SECTION 1

WHAT MATRICES ARE AND SOME BASIC OPERATIONS WITH THEM

1.1 INTRODUCTION

This section will introduce matrices and show how they are useful to represent data. It will review some basic matrix operations including matrix addition and multiplication. Some examples to illustrate why they are interesting and important for statistical applications will be given. The representation of a linear model using matrices will be shown.

1.2 WHAT ARE MATRICES AND WHY ARE THEY INTERESTING TO A STATISTICIAN?

Matrices are rectangular arrays of numbers. Some examples of such arrays are

$$A = \begin{bmatrix} 4 & -2 & 1 & 0 \\ 0 & 5 & 3 & -7 \end{bmatrix}, B = \begin{bmatrix} 1 \\ -2 \\ 6 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 0.2 & 0.5 & 0.6 \\ 0.7 & 0.1 & 0.8 \\ 0.9 & 0.4 & 0.3 \end{bmatrix}.$$

Often data may be represented conveniently by a matrix. We give an example to illustrate how.

Matrix Algebra for Linear Models, First Edition. Marvin H. J. Gruber.

Example 1.1 Representing Data by Matrices

An example that lends itself to statistical analysis is taken from the Economic Report of the President of the United States in 1988. The data represent the relationship between a dependent variable Y (personal consumption expenditures) and three other independent variables X_1, X_2, and X_3. The variable X_1 represents the gross national product, X_2 represents personal income (in billions of dollars), and X_3 represents the total number of employed people in the civilian labor force (in thousands). Consider this data for the years 1970–1974 in Table 1.1.

TABLE 1.1 Consumption expenditures in terms of gross national product, personal income, and total number of employed people

Obs	Year	Y	X_1	X_2	X_3
1	1970	640.0	1015.5	831.8	78,678
2	1971	691.6	1102.7	894.0	79,367
3	1972	757.6	1212.8	981.6	82,153
4	1973	837.2	1359.3	1101.7	85,064
5	1974	916.5	1472.4	1210.1	86,794

The dependent variable may be represented by a matrix with five rows and one column. The independent variables could be represented by a matrix with five rows and three columns. Thus,

$$Y = \begin{bmatrix} 640.0 \\ 691.6 \\ 757.6 \\ 837.2 \\ 916.5 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} 1015.5 & 831.8 & 78{,}678 \\ 1102.7 & 894.0 & 79{,}367 \\ 1212.8 & 981.6 & 82{,}153 \\ 1359.3 & 1101.7 & 85{,}064 \\ 1472.8 & 1210.1 & 86{,}794 \end{bmatrix}.$$

A matrix with m rows and n columns is an $m \times n$ matrix. Thus, the matrix Y in Example 1.1 is 5×1 and the matrix X is 5×3. A square matrix is one that has the same number of rows and columns. The individual numbers in a matrix are called the elements of the matrix. □

We now give an example of an application from probability theory that uses matrices.

Example 1.2 A “Musical Room” Problem

Another somewhat different example is the following. Consider a triangular-shaped building with four rooms one at the center, room 0, and three rooms around it numbered 1, 2, and 3 clockwise (Fig. 1.1).

There is a door from room 0 to rooms 1, 2, and 3 and doors connecting rooms 1 and 2, 2 and 3, and 3 and 1. There is a person in the building. The room that he/she is

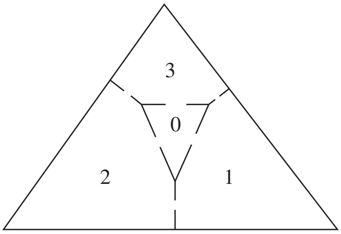

FIGURE 1.1 Building with four rooms.

in is the state of the system. At fixed intervals of time, he/she rolls a die. If he/she is in room 0 and the outcome is 1 or 2, he/she goes to room 1. If the outcome is 3 or 4, he/she goes to room 2. If the outcome is 5 or 6, he/she goes to room 3. If the person is in room 1, 2, or 3 and the outcome is 1 or 2, he/she advances one room in the clockwise direction. If the outcome is 3 or 4, he/she advances one room in the counterclockwise direction. An outcome of 5 or 6 will cause the person to return to room 0. Assume the die is fair.

Let p_{ij} be the probability that the person goes from room i to room j. Then we have the table of transitions

room	0	1	2	3
0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$
3	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0

that indicates

$$
\begin{aligned}
p_{00} &= p_{11} = p_{22} = p_{33} = 0 \\
p_{01} &= p_{02} = p_{03} = \frac{1}{3} \\
p_{12} &= p_{13} = p_{10} = \frac{1}{3} \\
p_{21} &= p_{23} = p_{20} = \frac{1}{3} \\
p_{31} &= p_{32} = p_{30} = \frac{1}{3}.
\end{aligned}
$$

Then the transition matrix would be

$$P=\begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix}.$$

□

Matrices turn out to be handy for representing data. Equations involving matrices are often used to study the relationship between variables.

