APPENDIX A

MATRIX ALGEBRA: AN INTRODUCTION

Matrix algebra is one of the most useful and powerful branches of mathematics for conceptualizing and analyzing psychological, sociological, and educational research data. As research becomes more and more multivariate, the need for a compact method of expressing data becomes greater. Certain problems require that sets of equations and subscripted variables be written. In many cases the use of matrix algebra simplifies and, when familiar, clarifies the mathematics and statistics. In addition, matrix algebra notation and thinking fit in nicely with the conceptualization of computer programming and use.

This chapter provides a brief introduction to matrix algebra. The emphasis is on those aspects that are related to subject matter covered in this book. Thus many matrix algebra techniques, important and useful in other contexts, are omitted. In addition, certain important derivations and proofs are neglected. Although the material presented here should suffice to enable you to follow the applications of matrix algebra in this book, it is strongly suggested that you expand your knowledge of this topic by studying one or more of the following texts: Dorf (1969), Green (1976), Hohn (1964), Horst (1963), and Searle (1966). In addition, you will find good introductions to matrix algebra in the books on multivariate analysis that were cited in Chapter 17.

BASIC DEFINITIONS

A *matrix* is an n-by-k rectangle of numbers or symbols that stand for numbers. The order of the matrix is n by k. It is customary to designate the rows first and the columns second. That is, n is the number of rows of the matrix and k the number of columns. A 2-by-3 matrix called **A** might be

$$\mathbf{A} = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 4 & 7 & 5 \\ 6 & 6 & 3 \end{bmatrix} \end{array}$$

Elements of a matrix are identified by reference to the row and column that they occupy. Thus, a_{11} refers to the element of the first row and first column of **A**, which in the above example is 4. Similarly, a_{23} is the element of the second row and third column of **A**, which in the above example is 3. In general, then a_{ij} refers to the element in row i and column j.

The *transpose* of a matrix is obtained simply by exchanging rows and columns. In the present case, the transpose of **A**, written **A**′, is

$$\mathbf{A}' = \begin{bmatrix} 4 & 6 \\ 7 & 6 \\ 5 & 3 \end{bmatrix}$$

If $n = k$, the matrix is square. A square matrix can be symmetric or asymmetric. A *symmetric* matrix has the same elements above the principal diagonal as below the diagonal except that they are transposed. The principal diagonal is the set of elements from the upper left corner to the lower right corner. Symmetric matrices are frequently encountered in multiple regression analysis and in multivariate analysis. The following is an example of a correlation matrix, which is symmetric:

$$\mathbf{R} = \begin{bmatrix} 1.00 & .70 & .30 \\ .70 & 1.00 & .40 \\ .30 & .40 & 1.00 \end{bmatrix}$$

Diagonal elements refer to correlations of variables with themselves, hence the 1's. Each off-diagonal element refers to a correlation between two variables and is identified by row and column numbers. Thus, $r_{12} = r_{21} = .70$; $r_{23} = r_{32} = .40$. And similarly for other elements.

A *column vector* is an n-by-1 array of numbers. For example:

$$\mathbf{b} = \begin{bmatrix} 8.0 \\ 1.3 \\ -2.0 \end{bmatrix}$$

A *row vector* is a 1-by-n array of numbers:

$$\mathbf{b}' = [8.0 \quad 1.3 \quad -.20]$$

$\mathbf{b}'$ is the *transpose* of $\mathbf{b}$. Note that vectors are designated by lowercase boldface letters, and that a prime is used to indicate a row vector.

A *diagonal* matrix is frequently encountered in statistical work. It is simply a matrix in which some values other than zero are in the principal diagonal of the matrix, and all the off-diagonal elements are zeros. Here is a diagonal matrix:

$$\begin{bmatrix} 2.759 & 0 & 0 \\ 0 & 1.643 & 0 \\ 0 & 0 & .879 \end{bmatrix}$$

A particularly important form of a diagonal matrix is an *identity* matrix, **I**, which has 1's in the principal diagonal:

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

MATRIX OPERATIONS

The power of matrix algebra becomes apparent when we explore the operations that are possible. The major operations are addition, subtraction, multiplication, and inversion. A large number of statistical operations can be done by knowing the basic rules of matrix algebra. Some matrix operations are now defined and illustrated.

