Chapter 2: Matrices

We want to solve

$$\begin{cases} 3x & + & 6y & & & = & 1, \\ & & y & + & z & = & 2, \\ -x & & & + & z & = & 3. \end{cases} \tag{$\heartsuit$}$$

To solve ($\heartsuit$), the approaches shown in Chapter 1 only involve the coefficients and the constant terms of the linear system. So we arrange those coefficients and constants as the following rectangular arrays (called matrices):

$$A = \begin{bmatrix} 3 & 6 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Also we form the unknowns as the array $\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. We want to represent the system as $A\boldsymbol{x} = \boldsymbol{b}$. But, how do we define the product of two matrices A and $\boldsymbol{x}$? What is the definition of two equal matrices?

2.1 Matrices

Definition 2.1.1: A *matrix* over $\mathbb{R}$ is a rectangular display of scalars (real numbers). A matrix with m rows and n columns is called an $m \times n$ *matrix* or *matrix of size* $m \times n$. If $m = n$, then the matrix is called a *square matrix of order* n (or *size* n). We use the notation

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

to describe an $m \times n$ matrix, where $a_{ij} \in \mathbb{R}$ for $1 \le i \le m$ and $1 \le j \le n$. For short, we use $A = [a_{ij}] = (a_{ij})$. This notation indicates that A is the matrix whose general (i, j)*-th entry* is a_{ij}. To avoid some confusion we shall use the notation $[A]_{i,j}$ to denote the (i, j)-th entry of A.

Remark 2.1.2:

1. Some textbooks use large parentheses instead of brackets– the distinction is not important. In this course, we shall adopt both.

2. Rows of a matrix will be referenced starting at the top and working down (i.e., row 1 is at the top) and columns will be referenced starting from the left (i.e., column 1 is at the left).

Example 2.1.1:

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

is a matrix with $m = 3$ rows and $n = 4$ columns, i.e., B is a 3×4 matrix. We can say that $[B]_{2,3} = -6$ while $[B]_{3,4} = -2$. ■

Definition 2.1.3:

1. A *column vector* of *size* or *length* n is an ordered list of n numbers, which is written in order vertically, starting at the top and proceeding to the bottom. At times, we will refer to a column vector as simply a *vector*.

2. Column vectors will be written in bold, usually with lower case Latin letter from the end of the alphabet such as $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$, $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$.

3. Some books like to write vectors with arrows, such as $\vec{u}$. Writing by hand, some like to put arrows on top of the symbol (I shall use this notation written on the white board), or a tilde underneath the symbol, as in $\underset{\sim}{u}$, or a line under the symbol, as $\underline{u}$.

4. To refer to the *entry* or *component* of vector $\boldsymbol{v}$ in location i of the list, we write $[\boldsymbol{v}]_i$.

2.2 Partition of Matrices

Sometimes we put horizontal lines or vertical lines to divide the matrix into different areas. It is same as the matrix without the lines.

Example 2.2.1: The matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 3.5 \\ 0 & -1 & 1 & 1.1 & 1 \\ 3 & 5.8 & 1 & 0 & -3 \\ 1 & 8 & 0 & 0 & 7 \end{bmatrix}$$

is same as the following matrices:

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 3.5 \\ 0 & -1 & 1 & 1.1 & 1 \\ 3 & 5.8 & 1 & 0 & -3 \\ 1 & 8 & 0 & 0 & 7 \end{array}\right], \quad \left[\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 3.5 \\ 0 & -1 & 1 & 1.1 & 1 \\ 3 & 5.8 & 1 & 0 & -3 \\ 1 & 8 & 0 & 0 & 7 \end{array}\right], \quad \left[\begin{array}{c|c|c|c|c} 1 & 2 & 3 & 4 & 3.5 \\ 0 & -1 & 1 & 1.1 & 1 \\ 3 & 5.8 & 1 & 0 & -3 \\ 1 & 8 & 0 & 0 & 7 \end{array}\right],$$