More explanation of how this is done will be offered in the sections of the book that follow.

The matrices to be studied in this book will have elements that are real numbers. This will suffice for the study of linear models and many other topics in statistics. We will not consider matrices whose elements are complex numbers or elements of an arbitrary ring or field.

We now consider some basic operations using matrices.

1.3 MATRIX NOTATION, ADDITION, AND MULTIPLICATION

We will show how to represent a matrix and how to add and multiply two matrices.

The elements of a matrix A are denoted by a_{ij} meaning the element in the ith row and the jth column. For example, for the matrix

$$C=\begin{bmatrix} 0.2 & 0.5 & 0.6 \\ 0.7 & 0.1 & 0.8 \\ 0.9 & 0.4 & 1.3 \end{bmatrix}$$

$c_{11}=0.2$, $c_{12}=0.5$, and so on. Three important operations include matrix addition, multiplication of a matrix by a scalar, and matrix multiplication. Two matrices A and B can be added only when they have the same number of rows and columns. For the matrix $C=A+B$, $c_{ij}=a_{ij}+b_{ij}$; in other words, just add the elements algebraically in the same row and column. The matrix $D=\alpha A$ where α is a real number has elements $d_{ij}=\alpha a_{ij}$; just multiply each element by the scalar. Two matrices can be multiplied only when the number of columns of the first matrix is the same as the number of rows of the second one in the product. The elements of the $n\times p$ matrix $E=AB$, assuming that A is $n\times m$ and B is $m\times p$, are

$$e_{ij}=\sum_{k=1}^{m} a_{ik}b_{kj}, \quad 1\le i\le m,\ 1\le j\le p.$$

Example 1.3 Illustration of Matrix Operations

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

Then

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

$$D = 3A = \begin{bmatrix} 3 & -3 \\ 6 & 9 \end{bmatrix},$$

and

$$E = AB = \begin{bmatrix} 1(-1)+(-1)(3) & 1(2)+(-1)(4) \\ 2(-1)+3(3) & 2(2)+3(4) \end{bmatrix} = \begin{bmatrix} -4 & -2 \\ 7 & 16 \end{bmatrix}. \qquad \square$$

Example 1.4 Continuation of Example 1.2

Suppose that elements of the row vector $\pi^{(0)} = \begin{bmatrix} \pi_0^{(0)} & \pi_1^{(0)} & \pi_2^{(0)} & \pi_3^{(0)} \end{bmatrix}$ where $\sum_{i=0}^{3} \pi_i^{(0)} = 1$ represent the probability that the person starts in room i. Then $\pi^{(1)} = \pi^{(0)}P$. For example, if

$$\pi^{(0)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{12} & \frac{1}{4} \end{bmatrix}$$

the probabilities the person is in room 0 initially are 1/2, room 1 1/6, room 2 1/12, and room 3 1/4, then

$$\pi^{(1)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{12} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & \frac{5}{18} & \frac{11}{36} & \frac{1}{4} \end{bmatrix}.$$

Thus, after one transition given the initial probability vector above the probabilities that the person ends up in room 0, room 1, room 2, or room 3 after one transition are 1/6, 5/18, 11/36, and 1/4, respectively. This example illustrates a discrete Markov chain. The possible transitions are represented as elements of a matrix.

Suppose we want to know the probabilities that a person goes from room i to room j after two transitions. Assuming that what happens at each transition is independent, we could multiply the two matrices. Then

$$P^2 = P \cdot P = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{9} & \frac{2}{9} & \frac{2}{9} \\ \frac{2}{9} & \frac{1}{3} & \frac{2}{9} & \frac{2}{9} \\ \frac{2}{9} & \frac{2}{9} & \frac{1}{3} & \frac{2}{9} \\ \frac{2}{9} & \frac{2}{9} & \frac{2}{9} & \frac{1}{3} \end{bmatrix}.$$

Thus, for example, if the person is in room 1, the probability that he/she returns there after two transitions is 1/3. The probability that he/she winds up in room 3 is 2/9. Also when $\pi^{(0)}$ is the initial probability vector, we have that $\pi^{(2)}=\pi^{(1)}P=\pi^{(0)}P^2$. The reader is asked to find $\pi^{(2)}$ in Exercise 1.17. □

Two matrices are equal if and only if their corresponding elements are equal. More formally, $A=B$ if and only if $a_{ij}=b_{ij}$ for all $1\leq i\leq m$ and $1\leq j\leq n$.