Addition and Subtraction

Two or more vectors can be added or subtracted provided they are of the same dimensionality. That is, they have the same number of elements. The following two vectors are added:

$$\underset{\mathbf{a}}{\begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}} + \underset{\mathbf{b}}{\begin{bmatrix} 7 \\ 7 \\ 4 \end{bmatrix}} = \underset{\mathbf{c}}{\begin{bmatrix} 11 \\ 10 \\ 9 \end{bmatrix}}$$

Similarly, matrices of the same dimensionality may be added or subtracted. The following two 3-by-2 matrices are added:

$$\underset{\mathbf{A}}{\begin{bmatrix} 6 & 4 \\ 5 & 6 \\ 9 & 5 \end{bmatrix}} + \underset{\mathbf{B}}{\begin{bmatrix} 7 & 4 \\ 7 & 4 \\ 1 & 3 \end{bmatrix}} = \underset{\mathbf{C}}{\begin{bmatrix} 13 & 8 \\ 12 & 10 \\ 10 & 8 \end{bmatrix}}$$

Now, **B** is subtracted from **A**:

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

Multiplication

To obtain the product of a row vector by a column vector, corresponding elements of each are multiplied and then added. For example, the multiplication of **a**′ by **b**, each consisting of three elements, is:

$$\underset{\mathbf{a}'}{[a_1 \quad a_2 \quad a_3]} \underset{\mathbf{b}}{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}} = a_1b_1 + a_2b_2 + a_3b_3$$

Note that the product of a row by a column is a single number called a *scalar*. This is why the product of a row by a column is referred to as the scalar product of vectors.

Here is a numerical example:

$$[4 \quad 1 \quad 3]\begin{bmatrix}1\\2\\5\end{bmatrix} = (4)(1) + (1)(2) + (3)(5) = 21$$

Scalar products of vectors are very frequently used in statistical analysis. For example, to obtain the sum of the elements of a column vector it is premultiplied by a unit row vector of the same dimensionality. Thus,

$$\Sigma X: \qquad [1 \quad 1 \quad 1 \quad 1 \quad 1]\begin{bmatrix}1\\4\\1\\3\\7\end{bmatrix} = 16$$

The sum of the squares of a column vector is obtained by premultiplying the vector by its transpose.

$$\Sigma X^2: \qquad [1 \quad 4 \quad 1 \quad 3 \quad 7]\begin{bmatrix}1\\4\\1\\3\\7\end{bmatrix} = 76$$

Similarly, the sum of the products of X and Y is obtained by multiplying the row of X by the column of Y, or the row of Y by the column of X.

$$\Sigma XY: \qquad [1 \quad 4 \quad 1 \quad 3 \quad 7]\begin{bmatrix}3\\-5\\7\\2\\-1\end{bmatrix} = -11$$

Scalar products of vectors are used frequently in this book, particularly in Chapters 17 and 18.

Instead of multiplying a row vector by a column vector, one may multiply a column vector by a row vector. The two operations are entirely different from each other. It was shown above that the former results in a scalar. The latter operation, on the other hand, results in a matrix. This is why it is referred to as the matrix product of vectors. For example,

$$\begin{bmatrix}3\\-5\\7\\2\\-1\end{bmatrix}[1 \quad 4 \quad 1 \quad 3 \quad 7] = \begin{bmatrix}3 & 12 & 3 & 9 & 21\\-5 & -20 & -5 & -15 & -35\\7 & 28 & 7 & 21 & 49\\2 & 8 & 2 & 6 & 14\\-1 & -4 & -1 & -3 & -7\end{bmatrix}$$

Note that each element of the column is multiplied, in turn, by each element of the row to obtain one element of the matrix. The products of the first element of the column by the row elements become the first row of the matrix. Those of the second element of the column by the row become the second row of the matrix, and so forth. Thus, the matrix product of a column vector of k elements and a row vector of k elements is a $k \times k$ matrix.

