}a_{13}a_{31}$$
$$-a_{11}a_{13}a_{32} + a_{12}a_{13}a_{31} = 0$$

The reason for this outcome lies in the fact that the sum of products in (5.9) can be considered as the result of the *regular* expansion by the second row of another determinant $|A^*| \equiv \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$, which differs from $|A|$ only in its second row and whose first two rows are identical. As an exercise, write out the cofactors of the second rows of $|A^*|$ and verify that these are precisely the cofactors which appeared in (5.9)—and with the correct signs. Since $|A^*| = 0$, because of its two identical rows, the expansion by alien cofactors shown in (5.9) will of necessity yield a value of zero also.

Property VI is valid for determinants of all orders and applies when a determinant is expanded by the alien cofactors of any row or any column. Thus we may state, in general, that for a determinant of order n the following holds:

$$(5.10) \qquad \begin{aligned} &\sum_{j=1}^{n} a_{ij}|C_{i'j}| = 0 \qquad (i \neq i') \qquad &&\text{[expansion by } i\text{th row and cofactors of } i'\text{th row]} \\ &\sum_{i=1}^{n} a_{ij}|C_{ij'}| = 0 \qquad (j \neq j') \qquad &&\text{[expansion by } j\text{th column and cofactors of } j'\text{th column]} \end{aligned}$$

Carefully compare (5.10) with (5.8). In the latter (regular Laplace expansion), the subscripts of a_{ij} and of $|C_{ij}|$ must be identical in each product term in the sum. In the expansion by alien cofactors, such as in (5.10), on the other hand, one of the two subscripts (a chosen value of i' or j') is inevitably "out of place."

Matrix Inversion

Property VI, as summarized in (5.10), is of direct help in developing a method of matrix inversion, i.e., of finding the inverse of a matrix.

Assume that an $n \times n$ nonsingular matrix A is given:

$$(5.11) \qquad \underset{(n\times n)}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \qquad (|A| \neq 0)$$

Since each element of A has a cofactor $|C_{ij}|$, it is possible to form a matrix of cofactors by replacing each element a_{ij} in (5.11) with its cofactor $|C_{ij}|$. Such a cofactor matrix, denoted by $C = [|C_{ij}|]$, must also be $n \times n$. For our present purposes, however, the transpose of C is of more interest. This transpose C' is commonly referred to as the *adjoint* of A and is symbolized by adj A. Written out, the adjoint takes the form

$$(5.12) \qquad \underset{(n\times n)}{C'} \equiv \text{adj } A \equiv \begin{bmatrix} |C_{11}| & |C_{21}| & \cdots & |C_{n1}| \\ |C_{12}| & |C_{22}| & \cdots & |C_{n2}| \\ \cdots & \cdots & \cdots & \cdots \\ |C_{1n}| & |C_{2n}| & \cdots & |C_{nn}| \end{bmatrix}$$

The matrices A and C' are conformable for multiplication, and their product AC' is another $n \times n$ matrix in which each element is a sum of products. By utilizing the formula for Laplace expansion as well as Property VI of determinants, the product AC' may be expressed as follows:

$$\underset{(n\times n)}{AC'} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}|C_{1j}| & \sum_{j=1}^{n} a_{1j}|C_{2j}| & \cdots & \sum_{j=1}^{n} a_{1j}|C_{nj}| \\ \sum_{j=1}^{n} a_{2j}|C_{1j}| & \sum_{j=1}^{n} a_{2j}|C_{2j}| & \cdots & \sum_{j=1}^{n} a_{2j}|C_{nj}| \\ \vdots & \vdots & & \vdots \\ \sum_{j=1}^{n} a_{nj}|C_{1j}| & \sum_{j=1}^{n} a_{nj}|C_{2j}| & \cdots & \sum_{j=1}^{n} a_{nj}|C_{nj}| \end{bmatrix}$$

$$= \begin{bmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |A| \end{bmatrix} \qquad [\text{by (5.8) and (5.10)}]$$

$$= |A| \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = |A| I_n \qquad [\text{factoring}]$$

As the determinant $|A|$ is a nonzero scalar, it is permissible to divide both sides of the equation $AC' = |A|I$ by $|A|$. The result is

$$\frac{AC'}{|A|} = I \qquad \text{or} \qquad A\frac{C'}{|A|} = I$$

Premultiplying both sides of the last equation by A^{-1}, and using the result that $A^{-1}A = I$, we can get $\frac{C'}{|A|} = A^{-1}$, or

$$A^{-1} = \frac{1}{|A|} \operatorname{adj} A \qquad \text{[by (5.12)]} \tag{5.13}$$

Now, we have found a way to invert the matrix A!

The general procedure for finding the inverse of a square matrix A thus involves the following steps: (1) find $|A|$ [we need to proceed with the subsequent steps if and only if $|A| \neq 0$, for if $|A| = 0$, the inverse in (5.13) will be undefined]; (2) find the cofactors of all the elements of A, and arrange them as a cofactor matrix $C = [|C_{ij}|]$; (3) take the transpose of C to get adj A; and (4) divide adj A by the determinant $|A|$. The result will be the desired inverse A^{-1}.

Example 2 Find the inverse of $A = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$. Since $|A| = -2 \neq 0$, the inverse A^{-1} exists. The cofactor of each element is in this case a 1×1 determinant, which is simply defined as the scalar element of that determinant itself (that is, $|a_{ij}| \equiv a_{ij}$). Thus, we have

$$C = \begin{bmatrix} |C_{11}| & |C_{12}| \\ |C_{21}| & |C_{22}| \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -2 & 3 \end{bmatrix}$$

Observe the minus signs attached to 1 and 2, as required for cofactors. Transposing the cofactor matrix yields

$$\operatorname{adj} A = \begin{bmatrix} 0 & -2 \\ -1 & 3 \end{bmatrix}$$

so the inverse A^{-1} can be written as

$$A^{-1} = \frac{1}{|A|} \operatorname{adj} A = -\frac{1}{2}\begin{bmatrix} 0 & -2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{3}{2} \end{bmatrix}$$

