CHAPTER 2

Vector and Matrix Operations for Multivariate Analysis

2.1 INTRODUCTION

Facility in the arithmetic of vectors and matrices, just like skill in applying ordinary arithmetic to daily affairs, is essential in multivariate analysis. In this chapter our purpose is to review the fundamentals of vector and matrix operations and the concept of the determinant of a matrix. The emphasis here is on defining vector and matrix operations and illustrating the mechanics of their application.

We begin the chapter with a description of vectors as ordered *n*-tuples of numbers that are subject to certain manipulative rules. Selected arithmetic operations on vectors are defined and illustrated numerically. A number of special-purpose vectors, such as the null vector, unit vector, and sign vector, are also described.

Matrices are then introduced and discussed from the same kind of viewpoint. Also, we describe various kinds of special matrices, such as symmetric, diagonal, scalar, and identity matrices and illustrate their application via small numerical examples.

The determinant of a matrix plays an important role in more advanced topics, such as matrix inversion, rank, and quadratic forms, that are introduced in later chapters. For this reason it seems appropriate to discuss determinants and some of their numerical properties at an early stage, and we do so in this section of the chapter.

We conclude the chapter with a discussion of certain matrices of particular interest to multivariate analysis, namely mean-corrected sums of squares and cross product (SSCP) matrices, covariance matrices, and correlation matrices. Computation of these major types of statistical matrices is carried out as a demonstration of how concise the matrix formulation of various arithmetic operations can be.

A word on notation: boldface lowercase letters, **a**, **b**, **c**, etc., will be used to denote vectors and boldface capitals, **A**, **B**, **C**, etc., will be used for matrices. The determinant of a matrix **A** will be expressed as $|\mathbf{A}|$. A prime, for example, $\mathbf{a}'$ or $\mathbf{A}'$, will denote the transpose of a vector or matrix, respectively. (The concept of transpose is taken up later in the chapter.)

The material of this chapter is presented rather crisply since our purpose here is to provide only the mechanics of vector and matrix operations before introducing the more conceptually oriented material of later chapters. However, sufficient numerical examples are presented to illustrate the computational aspects in some detail.

2.2 VECTOR REPRESENTATION

Multivariate analysis makes liberal use of vector concepts from linear algebra. Vectors can be defined in four major ways: (a) as strictly abstract entities on which certain relations and operations are specified, (b) as directed line segments in a geometric space, (c) as coordinate representations of points in a geometric space, or (d) as ordered n-tuples of numbers. We adopt the last viewpoint in this chapter in order to demonstrate the kinds of operations that can be performed on vectors. The geometric representations of (b) and (c), which can be illustrated graphically if two or three dimensions are involved, are discussed in Chapters 3–5.

2.3 BASIC DEFINITIONS AND OPERATIONS ON VECTORS

A vector a *of order n × 1 is an ordered set of n real*[1] *numbers (called scalars), which we can write as*

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

The a's denote real numbers and are called components, elements, or entries of a. The form above is called a *column* vector and consists of n rows and 1 column of elements (from which the designation $n \times 1$ derives). Alternatively we can write a vector a′ of order $1 \times n$ as

$$\mathbf{a}' = (a_1, a_2, \ldots, a_n)$$

and call this a *row* vector, consisting of 1 row and n columns of elements.

We shall use the notation a to denote a column vector and the notation a′, which is called the *transpose* of a, to denote a row vector. By vector transpose, generally, is meant that a column vector of order n by 1 becomes a row vector, involving the same ordered set of entries, but now of order 1 by n. Similarly, the transpose of a row vector of 1 by n is a column vector, involving the same ordered set of entries, but now of order n by 1.

Examples of column vectors are

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix}; \quad \begin{bmatrix} 4 \\ 2.6 \\ 5 \\ 0 \end{bmatrix}; \quad \begin{bmatrix} 2.718 \\ 5 \\ 1 \end{bmatrix}; \quad \begin{bmatrix} 18 \\ 21 \\ 0 \\ 0 \end{bmatrix}; \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Examples of row vectors are

$$(3, 1); \quad (18, 42, 6); \quad (\sqrt{\pi}, 13, 0, 5.2); \quad (0, 0, 2, 7); \quad \mathbf{t}' = (t_1, t_2)$$

[1] Throughout the book we shall always assume that the scalars are drawn from the set of real (as opposed to complex) numbers.

We can transpose the 3 x 1 column vector

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

to get the 1 x 3 row vector

$$\mathbf{a}' = (1, 2, 3)$$

Similarly, we can then find the transpose of the 1 x 3 row vector $\mathbf{a}' = (1, 2, 3)$ as follows:

$$(\mathbf{a}')' = \mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

and note that we are back where we started, that is, where a is a 3 x 1 column vector.

2.3.1 Null, Unit, Sign, and Zero–One Vectors

If all components of a vector are zero, we shall call this a *null* or zero vector, denoted as **0**. This should not be confused with the scalar 0. If all components of a vector are 1, this type of vector is called a *unit* vector, denoted as **1**. If the components consist of either 1's or −1's (with at least one of each type present), this is called a *sign* vector. If the components consist of either 1's or 0's (with at least one of each type present), this is called a zero–one vector. To illustrate:

	Column vectors	*Row vectors*
Null vectors	$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$	(0, 0, 0, 0)
Unit vectors	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	(1, 1, 1)
Sign vectors	$\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$	(1, −1)
Zero–one vectors	$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$	(0, 0, 1, 1, 0)

As will be shown in subsequent chapters, the zero vector frequently plays a role that is analogous to the scalar 0 in ordinary arithmetic. The unit vector is useful in certain kinds of summations, as is illustrated in Section 2.8. Sign and zero–one vectors are also useful in various kinds of operations involving either algebraic sums or the isolation of rows, columns, or elements of an array of numbers.

2.3.2 Vector Equality

Two vectors of the same order (either both $n \times 1$ or both $1 \times n$) are equal if they are equal component by component. Let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Then

$$\boxed{\begin{gathered} \mathbf{a} = \mathbf{b} \\ \text{if and only if} \\ a_i = b_i \qquad (i = 1, 2, \ldots, n) \end{gathered}}$$

For example,

$$\mathbf{a} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} = \mathbf{b} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$$

But

$$\mathbf{a} \neq \mathbf{c} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}; \qquad \mathbf{a} \neq \mathbf{d} = \begin{bmatrix} 3 \\ 0 \\ 4 \\ 9 \end{bmatrix}.$$

$$\mathbf{a} \neq \mathbf{e} = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}; \qquad \mathbf{a} \neq \mathbf{a}' = (3, 0, 4)$$

In the last case $\mathbf{a}$ and $\mathbf{a}'$ are not of the same order, since the first is a column vector and the second is a row vector.

Throughout most of this chapter, we shall present definitions in terms of column vectors, although our remarks will also hold true for row vectors.[2] Moreover, in discussing various operations on vectors, it will be assumed, unless otherwise specified, that the vectors are of common order—either all are $n \times 1$ or all are $1 \times n$.

[2] When row vectors are employed as numerical examples, emphasis is primarily on conserving space. The reader should remember that we could just as appropriately describe the operations in terms of column vectors.

2.3.3 Vector Addition and Subtraction

Two or more vectors of the same order can be added by adding correspondent components. That is

$$\mathbf{a}+\mathbf{b}=\begin{bmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{bmatrix}$$

Examples are

$$\mathbf{a}=\begin{bmatrix} 1 \\ 2 \\ 14 \end{bmatrix}; \quad \mathbf{b}=\begin{bmatrix} 3 \\ 0 \\ 9 \end{bmatrix}; \quad \mathbf{a}+\mathbf{b}=\begin{bmatrix} 4 \\ 2 \\ 23 \end{bmatrix}$$

$$\mathbf{c}'=(1,3); \quad \mathbf{d}'=(4,13); \quad \mathbf{c}'+\mathbf{d}'=(5,16)$$

But we cannot add

$$\mathbf{e}=\begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix} \text{ to } \mathbf{f}=\begin{bmatrix} 2 \\ 1 \end{bmatrix}; \quad \text{or} \quad \mathbf{g}=\begin{bmatrix} 5 \\ 1 \\ 2 \\ 7 \end{bmatrix} \text{ to } \mathbf{h}'=(6,3,0,2)$$

since in each case the order differs.

The difference between two vectors **a** *and* **b**, *of the same order, is defined to be that vector,* **a** – **b**, *which, when added to* **b**, *yields the vector* **a**. Again, subtraction is performed componentwise.[3] That is,

$$\mathbf{a}-\mathbf{b}=\begin{bmatrix} a_1-b_1 \\ a_2-b_2 \\ \vdots \\ a_n-b_n \end{bmatrix}$$

Examples are

$$\mathbf{a}=\begin{bmatrix} 1 \\ 2 \\ 14 \end{bmatrix}; \quad \mathbf{b}=\begin{bmatrix} 3 \\ 0 \\ 9 \end{bmatrix}; \quad \mathbf{a}-\mathbf{b}=\begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$$

$$\mathbf{c}'=(1,3); \quad \mathbf{d}'=(4,13); \quad \mathbf{c}'-\mathbf{d}'=(-3,-10)$$

[3] A more rigorous presentation would first define multiplication of a vector by a scalar (specifically multiplication by –1), followed by vector addition. Here, however, we follow the more natural presentation order of traditional arithmetic in which subtraction follows addition.

But we cannot subtract

$$\mathbf{e} = \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix} \quad \text{from} \quad \mathbf{f} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}; \qquad \text{or} \qquad \mathbf{g} = \begin{bmatrix} 5 \\ 1 \\ 2 \\ 7 \end{bmatrix} \quad \text{from} \quad \mathbf{h}' = (6, 3, 0, 2)$$

since in each case the order differs.

The operation of vector *addition*—of either column or row vectors—possesses the following properties:

1. The sum of two vectors **a** and **b** is a unique third vector **c**.

$$\text{Let} \quad \mathbf{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$\text{Then} \quad \mathbf{a} + \mathbf{b} = \mathbf{c} = \begin{bmatrix} 3 \\ 7 \end{bmatrix} \quad \text{is unique.}$$

2. Vector addition is commutative.

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$$

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

3. Vector addition is associative.

$(\mathbf{a} + \mathbf{b}) + \mathbf{d} = \mathbf{a} + (\mathbf{b} + \mathbf{d})$

$$\text{Let} \quad \mathbf{d} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \qquad \text{Then} \quad \underset{\mathbf{a}+\mathbf{b}}{\begin{bmatrix} 3 \\ 7 \end{bmatrix}} + \underset{\mathbf{d}}{\begin{bmatrix} 3 \\ 5 \end{bmatrix}} = \underset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} + \underset{\mathbf{b}+\mathbf{d}}{\begin{bmatrix} 5 \\ 9 \end{bmatrix}} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

4. There exists a null or zero vector **0** having the property $\mathbf{a} + \mathbf{0} = \mathbf{a}$ for any vector **a**.

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \mathbf{a}$$

5. Each vector **a** has a counterpart negative vector $-\mathbf{a}$ so that $\mathbf{a} + -\mathbf{a} = \mathbf{0}$.

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} -1 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

We shall have occasion to refer to one or more of these properties quite frequently in subsequent dicussions.

2.3.4 Scalar Multiplication of a Vector

Assume we have some real number k. As pointed out earlier, this is called a scalar in vector algebra. *Scalar multiplication of a vector involves multiplying each component of the vector by the scalar.*

$$k\mathbf{a} = k\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{bmatrix}$$

To illustrate the scalar multiplication of vectors, assume

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \text{and} \quad k = 3$$

Then

$$k\mathbf{a} = 3\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \times 1 \\ 3 \times 2 \\ 3 \times 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$$

Next, let $\mathbf{b}' = (4, 5, 6)$. Then

$$k\mathbf{b}' = 3(4, 5, 6) = (3 \times 4, 3 \times 5, 3 \times 6) = (12, 15, 18)$$

As in the case for vector addition, scalar multiplication of vectors exhibits a number of useful properties.[4]

1. If $\mathbf{a}$ is a vector and k is a scalar, the product $k\mathbf{a}$ is a uniquely defined vector.

$$\text{Let} \quad \mathbf{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \text{and} \quad k = 2$$

$$\text{Then} \quad 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \quad \text{is unique.}$$

2. Scalar multiplication is associative. For example, for two scalars k_1 and k_2, it is the case that $k_1(k_2\mathbf{a}) = (k_1 k_2)\mathbf{a}$.

$$\text{Let} \quad k_1 = 2 \quad \text{and} \quad k_2 = 3. \quad \text{Then} \quad 2\underbrace{\begin{bmatrix} 3 \\ 9 \end{bmatrix}}_{k_2\mathbf{a}} = 6\underbrace{\begin{bmatrix} 1 \\ 3 \end{bmatrix}}_{\mathbf{a}} = \begin{bmatrix} 6 \\ 18 \end{bmatrix}$$

[4] These properties and the properties listed in Section 2.3.3 collectively define a *vector space* for all vectors a, b, c, etc., and all scalars (real numbers) k_1, k_2, etc.

3. Scalar multiplication is distributive. For example, $k(\mathbf{a} + \mathbf{b}) = k\mathbf{a} + k\mathbf{b}$. Also, $(k_1 + k_2)\mathbf{a} = k_1\mathbf{a} + k_2\mathbf{a}$.

$$\text{Let} \quad \mathbf{b} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}. \quad \text{Then} \quad 2\overset{\mathbf{a}+\mathbf{b}}{\begin{bmatrix} 3 \\ 7 \end{bmatrix}} = 2\overset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} + 2\overset{\mathbf{b}}{\begin{bmatrix} 2 \\ 4 \end{bmatrix}} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

$$\text{Also} \quad 5\overset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} = 2\overset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} + 3\overset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} = \begin{bmatrix} 5 \\ 15 \end{bmatrix}$$

4. For any vector $\mathbf{a}$, we have the products $0 \cdot \mathbf{a} = \mathbf{0}$; $1\mathbf{a} = \mathbf{a}$ and $-1\mathbf{a} = -\mathbf{a}$. For example,

$$0\overset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \quad 1\overset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}; \quad -1\overset{\mathbf{a}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}} = \begin{bmatrix} -1 \\ -3 \end{bmatrix}$$

We can now consider an operation that generalizes both vector addition and scalar multiplication of vectors.

