Matrix

Liming Pang

1 Matrix

One motivation of inventing and studying matrices is to solve a system of linear equations of the form:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ .... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{cases} \tag{1.1}$$

For example, in middle school we have learnt to solve the system of equations

$$\begin{cases} 3x + 2y = 6 \\ 5x - y = 8 \end{cases}$$

But more generally, when there are more variables and more equations, how do we find the solutions? For example, how to solve for the following system of equations?

$$\begin{cases} 2x + 5y - z + w = 11 \\ 3x - y + z + 2w = 7 \\ 9x - 4y + z + 2w = 3 \\ 12x + 10y + 8z + 5w = 9 \end{cases}$$

In order to study a given system of equations, we introduce a way to denote the coefficients of the equations, which is the definition of matrix. Algebraic operations defined on matrices will help us solve for system of linear equations.

Definition 1. A $m \times n$ **matrix** is a collection of $m \times n$ numbers put into a rectangular shape of m horizontal lines and n vertical lines. Each horizontal line of numbers is called a **row**, which are labelled by $1, 2, ..., m$ from top to bottom. Each vertical line is called a **column**, which are labelled by $1, 2, ..., n$ from left to right. Each of the mn numbers is called an **entry**. The entry at i-th row and j-th column is called the ij-entry, and is labelled by ij.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \tag{1.2}$$

We also write $A = (a_{ij})$ to denote the matrix whose ij-entry is a_{ij}.

Given a system of linear equations, we can arrange the coefficients into a matrix as shown above. For example, the coefficients in the Equations (1.1) can be denoted by the matrix (0.2)

Example 2. *The following are matrices:*

$$\begin{bmatrix} 12 & 3 & 15 \\ -5 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 3 & 4 & 12 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -5 \end{bmatrix}$$

Remark 3. A vector in a coordinate system can be regarded as a matrix of one row or a matrix of one column, depending on whether it is written horizontally or vertically.

Example 4. *$A = (a_{ij})$ is a 3×2 matrix such that $a_{ij} = i + j$. Write A out explicitly.*

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{bmatrix}$$

Definition 5. A matrix is called a square matrix if it has same number of rows and columns.

2 Matrix Operations

Similar to matrices, we can also define some algebraic operations on matrices.

Definition 6. Given two $m \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$, define $A + B$ to be the matrix whose ij-entry is $a_{ij} + b_{ij}$, i.e. $A + B = (a_{ij} + b_{ij})$; define $A - B$ to be the matrix whose ij-entry is $a_{ij} - b_{ij}$, i.e. $A + B = (a_{ij} - b_{ij})$

Note that by the definition, two matrices can be added or subtracted only when they have the same number of rows and same number of columns.

Example 7.

$$\begin{bmatrix} 1 & 2 & 4 \\ -3 & 5 & -9 \end{bmatrix} + \begin{bmatrix} -1 & 4 & -2 \\ 4 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 6 & 2 \\ 1 & 12 & -7 \end{bmatrix}$$

Example 8.

$$\begin{bmatrix} 1 & 2 & 4 \\ -3 & 5 & -9 \end{bmatrix} - \begin{bmatrix} -1 & 4 & -2 \\ 4 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 6 \\ -7 & -2 & -11 \end{bmatrix}$$

Definition 9. If λ is a number and $A = (a_{ij})$ is a matrix, define the scalar multiplication to be $\lambda A = (\lambda a_{ij})$, i.e., all the entries of A are multiplied by λ. In particular, we denote $-A = (-1)A$.

Example 10.

$$-3 \begin{bmatrix} 1 & 2 & 4 \\ -3 & 5 & -9 \end{bmatrix} = \begin{bmatrix} -3 & -6 & -12 \\ 9 & -15 & 27 \end{bmatrix}$$

Definition 11. The $m \times n$ zero matrix is the $m \times n$ matrix all of whose entries are 0. It is usually denoted by $\mathbf{0}$.