$$\left[\begin{array}{ccccc} 1 & 2 & 3 & 4 & 3.5 \\ 0 & -1 & 1 & 1.1 & 1 \\ \hline 3 & 5.8 & 1 & 0 & -3 \\ 1 & 8 & 0 & 0 & 7 \end{array}\right], \quad \left[\begin{array}{cc|ccc} 1 & 2 & 3 & 4 & 3.5 \\ 0 & -1 & 1 & 1.1 & 1 \\ \hline 3 & 5.8 & 1 & 0 & -3 \\ 1 & 8 & 0 & 0 & 7 \end{array}\right], \quad \left[\begin{array}{cc|ccc} 1 & 2 & 3 & 4 & 3.5 \\ 0 & -1 & 1 & 1.1 & 1 \\ \hline 3 & 5.8 & 1 & 0 & -3 \\ \hline 1 & 8 & 0 & 0 & 7 \end{array}\right].$$

■

Example 2.2.2:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}, \quad \boldsymbol{u} = \begin{bmatrix} 9 \\ 10 \\ 11 \\ 12 \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} 13 \\ 14 \\ 15 \\ 16 \end{bmatrix}.$$

$$[A|\boldsymbol{u}] = \left[\begin{array}{cc|c} 1 & 2 & 9 \\ 3 & 4 & 10 \\ 5 & 6 & 11 \\ 7 & 8 & 12 \end{array}\right] = \begin{bmatrix} 1 & 2 & 9 \\ 3 & 4 & 10 \\ 5 & 6 & 11 \\ 7 & 8 & 12 \end{bmatrix}, \quad [A|\boldsymbol{u}|\boldsymbol{v}] = \left[\begin{array}{cc|c|c} 1 & 2 & 9 & 13 \\ 3 & 4 & 10 & 14 \\ 5 & 6 & 11 & 15 \\ 7 & 8 & 12 & 15 \end{array}\right] = \begin{bmatrix} 1 & 2 & 9 & 13 \\ 3 & 4 & 10 & 14 \\ 5 & 6 & 11 & 15 \\ 7 & 8 & 12 & 15 \end{bmatrix}.$$

■

Example 2.2.3:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix},$$

$$C = \begin{bmatrix} 11 & 12 \\ 13 & 14 \\ 15 & 16 \end{bmatrix}, \quad D = \begin{bmatrix} 21 & 22 & 23 \\ 24 & 25 & 26 \\ 27 & 28 & 29 \end{bmatrix}.$$

$$[A|B] = \left[\begin{array}{cc|ccc} 1 & 2 & 5 & 6 & 7 \\ 3 & 4 & 8 & 9 & 10 \end{array}\right] = \left[\begin{array}{ccccc} 1 & 2 & 5 & 6 & 7 \\ 3 & 4 & 8 & 9 & 10 \end{array}\right],$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|ccc} 1 & 2 & 5 & 6 & 7 \\ 3 & 4 & 8 & 9 & 10 \\ \hline 11 & 12 & 21 & 22 & 23 \\ 13 & 14 & 24 & 25 & 26 \\ 15 & 16 & 27 & 28 & 29 \end{array}\right] = \left[\begin{array}{ccccc} 1 & 2 & 5 & 6 & 7 \\ 3 & 4 & 8 & 9 & 10 \\ 11 & 12 & 21 & 22 & 23 \\ 13 & 14 & 24 & 25 & 26 \\ 15 & 16 & 27 & 28 & 29 \end{array}\right].$$

■

Definition 2.2.1: Suppose $A = (a_{ij})$ is an $m \times n$ matrix. For $1 \leq i \leq m$, the *i-th row* of A is the $1 \times n$ matrix $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$ (or sometimes is viewed as a *row vector* $(a_{i1}, \ldots, a_{in})$) which is usually denoted by A_{i*}.

For $1 \leq j \leq n$, the *j-th column* of A is the $m \times 1$ matrix $\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$ and is usually denoted by A_{*j}.

So A can be represented partition matrices as $\begin{bmatrix} A_{1*} \\ A_{2*} \\ \vdots \\ A_{m*} \end{bmatrix}$ or $\begin{bmatrix} A_{*1} & A_{*2} & \cdots & A_{*n} \end{bmatrix}$. Note that the horizontal or vertical lines are omitted.