2.3.5 Linear Combinations of Vectors

Most of our comments about vector addition and scalar multiplication in Sections 2.3.3 and 2.3.4 can be succinctly summarized in terms of the concept of a linear combination of a set of vectors. Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ denote a set of m vectors (each of order $n \times 1$) and let $k_1, k_2, \ldots, k_m$ denote a set of m scalars. *A linear combination of a set of vectors is defined as*

$$\boxed{\mathbf{v}_1 = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_m\mathbf{a}_m}$$

If we take another (arbitrary) linear combination involving another set of m scalars $k_1^*, k_2^*, \ldots, k_m^*$, we have

$$\boxed{\mathbf{v}_2 = k_1^*\mathbf{a}_1 + k_2^*\mathbf{a}_2 + \cdots + k_m^*\mathbf{a}_m}$$

Next, suppose we add the two linear combinations. If so, it will be the case that the following properties hold:

1. $\mathbf{v}_1 + \mathbf{v}_2 = (k_1 + k_1^*)\mathbf{a}_1 + (k_2 + k_2^*)\mathbf{a}_2 + \cdots + (k_m + k_m^*)\mathbf{a}_m$

2. Moreover, if c denotes still another scalar, then $c\mathbf{v}_1 = (ck_1)\mathbf{a}_1 + (ck_2)\mathbf{a}_2 + \ldots + (ck_m)\mathbf{a}_m$ is also a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$.

What this means is that linear combinations of vectors can be added together (such as $\mathbf{v}_1 + \mathbf{v}_2$) or can be multiplied by a scalar (such as $c\mathbf{v}_1$), resulting in new vectors that bear simple relationships to the old.

To illustrate, let

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \qquad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}; \qquad \mathbf{a}_3 = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

$$k_1 = 2; \quad k_2 = 3; \quad k_3 = 1; \quad k_1{}^* = 0; \quad k_2{}^* = 4; \quad k_3{}^* = 5; \quad c = 2$$

Then, we can first write $\mathbf{v}_1$ and $\mathbf{v}_2$ as

$$\mathbf{v}_1 = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + k_3\mathbf{a}_3$$

$$= 2\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 17 \\ 14 \end{bmatrix}$$

$$\mathbf{v}_2 = k_1{}^*\mathbf{a}_1 + k_2{}^*\mathbf{a}_2 + k_3{}^*\mathbf{a}_3$$

$$= 0\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 4\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 32 \\ 18 \end{bmatrix}$$

The first property above can now be illustrated as

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 3 \\ 17 \\ 14 \end{bmatrix} + \begin{bmatrix} 5 \\ 32 \\ 18 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 7\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} + 6\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} + \begin{bmatrix} 0 \\ 21 \\ 14 \end{bmatrix} + \begin{bmatrix} 6 \\ 24 \\ 12 \end{bmatrix} = \begin{bmatrix} 8 \\ 49 \\ 32 \end{bmatrix}$$

The second property above can now be illustrated as

$$c\mathbf{v}_1 = \begin{bmatrix} 6 \\ 34 \\ 28 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 6\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix} + \begin{bmatrix} 0 \\ 18 \\ 12 \end{bmatrix} + \begin{bmatrix} 2 \\ 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 34 \\ 28 \end{bmatrix}$$

The concept of a linear combination of a set of vectors is one of the most important aspects of vector algebra. We shall return to this topic in the next chapter dealing with the geometric aspects of vectors.

2.3.6 The Scalar Product of Two Vectors

The last operation involving vectors to be discussed in Section 2.3 is that of the scalar product (sometimes called inner product, or dot product) of two vectors. As is well known, when we multiply two numbers (scalars) together, we obtain another element of the same kind, namely, a number that represents their product. However, in vector algebra, multiplication of two vectors need not lead to a vector. For example, one way of multiplying two vectors (of the same order of course) yields a *number* rather than a vector. This number is called their scalar product. To illustrate the scalar product of two vectors, consider the column vectors

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Their scalar product is defined as

$$\mathbf{a}'\mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_kb_k + \cdots + a_nb_n = \sum_{k=1}^{n} a_kb_k$$

Notice that the first vector is treated as a row vector and the second vector is treated as a column vector. However, either one can serve as the first (row) vector. To illustrate,

$$\mathbf{a} = \begin{bmatrix} 1 \\ 4 \\ 0 \\ 3 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ 7 \\ 4 \end{bmatrix}$$

Hence

$$\begin{aligned} \mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a} &= (1 \times 0) + (4 \times 2) + (0 \times 7) + (3 \times 4) \\ &= (0 \times 1) + (2 \times 4) + (7 \times 0) + (4 \times 3) \\ &= 20 \end{aligned}$$

We might now check to see if the associative and distributive laws are valid for scalar products. As it turns out, the associative law, illustrated by $(\mathbf{a}'\mathbf{b})\mathbf{c}$ is *not* valid because the scalar product of a scalar, that results from $\mathbf{a}'\mathbf{b}$, and a vector **c** has *not* been defined. *That is, the scalar product idea is limited to the product of a row and column vector.* Of course, we earlier defined the operation of multiplying a vector by a scalar, but this is not a scalar product.

However, the distributive laws for the scalar product, with respect to addition, *are* valid:

$$a'(b + c) = (a'b) + (a'c)$$

Also,

$$(a + b)'c = (a'c) + (b'c)$$

To illustrate the distributive laws, let,

$$a = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \quad b = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}; \quad c = \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}$$

Then

$$a'(b + c) = (a'b) + (a'c)$$

$$(1,2,3)\begin{bmatrix} 7 \\ 2 \\ 3 \end{bmatrix} = (1,2,3)\begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} + (1,2,3)\begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}$$

$$20 = 7 + 13$$

Also,

$$(a + b)'c = (a'c) + (b'c)$$

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

$$27 = 13 + 14$$

2.3.7 Some Special Cases of the Scalar Product

In the definition of scalar product given above, no requirement was made that **a** had to differ from **b**. That is, one can legitimately compute the scalar product of a vector with itself. To illustrate,

$$\text{If} \quad a = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \text{then} \quad a'a = 1 + 4 + 9 = 14$$

Notice, then, that one obtains a *sum of squares* if one takes the scalar product of a vector with itself. And $a'a > 0$ unless, of course, $a = \mathbf{0}$.

Consider now the unit vector $\mathbf{1}' = (1, 1, 1)$ and the vector

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Their scalar product is

$$\mathbf{1}'\mathbf{a} = 1 + 2 + 3 = 6$$

Thus, the scalar product of the unit vector and a given vector results in the *sum* of the entries in the given vector. If the vector is a sign vector, then the algebraic sum is taken. For example,

$$(-1, 1, -1) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = -2$$

Finally, consider the relationships $\mathbf{a}'(k\mathbf{b}) = (k\mathbf{a}')\mathbf{b} = k(\mathbf{a}'\mathbf{b})$, where k is a scalar. To demonstrate that these relations hold, let

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}; \quad k = 2$$

Then

$$\mathbf{a}'(k\mathbf{b}) = (k\mathbf{a}')\mathbf{b}$$

$$(1, 2, 3) \begin{bmatrix} 8 \\ 0 \\ 2 \end{bmatrix} = (2, 4, 6) \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} = 14$$

Also

$$\mathbf{a}'(k\mathbf{b}) = k(\mathbf{a}'\mathbf{b})$$

$$(1, 2, 3) \begin{bmatrix} 8 \\ 0 \\ 2 \end{bmatrix} = 2(1, 2, 3) \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} = 14$$

We can sum up this part of the discussion by recapitulating the following properties of scalar products, as illustrated above:

1. $\mathbf{a}'(\mathbf{b} + \mathbf{c}) = \mathbf{a}'\mathbf{b} + \mathbf{a}'\mathbf{c}$
2. $(\mathbf{a} + \mathbf{b})'\mathbf{c} = \mathbf{a}'\mathbf{c} + \mathbf{b}'\mathbf{c}$
3. $\mathbf{a}'(k\mathbf{b}) = (k\mathbf{a}')\mathbf{b} = k(\mathbf{a}'\mathbf{b})$

We shall have more to say about the utility of scalar product multiplication in the concluding section of the chapter, which deals with the computation of various matrices derived from statistical data.

Finally, it should be mentioned that two types of vector-by-vector multiplication are *not* defined in matrix algebra. That is,

1. a row vector cannot be multiplied by a row vector;
2. a column vector cannot be multiplied by a column vector.

The remaining case—that of multiplying a column vector by a row vector— is covered in our discussion of matrices since in this instance their product is a matrix, not a scalar.

2.3.8 Some More Examples

To help review the vector operations described in this section, consider the following:

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \qquad \mathbf{a}' = (1, 2, 3); \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}; \qquad \mathbf{b}' = (0, 2, 5)$$

$$k_1 = 2; \qquad k_2 = 3$$

We can now illustrate the following operations.

Transpose of Vector We recall that the transpose of the 3 x 1 column vector

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

is the 1 x 3 row vector, written as

$$\mathbf{a}' = (1, 2, 3)$$

Moreover, were we to take the transpose of $\mathbf{a}'$, we would have

$$(\mathbf{a}')' = \mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

That is, the transpose of a transpose equals the original vector. Similarly, we can find

$$((\mathbf{a}')')' = \mathbf{a}' = (1, 2, 3)$$

Addition and Subtraction The sum and difference of $\mathbf{a}'$ and $\mathbf{b}'$, respectively, are simply

$$\mathbf{a}' + \mathbf{b}' = (1 + 0, 2 + 2, 3 + 5) = (1, 4, 8)$$

$$\mathbf{a}' - \mathbf{b}' = (1 - 0, 2 - 2, 3 - 5) = (1, 0, -2)$$

Scalar Multiplication of a Vector Some illustrations of scalar multiplication of a vector are

$$k_1\mathbf{a}' = (2\times 1, 2\times 2, 2\times 3); \qquad 0\mathbf{a} = \begin{bmatrix} 0\times 1 \\ 0\times 2 \\ 0\times 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}$$
$$= (2, 4, 6)$$

$$k_1(k_2\mathbf{a}') = (k_1k_2)\mathbf{a}'; \qquad k_2\mathbf{b} = 3\begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 15 \end{bmatrix}$$
$$= 6(1, 2, 3)$$
$$= (6, 12, 18)$$

Linear Combinations of Vectors Illustrations of linear combinations of vectors are

$$\mathbf{v}_1 = k_1\mathbf{a} + k_2\mathbf{b} = 2\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 10 \\ 21 \end{bmatrix}; \qquad \mathbf{v}_2' = k_1\mathbf{b}' - k_2\mathbf{a}'$$
$$= 2(0, 2, 5) - 3(1, 2, 3)$$
$$= (0, 4, 10) - (3, 6, 9)$$
$$= (-3, -2, 1)$$

If $c = 4$, then

$$c\mathbf{v}_2' = ck_1\mathbf{b}' - ck_2\mathbf{a}'$$
$$= 8(0, 2, 5) - 12(1, 2, 3)$$
$$= (0, 16, 40) - (12, 24, 36)$$
$$= (-12, -8, 4)$$

The Scalar Product Some scalar products of interest are

$$\mathbf{a}'\mathbf{b} = (1, 2, 3)\begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}$$
$$= (1\times 0) + (2\times 2) + (3\times 5)$$
$$= 19$$

$$\mathbf{b}'\mathbf{b} = (0, 2, 5)\begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} = 29; \qquad \mathbf{0}'\mathbf{b} = (0, 0, 0)\begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} = 0$$

$$\mathbf{a}'\mathbf{1} = (1, 2, 3)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \qquad \mathbf{a}'(\mathbf{1} + \mathbf{b}) = (1, 2, 3)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (1, 2, 3)\begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} = 6 + 19$$
$$= 6 \qquad\qquad = 25$$

Additional examples appear at the end of the chapter.

2.4 MATRIX REPRESENTATION

As in our introduction to vector arithmetic, our purpose here is to describe various operations involving matrices as they relate to subsequent discussion of multivariate procedures. Again, we attempt no definitive treatment of the topic but, rather, select those aspects of particular relevance to subsequent chapters.

We first present a discussion of elementary relations and operations on matrices and then turn to a description of special types of matrices. More advanced topics in matrix algebra are relegated to subsequent chapters and the appendixes.

2.5 BASIC DEFINITIONS AND OPERATIONS ON MATRICES

A matrix **A** *of order m by n, and with general entry* (a_{ij}), *consists of a rectangular array of real numbers (scalars) arranged in m rows and n columns.*

$$\mathbf{A}_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} = (a_{ij})_{m\times n}.$$

For example, a 4 x 5 matrix would be explicitly written, in brackets, as

$$\mathbf{A}_{4\times 5} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix}$$

where $i = 1, 2, 3, 4$ and $j = 1, 2, 3, 4, 5$. As is the case for vectors, matrices will appear in boldfaced type, such as **A**, **B**, **C**, etc.

A matrix can exhibit any relation between m, the number of rows, and n, the number of columns. For example, if either $m > n$ or $n > m$, we have a rectangular matrix. (The former is often called a vertical matrix, while the latter is often called a horizontal matrix.)

If $m = n$, the matrix is called square. To illustrate,

$$\mathbf{B}_{3\times 3} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

The set of elements on the diagonal, from upper left to lower right,

$$\{b_{11}, b_{22}, b_{33}\}$$

is called the main or principal diagonal of the square matrix **B**. Square matrices occur quite frequently as derived matrices in multivariate analysis. For example, a correlation matrix, to be described later in the chapter, is a square matrix.

