MATRIX ALGEBRA

Outlines

Matrix

Consider the set $\mathbb{R}^{m\times n}$ of all matrices of size $m \times n$ (m rows, n columns). Let

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

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \equiv [b_{ij}] \in \mathbb{R}^{m\times n}$$

in which a_{ij} (b_{ij}, resp.) is called the element in the ith row, jth column of A (B, resp.), $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$.

The Operations on Matrices in $\mathbb{R}^{m\times n}$

Let $A = [a_{ij}]$, $B = [b_{ij}] \in \mathbb{R}^{m\times n}$. We define

- The relation "equality" : $A = B$ if $a_{ij} = b_{ij}$, for all , .
- The algebraic operations : $A + B = [a_{ij} + b_{ij}]$, and $\alpha A = [\alpha a_{ij}]$, for all $\alpha \in \mathbb{R}$.

Multiplying Matrices

Let $A = [a_{ik}] \in \mathbb{R}^{m\times n}$ be a matrix of size $m \times n$ and let $B = [b_{kj}] \in \mathbb{R}^{n\times p}$ of size $n \times p$. The product $C = [c_{ij}] = AB \in \mathbb{R}^{m\times p}$, of size $m \times p$, is defined by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

for $i = 1, 2, \cdots , m$, $j = 1, 2, \cdots , p$.

$$\begin{array}{rc} \textit{row } i \rightarrow & \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \\ & & & \\ & & & \end{bmatrix} \end{array} \begin{array}{c} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} \\ \uparrow \\ \textit{column } j \end{array} =$$

$$\begin{array}{cc} \begin{bmatrix} \\ a_{i1}b_{ij} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \\ \\ \end{bmatrix} & \leftarrow \ \textit{row } i \\ \uparrow & \\ \textit{column } j & \end{array}$$

Moreover, by identifying a vector $\mathbf{u} = (u_1, u_2, \cdots, u_n) \in \mathbb{R}^n$ as a column matrix, $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$

we can define the product of a matrix with a vector, $A\mathbf{u} \in \mathbb{R}^{m\times 1}$, as

$$A\mathbf{u} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ \vdots \\ a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n \end{bmatrix} \in \mathbb{R}^{m\times 1}$$

and then, the system of linear equations

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

can be written as a matrix equation $Ax = b$, with the matrix of coefficients

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

the vector of unknowns

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

and the vector of the right-hand side coefficients,

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Furthermore, using the column vectors of A,

$$\mathbf{c}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, j = 1, 2, \cdots, n$$

the matrix by vector product (1) can be rewritten as

$$A\mathbf{u} = u_1\mathbf{c}_1 + u_2\mathbf{c}_2 + \cdots + u_n\mathbf{c}_n$$

Remark

It is noted that the matrix multiplication is (i) associative, $A(BC) = (AB)C$, $\alpha(AB) = (\alpha A)B = A(\alpha B)$, for all $\alpha \in \mathbb{R}$, (ii) distributive with respect to the addition, $A(B + C) = AB + AC$, $(B + C)A = BA + CA$, $(\alpha + \beta)A = \alpha A + \beta A$, and $\alpha(A + B) = \alpha A + \alpha B$, for all $\alpha, \beta \in \mathbb{R}$, but it is NOT commutative, i.e., $AB \neq BA$ in general. For example, with

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

we have $AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq BA = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Whenever $AB = BA$, we say that A and B are commutative.

For example, every square matrices $A \in \mathbb{R}^{n \times n}$ are commutative with zero matrix, , and with the identity matrix of size n, I_n , because $A0 = 0A = 0$ and $AI_n = I_n A = A$.

Definition 1.
Let $A, B \in \mathbb{R}^{n\times n}$. We say that B is the inverse matrix of A, noted $B = A^{-1}$, when $AB = BA = I_n$. The two matrices are called invertible.
By definition, A is also the inverse matrix of B, $A = B^{-1}$, and then, $(A^{-1})^{-1} = A$.
Example 1.

