2.2 Matrix Algebra

In the previous section we introduced the general idea of a matrix. The next step is to develop the algebra of matrices. Unless otherwise stated, we assume that all elements of the matrices that appear are real or complex numbers.

Addition and Subtraction of Matrices and Multiplication of a Matrix by a Scalar

Addition and subtraction of matrices is defined only for matrices with the same dimensions. We begin with addition.

**Definition 2.2.1**

If $A$ and $B$ are both $m \times n$ matrices, then we define the addition (or the sum) of $A$ and $B$, denoted by $A + B$, to be the $m \times n$ matrix whose elements are obtained by adding corresponding elements of $A$ and $B$. In index notation, if $A = [a_{ij}]$ and $B = [b_{ij}]$, then $A + B = [a_{ij} + b_{ij}]$.

**Example 2.2.2**

We have

$\begin{bmatrix} 2 & -1 & 3 \\ 4 & -5 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 5 \\ -5 & 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 8 \\ -1 & -3 & 7 \end{bmatrix}$.

Properties of Matrix Addition:

If $A$ and $B$ are both $m \times n$ matrices, then

- $A + B = B + A$ (matrix addition is commutative).
- $A + (B + C) = (A + B) + C$ (matrix addition is associative).

Both of these properties follow directly from Definition 2.2.1.

In order that we can model oscillatory physical phenomena, in much of the later work we will need to use complex as well as real numbers. Throughout the text we will use the term **scalar** to mean a real or complex number.

**Definition 2.2.3**

If $A$ is an $m \times n$ matrix and $s$ is a scalar, then we let $sA$ denote the matrix obtained by multiplying every element of $A$ by $s$. This procedure is called **scalar multiplication**.

In index notation, if $A = [a_{ij}]$, then $sA = [sa_{ij}]$.

**Example 2.2.4**

If $A = \begin{bmatrix} 2 & -1 \\ 4 & 6 \end{bmatrix}$, then $5A = \begin{bmatrix} 10 & -5 \\ 20 & 30 \end{bmatrix}$.
Example 2.2.5

If

\[ A = \begin{bmatrix} 1 + i & i \\ 2 & 3i & 4 \end{bmatrix} \]

and \( x = 1 - 2i \), where \( i = \sqrt{-1} \), find \( sA \).

Solution: We have

\[
\begin{align*}
    sA &= \begin{bmatrix} (1 - 2i)(1 + i) & (1 - 2i)i \\ (1 - 2i)(2 + 3i) & (1 - 2i)4 \end{bmatrix} \\
    &= \begin{bmatrix} 1 - i & 2 + i \\ 8 - i & 4 - 8i \end{bmatrix}.
\end{align*}
\]

□

DEFINITION 2.2.6

We define subtraction of two matrices with the same dimensions by

\[ A - B = A + (-1)B. \]

In index notation, \( A - B = [a_{ij} - b_{ij}] \). That is, we subtract corresponding elements.

Further properties satisfied by the operations of matrix addition and multiplication of a matrix by a scalar are as follows:

**Properties of Scalar Multiplication:** For any scalars \( s \) and \( t \), and for any matrices \( A \) and \( B \) of the same size,

- \( sA = A \) (unit property),
- \( s(A + B) = sA + sB \) (distributivity of scalars over matrix addition),
- \( (s + t)A = sA + tA \) (distributivity of scalar addition over matrices),
- \( s(tA) = (st)A = (ts)A = t(sA) \) (associativity of scalar multiplication).

The \( m \times n \) zero matrix, denoted \( 0_{m \times n} \) (or simply 0, if the dimensions are clear), is the \( m \times n \) matrix whose elements are all zeros. In the case of the \( n \times n \) zero matrix, we may write \( 0 \). We now collect a few properties of the zero matrix. The first of these below indicates that the zero matrix plays a similar role in matrix addition to that played by the number zero in the addition of real numbers.

**Properties of the Zero Matrix:** For all matrices \( A \) and the zero matrix of the same size, we have

- \( A + 0 = A \), \( A - A = 0 \), and \( 0A = 0 \).

Note that in the last property here, the zero on the left side of the equation is a scalar, while the zero on the right side of the equation is a matrix.

**Multiplication of Matrices**

The definition we introduced above for how to multiply a matrix by a scalar is essentially the only possibility if, in the case when \( s \) is a positive integer, we want \( sA \) to be the same matrix as the one obtained when \( A \) is added to itself \( s \) times. We now define how to multiply two matrices together. In this case the multiplication operation is by no means obvious. However, in Chapter 5 when we study linear transformations, the motivation for the matrix multiplication procedure we are defining here will become quite transparent (see Theorem 5.5.7).

We will build up to the general definition of matrix multiplication in three stages.

