1. Matrices (I)

1.1. Introduction.

A rectangular array of mn elements a_{ij} into m rows and n columns, where the elements a_{ij} belong to a field F, is said to be a *matrix* of order $m \times n$ (or an $m \times n$ matrix) over the field F. An $m \times n$ matrix is exhibited in the form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \text{ or in the form } \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Definitions.

1. **Equal matrices.** Two matrices A and B are said to be *equal* if A and B have the same order and their corresponding elements be equal. Thus if $A = (a_{ij})_{m,n}$ and $B = (b_{ij})_{m,n}$, then $A = B$ if and only if $a_{ij} = b_{ij}$ for $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$.

Scanned with CamScanner

2. Diagonal matrix. A square matrix is said to be a *diagonal matrix* if the elements other than the diagonal elements be all zero.

Examples of a real diagonal matrix are

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The diagonal matrix $(d_{ij})_{n,n}$ is denoted by diag $(d_{11}, d_{22}, \ldots, d_{nn})$.

3. Scalar matrix. A diagonal matrix is said to be a *scalar matrix* if all the diagonal elements be the same scalar.

4. Identity (or unit) matrix. A scalar matrix whose diagonal elements are all 1, the identity element of the ground field F, is said to be an *identity* matrix (or a *unit* matrix). The identity matrix of order n is denoted by I_n.

$$\text{Thus } I_n = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & & \ldots & \\ 0 & 0 & \ldots & 1 \end{pmatrix} = (\delta_{ij})_{n,n}, \text{ where } \delta_{ij} = 1, \text{ if } i = j,$$

$$\delta_{ij} = 0, \text{ if } i \neq j.$$

5. Triangular matrix.

A square matrix (a_{ij}) is said to be an *upper triangular* matrix if all the elements below the diagonal are 0. That is, $a_{ij} = 0$ if $i > j$.

A square matrix (a_{ij}) is said to be a *lower triangular* matrix if all the elements above the diagonal are 0. That is, $a_{ij} = 0$ if $i < j$.

A square matrix is said to be a *triangular* matrix if it is either upper triangular or lower triangular.

A triangular matrix $(a_{ij})_{n,n}$ is said to be *strictly triangular* if $a_{ii} = 0$ for $i = 1, 2, \ldots, n$.

Examples of a real triangular matrix are

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 4 & 1 & 6 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

A diagonal matrix is both upper triangular and lower triangular.

1.2. Algebraic operations on matrices.

We consider matrices over the same scalar field F.

1. Multiplication by a scalar. The product of an $m \times n$ matrix $A = (a_{ij})_{m,n}$ by a scalar c where $c \in F$, the field of scalars, is a matrix

$B = (b_{ij})_{m,n}$ defined by $b_{ij} = ca_{ij}$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$ and is written as cA. Thus we have $c(a_{ij})_{m,n} = (ca_{ij})_{m,n}$.

Let A be an $m \times n$ matrix and c, d are scalars. Then the following results are obvious.

(i) $c(dA) = (cd)A$,

(ii) $0A = O_{m,n}$, 0 being the zero element of F,

(iii) $cO_{m,n} = O_{m,n}$,

(iv) $cI_n = \begin{pmatrix} c & 0 & \ldots & 0 \\ 0 & c & \ldots & 0 \\ \ldots & & \ldots & \\ 0 & 0 & \ldots & c \end{pmatrix}$,

(v) $1A = A$, 1 being the identity element of F.

The scalar matrix of order n whose diagonal elements are all c can be expressed as cI_n.

2. Addition. Two matrices A and B are said to be *conformable for addition* if they have the same order.

If $A = (a_{ij})_{m,n}$ and $B = (b_{ij})_{m,n}$, then their sum $A + B$ is the matrix $C = (c_{ij})_{m,n}$, where $c_{ij} = a_{ij} + b_{ij}, i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$.

If A and B be matrices of different orders, then $A + B$ is not defined.

Let A, B be $m \times n$ matrices and c, d are scalars. Then the following results are obvious.

(i) $c(A + B) = cA + cB$, (ii) $(c + d)A = cA + dA$.

Theorem 1.2.1. Matrix addition is commutative.

This says that if A and B be two matrices such that $A + B$ is defined, then $A + B = B + A$.

Proof. Let $A = (a_{ij})_{m,n}$ $B = (b_{ij})_{m,n}$.