Matrix multiplication is done by multiplying rows by columns. An illustration is easier than verbal explanation. Suppose we want to multiply two matrices, **A** and **B,** to produce the product matrix, **C:**

$$\underset{\mathbf{A}}{\begin{bmatrix} 3 & 1 \\ 5 & 1 \\ 2 & 4 \end{bmatrix}} \times \underset{\mathbf{B}}{\begin{bmatrix} 4 & 1 & 4 \\ 5 & 6 & 2 \end{bmatrix}} = \underset{\mathbf{C}}{\begin{bmatrix} 17 & 9 & 14 \\ 25 & 11 & 22 \\ 28 & 26 & 16 \end{bmatrix}}$$

Following the rule of scalar product of vectors, we multiply and add as follows (follow the arrows):

$$\begin{array}{lll} (3)(4) + (1)(5) = 17 & (3)(1) + (1)(6) = 9 & (3)(4) + (1)(2) = 14 \\ (5)(4) + (1)(5) = 25 & (5)(1) + (1)(6) = 11 & (5)(4) + (1)(2) = 22 \\ (2)(4) + (4)(5) = 28 & (2)(1) + (4)(6) = 26 & (2)(4) + (4)(2) = 16 \end{array}$$

From the foregoing illustration it may be discerned that in order to multiply two matrices it is necessary that the number of columns of the first matrix be equal to the number of rows of the second matrix. This is referred to as the *conformability* condition. Thus, for example, an n-by-k matrix can be multiplied by a k-by-m matrix because the number of columns of the first (k) is equal to the number of rows of the second (k). In this context, the k's are referred to as the "interior" dimensions; n and m are referred to as the "exterior" dimensions.

Two matrices are conformable when they have the same "interior" dimensions. There are no restrictions on the "exterior" dimensions when two matrices are multiplied. It is useful to note that the "exterior" dimensions of two matrices being multiplied become the dimensions of the product matrix. For example, when a 3-by-2 matrix is multiplied by a 2-by-5 matrix, a 3-by-5 matrix is obtained:

$$(3\text{-by-}2) \times (2\text{-by-}5) = (3\text{-by-}5)$$

In general,

$$(n\text{-by-}k) \times (k\text{-by-}m) = (n\text{-by-}m)$$

A special case of matrix multiplication often encountered in statistical work is the multiplication of a matrix by its transpose to obtain a matrix of raw score, or deviation, Sums of Squares and Cross Products (SSCP). Assume that there are n subjects for whom measures on k variables are available. In other words, assume that the data matrix, **X**, is an n-by-k. To obtain the raw score SSCP calculate **X′X.** Here is a numerical example:

$$\underset{\mathbf{X}'}{k\begin{bmatrix} 1 & 4 & 1 & 3 & 7 \\ 2 & 3 & 3 & 4 & 6 \\ 2 & 5 & 1 & 3 & 5 \end{bmatrix}^{n}} \; \underset{\mathbf{X}}{n\begin{bmatrix} 1 & 2 & 2 \\ 4 & 3 & 5 \\ 1 & 3 & 1 \\ 3 & 4 & 3 \\ 7 & 6 & 5 \end{bmatrix}^{k}} = \underset{\mathbf{X'X}}{\begin{bmatrix} 76 & 71 & 67 \\ 71 & 74 & 64 \\ 67 & 64 & 64 \end{bmatrix}}$$

In statistical symbols, **X′X** is

$$\Sigma X_i X_j = \begin{bmatrix} \Sigma X_1^2 & \Sigma X_1 X_2 & \Sigma X_1 X_3 \\ \Sigma X_2 X_1 & \Sigma X_2^2 & \Sigma X_2 X_3 \\ \Sigma X_3 X_1 & \Sigma X_3 X_2 & \Sigma X_3^2 \end{bmatrix}$$

Using similar operations one may obtain deviation SSCP matrices. Such matrices are used frequently in this book (see, in particular, Chapters 4, 17, and 18).

A matrix can be multipled by a scalar: each element of the matrix is multiplied by the scalar. Suppose, for example, we want to calculate the mean of each of the elements of a matrix of sums of scores. Let $N = 10$. The operation is

$$\frac{1}{10}\begin{bmatrix} 20 & 48 \\ 30 & 40 \\ 35 & 39 \end{bmatrix} = \begin{bmatrix} 2.0 & 4.8 \\ 3.0 & 4.0 \\ 3.5 & 3.9 \end{bmatrix}$$

Each element of the matrix is multiplied by the scalar $^1/_{10}$.