Example 3 Find the inverse of $B = \begin{bmatrix} 4 & 1 & -1 \\ 0 & 3 & 2 \\ 3 & 0 & 7 \end{bmatrix}$. Since $|B| = 99 \neq 0$, the inverse B^{-1} also exists. The cofactor matrix is

$$\begin{bmatrix} \begin{vmatrix} 3 & 2 \\ 0 & 7 \end{vmatrix} & -\begin{vmatrix} 0 & 2 \\ 3 & 7 \end{vmatrix} & \begin{vmatrix} 0 & 3 \\ 3 & 0 \end{vmatrix} \\ -\begin{vmatrix} 1 & -1 \\ 0 & 7 \end{vmatrix} & \begin{vmatrix} 4 & -1 \\ 3 & 7 \end{vmatrix} & -\begin{vmatrix} 4 & 1 \\ 3 & 0 \end{vmatrix} \\ \begin{vmatrix} 1 & -1 \\ 3 & 2 \end{vmatrix} & -\begin{vmatrix} 4 & -1 \\ 0 & 2 \end{vmatrix} & \begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 21 & 6 & -9 \\ -7 & 31 & 3 \\ 5 & -8 & 12 \end{bmatrix}$$

Therefore,

$$\text{adj } B = \begin{bmatrix} 21 & -7 & 6 \\ 6 & 31 & -8 \\ -9 & 3 & 12 \end{bmatrix}$$

and the desired inverse matrix is

$$B^{-1} = \frac{1}{|B|} \text{adj } B = \frac{1}{99} \begin{bmatrix} 21 & -7 & 5 \\ 6 & 31 & -8 \\ -9 & 3 & 12 \end{bmatrix}$$

You can check that the results in the above two examples do satisfy $AA^{-1} = A^{-1}A = I$ and $BB^{-1} = B^{-1}B = I$, respectively.

EXERCISE 5.4

1 Suppose that we expand a fourth-order determinant by its *third column* and the cofactors of the *second-column* elements. How would you write the resulting sum of products in Σ notation? What will be the sum of products in Σ notation if we expand it by the *second row* and the cofactors of the *fourth-row* elements?

2 Find the inverse of each of the following matrices:

$(a)\ A = \begin{bmatrix} 5 & 2 \\ 0 & 1 \end{bmatrix}$ $\quad (c)\ C = \begin{bmatrix} 7 & 7 \\ 3 & -1 \end{bmatrix}$

$(b)\ B = \begin{bmatrix} 1 & 0 \\ 9 & 2 \end{bmatrix}$ $\quad (d)\ D = \begin{bmatrix} 7 & 6 \\ 0 & 3 \end{bmatrix}$

3 (*a*) Drawing on your answers to the preceding problem, formulate a two-step rule for finding the adjoint of a given 2×2 matrix A: In the first step, indicate what should be done to the two diagonal elements of A in order to get the diagonal elements of adj A; in the second step, indicate what should be done to the two off-diagonal elements of A. [*Warning:* This rule applies only to 2×2 matrices.]

(*b*) Add a third step which, in conjunction with the previous two steps, yields the 2×2 inverse matrix A^{-1}.

4 Find the inverse of each of the following matrices:

$(a)\ E = \begin{bmatrix} 4 & -2 & 1 \\ 7 & 3 & 3 \\ 2 & 0 & 1 \end{bmatrix}$ $\quad (c)\ G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

$(b)\ F = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 0 & 3 \\ 4 & 0 & 2 \end{bmatrix}$ $\quad (d)\ H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

5 Is it possible for a matrix to be its own inverse?

5.5 CRAMER'S RULE

The method of matrix inversion just discussed enables us to derive a convenient, practical way of solving a linear-equation system, known as *Cramer's rule*.

Derivation of the Rule

Given an equation system $Ax = d$, where A is $n \times n$, the solution can be written as

$$\bar{x} = A^{-1}d = \frac{1}{|A|}(\text{adj}\, A)d \qquad [\text{by } (5.13)]$$

provided A is nonsingular. According to (5.12), this means that

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} |C_{11}| & |C_{21}| & \cdots & |C_{n1}| \\ |C_{12}| & |C_{22}| & \cdots & |C_{n2}| \\ \cdots & \cdots & \cdots & \cdots \\ |C_{1n}| & |C_{2n}| & \cdots & |C_{nn}| \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

$$= \frac{1}{|A|}\begin{bmatrix} d_1|C_{11}| + d_2|C_{21}| + \cdots + d_n|C_{n1}| \\ d_1|C_{12}| + d_2|C_{22}| + \cdots + d_n|C_{n2}| \\ \cdots \\ d_1|C_{1n}| + d_2|C_{2n}| + \cdots + d_n|C_{nn}| \end{bmatrix}$$

$$= \frac{1}{|A|}\begin{bmatrix} \sum_{i=1}^{n} d_i|C_{i1}| \\ \sum_{i=1}^{n} d_i|C_{i2}| \\ \vdots \\ \sum_{i=1}^{n} d_i|C_{in}| \end{bmatrix}$$

Equating the corresponding elements on the two sides of the equation, we obtain the solution values

$$(5.14) \qquad \bar{x}_1 = \frac{1}{|A|}\sum_{i=1}^{n} d_i|C_{i1}| \qquad \bar{x}_2 = \frac{1}{|A|}\sum_{i=1}^{n} d_i|C_{i2}| \qquad \text{(etc.)}$$

The Σ terms in (5.14) look unfamiliar. What do they mean? From (5.8), we see that the Laplace expansion of a determinant $|A|$ by its first column can be expressed in the form $\sum_{i=1}^{n} a_{i1}|C_{i1}|$. If we replace the first column of $|A|$ by the column vector d but keep all the other columns intact, then a new determinant will result, which we can call $|A_1|$—the subscript 1 indicating that the first column has been replaced by d. The expansion of $|A_1|$ by its first column (the d column) will yield the expression $\sum_{i=1}^{n} d_i|C_{i1}|$, because the elements d_i now take the place of the elements a_{i1}. Returning to (5.14), we see therefore that

$$\bar{x}_1 = \frac{1}{|A|}|A_1|$$

Similarly, if we replace the second column of $|A|$ by the column vector d, while retaining all the other columns, the expansion of the new determinant $|A_2|$ by its second column (the d column) will result in the expression $\sum_{i=1}^{n} d_i|C_{i2}|$. When divided by $|A|$, this latter sum will give us the solution value $\bar{x}_2$; and so on.

