Euclidean Spaces $\mathbb{R}^n$ and spaces $\mathbb{C}^n$

We denote by $\mathbb{R}$ the set of all real numbers and by $\mathbb{C}$ the set

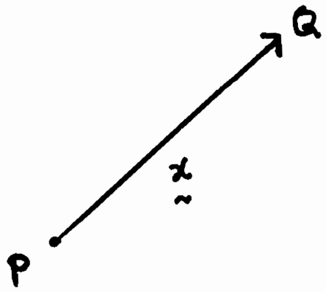
$$\left\{ z = x + iy \; : \; x, y \in \mathbb{R} \right\}$$

of all complex numbers. Here i denotes the <u>imaginary unit</u> $i^2 = -1$. Now let $n \geq 1$ be an integer, and let

$$\mathbb{R}^n := \left\{ \underset{\sim}{x} = \underbrace{(x_1, \cdots, x_n)}_{n\text{-tuple of real numbers}} \; : \; x_i \in \mathbb{R}, \; i = 1, \cdots, n \right\}$$

$$\mathbb{C}^n := \left\{ Z = \underbrace{(z_1, \cdots, z_n)}_{n\text{-tuple of complex numbers}} \; : \; z_i \in \mathbb{C}, \; i = 1, \cdots, n \right\}$$

Note that when $n = 1$, $\mathbb{R}^n$ is just written as $\mathbb{R}$ and $\mathbb{C}^n$ is written simply as $\mathbb{C}$. Let us recall that a <u>physical vector</u> is a directed line segment $\underset{\sim}{x}$.



If $P = (u_1, u_2, u_3)$ is the initial point and $Q = (v_1, v_2, v_3)$ is the terminal point of this vector, then the components of $\underset{\sim}{x}$ are $x_1 = v_1 - u_1$, $x_2 = v_2 - u_2$, $x_3 = v_3 - u_3$.

Thus the vector $\underset{\sim}{x}$ can be identified with the ordered triplet (x_1, x_2, x_3) of its components. If $\underset{\sim}{x} = (x_1, x_2, x_3)$ and $\underset{\sim}{y} = (y_1, y_2, y_3)$ are two vectors in the three space $\mathbb{R}^3$, then using 'parallelogram law', one easily verifies that

$$\underset{\sim}{x} + \underset{\sim}{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3).$$

In other words, the addition of (physical) vectors in the three space is 'component wise' addition.

Next, we regard real numbers as scalars, and note that multiplication of a vector $\underset{\sim}{x} = (x_1, x_2, x_3)$ by a scalar $\alpha \in \mathbb{R}$, called 'scalar multiplication' is also componentwise:

$$\alpha \underset{\sim}{x} = (\alpha x_1, \alpha x_2, \alpha x_3).$$

If we now call the elements of $\mathbb{R}^n$ as 'vectors', then the above observations make it natural to define the two basic operations of addition and scalar multiplication on these vectors componentwise:

$$\underset{\sim}{x} + \underset{\sim}{y} = (x_1+y_1, \ldots\ldots, x_n+y_n), \quad \text{and}$$

$$\alpha \underset{\sim}{x} = (\alpha x_1, \ldots\ldots, \alpha x_n), \quad \text{whenever}$$

$$\underset{\sim}{x} = (x_1, \ldots, x_n), \quad \underset{\sim}{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n \text{ and } \alpha \in \mathbb{R}.$$

It is now important to observe that these two basic operations satisfy the following properties:

I. Closure properties

1. (closure under addition) For every pair of elements $\underset{\sim}{x}, \underset{\sim}{y}$ in $\mathbb{R}^n$, there is precisely one element $\underset{\sim}{x} + \underset{\sim}{y} \in \mathbb{R}^n$ called the sum of $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

2. (closure under scalar multiplication)
For every pair of elements $\alpha, \underset{\sim}{x}$ where $\alpha \in \mathbb{R}$ and $\underset{\sim}{x} \in \mathbb{R}^n$, there is a unique element $\alpha \underset{\sim}{x}$, called the (scalar) multiple of $\underset{\sim}{x}$ by α.