Proposition 12. *If A, B, C are $m \times n$ matrices and λ, μ are numbers, then:*

1. $(A + B) + C = A + (B + C)$
2. $A + B = B + A$
3. $A + \mathbf{0} = \mathbf{0}$
4. $A + (-A) = \mathbf{0}$
5. $(\lambda + \mu)A = \lambda A + \mu A$
6. $\lambda(A + B) = \lambda A + \lambda B$

Definition 13. If $A = (a_{ij})$ is a $m \times n$ matrix and $B = \{b_{ij}\}$ is a $n \times p$ matrix, then define the product of A and B tp be the $m \times p$ matrix AB whose ij-entry is

$$\sum_{k=1}^{n} a_{ik} b_{kj}$$

Example 14.

$$\begin{aligned}
&\begin{bmatrix} -2 & 3 & 4 \\ -3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -3 & 5 \\ 2 & -1 \end{bmatrix} \\
&= \begin{bmatrix} -2\times 1 + 3\times(-3) + 4\times 2 & -2\times 2 + 3\times 5 + 4\times(-1) \\ -3\times 1 + 1\times(-3) + 0\times 2 & -3\times 2 + 1\times 5 + 0\times(-1) \end{bmatrix} \\
&= \begin{bmatrix} -3 & 7 \\ -6 & -1 \end{bmatrix}
\end{aligned}$$

Definition 15. The $n \times n$ identity matrix is the matrix $I_n = (a_{ij})$ such that

$$a_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

i.e., the $n \times n$ matrix whose diagonal entries are 1 and off-diagonal entries are 0.

Example 16. *The* 3×3 *identity matrix is*

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Proposition 17. *If* A *is a* $m \times n$ *matrix,* B *is a* $n \times p$ *matrix, If* C *is a* $p \times q$ *matrix, and* λ *is a number, then:*

1. $(AB)C = A(BC)$
2. $(\lambda A)BC = A(\lambda B)C = AB(\lambda C) = \lambda(ABC)$
3. $A \times I_n = A$, $I_n \times B = B$
4. $\mathbf{0}A = A\mathbf{0} = \mathbf{0}$

Similar to real numbers, the multiplication of matrices has Distributive Law over matrix addition and subtraction:

Proposition 18. *If A is a $m \times n$ matrix, B and C are $n \times p$ matrices, and D is a $p \times q$ matrix, then:*

$$A(B+C) = AB + AB, (B+C)D = BD + CD$$

Remark 19. There are some big difference between multiplication of numbers and multiplication of matrices.

1. The multiplication of matrices does not satisfy the Commutative Law for Multiplication!

   For example, If we take $A = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$, then

$$AB = \begin{bmatrix} 1 & 0 \\ 9 & 4 \end{bmatrix} \neq \begin{bmatrix} 4 & 9 \\ 0 & 1 \end{bmatrix} = BA$$

2. $AB = \mathbf{0}$ does not necessarily imply either A or B is $\mathbf{0}$. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

3. $AC = BC$ and $C \neq \mathbf{0}$ does not necessarily imply $A = B$. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

   But

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

Definition 20. If n is a positive integer, the n-th power of a matrix A is defined to be

$$A^n = \underbrace{A...A}_{n \text{ copies}}$$

the product of n copies of A.

Example 21. *Suppose A and B are two $n \times n$ matrices that satisfy $AB = B^2A$. Prove then that*

$$(AB)^2 = B^6A^2$$

$$\begin{aligned} (AB)^2 &= (B^2A)^2 \\ &= (B^2A)(B^2A) \\ &= B^2(AB)(BA) \\ &= B^2(B^2A)(BA) \\ &= B^4(AB)A \\ &= B^4(B^2A)A \\ &= B^6A^2 \end{aligned}$$

Definition 22. The **transpose** of a $m \times n$ matrix $A = (a_{ij})$ is the $n \times m$ matrix $A^T = (a_{ji})$, i.e., the ji-th entry of A^T equals to the ij-th entry of A for all i, j.

A matrix A is **symmetric** if $A = A^T$.

Example 23. *If* $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$*, then* $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

Example 24. $A^T = \begin{bmatrix} 1 & 4 & 3 \\ 4 & 5 & 6 \\ 3 & 6 & 0 \end{bmatrix}$ *is a symmetric matrix.*

Remark 25. The definition of a symmetric matrix implies that if a matrix is symmetric, then it has to be a square matrix.

Proposition 26. *If A is a $m \times n$ matrix and B is a $n \times p$ matrix, then:*

1. $(A^T)^T = A$
2. $(AB)^T = B^TA^T$

Remark 27. If we express a vector vertically as $\vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$, then the dot product of two vectors can be expressed as $\vec{u}.\vec{v} = \vec{u}^T\vec{v}$, the matrix multiplication of the transpose of $\vec{u}$ with $\vec{v}$.