Most, but not all, of the rules for addition and multiplication of real numbers hold true for matrices. The associative and commutative laws hold true for addition. The zero matrix is the matrix with all of the elements zero. An additive inverse of a matrix A would be $-A$, the matrix whose elements are $(-1)a_{ij}$. The distributive laws hold true.

However, there are several properties of real numbers that do not hold true for matrices. First, it is possible to have divisors of zero. It is not hard to find matrices A and B where $AB=0$ and neither A or B is the zero matrix (see Example 1.4).

In addition the cancellation rule does not hold true. For real nonzero numbers a, b, c, $ba=ca$ would imply that $b=c$. However (see Example 1.5) for matrices, $BA=CA$ may not imply that $B=C$.

Not every matrix has a multiplicative inverse. The identity matrix denoted by I has all ones on the longest (main) diagonal ($a_{ij}=1$) and zeros elsewhere ($a_{ij}=0,\ i\neq j$). For a matrix A, a multiplicative inverse would be a matrix such that $AB=I$ and $BA=I$. Furthermore, for matrices A and B, it is not often true that $AB=BA$. In other words, matrices do not satisfy the commutative law of multiplication in general.

The transpose of a matrix A is the matrix A' where the rows and the columns of A are exchanged. For example, for the matrix A in Example 1.3,

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

A matrix A is symmetric when $A=A'$. If $A=-A'$, the matrix is said to be skew symmetric. Symmetric matrices come up often in statistics.

Example 1.5 Two Nonzero Matrices Whose Product Is Zero

Consider the matrix

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

Notice that

$$AB=\begin{bmatrix}1(2)+2(-1) & 1(-2)+2(1)\\ 1(2)+2(-1) & 1(-2)+2(1)\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

□

Example 1.6 The Cancellation Law for Real Numbers Does Not Hold for Matrices

Consider matrices A, B, C where

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, B = \begin{bmatrix} 5 & 4 \\ 7 & 3 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 3 & 6 \\ 8 & 2 \end{bmatrix}.$$

Now

$$BA = CA = \begin{bmatrix} 9 & 18 \\ 10 & 20 \end{bmatrix}$$

but $B \neq C$. □

Matrix theory is basic to the study of linear models. Example 1.7 indicates how the basic matrix operations studied so far are used in this context.

Example 1.7 The Linear Model

Let Y be an n-dimensional vector of observations, an $n \times 1$ matrix. Let X be an $n \times m$ matrix where each column has the values of a prediction variable. It is assumed here that there are m predictors. Let β be an $m \times 1$ matrix of parameters to be estimated. The prediction of the observations will not be exact. Thus, we also need an n-dimensional column vector of errors ε. The general linear model will take the form

$$Y = X\beta + \varepsilon. \tag{1.1}$$

Suppose that there are five observations and three prediction variables. Then $n=5$ and $m=3$. As a result, we would have the multiple regression equation

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i, \quad 1 \leq i \leq 5. \tag{1.2}$$

Equation (1.2) may be represented by the matrix equation

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{31} \\ 1 & x_{12} & x_{22} & x_{32} \\ 1 & x_{13} & x_{23} & x_{33} \\ 1 & x_{14} & x_{24} & x_{34} \\ 1 & x_{15} & x_{25} & x_{35} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix}. \tag{1.3}$$

In experimental design models, the matrix is frequently zeros and ones indicating the level of a factor. An example of such a model would be

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \\ Y_{14} \\ Y_{21} \\ Y_{22} \\ Y_{23} \\ Y_{31} \\ Y_{32} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \\ \varepsilon_{31} \\ \varepsilon_{32} \end{bmatrix}. \tag{1.4}$$

This is an unbalanced one-way analysis of variance (ANOVA) model where there are three treatments with four observations of treatment 1, three observations of treatment 2, and two observations of treatment 3. Different kinds of ANOVA models will be studied in Part IV. □

1.4 SUMMARY

We have accomplished the following. First, we have explained what matrices are and illustrated how they can be used to summarize data. Second, we defined three basic matrix operations: addition, scalar multiplication, and matrix multiplication. Third, we have shown how matrices have some properties similar to numbers and do not share some properties that numbers have. Fourth, we have given some applications to probability and to linear models.

EXERCISES

1.1 Let

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

Find AB and BA.