In either the rectangular or square matrix case, the order or "dimensionality" specifies the number of rows and columns of the matrix. Sometimes this order is made explicit in the form of subscripts:

$$\mathbf{C}_{2 \times 3}$$

In other cases, the order is inferred from context, such as

$$\mathbf{D} = \begin{bmatrix} 1 & 2 & 7 & 9 & 3 \\ 0 & 4 & 3 & 1 & 1 \end{bmatrix}$$

While we note that no subscript appears on **D**, it is clear that this matrix is of order 2 x 5.

If $m = 1$, the matrix is equivalent to a row vector. If $n = 1$, the matrix is equivalent to a column vector. If $m = n = 1$, we have a 1 x 1 matrix.[5]

A column vector, written as

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

can now be viewed as an m by 1 matrix and a row vector $\mathbf{a}' = (a_1, a_2, \ldots, a_n)$ can now be viewed as a 1 by n matrix.

As will be shown later, analogous to earlier discussion of vectors, various kinds of special matrices can be defined. For the moment, however, we define the matrix that is a generalization of the **0** vector. *This matrix, called a null matrix, and denoted* $\boldsymbol{\phi}$*, consists of entries that are all zeros.*

Illustrations of null matrices of various orders are

$$\boldsymbol{\phi} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \boldsymbol{\phi} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \quad \boldsymbol{\phi} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Notice that each of the above null matrices is made up entirely of **0** vectors and all entries are 0's.

[5] It is often convenient to consider a scalar as a 1 x 1 matrix.

2.5.1 Matrix Transpose

Consider the 2 by 3 matrix:

$$\mathbf{A} = \begin{bmatrix} -1 & 4 & 3 \\ 0 & 5 & 2 \end{bmatrix}$$

Suppose we write the elements of each row of A as columns and obtain

$$\mathbf{A}' = \begin{bmatrix} -1 & 0 \\ 4 & 5 \\ 3 & 2 \end{bmatrix}$$

Note that this new matrix $\mathbf{A}'$ is a 3 by 2 matrix in which the entries of the first row of $\mathbf{A}'$ denote, in the same order, the first column of $\mathbf{A}$. This is also true of the elements in the second and third rows of $\mathbf{A}'$ compared, respectively, to the second and third columns of $\mathbf{A}$.

The new matrix $\mathbf{A}'$ represents the transpose of the original matrix $\mathbf{A}$. *A transpose of* $\mathbf{A}_{m \times n} = (a_{ij})_{m \times n}$ *is a matrix obtained from* $\mathbf{A}$ *by interchanging rows and columns so that*

$$\boxed{\mathbf{A}'_{m \times n} = (a_{ij})'_{m \times n} = (a_{ji})_{n \times m}}$$

To illustrate,

If $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$, then $\mathbf{A}' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

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

If $\boldsymbol{\phi} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$, then $\boldsymbol{\phi}' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Next, we can use the operation of matrix transpose to describe a symmetric matrix. *A square matrix* $\mathbf{A}$ *is called symmetric if*

$$\boxed{\mathbf{A} = (a_{ij}) = \mathbf{A}' = (a_{ji})}$$

That is, a symmetric matrix equals its transpose. To illustrate,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}; \qquad \mathbf{A}' = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

Finally, it should also be evident that the transpose of the transpose of a given matrix is the original matrix itself. That is,

$$\boxed{(\mathbf{A}')' = \mathbf{A}}$$

To illustrate,

$$\mathbf{A} = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \end{bmatrix}; \qquad \mathbf{A}' = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{bmatrix}; \qquad (\mathbf{A}')' = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \end{bmatrix}$$

2.5.2 Matrix Equality, Addition, Scalar Multiplication, and Subtraction

Two matrices **A** *and* **B** *are equal if and only if they are of the same order and each entry of the first is equal to the corresponding entry of the second.* That is,

$$\boxed{\begin{gathered} \mathbf{A} = \mathbf{B} \\ \text{if and only if} \\ (a_{ij}) = (b_{ij}) \\ \text{for} \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \end{gathered}}$$

To illustrate,

$$\mathbf{A} = \begin{bmatrix} -1 & 4 & 3 \\ 0 & 5 & 2 \end{bmatrix} = \mathbf{B} = \begin{bmatrix} -1 & 4 & 3 \\ 0 & 5 & 2 \end{bmatrix}$$

since they are of the same order and $a_{ij} = b_{ij}$, entry by entry.

In matrix addition each entry of a sum matrix is the sum of the corresponding entries of the two matrices being added, again assuming they are of the same order.

That is, we define the matrix **C**, denoting the result of adding **A** to **B** as

$$\boxed{\begin{gathered} \mathbf{C} = \mathbf{A} + \mathbf{B} \\ \text{if and only if} \\ (c_{ij}) = (a_{ij}) + (b_{ij}) \\ \text{for} \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \end{gathered}}$$

To illustrate,

$$\mathbf{A} = \begin{bmatrix} -1 & 4 & 3 \\ 0 & 5 & 2 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 0 & 2 & 1 \\ -1 & 4 & 3 \end{bmatrix}; \quad \mathbf{A} + \mathbf{B} = \mathbf{C} = \begin{bmatrix} -1 & 6 & 4 \\ -1 & 9 & 5 \end{bmatrix}$$

Next, we can consider the case of the transpose of the sum of two matrices.

If **A** *and* **B** *are of common order and if* **C** = **A** + **B**, *then*

$$\boxed{\mathbf{C}' = \mathbf{A}' + \mathbf{B}'}$$

To illustrate,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 1 \\ 5 & 1 & 3 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 3 & 4 \\ 9 & 6 & 9 \end{bmatrix}; \quad \mathbf{C}' = \begin{bmatrix} 1 & 9 \\ 3 & 6 \\ 4 & 9 \end{bmatrix} = \overset{\mathbf{A}'}{\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}} + \overset{\mathbf{B}'}{\begin{bmatrix} 0 & 5 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}}$$

Matrices can also be multiplied by a number (scalar), and this is called scalar multiplication of the matrix. The procedure is simple: One merely multiplies each entry of the matrix by the scalar k. That is,

$$\boxed{\begin{gathered} \mathbf{E} = k\mathbf{A} \\ \text{if and only if} \\ (e_{ij}) = k(a_{ij}) \\ \text{for} \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \end{gathered}}$$

For example, if we wish to multiply **A** by 3, we have

$$3\mathbf{A} = 3\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{bmatrix}$$

Subtraction of matrices is now defined as involving the case in which the matrix being subtracted is first multiplied by −1 and then the two matrices are added. That is

$$\boxed{\begin{gathered} C = A - B \\ \text{if and only if} \\ (c_{ij}) = (a_{ij}) - (b_{ij}) \\ \text{for} \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \end{gathered}}$$

To illustrate,

$$\mathbf{A} = \begin{bmatrix} -1 & 4 & 3 \\ 0 & 5 & 2 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 0 & 2 & 1 \\ -1 & 4 & 3 \end{bmatrix}; \quad \mathbf{A} - \mathbf{B} = \mathbf{C} = \begin{bmatrix} -1 & 2 & 2 \\ 1 & 1 & -1 \end{bmatrix}$$

2.5.3 Properties of Matrix Addition and Scalar Multiplication

Some of the properties exhibited by matrix addition[6] and scalar multiplication are listed below for future use:

1. Matrix addition is commutative:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

2. Matrix addition is associative:

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

3. Scalar multiplication of a matrix is commutative:

$$\mathbf{A}(k\mathbf{B}) = (k\mathbf{A})\mathbf{B}$$

4. Scalar multiplication is associative:

$$k_1(k_2\mathbf{A}) = (k_1k_2)\mathbf{A}$$

5. Scalar multiplication is distributive:

$$(k_1 + k_2)\mathbf{A} = k_1\mathbf{A} + k_2\mathbf{A}$$

6. There exists a null matrix (already defined) ϕ with the property that

$$\mathbf{A} + \phi = \mathbf{A}$$

7. Every matrix $\mathbf{A}$ has a counterpart matrix $-\mathbf{A}$ such that

$$\mathbf{A} + (-\mathbf{A}) = \phi$$

Not surprisingly, the preceding rules are similar to the ones discussed for vector operations in Sections 2.3.3 and 2.3.4 and could be numerically illustrated in similar fashion.

2.5.4 Matrix Multiplication

In discussing the multiplication of two (or more) matrices, *conformability* should first be pointed out. Similar to the previous discussion of matrix equality, addition, and subtraction, in which the matrices were assumed to be of common order before the relation or operation was meaningful, in multiplication the matrices must be conformable. If we wish to multiply **A** by **B**, they must exhibit commonality of *interior* dimensions.

For example, if **A** is of order m rows by n columns, it can be written as $\mathbf{A}_{m \times n}$. Next, suppose we have a second matrix **B** of order n rows by p columns, written as $\mathbf{B}_{n \times p}$. Using this form, we have

$$\boxed{\mathbf{A}_{m \times n} \quad \mathbf{B}_{n \times p} = \mathbf{C}_{m \times p}}$$

[6] It should be noted, however, that matrix subtraction is neither commutative nor associative.

Note that **A**, called the prefactor, has n columns. This is the "interior" dimension. If **A** has an interior dimension of n columns, then **B**, called the postfactor, must have an interior dimension of n rows. This is the condition of conformability and is necessary for matrix multiplication. Note that their product **C**, then, is of order m by p. These are the "exterior" dimensions of **A** and **B**, respectively.

A simple way to find the order of the matrix product is shown below:

$$\mathbf{A}_{m \times n} \cdot \mathbf{B}_{n \times p} = \mathbf{C}_{m \times p}$$

Note that the interior dimensions are the same (n columns of **A** and n rows of **B**) and that the exterior dimensions are obtained from the "outer" dimensions of **A** and **B**, respectively.

This same idea holds for more than two matrices, again assuming that all are conformable. For example, with three matrices, we have

$$\mathbf{A}_{m \times n} \quad \mathbf{B}_{n \times p} \quad \mathbf{C}_{p \times r} = \mathbf{D}_{m \times r}$$

Note further that the interior dimensions of the matrices conform.

Matrix multiplication follows a row-by-column rule, equivalent to the scalar product (Section 2.3.6) of each row of the first matrix with each column of the second. That is, we take the entries of each row of the prefactor, and these are multiplied by the corresponding entries of each column of the postfactor and then summed. If we use the first row of **A** and the first column of **B**, then the first element in **C** (i.e., c_{11}) will be the result of the preceding operation. The second element of **C** (i.e., c_{12}) is found by using the first row of **A** and the second column of **B**, and so on.

This can be summarized as follows:

$$\mathbf{C} = \mathbf{AB}$$

is defined as

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

for $i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, p$

To illustrate, we let

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

Then

$$\mathbf{C} = \mathbf{AB}$$

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

$$\mathbf{C}_{2\times 2} = \begin{bmatrix} 3 & 13 \\ 5 & 8 \end{bmatrix}$$

Now, let us reverse the order of multiplication. In this case we can do so since **B** is of order 3 by 2 and **A** is of order 2 by 3. In other instances, however (e.g., if **A** were of order 3 by 3), we could not multiply **B** by **A** since they would then not be conformable.

If we now multiply **BA** = **D**, we obtain the following:

$$\mathbf{D} = \mathbf{B}_{3\times 2}\ \mathbf{A}_{2\times 3} = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 3 & 2 \\ 2 & 0 & 1 \end{bmatrix}$$

$$\mathbf{D}_{3\times 3} = \begin{bmatrix} 4 & 6 & 7 \\ 7 & 3 & 6 \\ 3 & 3 & 4 \end{bmatrix}$$

Notice that $\mathbf{D} \neq \mathbf{C}$ and, as a matter of fact, they are not even of the same order. Even if two matrices are conformable, in general, $\mathbf{AB} \neq \mathbf{BA}$. That is, matrix multiplication, in general, is noncommutative. Hence, in discussing matrix multiplication we should *refer explicitly to the order in which they multiply.* For example, the matrix product **AB** can be described as "**A** is postmultiplied by **B**," or "**B** is premultiplied by **A**." Alternatively, we could use the terms "prefactor" and "postfactor" as mentioned earlier.

2.5.4.1 Multiplication of a Vector and a Matrix In some cases of interest we shall want to postmultiply some vector **a** by a matrix **B**. To illustrate, suppose we have the following:

$$\mathbf{a}' = (1, 0, 3); \qquad \mathbf{B} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}$$

where $\mathbf{a}'$ is a row vector of order 1 by 3. Note that $\mathbf{a}'$ and **B** are conformable, and we have

$$(1, 0, 3) \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} = (2, 6)$$

where their product displays the order of the exterior dimensions of $\mathbf{a}'$ and $\mathbf{B}$, namely, a 1×2 row vector. Alternatively, suppose we have the column vector $\mathbf{c} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Then we can find the product $\mathbf{B}$ as

$$\begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \\ 2 \end{bmatrix}$$

where their product is a 3×1 column vector. As can be seen, no new rules are involved for either of these cases.

2.5.4.2 Matrix Product of Two Vectors We now might ask what happens when two vectors are multiplied. As shown in Section 2.3.6, the scalar product of two vectors results in a single number (scalar) if row vector times column vector multiplication is performed. However, one might have the case of a column vector multiplying a row vector. In this case the results are quite different, and we consider it next.

