Q 01: Give Examples of
a] vector space b] finite dimensional vector space
c] Infinite dimensional vector space.

a] vector space:

Sol:
1] $X = R$ over R

2] X = Complex no's over complex numbers

3] X = set of complex no's over complex numbers

4] $C[0,1]$ on $[0,1]$

5] $X = P[0,1]$ over $[0,1]$

6] $X = P_n[0,1]$ $\forall$ degree $\leq n$

7] $X = \{0\}$: $K = R$

8] set of all $m \times n$ matrices usual matrix addition over K.

9] $Ax = 0$ [set of all homogeneous eqns]

b] Finite dimensional vector space

1] $\{0\}$
$\dim(\{0\}) = 0$

2] $P_n[0,1]$
$\dim[P_n[0,1]] = n+1$
Basis $= (1, x, x^2 \dots x^n)$

3. Standard Basis of R^n

$x = k_1(1, 0, 0 \dots 0) +$
$k_2(0, 1, 0 \dots 0) +$
$\dots$
$k_n(0, 0 \dots 1)$

c] Infinite dimension vector space.

1. set of continous functions $f: R \to R$.

2. ~~dim~~ $C[0,1]$
$\dim[C[0,1]] = \infty$

3. $P[0,1]$ & $P[-1,1]$
$\dim[P[0,1]] = \infty$, $\dim[P[-1,1]] = \infty$
Basis $= \{1, x, x^2 \dots \infty\}$

4. set of all non convergent sequences.

5. All the differential fn's from R to it self

3. Consider the following matrix

$$X = \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix}$$

a] Find the rank of X. On the basis of rank of X, discuss the existence of the inverse of X.

Sol: $X = \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} \Rightarrow |X| = 6 - 24 = -18$; $\boxed{\text{Rank}(X) = 2}$

As $|X| \neq 0 \rightarrow$ non-singular matrix.

$A^{-1} = \dfrac{\text{adj}A}{|A|} \Rightarrow X^{-1} = \dfrac{\text{adj}\,X}{|X|}$ → non-singular matrix

X^{-1} Exists; $|X| \neq 0$

∴ For all non-singular matrix inverse exists

b] Find the condition number

$X = \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix}$ $\quad X^{-1} = \dfrac{1}{-18}\begin{bmatrix} 2 & -4 \\ -6 & 3 \end{bmatrix}$

$\boxed{\text{Condition number} = \|A\|\,\|A^{-1}\|}$

Infinity norm:

Condition number $= \|A\|_\infty \cdot \|A^{-1}\|_\infty$

$\|A\|_\infty \Rightarrow$ Row sum matrix.

$\|A\|_\infty = 8$; $\|A^{-1}\|_\infty = 1/2$

∴ Condition number $= \|A\|_\infty \cdot \|A^{-1}\|_\infty = 8 \times \frac{1}{2} = 4$.

Since $\|A\|_\infty$ ~~...~~

$$\max \begin{cases} \rightarrow \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} \rightarrow 3+4 = 7 \\ \rightarrow \quad\quad\quad\quad \rightarrow 6+2 = 8 \end{cases}$$

$\max(7, 8) = 7$.

One norm:

Condition number $= \|A\|_1 \, \|A^{-1}\|_1$

$\|A\|_1 \Rightarrow$ column sum matrix.

$\|A\|_1 = \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix}$ → column sums 9, 6

$\|A^{-1}\|_1 = \begin{bmatrix} \frac{2}{-18} & \frac{-4}{-18} \\ \frac{-6}{-18} & \frac{3}{-18} \end{bmatrix}$ → 8/18, 7/18.

$\|A\|_1 = \max(9, 6) = 9$

$\|A\|_1 \cdot \|A^{-1}\|_1 = 9 \times \frac{8}{18} = 4$

$\boxed{\text{Condition number} = 4}$

(6) Give an Example of Monotonically increasing and decreasing function !

Sol: **Monotonic function:**

let f be a function defined on closed interval $[a, b]$. Then

① f is said to be monotonically increasing on $[a, b]$ if for $x_1, x_2 \in [a, b]$, $\boxed{x_1 > x_2}$

$\boxed{f(x_1) \geq f(x_2)}$

② f is said to be monotonically decreasing on $[a, b]$ if $x_1, x_2 \in [a, b]$ $\boxed{x_1 > x_2}$

$\boxed{f(x_1) \leq f(x_2)}$

③ f is said to be monotonic if f is either monotonically increasing (or) decreasing.

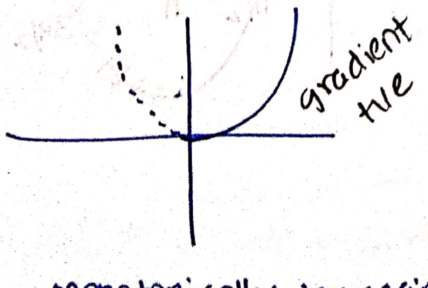
④. f is said to be strictly monotonically increasing on $[a, b]$ if for $x_1, x_2 \in [a, b]$ $\boxed{x_1 > x_2}$

$\boxed{f(x_1) > f(x_2)}$

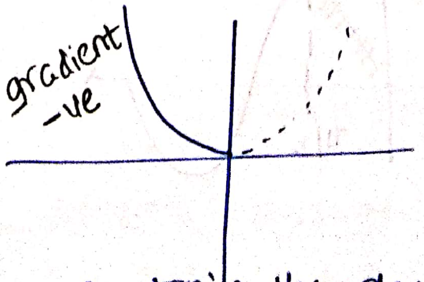
⑤ f is said to be strictly monotonically decreasing on $[a, b]$ if for $x_1, x_2 \in [a, b]$ $\boxed{x_1 > x_2}$

$\boxed{f(x_1) < f(x_2)}$

⑥ f is said to be strictly monotonic if f is either strictly monotonically increasing (or) strictly monotonically decreasing.



monotonically increasing

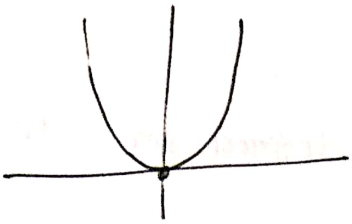


monotonically decreasing

Monotonic increasing :

1] $\boxed{y = x^2}$

function

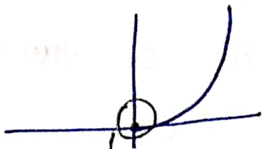


$x_1 = 4 > 0 \quad x_2 = 2 > 0$

$f(x_1) = 16 > f(x_2) = 4$

monotonically increasing

$y = x^2$; $\boxed{x > 0}$



At $x = 0 \Rightarrow$ stationary point

Neither increasing nor decreasing at $x = 0$

2] $\boxed{f(x) = \log x}$; $x > 0$

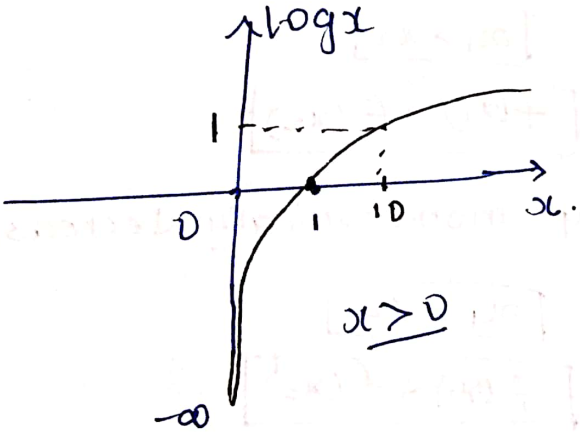
$x_1 = 1 > 0 \quad x_2 = 10 > 0$

$x_2 > x_1$

$f(x)_1 = \log(1) = 0$

$f(x_2) = \log(10) = 1$

$f(x_2) > f(x_1)$



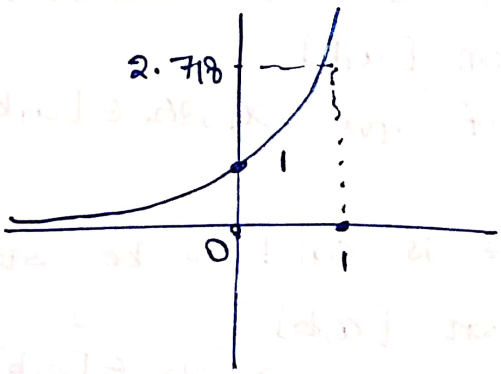
3] $\boxed{f(x) = e^x}$:

$x_1 = 0 \quad x_2 = 1$

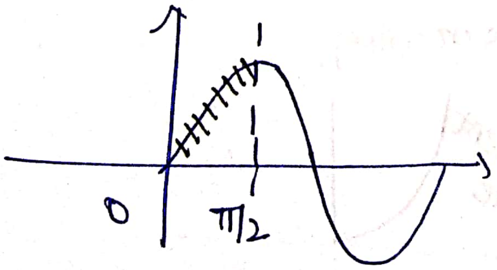
$x_2 > x_1$

$f(x_1) = e^0 = 1 \quad f(x_2) = e^1 = 2.718$

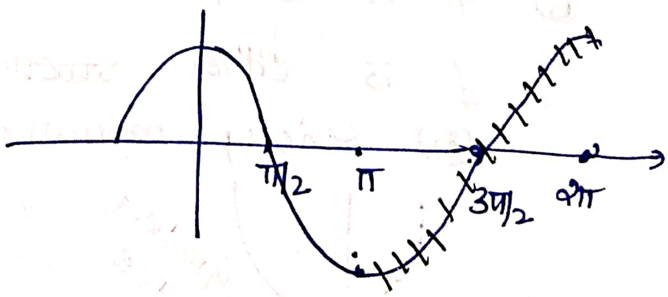
$f(x_2) > f(x_1)$



4) $\boxed{f(x) = \sin x \quad [0 \text{ to } \pi/2]}$



5] $\boxed{f(x) = \cos x \quad [\pi \text{ to } 2\pi]}$



6] polynomial with possitive coefficients.

$P_n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n . x^n$: for $x > 0$

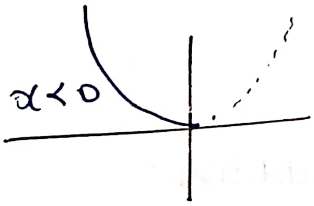
7] $f(x) = x . \ln(x)$ $\forall$ $x \in (1/e, \infty)$

Monotonic decreasing fns:

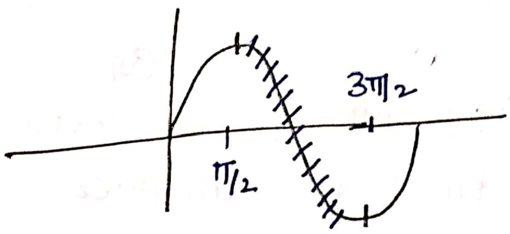
1) $\boxed{y = x^2}$: $x < 0$

$x_1 = -2 > x_2 = -4$

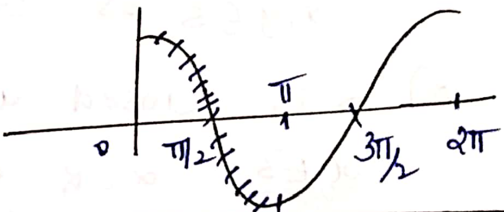
$f(x_1) = 4 < f(x_2) = 16$



2) $\boxed{f(x) = \sin x}$ $x \in [\pi/2, 3\pi/2]$



3) $f(x) = \cos x$ $x \in [0, \pi]$



4] $f(x) = x . \ln(x)$

$\forall x \in (0, 1/e)$

Q07) Find the projection of x onto $V = \text{span}\{u, v, w\}$

$x = (-1, 1, 1, 2)^T$ $u = (0, 2, -2, 1)^T$ $v = (1, 3, 0, 1)^T$ $w = (1, 1, 1, 1)^T$

Sol: The projection of y onto R(x).
p(y) is a vector in R(x) which is the smallest distance to y. That is

$$p(y) = \arg\min_{y' \in R(x)} \|y - y'\|$$

$$\|x - u\| = \sqrt{(-1)^2 + (-1)^2 + (3)^2 + (1)^2}$$

$$= \sqrt{1+1+9+1} = \sqrt{12}$$

$$\|x - v\| = \sqrt{(-2)^2 + (-2)^2 + (1)^2 + (1)^2}$$

$$= \sqrt{4+4+1+1} = \sqrt{10}$$

$$\|x - w\| = \sqrt{(-2)^2 + (0) + (0) + (1)^2}$$

$$= \sqrt{4+1} = \sqrt{5}$$

$\therefore$ projection of x onto V

$\boxed{w = (1, 1, 1, 1)^T}$

Q 08: Explain the term subspace of a vector space - Span of a set of vectors.

Sol:

(i) Subspace of a vector space:

→ If R^n is a vector space over R, then a subset S of R^n is said to be a subspace of vector space R^n : if S is also a vector space over field R.

(or)

→ If R^n is a vector space, then $S \subseteq R^n$ is said to be a subspace if

1) $0 \in S$

2) S is closed under vector addition:
$x, y \in S \Rightarrow x+y \in S$

3) S is closed under scalar multiplication
$x \in S \quad \alpha \in R \Rightarrow \alpha \cdot x \in S.$

(ii) Span of Set of vectors:

→ If $\boxed{S = \{v_1, v_2, \cdots v_k\}}$ is a set of vectors in a vector space V, the span of S is the $\boxed{\text{set of all linear combinations}}$ of the vectors in S.

$$\boxed{span(S) = \{c_1 v_1 + c_2 v_2 + \cdots + c_k v_k \mid \forall c_i \in R\}}$$

→ If Every vector in a given vector space can be written as a linear combination of vectors in a given set S, then S is called a Spanning set of vector space.

a] Range space of matrix Transformation $A: R^n \to R^m$ a subspace of R^m

Sol:. $y = wx$: $y \in R^m$, $w \in R^{n}$

let ~~see~~ $x.w = y$ $x ; w \longrightarrow y$

$R(x)$ [where $x: R^n \longrightarrow R^m$].

$y_1, y_2 \in R(x)$.

Since y_1 & y_2 are $R(x)$, they should have preimages.

$y_1 = xw' \in R^m$; $y_2 = xw'' \in R^m$ w' is preimage of y_1, w'' is preimage of y_2.