II. Properties of addition

3. (Commutative law) $\underset{\sim}{x} + \underset{\sim}{y} = \underset{\sim}{y} + \underset{\sim}{x}$, for all $\underset{\sim}{x}, \underset{\sim}{y}$ in $\mathbb{R}^n$

4. (associative law) $\underset{\sim}{x} + (\underset{\sim}{y} + \underset{\sim}{z}) = (\underset{\sim}{x} + \underset{\sim}{y}) + \underset{\sim}{z}$, for all $\underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{z}$ in $\mathbb{R}^n$

5. (existence of zero element) There exists a unique element $\underset{\sim}{0} = (0, \cdots, 0)$ in $\mathbb{R}^n$ such that

$$\underset{\sim}{x} + \underset{\sim}{0} = \underset{\sim}{0} + \underset{\sim}{x} = \underset{\sim}{x}, \quad \text{for all } \underset{\sim}{x} \in \mathbb{R}^n.$$

6. (existence of additive inverses or negatives) For every $\underset{\sim}{x} \in \mathbb{R}^n$, there exists a unique element written $-\underset{\sim}{x} \in \mathbb{R}^n$ such that $\underset{\sim}{x} + (-\underset{\sim}{x}) = \underset{\sim}{0}$.

III. Properties of scalar multiplication

7. (associativity) For all $\alpha, \beta \in \mathbb{R}$ and $\underset{\sim}{x} \in \mathbb{R}^n$,
$$\alpha(\beta \underset{\sim}{x}) = (\alpha\beta)\underset{\sim}{x}.$$

8. (distributive law for addition in $\mathbb{R}^n$) For all $\underset{\sim}{x}, \underset{\sim}{y}$ in $\mathbb{R}^n$ and α in $\mathbb{R}$,

$$\alpha(\underset{\sim}{x} + \underset{\sim}{y}) = \alpha\underset{\sim}{x} + \alpha\underset{\sim}{y}.$$

9. (distributive law for addition in $\mathbb{R}$) For all $\alpha, \beta \in \mathbb{R}$ and $\underset{\sim}{x} \in \mathbb{R}^n$,

$$(\alpha + \beta)\underset{\sim}{x} = \alpha\underset{\sim}{x} + \beta\underset{\sim}{x}.$$

10. (existence of identity for scalar multiplication) For all $\underset{\sim}{x} \in \mathbb{R}^n$, $1\underset{\sim}{x} = \underset{\sim}{x}$.

Definition A nonempty set X of objects (called elements or vectors) is called a vector space or a linear space if its elements $x, y, z, \ldots$ (in place of $\underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{z}, \ldots$) and real numbers $\alpha, \beta, \ldots$ satisfy properties 1–10 (to be called the axioms of vector space) for the operations of addition and scalar multiplication.

Remark Vector spaces defined as above are called vector spaces over $\mathbb{R}$ or real vector spaces. If we replace real numbers in the above axioms by complex numbers then we get the definition of a vector space over $\mathbb{C}$ or a complex vector space, in which the scalars are complex.

Examples (i) Let $X = \mathbb{R}^n$. Properties (1) $\to$ (10) make X into a real vector space. Note carefully, that as a particular case $X = \mathbb{R}$ becomes a real vector space.

(ii) Let $X = \mathbb{C}^n$. In $\mathbb{C}^n$ define 'addition' and 'scalar multiplication' componentwise:

$$Z + W = (z_1 + w_1, \ldots, z_n + w_n), \quad \alpha Z = (\alpha z_1, \ldots, \alpha z_n)$$

for $Z = (z_1, \ldots, z_n)$, $W = (w_1, \ldots, w_n) \in \mathbb{C}^n$. Firstly, take the 'scalars' α to be real numbers and verify that all the axioms (1)–(10) are satisfied. Hence

$\mathbb{C}^n$ becomes a real vector space. Next, note that if we allow scalars α to range over $\mathbb{C}$, the set of complex numbers, the axioms 1–10 continue to be satisfied. This observation enables us to regard $\mathbb{C}^n$ as a complex vector space as well. As a particular case, we note that $X = \mathbb{C}$ can be regarded as a real vector space as well as a complex vector space. We will give more examples of 'abstract' vector spaces a little later. First let us observe some elementary properties.

