5 Matrices and Matrix Operations

5.1 Matrix Operations（行列演算）

Definition 5.1 A *matrix*（行列）is an $m \times n$ rectangular array of numbers. The numbers in the array are called the *entries*（成分） in the matrix.

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

It is called an $m \times n$ matrix, a matrix with m rows and n columns, it is also denoted by $A = [a_{ij}]$. Two matrices are defined to be *equal* if they have the same size and their corresponding entries are equal. The entry $a_{i,j}$ in the i-th row j-th column of a matrix A is denoted by $(A)_{i,j}$.

An $n \times n$ matrix is called a *square matrix*（正方行列）.

Definition 5.2 [page 111] Let A and B be matrices of the same size and c a scalar. Then the *sum*（和） $A + B$ is the matrix obtained by adding the entries of B to the corresponding entries of A. The the *product* cA is the matrix obtained by multiplying each entry of the matrix A by c. The matrix cA is said to be a *scalar multiple*（スカラー倍） of A.

$$A + B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ & \cdots\cdots\cdots & & \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}, cA = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ & \cdots\cdots\cdots & & \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

Let A be an $m \times n$ matrix, and $B = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_p]$ an $n \times p$ matrix. If $\boldsymbol{x} \in \mathbb{R}^p$, then $B\boldsymbol{x} \in \mathbb{R}^n$ and hence $A(B\boldsymbol{x}) \in \mathbb{R}^m$. Then the composition（合成）of $T_1 : \mathbb{R}^p \to \mathbb{R}^n$ ($\boldsymbol{x} \to B\boldsymbol{x}$) and $T_2 : \mathbb{R}^n \to \mathbb{R}^m$ ($\boldsymbol{y} \to A\boldsymbol{y}$) is denoted by $T = T_2 \circ T_1$ and

$$T : \mathbb{R}^p \to \mathbb{R}^m \ (\boldsymbol{x} \mapsto (T_2 \circ T_1)(\boldsymbol{x}) = T_2(T_1(\boldsymbol{x})) = A(B\boldsymbol{x}))$$

is linear. The standard matrix（標準行列）C is

$$\begin{aligned} C &= [T(\boldsymbol{e}_1), T(\boldsymbol{e}_2), \ldots, T(\boldsymbol{e}_p)] = [T_2(T_1(\boldsymbol{e}_1)), T_2(T_1(\boldsymbol{e}_2)), \ldots, T_2(T_1(\boldsymbol{e}_p))] \\ &= [A(B\boldsymbol{e}_1), A(B\boldsymbol{e}_2), \ldots, A(B\boldsymbol{e}_p)] = [A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_n]. \end{aligned}$$

Definition 5.3 [page 113] Let A be an $m \times r$ matrix and $B = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n]$ be an $r \times n$ matrix whose j-th column is $\boldsymbol{b}_j$. Then

$$AB = A[\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n] = [A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_n].$$

If $A = (a_{i,j})$ is an $m \times r$ matrix and $B = (b_{k,l})$ is an $r \times n$ matrix, then the *product*（積）$C = AB$ is the $m \times n$ matrix whose (s,t) entry $c_{s,t}$ is defined as follows.

$$c_{s,t} = (A\boldsymbol{b}_t)_s = (s\text{th row of } A)\ \boldsymbol{b}_t = a_{s,1}b_{1,t} + a_{s,2}b_{2,t} + \cdots + a_{s,r}b_{r,t} = \sum_{u=1}^{r} a_{s,u}b_{u,t}.$$

$$C = AB = \begin{bmatrix} \sum_{u=1}^{r} a_{1,u}b_{u,1} & \sum_{u=1}^{r} a_{1,u}b_{u,2} & \cdots & \sum_{u=1}^{r} a_{1,u}b_{u,n} \\ \sum_{u=1}^{r} a_{2,u}b_{u,1} & \sum_{u=1}^{r} a_{2,u}b_{u,2} & \cdots & \sum_{u=1}^{r} a_{2,u}b_{u,n} \\ & \cdots\cdots\cdots & & \\ \sum_{u=1}^{r} a_{m,u}b_{u,1} & \sum_{u=1}^{r} a_{m,u}b_{u,2} & \cdots & \sum_{u=1}^{r} a_{m,u}b_{u,n} \end{bmatrix}.$$

Definition 5.4 [page 117] If A is an $m \times n$ matrix, then the *transpose*（転置） of A, denoted by $A^\top$, is defined to be the $n \times m$ matrix that results from interchanging the rows and columns of A, that is $(A^\top)_{i,j} = A_{j,i}$ $(1 \leq i \leq n,\ 1 \leq j \leq m)$.

Theorem 5.1 (Theorem 2 and 3 (page 115, 117)) *Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.*

(a) $A + B = B + A,\ A + (B + C) = (A + B) + C.$

(b) $A(BC) = (AB)C.$

(c) $A(B + C) = AB + AC,\ (B + C)A = BA + CA.$

(d) $a(B+C) = aB+aC,\ (a+b)C = aC+bC,\ a(bC) = (ab)C,\ a(BC) = (aB)C = B(aC).$

(e) $(A^\top)^\top = A.$

(f) $(A + B)^\top = A^\top + B^\top.$

(g) $(cA)^\top = cA^\top$, *where c is any scalar.*

5.2 Inverse of Matrices（逆行列）

Definition 5.5 [page 121] A square matrix with 1's on the main diagonal and 0's off the main diagonal is called an *identity matrix*（単位行列） and is denoted by I, or I_n when it is of size $n \times n$.

