11. Everitt, B. *Graphical Techniques for Multivariate Data.* New York: North-Holland, 1978.
12. Gable, G. G. "A Multidimensional Model of Client Success when Engaging External Consultants." *Management Science*, **42**, no. 8 (1996) 1175–1198.
13. Halinar, J. C. "Principal Component Analysis in Plant Breeding." Unpublished report based on data collected by Dr. F. A. Bliss, University of Wisconsin, 1979.
14. Johnson, R. A., and G. K. Bhattacharyya. *Statistics: Principles and Methods* (5th ed.). New York: John Wiley, 2005.
15. Kim, L., and Y. Kim. "Innovation in a Newly Industrializing Country: A Multiple Discriminant Analysis." *Management Science*, **31**, no. 3 (1985) 312–322.
16. Klatzky, S. R., and R. W. Hodge. "A Canonical Correlation Analysis of Occupational Mobility." *Journal of the American Statistical Association*, **66**, no. 333 (1971), 16–22.
17. Lee, J., "Relationships Between Properties of Pulp-Fibre and Paper." Unpublished doctoral thesis, University of Toronto. Faculty of Forestry (1992).
18. MacCrimmon, K., and D. Wehrung. "Characteristics of Risk Taking Executives." *Management Science*, **36**, no. 4 (1990), 422–435.
19. Marriott, F. H. C. *The Interpretation of Multiple Observations.* London: Academic Press, 1974.
20. Mather, P. M. "Study of Factors Influencing Variation in Size Characteristics in Fluvioglacial Sediments." *Mathematical Geology*, **4**, no. 3 (1972), 219–234.
21. McLaughlin, M., et al. "Professional Mediators' Judgments of Mediation Tactics: Multidimensional Scaling and Cluster Analysis." *Journal of Applied Psychology*, **76**, no. 3 (1991), 465–473.
22. Naik, D. N., and R. Khattree. "Revisiting Olympic Track Records: Some Practical Considerations in the Principal Component Analysis." *The American Statistician*, **50**, no. 2 (1996), 140–144.
23. Nason, G. "Three-dimensional Projection Pursuit." *Applied Statistics*, **44**, no. 4 (1995), 411–430.
24. Smith, M., and R. Taffler. "Improving the Communication Function of Published Accounting Statements." *Accounting and Business Research*, **14**, no. 54 (1984), 139–146.
25. Spenner, K. I. "From Generation to Generation: The Transmission of Occupation." Ph.D. dissertation, University of Wisconsin, 1977.
26. Tabakoff, B., et al. "Differences in Platelet Enzyme Activity between Alcoholics and Nonalcoholics." *New England Journal of Medicine*, **318**, no. 3 (1988), 134–139.
27. Timm, N. H. *Multivariate Analysis with Applications in Education and Psychology.* Monterey, CA: Brooks/Cole, 1975.
28. Trieschmann, J. S., and G. E. Pinches. "A Multivariate Model for Predicting Financially Distressed P-L Insurers." *Journal of Risk and Insurance*, **40**, no. 3 (1973), 327–338.
29. Tukey, J. W. *Exploratory Data Analysis.* Reading, MA: Addison-Wesley, 1977.
30. Wainer, H., and D. Thissen. "Graphical Data Analysis." *Annual Review of Psychology*, **32**, (1981), 191–241.
31. Wartzman, R. "Don't Wave a Red Flag at the IRS." *The Wall Street Journal* (February 24, 1993), C1, C15.
32. Weihs, C., and H. Schmidli. "OMEGA (On Line Multivariate Exploratory Graphical Analysis): Routine Searching for Structure." *Statistical Science*, **5**, no. 2 (1990), 175–226.

Chapter 2

MATRIX ALGEBRA AND RANDOM VECTORS

2.1 Introduction

We saw in Chapter 1 that multivariate data can be conveniently displayed as an array of numbers. In general, a rectangular array of numbers with, for instance, n rows and p columns is called a *matrix* of dimension $n \times p$. The study of multivariate methods is greatly facilitated by the use of matrix algebra.

The matrix algebra results presented in this chapter will enable us to concisely state statistical models. Moreover, the formal relations expressed in matrix terms are easily programmed on computers to allow the routine calculation of important statistical quantities.

