2

Submatrices and Partitioned Matrices

Two very important (and closely related) concepts are introduced in this chapter: that of a submatrix and that of a partitioned matrix. These concepts arise very naturally in statistics (especially in multivariate analysis and linear models) and in many other disciplines that involve probabilistic ideas. And results on submatrices and partitioned matrices, which can be found in Chapters 8, 9, 13, and 14 (and other of the subsequent chapters), have proved to be very useful. In particular, such results are almost indispensable in work involving the multivariate normal distribution—refer, for example, to Searle (1971, sec. 2.4f).

2.1 Some Terminology and Basic Results

A *submatrix* of a matrix **A** is a matrix that can be obtained by striking out rows and/or columns of **A**. For example, if we strike out the second row of the matrix

$$\begin{pmatrix} 2 & 4 & 3 & 6 \\ 1 & 5 & 7 & 9 \\ -1 & 0 & 2 & 2 \end{pmatrix},$$

we obtain the 2×4 submatrix

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

Alternatively, if we strike out the first and third columns, we obtain the 3×2 submatrix

$$\begin{pmatrix} 4 & 6 \\ 5 & 9 \\ 0 & 2 \end{pmatrix};$$

or, if we strike out the second row and the first and third columns, we obtain the 2×2 submatrix

$$\begin{pmatrix} 4 & 6 \\ 0 & 2 \end{pmatrix}.$$

Note that any matrix is a submatrix of itself; it is the submatrix obtained by striking out zero rows and zero columns.

Submatrices of a row or column vector, that is, of a matrix having one row or column, are themselves row or column vectors and are customarily referred to as *subvectors*.

Let $\mathbf{A}_*$ represent an $r \times s$ submatrix of an $m \times n$ matrix $\mathbf{A}$ obtained by striking out the $i_1, \ldots, i_{m-r}$th rows and $j_1, \ldots, j_{n-s}$th columns (of $\mathbf{A}$), and let $\mathbf{B}_*$ represent the $s \times r$ submatrix of $\mathbf{A}'$ obtained by striking out the $j_1, \ldots, j_{n-s}$th rows and $i_1, \ldots, i_{m-r}$th columns (of $\mathbf{A}'$). Then,

$$\mathbf{B}_* = \mathbf{A}'_*, \tag{1.1}$$

as is easily verified.

A submatrix of an $n \times n$ matrix is called a *principal submatrix* if it can be obtained by striking out the same rows as columns (so that the ith row is struck out whenever the ith column is struck out, and vice versa). The $r \times r$ (principal) submatrix of an $n \times n$ matrix obtained by striking out its last $n - r$ rows and columns is referred to as a *leading principal submatrix* ($r = 1, \ldots, n$). A principal submatrix of a symmetric matrix is symmetric, a principal submatrix of a diagonal matrix is diagonal, and a principal submatrix of an upper or lower triangular matrix is respectively upper or lower triangular, as is easily verified.

A matrix can be divided or partitioned into submatrices by drawing horizontal or vertical lines between various of its rows or columns, in which case the matrix is called a *partitioned matrix* and the submatrices are sometimes referred to as *blocks* (as in blocks of elements). For example,

$$\left(\begin{array}{r|rr|r} 2 & 4 & 3 & 6 \\ 1 & 5 & 7 & 9 \\ -1 & 0 & 2 & 2 \end{array}\right), \quad \left(\begin{array}{rrrr} 2 & 4 & 3 & 6 \\ 1 & 5 & 7 & 9 \\ \hline -1 & 0 & 2 & 2 \end{array}\right), \quad \left(\begin{array}{r|rr|r} 2 & 4 & 3 & 6 \\ 1 & 5 & 7 & 9 \\ \hline -1 & 0 & 2 & 2 \end{array}\right)$$

are various partitionings of the same matrix.

