Methods of Multivariate Analysis. Alvin C. Rencher
Copyright © 2002 John Wiley & Sons, Inc.
ISBN: 0-471-41889-7

CHAPTER 2

Matrix Algebra

2.1 INTRODUCTION

This chapter introduces the basic elements of matrix algebra used in the remainder of this book. It is essentially a review of the requisite matrix tools and is not intended to be a complete development. However, it is sufficiently self-contained so that those with no previous exposure to the subject should need no other reference. Anyone unfamiliar with matrix algebra should plan to work most of the problems entailing numerical illustrations. It would also be helpful to explore some of the problems involving general matrix manipulation.

With the exception of a few derivations that seemed instructive, most of the results are given without proof. Some additional proofs are requested in the problems. For the remaining proofs, see any general text on matrix theory or one of the specialized matrix texts oriented to statistics, such as Graybill (1969), Searle (1982), or Harville (1997).

2.2 NOTATION AND BASIC DEFINITIONS

2.2.1 Matrices, Vectors, and Scalars

A *matrix* is a rectangular or square array of numbers or variables arranged in rows and columns. We use uppercase boldface letters to represent matrices. All entries in matrices will be real numbers or variables representing real numbers. The elements of a matrix are displayed in brackets. For example, the ACT score and GPA for three students can be conveniently listed in the following matrix:

$$\mathbf{A} = \begin{pmatrix} 23 & 3.54 \\ 29 & 3.81 \\ 18 & 2.75 \end{pmatrix}. \tag{2.1}$$

The elements of $\mathbf{A}$ can also be variables, representing possible values of ACT and GPA for three students:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}. \tag{2.2}$$

In this double-subscript notation for the elements of a matrix, the first subscript indicates the row; the second identifies the column. The matrix $\mathbf{A}$ in (2.2) can also be expressed as

$$\mathbf{A} = (a_{ij}), \tag{2.3}$$

where a_{ij} is a general element.

With three rows and two columns, the matrix $\mathbf{A}$ in (2.1) or (2.2) is said to be 3×2. In general, if a matrix $\mathbf{A}$ has n rows and p columns, it is said to be $n \times p$. Alternatively, we say the *size* of $\mathbf{A}$ is $n \times p$.

A *vector* is a matrix with a single column or row. The following could be the test scores of a student in a course in multivariate analysis:

$$\mathbf{x} = \begin{pmatrix} 98 \\ 86 \\ 93 \\ 97 \end{pmatrix}. \tag{2.4}$$

Variable elements in a vector can be identified by a single subscript:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \tag{2.5}$$

We use lowercase boldface letters for column vectors. Row vectors are expressed as

$$\mathbf{x}' = (x_1, x_2, x_3, x_4) \quad \text{or as} \quad \mathbf{x}' = (x_1 \quad x_2 \quad x_3 \quad x_4),$$

where $\mathbf{x}'$ indicates the *transpose* of $\mathbf{x}$. The transpose operation is defined in Section 2.2.3.

Geometrically, a vector with p elements identifies a point in a p-dimensional space. The elements in the vector are the coordinates of the point. In (2.35) in Section 2.3.3, we define the distance from the origin to the point. In Section 3.12, we define the distance between two vectors. In some cases, we will be interested in a directed line segment or arrow from the origin to the point.

A single real number is called a *scalar*, to distinguish it from a vector or matrix. Thus 2, -4, and 125 are scalars. A variable representing a scalar is usually denoted by a lowercase nonbolded letter, such as $a = 5$. A product involving vectors and matrices may reduce to a matrix of size 1×1, which then becomes a scalar.

2.2.2 Equality of Vectors and Matrices

Two matrices are equal if they are the same size and the elements in corresponding positions are equal. Thus if $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$, then $\mathbf{A} = \mathbf{B}$ if $a_{ij} = b_{ij}$ for all i and j. For example, let

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & 4 \\ 1 & 3 & 7 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 3 & 1 \\ -2 & 3 \\ 4 & 7 \end{pmatrix},$$
$$\mathbf{C} = \begin{pmatrix} 3 & -2 & 4 \\ 1 & 3 & 7 \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} 3 & -2 & 4 \\ 1 & 3 & 6 \end{pmatrix}.$$

Then $\mathbf{A} = \mathbf{C}$. But even though $\mathbf{A}$ and $\mathbf{B}$ have the same elements, $\mathbf{A} \neq \mathbf{B}$ because the two matrices are not the same size. Likewise, $\mathbf{A} \neq \mathbf{D}$ because $a_{23} \neq d_{23}$. Thus two matrices of the same size are unequal if they differ in a single position.

2.2.3 Transpose and Symmetric Matrices

The *transpose* of a matrix $\mathbf{A}$, denoted by $\mathbf{A}'$, is obtained from $\mathbf{A}$ by interchanging rows and columns. Thus the columns of $\mathbf{A}'$ are the rows of $\mathbf{A}$, and the rows of $\mathbf{A}'$ are the columns of $\mathbf{A}$. The following examples illustrate the transpose of a matrix or vector:

$$\mathbf{A} = \begin{pmatrix} -5 & 2 & 4 \\ 3 & 6 & -2 \end{pmatrix}, \qquad \mathbf{A}' = \begin{pmatrix} -5 & 3 \\ 2 & 6 \\ 4 & -2 \end{pmatrix},$$
$$\mathbf{B} = \begin{pmatrix} 2 & -3 \\ 4 & 1 \end{pmatrix}, \qquad \mathbf{B}' = \begin{pmatrix} 2 & 4 \\ -3 & 1 \end{pmatrix},$$
$$\mathbf{a} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}, \qquad \mathbf{a}' = (2, \quad -3, \quad 1).$$

The transpose operation does not change a scalar, since it has only one row and one column.

If the transpose operator is applied twice to any matrix, the result is the original matrix:

$$(\mathbf{A}')' = \mathbf{A}. \tag{2.6}$$

If the transpose of a matrix is the same as the original matrix, the matrix is said to be *symmetric*; that is, $\mathbf{A}$ is symmetric if $\mathbf{A} = \mathbf{A}'$. For example,

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 10 & -7 \\ 4 & -7 & 9 \end{pmatrix}, \qquad \mathbf{A}' = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 10 & -7 \\ 4 & -7 & 9 \end{pmatrix}.$$

Clearly, all symmetric matrices are square.

2.2.4 Special Matrices

The *diagonal* of a $p \times p$ square matrix $\mathbf{A}$ consists of the elements $a_{11}, a_{22}, \ldots, a_{pp}$. For example, in the matrix

$$\mathbf{A} = \begin{pmatrix} 5 & -2 & 4 \\ 7 & 9 & 3 \\ -6 & 8 & 1 \end{pmatrix},$$

the elements 5, 9, and 1 lie on the diagonal. If a matrix contains zeros in all off-diagonal positions, it is said to be a *diagonal matrix*. An example of a diagonal matrix is

$$\mathbf{D} = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}.$$

This matrix can also be denoted as

$$\mathbf{D} = \operatorname{diag}(10, -3, 0, 7). \tag{2.7}$$

A diagonal matrix can be formed from any square matrix by replacing off-diagonal elements by 0's. This is denoted by diag($\mathbf{A}$). Thus for the preceding matrix $\mathbf{A}$, we have

$$\operatorname{diag}(\mathbf{A}) = \operatorname{diag}\begin{pmatrix} 5 & -2 & 4 \\ 7 & 9 & 3 \\ -6 & 8 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.8}$$

A diagonal matrix with a 1 in each diagonal position is called an *identity* matrix and is denoted by $\mathbf{I}$. For example, a 3×3 identity matrix is given by

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.9}$$

An *upper triangular matrix* is a square matrix with zeros below the diagonal, such as

$$\mathbf{T} = \begin{pmatrix} 8 & 3 & 4 & 7 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 6 \end{pmatrix}. \tag{2.10}$$

A *lower triangular matrix* is defined similarly.

A vector of 1's is denoted by $\mathbf{j}$:

$$\mathbf{j} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{2.11}$$

A square matrix of 1's is denoted by $\mathbf{J}$. For example, a 3×3 matrix $\mathbf{J}$ is given by

$$\mathbf{J} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \tag{2.12}$$

Finally, we denote a vector of zeros by $\mathbf{0}$ and a matrix of zeros by $\mathbf{O}$. For example,

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{O} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.13}$$

2.3 OPERATIONS

2.3.1 Summation and Product Notation

For completeness, we review the standard mathematical notation for sums and products. The sum of a sequence of numbers $a_1, a_2, \ldots, a_n$ is indicated by

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n.$$

If the n numbers are all the same, then $\sum_{i=1}^{n} a = a + a + \cdots + a = na$. The sum of all the numbers in an array with double subscripts, such as

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23}, \end{matrix}$$

is indicated by

$$\sum_{i=1}^{2}\sum_{j=1}^{3} a_{ij} = a_{11} + a_{12} + a_{13} + a_{21} + a_{22} + a_{23}.$$

This is sometimes abbreviated to

$$\sum_{i=1}^{2}\sum_{j=1}^{3} a_{ij} = \sum_{ij} a_{ij}.$$

The product of a sequence of numbers $a_1, a_2, \ldots, a_n$ is indicated by

$$\prod_{i=1}^{n} a_i = (a_1)(a_2)\cdots(a_n).$$

If the n numbers are all equal, the product becomes $\prod_{i=1}^{n} a = (a)(a)\cdots(a) = a^n$.

2.3.2 Addition of Matrices and Vectors

If two matrices (or two vectors) are the same size, their *sum* is found by adding corresponding elements; that is, if $\mathbf{A}$ is $n \times p$ and $\mathbf{B}$ is $n \times p$, then $\mathbf{C} = \mathbf{A} + \mathbf{B}$ is also $n \times p$ and is found as $(c_{ij}) = (a_{ij} + b_{ij})$. For example,

$$\begin{pmatrix} -2 & 5 \\ 3 & 1 \\ 7 & -6 \end{pmatrix} + \begin{pmatrix} 3 & -2 \\ 4 & 5 \\ 10 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 7 & 6 \\ 17 & -9 \end{pmatrix},$$

$$\begin{pmatrix} 1 \\ 3 \\ 7 \end{pmatrix} + \begin{pmatrix} 5 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 10 \end{pmatrix}.$$

Similarly, the *difference* between two matrices or two vectors of the same size is found by subtracting corresponding elements. Thus $\mathbf{C} = \mathbf{A} - \mathbf{B}$ is found as $(c_{ij}) = (a_{ij} - b_{ij})$. For example,

$$(3 \quad 9 \quad -4) - (5 \quad -4 \quad 2) = (-2 \quad 13 \quad -6).$$

If two matrices are identical, their difference is a zero matrix; that is, $\mathbf{A} = \mathbf{B}$ implies $\mathbf{A} - \mathbf{B} = \mathbf{O}$. For example,

$$\begin{pmatrix} 3 & -2 & 4 \\ 6 & 7 & 5 \end{pmatrix} - \begin{pmatrix} 3 & -2 & 4 \\ 6 & 7 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Matrix addition is commutative:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}. \tag{2.14}$$

The transpose of the sum (difference) of two matrices is the sum (difference) of the transposes:

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}', \tag{2.15}$$

$$(\mathbf{A} - \mathbf{B})' = \mathbf{A}' - \mathbf{B}', \tag{2.16}$$

$$(\mathbf{x} + \mathbf{y})' = \mathbf{x}' + \mathbf{y}', \tag{2.17}$$

$$(\mathbf{x} - \mathbf{y})' = \mathbf{x}' - \mathbf{y}'. \tag{2.18}$$

2.3.3 Multiplication of Matrices and Vectors

In order for the product **AB** to be defined, the number of columns in **A** must be the same as the number of rows in **B**, in which case **A** and **B** are said to be *conformable*. Then the (ij)th element of $\mathbf{C} = \mathbf{AB}$ is

$$c_{ij} = \sum_k a_{ik} b_{kj}. \tag{2.19}$$

Thus c_{ij} is the sum of products of the ith row of **A** and the jth column of **B**. We therefore multiply each row of **A** by each column of **B**, and the size of **AB** consists of the number of rows of **A** and the number of columns of **B**. Thus, if **A** is $n \times m$ and **B** is $m \times p$, then $\mathbf{C} = \mathbf{AB}$ is $n \times p$. For example, if

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

then

$$\mathbf{C} = \mathbf{AB} = \begin{pmatrix} 2 \cdot 1 + 1 \cdot 2 + 3 \cdot 3 & 2 \cdot 4 + 1 \cdot 6 + 3 \cdot 8 \\ 4 \cdot 1 + 6 \cdot 2 + 5 \cdot 3 & 4 \cdot 4 + 6 \cdot 6 + 5 \cdot 8 \\ 7 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 & 7 \cdot 4 + 2 \cdot 6 + 3 \cdot 8 \\ 1 \cdot 1 + 3 \cdot 2 + 2 \cdot 3 & 1 \cdot 4 + 3 \cdot 6 + 2 \cdot 8 \end{pmatrix}$$

$$= \begin{pmatrix} 13 & 38 \\ 31 & 92 \\ 20 & 64 \\ 13 & 38 \end{pmatrix}.$$

Note that **A** is 4×3, **B** is 3×2, and **AB** is 4×2. In this case, **AB** is of a different size than either **A** or **B**.

If **A** and **B** are both $n \times n$, then **AB** is also $n \times n$. Clearly, $\mathbf{A}^2$ is defined only if **A** is a square matrix.

In some cases **AB** is defined, but **BA** is not defined. In the preceding example, **BA** cannot be found because **B** is 3×2 and **A** is 4×3 and a row of **B** cannot be multiplied by a column of **A**. Sometimes **AB** and **BA** are both defined but are different in size. For example, if **A** is 2×4 and **B** is 4×2, then **AB** is 2×2 and **BA** is 4×4. If **A** and **B** are square and the same size, then **AB** and **BA** are both defined. However,

$$\mathbf{AB} \neq \mathbf{BA}, \tag{2.20}$$

except for a few special cases. For example, let

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & -2 \\ 3 & 5 \end{pmatrix}.$$

Then

$$\mathbf{AB} = \begin{pmatrix} 10 & 13 \\ 14 & 16 \end{pmatrix}, \qquad \mathbf{BA} = \begin{pmatrix} -3 & -5 \\ 13 & 29 \end{pmatrix}.$$

Thus we must be careful to specify the order of multiplication. If we wish to multiply both sides of a matrix equation by a matrix, we must multiply *on the left* or *on the right* and be consistent on both sides of the equation.

Multiplication is distributive over addition or subtraction:

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}, \tag{2.21}$$

$$\mathbf{A}(\mathbf{B} - \mathbf{C}) = \mathbf{AB} - \mathbf{AC}, \tag{2.22}$$

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}, \tag{2.23}$$

$$(\mathbf{A} - \mathbf{B})\mathbf{C} = \mathbf{AC} - \mathbf{BC}. \tag{2.24}$$

Note that, in general, because of (2.20),

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) \neq \mathbf{BA} + \mathbf{CA}. \tag{2.25}$$

Using the distributive law, we can expand products such as $(\mathbf{A} - \mathbf{B})(\mathbf{C} - \mathbf{D})$ to obtain

$$\begin{aligned} (\mathbf{A} - \mathbf{B})(\mathbf{C} - \mathbf{D}) &= (\mathbf{A} - \mathbf{B})\mathbf{C} - (\mathbf{A} - \mathbf{B})\mathbf{D} \qquad &&\text{[by (2.22)]} \\ &= \mathbf{AC} - \mathbf{BC} - \mathbf{AD} + \mathbf{BD} \qquad &&\text{[by (2.24)].} \end{aligned} \tag{2.26}$$

The transpose of a product is the product of the transposes in reverse order:

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'. \tag{2.27}$$

Note that (2.27) holds as long as $\mathbf{A}$ and $\mathbf{B}$ are conformable. They need not be square.

Multiplication involving vectors follows the same rules as for matrices. Suppose $\mathbf{A}$ is $n \times p$, $\mathbf{a}$ is $p \times 1$, $\mathbf{b}$ is $p \times 1$, and $\mathbf{c}$ is $n \times 1$. Then some possible products are $\mathbf{Ab}$, $\mathbf{c}'\mathbf{A}$, $\mathbf{a}'\mathbf{b}$, $\mathbf{b}'\mathbf{a}$, and $\mathbf{ab}'$. For example, let

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & 4 \\ 1 & 3 & 5 \end{pmatrix}, \qquad \mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \qquad \mathbf{c} = \begin{pmatrix} 2 \\ -5 \end{pmatrix}.$$

Then

$$\mathbf{Ab} = \begin{pmatrix} 3 & -2 & 4 \\ 1 & 3 & 5 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 16 \\ 31 \end{pmatrix},$$

$$\mathbf{c}'\mathbf{A} = (2 \quad -5)\begin{pmatrix} 3 & -2 & 4 \\ 1 & 3 & 5 \end{pmatrix} = (1 \quad -19 \quad -17),$$

$$\mathbf{c}'\mathbf{A}\mathbf{b} = (2 \quad -5)\begin{pmatrix} 3 & -2 & 4 \\ 1 & 3 & 5 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = (2 \quad -5)\begin{pmatrix} 16 \\ 31 \end{pmatrix} = -123,$$

$$\mathbf{a}'\mathbf{b} = (1 \quad -2 \quad 3)\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = 8,$$

$$\mathbf{b}'\mathbf{a} = (2 \quad 3 \quad 4)\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = 8,$$

$$\mathbf{a}\mathbf{b}' = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}(2 \quad 3 \quad 4) = \begin{pmatrix} 2 & 3 & 4 \\ -4 & -6 & -8 \\ 6 & 9 & 12 \end{pmatrix},$$

$$\mathbf{a}\mathbf{c}' = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}(2 \quad -5) = \begin{pmatrix} 2 & -5 \\ -4 & 10 \\ 6 & -15 \end{pmatrix}.$$

Note that $\mathbf{Ab}$ is a column vector, $\mathbf{c}'\mathbf{A}$ is a row vector, $\mathbf{c}'\mathbf{Ab}$ is a scalar, and $\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}$. The triple product $\mathbf{c}'\mathbf{Ab}$ was obtained as $\mathbf{c}'(\mathbf{Ab})$. The same result would be obtained if we multiplied in the order $(\mathbf{c}'\mathbf{A})\mathbf{b}$:

$$(\mathbf{c}'\mathbf{A})\mathbf{b} = (1 \quad -19 \quad -17)\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = -123.$$

This is true in general for a triple product:

$$\mathbf{ABC} = \mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}. \tag{2.28}$$

Thus multiplication of three matrices can be defined in terms of the product of two matrices, since (fortunately) it does not matter which two are multiplied first. Note that $\mathbf{A}$ and $\mathbf{B}$ must be conformable for multiplication, and $\mathbf{B}$ and $\mathbf{C}$ must be conformable. For example, if $\mathbf{A}$ is $n \times p$, $\mathbf{B}$ is $p \times q$, and $\mathbf{C}$ is $q \times m$, then both multiplications are possible and the product $\mathbf{ABC}$ is $n \times m$.

We can sometimes factor a sum of triple products on both the right and left sides. For example,

$$\mathbf{ABC} + \mathbf{ADC} = \mathbf{A}(\mathbf{B} + \mathbf{D})\mathbf{C}. \tag{2.29}$$

As another illustration, let $\mathbf{X}$ be $n \times p$ and $\mathbf{A}$ be $n \times n$. Then

$$\mathbf{X}'\mathbf{X} - \mathbf{X}'\mathbf{AX} = \mathbf{X}'(\mathbf{X} - \mathbf{AX}) = \mathbf{X}'(\mathbf{I} - \mathbf{A})\mathbf{X}. \tag{2.30}$$

If $\mathbf{a}$ and $\mathbf{b}$ are both $n \times 1$, then

$$\mathbf{a}'\mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \tag{2.31}$$

is a sum of products and is a scalar. On the other hand, $\mathbf{ab}'$ is defined for any size $\mathbf{a}$ and $\mathbf{b}$ and is a matrix, either rectangular or square:

$$\mathbf{ab}' = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} (b_1 \quad b_2 \quad \cdots \quad b_p) = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_p \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_p \\ \vdots & \vdots & & \vdots \\ a_n b_1 & a_n b_2 & \cdots & a_n b_p \end{pmatrix}. \tag{2.32}$$

Similarly,

$$\mathbf{a}'\mathbf{a} = a_1^2 + a_2^2 + \cdots + a_n^2, \tag{2.33}$$

$$\mathbf{aa}' = \begin{pmatrix} a_1^2 & a_1 a_2 & \cdots & a_1 a_n \\ a_2 a_1 & a_2^2 & \cdots & a_2 a_n \\ \vdots & \vdots & & \vdots \\ a_n a_1 & a_n a_2 & \cdots & a_n^2 \end{pmatrix}. \tag{2.34}$$

Thus $\mathbf{a}'\mathbf{a}$ is a sum of squares, and $\mathbf{aa}'$ is a square (symmetric) matrix. The products $\mathbf{a}'\mathbf{a}$ and $\mathbf{aa}'$ are sometimes referred to as the *dot product* and *matrix product*, respectively. The square root of the sum of squares of the elements of $\mathbf{a}$ is the *distance* from the origin to the point $\mathbf{a}$ and is also referred to as the *length* of $\mathbf{a}$:

$$\text{Length of } \mathbf{a} = \sqrt{\mathbf{a}'\mathbf{a}} = \sqrt{\textstyle\sum_{i=1}^n a_i^2}. \tag{2.35}$$

As special cases of (2.33) and (2.34), note that if $\mathbf{j}$ is $n \times 1$, then

$$\mathbf{j}'\mathbf{j} = n, \qquad \mathbf{jj}' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} = \mathbf{J}, \tag{2.36}$$

where $\mathbf{j}$ and $\mathbf{J}$ were defined in (2.11) and (2.12). If $\mathbf{a}$ is $n \times 1$ and $\mathbf{A}$ is $n \times p$, then

$$\mathbf{a}'\mathbf{j} = \mathbf{j}'\mathbf{a} = \sum_{i=1}^n a_i, \tag{2.37}$$

$$\mathbf{j}'\mathbf{A} = \left(\sum_i a_{i1}, \sum_i a_{i2}, \ldots, \sum_i a_{ip}\right), \qquad \mathbf{Aj} = \begin{pmatrix} \sum_j a_{1j} \\ \sum_j a_{2j} \\ \vdots \\ \sum_j a_{nj} \end{pmatrix}. \tag{2.38}$$

Thus $\mathbf{a}'\mathbf{j}$ is the sum of the elements in $\mathbf{a}$, $\mathbf{j}'\mathbf{A}$ contains the column sums of $\mathbf{A}$, and $\mathbf{Aj}$ contains the row sums of $\mathbf{A}$. In $\mathbf{a}'\mathbf{j}$, the vector $\mathbf{j}$ is $n \times 1$; in $\mathbf{j}'\mathbf{A}$, the vector $\mathbf{j}$ is $n \times 1$; and in $\mathbf{Aj}$, the vector $\mathbf{j}$ is $p \times 1$.