We begin by introducing some very basic concepts that are essential to both our geometrical interpretations and algebraic explanations of subsequent statistical techniques. If you have not been previously exposed to the rudiments of matrix algebra, you may prefer to follow the brief refresher in the next section by the more detailed review provided in Supplement 2A.

2.2 Some Basics of Matrix and Vector Algebra

Vectors

An array $\mathbf{x}$ of n real numbers $x_1, x_2, \ldots, x_n$ is called a *vector*, and it is written as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{or} \quad \mathbf{x}' = [x_1, x_2, \ldots, x_n]$$

where the prime denotes the operation of *transposing* a column to a row.

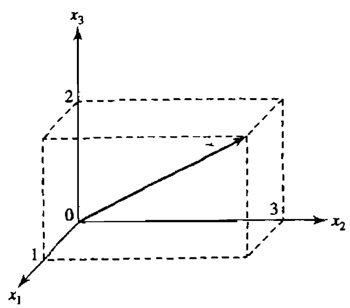

Figure 2.1 The vector $x' = [1, 3, 2]$.

A vector $\mathbf{x}$ can be represented geometrically as a directed line in n dimensions with component x_1 along the first axis, x_2 along the second axis, ..., and x_n along the nth axis. This is illustrated in Figure 2.1 for $n = 3$.

A vector can be *expanded* or *contracted* by multiplying it by a constant c. In particular, we define the vector $c\mathbf{x}$ as

$$c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$$

That is, $c\mathbf{x}$ is the vector obtained by multiplying each element of $\mathbf{x}$ by c. [See Figure 2.2(a).]

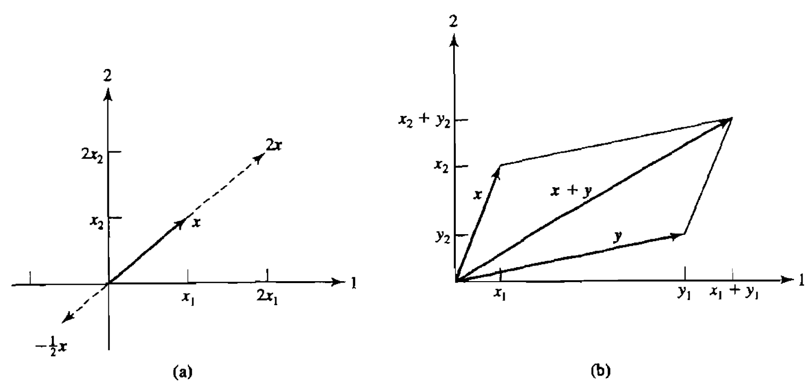

Figure 2.2 Scalar multiplication and vector addition.

Two vectors may be added. *Addition* of $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

so that $\mathbf{x} + \mathbf{y}$ is the vector with ith element $x_i + y_i$.

The sum of two vectors emanating from the origin is the diagonal of the parallelogram formed with the two original vectors as adjacent sides. This geometrical interpretation is illustrated in Figure 2.2(b).

A vector has both direction and length. In $n = 2$ dimensions, we consider the vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The length of $\mathbf{x}$, written $L_{\mathbf{x}}$, is defined to be

$$L_{\mathbf{x}} = \sqrt{x_1^2 + x_2^2}$$

Geometrically, the length of a vector in two dimensions can be viewed as the hypotenuse of a right triangle. This is demonstrated schematically in Figure 2.3.

The *length* of a vector $\mathbf{x}' = [x_1, x_2, \ldots, x_n]$, with n components, is defined by

$$L_{\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \tag{2-1}$$

Multiplication of a vector $\mathbf{x}$ by a scalar c changes the length. From Equation (2-1),

$$\begin{aligned} L_{c\mathbf{x}} &= \sqrt{c^2x_1^2 + c^2x_2^2 + \cdots + c^2x_n^2} \\ &= |c|\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = |c|L_{\mathbf{x}} \end{aligned}$$

Multiplication by c does not change the direction of the vector $\mathbf{x}$ if $c > 0$. However, a negative value of c creates a vector with a direction opposite that of $\mathbf{x}$. From

$$L_{c\mathbf{x}} = |c|L_{\mathbf{x}} \tag{2-2}$$

it is clear that $\mathbf{x}$ is expanded if $|c| > 1$ and contracted if $0 < |c| < 1$. [Recall Figure 2.2(a).] Choosing $c = L_{\mathbf{x}}^{-1}$, we obtain the *unit vector* $L_{\mathbf{x}}^{-1}\mathbf{x}$, which has length 1 and lies in the direction of $\mathbf{x}$.