Proposition Let X be a vector space, $x \in X$, and α be a scalar. We have:

(a) $0x = 0$

(b) $\alpha 0 = 0$

(c) $(-1)x = -x$

(d) $\alpha x = 0 \Longrightarrow$ either $\alpha = 0$ or $x = 0$

(Note that we denote by 0 both the zero scalar as well as the null vector.)

Proof (a) We have $0x + 0x = (0+0)x = 0x$.

There fore $(0x + 0x) + (-0x) = 0x - 0x$,

i.e., $0x + (0x + (-0x)) = 0x - 0x \Rightarrow 0x + 0 = 0 \Rightarrow 0x = 0$

(c) We have $x + (-1)x = 1 \cdot x + (-1)x = (1-1)x = 0x = 0$.

Hence, $(-1)x = -x$.

Proofs of (b) and (d) are left as exercises. ∎

Definition A subset W of a vector space X is called a subspace of X if W is itself a vector space under the addition and scalar multiplication defined on X.

Proposition A nonempty subset W of a vector space X is a subspace of X if and only if W satisfies the closure axioms.

Proof. If W is a subspace of X, then all the vector space axioms are satisfied; in particular (1) and (2) are satisfied. Conversely, assume that W satisfies the closure axioms 1-2. We need only prove the existence of inverses and the zero element in W. Indeed, by distributivity, for $x \in W$

$$(0+0)x = 0x = 0x + 0x \Rightarrow 0x = 0.$$

Hence by closure axioms, $0 \in W$. Also if $x \in W$, $(-1)x$ is the negative of x, which is in W by closure axioms. ∎

Remark : A nonempty subset W of a vector space X is a subspace $\Longleftrightarrow$ (i) $x, y \in W \Rightarrow x+y \in W$

(ii) $x \in W$ and α is a scalar $\Rightarrow \alpha x \in W$.

Examples (i) $W = \mathbb{R}$ is a subspace of the real vector space $X = \mathbb{C}$; but it is not a subspace of the complex vector space $\mathbb{C}$.

(ii) Let $W = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 : \; a_1 x_1 + a_2 x_2 + a_3 x_3 = 0 \right\}$ be a plane passing through $\underset{\sim}{0} = (0,0,0)$ in the vector space $X = \mathbb{R}^3$. Then W is a subspace of X. Indeed, if $\underset{\sim}{x} = (x_1, x_2, x_3)$, $\underset{\sim}{y} = (y_1, y_2, y_3) \in W$, then

$$\left.\begin{aligned} a_1 x_1 + a_2 x_2 + a_3 x_3 = 0 \\ a_1 y_1 + a_2 y_2 + a_3 y_3 = 0 \end{aligned}\right\} \Rightarrow a_1(x_1+y_1) + a_2(x_2+y_2) + a_3(x_3+y_3) = 0$$

$\Rightarrow$ $\underset{\sim}{x} + \underset{\sim}{y} \in W$. Also, $\alpha \underset{\sim}{x} \in W$ for $\alpha \in \mathbb{R}$.

Hence W is a subspace of $\mathbb{R}^3$.

(iii) $W = \left\{ (0, x_2, x_3, \cdots, x_{n-1}, 0) : x_i \in \mathbb{R}, \; i = 2, \cdots, n-1 \right\}$ is a subspace of $\mathbb{R}^n$.

(iv) $W = \left\{ (x_1, x_2, x_3) : \; x_i \geqslant 0, \; i = 1, 2, 3 \right\}$ is not a subspace of $\mathbb{R}^3$. It is easily seen that W is not closed under scalar multiplication.

For example $\underset{\sim}{x} = (1, 1, 1) \in W$, but $-1 \underset{\sim}{x} = (-1, -1, -1) \notin W$.

__Definition__ Let X be a vector space and $\{x_1,\cdots,x_k\} \subset X$. A vector $x \in X$ is called a __linear combination__ of the vectors $x_1,\cdots,x_k$ if it can be expressed as

$$x = \alpha_1 x_1 + \cdots + \alpha_k x_k,$$

for some scalars $\alpha_1,\cdots,\alpha_k$.

(Throughout we will __fix__, a priori, the set of scalars as __either__ $\mathbb{R}$ or $\mathbb{C}$.)

__Proposition__ Let $x_1,\cdots,x_k$ be vectors in a vector space X. We have:

(i) The set W of all linear combinations of $x_1,\cdots,x_k$ is a subspace of X;

(ii) W is the smallest subspace of X containing $x_1,\cdots,x_k$.