This procedure can now be generalized. To find the solution value of the jth variable $\bar{x}_j$, we can merely replace the jth column of the determinant $|A|$ by the constant terms $d_1 \cdots d_n$ to get a new determinant $|A_j|$ and then divide $|A_j|$ by the original determinant $|A|$. Thus, the solution of the system $Ax = d$ can be expressed as

$$\text{(5.15)} \qquad \bar{x}_j = \frac{|A_j|}{|A|} = \frac{1}{|A|}\begin{vmatrix} a_{11} & a_{12} & \cdots & d_1 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & d_2 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & d_n & \cdots & a_{nn} \end{vmatrix}$$

(jth column replaced by d)

The result in (5.15) is the statement of Cramer's rule. Note that, whereas the matrix inversion method yields the solution values of *all* the endogenous variables at once ($\bar{x}$ is a vector), Cramer's rule can give us the solution value of only a single endogenous variable at a time ($\bar{x}_j$ is a scalar).

Example 1 Find the solution of the equation system

$$5x_1 + 3x_2 = 30$$
$$6x_1 - 2x_2 = 8$$

The coefficients and the constant terms give the following determinants:

$$|A| = \begin{vmatrix} 5 & 3 \\ 6 & -2 \end{vmatrix} = -28 \qquad |A_1| = \begin{vmatrix} 30 & 3 \\ 8 & -2 \end{vmatrix} = -84$$

$$|A_2| = \begin{vmatrix} 5 & 30 \\ 6 & 8 \end{vmatrix} = -140$$

Therefore, by virtue of (5.15), we can immediately write

$$\bar{x}_1 = \frac{|A_1|}{|A|} = \frac{-84}{-28} = 3 \qquad \text{and} \qquad \bar{x}_2 = \frac{|A_2|}{|A|} = \frac{-140}{-28} = 5$$

Example 2 Find the solution of the equation system

$$7x_1 - x_2 - x_3 = 0$$
$$10x_1 - 2x_2 + x_3 = 8$$
$$6x_1 + 3x_2 - 2x_3 = 7$$

The relevant determinants $|A|$ and $|A_j|$ are found to be

$$|A| = \begin{vmatrix} 7 & -1 & -1 \\ 10 & -2 & 1 \\ 6 & 3 & -2 \end{vmatrix} = -61 \qquad |A_1| = \begin{vmatrix} 0 & -1 & -1 \\ 8 & -2 & 1 \\ 7 & 3 & -2 \end{vmatrix} = -61$$

$$|A_2| = \begin{vmatrix} 7 & 0 & -1 \\ 10 & 8 & 1 \\ 6 & 7 & -2 \end{vmatrix} = -183 \qquad |A_3| = \begin{vmatrix} 7 & -1 & 0 \\ 10 & -2 & 8 \\ 6 & 3 & 7 \end{vmatrix} = -244$$

thus the solution values of the variables are

$$\bar{x}_1 = \frac{|A_1|}{|A|} = \frac{-61}{-61} = 1 \qquad \bar{x}_2 = \frac{|A_2|}{|A|} = \frac{-183}{-61} = 3 \qquad \bar{x}_3 = \frac{|A_3|}{|A|} = \frac{-244}{-61} = 4$$

Notice that in each of these examples we find $|A| \neq 0$. This is a necessary condition for the application of Cramer's rule, as it is for the existence of the inverse A^{-1}. Cramer's rule is, after all, based upon the concept of the inverse matrix, even though in practice it bypasses the process of matrix inversion.

Note on Homogeneous-Equation Systems

The equation systems $Ax = d$ considered above can have any constants in the vector d. If $d = 0$, that is, if $d_1 = d_2 = \cdots = d_n = 0$, however, the equation system will become

$$Ax = 0$$

where 0 is a zero vector. This special case is referred to as a *homogeneous-equation system.**

If the matrix A is nonsingular, a homogeneous-equation system can yield only a "trivial solution," namely, $\bar{x}_1 = \bar{x}_2 = \cdots = \bar{x}_n = 0$. This follows from the fact that the solution $\bar{x} = A^{-1}d$ will in this case become

$$\underset{(n\times 1)}{\bar{x}} = \underset{(n\times n)}{A^{-1}} \underset{(n\times 1)}{0} = \underset{(n\times 1)}{0}$$

Alternatively, this outcome can be derived from Cramer's rule. The fact that $d = 0$ implies that $|A_j|$, for all j, must contain a whole column of zeros, and thus the solution will turn out to be

$$\bar{x}_j = \frac{|A_j|}{|A|} = \frac{0}{|A|} = 0 \qquad (j = 1, 2, \ldots, n)$$

Curiously enough, the *only* way to get a *non*trivial solution from a homogeneous-equation system is to have $|A| = 0$, that is, to have a *singular* coefficient

* The word "homogeneous" describes the property that when all the variables $x_1, \ldots, x_n$ are multiplied by the same number, the equation system will remain valid. This is possible only if the constant terms (those unattached to any x_i) are all zero.

matrix A! In that event, we have

$$\bar{x}_j = \frac{|A_j|}{|A|} = \frac{0}{0}$$

where the $0/0$ expression is not equal to zero but is, rather, something undefined. Consequently, Cramer's rule is not applicable. This does not mean that we cannot obtain solutions; it means only that we cannot get a unique solution.

Consider the homogeneous-equation system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned} \tag{5.16}$$

It is self-evident that $\bar{x}_1 = \bar{x}_2 = 0$ is a solution, but that solution is trivial. Now, assume that the coefficient matrix A is singular, so that $|A| = 0$. This implies that the row vector $[a_{11} \quad a_{12}]$ is a multiple of the row vector $[a_{21} \quad a_{22}]$; consequently, one of the two equations is redundant. By deleting, say, the second equation from (5.16), we end up with one (the first) equation in two variables, the solution of which is $\bar{x}_1 = (-a_{12}/a_{11})\bar{x}_2$. This solution is nontrivial and well defined if $a_{11} \neq 0$, but it really represents an infinite number of solutions because, for every possible value of $\bar{x}_2$, there is a corresponding value $\bar{x}_1$ such that the pair constitutes a solution. Thus no unique nontrivial solution exists for this homogeneous-equation system. This last statement is also generally valid for the n-variable case.

Solution Outcomes for a Linear-Equation System

Our discussion of the several variants of the linear-equation system $Ax = d$ reveals that as many as four different types of solution outcome are possible. For a better overall view of these variants, we list them in tabular form in Table 5.1.

