DAY 15: September 23rd

# Chapter II: **MATRIX ALGEBRA**

## 1 Matrix Operations

### Notation and Definition

Let $A$ be an $m \times n$ matrix, that is $A$ has $m$ rows and $n$ columns.

- The columns of $A$ are vectors in $\mathbb{R}^m$ and we often denote them by $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ so that

$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$$

- The entry in $i$th row and $j$th column us denoted by $a_{ij}$ and is called $(i, j)$-entry of $A$. So we have

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{2,1} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ a_{m1} & a_{i2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

- The **main diagonal entries** of $A$ are $a_{11}, a_{22}, a_{33} \ldots$.

- The matrix $A$ is a **diagonal matrix** if $A$ is a square $n \times n$ matrix whose nondiagonal entries are zero.

### Sums and Scalar Multiples

**Definition 1.1** (Sum of matrices)**.** Let $A$ and $B$ be both $m \times n$ matrices. Suppose $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$ and $B = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_n \end{bmatrix}$. Then the **sum** $A + B$ is defined by

$$A + B = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} + \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_n \end{bmatrix} = \begin{bmatrix} (\vec{a}_1 + \vec{b}_1) & (\vec{a}_2 + \vec{b}_2) & \cdots & (\vec{a}_n + \vec{b}_n) \end{bmatrix}$$

Alternatively, if $C = A + B$ then $C$ is again a $m \times n$ matrix with $c_{ij} = a_{ij} + b_{ij}$.

**Definition 1.2** (Scalar multiples)**.** Let $r$ be a scalar ($c \in \mathbb{R}$) and let $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$ be an $m \times n$ matrix. Then, the scalar multiple $rA$ is defined by

$$rA = r \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} = \begin{bmatrix} r\vec{a}_1 & r\vec{a}_2 & \cdots & r\vec{a}_n \end{bmatrix}.$$

**Example 1.3.**

(a) Let $A = \begin{bmatrix} 0 & 2 & -1 \\ -4 & -5 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 2 & -2 \\ -2 & -2 & -4 \end{bmatrix}$. Then

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

(b) Let $r = -2$ and $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 4 \end{bmatrix}$. Then

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

The sum of matrices have similar properties as the sum of vectors as the following shows.

**Proposition 1.4.** Let $A, B$ and $C$ be matrices of the same size, and let $r$ and $s$ be scalars. Denote by 0 the matrix whose entries are all zero (the zero matrix). Then

(a) $A + B = B + A$ (commutative)

(b) $(A + B) + c = A + (B + C)$ (associative)

(c) $A + 0 = 0 + A = A$

(d) $r(A + B) = rA + rB$ (distributive)

(e) $(r + s)A = rA + sA$ (distributive)

(f) $r(sA) = (rs)A$ (associative)

## Matrix Multiplication

**Definition 1.5** (Composition of linear transformations)**.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^p \to \mathbb{R}^n$ be a linear transformations with standard matrix $A \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$ and $B = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p \end{bmatrix}$ respectively. The **composition** $T \circ U$ of $T$ and $U$ is given by

$$T \circ U : \mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$$

with $T \circ U(\vec{x}) = T(U(\vec{x}))$ for every vector $\vec{x}$ in $\mathbb{R}^p$.

## Note

- Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$ be a vector in $\mathbb{R}^p$. We have

$$\begin{aligned} T \circ U(\vec{x}) &= T(U(\vec{x})) = A(B\vec{x}) \quad \text{note that } B\vec{x} \text{ is a vector in } \mathbb{R}^n \\ &= A(x_1\vec{b}_1 + x_2\vec{x}_2 + \cdots + x_p\vec{b}_p) \\ &= Ax_1\vec{b}_1 + Ax_2\vec{x}_2 + \cdots + Ax_p\vec{b}_p \\ &= x_1A\vec{b}_1 + x_2A\vec{x}_2 + \cdots + x_pA\vec{b}_p \quad \text{note that each } A\vec{b}_i \text{ is a vector in } \mathbb{R}^m \\ &= \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p \end{bmatrix} \vec{x} \end{aligned}$$

Hence, $T \circ U(\vec{x}) = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p \end{bmatrix} \vec{x}$

We define the product $AB$ of matrices $A$ and $B$ from the definition of the composition of the corresponding linear transformations.