In effect, a partitioned $m \times n$ matrix is an $m \times n$ matrix $\mathbf{A} = \{a_{ij}\}$ that has been reexpressed in the general form

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \ldots & \mathbf{A}_{1c} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \ldots & \mathbf{A}_{2c} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \ldots & \mathbf{A}_{rc} \end{pmatrix}.$$

Here, $\mathbf{A}_{ij}$ is an $m_i \times n_j$ matrix ($i = 1, \ldots, r$; $j = 1, \ldots, c$), where $m_1, \ldots, m_r$ and $n_1, \ldots, n_c$ are positive integers such that $m_1 + \cdots + m_r = m$ and $n_1 + \cdots + n_c = n$. Or, more explicitly,

$$\mathbf{A}_{ij} = \begin{pmatrix} a_{m_1+\cdots+m_{i-1}+1,\, n_1+\cdots+n_{j-1}+1} & \dots & a_{m_1+\cdots+m_{i-1}+1,\, n_1+\cdots+n_j} \\ \vdots & & \vdots \\ a_{m_1+\cdots+m_i,\, n_1+\cdots+n_{j-1}+1} & \dots & a_{m_1+\cdots+m_i,\, n_1+\cdots+n_j} \end{pmatrix}.$$

(When $i = 1$ or $j = 1$, interpret the degenerate sum $m_1 + \cdots + m_{i-1}$ or $n_1 + \cdots + n_{j-1}$ as zero.) Thus, a partitioned matrix can be regarded as an array or "matrix" of matrices.

Note that a matrix that has been divided by "staggered" lines, for example,

$$\left(\begin{array}{r|rr|r} 2 & 4 & 3 & 6 \\ \cline{4-4} 1 & 5 & 7 & 9 \\ \cline{1-3} -1 & 0 & 2 & 2 \end{array}\right),$$

does not satisfy our definition of a partitioned matrix. Thus, if a matrix, say

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1c} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2c} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \dots & \mathbf{A}_{rc} \end{pmatrix},$$

is introduced as a partitioned matrix, there is an implicit assumption that each of the submatrices $\mathbf{A}_{i1}, \mathbf{A}_{i2}, \dots, \mathbf{A}_{ic}$ in the ith "row" of submatrices contains the same number of rows ($i = 1, 2, \dots, r$) and similarly that each of the submatrices $\mathbf{A}_{1j}, \mathbf{A}_{2j}, \dots, \mathbf{A}_{rj}$ in the jth "column" of submatrices contains the same number of columns.

It is customary to identify each of the blocks in a partitioned matrix by referring to the row of blocks and the column of blocks in which it appears. Thus, the submatrix $\mathbf{A}_{ij}$ is referred to as the ijth block of the partitioned matrix

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1c} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2c} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \dots & \mathbf{A}_{rc} \end{pmatrix}.$$

In the case of a partitioned $m \times n$ matrix $\mathbf{A}$ of the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1r} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \dots & \mathbf{A}_{rr} \end{pmatrix} \tag{1.2}$$

(for which the number of rows of blocks equals the number of columns of blocks), the ijth block $\mathbf{A}_{ij}$ of $\mathbf{A}$ is called a *diagonal block* if $j = i$ and an *off-diagonal block* if $j \neq i$. If all of the off-diagonal blocks of $\mathbf{A}$ are null matrices, that is, if

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & & \mathbf{0} \\ \vdots & & \ddots & \\ \mathbf{0} & \mathbf{0} & & \mathbf{A}_{rr} \end{pmatrix},$$

then $\mathbf{A}$ is called a *block-diagonal matrix*, and sometimes $\mathrm{diag}(\mathbf{A}_{11}, \mathbf{A}_{22}, \dots, \mathbf{A}_{rr})$ is written for $\mathbf{A}$. If $\mathbf{A}_{ij} = \mathbf{0}$ for $j < i = 1, \dots, r$, that is, if

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1r} \\ \mathbf{0} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2r} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{A}_{rr} \end{pmatrix},$$

then $\mathbf{A}$ is called an *upper block-triangular matrix*. Similarly, if $\mathbf{A}_{ij} = \mathbf{0}$ for $j > i = 1, \dots, r$, that is, if

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & & \mathbf{0} \\ \vdots & \vdots & \ddots & \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & & \mathbf{A}_{rr} \end{pmatrix},$$

then $\mathbf{A}$ is called a *lower block-triangular matrix*. To indicate that $\mathbf{A}$ is either upper or lower block-triangular (without being more specific), $\mathbf{A}$ is referred to simply as *block-triangular*.

Note that a partitioned $m \times n$ matrix $\mathbf{A}$ of the form (1.2) is block-diagonal if and only if it is both upper block-triangular and lower block-triangular. Note also that, if $m = n = r$ (in which case each block of $\mathbf{A}$ consists of a single element), saying that $\mathbf{A}$ is block diagonal or upper or lower block triangular is equivalent to saying that $\mathbf{A}$ is diagonal or upper or lower triangular.

