ALGEBRA OF MATRICES (2.1 AND 2.2)

DR. ESLAM BADR

1. Definition and Examples

Matrices is a language to speak linear algebra. I would say that it is the most popular language to do what we aim to do in this course, especially for CS and Engineer students. Roughly speaking, we can solve "almost" every linear algebra problem by arranging the "data" we have in a matrix and selecting an appropriate "matrix technique".

Question

Motivated by this, we have two main questions to ask:

(1) What matrices are?
(2) How can we add, subtract, multiply matrices?

A *matrix* is an ordered rectangular array of real numbers, partitioned into rows and columns.

We use the letters A, B, C, ... for matrices. For example, we can write

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \\ a_{m1} & a_{m2} & \ddots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \ddots & \vdots & \\ b_{m1} & b_{m2} & \ddots & b_{mn} \end{bmatrix}$$

or simply we write $A = (a_{ij})$ and $B = (b_{ij})$ where a_{ij} (respectively b_{ij}) denotes the (i, j)-entry of the matrix A (respectively B); the elements lying in Row i and Column j of the matrix.

A matrix A has *size* $m \times n$ if it contains exactly m rows and n columns.

Example 1.1. Consider the matrix

$$A = \begin{bmatrix} 3 & 0 & e \\ \sqrt{2} & -1/2 & 5 \end{bmatrix}$$

of size 2×3.

Here $a_{11} = 3, a_{12} = 0,\ a_{13} = e,\ a_{21} = \sqrt{2},\ a_{22} = -1/2,$ and $a_{23} = 5$. In particular, the elements of a matrix can be any real numbers; positive as 3, negative as $-1/2$, zero, integers as 5, rationals as $-1/2$, algebraic as $\sqrt{2}$ or transcendental as e.

1.1. **Special Types of Matrices.**

- **The zero matrix:** $O_{m\times n}$ of size $m \times n$ where all entries are zeros.

For example,

$$O_{1\times 2} = \begin{bmatrix} 0 & 0 \end{bmatrix}, O_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, O_{3\times 2} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \cdots$$

- **Row-vector**; a matrix is a *row-vector* if it has a single row, equivalently, if it has size $1 \times n$.

For example, $R = \begin{bmatrix} 1 & -3 & 6 \end{bmatrix}$.

- **Column-vector**; a matrix is a *column-vector* if it has a single column, equivalently, if it has size $m \times 1$.

For example, $C = \begin{bmatrix} 1 \\ -3 \\ 6 \end{bmatrix}$.

- **Square Matrix**; a matrix is a *square matrix* if the number of rows equals the number of columns, equivalently, if it has size $n \times n$.

For example,

$$A = \begin{bmatrix} 1 & -3 \\ 0 & \sqrt{5} \end{bmatrix}, B = \begin{bmatrix} 1 & -3 & 1/3 \\ 0 & \sqrt{5} & 7 \\ 2 & 2/5 & -1 \end{bmatrix}, \cdots$$

By the "main-diagonal" of a square matrix $A = (a_{ij})$ we mean

$$\begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

- **Lower triangular matrix**; a square matrix $L = (l_{ij})$ is "lower Δ" if all elements above the main-diagonal are zeros. That is, if $l_{ij} = 0$ for each $i < j$.

  So L has the shape:

$$\begin{bmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ l_{n1} & l_{n2} & \dots & & l_{nn} \end{bmatrix}.$$

For example,

$$L = \begin{bmatrix} -1 & 0 & 0 \\ 2 & \sqrt{5} & 0 \\ 0 & 5 & -3 \end{bmatrix}$$

is lower Δ.

- **Upper triangular matrix**; a square matrix $U = (u_{ij})$ is "upper Δ" if all elements below the main-diagonal are zeros, that is, if $u_{ij} = 0$ for each $i > j$.

  So U has the shape:

$$\begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}$$

For example,

$$U = \begin{bmatrix} -1 & 2 & 0 \\ 0 & \sqrt{3} & 1 \\ 0 & 0 & -3 \end{bmatrix}$$

is upper Δ.

- **Diagonal matrix**; a square matrix $D = (d_{ij})$ is "diagonal" if it is both lower Δ and upper Δ. That is, if $d_{ij} = 0$ for each $i \neq j$.