**Case 1: Product of a row \( n \)-vector and a column \( n \)-vector.** We begin by generalizing a concept from elementary calculus. If \( a \) and \( b \) are either row or column \( n \)-vectors, with
2.2 Matrix Algebra

components $a_1, a_2, \ldots, a_n$, and $b_1, b_2, \ldots, b_n$, respectively, then their dot product, denoted $\mathbf{a} \cdot \mathbf{b}$, is the number

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$  

As we will see, this is the key formula in defining the product of two matrices. Now let $\mathbf{a}$ be a row $n$-vector, and let $\mathbf{x}$ be a column $n$-vector. Then their matrix product $\mathbf{a}\mathbf{x}$ is defined to be the $1 \times 1$ matrix whose single element is obtained by taking the dot product of the row vectors $\mathbf{a}$ and $\mathbf{x}^T$. Thus,

$$\mathbf{a}\mathbf{x} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$  

**Example 2.2.7** If $\mathbf{a} = \begin{bmatrix} 2 & -1 & 3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$, then

$$\mathbf{a}\mathbf{x} = \begin{bmatrix} 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = (2)(3) + (-1)(-2) + (3)(1) = 15.$$  

Case 2: Product of an $m \times n$ matrix and a column $n$-vector. If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{x}$ is a column $n$-vector, then the product $\mathbf{Ax}$ is defined to be the $m \times 1$ matrix whose $i$th element is obtained by taking the dot product of the $i$th row vector of $\mathbf{A}$ with $\mathbf{x}$. (See Figure 2.2.1.)

The $i$th row vector of $\mathbf{A}$ is $\mathbf{a}_i = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$, so that $\mathbf{Ax}$ has $i$th element

$$(\mathbf{Ax})_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n.$$  

Consequently the column vector $\mathbf{Ax}$ has elements

$$(\mathbf{Ax})_i = \sum_{k=1}^{m} a_{ik}x_k, \quad 1 \leq i \leq m. \quad (2.2.1)$$  

Figure 2.2.1: Multiplication of an $m \times n$ matrix with a column $n$-vector.
As illustrated in the next example, in practice, we do not use the formula (2.2.1); rather, we explicitly take the matrix products of the row vectors of $A$ with the column vector $x$.

**Example 2.2.8** Find $Ax$ if $A = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 4 & -6 \\ 5 & -2 & 0 \end{bmatrix}$ and $x = \begin{bmatrix} 7 \\ -3 \\ 1 \end{bmatrix}$.

**Solution:** We have

\[
Ax = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 4 & -6 \\ 5 & -2 & 0 \end{bmatrix} \begin{bmatrix} 7 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -11 \\ 41 \end{bmatrix}.
\]

The following result regarding multiplication of a column vector by a matrix will be used repeatedly in later chapters.

**Theorem 2.2.9** If $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ is an $m \times n$ matrix and $c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ is a column $n$-vector, then

\[
A c = c_1 a_{11} + c_2 a_{12} + \cdots + c_n a_{1n}.
\]

**Proof** The element $a_{ik}$ of $A$ is the $i$th component of the column $m$-vector $a_k$, so

\[
a_{ik} = (a_k)_i.
\]

Applying formula (2.2.1) for multiplication of a column vector by a matrix yields

\[
(Ac)_i = \sum_{k=1}^{n} a_{ik}c_k = \sum_{k=1}^{n} (a_k)_i c_k = \sum_{k=1}^{n} (c_k a_k)_i.
\]

Consequently,

\[
Ac = \sum_{k=1}^{n} c_k a_k = c_1 a_1 + c_2 a_2 + \cdots + c_n a_n
\]

as required.

If $x_1$, $x_2$, \ldots, $x_n$ are column $m$-vectors and $c_1$, $c_2$, \ldots, $c_n$ are scalars, then an expression of the form

\[
c_1 x_1 + c_2 x_2 + \cdots + c_n x_n
\]

is called a linear combination of the column vectors. Therefore, from Equation (2.2.2), we see that the vector $Ac$ is obtained by taking a linear combination of the column vectors of $A$. For example, if

\[
A = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix} \quad \text{and} \quad c = \begin{bmatrix} 5 \\ -1 \end{bmatrix},
\]

then

\[
Ac = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ -1 \end{bmatrix} = \begin{bmatrix} 9 \\ 14 \end{bmatrix}.
\]
then

\[ A \mathbf{c} = c_1 a_1 + c_2 a_2 = \begin{bmatrix} 5 \\ 4 \end{bmatrix} + (-1) \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 17 \end{bmatrix}. \]

Case 3: Product of an \( m \times n \) matrix and an \( n \times p \) matrix. If \( A \) is an \( m \times n \) matrix and \( B \) is an \( n \times p \) matrix, then the product \( AB \) has columns defined by multiplying the matrix \( A \) by the respective column vectors of \( B \), as described in Case 2. That is, if 

\[ B = [b_1, b_2, \ldots, b_p], \]

then \( AB \) is the \( m \times p \) matrix defined by 

\[ AB = [A b_1, A b_2, \ldots, A b_p]. \]

**Example 2.2.10** If 

\[ A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \] and  

\[ B = \begin{bmatrix} 2 & 3 \\ 5 & 2 \end{bmatrix}, \]

determine \( AB \).