**Definition 1.6** (Product of matrices)**.** Let $A$ be an $m \times n$ matrix, and if $B = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p \end{bmatrix}$ is an $n \times p$ matrix, the product $AB$ is the $m \times p$ matrix defined by

$$AB = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p \end{bmatrix}$$

## Row-Column Rule

If $A$ is an $m \times n$ matrix and if $B$ is $n \times p$ matrix, then $(i, j)$-entry $(AB)_{ij}$ of the product $AB$ is the sum of the products of corresponding entries from row $i$ of $A$ and row $j$ of $B$. That is

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

## Note:

- From the row-column rule, if $\text{row}_i(A)$ denote the $i$th row of the matrix $A$, then

$$\text{row}_i(AB) = \text{row}_i(A) \cdot B$$

- The product $AB$ is defined only when the number of the columns of $A$ is the same as the rows of $B$. If $A$ is of size $m \times n$ and if $B$ is of $n \times p$, then $AB$ is an $m \times p$ matrix.

**Example 1.7.** Let $A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 1 \\ 3 & 0 \end{bmatrix}$. Compute $AB$.

(a) By using Definition 1.6, we have $AB = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 \end{bmatrix}$, with

$$A\vec{b}_1 = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \\ -3 \end{bmatrix}, \quad A\vec{b}_2 = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

So

$$AB = \begin{bmatrix} 4 & 1 \\ 11 & -1 \\ -3 & -0 \end{bmatrix}$$

(b) By using the Row-Column Rule, we have

$$AB = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} (1)(-2) + (2)(3) & (1)(1) + (2)(0) \\ (-1)(-2) + (3)(3) & (-1)(1) + (3)(0) \\ (0)(-2) + (-1)(3) & (0)(1) + (-1)(0) \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 11 & -1 \\ -3 & -0 \end{bmatrix}$$

**Proposition 1.8.** Let $A$ be a $m \times n$ matrix, and let $B$ and $C$ be matrices of appropriate sizes so that the following sums and products are defined. Let $r$ be a scalar and let $I_n$ denote the identity matrix of size $n \times n$. Then

(a) $A(BC) = (AB)C$ (associative)

(b) $A(B+C) = AB + AC$ (left distributive)

(c) $(B+C)A = BA + CA$ (left distributive)

(d) $r(AB) = (rA)B = A(rB)$

(e) $I_m A = I_n A$.

The following is a list of important differences between the properties of matrices and that of real numbers.

**Remark 1.9.**

(a) In general $AB \neq BA$.

(b) In general, $AC = AB$ does not imply $B = C$. That is, the cancellation laws do not hold.

(c) In general, $AB = 0$ does not imply that either $A = 0$ or $B = 0$.

**Example 1.10.**

(a) Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, BA = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

So $AB \neq BA$, and $BA = 0$ yet $A \neq 0$ and $B \neq 0$.

(b) Let $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} -21 & -21 \\ 7 & 7 \end{bmatrix}, AC = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -3 & 5 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -21 & -21 \\ 7 & 7 \end{bmatrix}$$

We have $AB = AC$ yet $B \neq C$.

**Definition 1.11** (Powers of matrix)**.** Let $A$ be a $n \times n$ matrix and let $k$ be a positive integer. We define $A^k$ to be the $k$th power of $A$ by

$$A^k = \underbrace{A \cdot A \cdots A}_{k \text{ times}}$$

We put $A^0 = I_n$ and $A^1 = A$

**Example 1.12.** Let $A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$. Then

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

## The Transpose of a Matrix

**Definition 1.13** (Transpose of a matrix)**.** Let $A$ be an $m \times n$ matrix. The transpose $A^T$ of $A$ is the $n \times m$ matrix whose columns are formed from the corresponding rows of $A$. That is

$$(A^T)_{ij} = a_{ji}.$$

**Example 1.14.** Let $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 4 \end{bmatrix}$. Then

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

**Proposition 1.15.** Let $A$ and $B$ be matrices whose sizes are appropriate so the following operations are defined. Let $r$ be a scalar. Then

(a) $(A^T)^T = A$

(b) $(A + B)^T = A^T + B^T$

(c) $(rA)^T = rA^T$

(d) $(AB)^T = B^T A^T$. The transpose of a product of matrices equals the product of the transposes in the reverse order.

## Note

Note that generally $(AB)^T \neq A^T B^T$. Most of the time the product $A^T B^T$ is not even defined.