  So D has the shape:

$$\begin{bmatrix} d_{11} & & & \\ & d_{22} & \text{\Large 0} & \\ & \text{\Large 0} & \ddots & \\ & & & d_{nn} \end{bmatrix}$$

For example,

$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

is diagonal since all the off-diagonal elements are 0s.

- **Identity matrix** I_n; it is the diagonal matrix whose main-diagonal consists of $1s$.

  So, I_n has the shape:

$$I_n = \begin{bmatrix} 1 & & & \\ & 1 & \text{\Large 0} & \\ & \text{\Large 0} & \ddots & \\ & & & 1 \end{bmatrix}$$

For example,

$$I_1 = [1]\,,\; I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\; I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\cdots$$

2. Operations on matrices

Now we are ready to define operations on matrices.

(I) **Equality:** Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ are equal if and only if they have "the same size" and the "corresponding entires are equal".

Consider

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},\; B = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix},\; C = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix}.$$

We have $A \neq C$ and $B \neq C$ as A and B have size 2×2, but C has a different size 2×3. Also, $A \neq B$ because $a_{21} = 3 \neq b_{21} = 4$.

(II) **Addition/Subtraction:** Given two matrices $A = (a_{ij})$ and $B = (b_{ij})$, we have

"$A \pm B$ is defined if and only if A and B have the same size."

In this case, we add/sub element-wise. That is, $A \pm B = (a_{ij} + b_{ij})$ is the new matrix obtained from A and B by add/sub the corresponding elements for each i and j.

For example, let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},\; B = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix},\; C = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix}.$$

We have

$$\begin{aligned} A + B &= \begin{bmatrix} 1+1 & 2-2 \\ 3+0 & 4+3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 3 & 7 \end{bmatrix}, \\ B - A &= \begin{bmatrix} 1-1 & -2-2 \\ 0-3 & 3-4 \end{bmatrix} = \begin{bmatrix} 0 & -4 \\ -3 & -1 \end{bmatrix}, \\ B + C &= \text{undefined}, \\ A - C &= \text{undefined}. \end{aligned}$$

(III) **Scalar Multiplication:** Given a scalar (constant) $\alpha \in \mathbb{R}$ and a matrix $A = (a_{ij})$, we define the scalar multiplication $\alpha\, A$ to be the new matrix obtained from A by multiplying each of its element by the scalar α. That is,

$$\alpha\, A = (\alpha\, a_{ij})\,.$$

Take $A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix}$ and $\alpha = 3$, then

$$\alpha\, A = \begin{bmatrix} 3(1) & 3(0) & 3(-2) \\ 3(3) & 3(1) & 3(4) \end{bmatrix} = \begin{bmatrix} 3 & 0 & -6 \\ 9 & 3 & 12 \end{bmatrix}.$$

The interaction between add, sub, and scalar multiplication is described below.

Fact 2.1. *Let A, B and C be any $m \times n$ matrices, and let $\alpha, \beta \in \mathbb{R}$ be any scalars. Then, the following axioms are satisfied.*

- *Additive Identity Axiom; the zero matrix $O_{m\times n}$ behaves as an additive identity:*
$$A + O_{m\times n} = A \text{ and } = O_{m\times n} + A = A.$$
- *Commutativity Axiom; matrix addition is commutative:*
$$A + B = B + A.$$
- *Associativity Axiom; matrix addition is associative:*
$$(A + B) + C = A + (B + C),$$
- *Distributivity Axioms: we are allowed to distribute scalar multiplication to matrix add and vice versa. In other words, we have*
$$\begin{aligned} \alpha\,(A + B) &= \alpha\, A + \beta\, B, \\ (\alpha + \beta)\, A &= \alpha\, A + \beta\, A. \end{aligned}$$
- *Compatibility Axiom:* $(\alpha\beta)\, A = \alpha\,(\beta\, A)$.

(IV) **Matrix multiplication:** the first time we see matrix multiplication looks surprising because it is unclear why we may define it in the way it is. However, everything will make sense after Chapter four: Vector Spaces.

Given two matrices $A = (a_{ij})$ and $B = (b_{ij})$, then
"$A\,B$ is defined if and only if A has size $m \times n$ and B has size $n \times l$"

- To compute the product AB, we first verify that the number of columns in A (the left factor) equals the number of rows in B (the right factor).