Since $\mathbf{a}'\mathbf{b}$ is a scalar, it is equal to its transpose:

$$\mathbf{a}'\mathbf{b} = (\mathbf{a}'\mathbf{b})' = \mathbf{b}'(\mathbf{a}')' = \mathbf{b}'\mathbf{a}. \tag{2.39}$$

This allows us to write $(\mathbf{a}'\mathbf{b})^2$ in the form

$$(\mathbf{a}'\mathbf{b})^2 = (\mathbf{a}'\mathbf{b})(\mathbf{a}'\mathbf{b}) = (\mathbf{a}'\mathbf{b})(\mathbf{b}'\mathbf{a}) = \mathbf{a}'(\mathbf{bb}')\mathbf{a}. \tag{2.40}$$

From (2.18), (2.26), and (2.39) we obtain

$$(\mathbf{x} - \mathbf{y})'(\mathbf{x} - \mathbf{y}) = \mathbf{x}'\mathbf{x} - 2\mathbf{x}'\mathbf{y} + \mathbf{y}'\mathbf{y}. \tag{2.41}$$

Note that in analogous expressions with matrices, however, the two middle terms cannot be combined:

$$(\mathbf{A} - \mathbf{B})'(\mathbf{A} - \mathbf{B}) = \mathbf{A}'\mathbf{A} - \mathbf{A}'\mathbf{B} - \mathbf{B}'\mathbf{A} + \mathbf{B}'\mathbf{B},$$
$$(\mathbf{A} - \mathbf{B})^2 = (\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B}) = \mathbf{A}^2 - \mathbf{AB} - \mathbf{BA} + \mathbf{B}^2.$$

If $\mathbf{a}$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are all $p \times 1$ and $\mathbf{A}$ is $p \times p$, we obtain the following factoring results as extensions of (2.21) and (2.29):

$$\sum_{i=1}^{n} \mathbf{a}'\mathbf{x}_i = \mathbf{a}' \sum_{i=1}^{n} \mathbf{x}_i, \tag{2.42}$$

$$\sum_{i=1}^{n} \mathbf{A}\mathbf{x}_i = \mathbf{A} \sum_{i=1}^{n} \mathbf{x}_i, \tag{2.43}$$

$$\sum_{i=1}^{n} (\mathbf{a}'\mathbf{x}_i)^2 = \mathbf{a}' \left(\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' \right) \mathbf{a} \qquad \text{[by (2.40)]}, \tag{2.44}$$

$$\sum_{i=1}^{n} \mathbf{A}\mathbf{x}_i (\mathbf{A}\mathbf{x}_i)' = \mathbf{A} \left(\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' \right) \mathbf{A}'. \tag{2.45}$$

We can express matrix multiplication in terms of row vectors and column vectors. If $\mathbf{a}_i'$ is the ith row of $\mathbf{A}$ and $\mathbf{b}_j$ is the jth column of $\mathbf{B}$, then the (ij)th element of $\mathbf{AB}$

is $\mathbf{a}_i'\mathbf{b}_j$. For example, if $\mathbf{A}$ has three rows and $\mathbf{B}$ has two columns,

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \mathbf{a}_3' \end{pmatrix}, \qquad \mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2),$$

then the product $\mathbf{AB}$ can be written as

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1'\mathbf{b}_1 & \mathbf{a}_1'\mathbf{b}_2 \\ \mathbf{a}_2'\mathbf{b}_1 & \mathbf{a}_2'\mathbf{b}_2 \\ \mathbf{a}_3'\mathbf{b}_1 & \mathbf{a}_3'\mathbf{b}_2 \end{pmatrix}. \tag{2.46}$$

This can be expressed in terms of the rows of $\mathbf{A}$:

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1'(\mathbf{b}_1, \mathbf{b}_2) \\ \mathbf{a}_2'(\mathbf{b}_1, \mathbf{b}_2) \\ \mathbf{a}_3'(\mathbf{b}_1, \mathbf{b}_2) \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1'\mathbf{B} \\ \mathbf{a}_2'\mathbf{B} \\ \mathbf{a}_3'\mathbf{B} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \mathbf{a}_3' \end{pmatrix}\mathbf{B}. \tag{2.47}$$

Note that the first column of $\mathbf{AB}$ in (2.46) is

$$\begin{pmatrix} \mathbf{a}_1'\mathbf{b}_1 \\ \mathbf{a}_2'\mathbf{b}_1 \\ \mathbf{a}_3'\mathbf{b}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \mathbf{a}_3' \end{pmatrix}\mathbf{b}_1 = \mathbf{Ab}_1,$$

and likewise the second column is $\mathbf{Ab}_2$. Thus $\mathbf{AB}$ can be written in the form

$$\mathbf{AB} = \mathbf{A}(\mathbf{b}_1, \mathbf{b}_2) = (\mathbf{Ab}_1, \mathbf{Ab}_2).$$

This result holds in general:

$$\mathbf{AB} = \mathbf{A}(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p) = (\mathbf{Ab}_1, \mathbf{Ab}_2, \ldots, \mathbf{Ab}_p). \tag{2.48}$$

To further illustrate matrix multiplication in terms of rows and columns, let $\mathbf{A} = \binom{\mathbf{a}_1'}{\mathbf{a}_2'}$ be a $2 \times p$ matrix, $\mathbf{x}$ be a $p \times 1$ vector, and $\mathbf{S}$ be a $p \times p$ matrix. Then

$$\mathbf{Ax} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \end{pmatrix}\mathbf{x} = \begin{pmatrix} \mathbf{a}_1'\mathbf{x} \\ \mathbf{a}_2'\mathbf{x} \end{pmatrix}, \tag{2.49}$$

$$\mathbf{ASA}' = \begin{pmatrix} \mathbf{a}_1'\mathbf{Sa}_1 & \mathbf{a}_1'\mathbf{Sa}_2 \\ \mathbf{a}_2'\mathbf{Sa}_1 & \mathbf{a}_2'\mathbf{Sa}_2 \end{pmatrix}. \tag{2.50}$$

Any matrix can be multiplied by its transpose. If $\mathbf{A}$ is $n \times p$, then

$\mathbf{AA}'$ is $n \times n$ and is obtained as products of rows of $\mathbf{A}$ [see (2.52)].

Similarly,

$$\mathbf{A}'\mathbf{A} \text{ is } p \times p \text{ and is obtained as products of columns of } \mathbf{A} \text{ [see (2.54)].}$$

From (2.6) and (2.27), it is clear that both $\mathbf{AA}'$ and $\mathbf{A}'\mathbf{A}$ are symmetric.

In the preceding illustration for $\mathbf{AB}$ in terms of row and column vectors, the rows of $\mathbf{A}$ were denoted by $\mathbf{a}_i'$ and the columns of $\mathbf{B}$, by $\mathbf{b}_j$. If both rows and columns of a matrix $\mathbf{A}$ are under discussion, as in $\mathbf{AA}'$ and $\mathbf{A}'\mathbf{A}$, we will use the notation $\mathbf{a}_i'$ for rows and $\mathbf{a}_{(j)}$ for columns. To illustrate, if $\mathbf{A}$ is 3×4, we have

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \mathbf{a}_3' \end{pmatrix} = (\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \mathbf{a}_{(3)}, \mathbf{a}_{(4)}),$$

where, for example,

$$\mathbf{a}_2' = (a_{21} \;\; a_{22} \;\; a_{23} \;\; a_{24}),$$

$$\mathbf{a}_{(3)} = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}.$$

With this notation for rows and columns of $\mathbf{A}$, we can express the elements of $\mathbf{A}'\mathbf{A}$ or of $\mathbf{AA}'$ as products of the rows of $\mathbf{A}$ or of the columns of $\mathbf{A}$. Thus if we write $\mathbf{A}$ in terms of its rows as

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_n' \end{pmatrix},$$

then we have

$$\mathbf{A}'\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_n' \end{pmatrix} = \sum_{i=1}^{n} \mathbf{a}_i \mathbf{a}_i', \tag{2.51}$$

$$\mathbf{AA}' = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_n' \end{pmatrix} (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \begin{pmatrix} \mathbf{a}_1'\mathbf{a}_1 & \mathbf{a}_1'\mathbf{a}_2 & \cdots & \mathbf{a}_1'\mathbf{a}_n \\ \mathbf{a}_2'\mathbf{a}_1 & \mathbf{a}_2'\mathbf{a}_2 & \cdots & \mathbf{a}_2'\mathbf{a}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_n'\mathbf{a}_1 & \mathbf{a}_n'\mathbf{a}_2 & \cdots & \mathbf{a}_n'\mathbf{a}_n \end{pmatrix}. \tag{2.52}$$

Similarly, if we express $\mathbf{A}$ in terms of its columns as

$$\mathbf{A} = (\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \ldots, \mathbf{a}_{(p)}),$$

then

$$\mathbf{AA}' = (\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \ldots, \mathbf{a}_{(p)}) \begin{pmatrix} \mathbf{a}'_{(1)} \\ \mathbf{a}'_{(2)} \\ \vdots \\ \mathbf{a}'_{(p)} \end{pmatrix} = \sum_{j=1}^{p} \mathbf{a}_{(j)}\mathbf{a}'_{(j)}, \tag{2.53}$$

$$\begin{aligned} \mathbf{A}'\mathbf{A} &= \begin{pmatrix} \mathbf{a}'_{(1)} \\ \mathbf{a}'_{(2)} \\ \vdots \\ \mathbf{a}'_{(p)} \end{pmatrix} (\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \ldots, \mathbf{a}_{(p)}) \\ &= \begin{pmatrix} \mathbf{a}'_{(1)}\mathbf{a}_{(1)} & \mathbf{a}'_{(1)}\mathbf{a}_{(2)} & \cdots & \mathbf{a}'_{(1)}\mathbf{a}_{(p)} \\ \mathbf{a}'_{(2)}\mathbf{a}_{(1)} & \mathbf{a}'_{(2)}\mathbf{a}_{(2)} & \cdots & \mathbf{a}'_{(2)}\mathbf{a}_{(p)} \\ \vdots & \vdots & & \vdots \\ \mathbf{a}'_{(p)}\mathbf{a}_{(1)} & \mathbf{a}'_{(p)}\mathbf{a}_{(2)} & \ldots & \mathbf{a}'_{(p)}\mathbf{a}_{(p)} \end{pmatrix}. \end{aligned} \tag{2.54}$$

Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ matrix and $\mathbf{D}$ be a diagonal matrix, $\mathbf{D} = \text{diag}(d_1, d_2, \ldots, d_n)$. Then, in the product $\mathbf{DA}$, the ith row of $\mathbf{A}$ is multiplied by d_i, and in $\mathbf{AD}$, the jth column of $\mathbf{A}$ is multiplied by d_j. For example, if $n = 3$, we have

$$\begin{aligned} \mathbf{DA} &= \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= \begin{pmatrix} d_1 a_{11} & d_1 a_{12} & d_1 a_{13} \\ d_2 a_{21} & d_2 a_{22} & d_2 a_{23} \\ d_3 a_{31} & d_3 a_{32} & d_3 a_{33} \end{pmatrix}, \end{aligned} \tag{2.55}$$

$$\begin{aligned} \mathbf{AD} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \\ &= \begin{pmatrix} d_1 a_{11} & d_2 a_{12} & d_3 a_{13} \\ d_1 a_{21} & d_2 a_{22} & d_3 a_{23} \\ d_1 a_{31} & d_2 a_{32} & d_3 a_{33} \end{pmatrix}, \end{aligned} \tag{2.56}$$

$$\mathbf{DAD} = \begin{pmatrix} d_1^2 a_{11} & d_1 d_2 a_{12} & d_1 d_3 a_{13} \\ d_2 d_1 a_{21} & d_2^2 a_{22} & d_2 d_3 a_{23} \\ d_3 d_1 a_{31} & d_3 d_2 a_{32} & d_3^2 a_{33} \end{pmatrix}. \tag{2.57}$$

In the special case where the diagonal matrix is the identity, we have

$$\mathbf{IA} = \mathbf{AI} = \mathbf{A}. \tag{2.58}$$

If $\mathbf{A}$ is rectangular, (2.58) still holds, but the two identities are of different sizes.