To illustrate, assume we have the column vector

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

and the row vector $(1, 1, 2)$. Their matrix (or outer) product is obtained as

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} (1, 1, 2) = \begin{bmatrix} (2 \times 1) & (2 \times 1) & (2 \times 2) \\ (1 \times 1) & (1 \times 1) & (1 \times 2) \\ (3 \times 1) & (3 \times 1) & (3 \times 2) \end{bmatrix} = \begin{bmatrix} 2 & 2 & 4 \\ 1 & 1 & 2 \\ 3 & 3 & 6 \end{bmatrix}$$

which is a 3 by 3 *matrix*. Note in this case that each row of the first "matrix" has only a *single* element, as does each column of the second. Thus, the row-by-column rule is not violated in this special case.

2.5.4.3 Triple Product—Vector, Matrix, Vector Multiplication To round out the discussion we might wish to consider the triple product of

$$\mathbf{a}' = (1, 1, 2); \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix}; \qquad \mathbf{c} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

That is, we desire to find the product $\mathbf{a}'\mathbf{B}\mathbf{c}$. If so, we can proceed in stages. We first find the vector by matrix product:

$$\mathbf{a}'\mathbf{B} = (1, 1, 2) \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix} = (7, 3, 7)$$

Next, we find the scalar product:

$$\mathbf{a}'\mathbf{Bc} = (7, 3, 7)\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = 38$$

Hence, we see that the result of all this is a scalar. Note, of course, that it is of "order 1 x 1" and, thus, is in agreement with the order of the exterior dimensions of $\mathbf{a}'$ and $\mathbf{c}$.

A special case of vector, matrix, vector multiplication takes the form of $\mathbf{a}'\mathbf{Ba}$. This can be illustrated by

$$\mathbf{a}' = (1, 1, 2); \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix}; \qquad \mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Then

$$\mathbf{a}'\mathbf{B} = (1, 1, 2)\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix} = (7, 3, 7)$$

and

$$\mathbf{a}'\mathbf{Ba} = (7, 3, 7)\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 24$$

As noted, the result is also a scalar. Both of these cases are relevant to multivariate analysis and are discussed in later chapters.

2.5.5 Some Properties of Matrix Multiplication

We have already pointed out that matrix multiplication, in general, is noncommutative. However, matrix multiplication *does* obey certain other properties.

1. Associativity–assuming all matrices to be conformable we can state that

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

2. Distributivity–again assuming conformable matrices we can state that

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}; \qquad \text{left distributive law}$$

$$(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}; \qquad \text{right distributive law}$$

3. If k is a scalar, then we have the associativity property:

$$k(\mathbf{AB}) = (k\mathbf{A})\mathbf{B}$$

Finally, it is of interest to point out the rule involving the transpose of the product of two (or more) matrices. In the case of two matrices, the rule is

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

That is, the transpose of the product of two (or more) matrices is equal to the product of their respective transposes, multiplied in reverse order. To illustrate,

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}; \quad \mathbf{A}' = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \quad \mathbf{B}' = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$

Then

$$\mathbf{AB} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 9 \\ 10 & 14 \end{bmatrix}$$

$$(\mathbf{AB})' = \begin{bmatrix} 7 & 10 \\ 9 & 14 \end{bmatrix} = \mathbf{B}'\mathbf{A}' = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 9 & 14 \end{bmatrix}$$

2.5.6 Some Differences between Scalar and Matrix Arithmetic

Probably the most difficult temptation to suppress in working with matrix multiplication involves attributing properties to matrices that we associate with ordinary scalars. Table 2.1 shows some of the pitfalls that one should be wary of in doing

TABLE 2.1

Some Differences between Scalar and Matrix Relations

Scalars	Matrices
1. $ab = ba$	1. $\mathbf{AB} \neq \mathbf{BA}$, in general
2. If $ab = ac$ and $a \neq 0$, then $b = c$.	2. If $\mathbf{AB} = \mathbf{AC}$ and $\mathbf{A} \neq \boldsymbol{\phi}$, then it is not necessary that $\mathbf{B}$ equals $\mathbf{C}$.
3. If $ab = 0$, then either $a = 0$, or $b = 0$, or both $a, b = 0$.	3. If $\mathbf{AB} = \boldsymbol{\phi}$, then it is not necessarily the case that either $\mathbf{A} = \boldsymbol{\phi}$, $\mathbf{B} = \boldsymbol{\phi}$, or both $\mathbf{A}, \mathbf{B} = \boldsymbol{\phi}$.
4. If $ab = 0$, then $ba = 0$.	4. If $\mathbf{AB} = \boldsymbol{\phi}$, then $\mathbf{BA}$ does not necessarily equal $\boldsymbol{\phi}$.

arithmetic with matrices. As shown in the table, some marked differences exist. Not only does commutativity fail to hold in general for matrix multiplication, but other characteristics involving products equal to zero also do not hold generally. For example, if some matrix product **AB** equals the null matrix $\boldsymbol{\phi}$, we note that neither **A**, the prefactor, nor **B**, the postfactor, need to be equal to $\boldsymbol{\phi}$.

2.5.7 The Problem of Matrix Division

Up to this point we have discussed addition, subtraction, and multiplication of matrices, but division has been conspicuous by its absence. *And for good reason: division, as we know it in scalar arithmetic, is not defined in matrix algebra.*

What is defined is something more analogous to multiplication by a reciprocal. For example, in ordinary arithmetic, instead of dividing some number by 5, we could multiply the number by the reciprocal of 5:

$$1/5 = (5)^{-1}$$

assuming that the divisor is not equal to zero.

The analogous operation in matrix algebra is called *matrix inversion.* This operation is so special (and considerably more complex) that we defer discussion of it until Chapter 4. What can be said for now is that the inverse of a matrix **A**, if said inverse exists, is analogous to multiplication of **A** by a reciprocal in ordinary algebra. As such, in matrix algebra there is an analogy to the scalar relation:

$$a \times a^{-1} = 1$$

Needless to say, we shall spend a considerable amount of time on the topic of matrix inversion in subsequent chapters.

2.5.8 Some More Examples of Matrix Operations

To facilitate the review of matrix operations described in Section 2.5, consider the following:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 2 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

$$k = 3; \qquad \mathbf{a}' = (2, 1); \qquad \mathbf{b} = \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}$$

Matrix Transpose The transpose of **A** is

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

and the transpose of **B** is

$$\mathbf{B}' = \begin{bmatrix} 2 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

Addition of Matrices Addition of matrices is illustrated by

$$\mathbf{C} = \mathbf{A} + \mathbf{B}' = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 6 \\ 6 & 6 & 10 \end{bmatrix}$$

Subtraction of Matrices Matrix subtraction is illustrated by

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

Scalar Multiplication of a Matrix

$$\mathbf{C} = k\mathbf{A} = 3\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 6 & 3 & 12 \end{bmatrix}$$

Varieties of Multiplication

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 15 & 32 \\ 18 & 37 \end{bmatrix}$$

$$\mathbf{C}' = (\mathbf{AB})' = \mathbf{B}'\mathbf{A}' = \begin{bmatrix} 2 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 15 & 18 \\ 32 & 37 \end{bmatrix}$$

$$\mathbf{D} = k\mathbf{AB} = 3\begin{bmatrix} 15 & 32 \\ 18 & 37 \end{bmatrix} = \begin{bmatrix} 45 & 96 \\ 54 & 111 \end{bmatrix}$$

$$\mathbf{E} = \mathbf{a}'\mathbf{Ab} = (2, 1)\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix} = 60$$

$$\mathbf{F} = \mathbf{b}'\mathbf{B}(\mathbf{a}')' = (5, 2, 3)\begin{bmatrix} 2 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 94$$

Additional examples appear at the end of the chapter.

2.6 SOME SPECIAL MATRICES

So far we have been dealing mainly with arbitrary rectangular matrices. In a few cases we have used square matrices for illustrative purposes. In matrix algebra there are a number of special matrices that are encountered in multivariate analysis. We consider some of these here, particularly those that are frequently utilized in multivariate procedures.

2.6.1 Symmetric Matrices

Figure 2.1 shows, in schematic form, various special matrices of interest to multivariate analysis. The first property for categorizing types of matrices concerns whether they are square ($m = n$) or rectangular. In turn, rectangular matrices can be either vertical ($m > n$) or horizontal ($m < n$).

As we shall show in later chapters, square matrices play an important role in multivariate analysis. In particular, the notion of matrix symmetry is important. Earlier, a symmetric matrix was defined as a square matrix that satisfies the relation

$$\boxed{\mathbf{A} = \mathbf{A}' \qquad \text{or, equivalently,} \qquad (a_{ij}) = (a_{ji})}$$

That is, a symmetric matrix is a square matrix that is equal to its transpose. For example,

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & -5 \\ 4 & -5 & 1 \end{bmatrix}; \qquad \mathbf{A}' = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & -5 \\ 4 & -5 & 1 \end{bmatrix}$$

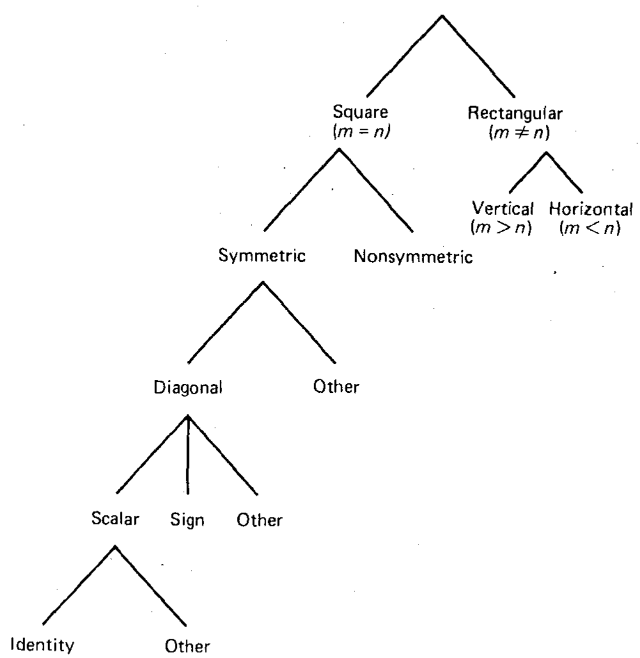


Fig. 2.1 Various types of matrices.

Symmetric matrices, such as correlation matrices and covariance matrices, are quite common in multivariate analysis, and we shall come across them repeatedly in later chapters.[7]

A few properties related to symmetry in matrices are of interest to point out:

1. The product of any (not necessarily symmetric) matrix and its transpose is symmetric; that is, both $\mathbf{AA}'$ and $\mathbf{A}'\mathbf{A}$ are symmetric matrices.
2. If $\mathbf{A}$ is any square (not necessarily symmetric) matrix, then $\mathbf{A} + \mathbf{A}'$ is symmetric.
3. If $\mathbf{A}$ is symmetric and k is a scalar, then $k\mathbf{A}$ is a symmetric matrix.
4. The sum of any number of symmetric matrices is also symmetric.
5. The product of two symmetric matrices is not necessarily symmetric.

Later chapters will discuss still other characteristics of symmetric matrices and the special role that they play in such topics as matrix eigenstructures and quadratic forms.

2.6.2 Diagonal, Scalar, Sign, and Identity Matrices

A special case of a symmetric matrix is a diagonal matrix. *A diagonal matrix is defined as a square matrix in which all off-diagonal entries are zero.* (Note that a diagonal matrix is necessarily symmetric.) Entries on the main diagonal may or may not be zero.

Examples of diagonal matrices are

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

If all entries on the main diagonal are equal scalars, then the diagonal matrix is called a *scalar matrix.*

Examples of scalar matrices are

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix}$$

If some of the entries on the main diagonal are -1 and the rest are $+1$, the diagonal matrix is called a *sign matrix.* Examples of sign matrices are

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

[7] While we do not go into detail here, a *skew symmetric* matrix $\mathbf{A}$ is a square matrix in which all main diagonal elements a_{ii} are zero and $A = -A'$. For example,

$$\text{If } \mathbf{A} = \begin{bmatrix} 0 & -3 & 1 \\ 3 & 0 & 2 \\ -1 & -2 & 0 \end{bmatrix}; \quad \text{then} \quad -\mathbf{A}' = \begin{bmatrix} 0 & -3 & -1 \\ 3 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix}$$

is skew symmetric.

If the entries on the diagonal of a scalar matrix are each equal to unity, then this type of scalar matrix is called an *identity matrix*, denoted **I**. Examples are

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

A number of useful properties are associated with diagonal matrices and, hence, the special cases of scalar, sign, and identity matrices.

* 1. The transpose of a diagonal matrix is equal to the original matrix.