A matrix can be multiplied by a vector. The first example given below is premultiplication by a vector, the second is postmultiplication:

$$\begin{bmatrix} 6 & 5 & 2 \end{bmatrix} \begin{bmatrix} 7 & 3 \\ 7 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 85 & 30 \end{bmatrix}$$

$$\begin{bmatrix} 7 & 7 & 4 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 85 \\ 30 \end{bmatrix}$$

Note that in the latter example, (2-by-3) × (3-by-1) becomes (2-by-1). This sort of multiplication of a matrix by a vector is done frequently in multiple regression analysis (see, for example, Chapter 4).

Thus far, nothing has been said about the operation of division in matrix algebra. In order to show how this is done it is necessary first to discuss some other concepts, to which we now turn.

DETERMINANTS

A *determinant* is a certain numerical value associated with a square matrix. The determinant of a matrix is indicated by vertical lines instead of brackets. For example, the determinant of a matrix **B** is written

$$\det \mathbf{B} = |\mathbf{B}| = \underset{\mathbf{B}}{\begin{vmatrix} 4 & 2 \\ 1 & 5 \end{vmatrix}}$$

The calculation of the determinant of a 2 × 2 matrix is very simple: it is the product of the elements of the principal diagonal minus the product of the remaining two elements. For the above matrix,

$$|\mathbf{B}| = \begin{vmatrix} 4 & 2 \\ 1 & 5 \end{vmatrix} = (4)(5) - (1)(2) = 20 - 2 = \mathit{18}$$

or, symbolically,

$$|\mathbf{B}| = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = b_{11}b_{22} - b_{12}b_{21}$$

The calculation of determinants for larger matrices is quite tedious, and will not be shown here (see references cited in the beginning of the chapter). In any event, matrix operations are most often done with the aid of a computer. The purpose here is solely to indicate the role played by determinants in some applications of statistical analysis.

Applications of Determinants

To give the flavor of the place and usefulness of determinants in statistical analysis, we turn first to two simple correlation examples. Suppose we have two correlation coefficients, r_{y1} and r_{y2}, calculated between a dependent variable, Y, and two variables, 1 and 2. The correlations are $r_{y1} = .80$ and $r_{y2} = .20$. We set up two matrices that express the two relations, but this is done immediately in the form of determinants, whose numerical values are calculated:

$$\begin{matrix} 1 & y \end{matrix}$$
$$\begin{vmatrix} 1.00 & .80 \\ .80 & 1.00 \end{vmatrix} = (1.00)(1.00) - (.80)(.80) = \mathit{.36}$$

and

$$\begin{matrix} 2 & y \end{matrix}$$
$$\begin{vmatrix} 1.00 & .20 \\ .20 & 1.00 \end{vmatrix} = (1.00)(1.00) - (.20)(.20) = .96$$

The two determinants are .36 and .96.

Now, to determine the percentage of variance shared by y and 1 and by y and 2, square the r's:

$$r_{y1}^2 = (.80)^2 = .64$$
$$r_{y2}^2 = (.20)^2 = .04$$

Subtract each of these from 1.00: 1.00 − .64 = *.36*, and 1.00 − .04 = *.96*. These are the determinants just calculated. They are $1 - r^2$, or the proportions of the variance not accounted for.

As an extension of the foregoing demonstration, it may be shown how the squared multiple correlation, R^2, can be calculated with determinants:

$$R_{y.12\ldots k}^2 = 1 - \frac{|\mathbf{R}|}{|\mathbf{R}_x|}$$

where $|\mathbf{R}|$ is the determinant of the correlation matrix of al the variables, that is, the independent variables as well as the dependent variable; $|\mathbf{R}_x|$ is the determinant of the correlation matrix of the independent variables. From the foregoing it can be seen that the ratio of the two determinants indicates the proportion of variance of the dependent variable, Y, *not* accounted by the independent variables, X's. (See Study Suggestions 6 and 7 at the end of this Appendix.)