**Solution:** We have 

\[ AB = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} (1)(2) + (4)(5) + (2)(8) & (1)(3) + (4)(-2) + (2)(4) \\ (3)(2) + (5)(5) + (7)(8) & (3)(3) + (5)(-2) + (7)(4) \end{bmatrix} = \begin{bmatrix} 38 & 3 \\ 87 & 27 \end{bmatrix}. \]

**Example 2.2.11** If 

\[ A = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \] and  

\[ B = \begin{bmatrix} 2 & 4 \end{bmatrix}, \]

determine \( AB \).

**Solution:** We have 

\[ AB = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 4 \end{bmatrix} = \begin{bmatrix} (2)(2) & (2)(4) \\ (-1)(2) & (-1)(4) \\ (3)(2) & (3)(4) \end{bmatrix} = \begin{bmatrix} 4 & 8 \\ -2 & -4 \\ 6 & 12 \end{bmatrix}. \]

Another way to describe \( AB \) is to note that the element \( (AB)_{ij} \) is obtained by computing the matrix product of the \( i \)th row vector of \( A \) and the \( j \)th column vector of \( B \). That is, 

\[ (AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}. \]

Expressing this using the summation notation yields the following result:

**Definition 2.2.12** If 

\[ A = \begin{bmatrix} a_{ij} \end{bmatrix} \]

is an \( m \times n \) matrix, 

\[ B = \begin{bmatrix} b_{ij} \end{bmatrix} \]

is an \( n \times p \) matrix, and 

\[ C = AB, \]

then 

\[ c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \]

for \( 1 \leq i \leq m, \ 1 \leq j \leq p. \) (2.2.3)

This is called the index form of the matrix product.

The formula (2.2.3) for the \( ij \)th element of \( AB \) is very important and will often be required in the future. The reader should memorize it.

In order for the product \( AB \) to be defined, we see that \( A \) and \( B \) must satisfy
number of columns of $A = \text{number of rows of } B$.

In such a case, if $C$ represents the product matrix $AB$, then the relationship between the dimensions of the matrices is

$$a_{m,n} = b_{1,p} = c_{m,p}$$

Now we give some further examples of matrix multiplication.

**Example 2.2.13**

If $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -2 \\ 0 & 5 \end{bmatrix}$, then

$AB = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 9 \\ 5 & 16 \end{bmatrix}$.

**Example 2.2.14**

If $A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$, then

$AB = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 2 & 1 \end{bmatrix}$.

**Example 2.2.15**

If $A = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 4 \\ -1 & -6 \\ 3 & 12 \end{bmatrix}$, then

$AB = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ -1 & -6 \\ 3 & 12 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ -1 & -12 \\ 3 & 12 \end{bmatrix}$.

**Example 2.2.16**

If $A = \begin{bmatrix} 1 - i & i \\ 2 + i & 1 + i \end{bmatrix}$ and $B = \begin{bmatrix} 3 + 2i & 1 + 4i \\ i & -1 + 2i \end{bmatrix}$, then

$AB = \begin{bmatrix} 1 - i & i \\ 2 + i & 1 + i \end{bmatrix} \begin{bmatrix} 3 + 2i & 1 + 4i \\ i & -1 + 2i \end{bmatrix} = \begin{bmatrix} 4 - i & 3 + 2i \\ 3 + 8i & -5 + 10i \end{bmatrix}$.

Notice that in Examples 2.2.13 and 2.2.14 above, the product $BA$ is not defined, since the number of columns of the matrix $B$ does not agree with the number of rows of the matrix $A$.

We can now establish some basic properties of matrix multiplication.

**Theorem 2.2.17**

If $A$, $B$ and $C$ have appropriate dimensions for the operations to be performed, then

$A(BC) = (AB)C$ (associativity of matrix multiplication), (2.2.4)

$A(B + C) = AB + AC$ (left distributivity of matrix multiplication), (2.2.5)

$(A + B)C = AC + BC$ (right distributivity of matrix multiplication). (2.2.6)
Proof The idea behind the proof of each of these results is to use the definition of matrix multiplication to show that the $ij$th element of the matrix on the left-hand side of each equation is equal to the $ij$th element of the matrix on the right-hand side. We illustrate by proving (2.2.6), but we leave the proofs of (2.2.4) and (2.2.5) as exercises. Suppose that $A$ and $B$ are $m \times n$ matrices and that $C$ is an $n \times p$ matrix. Then, from Equation (2.2.3),

$$[(A + B)C]_{ij} = \sum_{k=1}^{n}(a_{ik} + b_{ik})c_{kj} = \sum_{k=1}^{n}a_{ik}c_{kj} + \sum_{k=1}^{n}b_{ik}c_{kj} = (AC)_{ij} + (BC)_{ij}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq p.$$ 