The product of a scalar and a matrix is obtained by multiplying each element of the matrix by the scalar:

$$c\mathbf{A} = (ca_{ij}) = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1m} \\ ca_{21} & ca_{22} & \cdots & ca_{2m} \\ \vdots & \vdots & & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nm} \end{pmatrix}. \tag{2.59}$$

For example,

$$c\mathbf{I} = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}, \tag{2.60}$$

$$c\mathbf{x} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}. \tag{2.61}$$

Since $ca_{ij} = a_{ij}c$, the product of a scalar and a matrix is commutative:

$$c\mathbf{A} = \mathbf{A}c. \tag{2.62}$$

Multiplication of vectors or matrices by scalars permits the use of linear combinations, such as

$$\sum_{i=1}^{k} a_i \mathbf{x}_i = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k,$$

$$\sum_{i=1}^{k} a_i \mathbf{B}_i = a_1\mathbf{B}_1 + a_2\mathbf{B}_2 + \cdots + a_k\mathbf{B}_k.$$

If $\mathbf{A}$ is a symmetric matrix and $\mathbf{x}$ and $\mathbf{y}$ are vectors, the product

$$\mathbf{y}'\mathbf{A}\mathbf{y} = \sum_i a_{ii}y_i^2 + \sum_{i \neq j} a_{ij}y_iy_j \tag{2.63}$$

is called a *quadratic form*, whereas

$$\mathbf{x}'\mathbf{A}\mathbf{y} = \sum_{ij} a_{ij}x_iy_j \tag{2.64}$$

is called a *bilinear form*. Either of these is, of course, a scalar and can be treated as such. Expressions such as $\mathbf{x}'\mathbf{A}\mathbf{y}/\sqrt{\mathbf{x}'\mathbf{A}\mathbf{x}}$ are permissible (assuming $\mathbf{A}$ is positive definite; see Section 2.7).

2.4 PARTITIONED MATRICES

It is sometimes convenient to partition a matrix into submatrices. For example, a partitioning of a matrix $\mathbf{A}$ into four submatrices could be indicated symbolically as follows:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}.$$

For example, a 4×5 matrix $\mathbf{A}$ can be partitioned as

$$\mathbf{A} = \left(\begin{array}{rrr|rr} 2 & 1 & 3 & 8 & 4 \\ -3 & 4 & 0 & 2 & 7 \\ 9 & 3 & 6 & 5 & -2 \\ \hline 4 & 8 & 3 & 1 & 6 \end{array}\right) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

where

$$\mathbf{A}_{11} = \begin{pmatrix} 2 & 1 & 3 \\ -3 & 4 & 0 \\ 9 & 3 & 6 \end{pmatrix}, \qquad \mathbf{A}_{12} = \begin{pmatrix} 8 & 4 \\ 2 & 7 \\ 5 & -2 \end{pmatrix},$$
$$\mathbf{A}_{21} = (4 \quad 8 \quad 3), \qquad \mathbf{A}_{22} = (1 \quad 6).$$

If two matrices $\mathbf{A}$ and $\mathbf{B}$ are conformable and $\mathbf{A}$ and $\mathbf{B}$ are partitioned so that the submatrices are appropriately conformable, then the product $\mathbf{AB}$ can be found by following the usual row-by-column pattern of multiplication on the submatrices as if they were single elements; for example,

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}\begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{pmatrix}. \end{aligned} \tag{2.65}$$

It can be seen that this formulation is equivalent to the usual row-by-column definition of matrix multiplication. For example, the $(1, 1)$ element of $\mathbf{AB}$ is the product of the first row of $\mathbf{A}$ and the first column of $\mathbf{B}$. In the $(1, 1)$ element of $\mathbf{A}_{11}\mathbf{B}_{11}$ we have the sum of products of part of the first row of $\mathbf{A}$ and part of the first column of $\mathbf{B}$. In the $(1, 1)$ element of $\mathbf{A}_{12}\mathbf{B}_{21}$ we have the sum of products of the rest of the first row of $\mathbf{A}$ and the remainder of the first column of $\mathbf{B}$.

Multiplication of a matrix and a vector can also be carried out in partitioned form. For example,

$$\mathbf{Ab} = (\mathbf{A}_1, \mathbf{A}_2)\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \mathbf{A}_1\mathbf{b}_1 + \mathbf{A}_2\mathbf{b}_2, \tag{2.66}$$

where the partitioning of the columns of $\mathbf{A}$ corresponds to the partitioning of the elements of $\mathbf{b}$. Note that the partitioning of $\mathbf{A}$ into two sets of columns is indicated by a comma, $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$.

The partitioned multiplication in (2.66) can be extended to individual columns of $\mathbf{A}$ and individual elements of $\mathbf{b}$:

$$\begin{aligned}\mathbf{Ab} &= (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} \\ &= b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + \cdots + b_p\mathbf{a}_p. \end{aligned} \tag{2.67}$$

Thus $\mathbf{Ab}$ is expressible as a linear combination of the columns of $\mathbf{A}$, the coefficients being elements of $\mathbf{b}$. For example, let

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & 0 \\ 4 & 3 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}.$$

Then

$$\mathbf{Ab} = \begin{pmatrix} 11 \\ 10 \\ 28 \end{pmatrix}.$$

Using a linear combination of columns of $\mathbf{A}$ as in (2.67), we obtain

$$\begin{aligned}\mathbf{Ab} &= b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + b_3\mathbf{a}_3 \\ &= 4\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} + 2\begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 12 \\ 8 \\ 16 \end{pmatrix} + \begin{pmatrix} -4 \\ 2 \\ 6 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 11 \\ 10 \\ 28 \end{pmatrix}. \end{aligned}$$

We note that if $\mathbf{A}$ is partitioned as in (2.66), $\mathbf{A} = (\mathbf{A}_2, \mathbf{A}_2)$, the transpose is not equal to $(\mathbf{A}_1', \mathbf{A}_2')$, but rather

$$\mathbf{A}' = (\mathbf{A}_1, \mathbf{A}_2)' = \begin{pmatrix} \mathbf{A}_1' \\ \mathbf{A}_2' \end{pmatrix}. \tag{2.68}$$

2.5 RANK

Before defining the rank of a matrix, we first introduce the notion of linear independence and dependence. A set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ is said to be *linearly dependent* if constants $c_1, c_2, \ldots, c_n$ (not all zero) can be found such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n = \mathbf{0}. \tag{2.69}$$

If no constants $c_1, c_2, \ldots, c_n$ can be found satisfying (2.69), the set of vectors is said to be *linearly independent*.

If (2.69) holds, then at least one of the vectors $\mathbf{a}_i$ can be expressed as a linear combination of the other vectors in the set. Thus linear dependence of a set of vectors implies redundancy in the set. Among linearly independent vectors there is no redundancy of this type.

The *rank* of any square or rectangular matrix $\mathbf{A}$ is defined as

$$\begin{aligned} \text{rank}(\mathbf{A}) &= \text{number of linearly independent rows of } \mathbf{A} \\ &= \text{number of linearly independent columns of } \mathbf{A}. \end{aligned}$$

It can be shown that the number of linearly independent rows of a matrix is always equal to the number of linearly independent columns.

If $\mathbf{A}$ is $n \times p$, the maximum possible rank of $\mathbf{A}$ is the smaller of n and p, in which case $\mathbf{A}$ is said to be of *full rank* (sometimes said *full row rank* or *full column rank*). For example,

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

has rank 2 because the two rows are linearly independent (neither row is a multiple of the other). However, even though $\mathbf{A}$ is full rank, the columns are linearly dependent because rank 2 implies there are only two linearly independent columns. Thus, by (2.69), there exist constants c_1, c_2, and c_3 such that

$$c_1 \begin{pmatrix} 1 \\ 5 \end{pmatrix} + c_2 \begin{pmatrix} -2 \\ 2 \end{pmatrix} + c_3 \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.70}$$

By (2.67), we can write (2.70) in the form

$$\begin{pmatrix} 1 & -2 & 3 \\ 5 & 2 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

or

$$\mathbf{Ac} = \mathbf{0}. \tag{2.71}$$

A solution vector to (2.70) or (2.71) is given by any multiple of $\mathbf{c} = (14, -11, -12)'$. Hence we have the interesting result that a product of a matrix $\mathbf{A}$ and a vector $\mathbf{c}$ is equal to $\mathbf{0}$, even though $\mathbf{A} \neq \mathbf{O}$ and $\mathbf{c} \neq \mathbf{0}$. This is a direct consequence of the linear dependence of the column vectors of $\mathbf{A}$.

Another consequence of the linear dependence of rows or columns of a matrix is the possibility of expressions such as $\mathbf{AB} = \mathbf{CB}$, where $\mathbf{A} \neq \mathbf{C}$. For example, let

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 0 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 2 & 1 & 1 \\ 5 & -6 & -4 \end{pmatrix}.$$

Then

$$\mathbf{AB} = \mathbf{CB} = \begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix}.$$

All three matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are full rank; but being rectangular, they have a rank deficiency in either rows or columns, which permits us to construct $\mathbf{AB} = \mathbf{CB}$ with $\mathbf{A} \neq \mathbf{C}$. Thus in a matrix equation, we cannot, in general, cancel matrices from both sides of the equation.

There are two exceptions to this rule. One exception involves a nonsingular matrix to be defined in Section 2.6. The other special case occurs when the expression holds for all possible values of the matrix common to both sides of the equation. For example,

$$\text{If } \mathbf{Ax} = \mathbf{Bx} \text{ for all possible values of } \mathbf{x}, \text{ then } \mathbf{A} = \mathbf{B}. \tag{2.72}$$

To see this, let $\mathbf{x} = (1, 0, \ldots, 0)'$. Then the first column of $\mathbf{A}$ equals the first column of $\mathbf{B}$. Now let $\mathbf{x} = (0, 1, 0, \ldots, 0)'$, and the second column of $\mathbf{A}$ equals the second column of $\mathbf{B}$. Continuing in this fashion, we obtain $\mathbf{A} = \mathbf{B}$.