Table 5.1 Solution outcomes for a linear-equation system $Ax = d$

Determinant $\|A\|$ \ Vector d		$d \neq 0$ (nonhomogeneous system)	$d = 0$ (homogeneous system)
$\|A\| \neq 0$ (matrix A nonsingular)		There exists a unique, nontrivial solution $\bar{x} \neq 0$	There exists a unique, trivial solution $\bar{x} = 0$
$\|A\| = 0$ (matrix A singular)	Equations dependent	There exist an infinite number of solutions (not including the trivial one)	There exist an infinite number of solutions (including the trivial one)
	Equations inconsistent	No solution exists	[Not applicable]

As a first possibility, the system may yield a unique, nontrivial solution. This type of outcome can arise only when we have a nonhomogeneous system with a nonsingular coefficient matrix A. The second possible outcome is a unique, trivial solution, and this is associated with a homogeneous system with a nonsingular matrix A. As a third possibility, we may have an infinite number of solutions. This eventuality is linked exclusively to a system in which the equations are dependent (i.e., in which there are redundant equations). Depending on whether the system is homogeneous, the trivial solution may or may not be included in the set of infinite number of solutions. Finally, in the case of an inconsistent equation system, there exists no solution at all. From the point of view of a model builder, the most useful and desirable outcome is, of course, that of a unique, nontrivial solution $\bar{x} \neq 0$.

EXERCISE 5.5

1 Use Cramer's rule to solve the following equation systems:

(a) $\begin{aligned} 3x_1 - 2x_2 &= 11 \\ 2x_1 + x_2 &= 12 \end{aligned}$ (c) $\begin{aligned} 8x_1 - 7x_2 &= -6 \\ x_1 + x_2 &= 3 \end{aligned}$

(b) $\begin{aligned} -x_1 + 3x_2 &= -3 \\ 4x_1 - x_2 &= 12 \end{aligned}$ (d) $\begin{aligned} 6x_1 + 9x_2 &= 15 \\ 7x_1 - 3x_2 &= 4 \end{aligned}$

2 For each of the equation systems in the preceding problem, find the inverse of the coefficient matrix, and get the solution by the formula $\bar{x} = A^{-1}d$.

3 Use Cramer's rule to solve the following equation systems:

(a) $\begin{aligned} 8x_1 - x_2 &= 15 \\ x_2 + 5x_3 &= 1 \\ 2x_1 + 3x_3 &= 4 \end{aligned}$ (c) $\begin{aligned} 4x + 3y - 2z &= 7 \\ x + y &= 5 \\ 3x + z &= 4 \end{aligned}$

(b) $\begin{aligned} -x_1 + 3x_2 + 2x_3 &= 24 \\ x_1 + x_3 &= 6 \\ 5x_2 - x_3 &= 8 \end{aligned}$ (d) $\begin{aligned} -x + y + z &= a \\ x - y + z &= b \\ x + y - z &= c \end{aligned}$

4 Show that Cramer's rule can be derived alternatively by the following procedure. Multiply both sides of the first equation in the system $Ax = d$ by the cofactor $|C_{1j}|$, and then multiply both sides of the second equation by the cofactor $|C_{2j}|$, etc. Add all the newly obtained equations. Then assign the values $1, 2, \ldots, n$ to the index j, successively, to get the solution values $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ as shown in (5.14).

5.6 APPLICATION TO MARKET AND NATIONAL-INCOME MODELS

Simple equilibrium models such as those discussed in Chap. 3 can be solved with ease by Cramer's rule or by matrix inversion.

Market Model

The two-commodity model described in (3.12) can be written (after eliminating the quantity variables) as a system of two linear equations, as in (3.13′):

$$c_1P_1 + c_2P_2 = -c_0$$
$$\gamma_1P_1 + \gamma_2P_2 = -\gamma_0$$

The three determinants needed—$|A|$, $|A_1|$, and $|A_2|$—have the following values:

$$|A| = \begin{vmatrix} c_1 & c_2 \\ \gamma_1 & \gamma_2 \end{vmatrix} = c_1\gamma_2 - c_2\gamma_1$$

$$|A_1| = \begin{vmatrix} -c_0 & c_2 \\ -\gamma_0 & \gamma_2 \end{vmatrix} = -c_0\gamma_2 + c_2\gamma_0$$

$$|A_2| = \begin{vmatrix} c_1 & -c_0 \\ \gamma_1 & -\gamma_0 \end{vmatrix} = -c_1\gamma_0 + c_0\gamma_1$$

Therefore the equilibrium prices must be

$$\bar{P}_1 = \frac{|A_1|}{|A|} = \frac{c_2\gamma_0 - c_0\gamma_2}{c_1\gamma_2 - c_2\gamma_1} \qquad \bar{P}_2 = \frac{|A_2|}{|A|} = \frac{c_0\gamma_1 - c_1\gamma_0}{c_1\gamma_2 - c_2\gamma_1}$$

which are precisely those obtained in (3.14) and (3.15). The equilibrium quantities can be found, as before, by setting $P_1 = \bar{P}_1$ and $P_2 = \bar{P}_2$ in the demand or supply functions.

National-Income Model

The simple national-income model cited in (3.23) can also be solved by the use of Cramer's rule. As written in (3.23), the model consists of the following two simultaneous equations:

$$Y = C + I_0 + G_0$$
$$C = a + bY \qquad (a > 0, \quad 0 < b < 1)$$

These can be rearranged into the form

$$Y - C = I_0 + G_0$$
$$-bY + C = a$$

so that the endogenous variables Y and C appear only on the left of the equals signs, whereas the exogenous variables and the unattached parameter appear only on the right. The coefficient matrix now takes the form $\begin{bmatrix} 1 & -1 \\ -b & 1 \end{bmatrix}$, and the column vector of constants (data), $\begin{bmatrix} I_0 + G_0 \\ a \end{bmatrix}$. Note that the sum $I_0 + G_0$ is considered as a single entity, i.e., a single element in the constant vector.

Cramer's rule now leads immediately to the following solution:

$$\bar{Y} = \frac{\begin{vmatrix} (I_0 + G_0) & -1 \\ a & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -1 \\ -b & 1 \end{vmatrix}} = \frac{I_0 + G_0 + a}{1 - b}$$

$$\bar{C} = \frac{\begin{vmatrix} 1 & (I_0 + G_0) \\ -b & a \end{vmatrix}}{\begin{vmatrix} 1 & -1 \\ -b & 1 \end{vmatrix}} = \frac{a + b(I_0 + G_0)}{1 - b}$$

You should check that the solution values just obtained are identical with those shown in (3.24) and (3.25).