2.3 Matrix Representations of Linear Systems

In general, we will consider the problem of solving n unknowns $x_1, x_2, \ldots, x_n$ which satisfy the following m equations simultaneously:

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \right., \tag{2.1}$$

where $b_1, b_2, \ldots, b_m$ and a_{ij} $(1 \leq i \leq m,\ 1 \leq j \leq n)$ are given constants. For avoiding the confusion, sometimes we write $a_{i,j}$ to instead of a_{ij}. Let

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \tag{2.2}$$

A is called the *coefficient matrix* (or *matrix of coefficients*) of the system (2.1) and $\boldsymbol{b}$ is called the *vector of constants*. We will write $\mathcal{LS}(A,\,\boldsymbol{b})$ as a shorthand expression for the system of linear equations (2.1), which we will refer to as the *matrix representation* of the linear system.

A solution $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is called a *solution vector*. But for saving space, we sometimes written a vector as row form or an n-turple, for example, $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ or $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

$$[A|\boldsymbol{b}] = \left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right]$$ is called an *augmented matrix* of the system.

2.4 Algebra of Matrices

We denote $M_{m,n}(\mathbb{R})$ (or $M_{m,n}$) to be the set of all $m \times n$ matrices over $\mathbb{R}$. If $m = n$, then we sometimes use $M_n(\mathbb{R})$ to instead of $M_{n,n}(\mathbb{R})$. We also denote $M_{1,n}(\mathbb{R})$ as $\mathbb{R}^n$. But, mention again, we sometimes write element of $\mathbb{R}^n$ as row vector form for saving space.

Definition 2.4.1: Two matrices A and B are said to be *equal*, which is denoted by $A = B$, if they are both of the same size and $[A]_{i,j} = [B]_{i,j} \quad \forall i, j$.

The symbol $\forall$ means 'for every' or 'for each', but is read as 'for all'.

Definition 2.4.2: An $m \times n$ *zero matrix*, denoted by $\mathcal{O}$ (or $\mathcal{O}_{m,n}$, or $\mathcal{O}_{m\times n}$), is a matrix whose entries are all zero. If $m = n$, then we use $\mathcal{O}_n$ to instead of $\mathcal{O}_{n,n}$. If $A = (a_{ij}) \in M_n(\mathbb{R})$, then the sequence of entries $\{a_{11}, a_{22}, \ldots, a_{nn}\}$, is called the *diagonal* of A. A square matrix with zero entries everywhere except in the diagonal is called a *diagonal matrix*. The *identity matrix of order* n, denoted by I or I_n, is a diagonal matrix of order n with all entries in the diagonal are equal to 1.

For integers i, j, we define a notation δ_{ij} by

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

This is called the *Kronecker delta*. Then $[I_n]_{i,j} = \delta_{ij}$ for $1 \le i, j \le n$. Clearly, $\delta_{ij} = \delta_{ji}$.

Definition 2.4.3: Let U and L be $n \times n$ matrices. U is said to be *upper triangular* if $[U]_{i,j} = 0 \; \forall \; i > j$ and L is said to be *lower triangular* if $[L]_{i,j} = 0 \; \forall \; i < j$.

Therefore, a diagonal matrix is both upper and lower triangular matrix.

The *zero vector* $\mathbf{0} = \mathbf{0}_n$ of *size* (or *length*) n is a column vector of size n whose entries are 0, i.e.,

$\mathbf{0}_n = \mathcal{O}_{n,1}$. The *standard unit vectors* of length n are

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

That is, $[\boldsymbol{e}_i]_j = \delta_{i,j}$, $1 \le i, j \le n$.

Definition 2.4.4: Let $A, B \in M_{m,n}$. The *sum* of A and that of B, denoted by $A+B$, is an $m \times n$ matrix whose (i,j)-th entry is the sum of the (i,j)-th entries of A and B, i.e., $(A+B)_{i,j} = (A)_{i,j} + (B)_{i,j}\ \forall\, 1 \le i \le m$, $1 \le j \le n$. The operation "+" is called the *addition* (of matrices).