1.2 Let

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

A. Show that $CD = DC = 0$.

B. Does C or D have a multiplicative inverse? If yes, find it. If not, why not?

1.3 Let

$$E=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} \quad \text{and} \quad F=\begin{bmatrix}4 & 3\\ 2 & 1\end{bmatrix}.$$

Show that $EF \neq FE$.

1.4 Let

$$G=\begin{bmatrix}5 & 1\\ 1 & 5\end{bmatrix} \quad \text{and} \quad H=\begin{bmatrix}7 & -1\\ -1 & 7\end{bmatrix}.$$

A. Show that $GH=HG$.
B. Does this commutativity hold when

$$G=\begin{bmatrix}a & b\\ b & a\end{bmatrix}, H=\begin{bmatrix}c & -b\\ -b & c\end{bmatrix}$$

1.5 A diagonal matrix is a square matrix for which all of the elements that are not in the main diagonal are zero. Show that diagonal matrices commute.

1.6 Let P be a matrix with the property that $PP'=I$ and $P'P=I$. Let D_1 and D_2 be diagonal matrices. Show that the matrices $P'D_1P$ and $P'D_2P$ commute.

1.7 Show that any matrix is the sum of a symmetric matrix and a skew-symmetric matrix.

1.8 Show that in general for any matrices A and B that

A. $A'' = A$.
B. $(A+B)' = A' + B'$.
C. $(AB)' = B'A'$.

1.9 Show that if A and B commute, then

$$A'B' = B'A'.$$

1.10 Determine whether the matrices

$$A=\begin{bmatrix}\frac{5}{2} & -\frac{3}{2}\\ -\frac{3}{2} & \frac{5}{2}\end{bmatrix} \quad \text{and} \quad B=\begin{bmatrix}\frac{13}{2} & -\frac{5}{2}\\ -\frac{5}{2} & \frac{13}{2}\end{bmatrix}$$

commute.

1.11 For the model (1.3), write the entries of $X'X$ using the appropriate sum notation.

1.12 For the data on gross national product in Example 1.1

A. What is the X matrix? The Y matrix?
B. Write the system of equations $X'X\beta = X'Y$ with numbers.
C. Find the values of the β parameters that satisfy the system.

1.13 For the model in (1.4)

A. Write out the nine equations represented by the matrix equation.
B. Find $X'X$.
C. What are the entries of $X'Y$? Use the appropriate sum notation.
D. Write out the system of four equations $X'X\alpha = X'Y$.
E. Let

$$G = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

Show that $\alpha = GX'Y$ satisfies the system of equations in D. The matrix G is an example of a generalized inverse. Generalized inverses will be studied in Part III.

1.14 Show that for any matrix X, $X'X$, and XX' are symmetric matrices.

1.15 Let A and B be two 2×2 matrices where the rows and the columns add up to 1. Show that AB has this property.

1.16 Consider the linear model

$$\begin{aligned} y_{11} &= \mu + \alpha_1 + \beta_1 + \varepsilon_{11} \\ y_{12} &= \mu + \alpha_1 + \beta_2 + \varepsilon_{12} \\ y_{21} &= \mu + \alpha_2 + \beta_1 + \varepsilon_{21} \\ y_{22} &= \mu + \alpha_2 + \beta_2 + \varepsilon_{22}. \end{aligned}$$

A. Tell what the matrices should be for a model in the form

$$Y = X\gamma + \varepsilon$$

where Y is 4×1, X is 4×5, γ is 5×1, and ε is 4×1.
B. Find $X'X$.

1.17 **A.** Find P^3 for the transition matrix in Example 1.2.
B. Given the initial probability vector in Example 1.4, find

$$\pi^{(2)}, \pi^{(3)} = \pi^{(2)}P.$$

1.18 Suppose in Example 1.2 two coins are flipped instead of a die. A person in room 0 goes to room 1 if no heads are obtained, room 2 if one head is obtained, and room 3 if two heads are obtained. A person in rooms 1, 2, or 3 advances one room in the clockwise direction if no heads are obtained, goes to room 0 if one head is obtained, and advances one room in the counterclockwise direction if two heads are obtained. Assume that the coins are fair.

A. Find the transition matrix P.

B. Find P, P^2, and P^3.

C. Given $\pi^{(0)} = [1/2,\ 1/4,\ 1/16,\ 3/16]$, find $\pi^{(1)}, \pi^{(2)}, \pi^{(3)}$.

1.19 Give examples of nonzero 3×3 matrices whose product is zero and for which the cancellation rule fails.

1.20 A matrix A is nilpotent if there is an n such that $A^n = 0$. Show that

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

is a nilpotent matrix.