2. Sums and differences of diagonal matrices are also diagonal matrices.

3. Premultiplication of a matrix **A** by a diagonal matrix **D** results in a matrix in which each entry in a given row is the product of the original entry in **A** corresponding to that row and the diagonal element in the corresponding row of the diagonal matrix. To illustrate,

$$\underset{\mathbf{D}}{\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \underset{\mathbf{A}}{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 1 \end{bmatrix}} = \underset{\mathbf{DA}}{\begin{bmatrix} 3 & 6 & 9 & 12 \\ 4 & 2 & 8 & 6 \\ 3 & 2 & 1 & 1 \end{bmatrix}}$$

4. Postmultiplication of a matrix **A** by a diagonal matrix **D** results in a matrix in which each entry in a given column is the product of the original entry in **A** corresponding to that column and the diagonal element in the corresponding column of the diagonal matrix. To illustrate,

$$\underset{\mathbf{A}}{\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 4 & 1 \\ 4 & 3 & 1 \end{bmatrix}} \underset{\mathbf{D}}{\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}} = \underset{\mathbf{AD}}{\begin{bmatrix} 3 & 4 & 3 \\ 6 & 2 & 2 \\ 9 & 8 & 1 \\ 12 & 6 & 1 \end{bmatrix}}$$

5. Pre- and postmultiplication of a matrix **A** by diagonal matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ result in a matrix whose *ij*th entry is the product of the *i*th entry in the diagonal of the premultiplier, the *ij*th entry of **A**, and the *j*th entry of the postmultiplier. For example,

$$\underset{\mathbf{D}_1}{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}} \underset{\mathbf{A}}{\begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 4 & 1 \end{bmatrix}} \underset{\mathbf{D}_2}{\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}} = \underset{\mathbf{D}_1\mathbf{A}\mathbf{D}_2}{\begin{bmatrix} 2 & 18 \\ 2 & 6 \\ 12 & 9 \end{bmatrix}}$$

* 6. The product of any number of diagonal matrices is a diagonal matrix, each of whose entries is the product of the corresponding diagonal entries of the matrices. For example,

$$\underset{\mathbf{D}_1}{\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}} \underset{\mathbf{D}_2}{\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}} \underset{\mathbf{D}_3}{\begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}} = \underset{\mathbf{D}_1\mathbf{D}_2\mathbf{D}_3}{\begin{bmatrix} 8 & 0 \\ 0 & -6 \end{bmatrix}}$$

7. Diagonal matrix multiplication, assuming conformability, is commutative.
8. Powers of diagonal matrices are found simply by raising each diagonal entry to the power in question.[8] (Roots are found analogously.)
9. Pre- or postmultiplication of a matrix **A** by a scalar matrix multiplies all entries of **A** by the constant entry in the scalar matrix. It is equivalent to scalar multiplication of the matrix, by that scalar appearing on the diagonal.
10. As a special case, pre- or postmultiplication of a matrix **A** by **I**, the identity matrix, leaves the original matrix unchanged.
11. Powers of an identity matrix equal the original matrix.

While the above properties are by no means exhaustive of the characteristics of diagonal matrices, or the special cases of scalar, sign, and identity matrices, they do represent the main properties of interest to applied researchers.

2.6.3 Some Additional Examples

As an aid to integrating some of the discussion of this section, consider the following:

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix}; \quad \mathbf{D}_1 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \mathbf{D}_2 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Premultiplication by a Diagonal

$$\mathbf{D}_1\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 12 & 3 \\ 4 & 10 & 2 \\ 3 & 6 & 2 \end{bmatrix}$$

Postmultiplication by a Diagonal

$$\mathbf{A}\mathbf{D}_1 = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 1 \\ 6 & 10 & 1 \\ 9 & 12 & 2 \end{bmatrix}$$

[8] In general, a square matrix **A** can be raised to any power n that is a positive whole number by multiplying it by itself n times, denoted as $\mathbf{A}^n$. Roots can also be found for certain square matrices (not restricted to being diagonal). If a square matrix **A** has an nth root, then the matrix $\mathbf{A}^{1/n}$, when multiplied by itself n times, equals **A**. The topic of powers and roots of (square) matrices is covered in Chapter 5.

Pre- and Postmultiplication by Diagonals

$$\mathbf{D}_2\mathbf{A}\mathbf{D}_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 16 & 2 \\ 12 & 20 & 2 \\ 18 & 24 & 4 \end{bmatrix}$$

Scalar Matrix Multiplication

$$3\begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 12 & 3 \\ 6 & 15 & 3 \\ 9 & 18 & 6 \end{bmatrix}$$

Powers and Roots of a Diagonal with Positive Entries

$$\mathbf{D}_1^{\,2} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{D}_2^{\,1/2} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}^{1/2} = \begin{bmatrix} 1.414 & 0 & 0 \\ 0 & 1.414 & 0 \\ 0 & 0 & 1.414 \end{bmatrix}.$$

Identity and Sign Matrices

$$\mathbf{IA} = \mathbf{AI} = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} = \mathbf{A}$$

Let

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

Then

$$\mathbf{AF} = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 5 & 1 \\ 3 & 6 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -4 & 1 \\ -2 & -5 & 1 \\ -3 & -6 & 2 \end{bmatrix}$$

Additional examples appear at the end of the chapter.

2.7 DETERMINANTS OF MATRICES

The determinant of a matrix plays an important role in more advanced matrix concepts such as matrix inversion and matrix rank, as well as in multivariate analysis involving generalized measures of variance. Only square matrices have determinants. The determinant of a square matrix is a scalar function of the entries of the matrix. We denote the determinant of a matrix **A** by the symbol $|\mathbf{A}|$ and reiterate that the value of the determinant is expressed as a single number (scalar).

The early development of determinants was intimately connected with procedures for solving simultaneous equations. As historical background, and motivational interest, consider the two linear equations:

$$ax + by = c$$

$$dx + ey = f$$

These equations could be expressed in matrix times vector form as

$$\begin{bmatrix} a & b \\ d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ f \end{bmatrix}$$

Note that the left-hand side of the equations is simply the product of a 2 x 2 matrix times a 2 x 1 vector while the right-hand side is another 2 x 1 vector.

As may be recalled from elementary algebra, this system of equations can be solved, for, say, x by the formula

$$x = \frac{ce - fb}{ae - db}$$

assuming that the denominator of the above ratio is not equal to zero.

We can consider the right-hand side of the above equation in the context of determinants by expressing both numerator and denominator of the ratio as

$$x = \frac{\begin{vmatrix} c & b \\ f & e \end{vmatrix}}{\begin{vmatrix} a & b \\ d & e \end{vmatrix}}$$

In the simple case shown here, the determinants of $\begin{bmatrix} c & b \\ f & e \end{bmatrix}$ and $\begin{bmatrix} a & b \\ d & e \end{bmatrix}$ are easy to define. That is

$$\begin{vmatrix} c & b \\ f & e \end{vmatrix} = ce - fb \qquad \text{and} \qquad \begin{vmatrix} a & b \\ d & e \end{vmatrix} = ae - db$$

these are called *second-order* determinants. The unknown quantity x is the ratio of two determinants (scalars).

Historically, determinants were employed widely in the solution of simultaneous equations. With the advent of newer solution methods, however, their application in this context has diminished. Still, it is important to have some grasp of the rudiments of determinants, if only as a precursor to other procedures for solving equations that are developed in subsequent chapters.

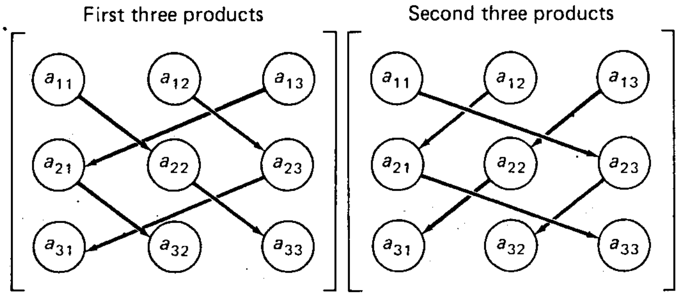
2.7.1 Operational Definition of a Determinant

The theoretical definition of a determinant for matrices of larger order than 2 x 2 is rather cumbersome and, therefore, as in the 2 x 2 case, we shall define it operationally as a series of computational steps. Assuming a square matrix **A** of order $m \times m$ with general entry (a_{ij}), the determinant of that matrix is found by carrying out the following sequence:

1. Form all possible products of m factors each, such that each factor is an entry of **A** and no two factors are drawn from the same row or column of **A**. There are $m!$ (m factorial) products of this type. For example, if **A** is of order 3 x 3, we have 3! or six products:

(i) $a_{11}\ a_{22}\ a_{33}$	(ii) $a_{12}\ a_{23}\ a_{31}$	(iii) $a_{13}\ a_{21}\ a_{32}$
(iv) $a_{13}\ a_{22}\ a_{31}$	(v) $a_{11}\ a_{23}\ a_{32}$	(vi) $a_{12}\ a_{21}\ a_{33}$

We note that each of the subscripts (1, 2, or 3) appears just once as a row subscript and just once as a column subscript in each of the six triple products. The connections shown below illustrate these six products.



2. Within each separate triple product arrange the factors so that row subscripts are in their natural order; this has been done above. Then count the number of inversions or transpositions involving *column* subscripts. In this case an inversion takes place when a larger column subscript precedes a smaller one. For the six triple products above, we have the following frequencies of inversions:

(i) 0 inversion	(ii) 2 inversions	(iii) 2 inversions
(iv) 3 inversions	(v) 1 inversion	(vi) 1 inversion

For example in case (ii), involving the product $a_{12}a_{23}a_{31}$, we note that column subscripts 1 and 2 need to be interchanged, followed by the interchange of column subscripts 3 and 2, in order to obtain the natural order.

3. Having done this for all $m!$ products, multiply each product that has an odd number of inversions by -1. If zero or an even number of inversions is involved, multiply by $+1$; that is, leave the product as is. In the above case the first three products (associated with an incidence of even-type inversions of 0, 2, and 2) will *not* be changed in sign, while the last three products will.

(i) $1(a_{11}a_{22}a_{33})$	(ii) $1(a_{12}a_{23}a_{31})$	(iii) $1(a_{13}a_{21}a_{32})$
(iv) $-1(a_{13}a_{22}a_{31})$	(v) $-1(a_{11}a_{23}a_{32})$	(vi) $-1(a_{12}a_{21}a_{33})$

4. Add the products (observing sign) together. This *sum* is the determinant.

$$|\mathbf{A}| = +(a_{11}a_{22}a_{33}) + (a_{12}a_{23}a_{31}) + (a_{13}a_{21}a_{32})$$
$$-(a_{13}a_{22}a_{31}) - (a_{11}a_{23}a_{32}) - (a_{12}a_{21}a_{33})$$

5. Notice, then, that three steps are involved in finding a determinant. The first step is to form all possible products that can be obtained by taking one element out of one row and column, another out of another row and column, and so on. A matrix of order $m \times m$ will yield $m!$ such products, each composed of m elements. The second step is to affix an algebraic sign to each product via the rule proposed above. The third step is to sum the $m!$ signed products.

6. The procedure can now be formalized by defining the determinant of $\mathbf{A}_{m \times m}$ as the sum of all $m!$ products (each with m factors) in $\mathbf{A}$ of the form

$$(-1)^t a_{1j_1} a_{2j_2} \ldots a_{mj_m}$$

where the sum is understood to be taken over all permutations of the second subscripts. The exponent t denotes the number of inversions required to bring the second subscripts into their natural sequence $(1, 2, \ldots, m)$.

Now let us illustrate the computation of determinants for two simple cases.

The 2 x 2 Case

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$|\mathbf{A}| = (-1)^0 a_{11}a_{22} + (-1)^1 a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21} = (1 \times 4) - (2 \times 3) = -2$$

The 3 x 3 Case

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 2 & 1 & 1 \end{bmatrix}$$

$$|\mathbf{A}| = (-1)^0 a_{11}a_{22}a_{33} + (-1)^2 a_{12}a_{23}a_{31} + (-1)^2 a_{13}a_{21}a_{32}$$
$$+ (-1)^3 a_{13}a_{22}a_{31} + (-1)^1 a_{11}a_{23}a_{32} + (-1)^1 a_{12}a_{21}a_{33}$$
$$= (1 \times -1 \times 1) + (2 \times 4 \times 2) + (3 \times 2 \times 1) - (3 \times -1 \times 2) - (1 \times 4 \times 1) - (2 \times 2 \times 1)$$
$$= -1 + 16 + 6 + 6 - 4 - 4$$

$$|\mathbf{A}| = 19$$

2.7.2 Expansion of Determinants by Cofactors

Even on the basis of the step-by-step demonstration shown above, the evaluation of a determinant (i.e., the process of finding the numerical value of the determinant) is rather complicated and prone to error. Understandably, we might seek some easier procedure in which the arithmetic is simpler and the computations more straightforward. Expansion by cofactors is one such method.

As shown above, a particularly simple determinant can be computed for the case of a 2×2 matrix:

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

In this case $|A| = a_{11}a_{22} - a_{21}a_{12}$.

We can take advantage of this simple 2×2 case in attempting to evaluate high-order determinants (i.e., those of matrices of order 3×3 and higher).

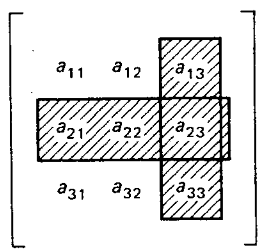
To do this, we first define the minor of an entry (a_{ij}) of a square matrix $A = (a_{ij})$ as the determinant of a submatrix obtained by deleting the ith row and jth column of A. For example, the minor of the entry a_{23} in the matrix.

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

is

$$\text{minor } (a_{23}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}.$$

Notice that this entails omitting those entries in the shaded area:



Similarly, we could find the minors of each of the other eight entries in A.

The cofactor of an entry a_{ij} of a square matrix $A = (a_{ij})$ is the product of the minor of (a_{ij}) and $(-1)^{i+j}$. The cofactor is also called a *signed* minor and is denoted by A_{ij}. In the above case,

$$A_{23} = (-1)^{2+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = -[(a_{11}a_{32}) - (a_{31}a_{12})]$$

Notice that the placement of signs follows an alternating pattern:

$$\begin{vmatrix} + & - & + & - & + & . & . & . \\ - & + & - & + & - & . & . & . \\ + & - & + & - & + & . & . & . \\ - & + & - & + & - & . & . & . \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Similarly, we could develop a cofactor (or a signed minor) for each of the eight remaining entries in **A**. However, to evaluate |**A**|, it turns out that we need only develop cofactors for any single *row*, or any single *column*, of the original matrix (rather than one each for all nine entries in, for example, the 3 x 3 matrix above).