Partitioned matrices having one row or one column are customarily referred to as *partitioned (row or column) vectors*. Thus, a partitioned m-dimensional column vector is an $m \times 1$ vector $\mathbf{a} = \{a_t\}$ that has been reexpressed in the general form

$$\begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_r \end{pmatrix}.$$

Here, $\mathbf{a}_i$ is an $m_i \times 1$ vector with elements $a_{m_1+\cdots+m_{i-1}+1}, \dots, a_{m_1+\cdots+m_{i-1}+m_i}$, respectively ($i = 1, \dots, r$), where $m_1, \dots, m_r$ are positive integers such that $m_1 + \cdots + m_r = m$. Similarly, a partitioned m-dimensional row vector is a $1 \times m$ vector $\mathbf{a}' = \{a_t\}$ that has been reexpressed in the general form $(\mathbf{a}'_1, \mathbf{a}'_2, \dots, \mathbf{a}'_r)$.

2.2 Scalar Multiples, Transposes, Sums, and Products of Partitioned Matrices

Let

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1c} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2c} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \dots & \mathbf{A}_{rc} \end{pmatrix}$$

represent a partitioned $m \times n$ matrix whose ijth block $\mathbf{A}_{ij}$ is of dimensions $m_i \times n_j$. Clearly, for any scalar k,

$$k\mathbf{A} = \begin{pmatrix} k\mathbf{A}_{11} & k\mathbf{A}_{12} & \dots & k\mathbf{A}_{1c} \\ k\mathbf{A}_{21} & k\mathbf{A}_{22} & \dots & k\mathbf{A}_{2c} \\ \vdots & \vdots & & \vdots \\ k\mathbf{A}_{r1} & k\mathbf{A}_{r2} & \dots & k\mathbf{A}_{rc} \end{pmatrix}. \tag{2.1}$$

In particular,

$$-\mathbf{A} = \begin{pmatrix} -\mathbf{A}_{11} & -\mathbf{A}_{12} & \dots & -\mathbf{A}_{1c} \\ -\mathbf{A}_{21} & -\mathbf{A}_{22} & \dots & -\mathbf{A}_{2c} \\ \vdots & \vdots & & \vdots \\ -\mathbf{A}_{r1} & -\mathbf{A}_{r2} & \dots & -\mathbf{A}_{rc} \end{pmatrix}. \tag{2.2}$$

Further, it is a simple exercise to show that

$$\mathbf{A}' = \begin{pmatrix} \mathbf{A}'_{11} & \mathbf{A}'_{21} & \dots & \mathbf{A}'_{r1} \\ \mathbf{A}'_{12} & \mathbf{A}'_{22} & \dots & \mathbf{A}'_{r2} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}'_{1c} & \mathbf{A}'_{2c} & \dots & \mathbf{A}'_{rc} \end{pmatrix}; \tag{2.3}$$

that is, $\mathbf{A}'$ can be expressed as a partitioned matrix, comprising c rows and r columns of blocks, the ijth of which is the transpose $\mathbf{A}'_{ji}$ of the jith block $\mathbf{A}_{ji}$ of $\mathbf{A}$.

Now, let

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \dots & \mathbf{B}_{1v} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \dots & \mathbf{B}_{2v} \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_{u1} & \mathbf{B}_{u2} & \dots & \mathbf{B}_{uv} \end{pmatrix}$$

represent a partitioned $p \times q$ matrix whose ijth block $\mathbf{B}_{ij}$ is of dimensions $p_i \times q_j$.

The matrices $\mathbf{A}$ and $\mathbf{B}$ are conformal (for addition) provided that $p = m$ and $q = n$. If $u = r$, $v = c$, $p_i = m_i$ $(i = 1, \dots, r)$, and $q_j = n_j$ $(j = 1, \dots, c)$, that is, if (besides $\mathbf{A}$ and $\mathbf{B}$ being conformal for addition) the rows and columns of $\mathbf{B}$ are partitioned in the same way as those of $\mathbf{A}$, then