Example 2.4.1: If $A = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & -7 \end{bmatrix}$ and $B = \begin{bmatrix} 6 & 2 & -4 \\ 3 & 5 & 2 \end{bmatrix}$, then

$$A + B = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & -7 \end{bmatrix} + \begin{bmatrix} 6 & 2 & -4 \\ 3 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 2+6 & -3+2 & 4+(-4) \\ 1+3 & 0+5 & -7+2 \end{bmatrix} = \begin{bmatrix} 8 & -1 & 0 \\ 4 & 5 & -5 \end{bmatrix}. \qquad \blacksquare$$

Proposition 2.4.5: *Let $A, B, C \in M_{m,n}$. Then we have*

(1) $A + B = B + A$. — Commutativity of addition

(2) $A + (B + C) = (A + B) + C$. — Associativity of addition

(3) $A + \mathcal{O} = A$. — Identity of addition

(4) there is a unique matrix A' such that $A + A' = \mathcal{O}$. — Inverse of addition

Since the additive inverse of A is unique, we use $-A$ to denote it.

Definition 2.4.6: Let $A \in M_{m,n}$, $c \in \mathbb{R}$. Define $cA \in M_{m,n}$ by $(cA)_{i,j} = c(A)_{i,j}, 1 \le i \le m$, $1 \le j \le n$. This multiplication is called *scalar multiplication* and cA is called the *scalar product* of A by c.

Example 2.4.2: If $A = \begin{bmatrix} 2 & 8 \\ -3 & 5 \\ 0 & 1 \end{bmatrix}$ and $c = 7$, then

$$cA = 7 \begin{bmatrix} 2 & 8 \\ -3 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7(2) & 7(8) \\ 7(-3) & 7(5) \\ 7(0) & 7(1) \end{bmatrix} = \begin{bmatrix} 14 & 56 \\ -21 & 35 \\ 0 & 7 \end{bmatrix}.$$

■

Proposition 2.4.7: *Let $A, B \in M_{m,n}$, $c, d \in \mathbb{R}$. Then we have*

(1) $c(A + B) = cA + cB$. Left distributive law for scalar multiplication

(2) $(c + d)A = cA + dA$. Right distributive law for scalar multiplication

(3) $(cd)A = c(dA)$. Associativity of scalar multiplication

(4) $1A = A$ *and* $(-1)A = -A$.

(5) Suppose $A \neq \mathcal{O}$ *and* $cA = \mathcal{O}$. *Then* $c = 0$.

Back to see the system (♡), we want to write the system as $A\boldsymbol{x} = \boldsymbol{b}$. So we have the following definition.

Definition 2.4.8: Suppose A is an $m \times n$ matrix with columns $A_{*1}, \dots, A_{*n}$ and $\boldsymbol{u}$ is a vector of size n. Then the *matrix-vector product* of A with $\boldsymbol{u}$ is the linear combination

$$A\boldsymbol{u} = [\boldsymbol{u}]_1 A_{*1} + [\boldsymbol{u}]_2 A_{*2} + \cdots + [\boldsymbol{u}]_n A_{*n} = \sum_{i=1}^{n} [\boldsymbol{u}]_i A_{*i}. \tag{2.3}$$

So, the matrix-vector product is yet another version of *multiplication,* at least in the sense that we have yet again overloaded juxtaposition of two symbols as our notation. Note that, an $m \times n$ matrix times a vector of size n will create a vector of size m. So if A is rectangular, then the size of the vector changes.

Let us write down (2.3) more precisely. Let $A = (a_{i,j})$ and $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$.

$$A\boldsymbol{u} = \begin{bmatrix} u_1a_{1,1} + u_2a_{1,2} + \cdots + u_na_{1,n} \\ u_1a_{2,1} + u_2a_{2,2} + \cdots + u_na_{2,n} \\ \vdots \\ u_1a_{m,1} + u_2a_{m,2} + \cdots + u_na_{m,n} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n u_ia_{1,i} \\ \sum_{i=1}^n u_ia_{2,i} \\ \vdots \\ \sum_{i=1}^n u_ia_{m,i} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1,i}u_i \\ \sum_{i=1}^n a_{2,i}u_i \\ \vdots \\ \sum_{i=1}^n a_{m,i}u_i \end{bmatrix}.$$