Consequently,

$$(A + B)C = AC + BC.$$ 

Theorem 2.2.17 states that matrix multiplication is associative and distributive (over addition). We now consider the question of commutativity of matrix multiplication. If $A$ is an $m \times n$ matrix and $B$ is an $n \times m$ matrix, we can form both of the products $AB$ and $BA$, which are $m \times m$ and $n \times n$, respectively. In the first of these, we say that $B$ has been premultiplied by $A$, whereas in the second, we say that $B$ has been postmultiplied by $A$. If $m \neq n$, then the matrices $AB$ and $BA$ will have different dimensions, so they cannot be equal. It is important to realize, however, that even if $m = n$, in general (that is, except for special cases)

$$AB \neq BA.$$ 

This is the statement that

matrix multiplication is not commutative.

With a little bit of thought this should not be too surprising, in view of the fact that the $ij$th element of $AB$ is obtained by taking the matrix product of the $i$th row vector of $A$ with the $j$th column vector of $B$, whereas the $ij$th element of $BA$ is obtained by taking the matrix product of the $i$th row vector of $B$ with the $j$th column vector of $A$. We illustrate with an example.

**Example 2.2.18** If $A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}$, find $AB$ and $BA$.

**Solution:** We have

$$AB = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 7 & -1 \\ 3 & -4 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 9 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 12 & 27 \\ 3 & -4 \end{bmatrix}.$$ 

Thus we see that in this example, $AB \neq BA$. □

As an exercise, the reader can calculate the matrix $BA$ in Examples 2.2.15 and 2.2.16 and again see that $AB \neq BA$.

For an $n \times n$ matrix we use the usual power notation to denote the operation of multiplying $A$ by itself. Thus,

$$A^2 = AA, \quad A^3 = AAA, \quad \text{and so on.}$$
The **identity matrix**, $I_n$ (or just $I$ if the dimensions are obvious), is the $n \times n$ matrix with ones on the main diagonal and zeros elsewhere. For example, 

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

**DEFINITION 2.2.19**

The elements of $I_n$ can be represented by the **Kronecker delta symbol**, $\delta_{ij}$, defined by

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Then, $I_n = [\delta_{ij}]$.

The following properties of the identity matrix indicate that it plays the same role in matrix multiplication as the number 1 does in the multiplication of real numbers.

**Properties of the Identity Matrix:**

1. $A_{m\times n}I_n = A_{m\times n}$.
2. $I_nA_{n\times p} = A_{n\times p}$.

**Proof** We establish property 1 and leave the proof of property 2 as an exercise (Problem 25). Using the index form of the matrix product, we have

$$(AI)_{ij} = \sum_{k=1}^{n} a_{ik} \delta_{kj} = a_{i1} \delta_{1j} + a_{i2} \delta_{2j} + \cdots + a_{in} \delta_{nj}.$$ 

But, from the definition of the Kronecker delta symbol, we see that all terms in the summation with $k \neq j$ vanish, so that we are left with

$$(AI)_{ij} = a_{ij} \delta_{jj} = a_{ij} \quad 1 \leq i \leq m, \quad 1 \leq j \leq n. \quad \blacksquare$$

The next example illustrates property 2 of the identity matrix.

**Example 2.2.20**

If $A = \begin{bmatrix} 2 & -1 \\ 3 & 5 \\ 0 & -2 \end{bmatrix}$, verify that $I_3A = A$.

**Solution:** We have

$$I_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 5 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 5 \\ 0 & -2 \end{bmatrix} = A. \quad \square$$
2.2 Matrix Algebra

Properties of the Transpose

The operation of taking the transpose of a matrix was introduced in the previous section. The next theorem gives three important properties satisfied by the transpose. These should be memorized.

**Theorem 2.2.21**
Let $A$ and $C$ be $m \times n$ matrices, and let $B$ be an $n \times p$ matrix. Then
1. $(A^T)^T = A$.
2. $(A + C)^T = A^T + C^T$.
3. $(AB)^T = B^T A^T$.

**Proof** For all three statements, our strategy is again to show that the $(i, j)$-elements of each side of the equation are the same. We prove statement 3 and leave the proofs of 1 and 2 for the exercises (Problem 24). From the definition of the transpose and the index form of the matrix product, we have

\[
[(AB)^T]_{ij} = (AB)_{ji} \quad \text{(definition of the transpose)}
\]

\[
= \sum_{k=1}^{n} a_{ik} b_{kj} \quad \text{(index form of the matrix product)}
\]

\[
= \sum_{k=1}^{n} b_{kj} a_{ik} = \sum_{k=1}^{n} b_{ji} a_{kj} = \sum_{k=1}^{n} a_{kj} b_{ji} = (B^T A^T)_{ij}.
\]

Consequently,

\[(AB)^T = B^T A^T.\]

Results for Triangular Matrices

Upper and lower triangular matrices play a significant role in the analysis of linear systems of equations. The following theorem and its corollary will be needed in Section 2.7.