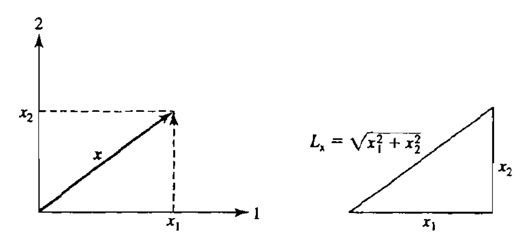

Figure 2.3
Length of $\mathbf{x} = \sqrt{x_1^2 + x_2^2}$.

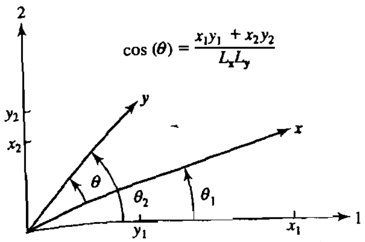

Figure 2.4 The angle θ between $\mathbf{x}' = [x_1, x_2]$ and $\mathbf{y}' = [y_1, y_2]$.

A second geometrical concept is *angle*. Consider two vectors in a plane and the angle θ between them, as in Figure 2.4. From the figure, θ can be represented as the difference between the angles θ_1 and θ_2 formed by the two vectors and the first coordinate axis. Since, by definition,

$$\cos(\theta_1) = \frac{x_1}{L_x} \qquad \cos(\theta_2) = \frac{y_1}{L_y}$$

$$\sin(\theta_1) = \frac{x_2}{L_x} \qquad \sin(\theta_2) = \frac{y_2}{L_y}$$

and

$$\cos(\theta) = \cos(\theta_2 - \theta_1) = \cos(\theta_2)\cos(\theta_1) + \sin(\theta_2)\sin(\theta_1)$$

the angle θ between the two vectors $\mathbf{x}' = [x_1, x_2]$ and $\mathbf{y}' = [y_1, y_2]$ is specified by

$$\cos(\theta) = \cos(\theta_2 - \theta_1) = \left(\frac{y_1}{L_y}\right)\left(\frac{x_1}{L_x}\right) + \left(\frac{y_2}{L_y}\right)\left(\frac{x_2}{L_x}\right) = \frac{x_1y_1 + x_2y_2}{L_xL_y} \qquad (2\text{-}3)$$

We find it convenient to introduce the *inner product* of two vectors. For $n = 2$ dimensions, the inner product of $\mathbf{x}$ and $\mathbf{y}$ is

$$\mathbf{x}'\mathbf{y} = x_1y_1 + x_2y_2$$

With this definition and Equation (2-3),

$$L_x = \sqrt{\mathbf{x}'\mathbf{x}} \qquad \cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{L_xL_y} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}}$$

Since $\cos(90°) = \cos(270°) = 0$ and $\cos(\theta) = 0$ only if $\mathbf{x}'\mathbf{y} = 0$, $\mathbf{x}$ and $\mathbf{y}$ are perpendicular when $\mathbf{x}'\mathbf{y} = 0$.

For an arbitrary number of dimensions n, we define the inner product of $\mathbf{x}$ and $\mathbf{y}$ as

$$\mathbf{x}'\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n \qquad (2\text{-}4)$$

The inner product is denoted by either $\mathbf{x}'\mathbf{y}$ or $\mathbf{y}'\mathbf{x}$.

Using the inner product, we have the natural extension of length and angle to vectors of n components:

$$L_x = \text{length of } \mathbf{x} = \sqrt{\mathbf{x}'\mathbf{x}} \qquad (2\text{-}5)$$

$$\cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{L_xL_y} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}} \qquad (2\text{-}6)$$

Since, again, $\cos(\theta) = 0$ only if $\mathbf{x}'\mathbf{y} = 0$, we say that $\mathbf{x}$ and $\mathbf{y}$ are *perpendicular* when $\mathbf{x}'\mathbf{y} = 0$.