Now, system ($\heartsuit$) can be written as

$$A\boldsymbol{x} = \begin{bmatrix} 3 & 6 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \boldsymbol{b}.$$

Example 2.4.3: Consider

$$A = \begin{bmatrix} 1 & 4 & 2 & 3 & 4 \\ -3 & 2 & 0 & 1 & -2 \\ 1 & 6 & -3 & -1 & 5 \end{bmatrix}, \quad \boldsymbol{u} = \begin{bmatrix} 2 \\ 1 \\ -2 \\ 3 \\ -1 \end{bmatrix}.$$

Then

$$A\boldsymbol{u} =$$

Proposition 2.4.9: *The set of solutions to the linear system* (2.1) *equals the set of solutions for* $\boldsymbol{x}$ *in the vector equation* $A\boldsymbol{x} = \boldsymbol{b}$*, where* A *and* $\boldsymbol{b}$ *are defined in* (2.2) *and* $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

Theorem 2.4.10: *Suppose that* A *and* B *are* $m \times n$ *matrices such that* $A\boldsymbol{x} = B\boldsymbol{x}$ *for every* $\boldsymbol{x} \in \mathbb{R}^n$*. Then* $A = B$.

Proof: Since $A\boldsymbol{x} = B\boldsymbol{x}\ \forall \boldsymbol{x} \in \mathbb{R}^n$,

Definition 2.4.11: Let $A \in M_{m,n}$ and $B \in M_{n,p}$. We define the *product* $AB \in M_{m,p}$ by

$$AB = A\left[\, B_{*1} \mid B_{*2} \mid \cdots \mid B_{*p} \,\right] = \left[\, AB_{*1} \mid AB_{*2} \mid \cdots \mid AB_{*p} \,\right].$$

How to memorize the formula: To find the (i, j)-th entry of AB.

(1) Find the i-th row of A (simply called the row below).

(2) Find the j-th column of B (simply called the column below).

(3) sum up the product the corresponding entries of the row and the column, i.e., (entry 1 of the row $\times$ entry 1 of the column) + (entry 2 of the row $\times$ entry 2 of the column) + $\cdots$

Example 2.4.4: Suppose

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

Find the $(3, 2)$-entry of AB and also find AB.

The 3-rd row of A is $\begin{bmatrix} -5 & 1 & 2 & -3 & 4 \end{bmatrix}$

The 2-nd column of B is $\begin{bmatrix} 6 \\ 4 \\ 1 \\ 4 \\ -2 \end{bmatrix}$.

Let us do the multiplication:

row	-5	1	2	-3	4
column	6	4	1	4	-2
product	-30	4	2	-12	-8

The sum is

$$-30 + 4 + 2 - 12 - 8 = -44.$$

Remark 2.4.12: Note that B and A must be of the proper size in order that BA is defined. Even if AB is defined, AB may not equal to BA. For example, $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ then $AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Thus matrix multiplication is *not commutative*. Also note that the product of two nonzero matrices may be a *zero* matrix as the above example shown.

Proposition 2.4.13: *For A, B and C are matrices (when the statement includes the matrix multiplication, the sizes of A, B, C, the identity matrix I and zero matrix $\mathcal{O}$ are chosen suitably), $c \in \mathbb{R}$, we have*

(1) $(cA)B = A(cB) = c(AB)$. Scalar pull through

(2) $(AB)C = A(BC)$. Associativity of multiplication

(3) $AI = A$, $IB = B$. Identity for multiplication

(4) $A(B + C) = AB + AC$. Left distributive law

(5) $(A + B)C = AC + BC$. Right distributive law

(6) $A\mathcal{O} = \mathcal{O}$ *and* $\mathcal{O}B = \mathcal{O}$. Zero matrix for multiplication

Proof:

It is because that the associative law holds on addition, scalar multiplication and multiplication, we usually omit to write the parentheses "( )".

Let A be a square matrix and n a positive integer. A^n denotes the product of n A's, i.e., $A^n = \overbrace{AA\cdots A}^{n \text{ times}}$. By convention, we let $A^0 = I$.