__Proof__ (i) It is easy to see that W is closed under addition and scalar multiplication. Indeed, if $x = \alpha_1 x_1 + \cdots + \alpha_k x_k$, $y = \beta_1 x_1 + \cdots + \beta_k x_k$ are elements of W, then $x+y = (\alpha_1+\beta_1)x_1 + \cdots + (\alpha_k+\beta_k)x_k$, $\alpha x = \alpha\alpha_1 x_1 + \cdots + \alpha\alpha_k x_k$ are linear combinations of $x_1,\cdots,x_k$. Hence $x+y, \alpha x \in W$.

(ii) Let $\tilde{W}$ be any subspace of X that contains $x_1,\cdots,x_k$. Since $\tilde{W}$ is closed under addition and scalar multiplication, it must contain all linear combinations of $x_1,\cdots,x_k$. Hence $W \subset \tilde{W}$ and W is the __smallest__ subspace containing $x_1,\cdots,x_k$. ∎

Definition (i) Let $\{x_1, \cdots, x_k\} \subset X$, where X is a vector space and $W = \{\underbrace{\alpha_1 x_1 + \cdots + \alpha_k x_k}_{\text{linear combination of } x_j\text{'s}} : \alpha_1, \cdots, \alpha_k \text{ scalars}\}$.

the subspace W is called the linear span of $\{x_1, \cdots, x_k\}$, denoted by $\operatorname{span}\{x_1, \cdots, x_k\}$.

(ii) If $S \subset X$, where X is a vector space, Then the linear span of S, is the subset

$$\operatorname{span}(S) := \Big\{ \underbrace{\sum_{j=1}^{k} \alpha_j x_j}_{\substack{\text{finite linear} \\ \text{combinations} \\ \text{of elements of } S}} : \alpha_j\text{'s scalars, } j = 1, \cdots, k, \text{ sums ranging over finite number of terms} \Big\}.$$

By convention, we set $\operatorname{span}(\phi) = \{0\}$.

Proposition With X and S as in the above definition, $\operatorname{span}(S)$ is the smallest subspace of X containing S.

Proof If $S \subset W \subset X$, where W is a subspace of X, Then by the closure axioms $\operatorname{span}(S) \subset W$. If we show that $\operatorname{span}(S)$ is itself a subspace, the proof will completed. Indeed, this follows from the fact that span(S) is closed under addition and scalar multiplication. ∎

Remark Different sets may span the same sub space.
For example

(i) $\operatorname{span}\left(\{(1,0,0),\ (0,1,0)\}\right) = \operatorname{span}\left(\{(1,0,0),(0,1,0),(1,1,0)\}\right)$

$$= \mathbb{R}^2 \times \{0\} = \left\{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\right\}$$

subspace of $\mathbb{R}^3$.

(ii) $\operatorname{span}\left(\{(1,0,0),(0,1,0),(0,0,1)\}\right)$

$$= \operatorname{span}\left(\{(1,1,0),(1,0,1),(0,1,1)\}\right) = \mathbb{R}^3$$

Bases and dimension of vector spaces

Definition Let X be a vector space and $\{x_1, \cdots, x_n\} \subset X$.
Consider the vector equation

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \theta,$$

where $\alpha_1, \cdots, \alpha_n$ are scalars. This equation always admits the trivial solution $\alpha_1 = \cdots = \alpha_n = 0$. In addition, if this equation has a nontrivial solution $\alpha_1, \cdots, \alpha_n$ (not all zero), then the set $\{x_1, \cdots, x_n\}$ is said to be linearly dependent (L.D.). Otherwise, this set is said to be linearly independent (L.I.). Put simply, the set $\{x_1, \cdots, x_n\}$ is said to be L.I. if $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0 \Rightarrow \alpha_1 = \cdots = \alpha_n = 0$ and it is said to be L.D. otherwise.

Definition Let X be a vector space. A subset $S \subset X$ is said to be linearly independent (L.I.) if every finite subset of S is linearly independent. Otherwise, it is said to be linearly dependent (L.D.)

Put differently, S is L.D. if there exist elements (distinct) $x_1, \ldots, x_n \in S$ and scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, not all zero, such that

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0.$$

<u>Convention</u> Empty set ϕ is linearly independent

and $\operatorname{span}(\phi) = \{0\}$.