The ratio of two determinants is also frequently used in multivariate analyses (see the sections of Chapters 17 and 18 dealing with Wilks' Λ).

Another important use of determinants is related to the concept of linear dependencies, to which we now turn.

Linear Dependence

Linear dependence means that one or more vectors of a matrix, rows or columns, are a linear combination of other vectors of the matrix. The vectors $\mathbf{a}' = [3 \quad 1 \quad 4]$ and $\mathbf{b}' = [6 \quad 2 \quad 8]$ are dependent since $2\mathbf{a}' = \mathbf{b}'$. If one vector is a function of another in this manner, the coefficient of correlation between them is 1.00. Dependence in a matrix can be defined by reference to its determinant. If the determinant of the matrix is zero it means that the matrix contains at least one linear dependency. Such a matrix is referred to as being *singular*. For example, calculate the determinant of the following matrix:

$$\begin{vmatrix} 3 & 1 \\ 6 & 2 \end{vmatrix} = (3)(2) - (1)(6) = 0$$

The matrix is singular, that is, it contains a linear dependency. Note that the values of the second row are twice the values of the first row.

A matrix whose determinant is not equal to zero is referred to as being *nonsingular*. The notions of singularity and nonsingularity of matrices play very important roles in statistical analysis. For example, in Chapter 8 issues regarding multicollinearity are discussed in reference to the determinant of the correlation matrix of the independent variables. As is shown below, a singular matrix has no inverse.

We turn now to the operation of division in matrix algebra, which is presented in the context of the discussion of matrix inversion.

MATRIX INVERSE

Recall that the division of one number into another number amounts to multiplying the dividend by the reciprocal of the divisor:

$$\frac{a}{b} = \frac{1}{b}a$$

For example, 12/4 = 12(1/4) = (12)(.25) = 3. Analogously, in matrix algebra, instead of dividing a matrix **A** by another matrix **B** to obtain matrix **C**, we multiply **A** by the *inverse* of **B** to obtain **C**. The inverse of **B** is written $\mathbf{B}^{-1}$. Suppose, in ordinary algebra, we had $ab = c$, and we wanted to find b. We would write

$$b = \frac{c}{a}$$

In matrix algebra, we write

$$\mathbf{B} = \mathbf{A}^{-1}\mathbf{C}$$

(Note that **C** is premultiplied by $\mathbf{A}^{-1}$ and not postmultiplied. In general, $\mathbf{A}^{-1}\mathbf{C} \neq \mathbf{C}\mathbf{A}^{-1}$.)

The formal definition of the inverse of a square matrix is: Given **A** and **B**, two square matrices, if $\mathbf{AB} = \mathbf{I}$, then **A** is the inverse of **B**.

Generally, the calculation of the inverse of a matrix is very laborious and, therefore, error prone. This is why it is best to use a computer program for such purposes (see below). Fortunately, however, the calculation of the inverse of a 2×2 matrix is very simple, and is shown here because: (1) it affords an illustration of the basic approach to the calculation of the inverse; (2) it affords the opportunity of showing the role played by the determinant in the calculation of the inverse; (3) inverses of 2×2 matrices are frequently calculated in some chapters of this book (see, in particular, Chapters 4, 17, and 18).

In order to show how the inverse of a 2×2 matrix is calculated it is necessary first to discuss briefly the *adjoint* of a such a matrix. This is shown in reference to the following matrix:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The adjoint of **A** is:

$$\text{adj } \mathbf{A} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Thus, to obtain the adjoint of a 2×2 matrix interchange the elements of its principal diagonal (a and d in the above example), and change the signs of the other two elements (b and c in the above example).[1]

Now the inverse of a matrix **A** is:

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|} = \frac{1}{|\mathbf{A}|} \text{adj } \mathbf{A}$$

where $|\mathbf{A}|$ is the determinant of **A.**

The inverse of the following matrix, **A,** is now calculated.

$$\mathbf{A} = \begin{bmatrix} 6 & 2 \\ 8 & 4 \end{bmatrix}$$

First, calculate the determinant of **A:**

$$|\mathbf{A}| = \begin{vmatrix} 6 & 2 \\ 8 & 4 \end{vmatrix} = (6)(4) - (2)(8) = 8$$