$$\mathbf{A}+\mathbf{B}=\begin{pmatrix}\mathbf{A}_{11}+\mathbf{B}_{11} & \mathbf{A}_{12}+\mathbf{B}_{12} & \dots & \mathbf{A}_{1c}+\mathbf{B}_{1c}\\ \mathbf{A}_{21}+\mathbf{B}_{21} & \mathbf{A}_{22}+\mathbf{B}_{22} & \dots & \mathbf{A}_{2c}+\mathbf{B}_{2c}\\ \vdots & \vdots & & \vdots\\ \mathbf{A}_{r1}+\mathbf{B}_{r1} & \mathbf{A}_{r2}+\mathbf{B}_{r2} & \dots & \mathbf{A}_{rc}+\mathbf{B}_{rc}\end{pmatrix}, \tag{2.4}$$

and the partitioning of **A** and **B** is said to be *conformal* (for addition). This result and terminology extend in an obvious way to the addition of any finite number of partitioned matrices.

If **A** and **B** are conformal (for addition), then

$$\mathbf{A}-\mathbf{B}=\begin{pmatrix}\mathbf{A}_{11}-\mathbf{B}_{11} & \mathbf{A}_{12}-\mathbf{B}_{12} & \dots & \mathbf{A}_{1c}-\mathbf{B}_{1c}\\ \mathbf{A}_{21}-\mathbf{B}_{21} & \mathbf{A}_{22}-\mathbf{B}_{22} & \dots & \mathbf{A}_{2c}-\mathbf{B}_{2c}\\ \vdots & \vdots & & \vdots\\ \mathbf{A}_{r1}-\mathbf{B}_{r1} & \mathbf{A}_{r2}-\mathbf{B}_{r2} & \dots & \mathbf{A}_{rc}-\mathbf{B}_{rc}\end{pmatrix}, \tag{2.5}$$

as is evident from results (2.4) and (2.2).

The matrix product **AB** is defined provided that $n = p$. If $c = u$ and $n_k = p_k$ $(k = 1, \dots, c)$ [in which case all of the products $\mathbf{A}_{ik}\mathbf{B}_{kj}$ $(i = 1, \dots, r;\ j = 1, \dots, v;\ k = 1, \dots, c)$, as well as the product **AB**, exist], then

$$\mathbf{AB}=\begin{pmatrix}\mathbf{F}_{11} & \mathbf{F}_{12} & \dots & \mathbf{F}_{1v}\\ \mathbf{F}_{21} & \mathbf{F}_{22} & \dots & \mathbf{F}_{2v}\\ \vdots & \vdots & & \vdots\\ \mathbf{F}_{r1} & \mathbf{F}_{r2} & \dots & \mathbf{F}_{rv}\end{pmatrix}, \tag{2.6}$$

where $\mathbf{F}_{ij} = \sum_{k=1}^{c} \mathbf{A}_{ik}\mathbf{B}_{kj} = \mathbf{A}_{i1}\mathbf{B}_{1j} + \mathbf{A}_{i2}\mathbf{B}_{2j} + \cdots + \mathbf{A}_{ic}\mathbf{B}_{cj}$, and the partitioning of **A** and **B** is said to be *conformal* (for the premultiplication of **B** by **A**).

In the special case where $r = c = u = v = 2$, that is, where

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

result (2.6) simplifies to

$$\mathbf{AB}=\begin{pmatrix}\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12}+\mathbf{A}_{12}\mathbf{B}_{22}\\ \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12}+\mathbf{A}_{22}\mathbf{B}_{22}\end{pmatrix}. \tag{2.7}$$

If $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_c)$ is an $m \times n$ matrix that has been partitioned only by columns (for emphasis, we sometimes insert commas between the submatrices of such a partitioned matrix), then

$$\mathbf{A}'=\begin{pmatrix}\mathbf{A}_1'\\ \mathbf{A}_2'\\ \vdots\\ \mathbf{A}_c'\end{pmatrix}, \tag{2.8}$$

and further if

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_c \end{pmatrix}$$

is an $n \times q$ partitioned matrix that has been partitioned only by rows (in a way that is conformal for its premultiplication by $\mathbf{A}$), then

$$\mathbf{AB} = \sum_{k=1}^{c} \mathbf{A}_k \mathbf{B}_k = \mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2 + \cdots + \mathbf{A}_c\mathbf{B}_c. \tag{2.9}$$