Examples (i) If a set S contains the zero vector 0, then S is linearly dependent since $1 \cdot 0 = 0$.

(ii) Put $\underset{\sim}{e}_i = (0, \ldots, 1, 0, \ldots 0) \in \mathbb{R}^n$, where 1 occurs in the ith component, $i = 1, 2, \ldots, n$. Let

$S = \{\underset{\sim}{e}_1, \ldots, \underset{\sim}{e}_n\}$. Then S is L.I.; Indeed

$$\alpha_1 \underset{\sim}{e}_1 + \alpha_2 \underset{\sim}{e}_2 + \cdots + \alpha_n \underset{\sim}{e}_n = 0$$

$$\Rightarrow (\alpha_1, \alpha_2, \ldots, \alpha_n) = (0, \ldots, 0) \Rightarrow \alpha_i = 0, \; i = 1, \ldots, n$$

Hence, S is L.I.

Definition A finite subset S of a vector space X is called a basis of X if S is linearly independent and spans X, i.e., span $(S) = X$. The space X is called finite dimensional if it has a finite basis. Otherwise, X is called infinite dimensional.

It is a fundamental fact that if X is a finite dimensional vector space, then every basis for X has the same number of elements. If a vector space X has a basis of n elements, then we write $\dim X = n$, where $\dim X$ denotes the dimension of X. By convention, If $X = \{0\}$, trivial space then we write $\dim X = 0$

Lemma Let $S = \{x_1, \dots, x_k\}$ be a linearly ~~an~~ independent set consisting of k elements in a vector space X. Then every set of $k+1$ element in span(S) is linearly dependent

The proof of this lemma is obtained by inducting on k. This is left to the reader.

Theorem Any two bases of a finite dimensional vector space X have the same number of elements.

Proof. Suppose S and T are bases of a finite dimensional vector space X. Let $|S| < |T|$. Here $|S|$ denotes the number of elements of S (cardinality of S).

Since $T \subset \operatorname{span}(S) = X$, by the previous lemma T is linearly dependent. This is a contradiction. ∎

<u>Examples</u> (i) Let $X = \mathbb{R}^n$ and let

$\underset{\sim}{e}_i = (0, \cdots, 1, \cdots 0)$, $i = 1, \cdots, n$. Then
↳ i^{th} place

$S = \{\underset{\sim}{e}_1, \cdots, \underset{\sim}{e}_n\}$ is a basis of X

We have seen earlier that S is L.I. Also since each vector $\underset{\sim}{x} = (x_1, \cdots, x_n)$ in $\mathbb{R}^n$ is expressible as

$$\underset{\sim}{x} = x_1 \underset{\sim}{e}_1 + \cdots + x_n \underset{\sim}{e}_n = \sum_{j=1}^{n} x_j \underset{\sim}{e}_j ,$$

$\operatorname{span}(S) = \mathbb{R}^n$. Thus S is a basis of $\mathbb{R}^n$ and we conclude that $\dim(\mathbb{R}^n) = n$.

(ii) Let $X = \mathbb{C}^n$. If we regard scalars as complex numbers, then as a vector space over $\mathbb{C}$ $S = \{ e_1, \dots, e_n \}$ is a basis for $\mathbb{C}^n$, where

$e_j = (0, \cdots, 1, \cdots 0)$. This follows exactly as in $\mathbb{R}^n$
$\quad$ └ j^{th} place

However, if we regard scalars as real numbers, then ~~the~~ a basis of $\mathbb{C}^n$ consists of $2n$ elements

$\{ e_1, ie_1, \dots\dots e_n, ie_n \}$

where $ie_j = (0, \cdots i, \cdots \cdot 0)$
$\quad$ ⟶ j^{th} place

We conclude that $\dim(\mathbb{C}^n) = n$ for complex scalars and $\dim(\mathbb{C}^n) = 2n$ for real scalars.