To see why this is so, we can rearrange the entries in |**A**| so that those in the second row appear first.

$$|\mathbf{A}| = a_{22}(a_{11}a_{33}) + a_{23}(a_{12}a_{31}) + a_{21}(a_{13}a_{32}) - a_{22}(a_{13}a_{31}) - a_{23}(a_{11}a_{32}) - a_{21}(a_{12}a_{33})$$

We then simplify as follows:

$$|\mathbf{A}| = a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) + a_{23}(a_{12}a_{31} - a_{11}a_{32})$$

However, since the cofactor A_{23} has already been defined as

$$A_{23} = (-1)^{2+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = -[(a_{11}a_{32} - a_{12}a_{31})] = (a_{12}a_{31} - a_{11}a_{32})$$

we can substitute the cofactor for the last term on the right in the above expression for |**A**|. Similarly, we can substitute the other two cofactors: A_{21} and A_{22} involving the second row of **A**.
These are

$$A_{21} = (-1)^{2+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = -[(a_{12}a_{33} - u_{13}a_{32})] = (a_{13}a_{32} - a_{12}a_{33})$$

$$A_{22} = (-1)^{2+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = (a_{11}a_{33} - a_{13}a_{31})$$

Having done all this, we can evaluate |**A**| via cofactor expansion as

$$|A| = a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23}$$

Alternatively, we can evaluate |**A**| by expanding along the first or third rows, or along any of the three columns of **A**.

In summary, in computing a determinant of a matrix of order m, expansion by cofactors transforms the problem into evaluating m determinants of order $m - 1$ and forming a linear combination of these. This procedure is continued along successive stages until second-order determinants are reached. For high-order matrices (e.g., $m \geqslant 4$), expansion by cofactors provides a simple stagewise, if still tedious, way to compute determinants by hand, ultimately arriving at computations involving second-order determinants.

Let us now illustrate the evaluation of determinants through cofactor expansion for the 3 x 3 and 4 x 4 cases:

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

We can now apply cofactor expansion using, illustratively, the second row of **A**:

$$A_{21} = (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1$$

$$A_{22} = (-1)^{2+2} \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} = -5$$

$$A_{23} = (-1)^{2+3} \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = 3$$

Then, by continued expansion we get

$$|A| = 2(1) - 1(-5) + 4(3) = 19$$

The same principle, illustrated above, applies in the case of fourth- or higher-ordered determinants. To illustrate, suppose we expand the determinant around the *first column* of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 0 & 1 & 3 & 3 \\ 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

As noted, the entries of the first column of **A** are 1, 0, 2, and 0. (By choosing a column with several zeros in it, the computations are simplified.)
We first find

$$|\mathbf{A}| = 1(-1)^{1+1} \begin{vmatrix} 1 & 3 & 3 \\ 1 & 0 & 1 \\ 1 & 2 & 2 \end{vmatrix} + 0 + 2(-1)^{3+1} \begin{vmatrix} 2 & 3 & 5 \\ 1 & 3 & 3 \\ 1 & 2 & 2 \end{vmatrix} + 0$$

We now continue to expand around the first column of each of the two 3 x 3 minors of **A**, above.

$$|\mathbf{A}| = 1\left\{1(-1)^{(1+1)} \begin{vmatrix} 0 & 1 \\ 2 & 2 \end{vmatrix} + 1(-1)^{(2+1)} \begin{vmatrix} 3 & 3 \\ 2 & 2 \end{vmatrix} + 1(-1)^{(3+1)} \begin{vmatrix} 3 & 3 \\ 0 & 1 \end{vmatrix}\right\}$$

$$+ 2\left\{2(-1)^{(1+1)} \begin{vmatrix} 3 & 3 \\ 2 & 2 \end{vmatrix} + 1(-1)^{(2+1)} \begin{vmatrix} 3 & 5 \\ 2 & 2 \end{vmatrix} + 1(-1)^{(3+1)} \begin{vmatrix} 3 & 5 \\ 3 & 3 \end{vmatrix}\right\}$$

$$= 1(-2 - 0 + 3) + 2(0 + 4 - 6)$$

$$|\mathbf{A}| = -3$$

While we stop our illustrations with the case of fourth-order determinants, the same principles can be applied to fifth and higher-ordered determinants. Fortunately, the availability of computer programs takes the labor out of finding determinants in problems of realistic size.

2.7.3 Some Properties of Determinants

A number of useful properties are associated with determinants. The most important of these are listed below:

1. If a matrix **B** is formed from a matrix **A** by interchanging a pair of rows (or a pair of columns), then $|\mathbf{A}| = -|\mathbf{B}|$.
2. If all entries of some row or column of **A** are zero, then $|\mathbf{A}| = 0$.
3. If two rows (or two columns) of **A** are equal, then $|\mathbf{A}| = 0$.
4. The determinant of **A** equals that of its transpose **A**′; that is, $|\mathbf{A}| = |\mathbf{A}'|$.
5. The determinant of the product of two (square) matrices of the same order equals the product of the determinants of the two matrices; that is, $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$.
6. If every entry of a row (or column) of **A** is multiplied by a scalar k, then the value of the determinant is $k|\mathbf{A}|$.
7. If the entries of a row (or column) of **A** are multiplied by a scalar and the results added or subtracted from the corresponding entries of another row (or column, respectively), then the determinant is unchanged.

Illustrations of these various properties follow:

Property 1

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

$$|\mathbf{A}| = -|\mathbf{B}| = 10$$

Property 2

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 4 \end{bmatrix}; \qquad |\mathbf{A}| = 0$$

Property 3

$$\mathbf{A} = \begin{bmatrix} 3 & 3 \\ 2 & 2 \end{bmatrix}; \qquad |\mathbf{A}| = 0$$

Property 4

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

$$|\mathbf{A}| = |\mathbf{A}'| = 10$$

Property 5

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 4 & 2 \\ 3 & 5 \end{bmatrix}; \qquad \mathbf{AB} = \begin{bmatrix} 15 & 11 \\ 20 & 24 \end{bmatrix}$$

$$|\mathbf{A}| = 10; \qquad |\mathbf{B}| = 14; \qquad |\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}| = 140$$

Property 6

$$\mathbf{A}=\begin{bmatrix}3 & 1\\ 2 & 4\end{bmatrix};\qquad k=3;\qquad \mathbf{B}=\begin{bmatrix}9 & 1\\ 6 & 4\end{bmatrix}$$

$$|\mathbf{A}|=10;\qquad |\mathbf{B}|=3|\mathbf{A}|=30$$

Property 7

$$\mathbf{A}=\begin{bmatrix}3 & 1\\ 2 & 4\end{bmatrix};\qquad k=3;\qquad \mathbf{b}=3\begin{bmatrix}3\\ 2\end{bmatrix}=\begin{bmatrix}9\\ 6\end{bmatrix}$$

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

$$|\mathbf{A}|=|\mathbf{C}|=10$$

In addition to the above (selected) properties of a determinant, we state two very important aspects of determinants that are relevant for discussion in subsequent chapters.

1. *A (square) matrix* **A** *is said to be singular if* $|\mathbf{A}| = 0$. *If* $|\mathbf{A}| \neq 0$, *it is said to be nonsingular.* This aspect of determinants will figure quite prominently in our future discussion of the regular inverse of a square matrix.
2. *The rank of a matrix is the order of the largest square submatrix whose determinant does not equal zero.*

To illustrate the characteristics of these definitions, consider the matrix:

$$\mathbf{A}=\begin{bmatrix}1 & 2 & 3\\ 0 & 1 & 2\\ 2 & 4 & 6\end{bmatrix}$$

Assume that we wish to find its determinant by cofactor expansion. We expand along the first column.

$$|\mathbf{A}|=1(-1)^{1+1}\begin{vmatrix}1 & 2\\ 4 & 6\end{vmatrix}+0+2(-1)^{1+3}\begin{vmatrix}2 & 3\\ 1 & 2\end{vmatrix}=-2+0+2$$

$$|\mathbf{A}|=0$$

We see that $|\mathbf{A}| = 0$ and, according to the definition above, **A** is singular. In this case we note that the entries of the third row of **A** are precisely twice their counterparts in the first row of **A**. *In general, if a particular row (or column) can be perfectly predicted from a linear combination of the other rows (columns), the matrix is said to be singular.*

Proceeding to the next topic (i.e., the rank of **A**), we check to see if a 2 x 2 submatrix exists whose determinant does not equal zero.

$$\text{minor}\,(a_{11})=\begin{vmatrix}1 & 2\\ 4 & 6\end{vmatrix}=-2$$

Note that we need go no further, since we have found a 2 x 2 submatrix whose determinant does not equal zero.

Notice, also, that even though the matrix is of order 3 x 3, the rank of A cannot be 3, since $|\mathbf{A}| = 0$. However, it does turn out that at least one submatrix of order 2 x 2, as illustrated above, has a nonzero determinant. Hence, the rank of A, in this case, is 2.

2.7.4 The Pivotal Method of Evaluating Determinants

In relatively large matrices, such as those of fourth and higher order, the evaluation of determinants even by cofactor expansion, becomes time consuming. Over the years mathematicians have developed a wide variety of numerical methods for evaluating determinants. One of these techniques, the pivotal method (Rao, 1952), has been chosen to illustrate this class of procedures. While we illustrate the method in the context of evaluating determinants, much more is obtained, as will be demonstrated in Chapter 4.

The easiest way to describe the pivotal method is by a numerical example. For illustrative purposes let us evaluate the determinant of a fourth-order matrix:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 & 2 \\ 4 & 2 & 3 & 4 \\ 1 & 4 & 2 & 2 \\ 3 & 1 & 0 & 1 \end{bmatrix}$$

Evaluating the determinant of A proceeds in a step-by-step fashion, with the aid of a work sheet similar to that appearing in Table 2.2.

The top, left-hand portion of Table 2.2 shows the original matrix A, whose determinant we wish to evaluate. To the right of this matrix is shown an identity matrix of the same order (4 x 4) as the matrix A. The last column (column 9) is a check sum column, each entry of which represents the algebraic sum of the specific row of interest. (Other than for arithmetic checking purposes, column 9 plays no role in the computations.)

The objective behind the pivotal method is to reduce the columnar entries in A successively so that for each column of interest we have only one entry and this single entry is unity. Specifically, the boxed entry in row 01 (the number 2) serves as the first pivot. Row 10 is obtained from row 01 by dividing each entry in row 01 by 2, the pivot item. Note that *all* entries in row 01 are divided by the pivot, including the entries under the identity matrix and the check sum column. Dividing 2 by itself, of course, produces the desired entry of unity in the first column of row 10.

Row 11 is obtained from the results of two operations. First, we multiply each entry of row 10 by 4, the first entry in row 02. This particular step is not shown in the work sheet, but the nine products are

$$4; \quad 6; \quad 2; \quad 4; \quad 2; \quad 0; \quad 0; \quad 0; \quad 18$$

These are then subtracted from their counterpart entries in row 02, so as to obtain row 11.

TABLE 2.2

Evaluating a Determinant by the Pivotal Method

Row No.	Original matrix				Identity matrix				Check sum column
0	1	2	3	4	5	6	7	8	9
01	[2]	3	1	2	1	0	0	0	9
02	4	2	3	4	0	1	0	0	14
03	1	4	2	2	0	0	1	0	10
04	3	1	0	1	0	0	0	1	6
10	1	1.5	0.5	1	0.5	0	0	0	4.5
11		[−4]	1	0	−2	1	0	0	−4
12		2.5	1.5	1	−0.5	0	1	0	5.5
13		−3.5	−1.5	−2	−1.5	0	0	1	−7.5
20		1	−0.25	0	0.5	−0.25	0	0	1
21			[2.125]	1	−1.75	0.625	1	0	3.0
22			−2.375	−2	0.25	−0.875	0	1	−4.0
30			1	0.471	−0.824	0.294	0.471	0	1.412
31				[−0.881]	−1.707	−0.177	1.119	1	−0.646
40				1	1.938	0.201	−1.270	−1.135	0.733
30*			1		−1.737	0.199	1.069	0.534	
20*		1			0.066	−0.200	0.267	0.134	1.267
10*	1				−0.668	−0.001	0.334	0.667	1.332

$|A| = (2)(-4)(2.125)(-0.881) = 15$

Note, particularly, that this subtraction has the desired effect of producing a zero (shown as a blank) in the first entry of row 11. Note further that the entries of row 11 add up to −4, the row check sum in the last column; the check sum column is provided for all rows and serves as an arithmetic check on the computations.

Row 12 is obtained in an analogous way; here, since the first element in row 03 is 1, we multiply row 10 by unity and then subtract the row 10 elements from their counterparts in row 03. Row 13 is also obtained in the same way. First, the row 10 entries are multiplied by 3, the first entry in row 04. Then these entries are subtracted from their counterparts in row 04. Finally, we see that in rows 10 through 13, all entries in column 1 are zero (and represented by blanks) except the first entry which is unity.

At the next stage in the computations, the first element in row 11 becomes the pivot. All entries in row 11 are divided by −4, the new pivot, and the results are shown in row

20. Row 20 now becomes the reference row. For example, row 21 is found in a way analogous to row 11. First, we multiply all entries of row 20 by 2.5, the first entry of row 12. Although not shown in the work sheet, these are

$$2.5; \quad -0.625; \quad 0; \quad 1.25; \quad -0.625; \quad 0; \quad 0; \quad 2.5$$

These elements are then subtracted from their counterparts in row 12 and the results appear in row 21.

The procedure is then repeated by multiplying row 20 by −3.5, the leading element in row 13 and subtracting these new entries from their counterparts in row 13. Note that in rows 20 through 22, entries in columns 1 and 2 are all zero, except for the leading element of 1 in row 20.

The third pivot item is the entry 2.125 in row 21. All entries in row 21 are divided by 2.125 and the results listed in row 30. Finally, the entries of row 30 are multiplied by −2.375, the leading entry in row 22. Although not shown in the work sheet, these entries are

$$-2.375; \quad -1.119; \quad 1.957; \quad -0.698; \quad -1.119; \quad 0; \quad -3.353$$

These entries are subtracted from their counterparts in row 22, providing row 31. The last pivot item is −0.881 and appears in row 31.