Let $A + B = C = (c_{ij})_{m,n}$ and $B + A = D = (d_{ij})_{m,n}$.

$$\begin{aligned} \text{Then } c_{ij} &= a_{ij} + b_{ij} \\ &= b_{ij} + a_{ij}, \quad \text{since } a_{ij}, b_{ij} \in F, \text{ the ground field} \\ &= d_{ij}. \end{aligned}$$

Since C and D are of the same order and $c_{ij} = d_{ij}$, $C = D$.

That is, $A + B = B + A$. This completes the proof.

Theorem 1.2.2. Matrix addition is associative.

This says that if A, B, C be matrices such that the matrices $B+C$, $A+(B+C)$, $A+B$, $(A+B)+C$ are defined, then $A+(B+C) = (A+B)+C$.

3. Multiplication of Matrices.

Two matrices A and B are said to be *conformable for the product* AB if the number of columns of A be equal to the number of rows of B. If $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{n,p}$ then the product AB is a matrix of order $m \times p$ and $AB = C = (c_{ij})_{m,p}$ where $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$, $i = 1, 2, \ldots, m; j = 1, 2, \ldots, p$.

The ijth element of the product AB is obtained by multiplying the corresponding elements of the ith row of A and the jth column of B and adding the products.

If the number of columns of A be not equal to number of rows of B, then AB is not defined.

It is obvious that the products AB and BA are two distinct entities. Indeed, one of them may exist while the other may not.

For an $m \times n$ matrix A, in order that both AB and BA should exist, B must be of order $n \times m$. In this case, however, AB and BA are matrices of different orders. In order that both AB and BA should exist as matrices of the same order, both A and B must be square matrices of the same order.

In the product AB, A is said to be a *pre-multiplier* and B is said to be a *post-multiplier.*

Note. Matrix multiplication is not commutative. That is, for two matrices A and B, $AB \neq BA$, in general.

First of all, if we choose the orders of A and B to be $m \times n$ and $n \times m$ respectively so that the conformability conditions for both the products AB and BA are satisfied, then we observe that the orders of AB and BA are $m \times m$ and $n \times n$ respectively and therefore AB cannot be equal to BA.

In order that AB and BA may be equal, both of them must be of the same order and this requires that A and B must be square matrices of the same order. However if we choose the orders of A and B to be $n \times n$ and $n \times n$, then although AB and BA become matrices of the same order, they may not be equal, in general.

This can be shown by taking at random

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \; B = \begin{pmatrix} 5 & 6 \\ 4 & 2 \end{pmatrix}.$$

Here $AB = \begin{pmatrix} 13 & 10 \\ 22 & 18 \end{pmatrix}$, $BA = \begin{pmatrix} 17 & 28 \\ 8 & 14 \end{pmatrix}$.

In some special cases, however, $AB = BA$.

For example, let $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $BA = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$.

Definition. Two matrices A and B are said to *commute* with each other if $AB = BA$. Since $AB = BA$, A and B must be square matrices of the same order.

Some Examples of commuting matrices.

1. Let A be a square matrix. Then A commutes with A itself.

2. Let A be a square matrix of order n. Then A commutes with I_n, because $A.I_n = I_n.A = A$.

3. Let A be a square matrix of order n. Then A commutes with $O_{n,n}$, because $A.O_{n,n} = O_{n,n}.A = O_{n,n}$.

4. Let A be a square matrix of order n. Then A commutes with the scalar matrix cI_n, because $A.cI_n = cI_n.A = cA$.

Definition. Divisor of zero. A non-zero matrix A of order $m \times n$ is said to be a *divisor of zero* if there exists a non-zero matrix B of order $n \times p$ such that $AB = O_{m,p}$, or if there exists a non-zero matrix C of order $p \times m$ such that $CA = O_{p,n}$.

When $AB = O$, A is said to be a *left divisor of zero* and B is said to be a *right divisor of zero.*

Example.

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 6 & -4 \\ -3 & 2 \end{pmatrix}$,

$C = \begin{pmatrix} 4 & 2 & 6 \\ 6 & 3 & 9 \end{pmatrix}$, $D = \begin{pmatrix} 6 & 0 & 8 \\ 5 & 4 & 8 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. So A is a left divisor of zero and B is a right divisor of zero. $BA \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Also $BC = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

So B is a left divisor of zero and C is a right divisor of zero. CB is not defined.