Let us now try to solve this model by inverting the coefficient matrix. Since the coefficient matrix is $A = \begin{bmatrix} 1 & -1 \\ -b & 1 \end{bmatrix}$, its cofactor matrix will be $\begin{bmatrix} 1 & b \\ 1 & 1 \end{bmatrix}$, and we therefore have adj $A = \begin{bmatrix} 1 & 1 \\ b & 1 \end{bmatrix}$. It follows that the inverse matrix is

$$A^{-1} = \frac{1}{|A|} \text{adj } A = \frac{1}{1 - b}\begin{bmatrix} 1 & 1 \\ b & 1 \end{bmatrix}$$

We know that, for the equation system $Ax = d$, the solution is expressible as $\bar{x} = A^{-1}d$. Applied to the present model, this means that

$$\begin{bmatrix} \bar{Y} \\ \bar{C} \end{bmatrix} = \frac{1}{1 - b}\begin{bmatrix} 1 & 1 \\ b & 1 \end{bmatrix}\begin{bmatrix} I_0 + G_0 \\ a \end{bmatrix} = \frac{1}{1 - b}\begin{bmatrix} I_0 + G_0 + a \\ b(I_0 + G_0) + a \end{bmatrix}$$

It is easy to see that this is again the same solution as obtained before.

Matrix Algebra versus Elimination of Variables

The two economic models used for illustration here both involve two equations only, and thus only second-order determinants need to be evaluated. For large equation systems, higher-order determinants will appear, and their evaluation may prove to be no simple task. Nor is the inversion of large matrices exactly child's play. From the computational point of view, in fact, matrix inversion and Cramer's rule are not necessarily more efficient than the method of successive elimination of variables.

If so, one may ask, why use the matrix methods at all? As we have seen from the preceding pages, matrix algebra has given us a compact notation for any linear-equation system, and also furnishes a determinantal criterion for testing the existence of a unique solution. These are advantages not otherwise available. In addition to these, it may be mentioned that, unlike the elimination-of-variable method, which affords no means of analytically expressing the solution, the matrix-inversion method and Cramer's rule do provide the handy solution expressions $\bar{x} = A^{-1}d$ and $\bar{x}_j = |A_j|/|A|$. Such analytical expressions of the solution are useful not only because they are in themselves a summary statement of the actual solution procedure, but also because they make possible the performance of further mathematical operations on the solution as written, if called for.

Under certain circumstances, matrix methods can even claim a computational advantage, such as when the task is to solve at the same time several equation systems having an identical coefficient matrix A but different constant-term vectors. In such cases, the elimination-of-variable method would require that the computational procedure be repeated each time a new equation system is considered. With the matrix-inversion method, however, we are required to find the common inverse matrix A^{-1} *only once*; then the same inverse can be used to premultiply all the constant-term vectors pertaining to the various equation systems involved, in order to obtain their respective solutions. This particular computational advantage will take on great practical significance when we consider the solution of the Leontief input-output models in the next section.

EXERCISE 5.6

1 Solve the national-income model in Exercise 3.5-1:
(*a*) by matrix inversion (*b*) by Cramer's rule
(List the variables in the order Y, C, T.)

2 Solve the national-income model in Exercise 3.5-2:
(*a*) by matrix inversion (*b*) by Cramer's rule
(List the variables in the order Y, C, G.)

5.7 LEONTIEF INPUT-OUTPUT MODELS

In its "static" version, Professor Leontief's input-output analysis* deals with this particular question: "What level of output should each of the n industries in an economy produce, in order that it will just be sufficient to satisfy the total demand for that product?"

The rationale for the term *input-output analysis* is quite plain to see. The output of any industry (say, the steel industry) is needed as an input in many other industries, or even for that industry itself; therefore the "correct" (i.e., shortage-free as well as surplus-free) level of steel output will depend on the input requirements of all the n industries. In turn, the output of many other industries will enter into the steel industry as inputs, and consequently the "correct" levels of the other products will in turn depend partly upon the input requirements of the steel industry. In view of this interindustry dependence, any set of "correct"

* Wassily W. Leontief, *The Structure of American Economy 1919–1939*, 2d ed., Oxford University Press, Fair Lawn, N.J., 1951.

output levels for the n industries must be one that is consistent with all the input requirements in the economy, so that no bottlenecks will arise anywhere. In this light, it is clear that input-output analysis should be of great use in production planning, such as in planning for the economic development of a country or for a program of national defense.

Strictly speaking, input-output analysis is not a form of the general equilibrium analysis as discussed in Chap. 3. Although the interdependence of the various industries is emphasized, the "correct" output levels envisaged are those which satisfy technical input-output relationships rather than market equilibrium conditions. Nevertheless, the problem posed in input-output analysis also boils down to one of solving a system of simultaneous equations, and matrix algebra can again be of service.

Structure of an Input-Output Model

Since an input-output model normally encompasses a large number of industries, its framework is of necessity rather involved. To simplify the problem, the following assumptions are as a rule adopted: (1) each industry produces only one homogeneous commodity (broadly interpreted, this does permit the case of two or more jointly produced commodities, provided they are produced in a fixed proportion to one another); (2) each industry uses a fixed input ratio (or factor combination) for the production of its output; and (3) production in every industry is subject to constant returns to scale, so that a k-fold change in every input will result in an exactly k-fold change in the output. These assumptions are, of course, unrealistic. A saving grace is that, if an industry produces two different commodities or uses two different possible factor combinations, then that industry may—at least conceptually—be broken down into two separate industries.

From these assumptions we see that, in order to produce each unit of the jth commodity, the input need for the ith commodity must be a fixed amount, which we shall denote by a_{ij}. Specifically, the production of each unit of the jth commodity will require a_{1j} (amount) of the first commodity, a_{2j} of the second commodity,..., and a_{nj} of the nth commodity. (The order of the subscripts in a_{ij} is easy to remember: the first subscript refers to the input, and the second to the output, so that a_{ij} indicates how much of the ith commodity is used for the production of each unit of the jth commodity.) For our purposes, we may assume prices to be given and, thus, adopt "a dollar's worth" of each commodity as its unit. Then the statement $a_{32} = 0.35$ will mean that 35 cents' worth of the third commodity is required as an input for producing a dollar's worth of the second commodity. The a_{ij} symbol will be referred to as an *input coefficient*.