Definition 2.4.14: The *transpose* A^t of a matrix $A = [a_{ij}] \in M_{m,n}$ is the matrix in $M_{n,m}$ that whose (i,j)-th entry is a_{ji}. That is,

$$[A^t]_{i,j} = [A]_{j,i} \;\forall\; i = 1,\ldots,n,\; j = 1,\ldots,m.$$

Proposition 2.4.15: *Let A and B be matrices (when the statement includes the matrix multiplication, the sizes of A and B are chosen suitably). Then*

(1) $(A^t)^t = A$. Transpose of the transpose

(2) $(A+B)^t = A^t + B^t$. Transpose of a sum

(3) $(AB)^t = B^t A^t$. Transpose of a product

(4) $(cA)^t = cA^t$ *for* $c \in \mathbb{R}$.

Proof:

Definition 2.4.16: A square matrix S is called *symmetric* if $S^t = S$, i.e., $[S]_{i,j} = [S]_{j,i} \;\forall\; i,j$. A square matrix A is called *skew-symmetric* or *anti-symmetric* if $A^t = -A$, i.e., $[A]_{i,j} = -[A]_{j,i} \;\forall\; i,j$.

Example 2.4.5: The matrix

$$A = \begin{bmatrix} 2 & 3 & -9 & 5 & 7 \\ 3 & 1 & 6 & -2 & -3 \\ -9 & 6 & 0 & -1 & 9 \\ 5 & -2 & -1 & 4 & -8 \\ 7 & -3 & 9 & -8 & -3 \end{bmatrix}$$

is symmetric.

The matrix

$$B = \begin{bmatrix} 0 & 3 & -9 & 5 & 7 \\ -3 & 0 & 6 & -2 & -3 \\ 9 & -6 & 0 & -1 & 9 \\ -5 & 2 & 1 & 0 & -8 \\ -7 & 3 & -9 & 8 & 0 \end{bmatrix}$$

is skew-symmetric. ■

Proposition 2.4.17: *Let $A \in M_n(\mathbb{R})$ be a skew-symmetric matrix. Then each entry in the diagonal of A is zero, i.e., $[A]_{i,i} = 0$ for each i.*

2.5 Block Multiplication

Let $A \in M_{m,n}$ and $B = (B_1 \mid B_2)$, where $B_1 \in M_{n,p_1}$ and $B_2 \in M_{n,p_2}$. By Definition 2.4.11, we have $AB = (AB_1 \mid AB_2)$. It can be generalized to block (matrix) multiplication.

Let $A \in M_{m,n}$. Suppose that there are two partitions of m and n. Namely, $m = m_1 + \cdots + m_r$ and $n = n_1 + \cdots + n_s$ for some positive integers m_i, n_j $(1 \le i \le r,\ 1 \le j \le s)$. Then A can be partitioned into rs submatrices as a *partitioned matrix*:

$$A = \left(\begin{array}{c|c|c} A^{1,1} & \cdots & A^{1,s} \\ \hline \vdots & \vdots & \vdots \\ \hline A^{r,1} & \cdots & A^{r,s} \end{array}\right),$$

where each $A^{i,j}$ is an $m_i \times n_j$ submatrices of A.

Suppose B is an $n \times p$ matrix of a partitioned matrix

$$B = \left(\begin{array}{c|c|c} B^{1,1} & \cdots & B^{1,t} \\ \hline \vdots & \vdots & \vdots \\ \hline B^{s,1} & \cdots & B^{s,t} \end{array}\right),$$

where $p = p_1 + \cdots + p_t$, $p_1, \ldots, p_t$ are positive integers and each $B^{j,k}$ is an $n_j \times p_k$ submatrices of B.

Let $C = AB$. Then C is an $m \times p$ matrix which may be partitioned in a partitioned matrix:

$$C = \left(\begin{array}{c|c|c} C^{1,1} & \cdots & C^{1,t} \\ \hline \vdots & \vdots & \vdots \\ \hline C^{r,1} & \cdots & C^{r,t} \end{array}\right),$$

where each $C^{i,k}$ is an $m_i \times p_k$ submatrices of C.

By a tedious but straightforward verification, one can show that $C^{i,k} = \sum\limits_{j=1}^{s} A^{i,j}B^{j,k}$ for each i, k.