Suppose a rectangular matrix $\mathbf{A}$ is $n \times p$ of rank p, where $p < n$. We typically shorten this statement to "$\mathbf{A}$ is $n \times p$ of rank $p < n$."

2.6 INVERSE

If a matrix $\mathbf{A}$ is square and of full rank, then $\mathbf{A}$ is said to be *nonsingular*, and $\mathbf{A}$ has a unique *inverse*, denoted by $\mathbf{A}^{-1}$, with the property that

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{2.73}$$

For example, let

$$\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 2 & 6 \end{pmatrix}.$$

Then

$$\mathbf{A}^{-1} = \begin{pmatrix} .6 & -.4 \\ -.2 & .3 \end{pmatrix},$$

$$\mathbf{A}\mathbf{A}^{-1} = \begin{pmatrix} 3 & 4 \\ 2 & 6 \end{pmatrix}\begin{pmatrix} .6 & -.4 \\ -.2 & .3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

If $\mathbf{A}$ is square and of less than full rank, then an inverse does not exist, and $\mathbf{A}$ is said to be *singular*. Note that rectangular matrices do not have inverses as in (2.73), even if they are full rank.

If $\mathbf{A}$ and $\mathbf{B}$ are the same size and nonsingular, then the inverse of their product is the product of their inverses in reverse order,

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{2.74}$$

Note that (2.74) holds only for nonsingular matrices. Thus, for example, if $\mathbf{A}$ is $n \times p$ of rank $p < n$, then $\mathbf{A}'\mathbf{A}$ has an inverse, but $(\mathbf{A}'\mathbf{A})^{-1}$ is not equal to $\mathbf{A}^{-1}(\mathbf{A}')^{-1}$ because $\mathbf{A}$ is rectangular and does not have an inverse.

If a matrix is nonsingular, it can be canceled from both sides of an equation, provided it appears on the left (or right) on both sides. For example, if $\mathbf{B}$ is nonsingular, then

$$\mathbf{AB} = \mathbf{CB} \quad \text{implies} \quad \mathbf{A} = \mathbf{C},$$

since we can multiply on the right by $\mathbf{B}^{-1}$ to obtain

$$\mathbf{ABB}^{-1} = \mathbf{CBB}^{-1},$$

$$\mathbf{AI} = \mathbf{CI},$$

$$\mathbf{A} = \mathbf{C}.$$

Otherwise, if $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are rectangular or square and singular, it is easy to construct $\mathbf{AB} = \mathbf{CB}$, with $\mathbf{A} \neq \mathbf{C}$, as illustrated near the end of Section 2.5.

The inverse of the transpose of a nonsingular matrix is given by the transpose of the inverse:

$$(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'. \tag{2.75}$$

If the symmetric nonsingular matrix $\mathbf{A}$ is partitioned in the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{a}_{12} \\ \mathbf{a}'_{12} & a_{22} \end{pmatrix},$$

then the inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{b}\begin{pmatrix} b\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{a}_{12}\mathbf{a}_{12}'\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{a}_{12} \\ -\mathbf{a}_{12}'\mathbf{A}_{11}^{-1} & 1 \end{pmatrix}, \qquad (2.76)$$

where $b = a_{22} - \mathbf{a}_{12}'\mathbf{A}_{11}^{-1}\mathbf{a}_{12}$. A nonsingular matrix of the form $\mathbf{B} + \mathbf{cc}'$, where $\mathbf{B}$ is nonsingular, has as its inverse

$$(\mathbf{B} + \mathbf{cc}')^{-1} = \mathbf{B}^{-1} - \frac{\mathbf{B}^{-1}\mathbf{cc}'\mathbf{B}^{-1}}{1 + \mathbf{c}'\mathbf{B}^{-1}\mathbf{c}}. \qquad (2.77)$$

2.7 POSITIVE DEFINITE MATRICES

The symmetric matrix $\mathbf{A}$ is said to be *positive definite* if $\mathbf{x}'\mathbf{Ax} > 0$ for all possible vectors $\mathbf{x}$ (except $\mathbf{x} = \mathbf{0}$). Similarly, $\mathbf{A}$ is *positive semidefinite* if $\mathbf{x}'\mathbf{Ax} \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$. [A quadratic form $\mathbf{x}'\mathbf{Ax}$ was defined in (2.63).] The diagonal elements a_{ii} of a positive definite matrix are positive. To see this, let $\mathbf{x}' = (0, \ldots, 0, 1, 0, \ldots, 0)$ with a 1 in the ith position. Then $\mathbf{x}'\mathbf{Ax} = a_{ii} > 0$. Similarly, for a positive semidefinite matrix $\mathbf{A}$, $a_{ii} \geq 0$ for all i.

One way to obtain a positive definite matrix is as follows:

If $\mathbf{A} = \mathbf{B}'\mathbf{B}$, where $\mathbf{B}$ is $n \times p$ of rank $p < n$, then $\mathbf{B}'\mathbf{B}$ is positive definite. (2.78)

This is easily shown:

$$\mathbf{x}'\mathbf{Ax} = \mathbf{x}'\mathbf{B}'\mathbf{Bx} = (\mathbf{Bx})'(\mathbf{Bx}) = \mathbf{z}'\mathbf{z},$$

where $\mathbf{z} = \mathbf{Bx}$. Thus $\mathbf{x}'\mathbf{Ax} = \sum_{i=1}^{n} z_i^2$, which is positive ($\mathbf{Bx}$ cannot be $\mathbf{0}$ unless $\mathbf{x} = \mathbf{0}$, because $\mathbf{B}$ is full rank). If $\mathbf{B}$ is less than full rank, then by a similar argument, $\mathbf{B}'\mathbf{B}$ is positive semidefinite.

Note that $\mathbf{A} = \mathbf{B}'\mathbf{B}$ is analogous to $a = b^2$ in real numbers, where the square of any number (including negative numbers) is positive.

In another analogy to positive real numbers, a positive definite matrix can be factored into a "square root" in two ways. We give one method in (2.79) and the other in Section 2.11.8.

A positive definite matrix $\mathbf{A}$ can be factored into

$$\mathbf{A} = \mathbf{T}'\mathbf{T}, \qquad (2.79)$$

where $\mathbf{T}$ is a nonsingular upper triangular matrix. One way to obtain $\mathbf{T}$ is the *Cholesky decomposition*, which can be carried out in the following steps.

Let $\mathbf{A} = (a_{ij})$ and $\mathbf{T} = (t_{ij})$ be $n \times n$. Then the elements of $\mathbf{T}$ are found as follows:

$$
\begin{aligned}
t_{11} &= \sqrt{a_{11}}, \quad t_{1j} = \frac{a_{1j}}{t_{11}} && 2 \le j \le n, \\
t_{ii} &= \sqrt{a_{ii} - \sum_{k=1}^{i-1} t_{ki}^2} && 2 \le i \le n, \\
t_{ij} &= \frac{a_{ij} - \sum_{k=1}^{i-1} t_{ki} t_{kj}}{t_{ii}} && 2 \le i < j \le n, \\
t_{ij} &= 0 && 1 \le j < i \le n.
\end{aligned}
$$

For example, let

$$
\mathbf{A} = \begin{pmatrix} 3 & 0 & -3 \\ 0 & 6 & 3 \\ -3 & 3 & 6 \end{pmatrix}.
$$

Then by the Cholesky method, we obtain

$$
\mathbf{T} = \begin{pmatrix} \sqrt{3} & 0 & -\sqrt{3} \\ 0 & \sqrt{6} & \sqrt{1.5} \\ 0 & 0 & \sqrt{1.5} \end{pmatrix},
$$

$$
\begin{aligned}
\mathbf{T}'\mathbf{T} &= \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{6} & 0 \\ -\sqrt{3} & \sqrt{1.5} & \sqrt{1.5} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & -\sqrt{3} \\ 0 & \sqrt{6} & \sqrt{1.5} \\ 0 & 0 & \sqrt{1.5} \end{pmatrix} \\
&= \begin{pmatrix} 3 & 0 & -3 \\ 0 & 6 & 3 \\ -3 & 3 & 6 \end{pmatrix} = \mathbf{A}.
\end{aligned}
$$

2.8 DETERMINANTS

The *determinant* of an $n \times n$ matrix $\mathbf{A}$ is defined as the sum of all $n!$ possible products of n elements such that

1. each product contains one element from every row and every column, and
2. the factors in each product are written so that the column subscripts appear in order of magnitude and each product is then preceded by a plus or minus sign according to whether the number of inversions in the row subscripts is even or odd.

An *inversion* occurs whenever a larger number precedes a smaller one. The symbol $n!$ is defined as

$$
n! = n(n-1)(n-2)\cdots 2 \cdot 1. \tag{2.80}
$$

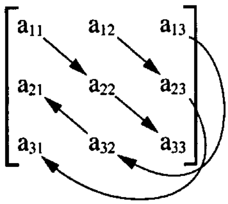
The determinant of $\mathbf{A}$ is a scalar denoted by $|\mathbf{A}|$ or by $\det(\mathbf{A})$. The preceding definition is not useful in evaluating determinants, except in the case of 2×2 or 3×3 matrices. For larger matrices, other methods are available for manual computation, but determinants are typically evaluated by computer. For a 2×2 matrix, the determinant is found by

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}. \tag{2.81}$$

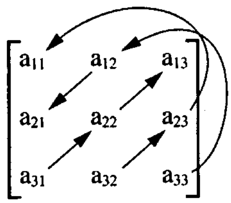
For a 3×3 matrix, the determinant is given by

$$|\mathbf{A}| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{12}a_{21}. \tag{2.82}$$

This can be found by the following scheme. The three positive terms are obtained by



and the three negative terms, by



The determinant of a diagonal matrix is the product of the diagonal elements; that is, if $\mathbf{D} = \operatorname{diag}(d_1, d_2, \ldots, d_n)$, then

$$|\mathbf{D}| = \prod_{i=1}^{n} d_i. \tag{2.83}$$

As a special case of (2.83), suppose all diagonal elements are equal, say,

$$\mathbf{D} = \operatorname{diag}(c, c, \ldots, c) = c\mathbf{I}.$$

Then

$$|\mathbf{D}| = |c\mathbf{I}| = \prod_{i=1}^{n} c = c^n. \tag{2.84}$$

The extension of (2.84) to any square matrix $\mathbf{A}$ is

$$|c\mathbf{A}| = c^n |\mathbf{A}|. \tag{2.85}$$

Since the determinant is a scalar, we can carry out operations such as

$$|\mathbf{A}|^2, \qquad |\mathbf{A}|^{1/2}, \qquad \frac{1}{|\mathbf{A}|},$$

provided that $|\mathbf{A}| > 0$ for $|\mathbf{A}|^{1/2}$ and that $|\mathbf{A}| \neq 0$ for $1/|\mathbf{A}|$.