$\alpha y_1 + \beta y_2 = \alpha x w' + \beta x w''$ $w', w'' \in R^n$

$= x[\alpha w' + \beta w'']$

where $\alpha w' + \beta w'' = w \in R^n$

$\therefore \alpha y_1 + \beta y_2 = x \cdot w \in R^m$

$\therefore$ Range space of matrix Transformation $A: R^n \longrightarrow R^m$ a subspace of R^m

b) All vectors $(v_1, v_2, v_3)^T$ in R^3 with $v_1 - v_2 + 2v_3 = 0$ a subspace of R^3

let $(v_1, v_2, v_3)^T \in R^3$ & $(u_1, u_2, u_3)^T \in R^3$

$v_1 - v_2 + 2v_3 = 0$ $u_1 - u_2 + 2u_3 = 0$

$\alpha v + \beta u \to (\underbrace{\alpha v_1 + \beta u_1}_{z_1}, \underbrace{\alpha v_2 + \beta u_2}_{z_2}, \underbrace{\alpha v_3 + \beta u_3}_{z_3})$

$z_1 - z_2 + 2z_3 = \alpha v_1 + \beta u_1 - [\alpha v_2 + \beta u_2] + 2 \cdot (\alpha v_3 + \beta u_3)$

$= \alpha[\cancelto{0}{v_1 - v_2 + 2v_3}] + \beta[\cancelto{0}{u_1 - u_2 + 2u_3}] = 0$

$\therefore (v_1, v_2, v_3)^T$ forms a sub space of R^3 with $v_1 - v_2 + 2v_3 = 0$

c] All vector $(v_1, v_2)^T$ in R^2 $v_1 \geq v_2$ a subspace of R^2

sol. let $(v_1, v_2)^T$ & $(u_1, u_2)^T$
$v_1 \geq v_2$ $u_1 \geq u_2$.

$\alpha v + \beta u \Rightarrow (\alpha v_1 + \beta u_1, \alpha v_2 + \beta u_2)$

$v_1 - v_2 \geq 0$ $\alpha v_1 + \beta u_1 \geq \alpha v_2 + \beta u_2$

$u_1 - u_2 \geq 0$ $\alpha (v_1 - v_2) \geq \beta (u_2 - u_1)$

$\alpha \underbrace{(v_1 - v_2)}_{v_1 - v_2 \geq 0} \geq -\beta \underbrace{(u_1 - u_2)}_{u_1 - u_2 \geq 0} \Rightarrow$ satisfied.

$\therefore$ $(v_1, v_2)^T$ in R^2 $v_1 \geq v_2$ is a subspace of R^2.

d] $v = (v_1, v_2, v_3)^T \in R^3$ with v_1, v_2, v_3 positive

let $(v_1, v_2, v_3)^T$ & $(u_1, u_2, u_3)^T$
$v_1 > 0$ $v_2 > 0$ $v_3 > 0$ $u_1 > 0$ $u_2 > 0$ $u_3 > 0$

$(\alpha v + \beta u) \Rightarrow (\alpha v_1 + \beta u_1, \alpha v_2 + \beta u_2, \alpha v_3 + \beta u_3)$

$\alpha v_1 + \beta u_1 > 0$ $\alpha v_2 + \beta u_2 > 0$ $\alpha v_3 + \beta u_3 > 0$

B) c $v_1 > 0$, $u_1 > 0$ B) c $v_2 > 0$, $u_2 > 0$ B) c $v_3 > 0$, $u_3 > 0$

$\therefore$ $v = (v_1, v_2, v_3)^T \in R^3$ with v_1, v_2, v_3 positive is subspace of R^3.

e] $v = [v_1, v_2, v_3]^T \in R^3$ with positive $v_1 - v_2 + v_3 = K$

let $[v_1, v_2, v_3]^T$ & $[u_1, u_2, u_3]^T$
$v_1 - v_2 + v_3 = K$ $u_1 - u_2 + u_3 = K$

$\alpha v + \beta u \Rightarrow (\alpha v_1 + \beta u_1, \alpha v_2 + \beta u_2, \alpha v_3 + \beta u_3)$

$\alpha v_1 + \beta u_1 - \alpha v_2 - \beta u_2 + \alpha v_3 + \beta u_3$

$= \alpha [v_1 - v_2 + v_3] + \beta [u_1 - u_2 + u_3] = \beta K + \alpha K$

$\alpha V + \beta u = K(\alpha+\beta) \neq K$.

$\therefore$ $V = (V_1, V_2, V_3)^T \in R^3$ with $V_1 - V_2 + V_3 = K$ where K is a constant is not a subspace of R^3.

Q 09: Consider $Xw = y$ where X is a 7×5 matrix

a] Is X onto justify your answer:

$$\underset{7\times1}{[y]} = \underset{7\times5}{[\;X\;]}\;\underset{5\times1}{[w]}$$

$N = 7$

$n+1 = 5$.

$\boxed{N > n+1}$ B/c no of equations are more than no of unknowns.

$\therefore$ therefore there exists some y's which doesn't have solution.

$\therefore$ X is not onto.

b) what is the maximum dimension the Range of X can have? Justify your answer.

Sol: given X is a 7×5

$$y = Xw \Rightarrow \underset{7\times1}{(y)} = \underset{7\times5}{[\;X\;]}\;\underset{5\times1}{[w]}$$

$X : R^5 \rightarrow R^7$ & X is not onto

$$y = w_0 \cdot X_0 + w_1 X_1 + w_2 X_2 + w_3 X_3 + w_4 X_4$$

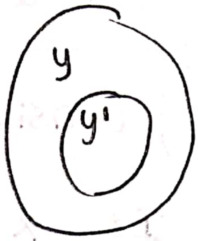
y can be expressed as unique combination of $S = \{X_0, X_1, X_2, X_3, X_4\}$.

$\therefore$ S forms the Basis

and $\boxed{\text{max dimension of } R(X) = 5}$

c) If it is possible to find w, what should be the rank of X? In this case is X one to one? Justify your ans?

Sol: As X is not onto.



→ we have to find y' i.e projection of y on to $R(X)$

such that $$\boxed{P(y) = \min_{y_i \in R(X)} \| y - y_i \|}$$

In this case matrix equation

$$Xw = P(y)$$

→ The preimage of $P(y)$ onto the range space of X is to be found out

→ given If it is possible to w

→ For getting a unique preimage in this case

$$X : \mathbb{R}^{n+1} \to \mathbb{R}^N$$

dimension of $R(X)$ should be $n+1$.

That is columns of X should be Linearly Independent

$\therefore$ $\boxed{X \text{ is one-one.}}$

$\boxed{\text{Rank}(X) = 5.}$

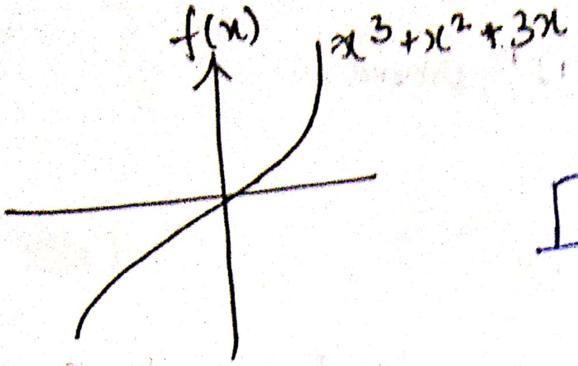
---

10. Find $\arg\min f(x)$ and $\min f(x)$ where

a) $f(x) = x^3 + x^2 + 3x$, $x \in \mathbb{R}$.

Sol: $$f'(x) = 3x^2 + 2x + 3 = 0$$

$$x = -1/3 \pm \frac{2\sqrt{2}}{2} i \Rightarrow$$ stationary point is complex



$\boxed{\text{min at } -\infty}$

$\boxed{\text{arg min at } -\infty}$

b] $f(x) = \sin x + \cos x \quad , x \in \mathbb{R}.$

$f'(x) = \cos x - \sin x = 0 \Rightarrow \quad \cos x = \sin x$

$\tan x = 1 \Rightarrow x = \pi/4.$

$x = [4n+1]\,\pi/4$

$\cancel{f''(x) = -\sin x - \cos x}$

$\Rightarrow \; -\sqrt{1^2+1^2} \leq \sin x + \cos x \leq \sqrt{1^2+1^2}$

$-\sqrt{2} \leq \sin x + \cos x \leq \sqrt{2}$

$\min F(x) = -\sqrt{2}$ for $\operatorname{argmin} F(x) = \boxed{4n+1} \cdot \pi/4.$

c] $f(x) = \exp(\|x\|^2) \quad , x \in \mathbb{R}^2.$

$f'(x) = \frac{d}{dx} \exp(\langle x, x \rangle) = e^{\|x\|^2} \cdot 2x = 0$

$x = 0$

$f''(x) = 2x\, e^{\|x\|^2} \cdot 2x + 2 \cdot e^{\|x\|^2}$

$f''(0) = 0 + 2 > 0$

$\therefore \; x = (0,0)$ is $\operatorname{argmin}(f(x))$

$f(x) = 1$ is min.