- Second, if we call $AB = C = (c_{ij})$, then C has size $m \times l$. Moreover, the (i, j)-entry of $C = AB$ is obtianed by expanding the Row-Column vectors product:

$$c_{ij} = \text{Row}_{i,A} \cdot \text{Col}_{j,B}\,.$$

Example 2.2. Let $A = \begin{bmatrix} 1 & 0 \\ 3 & 1 \\ 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$.

- The matrix A has size 3×2 and B has size 2×2. In particular, the number of columns in A equals the number of rows in B. Therefore, the product AB exists and has size 3×2; the number of rows in $A \times$ the number of columns in B.

- The product BA is undefined since the number of columns in B (the left factor) is not equal the number of rows in A (the right factor).

- To compute AB (of size 3×2), we first write

$$AB = C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}.$$

Second, we obtain

$$\begin{aligned}
c_{11} &= \text{Row}_{1,A} \cdot \text{Col}_{1,B} = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = (1)(1) + (0)(3) = 1, \\
c_{12} &= \text{Row}_{1,A} \cdot \text{Col}_{2,B} = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (1)(0) + (0)(1) = 0, \\
c_{21} &= \text{Row}_{2,A} \cdot \text{Col}_{1,B} = \begin{bmatrix} 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = (3)(1) + (1)(3) = 6, \\
c_{22} &= \text{Row}_{2,A} \cdot \text{Col}_{2,B} = \begin{bmatrix} 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (3)(0) + (1)(1) = 1, \\
c_{31} &= \text{Row}_{3,A} \cdot \text{Col}_{1,B} = \begin{bmatrix} 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = (2)(1) + (5)(3) = 17, \\
c_{32} &= \text{Row}_{3,A} \cdot \text{Col}_{2,B} = \begin{bmatrix} 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (2)(0) + (5)(1) = 5.
\end{aligned}$$

Summing up, we obtain $AB = \begin{bmatrix} 1 & 0 \\ 6 & 1 \\ 17 & 5 \end{bmatrix}$.

Question

For any two numbers a and b, we know that $ab = ba$.
Do you think that we can extend this property for matrices?

The answer is **NO** as shown by the next example.

Example 2.3. Decide whether the following statement is True or False.

"We have $AB = BA$ for any matrices A and B."

This is **FALSE**, since, in Example 2.2, we have A and B such that AB exists but BA does not exist. So it is an example showing that matrix multiplication is not commutative.

Question

For any two numbers a and b, we know that $ab = 0$ implies that $a = 0$ or $b = 0$.
Do you think that we can extend this property for matrices?

Again, the answer is **NO**, see the next example.

Example 2.4. Decide whether the following statement is True or False.

"if $AB = O$ then $A = O$ or $B = O$."

This is **FALSE**, since there are many counter examples. For instance, if we take

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

then $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = O$ but neither A nor B equals O.

Question

We have a "cancellation property" valid for numbers, in the sense that if $c\,a = c\,b$ with $c \neq 0$ then $a = b$.
Do you think we can extend this property for matrices?

The answer still **NO**.

Example 2.5. Decide whether the following statement is True or False.

"if $C\,A = C\,B$ such that $C \neq O$, then $A = B$."

This is **FALSE**. Take

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

then $CA = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = CB$, but $A \neq B$.

So having a non-zero common factor for matrices is not sufficient to do cancellation!
Can you find a sufficient condition on C that allows us to cancel it?
No worries as we will discuss this later :)

However, we can't say that matrix multiplication is so bad because it still have some nice properties as demonstrated below.

Fact 2.6. *The following axioms hold for matrix multiplication.*

- *Multiplicative identity Axiom; if A is an $m \times n$ matrix, then*
$$A\,I_n = A, \text{ and } I_m\,A = A.$$
- *Associativity Axiom; matrix multiplication is associative:*
$$(A\,B)\,C = A\,(B\,C).$$
- *Compatibility Axiom; we have that*
$$\alpha\,(A\,B) = A\,(\alpha\,B) = (\alpha\,A)\,B.$$
- *Distributivity Axioms; we are allowed to distribute multiplication to add/sub from both directions. In other words, we have*
$$\begin{aligned} A\,(B \pm C) &= A\,B \pm A\,C, \\ (B \pm C)\,A &= B\,A \pm C\,A. \end{aligned}$$

(V) **Powers:** For a square matrix A of size $n \times n$, we define its powers by

$$A^0 = I_n,\ A^1 = A,\ A^2 = A\,A,\ \cdots,\ A^k = A\,A\,\cdots\,A(k\,times)$$

(VI) **Transpose:** For any matrix A of size $m \times n$, the "transpose" of A, denoted by A^T, is the matrix obtained by writing the rows of A as columns.
Therefore, A^T has size $n \times m$.