If the square matrix $\mathbf{A}$ is singular, its determinant is 0:

$$|\mathbf{A}| = 0 \text{ if } \mathbf{A} \text{ is singular.} \tag{2.86}$$

If $\mathbf{A}$ is *near singular*, then there exists a linear combination of the columns that is close to $\mathbf{0}$, and $|\mathbf{A}|$ is also close to 0. If $\mathbf{A}$ is nonsingular, its determinant is nonzero:

$$|\mathbf{A}| \neq 0 \text{ if } \mathbf{A} \text{ is nonsingular.} \tag{2.87}$$

If $\mathbf{A}$ is positive definite, its determinant is positive:

$$|\mathbf{A}| > 0 \text{ if } \mathbf{A} \text{ is positive definite.} \tag{2.88}$$

If $\mathbf{A}$ and $\mathbf{B}$ are square and the same size, the determinant of the product is the product of the determinants:

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|. \tag{2.89}$$

For example, let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -3 & 5 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}.$$

Then

$$\mathbf{AB} = \begin{pmatrix} 6 & 8 \\ -7 & 9 \end{pmatrix}, \qquad |\mathbf{AB}| = 110,$$

$$|\mathbf{A}| = 11, \qquad |\mathbf{B}| = 10, \qquad |\mathbf{A}||\mathbf{B}| = 110.$$

The determinant of the transpose of a matrix is the same as the determinant of the matrix, and the determinant of the the inverse of a matrix is the reciprocal of the determinant:

$$|\mathbf{A}'| = |\mathbf{A}|, \tag{2.90}$$

$$|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|} = |\mathbf{A}|^{-1}. \tag{2.91}$$

If a partitioned matrix has the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_{22} \end{pmatrix},$$

where $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are square but not necessarily the same size, then

$$|\mathbf{A}| = \begin{vmatrix} \mathbf{A}_{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_{22} \end{vmatrix} = |\mathbf{A}_{11}||\mathbf{A}_{22}|. \tag{2.92}$$

For a general partitioned matrix,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

where $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are square and nonsingular (not necessarily the same size), the determinant is given by either of the following two expressions:

$$\begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{vmatrix} = |\mathbf{A}_{11}||\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}| \tag{2.93}$$

$$= |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|. \tag{2.94}$$

Note the analogy of (2.93) and (2.94) to the case of the determinant of a 2×2 matrix as given by (2.81):

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}a_{22} - a_{21}a_{12} \\ &= a_{11}\left(a_{22} - \frac{a_{21}a_{12}}{a_{11}}\right) \\ &= a_{22}\left(a_{11} - \frac{a_{12}a_{21}}{a_{22}}\right). \end{aligned}$$

If $\mathbf{B}$ is nonsingular and $\mathbf{c}$ is a vector, then

$$|\mathbf{B} + \mathbf{c}\mathbf{c}'| = |\mathbf{B}|(1 + \mathbf{c}'\mathbf{B}^{-1}\mathbf{c}). \tag{2.95}$$

2.9 TRACE

A simple function of an $n \times n$ matrix $\mathbf{A}$ is the *trace*, denoted by tr($\mathbf{A}$) and defined as the sum of the diagonal elements of $\mathbf{A}$; that is, $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$. The trace is, of course, a scalar. For example, suppose

$$\mathbf{A} = \begin{pmatrix} 5 & 4 & 4 \\ 2 & -3 & 1 \\ 3 & 7 & 9 \end{pmatrix}.$$

Then

$$\mathrm{tr}(\mathbf{A}) = 5 + (-3) + 9 = 11.$$

The trace of the sum of two square matrices is the sum of the traces of the two matrices:

$$\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B}). \tag{2.96}$$

An important result for the product of two matrices is

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA}). \tag{2.97}$$

This result holds for any matrices $\mathbf{A}$ and $\mathbf{B}$ where $\mathbf{AB}$ and $\mathbf{BA}$ are both defined. It is not necessary that $\mathbf{A}$ and $\mathbf{B}$ be square or that $\mathbf{AB}$ equal $\mathbf{BA}$. For example, let

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 4 & 6 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 4 & 5 \end{pmatrix}.$$

Then

$$\mathbf{AB} = \begin{pmatrix} 9 & 10 & 16 \\ 4 & -8 & -3 \\ 24 & 16 & 34 \end{pmatrix}, \qquad \mathbf{BA} = \begin{pmatrix} 3 & 17 \\ 30 & 32 \end{pmatrix},$$

$$\mathrm{tr}(\mathbf{AB}) = 9 - 8 + 34 = 35, \qquad \mathrm{tr}(\mathbf{BA}) = 3 + 32 = 35.$$

From (2.52) and (2.54), we obtain

$$\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \mathrm{tr}(\mathbf{AA}') = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij}^2, \tag{2.98}$$

where the a_{ij}'s are elements of the $n \times p$ matrix $\mathbf{A}$.

2.10 ORTHOGONAL VECTORS AND MATRICES

Two vectors $\mathbf{a}$ and $\mathbf{b}$ of the same size are said to be *orthogonal* if

$$\mathbf{a}'\mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = 0. \tag{2.99}$$

Geometrically, orthogonal vectors are perpendicular [see (3.14) and the comments following (3.14)]. If $\mathbf{a}'\mathbf{a} = 1$, the vector $\mathbf{a}$ is said to be *normalized*. The vector $\mathbf{a}$ can always be normalized by dividing by its length, $\sqrt{\mathbf{a}'\mathbf{a}}$. Thus

$$\mathbf{c} = \frac{\mathbf{a}}{\sqrt{\mathbf{a}'\mathbf{a}}} \tag{2.100}$$

is normalized so that $\mathbf{c}'\mathbf{c} = 1$.

A matrix $\mathbf{C} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p)$ whose columns are normalized and mutually orthogonal is called an *orthogonal* matrix. Since the elements of $\mathbf{C}'\mathbf{C}$ are products of columns of $\mathbf{C}$ [see (2.54)], which have the properties $\mathbf{c}_i'\mathbf{c}_i = 1$ for all i and $\mathbf{c}_i'\mathbf{c}_j = 0$ for all $i \neq j$, we have

$$\mathbf{C}'\mathbf{C} = \mathbf{I}. \tag{2.101}$$

If $\mathbf{C}$ satisfies (2.101), it necessarily follows that

$$\mathbf{C}\mathbf{C}' = \mathbf{I}, \tag{2.102}$$

from which we see that the rows of $\mathbf{C}$ are also normalized and mutually orthogonal. It is clear from (2.101) and (2.102) that $\mathbf{C}^{-1} = \mathbf{C}'$ for an orthogonal matrix $\mathbf{C}$.

We illustrate the creation of an orthogonal matrix by starting with

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{pmatrix},$$

whose columns are mutually orthogonal. To normalize the three columns, we divide by the respective lengths, $\sqrt{3}$, $\sqrt{6}$, and $\sqrt{2}$, to obtain

$$\mathbf{C} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix}.$$

Note that the rows also became normalized and mutually orthogonal so that $\mathbf{C}$ satisfies both (2.101) and (2.102).

Multiplication by an orthogonal matrix has the effect of rotating axes; that is, if a point $\mathbf{x}$ is transformed to $\mathbf{z} = \mathbf{C}\mathbf{x}$, where $\mathbf{C}$ is orthogonal, then

$$\mathbf{z}'\mathbf{z} = (\mathbf{Cx})'(\mathbf{Cx}) = \mathbf{x}'\mathbf{C}'\mathbf{Cx} = \mathbf{x}'\mathbf{Ix} = \mathbf{x}'\mathbf{x}, \tag{2.103}$$

and the distance from the origin to $\mathbf{z}$ is the same as the distance to $\mathbf{x}$.

2.11 EIGENVALUES AND EIGENVECTORS

2.11.1 Definition

For every square matrix $\mathbf{A}$, a scalar λ and a nonzero vector $\mathbf{x}$ can be found such that

$$\mathbf{Ax} = \lambda\mathbf{x}. \tag{2.104}$$

In (2.104), λ is called an *eigenvalue* of $\mathbf{A}$, and $\mathbf{x}$ is an *eigenvector* of $\mathbf{A}$ corresponding to λ. To find λ and $\mathbf{x}$, we write (2.104) as

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}. \tag{2.105}$$

If $|\mathbf{A} - \lambda\mathbf{I}| \neq 0$, then $(\mathbf{A} - \lambda\mathbf{I})$ has an inverse and $\mathbf{x} = \mathbf{0}$ is the only solution. Hence, in order to obtain nontrivial solutions, we set $|\mathbf{A} - \lambda\mathbf{I}| = 0$ to find values of λ that can be substituted into (2.105) to find corresponding values of $\mathbf{x}$. Alternatively, (2.69) and (2.71) require that the columns of $\mathbf{A} - \lambda\mathbf{I}$ be linearly dependent. Thus in $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, the matrix $\mathbf{A} - \lambda\mathbf{I}$ must be singular in order to find a solution vector $\mathbf{x}$ that is not $\mathbf{0}$.

The equation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ is called the *characteristic equation*. If $\mathbf{A}$ is $n \times n$, the characteristic equation will have n roots; that is, $\mathbf{A}$ will have n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. The λ's will not necessarily all be distinct or all nonzero. However, if $\mathbf{A}$ arises from computations on real (continuous) data and is nonsingular, the λ's will all be distinct (with probability 1). After finding $\lambda_1, \lambda_2, \ldots, \lambda_n$, the accompanying eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ can be found using (2.105).

If we multiply both sides of (2.105) by a scalar k and note by (2.62) that k and $\mathbf{A} - \lambda\mathbf{I}$ commute, we obtain

$$(\mathbf{A} - \lambda\mathbf{I})k\mathbf{x} = k\mathbf{0} = \mathbf{0}. \tag{2.106}$$

Thus if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, $k\mathbf{x}$ is also an eigenvector, and eigenvectors are unique only up to multiplication by a scalar. Hence we can adjust the length of $\mathbf{x}$, but the direction from the origin is unique; that is, the relative values of (ratios of) the components of $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ are unique. Typically, the eigenvector $\mathbf{x}$ is scaled so that $\mathbf{x}'\mathbf{x} = 1$.

To illustrate, we will find the eigenvalues and eigenvectors for the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}.$$

The characteristic equation is

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) + 2 = 0,$$

$$\lambda^2 - 5\lambda + 6 = (\lambda - 3)(\lambda - 2) = 0,$$

from which $\lambda_1 = 3$ and $\lambda_2 = 2$. To find the eigenvector corresponding to $\lambda_1 = 3$, we use (2.105),

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0},$$

$$\begin{pmatrix} 1-3 & 2 \\ -1 & 4-3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$-2x_1 + 2x_2 = 0$$

$$-x_1 + x_2 = 0.$$

As expected, either equation is redundant in the presence of the other, and there remains a single equation with two unknowns, $x_1 = x_2$. The solution vector can be written with an arbitrary constant,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} = c\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

If c is set equal to $1/\sqrt{2}$ to normalize the eigenvector, we obtain

$$\mathbf{x}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

Similarly, corresponding to $\lambda_2 = 2$, we have

$$\mathbf{x}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}.$$

2.11.2 I + A and I − A

If λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ is the corresponding eigenvector, then $1 + \lambda$ is an eigenvalue of $\mathbf{I} + \mathbf{A}$ and $1 - \lambda$ is an eigenvalue of $\mathbf{I} - \mathbf{A}$. In either case, $\mathbf{x}$ is the corresponding eigenvector.