Q. 11 a] ∇f at the point $x = (-1, 1)^T$ where

$f(x) = \sin(\cos(||x||^2))$

$$||x||^2 = \langle x, x \rangle = x^T \cdot x$$

$$\frac{d}{dx} ||x||^2 = \frac{d}{dx} \langle x, x \rangle = \frac{d}{dx} (x^T \cdot x) = \frac{d}{dx} x^T \cdot x + \frac{d}{dx} x^T \cdot x$$

$$= 2x.$$

$$\nabla f(x) = \frac{d}{dx} \left[\sin\left[\cos\left[x^T \cdot x\right]\right]\right]$$

$$= \cos\left[\cos\left[x^T x\right]\right]\left[-\sin\left[x^T x\right]\right] \cdot \frac{d}{dx}\left[x^T x\right]$$

$$\nabla f(x) = \cos\left[\cos\left[x^T x\right]\right]\left[-\sin\left[x^T x\right]\right] \cdot 2x.$$

$$x^T x = [-1, 1] \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 1 + 1 = 2$$

$$x^T x = 2$$

$$= \underbrace{\cos\left[\cos[2]\right]\left[-\sin[2]\right]}_{-0.8316} \cdot 2 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$= -1.663 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow \boxed{\nabla f(x) = \begin{bmatrix} 1.663 \\ -1.663 \end{bmatrix}}$$

b] $\nabla_w f(w)$ where. $f(w) = \exp(-||xw - y||^2)$

$$J(w) = \frac{1}{2} ||xw - y||^2$$

$$\frac{d}{dw} J(w) = \nabla_w w = \frac{1}{2} \cdot \langle xw - y, xw - y \rangle$$

$$= \frac{1}{2} \left[\langle xw, xw \rangle - \langle xw, y \rangle - \langle y, xw \rangle + \langle y, y \rangle \right]$$

$$J = \frac{1}{2}\left[(xw)^T \cdot (xw) - (xw)^T y - y^T xw + y^T y \right]$$

$$J = \frac{1}{2}\left[w^T x^T x w - w^T x^T y - y^T x w + y^T y \right]$$

$$\nabla_w J = \frac{1}{2}\left[\frac{d}{dw} w^T (x^T x w) + \frac{d}{dw} (x^T x w)^T w - \frac{d}{dw} w^T (x^T y) - \frac{d}{dw} (x^T y)^T \cdot w + \frac{d}{dw} (y^T y) \right]$$

$$= \frac{1}{2}\left[x^T x w + x^T x w - x^T y - x^T y \right]$$

$$\nabla_w J = x^T x w - x^T y$$

$$f(w) = \exp(-\|xw - y\|^2) = \exp\left(-\underbrace{\frac{1}{2} \cdot \|xw - y\|^2}_{J} \cdot 2\right)$$

$$= \exp(-2 \cdot J) = ~~\text{exp}~~$$

$$\nabla_w f(w) = \exp(-2J) \times -2 \times \nabla(J).$$

$$= \exp(-2J) \times -2 \cdot [x^T x w - x^T y]$$

~~$\nabla_w f(w)$~~

$$\boxed{\nabla_w f(w) = -2[f(w)]^2 \cdot \nabla_w J}$$

where $J = \frac{1}{2} \cdot \|xw - y\|^2$

---

function:

There should not be multiple outputs of y for single x.

ONE-ONE: ~~[illegible]~~ Element is set-B has a **single preimage** in set-A

Need not be Every element

ONTO: **Every Element** has **atleast one preimage** in A

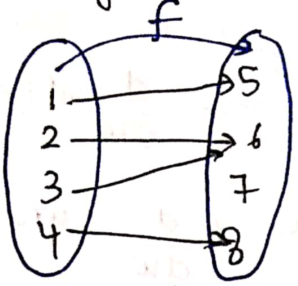
CODOMAIN = RANGE → ONTO

CONCEPTS

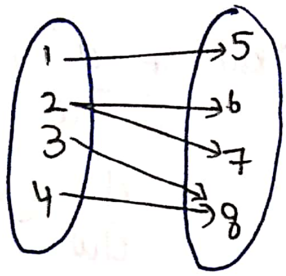
Function:

There should not be multiple outputs of y for single x. Then given Relation is a function

Ex:



fn ✓



Not a fn

As single x i.e 2 has multiple outputs in y i.e (6,7)

Note:

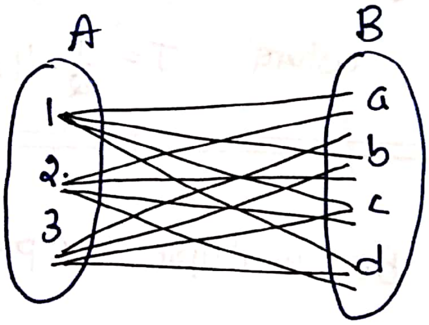
SETS → Cartesian product → Relation → Function

SET: ~~Red~~ Well defined collection of objects

Cartesian product: $A \times B$

$\{(a,b) : a \in A, b \in B\}$

All possible ordered pairs from A and B



$$A \times B = \begin{Bmatrix} (1,a), (1,b), (1,c), (1,d) \\ (2,a), (2,b), (2,c), (2,d) \\ (3,a), (3,b), (3,c), (3,d) \end{Bmatrix}$$

Relation: $R: A \to B$; $R \subseteq B$

Relation: $R: A \to B$

$R \subseteq B$

Ex: $A = \{1, 2, 3\}$ $B = \{a, b, c, d\}$

R : all ordered pairs begining with an odd no

$\{(1,a), (1,b), (1,c), (1,d)$

$(3,a), (3,b), (3,c), (3,d)\}$

∴ Relation happens to be a part of cartesian product

~~Domain~~: Ex: $A = \{1, 2, 3, 5\}$ $B = \{4, 6, 9\}$

set builder form → $R: \{(x,y):$ diff b/w x and y is odd $x \in A$ and $y \in B\}$

Sol:

Domain : set of all 1st elements in the ordered pair

Roaster form

$\{(1,4), (1,6), (2,9)$

$(3,4), (3,6), (5,4)$

$(5,6)\}$

Domain: $\{1, 2, 3, 5\}$

Range : set of all second elements in the order pair

Range: $\{4, 6, 9\}$

CO-Domain: codomain is always the second set

Codomain : B

Ex: $R: \{(x,y) : y = x + \frac{6}{x}$ where $\boxed{x, y \in N}$ & $\boxed{x < 6}\}$

$\{(1,7), (2,5), (3,5)\}$

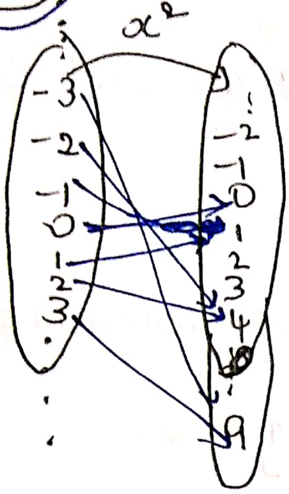
Domain : $\{1, 2, 3\}$

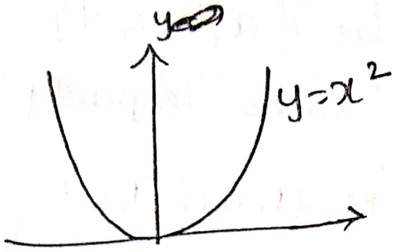
Range : $\{4, 5, 7\}$

Codomain : N.