Example 2.1 (Calculating lengths of vectors and the angle between them) Given the vectors $\mathbf{x}' = [1, 3, 2]$ and $\mathbf{y}' = [-2, 1, -1]$, find $3\mathbf{x}$ and $\mathbf{x} + \mathbf{y}$. Next, determine the length of $\mathbf{x}$, the length of $\mathbf{y}$, and the angle between $\mathbf{x}$ and $\mathbf{y}$. Also, check that the length of $3\mathbf{x}$ is three times the length of $\mathbf{x}$.

First,

$$3\mathbf{x} = 3\begin{bmatrix}1\\3\\2\end{bmatrix} = \begin{bmatrix}3\\9\\6\end{bmatrix}$$

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix}1\\3\\2\end{bmatrix} + \begin{bmatrix}-2\\1\\-1\end{bmatrix} = \begin{bmatrix}1-2\\3+1\\2-1\end{bmatrix} = \begin{bmatrix}-1\\4\\1\end{bmatrix}$$

Next, $\mathbf{x}'\mathbf{x} = 1^2 + 3^2 + 2^2 = 14$, $\mathbf{y}'\mathbf{y} = (-2)^2 + 1^2 + (-1)^2 = 6$, and $\mathbf{x}'\mathbf{y} = 1(-2) + 3(1) + 2(-1) = -1$. Therefore,

$$L_x = \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{14} = 3.742 \qquad L_y = \sqrt{\mathbf{y}'\mathbf{y}} = \sqrt{6} = 2.449$$

and

$$\cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{L_xL_y} = \frac{-1}{3.742 \times 2.449} = -.109$$

so $\theta = 96.3°$. Finally,

$$L_{3x} = \sqrt{3^2 + 9^2 + 6^2} = \sqrt{126} \quad \text{and} \quad 3L_x = 3\sqrt{14} = \sqrt{126}$$

showing $L_{3x} = 3L_x$. ■

A pair of vectors $\mathbf{x}$ and $\mathbf{y}$ of the same dimension is said to be *linearly dependent* if there exist constants c_1 and c_2, both not zero, such that

$$c_1\mathbf{x} + c_2\mathbf{y} = \mathbf{0}$$

A set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ is said to be *linearly dependent* if there exist constants $c_1, c_2, \ldots, c_k$, not all zero, such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0} \qquad (2\text{-}7)$$

Linear dependence implies that at least one vector in the set can be written as a linear combination of the other vectors. Vectors of the same dimension that are not linearly dependent are said to be *linearly independent*.

Example 2.2 (Identifying linearly independent vectors) Consider the set of vectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Setting

$$c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + c_3 \mathbf{x}_3 = \mathbf{0}$$

implies that

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ 2c_1 \qquad - 2c_3 &= 0 \\ c_1 - c_2 + c_3 &= 0 \end{aligned}$$

with the unique solution $c_1 = c_2 = c_3 = 0$. As we cannot find three constants c_1, c_2, and c_3, *not all zero*, such that $c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + c_3 \mathbf{x}_3 = \mathbf{0}$, the vectors $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$ are *linearly independent.* ■

The *projection* (or shadow) of a vector $\mathbf{x}$ on a vector $\mathbf{y}$ is

$$\text{Projection of } \mathbf{x} \text{ on } \mathbf{y} = \frac{(\mathbf{x}'\mathbf{y})}{\mathbf{y}'\mathbf{y}}\mathbf{y} = \frac{(\mathbf{x}'\mathbf{y})}{L_\mathbf{y}}\frac{1}{L_\mathbf{y}}\mathbf{y} \tag{2-8}$$

where the vector $L_\mathbf{y}^{-1}\mathbf{y}$ has unit length. The *length* of the projection is

$$\text{Length of projection} = \frac{|\mathbf{x}'\mathbf{y}|}{L_\mathbf{y}} = L_\mathbf{x}\left|\frac{\mathbf{x}'\mathbf{y}}{L_\mathbf{x}L_\mathbf{y}}\right| = L_\mathbf{x}|\cos(\theta)| \tag{2-9}$$

where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. (See Figure 2.5.)

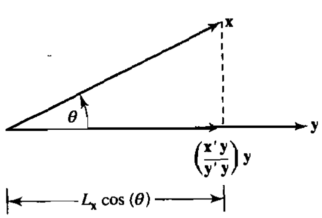

Figure 2.5 The projection of **x** on **y**.