We demonstrate this for $\mathbf{I} + \mathbf{A}$:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x},$$

$$\mathbf{x} + \mathbf{A}\mathbf{x} = \mathbf{x} + \lambda\mathbf{x},$$

$$(\mathbf{I} + \mathbf{A})\mathbf{x} = (1 + \lambda)\mathbf{x}.$$

2.11.3 tr(A) and |A|

For any square matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, we have

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i, \tag{2.107}$$

$$|\mathbf{A}| = \prod_{i=1}^{n} \lambda_i. \tag{2.108}$$

Note that by the definition in Section 2.9, tr($\mathbf{A}$) is also equal to $\sum_{i=1}^{n} a_{ii}$, but $a_{ii} \neq \lambda_i$.

We illustrate (2.107) and (2.108) using the matrix

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

from the illustration in Section 2.11.1, for which $\lambda_1 = 3$ and $\lambda_2 = 2$. Using (2.107), we obtain

$$\text{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 = 3 + 2 = 5,$$

and from (2.108), we have

$$|\mathbf{A}| = \lambda_1\lambda_2 = 3(2) = 6.$$

By definition, we obtain

$$\text{tr}(\mathbf{A}) = 1 + 4 = 5 \quad \text{and} \quad |\mathbf{A}| = (1)(4) - (-1)(2) = 6.$$

2.11.4 Positive Definite and Semidefinite Matrices

The eigenvalues and eigenvectors of positive definite and positive semidefinite matrices have the following properties:

1. The eigenvalues of a positive definite matrix are all positive.
2. The eigenvalues of a positive semidefinite matrix are positive or zero, with the number of positive eigenvalues equal to the rank of the matrix.

It is customary to list the eigenvalues of a positive definite matrix in descending order: $\lambda_1 > \lambda_2 > \cdots > \lambda_p$. The eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are listed in the same order; $\mathbf{x}_1$ corresponds to λ_1, $\mathbf{x}_2$ corresponds to λ_2, and so on.

The following result, known as the Perron–Frobenius theorem, is of interest in Chapter 12: If all elements of the positive definite matrix $\mathbf{A}$ are positive, then all elements of the first eigenvector are positive. (The first eigenvector is the one associated with the first eigenvalue, λ_1.)

2.11.5 The Product AB

If **A** and **B** are square and the same size, the eigenvalues of **AB** are the same as those of **BA**, although the eigenvectors are usually different. This result also holds if **AB** and **BA** are both square but of different sizes, as when **A** is $n \times p$ and **B** is $p \times n$. (In this case, the nonzero eigenvalues of **AB** and **BA** will be the same.)

2.11.6 Symmetric Matrix

The eigenvectors of an $n \times n$ symmetric matrix **A** are mutually orthogonal. It follows that if the n eigenvectors of **A** are normalized and inserted as columns of a matrix $\mathbf{C} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, then **C** is orthogonal.

2.11.7 Spectral Decomposition

It was noted in Section 2.11.6 that if the matrix $\mathbf{C} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ contains the normalized eigenvectors of an $n \times n$ symmetric matrix **A**, then **C** is orthogonal. Therefore, by (2.102), $\mathbf{I} = \mathbf{CC}'$, which we can multiply by **A** to obtain

$$\mathbf{A} = \mathbf{ACC}'.$$

We now substitute $\mathbf{C} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$:

$$\begin{aligned}
\mathbf{A} &= \mathbf{A}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\mathbf{C}' \\
&= (\mathbf{Ax}_1, \mathbf{Ax}_2, \ldots, \mathbf{Ax}_n)\mathbf{C}' && \text{[by (2.48)]} \\
&= (\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \ldots, \lambda_n\mathbf{x}_n)\mathbf{C}' && \text{[by (2.104)]} \\
&= \mathbf{CDC}' && \text{[by (2.56)]}, \qquad (2.109)
\end{aligned}$$

where

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \qquad (2.110)$$

The expression $\mathbf{A} = \mathbf{CDC}'$ in (2.109) for a symmetric matrix **A** in terms of its eigenvalues and eigenvectors is known as the *spectral decomposition* of **A**.

Since **C** is orthogonal and $\mathbf{C}'\mathbf{C} = \mathbf{CC}' = \mathbf{I}$, we can multiply (2.109) on the left by $\mathbf{C}'$ and on the right by **C** to obtain

$$\mathbf{C}'\mathbf{AC} = \mathbf{D}. \qquad (2.111)$$

Thus a symmetric matrix **A** can be *diagonalized* by an orthogonal matrix containing normalized eigenvectors of **A**, and by (2.110) the resulting diagonal matrix contains eigenvalues of **A**.

2.11.8 Square Root Matrix

If $\mathbf{A}$ is positive definite, the spectral decomposition of $\mathbf{A}$ in (2.109) can be modified by taking the square roots of the eigenvalues to produce a *square root matrix*,

$$\mathbf{A}^{1/2} = \mathbf{C}\mathbf{D}^{1/2}\mathbf{C}', \tag{2.112}$$

where

$$\mathbf{D}^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}. \tag{2.113}$$

The square root matrix $\mathbf{A}^{1/2}$ is symmetric and serves as the square root of $\mathbf{A}$:

$$\mathbf{A}^{1/2}\mathbf{A}^{1/2} = (\mathbf{A}^{1/2})^2 = \mathbf{A}. \tag{2.114}$$

2.11.9 Square Matrices and Inverse Matrices

Other functions of $\mathbf{A}$ have spectral decompositions analogous to (2.112). Two of these are the square and inverse of $\mathbf{A}$. If the square matrix $\mathbf{A}$ has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and accompanying eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, then $\mathbf{A}^2$ has eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$ and eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. If $\mathbf{A}$ is nonsingular, then $\mathbf{A}^{-1}$ has eigenvalues $1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n$ and eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. If $\mathbf{A}$ is also symmetric, then

$$\mathbf{A}^2 = \mathbf{C}\mathbf{D}^2\mathbf{C}', \tag{2.115}$$

$$\mathbf{A}^{-1} = \mathbf{C}\mathbf{D}^{-1}\mathbf{C}', \tag{2.116}$$

where $\mathbf{C} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ has as columns the normalized eigenvectors of $\mathbf{A}$ (and of $\mathbf{A}^2$ and $\mathbf{A}^{-1}$), $\mathbf{D}^2 = \text{diag}(\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2)$, and $\mathbf{D}^{-1} = \text{diag}(1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n)$.

2.11.10 Singular Value Decomposition

In (2.109) in Section 2.11.7, we expressed a symmetric matrix $\mathbf{A}$ in terms of its eigenvalues and eigenvectors in the spectral decomposition $\mathbf{A} = \mathbf{C}\mathbf{D}\mathbf{C}'$. In a similar manner, we can express any (real) matrix $\mathbf{A}$ in terms of eigenvalues and eigenvectors of $\mathbf{A}'\mathbf{A}$ and $\mathbf{A}\mathbf{A}'$. Let $\mathbf{A}$ be an $n \times p$ matrix of rank k. Then the *singular value decomposition* of $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}', \tag{2.117}$$

where $\mathbf{U}$ is $n \times k$, $\mathbf{D}$ is $k \times k$, and $\mathbf{V}$ is $p \times k$. The diagonal elements of the nonsingular diagonal matrix $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$ are the positive square roots of

$\lambda_1^2, \lambda_2^2, \ldots, \lambda_k^2$, which are the nonzero eigenvalues of $\mathbf{A}'\mathbf{A}$ or of $\mathbf{A}\mathbf{A}'$. The values $\lambda_1, \lambda_2, \ldots, \lambda_k$ are called the *singular values* of $\mathbf{A}$. The k columns of $\mathbf{U}$ are the normalized eigenvectors of $\mathbf{A}\mathbf{A}'$ corresponding to the eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_k^2$. The k columns of $\mathbf{V}$ are the normalized eigenvectors of $\mathbf{A}'\mathbf{A}$ corresponding to the eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_k^2$. Since the columns of $\mathbf{U}$ and of $\mathbf{V}$ are (normalized) eigenvectors of symmetric matrices, they are mutually orthogonal (see Section 2.11.6), and we have $\mathbf{U}'\mathbf{U} = \mathbf{V}'\mathbf{V} = \mathbf{I}$.

PROBLEMS

2.1 Let

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 3 \\ 7 & 5 & 8 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 3 & -2 & 4 \\ 6 & 9 & -5 \end{pmatrix}.$$

(a) Find $\mathbf{A} + \mathbf{B}$ and $\mathbf{A} - \mathbf{B}$.

(b) Find $\mathbf{A}'\mathbf{A}$ and $\mathbf{A}\mathbf{A}'$.

2.2 Use the matrices $\mathbf{A}$ and $\mathbf{B}$ in Problem 2.1:

(a) Find $(\mathbf{A} + \mathbf{B})'$ and $\mathbf{A}' + \mathbf{B}'$ and compare them, thus illustrating (2.15).

(b) Show that $(\mathbf{A}')' = \mathbf{A}$, thus illustrating (2.6).

2.3 Let

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 0 \\ 1 & 5 \end{pmatrix}.$$

(a) Find $\mathbf{AB}$ and $\mathbf{BA}$.

(b) Find $|\mathbf{AB}|$, $|\mathbf{A}|$, and $|\mathbf{B}|$ and verify that (2.89) holds in this case.

2.4 Use the matrices $\mathbf{A}$ and $\mathbf{B}$ in Problem 2.3:

(a) Find $\mathbf{A} + \mathbf{B}$ and $\mathrm{tr}(\mathbf{A} + \mathbf{B})$.

(b) Find $\mathrm{tr}(\mathbf{A})$ and $\mathrm{tr}(\mathbf{B})$ and show that (2.96) holds for these matrices.

2.5 Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 3 & -2 \\ 2 & 0 \\ -1 & 1 \end{pmatrix}.$$

(a) Find $\mathbf{AB}$ and $\mathbf{BA}$.

(b) Compare $\mathrm{tr}(\mathbf{AB})$ and $\mathrm{tr}(\mathbf{BA})$ and confirm that (2.97) holds here.

2.6 Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 5 & 10 & 15 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -1 & 1 & -2 \\ -1 & 1 & -2 \\ 1 & -1 & 2 \end{pmatrix}.$$

(a) Show that $\mathbf{AB} = \mathbf{O}$.

(b) Find a vector $\mathbf{x}$ such that $\mathbf{Ax} = \mathbf{0}$.

(c) Show that $|\mathbf{A}| = 0$.

2.7 Let

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 4 \\ -1 & 1 & 3 \\ 4 & 3 & 2 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 3 & -2 & 4 \\ 7 & 1 & 0 \\ 2 & 3 & 5 \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Find the following:

(a) $\mathbf{Bx}$
(b) $\mathbf{y}'\mathbf{B}$
(c) $\mathbf{x}'\mathbf{Ax}$
(d) $\mathbf{x}'\mathbf{Ay}$
(e) $\mathbf{x}'\mathbf{x}$
(f) $\mathbf{x}'\mathbf{y}$
(g) $\mathbf{xx}'$
(h) $\mathbf{xy}'$
(i) $\mathbf{B}'\mathbf{B}$

2.8 Use $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{A}$ as defined in Problem 2.7:

(a) Find $\mathbf{x} + \mathbf{y}$ and $\mathbf{x} - \mathbf{y}$.