Q 13

a] $f: R \to R$; $f(x) = x^2$





→ There are no multiple outputs of y for single x.

∴ $f(x) = x^2$ is a $\boxed{fn}$

→ $\boxed{\text{Domain} = R}$

→ $\boxed{\text{Range} = R^+}$ i.e $\{x^2 \in R, \forall x \in R\}$

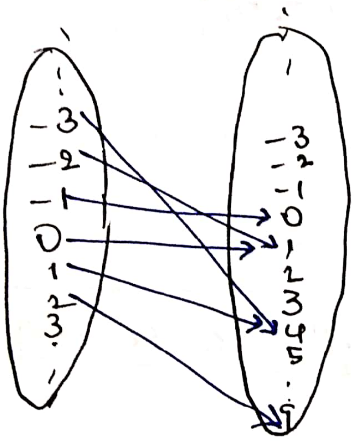
→ $\boxed{\text{Codomain} = R}$

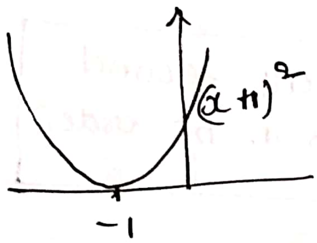
$\boxed{\text{Not one - one}}$ $\boxed{\text{Not - onto}}$

b] $f: R \to R$ $f(x) = (x+1)^2$

$\boxed{\text{Not one-one}}$

$\boxed{\text{Not onto}}$





→ There a no multiple outputs of y for single x.

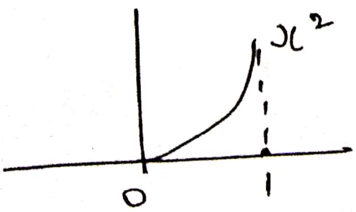
$f(x) = (x+1)^2$ is a $\boxed{\text{function}}$

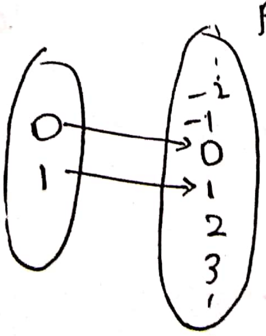
→ $\boxed{\text{Domain} = R}$; $\boxed{\text{Range} = R^+}$

$\{(x+1)^2 \in R, \forall x \in R\}$

→ $\boxed{\text{Codomain} = R}$

c] $f: [0,1] \to R$ $f(x) = x^2$ $\boxed{\text{one-one}}$ $\boxed{\text{Not onto}}$





→ there are No multiple o/p's for single x

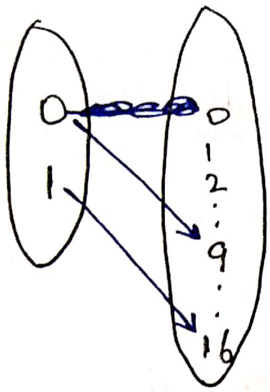
$f: [0,1) \to R$ is a function

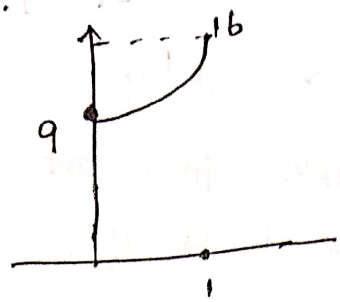
Domain : $[0,1]$

Range : $[0,1)$

codomain : R

d] $f:[0,1] \to R$ where $f(x) = (x+3)^2$.





→ There are no multiple o/p's for single x

$\therefore f(x) = (x+3)^2$ is a function

→ Domain = $[0,1]$

→ Range = $\{9, 16\}$

→ codomain = R

→ $\boxed{\text{One - one}}$

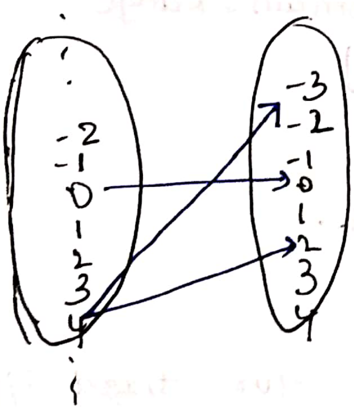
→ $\boxed{\text{Not onto}}$

e] $f: R \to R$: $f(x) = \sqrt{x}$.

→ There are multiple y's for single x.

$\therefore f(x) = \sqrt{x}$ is not a function.

~~Domain~~



f] $f: R^2 \to R$ $f(x) = \langle x, a \rangle$, $a \in R^2$.

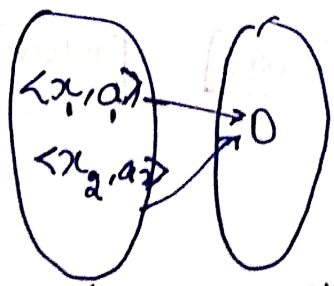
Sol: $\langle x, a \rangle = x^T a$

As there are no multiple outputs for a fixed i/p. It is a function.

Domain = R^2

Range = R

codomain = R



for different combinations of $\langle x, a \rangle$ will also lead to a single o/p. So $\boxed{\text{not One-one}}$

→ $\langle x, a \rangle$ Every Element in R is mapped $\boxed{\text{ONTO}}$

g] $f: R^2 \times R^2 \longrightarrow R$ where $f(x,y) = \langle x, y \rangle$:

As $f(x,y) = \langle x,y \rangle = x^T y$

There are no multiple o/p's for single i/p.

$\therefore$ $f(x,y) = \langle x,y \rangle$ is a fn.

Domain : $R^2 \times R^2$

Range : R

codomain : R

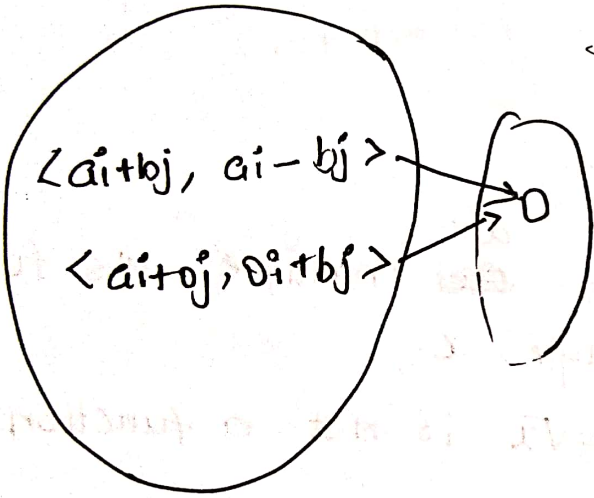
Ex:

$x = ai + bj$ $y = bi - aj$

$\langle x,y \rangle = ab - ab = 0$

$x = ai + 0.bj$ $y = 0.i + bj$

$\langle x,y \rangle = 0$



$\therefore$ $\boxed{\text{Not one-one.}}$

As codomain = Range

$\boxed{\text{ONTO}}$

h) $f: R^3 \times R$ where $f(x) = \|x\|^2$.