Matrices

A *matrix* is any rectangular array of real numbers. We denote an arbitrary array of n rows and p columns by

$$\underset{(n\times p)}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

Many of the vector concepts just introduced have direct generalizations to matrices.

The *transpose* operation $\mathbf{A}'$ of a matrix changes the columns into rows, so that the first column of $\mathbf{A}$ becomes the first row of $\mathbf{A}'$, the second column becomes the second row, and so forth.

Example 2.3 (The transpose of a matrix) If

$$\underset{(2\times 3)}{\mathbf{A}} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 5 & 4 \end{bmatrix}$$

then

$$\underset{(3\times 2)}{\mathbf{A}'} = \begin{bmatrix} 3 & 1 \\ -1 & 5 \\ 2 & 4 \end{bmatrix}$$

■

A matrix may also be multiplied by a constant c. The product $c\mathbf{A}$ is the matrix that results from multiplying each element of $\mathbf{A}$ by c. Thus

$$\underset{(n\times p)}{c\mathbf{A}} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1p} \\ ca_{21} & ca_{22} & \cdots & ca_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{np} \end{bmatrix}$$

Two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimensions can be added. The sum $\mathbf{A} + \mathbf{B}$ has (i, j)th entry $a_{ij} + b_{ij}$.

Example 2.4 (The sum of two matrices and multiplication of a matrix by a constant) If

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

then

$$\underset{(2\times 3)}{4\mathbf{A}} = \begin{bmatrix} 0 & 12 & 4 \\ 4 & -4 & 4 \end{bmatrix} \quad \text{and}$$

$$\underset{(2\times 3)}{\mathbf{A}} + \underset{(2\times 3)}{\mathbf{B}} = \begin{bmatrix} 0+1 & 3-2 & 1-3 \\ 1+2 & -1+5 & 1+1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -2 \\ 3 & 4 & 2 \end{bmatrix}$$

■

It is also possible to define the multiplication of two matrices if the dimensions of the matrices conform in the following manner: When $\mathbf{A}$ is $(n \times k)$ and $\mathbf{B}$ is $(k \times p)$, so that the number of elements in a row of $\mathbf{A}$ is the same as the number of elements in a column of $\mathbf{B}$, we can form the matrix product $\mathbf{AB}$. An element of the new matrix $\mathbf{AB}$ is formed by taking the inner product of each row of $\mathbf{A}$ with each column of $\mathbf{B}$.

The *matrix product* **AB** is

$$\underset{(n\times k)(k\times p)}{\mathbf{A}\ \mathbf{B}} = \text{the } (n \times p) \text{ matrix whose entry in the } i\text{th row and } j\text{th column is the inner product of the } i\text{th row of } \mathbf{A} \text{ and the } j\text{th column of } \mathbf{B}$$

or

$$(i,j) \text{ entry of } \mathbf{AB} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} = \sum_{\ell=1}^{k} a_{i\ell}b_{\ell j} \qquad (2\text{-}10)$$

When $k = 4$, we have four products to add for each entry in the matrix **AB**. Thus,

$$\underset{(n\times 4)(4\times p)}{\mathbf{A}\ \mathbf{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2p} \\ b_{31} & \cdots & b_{3j} & \cdots & b_{3p} \\ b_{41} & \cdots & b_{4j} & \cdots & b_{4p} \end{bmatrix}$$

Column j

$$= \text{Row } i \begin{bmatrix} & \vdots & \\ \cdots & (a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + a_{i4}b_{4j}) & \cdots \\ & \vdots & \end{bmatrix}$$

Example 2.5 (Matrix multiplication) If

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 5 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -2 \\ 7 \\ 9 \end{bmatrix}, \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}$$

then

$$\underset{(2\times 3)(3\times 1)}{\mathbf{A}\ \mathbf{B}} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 5 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 7 \\ 9 \end{bmatrix} = \begin{bmatrix} 3(-2) + (-1)(7) + 2(9) \\ 1(-2) + 5(7) + 4(9) \end{bmatrix}$$

$$= \underset{(2\times 1)}{\begin{bmatrix} 5 \\ 69 \end{bmatrix}}$$

and

$$\underset{(2\times 2)(2\times 3)}{\mathbf{C}\ \mathbf{A}} = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 2 \\ 1 & 5 & 4 \end{bmatrix}$$

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

$$= \underset{(2\times 3)}{\begin{bmatrix} 6 & -2 & 4 \\ 2 & -6 & -2 \end{bmatrix}}$$

■

When a matrix **B** consists of a single column, it is customary to use the lowercase **b** vector notation.