Finally, the four pivots are multiplied together, leading to the determinant

$$|\mathbf{A}| = (2)(-4)(2.125)(-0.881) = 1.5$$

At this point the reader may well wonder what is the role played by the various changes being made in the identity matrix as the pivot procedure is applied. Moreover, we have not discussed the various calculations appearing in rows 40 through 10*.

As it turns out, the pivotal method is much more versatile and useful than illustrated here. While the determinant of the matrix is, indeed, obtained, the pivotal method can be employed for three important purposes:

1. computing determinants (as the product of pivot elements);[9]
2. solving a set of simultaneous equations;
3. finding the inverse of a matrix.

Here we have only described the first objective. Later on (in Chapter 4) we review the pivotal method in terms of all three of the above objectives and, in the process, discuss the remaining computations in Table 2.2.

The reader may also have wondered about what happens when a candidate pivot is zero (which, fortunately, did not happen in the preceding example). Clearly, we cannot divide the other entries of that row by zero. It turns out, however, that there is a straightforward way of dealing with this problem. We shall illustrate it in the continued discussion of this method in the context of matrix inversion in Chapter 4.

[9] In general, the determinant of an upper triangular matrix (i.e., a square matrix, all of whose elements below the main diagonal are zero) is given by the product of its main diagonal elements. Similar remarks pertain to the determinant of a lower triangular matrix (i.e., a square matrix, all of whose elements above the main diagonal are zero). The pivot procedure produces a derived triangular matrix via transformation.

In summary, our discussion of determinants does not end here. Since determinants figure quite prominently in other topics such as matrix inversion and matrix rank, we shall return to further discussion of them in subsequent chapters.

2.8 APPLYING MATRIX OPERATIONS TO STATISTICAL DATA

Much of the foregoing discussion has been introduced for a specific purpose, namely, to describe matrix and vector operations that are relevant for multivariate procedures. One of the main virtues of matrix algebra is its conciseness, that is, the succinct way in which many statistical operations can be described.

To illustrate the compactness of matrix formulation, consider the artificial data of Table 2.3. For ease of comparison these are the same data that appeared in the sample problem of Table 1.2 in Chapter 1. That is, Y denotes the employee's number of days

TABLE 2.3

Computing Various Types of Cross-Product Matrices from Sample Data

Employee	Y	Y^2	X_1	X_1^2	X_2	X_2^2	YX_1	YX_2	X_1X_2
a	1	1	1	1	1	1	1	1	1
b	0	0	2	4	1	1	0	0	2
c	1	1	2	4	2	4	2	2	4
d	4	16	3	9	2	4	12	8	6
e	3	9	5	25	4	16	15	12	20
f	2	4	5	25	6	36	10	12	30
g	5	25	6	36	5	25	30	25	30
h	6	36	7	49	4	16	42	24	28
i	9	81	10	100	8	64	90	72	80
j	13	169	11	121	7	49	143	91	77
k	15	255	11	121	9	81	165	135	99
l	16	256	12	144	10	100	192	160	120
	75	823	75	639	59	397	702	542	497

Raw cross-product matrix

$$\mathbf{B} = \begin{array}{c} Y \\ X_1 \\ X^2 \end{array} \overset{\begin{array}{ccc} Y & X_1 & X_2 \end{array}}{\begin{bmatrix} 823 & 702 & 542 \\ 702 & 639 & 497 \\ 542 & 497 & 397 \end{bmatrix}}$$

SSCP matrix

$$\mathbf{S} = \begin{array}{c} Y \\ X_1 \\ X_2 \end{array} \overset{\begin{array}{ccc} Y & X_1 & X_2 \end{array}}{\begin{bmatrix} 354.25 & 233.25 & 173.25 \\ 233.25 & 170.25 & 128.25 \\ 173.25 & 128.25 & 106.92 \end{bmatrix}}$$

Covariance matrix

$$\mathbf{C} = \begin{array}{c} Y \\ X_1 \\ X_2 \end{array} \overset{\begin{array}{ccc} Y & X_1 & X_2 \end{array}}{\begin{bmatrix} 29.52 & 19.44 & 14.44 \\ 19.44 & 14.19 & 10.69 \\ 14.44 & 10.69 & 8.91 \end{bmatrix}}$$

Correlation matrix

$$\mathbf{R} = \begin{array}{c} Y \\ X_1 \\ X_2 \end{array} \overset{\begin{array}{ccc} Y & X_1 & X_2 \end{array}}{\begin{bmatrix} 1.00 & 0.95 & 0.89 \\ 0.95 & 1.00 & 0.95 \\ 0.89 & 0.95 & 1.00 \end{bmatrix}}$$

absent during the past year; X_1 denotes his attitude rating (the higher the score the less favorable his attitude toward the firm); and X_2 denotes the number of years he has been employed by the firm.

As recalled from Chapter 1, this miniature data bank will be used later on in the book to demonstrate several multivariate techniques, including multiple regression, principal components analysis, and multiple discriminant analysis. For the moment, however, let us consider the role of matrix algebra in the development of data summaries *prior* to employing specific analytical techniques.

The computation of means, variances, covariances, correlations, etc., is a necessary preliminary to subsequent multivariate analyses in addition to being useful in its own right as a way to summarize aspects of variation in the data.

2.8.1 Sums, Sums of Squares, and Cross Products

To demonstrate the compactness of matrix notation, suppose we are concerned with computing the usual sums, sums of squares, and sums of cross products of the "raw" scores involving, for example, Y and X_1 in Table 2.3:

$$\Sigma Y; \quad \Sigma X_1; \quad \Sigma Y^2; \quad \Sigma X_1{}^2; \quad \Sigma YX_1$$

In scalar products form, the first two expressions are simply

$$\Sigma Y = \mathbf{1}'\mathbf{y} = 75; \qquad \Sigma X_1 = \mathbf{1}'\mathbf{x}_1 = 75$$

where $\mathbf{1}'$ is a 1×12 unit vector, with all entries unity, and $\mathbf{y}$ and $\mathbf{x}_1$ are the Y and X observations expressed as vectors. Notice in each case that a scalar product of two vectors is involved.

Similarly, the scalar product notion can be employed to compute three other quantities involving Y and X_1:

$$\Sigma Y^2 = \mathbf{y}'\mathbf{y} = 823 \qquad \Sigma X_1{}^2 = \mathbf{x}'_1\mathbf{x}_1 = 639 \qquad \Sigma YX_1 = \mathbf{y}'\mathbf{x}_1 = 702$$

Table 2.3 lists the numerical values for all of these products and, in addition, the products involving X_2 as well.

As a matter of fact, if we designate the matrix $\mathbf{A}$ to be the 12×3 matrix of original data involving variables Y, X_1, and X_2, the following expression

$$\boxed{\mathbf{B} = \mathbf{A}'\mathbf{A}}$$

which is often called the *minor product moment* (of $\mathbf{A}$), will yield a symmetric matrix $\mathbf{B}$ of order 3×3. The diagonal entries of the matrix $\mathbf{B}$ denote the raw sums of squares of each variable, and the off-diagonal elements denote the raw sums of cross products as shown in Table 2.3.

2.8.2 Mean-Corrected (SSCP) Matrix

We can also express the sums of squares and cross products as deviations about the means of Y, X_1, and X_2. The mean-corrected sums of squares and cross-products matrix is often more simply called the SSCP (sums of squares and cross products) matrix and is expressed in matrix notation as

$$\boxed{\mathbf{S} = \mathbf{A}'\mathbf{A} - \frac{1}{m}(\mathbf{A}'\mathbf{1})(\mathbf{1}'\mathbf{A})}$$

where **1** denotes a 12 x 1 unit vector and m denotes the number of observations; $m = 12$. The last term on the right-hand side of the equation represents the correction term and is a generalization of the usual scalar formula for computing sums of squares about the mean:

$$\Sigma x^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{m}$$

where $x = X - \bar{X}$; that is, where x denotes deviation-from-mean form. Alternatively, if the columnar means are subtracted out of **A** to begin with, yielding the mean-corrected matrix $\mathbf{A}_d$, then

$$\boxed{\mathbf{S} = \mathbf{A}_d'\mathbf{A}_d}$$

For example, the mean-corrected sums of squares and cross products for Y and X_1 are

$$\Sigma y^2 = \Sigma Y^2 - \frac{(\Sigma Y)^2}{m} = 823 - \frac{(75)^2}{12} = 354.25$$

$$\Sigma x_1{}^2 = \Sigma X_1{}^2 - \frac{(\Sigma X_1)^2}{m} = 639 - \frac{(75)^2}{12} = 170.25$$

$$\Sigma y x_1 = \Sigma Y X_1 - \frac{(\Sigma Y \Sigma X_1)}{m} = 702 - \frac{(75 \times 75)}{12} = 233.25$$

The SSCP matrix for all three variables appears in Table 2.3.

2.8.3 Covariance and Correlation Matrices

The covariance matrix, shown in Table 2.3, is obtained from the (mean-corrected) SSCP matrix by simply dividing each entry of **S** by the scalar m, the sample size. That is,

$$\boxed{\mathbf{C} = \frac{1}{m}\mathbf{S}}$$

In summational form the off-diagonal elements of **C** can be illustrated for the variables Y and X_1 by the notation

$$\text{cov}(YX_1) = \Sigma y x_1 / m = 233.25/12 = 19.44$$

Note that a covariance, then, is merely an averaged cross product of mean-corrected scores. The diagonals of **C** are, of course, variances; for example,

$$s_y{}^2 = \Sigma y^2 / m$$

(In some applications we may wish to obtain an unbiased estimate of the population covariance matrix; if so, we use the divisor $m - 1$ instead of m).

The *correlation* between two variables, y and x_1, is often obtained as

$$r_{yx} = \frac{\Sigma y x_1}{\sqrt{\Sigma y^2}\sqrt{\Sigma x_1{}^2}}$$

where y and x_1 are each expressed in deviation-from-mean form (as noted above).

Not surprisingly, **R** the correlation matrix is related to **S**, the SSCP matrix, and **C**, the covariance matrix. For example, let us return to **S**. The entries on the main diagonal of **S** represent mean-corrected sums of squares of the three variables Y, X_1, and X_2.

If we take the square roots of these three entries and enter the reciprocals of these square roots in a diagonal matrix, we have

$$\mathbf{D} = \begin{bmatrix} 1/\sqrt{\Sigma y^2} & 0 & 0 \\ 0 & 1/\sqrt{\Sigma x_1{}^2} & 0 \\ 0 & 0 & 1/\sqrt{\Sigma x_2{}^2} \end{bmatrix}$$

Then, by pre- and postmultiplying **S** by **D** we can obtain the correlation matrix **R**.

$$\mathbf{R} = \mathbf{DSD}$$

$$\mathbf{R} = \begin{bmatrix} \dfrac{\Sigma y^2}{\sqrt{\Sigma y^2}\sqrt{\Sigma y^2}} & \dfrac{\Sigma y x_1}{\sqrt{\Sigma y^2}\sqrt{\Sigma x_1{}^2}} & \dfrac{\Sigma y x_2}{\sqrt{\Sigma y^2}\sqrt{\Sigma x_2{}^2}} \\ \dfrac{\Sigma y x_1}{\sqrt{\Sigma y^2}\sqrt{\Sigma x_1{}^2}} & \dfrac{\Sigma x_1{}^2}{\sqrt{\Sigma x_1{}^2}\sqrt{\Sigma x_1{}^2}} & \dfrac{\Sigma x_1 x_2}{\sqrt{\Sigma x_1{}^2}\sqrt{\Sigma x_2{}^2}} \\ \dfrac{\Sigma y x_2}{\sqrt{\Sigma y^2}\sqrt{\Sigma x_2{}^2}} & \dfrac{\Sigma x_1 x_2}{\sqrt{\Sigma x_1{}^2}\sqrt{\Sigma x_2{}^2}} & \dfrac{\Sigma x_2{}^2}{\sqrt{\Sigma x_2{}^2}\sqrt{\Sigma x_2{}^2}} \end{bmatrix}$$

The above matrix is the derived matrix of correlations between each pair of variables and is also shown in Table 2.3

Ordinarily, we could then go on and use **R** in further calculation, for example, to find the regression of Y on X_1 and X_2. Since our purpose here is only to show the conciseness of matrix notation, we defer these additional steps until later. In future chapters we shall have occasion to discuss all four of the preceding matrices: (a) the raw sums and cross-products matrix, (b) the (mean-corrected) SSCP matrix, (c) the covariance matrix, and (d) the correlation matrix.

At this point, however, we should note that they are all variations on a common theme: All involve computing the minor product moment of some matrix.

1. Raw sums of squares and cross-products matrix:

$$\mathbf{B} = \mathbf{A}'\mathbf{A}$$

2. The (mean-corrected) SSCP matrix:

$$\mathbf{S} = \mathbf{A}_d'\mathbf{A}_d$$

where $\mathbf{A}_d$ is the matrix of deviation-from-mean scores; that is, each column of $\mathbf{A}_d$ sums to zero since each columnar mean has been subtracted from each datum.

3. The covariance matrix

$$\mathbf{C} = 1/m\,\mathbf{A}_d'\mathbf{A}_d$$

4. The correlation matrix

$$\mathbf{R} = 1/m\,\mathbf{A}_s'\mathbf{A}_s$$

where $\mathbf{A}_s$ is the matrix of standardized scores.

As can be found from Table 1.2, in which the sample problem data first appear, both deviation-from-mean and standardized scores are shown along with the original scores.