Consider the two matrices, $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & 1 \\ -2 & 3 & 1 \end{bmatrix}$ and

$B = \begin{bmatrix} -1 & 4 & -3 \\ -1 & 3 & -2 \\ 1 & -1 & 1 \end{bmatrix}$. We have $AB = BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$
Hence, $A = B^{-1}$ and $B = A^{-1}$.

Moreover, let A, B be two invertible matrix, $AA^{-1} = A^{-1}A = I$ and $BB^{-1} = B^{-1}B = I$. Since
$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AA^{-1} = I$.
It follows that the product matrix AB is also invertible and
$(AB)^{-1} = B^{-1}A^{-1}$

Definition 2. A matrix E is called an elementary matrix if it can be obtained from an identity matrix by performing a single elementary row operation.

The following theorem gives a connection between the matrix multiplication and the elementary row operations.

Theorem 3. Let $A \in \mathbb{R}^{n\times n}$. If we perform a single elementary row operation as $[A|\, I_n] \rightarrow [B|\, E]$ then $B = EA$.

Example 2

By changing the first and third row,

$$[A|\, I_3] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 7 & 8 & 9 & 0 & 0 & 1 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 1 & 2 & 3 & 1 & 0 & 0 \end{array}\right] = [B|\, E$$

we get

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

By multiplying the first row by 2,

$$[A|\, I_3] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 2 & 4 & 6 & 2 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right] = [B|\, E$$

By substracting the second row with 4 times the first row,

$$[A|\, I_3] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right] =$$

we get

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

As an application of Theorem 1, we define the inverse operations for the elementary row operations as follows,
1. The inverse operation for the "interchange rows i and j" is itself, "interchange rows i and j".
2. The inverse operation for the "multiply the row i by " is the operation "multiply the row i by ".
3. The inverse operation for the "add time row j to row i" is the operation "add time row j to row i".

Corollary 4

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

Let E be an elementary matrix which is the result of some single elementary operation on the identity matrix, and let E_0 be the matrix that results by performing the corresponding inverse elementary operation. It is clear that $E_0 E = I$ and $E E_0 = I$

Outlines

Inversion Algorithm

By combining the above result, we have
If we can perfom a sequence of elementary row operations that reduces a matrix $A \in \mathbb{R}^{n \times n}$ to the identity matrix I_n , then A is invertible and its inverse A^{-1} is the result of performing the same sequence of operations on I_n . That is

$$[A| \, I_n] \rightarrow [\, I_n| \, B] \tag{1}$$

gives $B = A^{-1}$. Precisely, the sequence of elementary row operations of (1) is similar to that of the Gauss-Jordan algorithm for converting $A \rightarrow I_n$ with the only difference is that all leading 1 must be on the diagonal.

Example 3. Let

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & 1 \\ -2 & 3 & 1 \end{bmatrix}$$

By performing

$$[A|\, I_3] = \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 1 & 0 \\ -2 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 3 & 2 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 4 & -3 \\ 0 & 1 & 0 & -1 & 3 & -2 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right] = [I_3|\, A^{-1}]$$

we get $A^{-1} = \begin{bmatrix} -1 & 4 & -3 \\ -1 & 3 & -2 \\ 1 & -1 & 1 \end{bmatrix}$

If in the ith step of the Gauss-Jordan algorithm, the ith element of row i and all elements below are zeros, this matrix can not convert into the identity matrix and then is not invertible.

Example 4. Let

$$A = \begin{bmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{bmatrix}$$

Perform $[A|\, I_3] = \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 1 & 5 & -1 & 0 & 1 & 0 \\ 3 & 13 & -6 & 0 & 0 & 1 \end{array}\right]$

$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 2 & 3 & -1 & 1 & 0 \\ 0 & 4 & 6 & -3 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{17}{2} & \frac{5}{2} & -\frac{3}{2} & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -1 & -2 & 1 \end{array}\right]$

The matrix A can not convert into the identity matrix. So it is not