**Theorem 2.2.22**
The product of two lower (upper) triangular matrices is a lower (upper) triangular matrix.

**Proof** Suppose that $A$ and $B$ are $n \times n$ lower triangular matrices. Then, $a_{ik} = 0$ whenever $i > k$, and $b_{kj} = 0$ whenever $k < j$. If we let $C = AB$, then we must prove that

\[c_{ij} = 0 \quad \text{whenever} \quad i < j.
\]

Using the index form of the matrix product, we have

\[c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=j}^{n} a_{ij} b_{kj} \quad \text{(since $b_{kj} = 0$ if $k < j$)}, \quad (2.2.7)\]
CHAPTER 2 Matrices and Systems of Linear Equations

We now impose the condition that \( i < j \). Then, since \( k \geq j \) in (2.2.7), it follows that \( k > i \). However, this implies that \( a_{ik} = 0 \) (since \( A \) is lower triangular), and hence, from (2.2.7), that

\[
\begin{align*}
c_{ij} &= 0 \quad \text{whenever} \quad i < j.
\end{align*}
\]

as required.

To establish the result for upper triangular matrices, either we can give an argument similar to that presented above for lower triangular matrices, or we can use the fact that the transpose of a lower triangular matrix is an upper triangular matrix, and vice versa. Hence, if \( A \) and \( B \) are \( n \times n \) upper triangular matrices, then \( A^T \) and \( B^T \) are lower triangular, and therefore by what we proved above, \((AB)^T = B^TA^T\) remains lower triangular. Thus, \( AB \) is upper triangular.

**Corollary 2.2.23**

The product of two unit lower (upper) triangular matrices is a unit lower (upper) triangular matrix.

**Proof** Let \( A \) and \( B \) be unit lower triangular \( n \times n \) matrices. We know from Theorem 2.2.22 that \( C = AB \) is a lower triangular matrix. We must establish that \( c_{ii} = 1 \) for each \( i \). The elements on the main diagonal of \( C \) can be obtained by setting \( j = i \) in (2.2.7):

\[
\begin{align*}
c_{ii} &= \sum_{k=1}^{n} a_{ik}b_{ki}.
\end{align*}
\]

Since \( a_{ik} = 0 \) whenever \( k > i \), the only nonzero term in the summation in (2.2.8) occurs when \( k = i \). Consequently,

\[
\begin{align*}
c_{ii} &= a_{ii}b_{ii} = 1 \cdot 1 = 1, \quad i = 1, 2, \ldots, n.
\end{align*}
\]

The proof for unit upper triangular matrices is similar and left as an exercise.

**The Algebra and Calculus of Matrix Functions**

By and large, the algebra of matrix and vector functions is the same as that for matrices and vectors of real or complex numbers. Since vector functions are a special case of matrix functions, we focus here on matrix functions. The main comment here pertains to scalar multiplication. In the description of scalar multiplication of matrices of numbers, the scalars were required to be real or complex numbers. However, for matrix functions, we can scalar multiply by any scalar function \( s(t) \).

**Example 2.2.24**

If \( s(t) = e^t \) and \( A(t) = \begin{bmatrix} -2 + t & e^{2t} \\ -4 & t \cos t \end{bmatrix} \), then

\[
\begin{align*}
s(t)A(t) &= \begin{bmatrix} e^{t(-2 + t)} & e^{3t} \\ -4e^t & e^{t \cos t} \end{bmatrix}.
\end{align*}
\]

**Example 2.2.25**

Referring to \( A \) and \( B \) from Example 2.1.14, find \( 2A - tB^T \).
2.2 Matrix Algebra

Solution: We have
\[ 2A - tB^T = \begin{bmatrix} 2t^3 & 2t - 2 \cos t & 10 \\ 2e^{t^2} & 2 \ln (t + 1) & 2e^{t^2} \end{bmatrix} - \begin{bmatrix} t^3 + t^2 - t - \ln(t) \\ 3t - 2 \cos t \\ 2e^{t^2} - 2 \ln t + 10 \end{bmatrix} \]

\[ = \begin{bmatrix} 2t^3 + t^2 - 2t - 2 \cos t \\ 3t - 2 \cos t \\ 2e^{t^2} - 2 \ln t + 10 \end{bmatrix}. \]

We can also perform calculus operations on matrix functions. In particular we can differentiate and integrate them. The rules for doing so are as follows:

1. The derivative of a matrix function is obtained by differentiating every element of the matrix. Thus, if \( A(t) = [a_{ij}(t)] \), then
\[
\frac{dA}{dt} = \frac{d[a_{ij}(t)]}{dt},
\]
provided that each of the \( a_{ij} \) is differentiable.

2. It follows from (1) and the index form of the matrix product that if \( A \) and \( B \) are both differentiable and the product \( AB \) is defined, then
\[
\frac{d}{dt}(AB) = A \frac{dB}{dt} + \frac{dA}{dt} B.
\]
The key point to notice is that the order of the multiplication must be preserved.