Example 2.6 (Some typical products and their dimensions) Let

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 7 \\ -3 \\ 6 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} 5 \\ 8 \\ -4 \end{bmatrix} \quad \mathbf{d} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$$

Then **Ab**, **bc′**, **b′c**, and **d′Ab** are typical products.

$$\mathbf{A}\,\mathbf{b} = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \end{bmatrix} \begin{bmatrix} 7 \\ -3 \\ 6 \end{bmatrix} = \begin{bmatrix} 31 \\ -4 \end{bmatrix}$$

The product **A b** is a vector with dimension equal to the number of rows of **A**.

$$\mathbf{b}'\mathbf{c} = [7 \quad -3 \quad 6] \begin{bmatrix} 5 \\ 8 \\ -4 \end{bmatrix} = [-13]$$

The product **b′c** is a 1×1 vector or a single number, here -13.

$$\mathbf{b}\,\mathbf{c}' = \begin{bmatrix} 7 \\ -3 \\ 6 \end{bmatrix} [5 \quad 8 \quad -4] = \begin{bmatrix} 35 & 56 & -28 \\ -15 & -24 & 12 \\ 30 & 48 & -24 \end{bmatrix}$$

The product **b c′** is a matrix whose row dimension equals the dimension of **b** and whose column dimension equals that of **c**. This product is unlike **b′c**, which is a single number.

$$\mathbf{d}'\mathbf{A}\,\mathbf{b} = [2 \quad 9] \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \end{bmatrix} \begin{bmatrix} 7 \\ -3 \\ 6 \end{bmatrix} = [26]$$

The product **d′A b** is a 1×1 vector or a single number, here 26. ■

Square matrices will be of special importance in our development of statistical methods. A square matrix is said to be *symmetric* if $\mathbf{A} = \mathbf{A}'$ or $a_{ij} = a_{ji}$ for all i and j.

Example 2.7 (A symmetric matrix) The matrix

$$\begin{bmatrix} 3 & 5 \\ 5 & -2 \end{bmatrix}$$

is symmetric; the matrix

$$\begin{bmatrix} 3 & 6 \\ 4 & -2 \end{bmatrix}$$

is not symmetric. ■

When two square matrices **A** and **B** are of the same dimension, both products **AB** and **BA** are defined, although they need not be equal. (See Supplement 2A.) If we let **I** denote the square matrix with ones on the diagonal and zeros elsewhere, it follows from the definition of matrix multiplication that the (i, j)th entry of **AI** is $a_{i1} \times 0 + \cdots + a_{i,j-1} \times 0 + a_{ij} \times 1 + a_{i,j+1} \times 0 + \cdots + a_{ik} \times 0 = a_{ij}$, so $\mathbf{AI} = \mathbf{A}$. Similarly, $\mathbf{IA} = \mathbf{A}$, so

$$\underset{(k\times k)(k\times k)}{\mathbf{I}\ \mathbf{A}} = \underset{(k\times k)(k\times k)}{\mathbf{A}\ \mathbf{I}} = \underset{(k\times k)}{\mathbf{A}} \quad \text{for any } \underset{(k\times k)}{\mathbf{A}} \tag{2-11}$$

The matrix **I** acts like 1 in ordinary multiplication $(1 \cdot a = a \cdot 1 = a)$, so it is called the *identity* matrix.

The fundamental scalar relation about the existence of an inverse number a^{-1} such that $a^{-1}a = aa^{-1} = 1$ if $a \neq 0$ has the following matrix algebra extension: If there exists a matrix **B** such that

$$\underset{(k\times k)(k\times k)}{\mathbf{B}\ \mathbf{A}} = \underset{(k\times k)(k\times k)}{\mathbf{A}\ \mathbf{B}} = \underset{(k\times k)}{\mathbf{I}}$$

then **B** is called the *inverse* of **A** and is denoted by $\mathbf{A}^{-1}$.