Second, form the adjoint of **A:**

$$\text{adj } \mathbf{A} = \begin{bmatrix} 4 & -2 \\ -8 & 6 \end{bmatrix}$$

Third, multiply the adj **A** by the reciprocal of $|\mathbf{A}|$ to obtain the inverse of **A:**

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{adj } \mathbf{A} = \frac{1}{8} \begin{bmatrix} 4 & -2 \\ -8 & 6 \end{bmatrix} = \begin{bmatrix} .50 & -.25 \\ -1.00 & .75 \end{bmatrix}$$

[1]For a general definition of the adjoint of a matrix, see references cited in the beginning of this chapter. Adjoints of 2×2 matrices are used frequently in Chapters 17 and 18.

It was said above that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. For the present example,

$$\mathbf{A}^{-1}\mathbf{A} = \underset{\mathbf{A}^{-1}}{\begin{bmatrix} .50 & -.25 \\ -1.00 & .75 \end{bmatrix}} \underset{\mathbf{A}}{\begin{bmatrix} 6 & 2 \\ 8 & 4 \end{bmatrix}} = \underset{\mathbf{I}}{\begin{bmatrix} 1.00 & 0 \\ 0 & 1.00 \end{bmatrix}}$$

It was said above that a matrix whose determinant is zero is singular. From the foregoing demonstration of the calculation of the inverse it should be clear that a singular matrix has no inverse. Although one does not generally encounter singular matrices in social science research, an unwary researcher may introduce singularity in the treatment of the data. For example, suppose that a test battery consisting of five subtests is used to predict a given criterion. If, under such circumstances, the researcher uses not only the scores on the five subtests but also a total score, obtained as the sum of the five subscores, he or she has introduced a linear dependency (see above), thereby rendering the matrix singular. Similarly, when one uses scores on two scales as well as the differences between them in the same matrix. Other situations when one should be on guard not to introduce linear dependencies in a matrix occur when coded vectors are used to represent categorical variables (see Chapter 9).

CONCLUSION

It is realized that this brief introduction to matrix algebra cannot serve to demonstrate its great power and elegance. To do this, it would be necessary to use matrices whose dimensions are larger than the ones used here for simplicity of presentation. To begin to appreciate the power of matrix algebra it is suggested that you think of the large data matrices frequently encountered in behavioral research. Using matrix algebra one can manipulate and operate upon large matrices with relative ease, when ordinary algebra will simply not do. For example, when in multiple regression analysis only two independent variables are used, it is relatively easy to do the calculations by ordinary algebra (see Chapter 3). But with increasing numbers of independent variables, the use of matrix algebra for the calculation of multiple regression analysis becomes a must. And, as is amply demonstrated in Parts 3 and 4 of this book, matrix algebra is the language of linear structural equation models and multivariate analysis. In short, to understand and be able to intelligently apply these methods it is essential that you develop a working knowledge of matrix algebra. It is therefore strongly suggested that you do some or all the calculations of the matrix operations presented in the various chapters, particularly those in Chapter 4, and in Parts 3 and 4 of the book. Furthermore, it is suggested that you learn to use computer programs when you have to manipulate relatively large matrices. Of the various computer programs for matrix manipulations, one of the best and most versatile is the MATRIX program of the SAS package. Examples of applications of this program are given in Chapters 8 and 15.

STUDY SUGGESTIONS

1. You will find it useful to work through some of the rules of matrix algebra. Use of the rules occurs again and again in multiple regression, factor analysis, discriminant analysis, canonical correlation, and multivariate analysis of variance. The most important of the rules are as follows:

(1) $\mathbf{ABC} = (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

This is the *associative rule* of matrix multiplication. It simply indicates that the multiplication of three (or more) matrices can be done by pairing and multiplying the first two matrices and then multiplying the product by the remaining matrix, or by pairing and multiplying the second two and then multiplying the product by the first matrix. Or we can regard the rule in the following way:

$$\mathbf{AB} = \mathbf{D}, \text{ then } \mathbf{DC}$$
$$\mathbf{BC} = \mathbf{E}, \text{ then } \mathbf{AE}$$

(2) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$

That is, the order of addition makes no difference. And the associative rule applies:

$$\mathbf{A} + \mathbf{B} + \mathbf{C} = (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$$

(3) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

This is the *distributive rule* of ordinary algebra.