3. If \( A(t) = [a_{ij}(t)] \), where each \( a_{ij}(t) \) is integrable on an interval \([a, b]\), then
\[
\int_a^b A(t) \, dt = \left[ \int_a^b a_{ij}(t) \, dt \right].
\]

Example 2.2.26

If \( A(t) = \begin{bmatrix} 2t & 1 \\ 6e^{2t} & 4e^{2t} \end{bmatrix} \), determine \( dA/dt \) and \( \int_0^1 A(t) \, dt \).

Solution: We have
\[
\frac{dA}{dt} = \begin{bmatrix} 2 & 0 \\ 12te^{2t} & 8e^{2t} \end{bmatrix},
\]
whereas
\[
\int_0^1 A(t) \, dt = \begin{bmatrix} \int_0^1 2t \, dt & \int_0^1 1 \, dt \\ \int_0^1 6e^{2t} \, dt & \int_0^1 4e^{2t} \, dt \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 (e^2 - 1) & 2 (e^2 - 1) \end{bmatrix}.
\]

Exercises for 2.2

Key Terms
Matrix addition and subtraction, Scalar multiplication, Matrix multiplication, Dot product, Linear combination of column vectors, Index form, Premultiplication, Postmultiplication, Zero matrix, Identity matrix, Kronecker delta symbol.

Skills
- Be able to perform matrix addition, subtraction, and multiplication.
- Know the basic relationships between the dimensions of two matrices \( A \) and \( B \) in order for \( A + B \) to be defined, and in order for \( AB \) to be defined.
- Be able to multiply a matrix by a scalar.
- Be able to express the product \( Ax \) of a matrix and a column vector as a linear combination of the columns of \( A \).
True-False Review

For Questions 1–12, decide if the given statement is true or false, and give a brief justification for your answer. If true, you can quote a relevant definition or theorem from the text. If false, provide an example, illustration, or brief explanation of why the statement is false.

1. For all matrices \( A, B, \) and \( C \) of the appropriate dimensions, we have
\[
(AB)C = (CA)B.
\]

2. If \( A \) is an \( m \times n \) matrix, \( B \) is an \( n \times p \) matrix, and \( C \) is a \( p \times q \) matrix, then \( ABC \) is an \( m \times q \) matrix.

3. If \( A \) and \( B \) are symmetric \( n \times n \) matrices, then so is \( A + B \).

4. If \( A \) and \( B \) are skew-symmetric \( n \times n \) matrices, then \( AB \) is a symmetric matrix.

5. For \( n \times n \) matrices \( A \) and \( B \), we have
\[
(A + B)^2 = A^2 + 2AB + B^2.
\]

6. If \( AB = 0 \), then either \( A = 0 \) or \( B = 0 \).

7. If \( A \) and \( B \) are square matrices such that \( AB \) is upper triangular, then \( A \) and \( B \) must both be upper triangular.

8. If \( A \) is a square matrix such that \( A^2 = A \), then \( A \) must be the zero matrix or the identity matrix.

9. If \( A \) is a matrix of numbers, then if we consider \( A \) as a matrix function, its derivative is the zero matrix.

10. If \( A \) and \( B \) are matrix functions whose product \( AB \) is defined, then
\[
\frac{d}{dt}(AB) = A\frac{dB}{dt} + B\frac{dA}{dt}.
\]

11. If \( A \) is an \( n \times n \) matrix function such that \( A \) and \( dA/dt \) are the same function, then \( A = ce^{tI} \) for some constant \( c \).

12. If \( A \) and \( B \) are matrix functions whose product \( AB \) is defined, then the matrix functions \((AB)^T\) and \(B^T A^T \) are the same.

Problems

1. If
\[
A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & 5 \end{bmatrix}
\]
find \( 2A, -3B, A - 2B, \) and \( 3A + 4B \).

2. If
\[
A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 1 & 2 \\ -1 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 2 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}
\]
find the matrix \( D \) such that \( 2A + B - 3C + 2D = A + 4C \).

3. Let
\[
A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 & 3 \\ 4 & 6 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & -2 & 3 \end{bmatrix}
\]
Find, if possible, \( AB, BC, CA, DC, DB, AD, \) and \( CD \).

For Problems 4–6, determine \( AB \) for the given matrices. In these problems \( i \) denotes \( \sqrt{-1} \).