For an n-industry economy, the input coefficients can be arranged into a matrix $A = [a_{ij}]$, as in Table 5.2, in which each *column* specifies the input requirements for the production of one unit of the output of a particular industry. The second column, for example, states that to produce a unit (a dollar's worth) of commodity II, the inputs needed are: a_{12} units of commodity I, a_{22} units of

Table 5.2 Input-coefficient matrix

Input	Output				
	I	II	III	$\cdots$	N
I	a_{11}	a_{12}	a_{13}	$\cdots$	a_{1n}
II	a_{21}	a_{22}	a_{23}	$\cdots$	a_{2n}
III	a_{31}	a_{32}	a_{33}	$\cdots$	a_{3n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
N	a_{n1}	a_{n2}	a_{n3}	$\cdots$	a_{nn}

commodity II, etc. If no industry uses its own product as an input, then the elements in the principal diagonal of matrix A will all be zero.

The Open Model

If, besides the n industries, the model also contains an "open" sector (say, households) which exogenously determines a *final demand* (noninput demand) for the product of each industry and which supplies a *primary input* (say, labor service) not produced by the n industries themselves, the model is an *open model.*

In view of the presence of the open sector, the sum of the elements in each column of the input-coefficient matrix A (or *input matrix* A, for short) must be less than 1. Each column sum represents the *partial* input cost (not including the cost of the primary input) incurred in producing a dollar's worth of some commodity; if this sum is greater than or equal to \$1, therefore, production will not be economically justifiable. Symbolically, this fact may be stated thus:

$$\sum_{i=1}^{n} a_{ij} < 1 \qquad (j = 1, 2, \ldots, n)$$

where the summation is over i, that is, over the elements appearing in the various *rows* of a specific column j. Carrying this line of thought a step further, it may also be stated that, since the value of output (\$1) must be fully absorbed by the payments to all factors of production, the amount by which the column sum falls short of \$1 must represent the payment to the primary input of the open sector. Thus the value of the primary input needed in producing a unit of the jth commodity should be $1 - \sum_{i=1}^{n} a_{ij}$.

If industry I is to produce an output just sufficient to meet the input requirements of the n industries as well as the final demand of the open sector, its output level x_1 must satisfy the following equation:

$$x_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + d_1$$

or $$(1 - a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n = d_1$$

where d_1 denotes the final demand for its output and $a_{1j}x_j$ represents the input demand from the jth industry.* Note that, aside from the first coefficient, $(1 - a_{11})$, the remaining coefficients in the last equation are transplanted directly from the first row of Table 5.2, except that they are now all prefixed with minus signs. Similarly, the corresponding equation for industry II will have the same coefficients as in the second row of Table 5.1 (again with minus signs added), except that the variable x_2 will have the coefficient $(1 - a_{22})$ instead of $-a_{22}$. For the entire set of n industries, the "correct" output levels can therefore be summarized by the following system of n linear equations:

$$
\begin{aligned}
(1 - a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n &= d_1 \\
-a_{21}x_1 + (1 - a_{22})x_2 - \cdots - a_{2n}x_n &= d_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
-a_{n1}x_1 - a_{n2}x_2 - \cdots + (1 - a_{nn})x_n &= d_n
\end{aligned}
\tag{5.17}
$$

In matrix notation, this may be written as

$$
\begin{bmatrix}
(1 - a_{11}) & -a_{12} & \cdots & -a_{1n} \\
-a_{21} & (1 - a_{22}) & \cdots & -a_{2n} \\
\vdots & \vdots & & \vdots \\
-a_{n1} & -a_{n2} & \cdots & (1 - a_{nn})
\end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
=
\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}
\tag{5.17'}
$$

If the 1s in the principal diagonal of the matrix on the left are ignored, the matrix is simply $-A = [-a_{ij}]$. As it is, on the other hand, the matrix is the *sum* of the identity matrix I_n (with 1s in its principal diagonal and with 0s everywhere else) and the matrix $-A$. Thus (5.17′) can also be written as

$$(I - A)x = d \tag{5.17''}$$

where x and d are, respectively, the variable vector and the final-demand (constant-term) vector. The matrix $(I - A)$ is called the *technology matrix*, and we may denote it by T. Thus the system can also be written as

$$Tx = d \tag{5.17'''}$$

As long as T is nonsingular—and there is no a priori reason why it should not be—we shall be able to find its inverse T^{-1}, and obtain the unique solution of the system from the equation

$$\bar{x} = T^{-1}d = (I - A)^{-1}d \tag{5.18}$$

* Do not ever add up the input coefficients across a row; such a sum—say, $a_{11} + a_{12} + \cdots + a_{1n}$—is devoid of economic meaning. The sum of the products $a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n$, on the other hand, does have an economic meaning; it represents the total amount of x_1 needed as input for all the n industries.

A Numerical Example

For purposes of illustration, suppose that there are only three industries in the economy and that the input-coefficient matrix is as follows (let us use decimal values this time):

$$(5.19)\qquad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 0.2 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.2 \\ 0.1 & 0.3 & 0.2 \end{bmatrix}$$

Note that in A each column sum is less than 1, as it should be. Further, if we denote by a_{0j} the dollar amount of the primary input used in producing a dollar's worth of the jth commodity, we can write [by subtracting each column sum in (5.19) from 1]:

$$(5.20)\qquad a_{01} = 0.3 \qquad a_{02} = 0.3 \qquad \text{and} \qquad a_{03} = 0.4$$

With the matrix A above, the open input-output system can be expressed in the form $Tx = (I - A)x = d$ as follows:

$$(5.21)\qquad \begin{bmatrix} 0.8 & -0.3 & -0.2 \\ -0.4 & 0.9 & -0.2 \\ -0.1 & -0.3 & 0.8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

Here we have deliberately not given specific values to the final demands d_1, d_2, and d_3. In this way, by keeping the vector d in parametric form, our solution will appear as a "formula" into which we can feed various specific d vectors to obtain various corresponding specific solutions.