Similarly, if

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_r \end{pmatrix}$$

is an $m \times n$ matrix that has been partitioned only by rows, then

$$\mathbf{A}' = (\mathbf{A}_1', \mathbf{A}_2', \ldots, \mathbf{A}_r'), \tag{2.10}$$

and further if $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_v)$ is an $n \times q$ matrix that has been partitioned only by columns, then

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_1\mathbf{B}_1 & \mathbf{A}_1\mathbf{B}_2 & \ldots & \mathbf{A}_1\mathbf{B}_v \\ \mathbf{A}_2\mathbf{B}_1 & \mathbf{A}_2\mathbf{B}_2 & \ldots & \mathbf{A}_2\mathbf{B}_v \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_r\mathbf{B}_1 & \mathbf{A}_r\mathbf{B}_2 & \ldots & \mathbf{A}_r\mathbf{B}_v \end{pmatrix}. \tag{2.11}$$

2.3 Some Results on the Product of a Matrix and a Column Vector

Let $\mathbf{A}$ represent an $m \times n$ matrix and $\mathbf{x}$ an $n \times 1$ vector. Writing $\mathbf{A}$ as $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, where $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are the columns of $\mathbf{A}$, and $\mathbf{x}$ as $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$, where $x_1, x_2, \ldots, x_n$ are the elements of $\mathbf{x}$, we find, as a special case of result (2.9), that

$$\mathbf{Ax} = \sum_{k=1}^{n} x_k \mathbf{a}_k = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n. \tag{3.1}$$

Thus, the effect of postmultiplying a matrix by a column vector is to form a linear combination of the columns of the matrix, where the coefficients in the linear combination are the elements of the column vector. Similarly, the effect of premultiplying a matrix by a row vector is to form a linear combination of the rows of the matrix, where the coefficients in the linear combination are the elements of the row vector.

Representation (3.1) is helpful in establishing the elementary results expressed in the following two lemmas.

Lemma 2.3.1. For any column vector $\mathbf{y}$ and nonnull column vector $\mathbf{x}$, there exists a matrix $\mathbf{A}$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$.

Proof. Since $\mathbf{x}$ is nonnull, one of its elements, say x_j, is nonzero. Take $\mathbf{A}$ to be the matrix whose jth column is $(1/x_j)\mathbf{y}$ and whose other columns are null. Then, $\mathbf{y} = \mathbf{A}\mathbf{x}$, as is evident from result (3.1). Q.E.D.

Lemma 2.3.2. For any two $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} = \mathbf{B}$ if and only if $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}$ for every $n \times 1$ vector $\mathbf{x}$.

Proof. It is obvious that, if $\mathbf{A} = \mathbf{B}$, then $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}$ for every vector $\mathbf{x}$.

To prove the converse, suppose that $\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}$ for every $\mathbf{x}$. Taking $\mathbf{x}$ to be the $n \times 1$ vector whose ith element is 1 and whose other elements are 0, and letting $\mathbf{a}_i$ and $\mathbf{b}_i$ represent the ith columns of $\mathbf{A}$ and $\mathbf{B}$, respectively, it is clear from result (3.1) that

$$\mathbf{a}_i = \mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x} = \mathbf{b}_i$$

$(i = 1, \ldots, n)$. We conclude that $\mathbf{A} = \mathbf{B}$. Q.E.D.

Note that Lemma 2.3.2 implies, in particular, that $\mathbf{A} = \mathbf{0}$ if and only if $\mathbf{A}\mathbf{x} = \mathbf{0}$ for every $\mathbf{x}$.

2.4 Expansion of a Matrix in Terms of Its Rows, Columns, or Elements

An $m \times n$ matrix $\mathbf{A} = \{a_{ij}\}$ can be expanded in terms of its rows, columns, or elements by making use of formula (2.9). Denote the ith row of $\mathbf{A}$ by $\mathbf{r}'_i$ and the ith column of $\mathbf{I}_m$ by $\mathbf{e}_i$ $(i = 1, 2, \ldots, m)$. Then, writing $\mathbf{I}_m$ as $\mathbf{I}_m = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m)$ and $\mathbf{A}$ as

$$\mathbf{A} = \begin{pmatrix} \mathbf{r}'_1 \\ \mathbf{r}'_2 \\ \vdots \\ \mathbf{r}'_m \end{pmatrix}$$

and applying formula (2.9) to the product $\mathbf{I}_m\mathbf{A}$, we obtain the expansion

$$\mathbf{A} = \sum_{i=1}^{m} \mathbf{e}_i\mathbf{r}'_i = \mathbf{e}_1\mathbf{r}'_1 + \mathbf{e}_2\mathbf{r}'_2 + \cdots + \mathbf{e}_m\mathbf{r}'_m. \tag{4.1}$$