$AC = \begin{pmatrix} 16 & 8 & 24 \\ 32 & 16 & 48 \end{pmatrix}$, $AD = \begin{pmatrix} 16 & 8 & 24 \\ 32 & 16 & 48 \end{pmatrix}$. $AC = AD$, but $C \neq D$. This happens because $A(C-D) = O$ does not imply $C - D = O$.

Theorem 1.2.3. Matrix multiplication is associative.

1.3. Transpose of a matrix.

Let A be an $m \times n$ matrix. Then the $n \times m$ matrix obtained by interchanging rows and columns of A is said to be the *transpose* of A and is denoted by A^t (or A^T).

Thus if $A = (a_{ij})_{m,n}$ then $A^t = B = (b_{ij})_{n,m}$, where $b_{ij} = a_{ji}$, $i = 1, 2, ..., n$; $j = 1, 2, ..., m$.

Theorem 1.3.1. $(A^t)^t = A$.

The proof is obvious.

Theorem 1.3.2. If A and B be two matrices such that $A + B$ is defined then $(A + B)^t = A^t + B^t$.

Proof. Let $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{m,n}$. Then $A + B$ is defined.

The order of $A + B$ is $m \times n$ and the order of $(A + B)^t$ is $n \times m$.

Also, the order of A^t is $n \times m$, the order of B^t is $n \times m$.

Therefore the order of $A^t + B^t$ is $n \times m$.

Thus the order of $(A + B)^t$ = the order of $(A^t + B^t)$ ... (i)

Scanned with CamScanner

Again, the ijth element of $(A+B)^t$
$=$ the jith element of $(A+B)$
$=$ the jith element of A + jith element of B
$=$ the ijth element of A^t + ijth element of B^t
$=$ the ijth element of (A^t+B^t) $\quad\dots\quad$ (ii)

From (i) and (ii) it follows that $(A+B)^t = A^t + B^t$.

Theorem 1.3.3. If c be a scalar, $(cA)^t = cA^t$.

Proof. Let $A = (a_{ij})_{m,n}$. Then the order of $(cA)^t$ is $n \times m$ and the order of cA^t is $n \times m$.

Thus the order of $(cA)^t$ = the order of cA^t $\quad\dots\quad$ (i)

Again, the ijth element of $(cA)^t$
$=$ the jith element of cA
$= c$(jith element of A)
$= c$(ijth element of A^t)
$=$ the ijth element of cA^t $\quad\dots\quad$ (ii)

From(i) and (ii) it follows that $(cA)^t = cA^t$.

Corollary. If A and B be two matrices of the same order $(cA+dB)^t = cA^t + dB^t$, where c, d are scalars.

Theorem 1.3.4. If A and B be two matrices such that AB is defined, then $(AB)^t = B^tA^t$.

Proof. Let $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{n,p}$. Then AB is defined.
The order of AB is $m \times p$. So the order of $(AB)^t$ is $p \times m$.
Also, the order of B^t is $p \times n$, the order of A^t is $n \times m$.
So the order of B^tA^t is $p \times m$.

Thus the order of $(AB)^t$ = the order of B^tA^t $\quad\dots\quad$ (i)

Again, the ijth element of $(AB)^t$
$=$ the jith element of AB
$=$ the sum of the products of corresponding elements of the jth row of A and the ith column of B
$=$ the sum of the products of corresponding elements of the jth column of A^t and the ith row of B^t
$=$ the sum of the products of corresponding elements of the ith row of B^t and the jth column of A^t
$=$ the ijth element of B^tA^t $\quad\dots\quad$ (ii)

From (i) and (ii) it follows that $(AB)^t = B^tA^t$.

Note. If A, B, C be three matrices such that ABC is defined, then $(ABC)^t = C^tB^tA^t$.

In general, if $A_1, A_2, ..., A_n$ be n matrices such that the product $A_1A_2...A_n$ is defined, then $(A_1A_2...A_n)^t = A_n^t...A_2^tA_1^t$.

1.4. Symmetric and skew symmetric matrices.

A square matrix A is said to be *symmetric* if $A = A^t$. Therefore $A = (a_{ij})$ is symmetric if $a_{ij} = a_{ji}$.

A square matrix A is said to be *skew symmetric* if $A = -A^t$. Therefore $A = (a_{ij})$ is skew symmetric if $a_{ij} = -a_{ji}$.