The technical condition that an inverse exists is that the k columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ of **A** are linearly independent. That is, the existence of $\mathbf{A}^{-1}$ is equivalent to

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_k\mathbf{a}_k = \mathbf{0} \quad \text{only if} \quad c_1 = \cdots = c_k = 0 \tag{2-12}$$

(See Result 2A.9 in Supplement 2A.)

Example 2.8 (The existence of a matrix inverse) For

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$$

you may verify that

$$\begin{bmatrix} -.2 & .4 \\ .8 & -.6 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} (-.2)3 + (.4)4 & (-.2)2 + (.4)1 \\ (.8)3 + (-.6)4 & (.8)2 + (-.6)1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so

$$\begin{bmatrix} -.2 & .4 \\ .8 & -.6 \end{bmatrix}$$

is $\mathbf{A}^{-1}$. We note that

$$c_1\begin{bmatrix} 3 \\ 4 \end{bmatrix} + c_2\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

implies that $c_1 = c_2 = 0$, so the columns of **A** are linearly independent. This confirms the condition stated in (2-12). ■

A method for computing an inverse, when one exists, is given in Supplement 2A. The routine, but lengthy, calculations are usually relegated to a computer, especially when the dimension is greater than three. Even so, you must be forewarned that if the column sum in (2-12) is *nearly* **0** for some constants $c_1, \ldots, c_k$, then the computer may produce incorrect inverses due to extreme errors in rounding. It is always good to check the products $\mathbf{AA}^{-1}$ and $\mathbf{A}^{-1}\mathbf{A}$ for equality with **I** when $\mathbf{A}^{-1}$ is produced by a computer package. (See Exercise 2.10.)

Diagonal matrices have inverses that are easy to compute. For example,

$$\begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 & 0 \\ 0 & 0 & a_{33} & 0 & 0 \\ 0 & 0 & 0 & a_{44} & 0 \\ 0 & 0 & 0 & 0 & a_{55} \end{bmatrix} \text{ has inverse } \begin{bmatrix} \frac{1}{a_{11}} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{a_{22}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{a_{33}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{a_{44}} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{a_{55}} \end{bmatrix}$$

if all the $a_{ii} \neq 0$.

Another special class of square matrices with which we shall become familiar are the *orthogonal* matrices, characterized by

$$\mathbf{QQ}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I} \quad \text{or} \quad \mathbf{Q}' = \mathbf{Q}^{-1} \tag{2-13}$$

The name derives from the property that if **Q** has ith row $\mathbf{q}_i'$, then $\mathbf{QQ}' = \mathbf{I}$ implies that $\mathbf{q}_i'\mathbf{q}_i = 1$ and $\mathbf{q}_i'\mathbf{q}_j = 0$ for $i \neq j$, so the rows have unit length and are mutually perpendicular (orthogonal). According to the condition $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$, the columns have the same property.

We conclude our brief introduction to the elements of matrix algebra by introducing a concept fundamental to multivariate statistical analysis. A square matrix **A** is said to have an *eigenvalue* λ, with corresponding *eigenvector* $\mathbf{x} \neq \mathbf{0}$, if

$$\mathbf{Ax} = \lambda\mathbf{x} \tag{2-14}$$

Ordinarily, we normalize $\mathbf{x}$ so that it has length unity; that is, $1 = \mathbf{x}'\mathbf{x}$. It is convenient to denote normalized eigenvectors by $\mathbf{e}$, and we do so in what follows. Sparing you the details of the derivation (see [1]), we state the following basic result:

Let $\mathbf{A}$ be a $k \times k$ square symmetric matrix. Then $\mathbf{A}$ has k pairs of eigenvalues and eigenvectors namely,

$$\lambda_1, \mathbf{e}_1 \quad \lambda_2, \mathbf{e}_2 \quad \cdots \quad \lambda_k, \mathbf{e}_k \tag{2-15}$$

The eigenvectors can be chosen to satisfy $1 = \mathbf{e}_1'\mathbf{e}_1 = \cdots = \mathbf{e}_k'\mathbf{e}_k$ and be mutually perpendicular. The eigenvectors are unique unless two or more eigenvalues are equal.