Similarly, denote the jth column of $\mathbf{A}$ by $\mathbf{a}_j$ and the jth row of $\mathbf{I}_n$ by $\mathbf{u}'_j$ $(j = 1, 2, \ldots, n)$. Then, writing $\mathbf{A}$ as $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ and $\mathbf{I}_n$ as

$$\mathbf{I}_n = \begin{pmatrix} \mathbf{u}_1' \\ \mathbf{u}_2' \\ \vdots \\ \mathbf{u}_n' \end{pmatrix}$$

and applying formula (2.9) to the product $\mathbf{AI}_n$, we obtain the alternative expansion

$$\mathbf{A} = \sum_{j=1}^{n} \mathbf{a}_j \mathbf{u}_j' = \mathbf{a}_1 \mathbf{u}_1' + \mathbf{a}_2 \mathbf{u}_2' + \cdots + \mathbf{a}_n \mathbf{u}_n'. \tag{4.2}$$

Moreover, the application of formula (3.1) to the product $\mathbf{I}_m \mathbf{a}_j$ gives the expansion

$$\mathbf{a}_j = \sum_{i=1}^{m} a_{ij} \mathbf{e}_i .$$

Upon substituting this expansion into expansion (4.2), we obtain the further expansion

$$\mathbf{A} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \mathbf{U}_{ij}, \tag{4.3}$$

where $\mathbf{U}_{ij} = \mathbf{e}_i \mathbf{u}_j'$ is an $m \times n$ matrix whose ijth element equals 1 and whose remaining $mn - 1$ elements equal 0. In the special case where $n = m$ (i.e., where $\mathbf{A}$ is square), $\mathbf{u}_j = \mathbf{e}_j$ and hence $\mathbf{U}_{ij} = \mathbf{e}_i \mathbf{e}_j'$, and in the further special case where $\mathbf{A} = \mathbf{I}_m$, expansion (4.3) reduces to

$$\mathbf{I}_m = \sum_{i=1}^{m} \mathbf{e}_i \mathbf{e}_i'. \tag{4.4}$$

Note that, as a consequence of result (4.3), we have that

$$\mathbf{e}_i' \mathbf{A} \mathbf{u}_j = \mathbf{e}_i' \left(\sum_{k=1}^{m} \sum_{s=1}^{n} a_{ks} \mathbf{e}_k \mathbf{u}_s' \right) \mathbf{u}_j = \sum_{k=1}^{m} \sum_{s=1}^{n} a_{ks} \mathbf{e}_i' \mathbf{e}_k \mathbf{u}_s' \mathbf{u}_j ,$$

which (since $\mathbf{e}_i' \mathbf{e}_k$ equals 1, if $k = i$, and equals 0, if $k \neq i$, and since $\mathbf{u}_s' \mathbf{u}_j$ equals 1, if $s = j$, and equals 0, if $s \neq j$) simplifies to

$$\mathbf{e}_i' \mathbf{A} \mathbf{u}_j = a_{ij}. \tag{4.5}$$

Exercises

Section 2.1

1. Verify result (1.1).

2. Verify (a) that a principal submatrix of a symmetric matrix is symmetric, (b) that a principal submatrix of a diagonal matrix is diagonal, and (c) that a principal submatrix of an upper triangular matrix is upper triangular.

3. Let
$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1r} \\ \mathbf{0} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2r} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_{rr} \end{pmatrix}$$
represent an $n \times n$ upper block-triangular matrix whose ijth block $\mathbf{A}_{ij}$ is of dimensions $n_i \times n_j$ ($j \geq i = 1, \dots, r$). Show that $\mathbf{A}$ is upper triangular if and only if each of its diagonal blocks $\mathbf{A}_{11}, \mathbf{A}_{22}, \dots, \mathbf{A}_{rr}$ is upper triangular.

Section 2.2

4. Verify results (2.3) and (2.6).

Matrix Algebra From a Statistician's Perspective
Harville, D.A.
1997, XVI, 634 p., Softcover
ISBN: 978-0-387-78356-7