By inverting the 3×3 technology matrix T, the solution of (5.21) can be found, approximately (because of rounding of decimal figures), to be:

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = T^{-1}d = \frac{1}{0.384} \begin{bmatrix} 0.66 & 0.30 & 0.24 \\ 0.34 & 0.62 & 0.24 \\ 0.21 & 0.27 & 0.60 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

If the specific final-demand vector (say, the final-output target of a development program) happens to be $d = \begin{bmatrix} 10 \\ 5 \\ 6 \end{bmatrix}$, in billions of dollars, then the following specific solution values will emerge (again in billions of dollars):

$$\bar{x}_1 = \frac{1}{0.384}[0.66(10) + 0.30(5) + 0.24(6)] = \frac{9.54}{0.384} = 24.84$$

and similarly,

$$\bar{x}_2 = \frac{7.94}{0.384} = 20.68 \qquad \text{and} \qquad \bar{x}_3 = \frac{7.05}{0.384} = 18.36$$

An important question now arises. The production of the output mix $\bar{x}_1$, $\bar{x}_2$, and $\bar{x}_3$ must entail a definite required amount of the primary input. Would the amount *required* be consistent with what is *available* in the economy? On the basis of (5.20), the required primary input may be calculated as follows:

$$\sum_{j=1}^{3} a_{0j}\bar{x}_j = 0.3(24.84) + 0.3(20.68) + 0.4(18.36) = \$21.00 \text{ billion}$$

Therefore, the specific final demand $d = \begin{bmatrix} 10 \\ 5 \\ 6 \end{bmatrix}$ will be feasible if and only if the available amount of the primary input is at least \$21 billion. If the amount available falls short, then that particular production target will, of course, have to be revised downward accordingly.

One important feature of the above analysis is that, as long as the input coefficients remain the same, the inverse $T^{-1} = (I - A)^{-1}$ will not change; therefore only *one* matrix inversion needs to be performed, even if we are to consider a hundred or a thousand different final-demand vectors—such as a spectrum of alternative development targets. This can mean considerable savings in computational effort as compared with the elimination-of-variable method, especially if large equation systems are involved. Note that this advantage is not shared by Cramer's rule. By the latter rule, the solution will be calculated according to the formula $\bar{x}_j = |T_j|/|T|$, but each time a different final-demand vector d is used, we must reevaluate the determinants $|T_j|$. This would be more time-consuming than the multiplication of a known T^{-1} by a new vector d.

Finding the Inverse by Approximation

For large equation systems, the task of inverting a matrix can be exceedingly lengthy and tedious. Even though computers can aid us, simpler computational schemes would still be desirable. For the input-output models under consideration, there does exist a method of finding an approximation to the inverse $T^{-1} = (I - A)^{-1}$ to any desired degree of accuracy; thus it is possible to avoid the process of matrix inversion entirely.

Let us first consider the following matrix multiplication (m = a positive integer):

$$\begin{aligned}
&(I - A)(I + A + A^2 + \cdots + A^m) \\
&\quad = I(I + A + A^2 + \cdots + A^m) - A(I + A + A^2 + \cdots + A^m) \\
&\quad = (I + A + A^2 + \cdots + A^m) - (A + A^2 + \cdots + A^m + A^{m+1}) \\
&\quad = I - A^{m+1}
\end{aligned}$$

Had the result of the multiplication been the identity matrix I alone, we could have taken the matrix sum $(I + A + A^2 + \cdots + A^m)$ as the inverse of $(I - A)$. It is the presence of the $-A^{m+1}$ term that spoils things! Fortunately, though, there remains for us a second-best course, for if the matrix A^{m+1} can be made to approach an $n \times n$ null matrix, then $I - A^{m+1}$ will approach I, and accordingly the said sum matrix $(I + A + A^2 + \cdots + A^m)$ will approach the desired inverse

$(I - A)^{-1}$. By making A^{m+1} approach a null matrix, therefore, we can obtain an *approximation inverse* by adding the matrices $I, A, A^2, \ldots, A^m$.

But can we make A^{m+1} approach a null matrix? And if so, how? The answer to the first question is yes if—as is true of the input-output models under consideration—the elements in *each* column of matrix A are nonnegative numbers adding up to *less than* 1, such as illustrated in (5.19). For such cases, A^{m+1} can be made to approach a null matrix by making the power m sufficiently large, i.e., by a long-enough process of repeated self-multiplication of matrix A. We shall sketch the proof for this statement presently, but if for now its validity is granted, the procedure of computing the approximation inverse becomes very clear: we can simply calculate the successive matrices $A^2, A^3, \ldots,$ until there emerges a matrix A^{m+1} whose elements are, by a preselected standard, all of a negligible order of magnitude ("approaching zero"). When that happens, we can terminate the multiplication process and add up all the matrices already obtained, to form the approximation inverse $(I + A + A^2 + \cdots + A^m)$.*

Note that, when the matrix A is such that A^{m+1} approaches the null matrix as m is increased indefinitely, the approximation inverse $(I + A + A^2 + \cdots + A^m)$ will also have the property that all its elements are nonnegative. The first two terms in the sum, I and A, obviously contain nonnegative elements only. But so do all powers of A, because the self-multiplication of A involves nothing other than the multiplication and addition of the nonnegative elements of A itself. Inasmuch as the final-demand vector d also contains only nonnegative elements, it should be clear from (5.18) that the solution output levels must also be nonnegative. This, of course, is precisely what we wanted them to be.

Let us now sketch the proof for the assertion that, given a nonnegative input-coefficient matrix $A = [a_{ij}]$ whose column sums are each less than 1, the matrix A^{m+1} will approach a null matrix as m is increased indefinitely.† For this purpose, we shall need the concept of the *norm* of a matrix A, which is defined as the *largest column sum* in A and is denoted by $N(A)$. In the matrix of (5.19), for instance, we have $N(A) = 0.7$; this is the first column sum, which happens also to be equal to the second column sum. It is immediately clear that no element in a matrix can ever exceed the value of the norm; that is,

$$a_{ij} \leq N(A) \qquad \text{(for all } i, j\text{)}$$

In the input-output context, we have $N(A) < 1$, and all $a_{ij} < 1$. Actually, the

* The approximation of $(I - A)^{-1}$ by $(I + A + A^2 + \cdots + A^m)$ is analogous to the approximation of the infinite series

$$(1 - r)^{-1} = \frac{1}{1 - r} = 1 + r + r^2 + \cdots \qquad (0 < r < 1)$$

by the sum $(1 + r + r^2 + \cdots + r^n)$. Since the subsequent terms in the series become progressively smaller, we can approximate $(1 - r)^{-1}$ to any desired degree of accuracy by an appropriate choice of the number n.