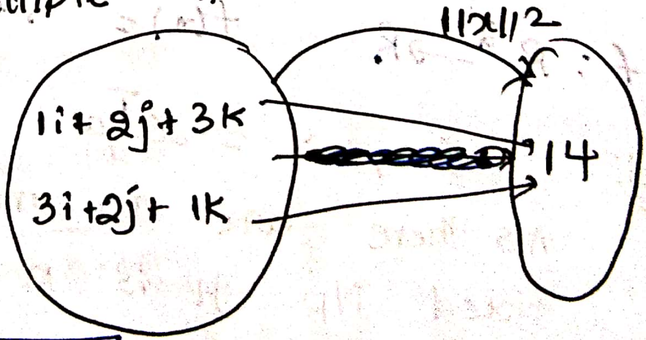
As $f(x) = \|x\|^2$

There are no multiple o/p's for fixed i/p.

Domain : R^3

Range : R^+

codomain : R



$\boxed{\text{Not one}}$, $\boxed{\text{Not onto}} \rightarrow$ codomain $\neq$ Range.

i] $f: R^4 \times R^4 \rightarrow R$, $f(x,y) = \exp(\|x-y\|^2)$

Sol: $f(x,y) = \exp(\|x-y\|^2) \rightarrow$ function.

There are no mupltiple o/p's for single i/p.

$\therefore$ Domain : $R^4 \times R^4$

Range : N $\Rightarrow$ codomain $\neq$ Range

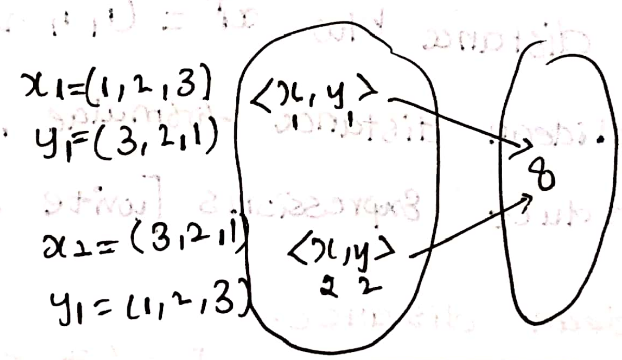
Codomain : R

$\boxed{\text{NOT ONTO}}$

Ex: $x_1 = (1,2,3)$ $y_1 = (3,2,1)$ $x_2 = (3,2,1)$ $y_1 = (1,2,3)$

$\|x_1 - y_1\|^2 = (2)^2 + (0)^2 + (-2)^2 = 8$

$\|x_2 - y_2\|^2 = (-2)^2 + (0)^2 + (+2)^2 = 8$



$\boxed{\text{NOT ONE - ONE}}$

ii] $f: R^2 \times R^2 \rightarrow M(2,R)$, $f(x,y) = xy^T$

where $M(2,R)$ is set of all 2×2 Real matrices

Sol: $f(x,y) = x \cdot y^T \rightarrow$ function.

There are no multiple o/p's for a single i/p

Domain = $(R^2 \times R^2)$

Range = $M(2,R)$

codomain = $M(2,R)$

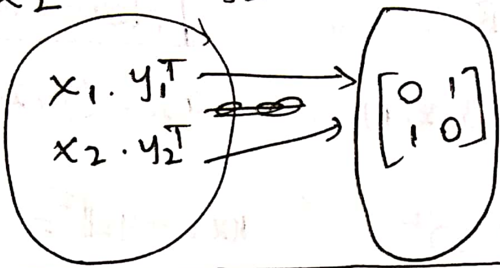
Range = codomain

$\boxed{\text{ONTO function}}$

Ex:

$$\underset{x_1}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \underset{y_1^T}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\underset{x_2}{\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}} \underset{y_2^T}{\begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



NOT ONE-ONE

Q14: Find distance b/w $x_1^T = (1, 1, -2)$ and $x_2^T = (7, -3, 2)$ using Euclidean distance formulae, norm and innper product expressions [write all relevant steps]

Sol: Euclidean distance.

$x_1^T = (\underset{x_1}{1}, \underset{x_2}{1}, \underset{x_3}{-2})$ $\qquad$ $x_2^T = (\underset{y_1}{7}, \underset{y_2}{-3}, \underset{y_3}{2})$

$$\text{Euclidean distance} = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2}$$

$$= \sqrt{(6)^2 + (-4)^2 + (4)^2}$$

$$= \sqrt{36 + 16 + 16} = \sqrt{68}$$

$$\text{norm} = \|x_2 - x_1\|$$

$$x_2^T - x_1^T = (6, -4, 4)$$

$$\|x_2^T - x_1^T\| = \sqrt{6^2 + (-4)^2 + (4)^2} = \sqrt{68}$$

Inner product $= \langle x_1, x_2 \rangle = x_1^T x_2$

$= [1, 1, -2] \cdot \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}$

$= 7 - 3 - 4 = 0$

$\langle x_1, x_2 \rangle = x_1^T \cdot x_2$

$\|x_1\|_2 = \sqrt{1^2 + 1^2 + (-2)^2} = \sqrt{6}$

$\|x_2\|_2 = \sqrt{(7)^2 + (-3)^2 + (2)^2} = \sqrt{62}$

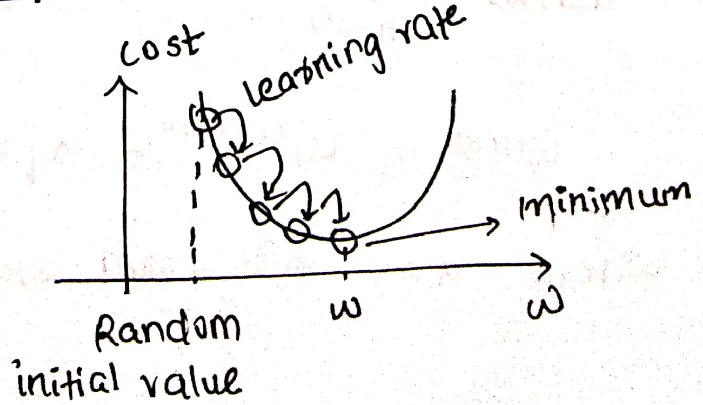
Q15: write a short notes on iterative techniques and gradient optimization approach.

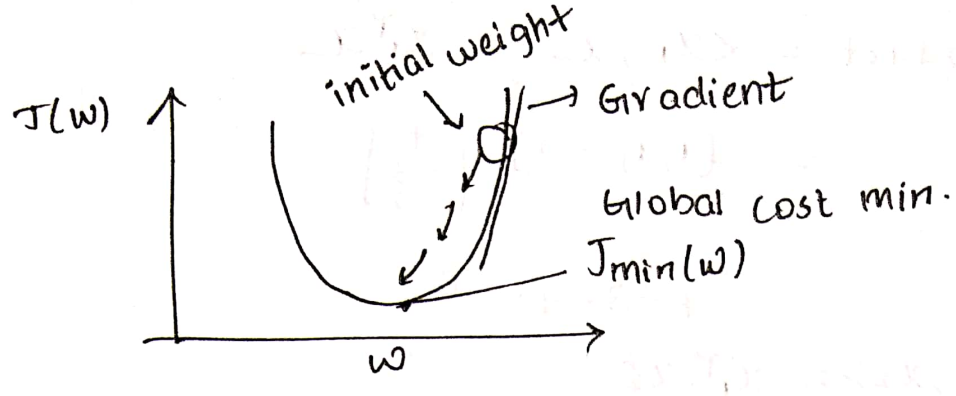
→ If the size of X is very large. Then it becomes computationally expensive to find X in direct method. so to avoid this we use iterative techniques to approach towards the solution & to find the parameter w.