An $n \times n$ matrix A is is said to be *invertible*（可逆） (or *nonsingular*（正則）), if there is an $n \times n$ matrix C such that

$$CA = I \text{ and } AC = I,$$

where I is the $n \times n$ identity matrix. In this case, C is called the *inverse*（逆行列） of A. If no such matrix C can be found, then A is said to be *singular*（非正則）.

When A is invertible, the inverse is unique. In fact, if

$$CA = I = AC, \text{ and } BA = I = AB, B = BI = B(AC) = (BA)C = IC = C.$$

Theorem 5.2 (Theorem 4 (page 121)) *Let*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \textit{ Then } \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = (ad - bc)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hence A is invertible if and only if $ad - bc \neq 0$ and

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Theorem 5.3 (Theorem 7 (page 125)) *Let A be an $n \times n$ square matrix, and $I = I_n$ the identity matrix of size n. Set $C = [\,A, I\,]$. If the reduced row echelon form of C is of form $[\,I, B\,]$, then $B = A^{-1}$, otherwise the inverse of A does not exist. Thus a square matrix A is invertible if and only if the reduced row echelon form of A is I.*

Example 5.1 For a matrix

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 2 \\ 3 & 2 & 1 \end{bmatrix}, \quad \text{set} \quad C = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

We perform a sequence of elementary row operations to obtain the reduced row echelon form of C.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{bmatrix} \overset{[1,2]}{=} \begin{bmatrix} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{bmatrix} \overset{[3,1;-3]}{=} \begin{bmatrix} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & -4 & -5 & 0 & -3 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & -4 & -5 & 0 & -3 & 1 \end{bmatrix} \overset{[1,2;-2]}{=} \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & -4 & -5 & 0 & -3 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & -4 & -5 & 0 & -3 & 1 \end{bmatrix} \overset{[3,2;4]}{=} \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 4 & -3 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 4 & -3 & 1 \end{bmatrix} \overset{[3;-1]}{=} \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -4 & 3 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -4 & 3 & -1 \end{bmatrix} \overset{[2,3;-1]}{=} \begin{bmatrix} 1 & 0 & 0 & -2 & 1 & 0 \\ 0 & 1 & 0 & 5 & -3 & 1 \\ 0 & 0 & 1 & -4 & 3 & -1 \end{bmatrix}.$$

Hence we can tell that the matrix A is invertible and its inverse matrix is

$$A^{-1} = \begin{bmatrix} -2 & 1 & 0 \\ 5 & -3 & 1 \\ -4 & 3 & -1 \end{bmatrix}.$$

Proposition 5.4 (Theorem 6 (page 123)) (a) *If both A and B are invertible matrices. Then AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*

(b) *If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then $(AB)^\top = B^\top A^\top$.*

(c) *If A is an invertible matrix, then $A^\top$ is invertible and $(A^\top)^{-1} = (A^{-1})^\top$.*

Proof. (a) Since $(AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB)$, $(AB)^{-1} = B^{-1}A^{-1}$.
(b) For a matrix C, (i,j)-entry of C is denoted by $C_{i,j}$. Then

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

Using this notation let us show $(AB)^\top = B^\top A^\top$ (Theorem 3 in page 99).

$$((AB)^\top)_{ij} = (AB)_{ji} = \sum_{h=1}^{n} A_{jh}B_{hi} = \sum_{h=1}^{n} B_{hi}A_{jh} = \sum_{h=1}^{n} (B^\top)_{ih}(A^\top)_{hj} = (B^\top A^\top)_{ij}.$$

Thus $(AB)^\top = B^\top A^\top$.
(c) If $AB = I = BA$, then $B^\top A^\top = (AB)^\top = I^\top = (BA)^\top = A^\top B^\top$. Since $I^\top = I$, we have the assertion. ■

Definition 5.6 An $n \times n$ matrix is called an *elementary matrix*（基本行列） if it can be obtained from the $n \times n$ identity matrix I_n by performing a single elementary row operation.

1. $E(i;c)$: the matrix obtained from I_n by performing $[i;c]$ $(c \neq 0)$.
2. $E(i,j)$: the matrix obtained from I_n by performing $[i,j]$.
3. $E(i,j;c)$: the matrix obtained from I_n by performing $[i,j;c]$.

Proposition 5.5 (page 125) *If the elementary matrix E results from performing a certain row operation on I_m and A is an $m \times n$ matrix, then the product EA is the matrix that results when this same row operation is performed on A.*

Examples of Elementary Matrices

$$E(3;c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix},\ E(1,2) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ E(3,1;c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c & 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ cz \end{bmatrix},\ \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ x \\ z \end{bmatrix},\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ cx+z \end{bmatrix}.$$

Proposition 5.6 *Every elementary matrix is invertible, and the inverse is also an elementary matrix.*

$$E(i;c)^{-1} = E(i;1/c),\ E(i,j)^{-1} = E(i,j),\ \textit{and } E(i,j;c)^{-1} = E(i,j;-c).$$

Sketch of a Proof of Theorem 4.3 using Example 4.1. Let $C = [\,A, I\,]$. Then

$$[\,I, B\,] = E(2,3;-1)E(3;-1)E(3,2;4)E(1,2;-2)E(3,1;-3)E(1,2)[\,A, I\,] = [\,PA, P\,],$$

where $P = E(2,3;-1)E(3;-1)E(3,2;4)E(1,2;-2)E(3,1;-3)$. Hence $PA = I$ and $P = B$. So $BA = I$. Since $B = P$ is a product of invertible matrix, B is also invertible and

$$B^{-1} = E(2,1)E(3,1;3)E(1,2;2)E(3,2;-4)E(3,-1)E(2,3;1).$$

We have $B^{-1} = B^{-1}I = B^{-1}(BA) = (B^{-1}B)A = IA = A$.