Example 2.9 (Verifying eigenvalues and eigenvectors) Let

$$\mathbf{A} = \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix}$$

Then, since

$$\begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = 6 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$\lambda_1 = 6$ is an eigenvalue, and

$$\mathbf{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

is its corresponding normalized eigenvector. You may wish to show that a second eigenvalue–eigenvector pair is $\lambda_2 = -4$, $\mathbf{e}_2' = [1/\sqrt{2}, 1/\sqrt{2}]$. ■

A method for calculating the λ's and $\mathbf{e}$'s is described in Supplement 2A. It is instructive to do a few sample calculations to understand the technique. We usually rely on a computer when the dimension of the square matrix is greater than two or three.

2.3 Positive Definite Matrices

The study of the variation and interrelationships in multivariate data is often based upon distances and the assumption that the data are multivariate normally distributed. Squared distances (see Chapter 1) and the multivariate normal density can be expressed in terms of matrix products called *quadratic forms* (see Chapter 4). Consequently, it should not be surprising that quadratic forms play a central role in

multivariate analysis. In this section, we consider quadratic forms that are always nonnegative and the associated *positive definite* matrices.

Results involving quadratic forms and symmetric matrices are, in many cases, a direct consequence of an expansion for symmetric matrices known as the *spectral decomposition*. The spectral decomposition of a $k \times k$ symmetric matrix $\mathbf{A}$ is given by[1]

$$\underset{(k\times k)}{\mathbf{A}} = \lambda_1 \underset{(k\times1)}{\mathbf{e}_1} \underset{(1\times k)}{\mathbf{e}_1'} + \lambda_2 \underset{(k\times1)}{\mathbf{e}_2} \underset{(1\times k)}{\mathbf{e}_2'} + \cdots + \lambda_k \underset{(k\times1)}{\mathbf{e}_k} \underset{(1\times k)}{\mathbf{e}_k'} \tag{2-16}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the eigenvalues of $\mathbf{A}$ and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ are the associated normalized eigenvectors. (See also Result 2A.14 in Supplement 2A). Thus, $\mathbf{e}_i'\mathbf{e}_i = 1$ for $i = 1, 2, \ldots, k$, and $\mathbf{e}_i'\mathbf{e}_j = 0$ for $i \neq j$.

Example 2.10 (The spectral decomposition of a matrix) Consider the symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$

The eigenvalues obtained from the characteristic equation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ are $\lambda_1 = 9$, $\lambda_2 = 9$, and $\lambda_3 = 18$ (Definition 2A.30). The corresponding eigenvectors $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$ are the (normalized) solutions of the equations $\mathbf{A}\mathbf{e}_i = \lambda_i\mathbf{e}_i$ for $i = 1, 2, 3$. Thus, $\mathbf{A}\mathbf{e}_1 = \lambda\mathbf{e}_1$ gives

$$\begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix} \begin{bmatrix} e_{11} \\ e_{21} \\ e_{31} \end{bmatrix} = 9 \begin{bmatrix} e_{11} \\ e_{21} \\ e_{31} \end{bmatrix}$$

or

$$\begin{aligned} 13e_{11} - 4e_{21} + 2e_{31} &= 9e_{11} \\ -4e_{11} + 13e_{21} - 2e_{31} &= 9e_{21} \\ 2e_{11} - 2e_{21} + 10e_{31} &= 9e_{31} \end{aligned}$$

Moving the terms on the right of the equals sign to the left yields three homogeneous equations in three unknowns, but two of the equations are redundant. Selecting one of the equations and arbitrarily setting $e_{11} = 1$ and $e_{21} = 1$, we find that $e_{31} = 0$. Consequently, the normalized eigenvector is $\mathbf{e}_1' = [1/\sqrt{1^2 + 1^2 + 0^2}, 1/\sqrt{1^2 + 1^2 + 0^2}, 0/\sqrt{1^2 + 1^2 + 0^2}] = [1/\sqrt{2}, 1/\sqrt{2}, 0]$, since the sum of the squares of its elements is unity. You may verify that $\mathbf{e}_2' = [1/\sqrt{18}, -1/\sqrt{18}, -4/\sqrt{18}]$ is also an eigenvector for $9 = \lambda_2$, and $\mathbf{e}_3' = [2/3, -2/3, 1/3]$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_3 = 18$. Moreover, $\mathbf{e}_i'\mathbf{e}_j = 0$ for $i \neq j$.

[1] A proof of Equation (2-16) is beyond the scope of this book. The interested reader will find a proof in [6], Chapter 8.