(iii)

Let $X = \left\{ \underset{\sim}{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 : \quad x_1 + x_2 - x_3 = 0 \right\}$

It is easily checked that X is a subspace of $\mathbb{R}^3$. What is a basis for X and what is its dimension? We can write the elements of X as

$$\underset{\sim}{x} = (x_1, x_2, x_1 + x_2) = x_1 (1, 0, 1) + x_2 (0, 1, 1)$$

which shows that $\qquad x_1, x_2 \in X$

$$X = \operatorname{span}\left(\{ (1,0,1), (0,1,1) \} \right)$$

It is also easily checked that $S = \{ (1,0,1), (0,1,1) \}$ is L.I. Hence it is a basis for X and $\dim X = 2$.

Topic 13 : Matrix Operations

Definition :

- Let m, n be positive integers.

  An $m \times n$ **matrix** M is a collection of mn numbers arranged in a rectangular array

$$M = \begin{array}{cc} \begin{bmatrix} a_{11} \dots & \vdots & \dots a_{1n} \\ \vdots & \vdots & \vdots \\ \dots & a_{ij} & \dots \\ \vdots & \vdots & \vdots \\ a_{m1} \dots & \vdots & \dots a_{mn} \end{bmatrix} & \text{ith row} \\ \text{jth column} & \end{array}$$

The entry in the ith row and jth column is denoted by a_{ij}.

We also write $M = (a_{ij})$,

where $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.

$$M = (C_1, C_2, \cdots C_n) = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix}$$

$\mathbb{R}$

$\mathbb{C}$

$$C_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

$$R_i = [a_{i1} \cdots a_{in}]$$

- A **row vector** is a $1 \times n$ matrix denoted by $M = [a_1 a_2 \ldots a_n]$ or $M = (a_1 a_2 \ldots a_n)$.

  A **column vector** is an $m \times 1$ matrix denoted it by

$$M = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Operations On Matrices

(1) **Addition.** Let $A = (a_{ij})$ and $B = (b_{ij})$ be two matrices of size $m \times n$.

Then $A + B$ is defined as the matrix (s_{ij}) where

$$s_{ij} = a_{ij} + b_{ij} \quad \forall \; i, j.$$

(2) **Scalar Multiplication.** If c is a scalar and $A = (a_{ij})$ then cA denotes the matrix whose entries are ca_{ij} for all i,j . The matrix cA is called a scalar multiple of A.

$\mathbb{K}^{m \times n}$ $\quad \mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$

$\mathcal{M}_{m \times n}$

(3) **Matrix Multiplication.** Let $A = (a_{ij})$ be an $m \times n$ matrix and $B = (b_{ij})$ be a $p \times q$ matrix.

Then AB is defined if $n = p$, i.e., the number of columns of $A =$ the number of rows of B.

In this case $AB = (c_{ij})$ is an $m \times q$ matrix and

$$c_{ij} = R_i \cdot \mathrm{C}_j = \sum_{k=1}^{n} a_{ik} b_{kj}$$

where R_i is the ith row vector of A and C_j is the jth column vector of B

$$\begin{array}{cc} & \\ i & \boxed{a_{i1} \cdots a_{in}} \end{array} \begin{array}{c} j \\ \begin{array}{|c|}\hline b_{1j} \\ \vdots \\ \vdots \\ b_{nj} \\ \hline \end{array} \end{array} = \begin{array}{|ccc|}\hline & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \\ \hline \end{array} = \text{A B}$$

$R_i \cdot C_j$

Consider the system of linear equations :

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

We can write this in matrix notation as

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

We write the above matrix equation briefly as :

$$AX = B \qquad \underset{\sim}{A}\,\underset{\sim}{x} = \underset{\sim}{b}$$

where $A = (a_{ij})$ is called the **coefficient matrix** and X and B are the column vectors. For example :

$$\begin{aligned} x_2 + 3x_3 &= 0 \\ x_1 + 2x_3 &= 1 \\ x_1 + x_2 &= 2 \end{aligned} \quad \text{can be written as} \quad \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 2 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Properties Of Matrix Operations

(1) $A(B+C) = AB + AC$.

(2) $(P+Q)R = PR + QR$.

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

(4) $c(AB) = (cA)B = A(cB)$.

The above laws hold whenever the sizes of matrices involved are suitable for the operations. Matrix **multiplication** is not commutative. For example :

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

But

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Definitions :

1. A matrix whose all entries are zero is called the **zero matrix.**
2. The **entries** a_{ii} of a square matrix $A = (a_{ij})$ are called **diagonal entries**.
3. If the only nonzero entries are diagonal entries in a matrix then A is called a **diagonal matrix.**
4. An $n \times n$ diagonal matrix whose diagonal entries are 1 is called the $n \times n$ **identity matrix**. It is denoted by I_n.