(4) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

The transpose of the product of two matrices is equal to the transpose of their product in reverse order.

(5) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

This rule is the same as that in (4), above, except that it is applied to matrix inverses.

(6) $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$

This rule can be used as a proof that the calculation of the inverse of a matrix is correct.

(7) $\mathbf{AB} \neq \mathbf{BA}$

This is actually not a rule. It is included to emphasize that the order of the multiplication of matrices is important.

Here are three matrices, **A**, **B**, and **C**.

$$\underset{\mathbf{A}}{\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}} \quad \underset{\mathbf{B}}{\begin{pmatrix} 3 & 4 \\ 0 & 1 \end{pmatrix}} \quad \underset{\mathbf{C}}{\begin{pmatrix} 0 & 2 \\ 5 & 3 \end{pmatrix}}$$

(a) Demonstrate the associative rule by multiplying:

$$\mathbf{A} \times \mathbf{B}; \text{ then } \mathbf{AB} \times \mathbf{C}$$
$$\mathbf{B} \times \mathbf{C}; \text{ then } \mathbf{A} \times \mathbf{BC}$$

(b) Demonstrate the distributive rule using **A**, **B**, and **C** of (a), above.

(c) Using **B** and **C**, above, show that $\mathbf{BC} \neq \mathbf{CB}$.

2. What are the dimensions of the matrix that will result from multiplying a 3-by-6 matrix **A** by a 6-by-2 matrix **B**?
3. Given:

$$\mathbf{A} = \begin{bmatrix} 1.26 & -.73 \\ 2.12 & 1.34 \\ 4.61 & -.31 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 4.11 & 1.12 \\ & \\ -2.30 & -.36 \end{bmatrix}$$

What is **AB**?

4. When it is said that a matrix is singular, what does it imply about its determinant?
5. Calculate the inverse of the following matrix:

$$\begin{bmatrix} 15 & -3 \\ 6 & 12 \end{bmatrix}$$

6. In a study of Holtzman and Brown (1968), the correlations among measures of study habits and attitudes, scholastic aptitude, and grade-point averages were reported as follows:

	SHA	SA	GPA
SHA	1.00	.32	.55
SA	.32	1.00	.61
GPA	.55	.61	1.00

The determinant of this matrix is .4377. Calculate R^2 for GPA with SHA and SA. (*Hint:* You need to calculate the determinant of the matrix of the independent variables, and then use the two determinants for the calculation of R^2.)

7. Liddle (1958) reported the following correlations among intellectual ability, leadership ability, and withdrawn maladjustment:

	IA	LA	WM
IA	1.00	.37	−.28
LA	.37	1.00	−.61
WM	−.28	−.61	1.00

The determinant of this matrix is .5390. Calculate:
(a) the proportion of variance of WM *not* accounted for by IA and LA.
(b) R^2 of WM with IA and LA.
(c) using matrix algebra, the regression equation of WM on IA and LA. (See Chapter 4 for matrix equation.)

8. It is strongly suggested that you study one or more of the references cited in the beginning of this Appendix.

ANSWERS

1. (a)

$$\mathbf{ABC} = \begin{bmatrix} 55 & 45 \\ 30 & 24 \end{bmatrix}$$

(b)

$$\mathbf{A(B + C)} = \begin{bmatrix} 21 & 24 \\ 13 & 14 \end{bmatrix}$$

(c)

$$\mathbf{BC} = \begin{bmatrix} 20 & 18 \\ 5 & 3 \end{bmatrix} \qquad \mathbf{CB} = \begin{bmatrix} 0 & 2 \\ 15 & 23 \end{bmatrix}$$

2. 3-by-2

3.

$$\mathbf{AB} = \begin{bmatrix} 6.8576 & 1.6740 \\ 5.6312 & 1.8920 \\ 19.6601 & 5.2748 \end{bmatrix}$$

4. The determinant is zero.