4. \( A = \begin{bmatrix} 2 - i & 1 + i \\ -i & 2 + 4i \end{bmatrix}, \quad B = \begin{bmatrix} 1 - 3i \\ 0 & 4 + i \end{bmatrix} \)

5. \( A = \begin{bmatrix} 3 + 2i & 2 - 4i \\ 5 + i & -1 + 3i \end{bmatrix}, \quad B = \begin{bmatrix} -1 + i & 3 + 2i \\ 4 - 3i & 1 + i \end{bmatrix} \)

6. \( A = \begin{bmatrix} 3 - 2i & i \\ -i & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 + i & 2 - i \\ 1 + 5i & 0 & 3 - 2i \end{bmatrix} \)
7. Let
\[ A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 3 & 4 \\ 1 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 2 \\ 1 & 5 \\ -1 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} -3 & 2 \\ 1 & -4 \end{bmatrix}. \]
Find \( ABC \) and \( CAB \).

8. If
\[ A = \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 \\ 5 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 3 \\ -1 \end{bmatrix}. \]
find \( (2A - 3B)C \).

For Problems 9–11, determine \( A \) by computing an appropriate linear combination of the column vectors of \( A \).

9. \[ A = \begin{bmatrix} 1 \\ 3 \\ -5 \\ 4 \end{bmatrix}, \quad e = \begin{bmatrix} 6 \\ -2 \end{bmatrix}. \]

10. \[ A = \begin{bmatrix} 3 & -1 & 4 \\ 2 & 1 & 5 \\ 7 & -6 & 3 \end{bmatrix}, \quad e = \begin{bmatrix} 2 \\ -4 \end{bmatrix}. \]

11. \[ A = \begin{bmatrix} -1 & 2 \\ 4 & 7 \\ 5 & -4 \end{bmatrix}, \quad e = \begin{bmatrix} 5 \\ -1 \end{bmatrix}. \]

12. If \( A \) is an \( m \times n \) matrix and \( C \) is an \( r \times s \) matrix, what must be the dimensions of \( B \) in order for the product \( ABC \) to be defined? Write an expression for the \((i, j)\)th element of \( ABC \) in terms of the elements of \( A, B \), and \( C \).

13. Find \( A^2, A^3 \), and \( A^4 \) if
(a) \[ A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}. \]
(b) \[ A = \begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix}. \]

14. If \( A \) and \( B \) are \( n \times n \) matrices, prove that
(a) \[ (A + B)^2 = A^2 + AB + BA + B^2. \]
(b) \[ (A - B)^2 = A^2 - AB - BA + B^2. \]

15. If \[ A = \begin{bmatrix} 3 & -1 \\ 4 & 5 \end{bmatrix}, \]
calculate \( A^2 \) and verify that \( A \) satisfies \( A^2 = 2A - 8I_2 \).

16. Find a matrix \[ A = \begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \]
such that
\[ A^2 + \begin{bmatrix} 0 & -1 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} = I_3. \]

17. If
\[ A = \begin{bmatrix} x & 1 \\ -2 & y \end{bmatrix}. \]
determine all values of \( x \) and \( y \) for which \( A^2 = A \).

18. The Pauli spin matrices \( \sigma_1, \sigma_2, \) and \( \sigma_3 \) are defined by
\[ \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \]
and
\[ \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \]
Verify that they satisfy
\[ \sigma_i \sigma_j = i \epsilon_{ijk} \sigma_k, \quad \sigma_i \sigma_j = i \epsilon_{ijk} \sigma_k, \quad \sigma_3 \sigma_1 = i \sigma_2. \]

If \( A \) and \( B \) are \( n \times n \) matrices, we define their commutator, denoted \( [A, B] \), by
\[ [A, B] = AB - BA. \]

Thus, \( [A, B] = 0 \) if and only if \( A \) and \( B \) commute. That is, \( AB = BA \). Problems 19–22 require the commutator.

19. If
\[ A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}. \]
find \( [A, B] \).

20. If
\[ A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \]
compute all of the commutators \( [A_1, A_2], [A_1, A_3], \) and determine which of the matrices commute.

21. If
\[ A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \]
verify that \( [A_1, A_2] = 0, [A_2, A_3] = A_1, [A_3, A_1] = A_2. \)
22. If \( A, B \) and \( C \) are \( n \times n \) matrices, find \([A, [B, C]]\) and prove the Jacobi identity
\[
\]
23. Use the index form of the matrix product to prove properties (2.2.4) and (2.2.5).
24. Prove parts 1 and 2 of Theorem 2.2.21.
25. Prove property 2 of the identity matrix.
26. If \( A \) and \( B \) are \( n \times n \) matrices, prove that \( \text{tr}(AB) = \text{tr}(BA) \).
27. If
\[
A = \begin{bmatrix} 1 & -1 & 1 & 4 \\ 2 & 0 & 2 & -3 \\ 3 & 4 & -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix},
\]
find \( A^T, B^T, AAT, AB \) and \( B^TA^T \).
28. Let \( A = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 2 \end{bmatrix} \) and let \( S \) be the matrix with column vectors
\[
s_1 = \begin{bmatrix} -x \\ x \end{bmatrix}, \quad s_2 = \begin{bmatrix} -y \\ y \end{bmatrix}, \quad s_3 = \begin{bmatrix} -z \\ z \end{bmatrix},
\]
where \( x, y, z \) are constants.
(a) Show that \( AS = [s_1, s_2, s_3] \).
(b) Find all values of \( x, y, z \) such that \( S^TAS = \text{diag}(1, 1, 7) \).
29. A matrix that is a multiple of \( I_n \) is called an \( n \times n \) scalar matrix.
(a) Determine the \( 4 \times 4 \) scalar matrix whose trace is 8.
(b) Determine the \( 3 \times 3 \) scalar matrix such that the product of the elements on the main diagonal is 243.
30. Prove that for each positive integer \( n \), there is a unique scalar matrix whose trace is a given constant \( k \).