→ Initially By assuming & calculating the error function by adjusting the hyper parameter or learning rate. We update the value of w in Each iteration to reduce error.

Iterative Techniques:

Gradient descent:





→ The gradient vector can be interpreted as the "direction and the rate of fastest increase". If the gradient of the function is non-zero at a point p, the direction of the gradient is in the direction in which function increases most quickly from p. and the magnitude of the gradient is the rate of increase in that direction.

→ The size of steps take to reach minimum or bottom is learning rate.

<u>Gradient</u> <u>descent</u>: If a real valued function F(x) is defined and differentiable in a neighbourhood of point a, then F(x) decreases fastest if one goes from a in the direction of negative gradient of F at a, $\nabla F(a)$

$$\nabla J(w) = \lim_{h \to 0} \frac{J(w+h) - J(w)}{h}$$

$$w_{new} = w_{current} - \alpha [\nabla J(w)]_{w_{current}}$$

where $\alpha > 0$ is called step length.

Updation of w:

For applying Gradient descent, consider the following steps.

→ Choose the initial value $w = (w_0, w_1 \ldots w_n)^T \in R^{n+1}$.

Then repeatedly perform the update.

$$w := w - \alpha \cdot \nabla J$$

J is a function of $w_0, w_1 \ldots w_n$.

$$\nabla J = \left(\frac{\partial J}{\partial w_0}, \frac{\partial J}{\partial w_1}, \ldots \frac{\partial J}{\partial w_n} \right)^T$$

$$(w_0, w_1, w_2 \ldots w_n)^T := (w_0, w_1 \ldots w_n)^T - \alpha \cdot \left[\frac{\partial J}{\partial w_0}, \frac{\partial J}{\partial w_1} \ldots \frac{\partial J}{\partial w_n} \right]^T$$

Batch Gradient Descent:

→ For updating the parameter, the algorithm looks at every data point in the training set at every step and hence it is called Batch Gradient descent.

→ In general, gradient descent doesnot guarentee a global minimum. since J is a convex quadratic function, the algorithm converges to global minimum [assuming α is not too large]

Algorithm 1: Updation of w: Gradient Descent.

Initialize w

Iterate until convergence {

$$w := w + \alpha \cdot \sum_{i=1}^{N} (y_i - f(x_i)) \cdot x_i$$

}

Algorithm 2: updation of w: Gradient descent:

Initialize w.

Iterate until convergence {

$w_j := w_j + \alpha \sum_{i=1}^{N} (y_i - f(x_i)) x_{ij}$, $j = 0, 1, \ldots n$

}

Stochastic Gradient descent:

→ The online version of gradient descent called stochastic gradient descent.

→ In contrast to batch gradient descent technique, stochastic gradient process only one training point at each step. Hence when N becomes large, that is for large data sets, stochastic gradient descent is more computationally efficient than batch gradient descent.

Algorithm 3: updating w using stochastic Gradient descent.

Iterate until convergence {

for $i = 1, 2 \ldots N$ {

$w := w + \alpha (y_i - f(x_i)) x_i$

}

}

Algorithm 4: updating w using stochastic Gradient descent.

Iterate until convergence {

for $i = 1, 2 \ldots N$ {

$w_j := w_j + \alpha (y_i - f(x_i)) x_{ij}$, $j = 0, 1, 2 \ldots n$

}

}

Q2: (i) $\{(1,-1),(2,11),(4,19)\} = \{x_i, y_i\}$

(ii) $x_1 = (2,3,2)^T, \; y_1 = 7$

$x_2 = (4,-5,5)^T, \; y_2 = 28$

$x_3 = (-3,7,-2)^T, \; y_3 = -19.$

a) write $f(x_i)$, $i = 1, 2, \dots N$

$$\boxed{\begin{aligned} f(x_i) &= w_0 \cdot x_{i0} + w_1 \cdot x_{i1} + w_2 \cdot x_{i2} + \dots + w_n \cdot x_{in} = w^T \cdot x_i \\ \text{where } x_i &= (1, x_{i1}, x_{i2}, \dots x_{in})^T \in R^{n+1} \\ w &= (w_0, w_1, w_2 \dots w_n)^T \in R^{n+1} \end{aligned}}$$

Here

$w_0, w_1, w_2 \dots w_n$ are the unknown parameters

(i) $f(x_i) = w^T \cdot x_i$

$w_0 + w_1(+1) = -1$

$w_0 + w_1(2) = 11$

$w_0 + w_1(4) = 19$

(ii) $f(x_i) = w^T \cdot x_i$

$w_0 + w_1(2) + w_2(3) + w_3(2) = 7$

$w_0 + w_1(4) + w_2(-5) + w_3(5) = 28$

$w_0 + w_1(-3) + w_2(7) + w_3(-2) = -19$

b) Express data in $xw = y$

(i)
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} -1 \\ 11 \\ 19 \end{bmatrix}$$

(ii)
$$\begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & 4 & -5 & 5 \\ 1 & -3 & 7 & -2 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 28 \\ -19 \end{bmatrix}$$

c) Set of vectors that spans Range of X.

(i)
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \Rightarrow \begin{matrix} R_2 = R_2 - R_1 \\ R_3 = R_3 - R_1 \end{matrix} \Rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 3 \end{bmatrix} \Rightarrow R_3 = R_3 - 3R_2$$

$$\Rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \boxed{\dim = 2}$$

$\boxed{\{(1,1),(1,2),(1,4)\} \text{ spans range of } X}$

$\therefore$ vectors $\{ \cancel{(1,1)}\, \cancel{(1,2)} \}$

(ii) $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & 4 & -5 & 5 \\ 1 & -3 & 7 & -2 \end{bmatrix} \Rightarrow \begin{matrix} R_2 = R_2 - R_1 \\ R_3 = R_3 - R_1 \end{matrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 2 & -8 & 3 \\ 0 & -5 & 4 & -4 \end{bmatrix}$

$R_3 = 2R_3 + 5R_2 \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 2 & -8 & 3 \\ 0 & 0 & -32 & 7 \end{bmatrix}$ $\boxed{\dim = 3.}$

As number of unknowns 4 (w_0, w_1, w_2, w_3)

But number of non-zero rows in Echelon form = 3

$\therefore$ The set of vectors are L.D

$\therefore$ ~~$\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ -8 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ -2 \end{pmatrix} \right\}$ doesn't span~~

Range of X.

$\Rightarrow \{ (1,2,3,2), (1,4,-5,5), (1,-3,7,-2) \}$ doesn't span range of X.