† For a more detailed discussion, see Frederick V. Waugh, "Inversion of the Leontief Matrix by Power Series," *Econometrica*, April, 1950, pp. 142–154.

matrix A being nonnegative, we must have

$$0 < N(A) < 1$$

Regarding norms of matrices, there is a theorem stating that, given any two (conformable) matrices A and B, the norm of the product matrix AB can never exceed the product of $N(A)$ and $N(B)$:

$$N(AB) \leq N(A)N(B) \tag{5.22}$$

In the special case of $A = B$, where the matrix is square, this result means that

$$N(A^2) \leq [N(A)]^2 \tag{5.23}$$

When $B = A^2$, (5.22) and (5.23) together imply that

$$N(A^3) \leq N(A)N(A^2) \leq N(A)[N(A)]^2 = [N(A)]^3$$

The generalized version of the last result is

$$N(A^m) \leq [N(A)]^m \tag{5.24}$$

It is in this light that the fact $0 < N(A) < 1$ acquires significance, for as m becomes infinite, $[N(A)]^m$ must approach zero if $N(A)$ is a positive fraction. By (5.24), this means that $N(A^m)$ must also approach zero, since $N(A^m)$ is at most as large as $[N(A)]^m$. If so, however, the elements in the matrix A^m must approach zero also when m is increased indefinitely, because no element in the latter matrix can exceed the value of the norm $N(A^m)$. Thus, by making m sufficiently large, the matrix A^{m+1} can be made to approach a null matrix, when the condition $0 < N(A) < 1$ is satisfied.

The Closed Model

If the exogenous sector of the open input-output model is absorbed into the system as just another *industry*, the model will become a *closed model*. In such a model, final demand and primary input do not appear; in their place will be the input requirements and the output of the newly conceived industry. All goods will now be *intermediate* in nature, because everything that is produced is produced only for the sake of satisfying the input requirements of the $(n + 1)$ industries in the model.

At first glance, the conversion of the open sector into an additional industry would not seem to create any significant change in the analysis. Actually, however, since the new industry is assumed to have a fixed input ratio as does any other industry, the supply of what used to be the primary input must now bear a fixed proportion to what used to be called the *final demand*. More concretely, this may mean, for example, that households will consume each commodity in a fixed proportion to the labor service they supply. This certainly constitutes a significant change in the analytical framework involved.

Mathematically, the disappearance of the final demands means that we will now have a homogeneous-equation system. Assuming four industries only (including the new one, designated by the subscript 0), the "correct" output levels will, by analogy to (5.17′), be those which satisfy the equation system:

$$\begin{bmatrix} (1-a_{00}) & -a_{01} & -a_{02} & -a_{03} \\ -a_{10} & (1-a_{11}) & -a_{12} & -a_{13} \\ -a_{20} & -a_{21} & (1-a_{22}) & -a_{23} \\ -a_{30} & -a_{31} & -a_{32} & (1-a_{33}) \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Because this equation system is homogeneous, it can have a nontrivial solution if and only if the 4×4 technology matrix $(I - A)$ has a vanishing determinant. The latter condition is indeed always satisfied: In a closed model, there is no more primary input; hence each column sum in the input-coefficient matrix A must now be exactly equal to (rather than less than) 1; that is, $a_{0j} + a_{1j} + a_{2j} + a_{3j} = 1$, or

$$a_{0j} = 1 - a_{1j} - a_{2j} - a_{3j}$$

But this implies that, in every column of the matrix $(I - A)$ above, the top element is always equal to the negative of the sum of the other three elements. Consequently, the four rows are linearly dependent, and we must find $|I - A| = 0$. This guarantees that the system does possess nontrivial solutions; in fact, as indicated in Table 5.1, it has an infinite number of them. This means that in a closed model, with a homogeneous-equation system, no unique "correct" output mix exists. We can determine the output levels $\bar{x}_1, \ldots, \bar{x}_4$ in proportion to one another, but cannot fix their absolute levels unless additional restrictions are imposed on the model.

EXERCISE 5.7

1 On the basis of the model in (5.21), if the final demands are $d_1 = 30$, $d_2 = 15$, and $d_3 = 10$ (all in billions of dollars), what will be the solution output levels for the three industries? (Round off answers to two decimal places.)

2 Using the information in (5.20), calculate the total amount of primary input required to produce the solution output levels of the preceding problem.

3 In a two-industry economy, it is known that industry I uses 10 cents of its own product and 60 cents of commodity II to produce a dollar's worth of commodity I; industry II uses none of its own product but uses 50 cents of commodity I in producing a dollar's worth of commodity II; and the open sector demands \$1000 billion of commodity I and \$2000 billion of commodity II.

(*a*) Write out the input matrix, the technology matrix, and the specific input-output matrix equation for this economy.

(*b*) Find the solution output levels by Cramer's rule.

4 Given the input matrix and the final-demand vector

$$A = \begin{bmatrix} 0.05 & 0.25 & 0.34 \\ 0.33 & 0.10 & 0.12 \\ 0.19 & 0.38 & 0 \end{bmatrix} \qquad d = \begin{bmatrix} 1800 \\ 200 \\ 900 \end{bmatrix}$$

(*a*) Explain the economic meaning of the elements 0.33, 0, and 200.
(*b*) Explain the economic meaning (if any) of the third-column sum.
(*c*) Explain the economic meaning (if any) of the third-row sum.
(*d*) Write out the specific input-output matrix equation for this model.

5 Find the solution output levels of the three industries in the preceding problem by Cramer's rule. (Round off answers to two decimal places.)

5.8 LIMITATIONS OF STATIC ANALYSIS

In the discussion of static equilibrium in the market or in the national income, our primary concern has been to find the equilibrium values of the endogenous variables in the model. A fundamental point that was ignored in such an analysis is the actual process of adjustments and readjustments of the variables ultimately leading to the equilibrium state (if it is at all attainable). We asked only about where we shall arrive but did not question when or what may happen along the way.

The static type of analysis fails, therefore, to take into account two problems of importance. One is that, since the adjustment process may take a long time to complete, an equilibrium state as determined within a particular frame of static analysis may have lost its relevance before it is even attained, if the exogenous forces in the model have undergone some changes in the meantime. This is the problem of shifts of the equilibrium state. The second is that, even if the adjustment process is allowed to run its course undisturbed, the equilibrium state envisaged in a static analysis may be altogether unattainable. This would be the case of a so-called "unstable equilibrium," which is characterized by the fact that the adjustment process will drive the variables further away from, rather than progressively closer to, that equilibrium state. To disregard the adjustment process, therefore, is to assume away the problem of attainability of equilibrium.

The shifts of the equilibrium state (in response to exogenous changes) pertain to a type of analysis called *comparative statics*, and the question of attainability and stability of equilibrium falls within the realm of *dynamic analysis*. Each of these clearly serves to fill a significant gap in the static analysis, and it is thus imperative to inquire into those areas of analysis also. We shall leave the study of dynamic analysis to Part 5 of the book and shall next turn our attention to the problem of comparative statics.