Such multiplication is referred as *partitioned multiplication* or *block multiplication*. Thus when we multiply two partitioned matrices, we may regard blocks as entries and multiply in the usual way.

Example 2.5.1: Let

$$A = \left(\begin{array}{ccc|cc|cc} 1 & 1 & 0 & 1 & 2 & 2 & 1 \\ 0 & -1 & -1 & -1 & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ \hline 0 & 1 & 0 & 2 & 3 & 1 & 2 \end{array}\right)$$

$$B = \left(\begin{array}{c|cc|cccc} 1 & 2 & 3 & 1 & 1 & 0 & 1 \\ 0 & 1 & 2 & -3 & -1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & -3 & 0 & 0 & 0 \\ -2 & 3 & -1 & 0 & -1 & 1 & 2 \\ \hline 0 & 1 & 0 & 1 & 3 & 0 & 2 \\ 1 & 0 & 0 & 2 & 0 & -1 & 0 \end{array}\right).$$

Let $C = AB$. Then C can be written as a block form:

$$C = \left(\begin{array}{c|c|c} C^{1,1} & C^{1,2} & C^{1,3} \\ \hline C^{2,1} & C^{2,2} & C^{2,3} \\ \hline C^{3,1} & C^{3,2} & C^{3,3} \end{array}\right),$$

where $C^{1,1}$ is a 2×1 matrix, $C^{1,2}$ is a 2×2 matrix, $C^{1,3}$ is 2×4 matrix and so on.

We consider the matrix $C^{1,3}$. Since $C^{1,3} = A^{1,1}B^{1,3} + A^{1,2}B^{2,3} + A^{1,3}B^{3,3}$,

$$\begin{aligned} C^{1,3} &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 1 \\ -3 & -1 & 0 & -1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \\ &\quad + \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -3 & 0 & 0 & 0 \\ 0 & -1 & 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 & 2 \\ 2 & 0 & -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 0 & 0 & 0 \\ 3 & 0 & -1 & 0 \end{pmatrix} + \begin{pmatrix} -3 & -2 & 2 & 4 \\ 3 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 4 & 6 & -1 & 4 \\ 3 & 3 & -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 4 & 1 & 8 \\ 9 & 3 & -2 & 2 \end{pmatrix}. \end{aligned}$$

One can compute all the $C^{i,j}$'s and obtains

$$C = \left(\begin{array}{c|cc|cccc} 0 & 13 & 4 & -1 & 4 & 1 & 8 \\ -2 & -2 & -3 & 9 & 3 & -2 & 2 \\ \hline 5 & 5 & 4 & 1 & 5 & 0 & 4 \\ 0 & 3 & -1 & 2 & 0 & 1 & 3 \\ \hline 0 & 15 & 1 & -4 & -1 & 1 & 7 \end{array}\right).$$

Example 2.5.2: Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. By using the block multiplication we have

$$AB = \begin{pmatrix} A_{1*}B \\ A_{2*}B \\ \vdots \\ A_{m*}B \end{pmatrix}.$$

Example 2.5.3: Compute A^3, where $A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$.

$A = \left(\begin{array}{ccc|cc} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & -1 & 1 \end{array}\right)$. Let $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}$.

Then $A = \begin{pmatrix} D & O_{3,2} \\ O_{2,3} & B \end{pmatrix}$. And then $A^2 = \begin{pmatrix} D^2 & O_{3,2} \\ O_{2,3} & B^2 \end{pmatrix}$, $A^3 = \begin{pmatrix} D^3 & O_{3,2} \\ O_{2,3} & B^3 \end{pmatrix}$.

It is easy to see that $D^3 = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

We only need to compute B^3.

$B^2 = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 9 \\ -3 & -2 \end{pmatrix}$

$B^3 = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 9 \\ -3 & -2 \end{pmatrix} = \begin{pmatrix} -7 & 12 \\ -4 & -11 \end{pmatrix}$.

$D^3 = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $B^3 = \begin{pmatrix} -7 & 12 \\ -4 & -11 \end{pmatrix}$.

Then

$$A^3 = \left(\begin{array}{ccc|cc} 8 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & -7 & 12 \\ 0 & 0 & 0 & -4 & -11 \end{array}\right).$$