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

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

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

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 $\mathbf{A}\mathbf{j}$ 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 $\mathbf{A}\mathbf{j}$, 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

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

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:

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

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,

$\mathbf{A}'\mathbf{A}$ is $p \times p$ and is obtained as products of columns of $\mathbf{A}$ [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

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

$$\mathbf{A}'\mathbf{A} = \begin{pmatrix} \mathbf{a}'_{(1)} \\ \mathbf{a}'_{(2)} \\ \vdots \\ \mathbf{a}'_{(p)} \end{pmatrix} (\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \ldots, \mathbf{a}_{(p)})$$

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

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

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

$$= \begin{pmatrix} d_1a_{11} & d_1a_{12} & d_1a_{13} \\ d_2a_{21} & d_2a_{22} & d_2a_{23} \\ d_3a_{31} & d_3a_{32} & d_3a_{33} \end{pmatrix}, \tag{2.55}$$

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

$$= \begin{pmatrix} d_1a_{11} & d_2a_{12} & d_3a_{13} \\ d_1a_{21} & d_2a_{22} & d_3a_{23} \\ d_1a_{31} & d_2a_{32} & d_3a_{33} \end{pmatrix}, \tag{2.56}$$

$$\mathbf{DAD} = \begin{pmatrix} d_1^2a_{11} & d_1d_2a_{12} & d_1d_3a_{13} \\ d_2d_1a_{21} & d_2^2a_{22} & d_2d_3a_{23} \\ d_3d_1a_{31} & d_3d_2a_{32} & d_3^2a_{33} \end{pmatrix}. \tag{2.57}$$

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

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

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

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