Finally, the matrices $\mathbf{A}_d$ of mean-corrected scores and $\mathbf{A}_s$ of standardized scores are derived from $\mathbf{A}$, the matrix of original scores, in the following way. We first find

$$\mathbf{A}_d = \mathbf{A} - \mathbf{1}\bar{\mathbf{a}}'$$

where $\mathbf{1}$ is a 12×1 unit column vector and $\bar{\mathbf{a}}'$ is a 1×3 row vector of variable means. The vector of means is, itself, obtained from

$$\bar{\mathbf{a}}' = \mathbf{1}'\mathbf{A}/m$$

where $\mathbf{1}'$ is now a 1×12 row vector. Next, we find the matrix of standardized scores from $\mathbf{A}_d$ as follows:

$$\mathbf{A}_s = \mathbf{A}_d\mathbf{D}$$

where $\mathbf{D}$ is a diagonal matrix whose entries along the main diagonal are the reciprocals of the standard deviations of the variables in $\mathbf{A}$.

The standard deviation of any column of $\mathbf{A}_d$, say $\mathbf{a}_{dj}$, is simply

$$s_{a_{dj}} = \sqrt{\mathbf{a}'_{dj}\mathbf{a}_{dj}/m}$$

In summary, any of the operations needed to find various cross-product matrices are readily expressible in matrix format. In so doing we arrive at a very compact and graceful way to portray some otherwise cumbersome operations.

2.9 SUMMARY

The purpose of this chapter has been to introduce the reader to relations and operations on vectors and matrices. Our emphasis has been on defining various operations and describing the mechanics by which one manipulates vectors and matrices. Such elementary operations as addition and subtraction, multiplication of vectors and matrices by scalars, the scalar product of two vectors, vector times matrix multiplication, etc., were described and illustrated numerically. Various properties of these operations were also described.

Special kinds of vectors (e.g., null, sign, unit) and special kinds of matrices (e.g., diagonal, scalar, identity) were also defined and illustrated numerically. We then turned to an introductory discussion of determinants of (square) matrices. Evaluation of determinants via expansion by cofactors and the pivotal method was described and illustrated.

We concluded the chapter with a demonstration of how matrix algebra can be used to provide concise descriptions of various statistical operations that are preparatory to specific multivariate analyses. These matrix operations are particularly amenable to computer programming and are used extensively in programs that deal with multivariate procedures.

REVIEW QUESTIONS

1. Write the following equations in matrix form:

a. $4x + y - z = 0$ b. $2x + 3y + z = 11$

$3x - 4y + 2z = 1$ $x + y + 7z = 24$

$5x - y - 2z = 7$ $3x + 5y + 4z = 25$

2. Given the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 \\ 4 & 0 & 1 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 2 & 0 \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} 0 & 1 & 0 \\ 4 & -1 & -2 \end{bmatrix}$$

find

a. $\mathbf{A} + \mathbf{B}$ b. $(\mathbf{A} + \mathbf{C}) + \mathbf{B}$ c. $\mathbf{A} + (\mathbf{B} + \mathbf{C})$

d. $\mathbf{A} - (\mathbf{B} + \mathbf{C})$ e. $-(\mathbf{A} + \mathbf{B})$ f. $(\mathbf{A} - \mathbf{B}) + \mathbf{C}$

3. Given the vectors

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$$

and the scalars

$$k_1 = 2; \quad k_2 = 5$$

find

a. $\mathbf{b}'\mathbf{b}$ b. $-k_1\mathbf{a}$ c. $k_2\mathbf{b}'$ d. $\mathbf{a}'\mathbf{b}$ e. $k_1k_2(\mathbf{a}'\mathbf{a})$ f. $\frac{1}{k_1}(\mathbf{b}'\mathbf{a})$

4. Given the matrices, vectors, and scalars of Problems 2 and 3, find

a. $\mathbf{a}'\mathbf{A}'$ b. $k_1\mathbf{B}$ c. $(\mathbf{A}\mathbf{B}')'$ d. $k_1\mathbf{C}$ e. $k_2\mathbf{B}\mathbf{A}'\mathbf{C}$ f. $\mathbf{a}\mathbf{b}'$

5. Examine the relationships among $(\mathbf{DE})'$, $\mathbf{D}'\mathbf{E}'$, and $\mathbf{E}'\mathbf{D}'$ under the following two sets of conditions.

Let:

a. $$\mathbf{D} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad \mathbf{E} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

Let:

b. $$\mathbf{D} = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{E} = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}.$$

6. Given the matrices

$$\mathbf{F} = \begin{bmatrix} -1 & 3 & 5 \\ 1 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix} \quad \text{and} \quad \mathbf{G} = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix}$$

demonstrate that

a. $\mathbf{FG} = \phi$ does not imply that either $\mathbf{F} = \phi$ or $\mathbf{G} = \phi$.
b. Find $\mathbf{GF}$. Is this product equal to ϕ?

7. Given the matrices and vectors in Problems 3 and 6, find the products:

a. $\mathbf{a}'\mathbf{Fb}$ b. $\mathbf{b}'\mathbf{Gb}$ c. $\mathbf{a}'\mathbf{FGa}$ d. $\mathbf{b}'\mathbf{FGa}$

8. Consider the diagonal matrices

$$\mathbf{H}_1 = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{H}_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

and the vectors and matrices of Problems 3 and 6. Find

a. $\mathbf{a}'(\mathbf{H}_1\mathbf{F})$ b. $\mathbf{b}'(\mathbf{H}_1\mathbf{G}\mathbf{H}_2)$ c. $\mathbf{a}'(\mathbf{H}_1\mathbf{H}_2)\mathbf{b}$ d. $\mathbf{a}'(\mathbf{FGH}_2)$

9. In ordinary algebra, we have the relationship

$$x^2 - x - 2 = (x + 1)(x - 2)$$

In matrix algebra, if

$$\mathbf{X} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

see if the following holds:

$$\mathbf{X}^2 - \mathbf{X} - 2\mathbf{I} = (\mathbf{X} + \mathbf{I})(\mathbf{X} - 2\mathbf{I})$$

10. If

$$\mathbf{J} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{K} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Find

a. $\mathbf{J}^2$ b. $\mathbf{K}^2$ c. $(\mathbf{JK})^2$ d. $(\mathbf{KJ})^2 + (\mathbf{JK})'$

11. Evaluate the determinants of the following 2 x 2 matrices:

a. $\mathbf{L}_1 = \begin{bmatrix} x^2 & x \\ x^4 & x^3 \end{bmatrix}$ b. $\mathbf{L}_2 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}$

c. $\mathbf{L}_3 = \begin{bmatrix} 1/2 & 1/3 \\ 1/4 & 1/6 \end{bmatrix}$ d. $\mathbf{L}_4 = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

12. By means of cofactor expansion, evaluate the determinants of

a. $\mathbf{M}_1 = \begin{bmatrix} 4 & -12 & -4 \\ 2 & 1 & 3 \\ -1 & -3 & 2 \end{bmatrix}$

b. $\mathbf{M}_2 = \begin{bmatrix} 0 & 3 & 5 \\ 2 & 6 & 7 \\ 4 & 1 & 1 \end{bmatrix}$

c. $\mathbf{M}_3 = \begin{bmatrix} 1 & 5 & 2 & 1 \\ 3 & 7 & 4 & 5 \\ 2 & 9 & 1 & 2 \\ 4 & 0 & 1 & 3 \end{bmatrix}$

13. Evaluate the determinant of the fourth-order matrix used in Section 2.7.4 via cofactor expansion and check to see that it equals the value of the determinant found from the pivotal method.

14. Apply the pivotal method to matrix $\mathbf{M}_3$ in Problem 12 and check your answer with that found by cofactor expansion.

15. Assume the following data bank:

	Y	X_1	X_2	X_3
a	2	1	0	9
b	4	2	3	8
c	3	5	2	4
d	7	3	4	5
e	8	7	7	2
f	9	8	7	1

Find, via matrix methods,

a. ΣY; $\bar{X}_1$; ΣYX_2; $\Sigma X_3^2 - (\Sigma X_3)^2/m$

b. the 4 x 4 (mean-corrected) SSCP matrix **S**

c. the covariance matrix **C** d. the correlation matrix **R**

e. the matrix of mean-corrected scores

f. show that the sum of the deviations about the mean equals zero for the first column Y.

Answers to Numerical Problems

CHAPTER 2

2.1a $\begin{bmatrix} 4 & 1 & -1 \\ 3 & -4 & 2 \\ 5 & -1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 7 \end{bmatrix}$ 2.1b $\begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 7 \\ 3 & 5 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 11 \\ 24 \\ 25 \end{bmatrix}$

2.2a $\begin{bmatrix} 3 & 5 & 1 \\ 3 & 2 & 1 \end{bmatrix}$ 2.2b $\begin{bmatrix} 3 & 6 & 1 \\ 7 & 1 & -1 \end{bmatrix}$ 2.2c $\begin{bmatrix} 3 & 6 & 1 \\ 7 & 1 & -1 \end{bmatrix}$

2.2d $\begin{bmatrix} -1 & -2 & -7 \\ 1 & -1 & 3 \end{bmatrix}$ 2.2e $\begin{bmatrix} -3 & -5 & -1 \\ -3 & -2 & -1 \end{bmatrix}$ 2.2f $\begin{bmatrix} -1 & 0 & -7 \\ 9 & -3 & -1 \end{bmatrix}$

2.3a 26 2.3b $\begin{bmatrix} -2 \\ -4 \\ -8 \end{bmatrix}$ 2.3c (5, 15, 20) 2.3d 23 2.3e 210

2.3f 23/2

2.4a (−7, 8) 2.4b $\begin{bmatrix} 4 & 6 & 8 \\ -2 & 4 & 0 \end{bmatrix}$ 2.4c $\begin{bmatrix} -4 & 12 \\ 3 & -4 \end{bmatrix}$

2.4d $\begin{bmatrix} 0 & 2 & 0 \\ 8 & -2 & -4 \end{bmatrix}$ 2.4e $\begin{bmatrix} 240 & -80 & -120 \\ -80 & 35 & 40 \end{bmatrix}$ 2.4f $\begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 8 \\ 4 & 12 & 16 \end{bmatrix}$

2.5a

(i) $(\mathbf{DE})' = \mathbf{E}'\mathbf{D}' = \begin{bmatrix} ea + gb & ec + gd \\ fa + hb & fc + hd \end{bmatrix}$

(ii) $\mathbf{D}'\mathbf{E}' = \begin{bmatrix} ae + cf & ag + ch \\ be + df & bg + dh \end{bmatrix}$ (iii) $\mathbf{E}'\mathbf{D}'$ = See (i) above.

2.5b

(i) $(\mathbf{DE})' = \mathbf{E}'\mathbf{D}' = \begin{bmatrix} 3 & 0 \\ 10 & 4 \end{bmatrix}$ (ii) $\mathbf{D}'\mathbf{E}' = \begin{bmatrix} 3 & 0 \\ 17 & 4 \end{bmatrix}$

(iii) $\mathbf{E}'\mathbf{D}'$ = See (i) above.

2.6a $\mathbf{F} \neq \phi$; $\mathbf{G} \neq \phi$; $\mathbf{FG} = \phi$ 2.6b $\mathbf{GF} = \begin{bmatrix} -10 & 30 & 50 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq \phi$

2.7a 84 2.7b −30 2.7c 0 **2.7d** 0

2.8a (−2, 6, 10) **2.8b** (−3, 0, 18) **2.8c** 28 **2.8d** (0, 0, 0)

2.9 $\mathbf{X}^2 - \mathbf{X} - 2\mathbf{I} = (\mathbf{X} + \mathbf{I})(\mathbf{X} - 2\mathbf{I}) = \begin{bmatrix} a^2 + bc - a - 2 & ab + bd - b \\ ac + cd - c & bc + d^2 - d - 2 \end{bmatrix}$

2.10a $\begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix}$ **2.10b** $\begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$ **2.10c** $\begin{bmatrix} 28 & 66 \\ 44 & 105 \end{bmatrix}$ **2.10d** $\begin{bmatrix} 30 & 48 \\ 72 & 114 \end{bmatrix}$

2.11a 0 **2.11b** 0 **2.11c** 0 **2.11d** $a^2 + b^2$

2.12a 148 **2.12b** −32 **2.12c** −128

2.13 $|\mathbf{A}| = 3(10) + 2 - 17 = 15$

2.14 $|\mathbf{M}| = (1)(-8)(-2.75)(-5.8182) = -128$

2.15a $\Sigma Y = 33$ $\bar{X}_1 = 4.33$ $\Sigma Y X_2 = 165$ $\Sigma X_3{}^2 - (\Sigma X_3)^2/m = 50.83$

2.15b

$$\mathbf{S} = \begin{bmatrix} 41.5 & 31 & 38 & -37.5 \\ 31 & 39.3 & 33.3 & -43.7 \\ 38.5 & 33.3 & 38.3 & -38.2 \\ -37.5 & -43.7 & -38.2 & 50.8 \end{bmatrix}$$

2.15c

$$\mathbf{C} = \begin{bmatrix} 6.9 & 5.2 & 6.3 & -6.3 \\ 5.2 & 6.6 & 5.6 & -7.3 \\ 6.3 & 5.6 & 6.4 & -6.4 \\ -6.3 & -7.3 & -6.4 & 8.5 \end{bmatrix}$$

2.15d

$$\mathbf{R} = \begin{bmatrix} 1 & 0.76 & 0.96 & -0.82 \\ 0.76 & 1 & 0.85 & -0.98 \\ 0.96 & 0.85 & 1 & -0.86 \\ -0.82 & -0.98 & -0.86 & 1 \end{bmatrix}$$

2.15e

$$\mathbf{A_d} = \begin{bmatrix} -3.5 & -3.3 & -3.8 & 4.2 \\ -1.5 & -2.3 & -0.8 & 3.2 \\ -2.5 & 0.7 & -1.8 & -0.8 \\ 1.5 & -1.3 & 0.2 & 0.2 \\ 2.5 & 2.7 & 3.2 & -2.8 \\ 3.5 & 3.7 & 3.2 & -3.8 \end{bmatrix}$$

2.15f $-3.5 - 1.5 - 2.5 + 1.5 + 2.5 + 3.5 = 0$