If \( A \) is an \( n \times n \) matrix, then the matrices \( S \) and \( T \) defined by
\[
S = \frac{1}{2}(A + A^T), \quad T = \frac{1}{2}(A - A^T)
\]
are referred to as the symmetric and skew-symmetric parts of \( A \), respectively. Problems 31–34 investigate properties of \( S \) and \( T \).
31. Use the properties of the transpose to show that \( S \) and \( T \) are symmetric and skew-symmetric, respectively.
32. Find \( S \) and \( T \) for the matrix
\[
A = \begin{bmatrix} 1 & -5 & 3 \\ 3 & 2 & 4 \\ 7 & -2 & 6 \end{bmatrix}.
\]
33. If \( A \) is an \( n \times n \) symmetric matrix, prove that \( T = 0 \). What is the corresponding result for skew-symmetric matrices?
34. Show that every \( n \times n \) matrix can be written as the sum of a symmetric and a skew-symmetric matrix.
35. Prove that if \( A \) is an \( n \times p \) matrix and \( D = \text{diag}(d_1, d_2, \ldots, d_n) \), then \( DA \) is the matrix obtained by multiplying the \( i \)th row vector of \( A \) by \( d_i \) (\( 1 \leq i \leq n \)).
36. Use the properties of the transpose to prove that
(a) \( AA^T \) is a symmetric matrix.
(b) \( (ABC)^T = C^T B^T A^T \).

For Problems 37–40, determine the derivative of the given matrix function.

37. \( A(t) = e^{-t} \sin t \)
38. \( A(t) = \begin{bmatrix} t \sin t \\ \cos t \end{bmatrix} \)
39. \( A(t) = \begin{bmatrix} e^t & e^{2t} \\ 2e^t & 4e^t \end{bmatrix} \)
40. \( A(t) = \begin{bmatrix} \sin t \cos t \\ -\cos t \sin t \end{bmatrix} \)

41. Let \( A = [a_{ij}(t)] \) be an \( m \times n \) matrix function and let \( B = [b_{ij}(t)] \) be an \( n \times p \) matrix function. Use the definition of matrix multiplication to prove that
\[
\frac{d}{dt}(AB) = A \frac{dB}{dt} + B \frac{dA}{dt}.
\]
42. For Problems 42–45, determine \( \int_0^1 A(t) \, dt \) for the given matrix function.

43. \( A(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix} \), \( a = 0, b = \pi/2 \).
2.3 Terminology for Systems of Linear Equations

As we mentioned in Section 2.1, a main aim of this chapter is to apply matrices to determine the solution properties of any system of linear equations. We are now in a position to pursue that aim. We begin by introducing some notation and terminology.

DEFINITION 2.3.1

The general \( m \times n \) system of linear equations is of the form

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\
    &\vdots \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m,
\end{align*}
\]

(2.3.1)

where the system coefficients \( a_{ij} \) and the system constants \( b_i \) are given scalars and \( x_1, x_2, \ldots, x_n \) denote the unknowns in the system. If \( b_i = 0 \) for all \( i \), then the system is called homogeneous; otherwise it is called nonhomogeneous.

DEFINITION 2.3.2

By a solution to the system (2.3.1) we mean an ordered \( n \)-tuple of scalars, \( (c_1, c_2, \ldots, c_n) \), which, when substituted for \( x_1, x_2, \ldots, x_n \) into the left-hand side of system (2.3.1), yield the values on the right-hand side. The set of all solutions to system (2.3.1) is called the solution set to the system.

Remarks

1. Usually the \( a_{ij} \) and \( b_i \) will be real numbers, and we will then be interested in determining only the real solutions to system (2.3.1). However, many of the problems that arise in the later chapters will require the solution to systems with complex coefficients, in which case the corresponding solutions will also be complex.

2. If \( (c_1, c_2, \ldots, c_n) \) is a solution to the system (2.3.1), we will sometimes specify this solution by writing \( x_1 = c_1, x_2 = c_2, \ldots, x_n = c_n \). For example, the ordered pair of numbers \((1, 2)\) is a solution to the system

\[
\begin{align*}
    x_1 + x_2 &= 3, \\
    3x_1 - 2x_2 &= -1,
\end{align*}
\]

and we could express this solution in the equivalent form \( x_1 = 1, x_2 = 2 \).