For $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 1 & 1 \end{bmatrix}$, we get $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$.

The $(\cdot)^T$ operation behaves nicely with addition/subtraction, scalar multiplication, and multiplication as illustrated by the next fact.

Fact 2.7. *The following properties are true.*
- $(A \pm B)^T = A^T \pm B^T, (A \pm B \pm C)^T = A^T \pm B^T \pm C^T, \ldots$
- $(\alpha\,A)^T = \alpha\,A^T,$
- $(A^T)^T = A,$
- $(A\,B)^T = B^T\,A^T, (A\,B\,C)^T = C^T\,B^T\,A^T, \ldots$

Next, we can use the transpose operation to determine whether a given matrix is **symmetric**, **skew-symmetric** or **neither**.

A square matrix A is said to be "symmetric" if $A^T = A$. This means that we have reflection around the main-diagonal of A. Equivalently, $a_{ij} = a_{ji}$ for each i and j.

Example 2.8. The matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 8 \\ 3 & 8 & 0 \end{bmatrix}$ is symmetric as $a_{12} = a_{21}$, $a_{13} = a_{31}$, $a_{23} = a_{32}$, or simply because $A^T = A$.

The matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ -5 & -1 & 8 \\ 3 & 8 & 0 \end{bmatrix}$ is **NOT** symmetric as $a_{12} = 2 \neq a_{21} = -5$, in particular, $A^T \neq A$.

A square matrix A is said to be "skew-symmetric" if $A^T = -A$. This means that we have reflection around the main-diagonal up to a negative sign. That is, $a_{ij} = -a_{ji}$ for each i and j. In particular, the main-diagonal elements must be zeros.

Example 2.9. The matrix $A = \begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 8 \\ -3 & -8 & 0 \end{bmatrix}$ is skew-symmetric, since $a_{11} = a_{22} = a_{33} = 0, a_{12} = -a_{21}$, $a_{13} = -a_{31}$, $a_{23} = -a_{32}$, or simply because $A^T = -A$.

The matrix $A = \begin{bmatrix} 0 & 2 & 3 \\ 5 & 0 & 8 \\ -2 & -8 & 0 \end{bmatrix}$ is **NOT** skew-symmetric as $a_{12} = 2 \neq -a_{21} = -5$, in particular, $A^T \neq -A$.

Now I want to convince you that these (skew)-symmetric matrices can be used to understand any other matrix in a sense.

Example 2.10. Decide whether the following statement is True or False.

"For any square matrix A, the matrix $A + A^T$ is symmetric."

This is **TRUE**. To see this, let us call $B = A + A^T$. To show that B is symmetric, you need to verify that $B^T = B$.
As a consequence of Fact 2.1 and Fact 2.7, we have that

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

Example 2.11. Decide whether the following statement is True or False.

"For any square matrix A, the matrix $A - A^T$ is skew-symmetric."

This is **TRUE**. To see this, let us call $C = A - A^T$. To show that C is skew-symmetric, you need to verify that $C^T = -C$.
As a consequence of Fact 2.1 and Fact 2.7, we deduce that

$$C^T = (A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T) = -C.$$

As a result of Examples 2.10 and 2.11:

Fact 2.12. *Any square matrix A can be written as the sum $B + C$ of two matrices; a symmetric matrix B and a skew-symmetric matrix C.*

Proof. By Example 2.10 and Example 2.11, we know that $A+A^T$ is symmetric and $A - A^T$ is skew-symmetric. Hence $\frac{1}{2}(A + A^T)$ is also symmetric and $\frac{1}{2}(A - A^T)$ is skew-symmetric. Finally, we always can write

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) = B + C$$

with $B = \frac{1}{2}(A + A^T)$ symmetric and $C = \frac{1}{2}(A - A^T)$ skew-symmetric.

This proves the result. $\square$

Question

Enjoy yourself by answering the following exercise.

True/False: "For any $m \times n$ matrix A, we have that $A\,A^T$ and $A^T A$ are symmetric."