Inverse Of A Matrix

Definition :

Let A be an $n \times n$ matrix.
If there is an $n \times n$ matrix B such that $AB = BA = I_n$ then we say A is **invertible** and B is the inverse of A.
The inverse of A is denoted by A^{-1}. $B = A^{-1}$

$$\mathrm{diag}(a_{11}, a_{22}, \cdots a_{nn}) = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$$

Remarks :

(1) Inverse of a matrix is uniquely determined. Indeed, if B and C are inverses of A then

$$B = BI = B(AC) = (BA)C = IC = C.$$

(2) If A and B are invertible $n \times n$ matrices. Then AB is also invertible. Indeed,

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I.$$

Similarly $(AB)(B^{-1}A^{-1}) = I$. $(AB)^{-1} = B^{-1}A^{-1}$

(3) We will see later that if there exists an $n \times n$ matrix B for an $n \times n$ matrix A such that $AB = I$ or $BA = I$ then A is invertible. This fact fails for non-square matrices. For example

$$[12]\begin{bmatrix}1\\0\end{bmatrix} = [1] = I_1, \text{ but } \begin{bmatrix}1\\0\end{bmatrix}[12] = \begin{bmatrix}1 & 2\\0 & 0\end{bmatrix} \neq I_2.$$

(4) Inverse of a square matrix need not exist.

For example, let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. If $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is an inverse of A then

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \neq I_2$$

for any a, b, c, d.

Transpose Of A Matrix

Definition :

Let $A = (a_{ij})$ an $m \times n$ matrix. Then the **transpose** of A, denoted by A^t, is the matrix (b_{ij}) such that $b_{ij} = a_{ji}$ for all i, j. Thus rows of A become columns of A^t and columns of A become rows of A^t.

For example if

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \text{ then } A^t = \begin{bmatrix} 2 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

$A^* = [\overline{a_{ji}}]$

$\overline{A}^t$

$= \overline{A^t}$

$A^* = (\overline{a}_{ji})$ is called the conjugate transpose of A

$\overline{A} = [\overline{a_{ij}}]$

Proposition :

For matrices A and B of suitable sizes $(AB)^t = B^t A^t$.
For any invertible square matrix A, $(A^{-1})^t = (A^t)^{-1}$.

Proof : (i) For any matrix C, let C_{ij} denote its ij-entry.
Let $A = (a_{ij}), B = (b_{ij})$.

Then for all i, j;

$$\begin{aligned} ((AB)^t)_{ij} &= (AB)_{ji} \\ &= \sum a_{jk} b_{ki} \\ &= \sum (A^t)_{kj} (B^t)_{ik} \\ &= \sum (B^t)_{ik} (A^t)_{kj} \\ &= (B^t A^t)_{ij} \end{aligned}$$

(ii) $$AA^{-1} = I, (AA^{-1})^t = I.$$

$$\implies (A^{-1})^t A^t = I.$$

$$\implies (A^t)^{-1} = (A^{-1})^t.$$

∎

$(AB)^* = B^* A^*$

For any invertible square matrix A, $(A^{-1})^* = (A^*)^{-1}$

Definition : A square matrix A is called **symmetric** if $A = A^t$. It is called **skew-symetric** if $A^t = -A$.

Proposition :

(i) If an invertible matrix A is symmetric then so is A^{-1}.

(ii) Every square matrix A is a sum of symmetric and a skew - symmetric matrix in a unique way.

Proof : (i) $(A^{-1})^t = (A^t)^{-1} = A^{-1}$.

(ii) Since

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t),$$

every matrix is a sum of a symmetric and a skew-symmetric matrix. To see the uniqueness of expression, suppose that P, B are symmetric matrices and Q, C are skew-symmetric matrices and

$$A = B + C = P + Q.$$

Then $B - P = Q - C := D$. Then D is both symmetric and skew-symmetric. But this is possible only for the zero matrix. ∎

A is said to be Hermitian if $A^* = A$ (Self-adjoint)

A is said to be skew-Hermitian if $A^* = -A$