Examples of a symmetric matrix are

$$\begin{pmatrix} 1 & 3 & 5 \\ 3 & 0 & 7 \\ 5 & 7 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2+i & 3 \\ 2+i & i & 1-i \\ 3 & 1-i & 0 \end{pmatrix}, \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}.$$

Examples of a skew symmetric matrix are

$$\begin{pmatrix} 0 & 1 & -8 \\ -1 & 0 & 2 \\ 8 & -2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2+i & 3i \\ -2-i & 0 & 2 \\ -3i & -2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Note. An $n \times n$ null matrix is both symmetric and skew symmetric.

Theorem 1.4.1. If A and B be two symmetric matrices of the same order then $A + B$ is symmetric.

Proof. $(A+B)^t = A^t + B^t = A + B$, since $A^t = A$, $B^t = B$.
This proves that $A + B$ is symmetric.

Theorem 1.4.2. If A and B be two symmetric matrices of the same order then AB is symmetric if and only if $AB = BA$.

Proof. Let AB be symmetric.

$$\begin{aligned} \text{Then } AB &= (AB)^t \\ &= B^tA^t = BA, \text{ since } B^t = B,\ A^t = A. \end{aligned}$$

Conversely, let $AB = BA$.

$$\begin{aligned} \text{Then } (AB)^t &= B^tA^t = BA, \text{ since } B^t = B,\ A^t = A \\ &= AB, \text{ by the assumed condition.} \end{aligned}$$

Therefore AB is symmetric.
This completes the proof.

Theorem 1.4.3. If A be an $m \times n$ matrix, then the matrices AA^t and A^tA are both symmetric.

AA^t and A^tA are square matrices of order m and n respectively.
$(AA^t)^t = (A^t)^tA^t = AA^t$ and $(A^tA)^t = A^t(A^t)^t = A^tA$.
This shows that AA^t and A^tA are both symmetric matrices.

Theorem 1.4.6. A real (or complex) square matrix can be uniquely expressed as the sum of a symmetric matrix and a skew symmetric matrix.

Proof. Let A be a given matrix. Then A can be expressed as
$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$.

Now $[\frac{1}{2}(A + A^t)]^t = \frac{1}{2}[A^t + (A^t)^t] = \frac{1}{2}[A + A^t]$

and $[\frac{1}{2}(A - A^t)]^t = \frac{1}{2}[A^t - (A^t)^t] = \frac{1}{2}[A^t - A] = -\frac{1}{2}[A - A^t]$.

This shows that $\frac{1}{2}(A + A^t)$ is a symmetric matrix and $\frac{1}{2}(A - A^t)$ is a skew symmetric mayrix.

Therefore A is expressed as the sum of a symmetric matrix and a skew symmetric matrix.

We now show that this decomposition is unique.

Let $A = P + Q$ where P is symmetric and Q is skew symmetric.
Then $A^t = P^t + Q^t = P - Q$.
We have $A + A^t = 2P$, $A - A^t = 2Q$.
So $P = \frac{1}{2}(A + A^t)$, $Q = \frac{1}{2}(A - A^t)$ and this proves the theorem.

Note. The theorem does not hold if the ground field F be of characteristic 2.

Worked Example (continued).

2. Express $A = \begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 1 & 6 & 8 \end{pmatrix}$ as the sum of a symmetric matrix and a skew symmetric matrix.

Let $A = P + Q$ where P is symmetric and Q is skew symmetric.
Then $A^t = P^t + Q^t = P - Q$.
We have $P = \frac{1}{2}(A + A^t)$, $Q = \frac{1}{2}(A - A^t)$.

$$P = \frac{1}{2}\left[\begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 1 & 6 & 8 \end{pmatrix} + \begin{pmatrix} 4 & 3 & 1 \\ 5 & 7 & 6 \\ 1 & 2 & 8 \end{pmatrix}\right] = \begin{pmatrix} 4 & 4 & 1 \\ 4 & 7 & 4 \\ 1 & 4 & 8 \end{pmatrix},$$

$$Q = \frac{1}{2}\left[\begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 1 & 6 & 8 \end{pmatrix} - \begin{pmatrix} 4 & 3 & 1 \\ 5 & 7 & 6 \\ 1 & 2 & 8 \end{pmatrix}\right] = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}.$$

$$\text{Therefore } A = \begin{pmatrix} 4 & 4 & 1 \\ 4 & 7 & 4 \\ 1 & 4 & 8 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}.$$

Definition. A square matrix A is said to be *idempotent* if $A^2 = A$.