(b) Find $(\mathbf{x} - \mathbf{y})'\mathbf{A}(\mathbf{x} - \mathbf{y})$.

2.9 Using $\mathbf{B}$ and $\mathbf{x}$ in Problem 2.7, find $\mathbf{Bx}$ as a linear combination of columns of $\mathbf{B}$ as in (2.67) and compare with $\mathbf{Bx}$ found in Problem 2.7(a).

2.10 Let

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 4 & 2 \\ 5 & 0 & 3 \end{pmatrix}, \qquad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(a) Show that $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ as in (2.27).

(b) Show that $\mathbf{AI} = \mathbf{A}$ and that $\mathbf{IB} = \mathbf{B}$.

(c) Find $|\mathbf{A}|$.

2.11 Let

$$\mathbf{a} = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

(a) Find $\mathbf{a}'\mathbf{b}$ and $(\mathbf{a}'\mathbf{b})^2$.

(b) Find $\mathbf{bb}'$ and $\mathbf{a}'(\mathbf{bb}')\mathbf{a}$.

(c) Compare $(\mathbf{a}'\mathbf{b})^2$ with $\mathbf{a}'(\mathbf{bb}')\mathbf{a}$ and thus illustrate (2.40).

2.12 Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}.$$

Find **DA**, **AD**, and **DAD**.

2.13 Let the matrices **A** and **B** be partitioned as follows:

$$\mathbf{A} = \left(\begin{array}{cc|c} 2 & 1 & 2 \\ 3 & 2 & 0 \\ \hline 1 & 0 & 1 \end{array}\right), \qquad \mathbf{B} = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & 1 & 2 \\ \hline 2 & 3 & 1 & 2 \end{array}\right).$$

(a) Find **AB** as in (2.65) using the indicated partitioning.

(b) Check by finding **AB** in the usual way, ignoring the partitioning.

2.14 Let

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 0 & -1 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathbf{C} = \begin{pmatrix} 2 & 1 & 1 \\ 5 & -6 & -4 \end{pmatrix}.$$

Find **AB** and **CB**. Are they equal? What is the rank of **A**, **B**, and **C**?

2.15 Let

$$\mathbf{A} = \begin{pmatrix} 5 & 4 & 4 \\ 2 & -3 & 1 \\ 3 & 7 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

(a) Find $\text{tr}(\mathbf{A})$ and $\text{tr}(\mathbf{B})$.

(b) Find $\mathbf{A} + \mathbf{B}$ and $\text{tr}(\mathbf{A} + \mathbf{B})$. Is $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$?

(c) Find $|\mathbf{A}|$ and $|\mathbf{B}|$.

(d) Find **AB** and $|\mathbf{AB}|$. Is $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$?

2.16 Let

$$\mathbf{A} = \begin{pmatrix} 3 & 4 & 3 \\ 4 & 8 & 6 \\ 3 & 6 & 9 \end{pmatrix}.$$

(a) Show that $|\mathbf{A}| > 0$.

(b) Using the Cholesky decomposition in Section 2.7, find an upper triangular matrix **T** such that $\mathbf{A} = \mathbf{T}'\mathbf{T}$.

2.17 Let

$$\mathbf{A} = \begin{pmatrix} 3 & -5 & -1 \\ -5 & 13 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

(a) Show that $|\mathbf{A}| > 0$.

(b) Using the Cholesky decomposition in Section 2.7, find an upper triangular matrix $\mathbf{T}$ such that $\mathbf{A} = \mathbf{T}'\mathbf{T}$.

2.18 The columns of the following matrix are mutually orthogonal:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 0 \\ 1 & -1 & -1 \end{pmatrix}.$$

(a) Normalize the columns of $\mathbf{A}$ by dividing each column by its length; denote the resulting matrix by $\mathbf{C}$.

(b) Show that $\mathbf{C}$ is an orthogonal matrix, that is, $\mathbf{C}'\mathbf{C} = \mathbf{C}\mathbf{C}' = \mathbf{I}$.

2.19 Let

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

(a) Find the eigenvalues and associated normalized eigenvectors.

(b) Find $\text{tr}(\mathbf{A})$ and $|\mathbf{A}|$ and show that $\text{tr}(\mathbf{A}) = \sum_{i=1}^{3} \lambda_i$ and $|\mathbf{A}| = \prod_{i=1}^{3} \lambda_i$.

2.20 Let

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}.$$

(a) The eigenvalues of $\mathbf{A}$ are $1, 4, -2$. Find the normalized eigenvectors and use them as columns in an orthogonal matrix $\mathbf{C}$.

(b) Show that $\mathbf{C}'\mathbf{A}\mathbf{C} = \mathbf{D}$ as in (2.111), where $\mathbf{D}$ is diagonal with the eigenvalues of $\mathbf{A}$ on the diagonal.

(c) Show that $\mathbf{A} = \mathbf{C}\mathbf{D}\mathbf{C}'$ as in (2.109).

2.21 For the positive definite matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

calculate the eigenvalues and eigenvectors and find the square root matrix $\mathbf{A}^{1/2}$ as in (2.112). Check by showing that $(\mathbf{A}^{1/2})^2 = \mathbf{A}$.

2.22 Let

$$\mathbf{A} = \begin{pmatrix} 3 & 6 & -1 \\ 6 & 9 & 4 \\ -1 & 4 & 3 \end{pmatrix}.$$

(a) Find the spectral decomposition of $\mathbf{A}$ as in (2.109).

(b) Find the spectral decomposition of $\mathbf{A}^2$ and show that the diagonal matrix of eigenvalues is equal to the square of the matrix $\mathbf{D}$ found in part (a), thus illustrating (2.115).

(c) Find the spectral decomposition of $\mathbf{A}^{-1}$ and show that the diagonal matrix of eigenvalues is equal to the inverse of the matrix $\mathbf{D}$ found in part (a), thus illustrating (2.116).

2.23 Find the singular value decomposition of $\mathbf{A}$ as in (2.117), where

$$\mathbf{A} = \begin{pmatrix} 4 & -5 & -1 \\ 7 & -2 & 3 \\ -1 & 4 & -3 \\ 8 & 2 & 6 \end{pmatrix}.$$

2.24 If $\mathbf{j}$ is a vector of 1's, as defined in (2.11), show that the following hold:

(a) $\mathbf{j}'\mathbf{a} = \mathbf{a}'\mathbf{j} = \sum_i a_i$ as in (2.37).

(b) $\mathbf{j}'\mathbf{A}$ is a row vector whose elements are the column sums of $\mathbf{A}$ as in (2.38).

(c) $\mathbf{Aj}$ is a column vector whose elements are the row sums of $\mathbf{A}$ as in (2.38).

2.25 Verify (2.41); that is, show that $(\mathbf{x} - \mathbf{y})'(\mathbf{x} - \mathbf{y}) = \mathbf{x}'\mathbf{x} - 2\mathbf{x}'\mathbf{y} + \mathbf{y}'\mathbf{y}$.

2.26 Show that $\mathbf{A}'\mathbf{A}$ is symmetric, where $\mathbf{A}$ is $n \times p$.

2.27 If $\mathbf{a}$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are all $p \times 1$ and $\mathbf{A}$ is $p \times p$, show that (2.42)–(2.45) hold:

(a) $\sum_{i=1}^n \mathbf{a}'\mathbf{x}_i = \mathbf{a}' \sum_{i=1}^n \mathbf{x}_i$.

(b) $\sum_{i=1}^n \mathbf{A}\mathbf{x}_i = \mathbf{A} \sum_{i=1}^n \mathbf{x}_i$.

(c) $\sum_{i=1}^n (\mathbf{a}'\mathbf{x}_i)^2 = \mathbf{a}'(\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i')\mathbf{a}$.

(d) $\sum_{i=1}^n \mathbf{A}\mathbf{x}_i(\mathbf{A}\mathbf{x}_i)' = \mathbf{A}(\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i')\mathbf{A}'$.

2.28 Assume that $\mathbf{A} = \binom{\mathbf{a}_1'}{\mathbf{a}_2'}$ is $2 \times p$, $\mathbf{x}$ is $p \times 1$, and $\mathbf{S}$ is $p \times p$.

(a) Show that

$$\mathbf{Ax} = \begin{pmatrix} \mathbf{a}_1'\mathbf{x} \\ \mathbf{a}_2'\mathbf{x} \end{pmatrix},$$

as in (2.49).

(b) Show that

$$\mathbf{ASA}' = \begin{pmatrix} \mathbf{a}_1'\mathbf{S}\mathbf{a}_1 & \mathbf{a}_1'\mathbf{S}\mathbf{a}_2 \\ \mathbf{a}_2'\mathbf{S}\mathbf{a}_1 & \mathbf{a}_2'\mathbf{S}\mathbf{a}_2 \end{pmatrix},$$

as in (2.50).

2.29 **(a)** If the rows of $\mathbf{A}$ are denoted by $\mathbf{a}_i'$, show that $\mathbf{A}'\mathbf{A} = \sum_{i=1}^n \mathbf{a}_i\mathbf{a}_i'$ as in (2.51).

(b) If the columns of $\mathbf{A}$ are denoted by $\mathbf{a}_{(j)}$, show that $\mathbf{AA}' = \sum_{j=1}^p \mathbf{a}_{(j)}\mathbf{a}_{(j)}'$ as in (2.53).

2.30 Show that $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$ as in (2.75).

2.31 Show that the inverse of the partitioned matrix given in (2.76) is correct by multiplying by

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{a}_{12} \\ \mathbf{a}_{12}' & a_{22} \end{pmatrix}$$

to obtain an identity.

2.32 Show that the inverse of $\mathbf{B} + \mathbf{cc}'$ given in (2.77) is correct by multiplying by $\mathbf{B} + \mathbf{cc}'$ to obtain an identity.

2.33 Show that $|c\mathbf{A}| = c^n|\mathbf{A}|$ as in (2.85).

2.34 Show that $|\mathbf{A}^{-1}| = 1/|\mathbf{A}|$ as in (2.91).

2.35 If $\mathbf{B}$ is nonsingular and $\mathbf{c}$ is a vector, show that $|\mathbf{B} + \mathbf{cc}'| = |\mathbf{B}|(1 + \mathbf{c}'\mathbf{B}^{-1}\mathbf{c})$ as in (2.95).

2.36 Show that $\mathrm{tr}(\mathbf{A}'\mathbf{A}) = \mathrm{tr}(\mathbf{AA}') = \sum_{ij} a_{ij}^2$ as in (2.98).

2.37 Show that $\mathbf{CC}' = \mathbf{I}$ in (2.102) follows from $\mathbf{C}'\mathbf{C} = \mathbf{I}$ in (2.101).

2.38 Show that the eigenvalues of $\mathbf{AB}$ are the same as those of $\mathbf{BA}$, as noted in Section 2.11.5.

2.39 If $\mathbf{A}^{1/2}$ is the square root matrix defined in (2.112), show that

(a) $(\mathbf{A}^{1/2})^2 = \mathbf{A}$ as in (2.114),

(b) $|\mathbf{A}^{1/2}|^2 = |\mathbf{A}|$,

(c) $|\mathbf{A}^{1/2}| = |\mathbf{A}|^{1/2}$.