$\Rightarrow \{ (1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1) \}$ spans Range(x) $\in R^{n+1} = R^4$.

d) Find the dimension of Range (x)

(i) $\boxed{\dim = 2}$

No of non-zero Rows in Echelon form = 2

(ii) $\boxed{\dim = 3}$

No of non-zero Rows in Echelon form = 3

e] (i) $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} -1 \\ 11 \\ 19 \end{bmatrix}$

$X \quad w \quad = \quad y$ $\qquad N = 3$

$3\times 2 \quad 2\times 1 \quad 3\times 1$ $\qquad n+1 = 2$

$N\times(n+1) \quad (n+1)\times 1 \quad N\times 1$

$\therefore$ $N > (n+1) \Rightarrow$ over determined system

↳ Apply Least square Approximation

ie Right pseudo inverse

To find w.

$$\boxed{w = (X^T X)^{-1} X^T y}$$ → Right pseudo inverse By least square Approximation

$$X^T X = \underset{2\times 3}{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}} = \underset{2\times 2}{\begin{bmatrix} 3 & 7 \\ 7 & 21 \end{bmatrix}}$$

$$(X^T X)^{-1} = \frac{1}{14} \cdot \begin{bmatrix} 21 & -7 \\ -7 & 3 \end{bmatrix}$$

$$(X^T X)^{-1} \cdot X^T = \frac{1}{14} \cdot \begin{bmatrix} 21 & -7 \\ -7 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}$$

$$= \frac{1}{14} \begin{bmatrix} 14 & 7 & -7 \\ -4 & -1 & 5 \end{bmatrix}$$

$$(X^T X)^{-1} \cdot X^T \cdot y = \frac{1}{14} \cdot \begin{bmatrix} 14 & 7 & -7 \\ -4 & -1 & 5 \end{bmatrix} \begin{bmatrix} -1 \\ 11 \\ 19 \end{bmatrix}$$

$$= \frac{1}{14} \cdot \begin{bmatrix} -70 \\ 88 \end{bmatrix} = \begin{bmatrix} -5 \\ 6.28 \end{bmatrix}$$

$$\boxed{w = \begin{bmatrix} -5 \\ 6.28 \end{bmatrix}} \quad \boxed{w_0 = -5} \quad \boxed{w_1 = 6.28}$$

Hyperplane:

Regression:

$$\boxed{y_i = w^T \cdot x_i + w_0}$$

Classification:

Decission boundary

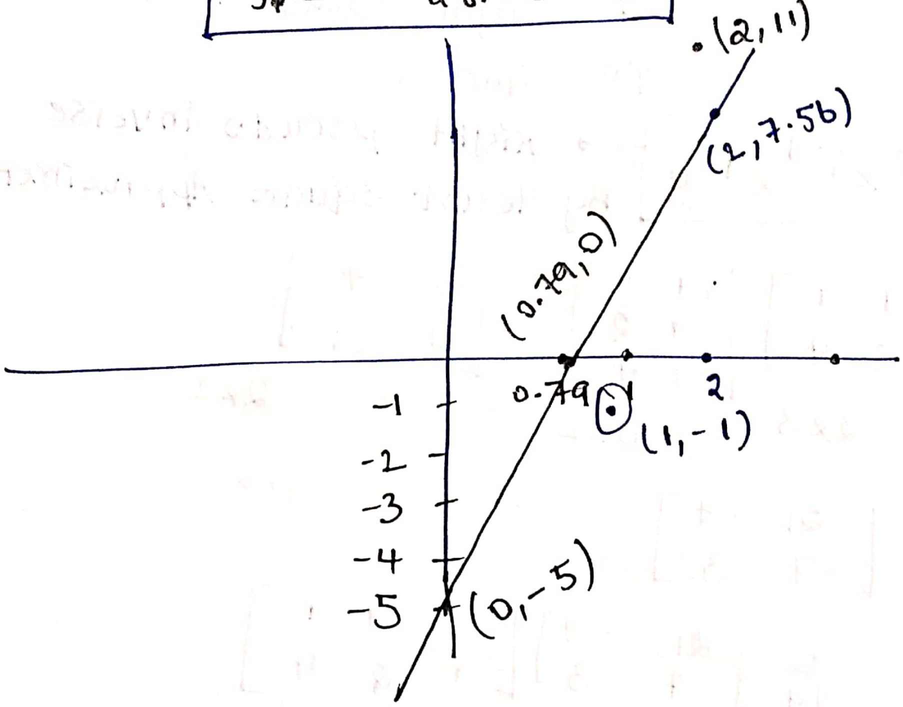
$$\boxed{w^T \cdot x_i + w_0 = 0}$$

For the given Regression problem Hyperplane

$$y_i = w^T \cdot x_i + w_0$$

· (3,19)

$$\boxed{y_i = 6.28 \cdot x - 5} \rightarrow \text{hyper plane}$$



f) Report values of parameter

$\boxed{w_0 = -5}$ $\boxed{w_1 = 6.28}$

g) projection of y onto Range of x

$$\boxed{y' = X w}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -5 \\ 6.28 \end{bmatrix} =$$

$$\boxed{y' = \begin{bmatrix} 1.28 \\ 7.56 \\ 20.12 \end{bmatrix} = w \cdot X} \rightarrow \text{projection of } y \text{ onto Range of } X.$$

Q5: consider the following set of points $\{(x_i, y_i)\}$: $\{(-1,0), (2,2), (4,3)\}$.
Is that possible to find hyperplane that generates data justify your answer.

Sol:

$$\underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}}_{3\times 2} \underbrace{\begin{bmatrix} w_0 \\ w_1 \end{bmatrix}}_{2\times 1} = \underbrace{\begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}}_{3\times 1} \qquad w = (x^T x)^{-1} x^T y$$

$$x \qquad w = y$$

$N = 3 \quad (n+1) = 2$

$\boxed{N > (n+1) \Rightarrow \text{overdetermined system}}$
↳ solved by least square approximation [Right pseudo inverse]

$$X = \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}$$

$$x^T x = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}$$

$$x^T x = \begin{bmatrix} 3 & 5 \\ 5 & 21 \end{bmatrix} \qquad (x^T x)^{-1} = \frac{1}{38} \begin{bmatrix} 21 & -5 \\ -5 & 3 \end{bmatrix}$$

$$(x^T x)^{-1} \cdot x^T = \frac{1}{38} \underbrace{\begin{bmatrix} 21 & -5 \\ -5 & 3 \end{bmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ -1 & 2 & 4 \end{bmatrix}}_{2\times 3}$$

$$= \frac{1}{38} \begin{bmatrix} 26 & 11 & 1 \\ -8 & 1 & 7 \end{bmatrix}$$

$w = (x^T x)^{-1} x^T y$

$= \frac{1}{38} \begin{bmatrix} 26 & 11 & 1 \\ -8 & 1 & 7 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$

$w = \frac{1}{38} \begin{bmatrix} 25 \\ 23 \end{bmatrix} \Rightarrow w = \begin{bmatrix} 0.657 \\ 0.605 \end{bmatrix}$

$\boxed{w_0 = 0.657}$ $\boxed{w_1 = 0.605}$

hyperplane for given Regression pblm $\rightarrow$ $\boxed{y = w^T x_i + w_0}$

$\boxed{y = 0.605 x_1 + 0.657}$

↳ Hyperplane for the given pblm

Q 04: Consider the following data

$w_0 + 2w_1 - w_2 = 1$

$w_0 - w_1 + 2w_2 = 1$

$w_0 + 3w_1c + cw_2 = 1$

a] Express $y = [1, 1, 1]^T$ as a linear combination of vectors. Are those vectors linearly independent?

Is ~~this~~ that possible to find the ~~unknown~~ unknown parameters of eqn? Justify your answer?

Sol:

$$\begin{bmatrix} 1 & 2 & -1 \\ 1 & -1 & 2 \\ 1 & 3c & c \end{bmatrix} \rightarrow \begin{array}{l} R_2 = R_2 - R_1 \\ R_3 = R_3 - R_1 \end{array}$$

$$\begin{array}{rrrr} & 1 & -1 & 2 \\ - & 1 & 2 & -1 \\ \hline R_2 \rightarrow & 0 & -3 & 3 \end{array} \qquad \begin{array}{rrrr} & 1 & 3c & c \\ - & 1 & 2 & -1 \\ \hline R_3 \rightarrow & 0 & 3c-2 & c+1 \end{array} \Rightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 3c-2 & c+1 \end{bmatrix}$$

$R_3 = 3 \times R_3 + (3c-2) R_2$

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 12c-3 \end{bmatrix}$$

$$\begin{array}{ccc} 0 & 3[3c-2] & 3c+3 \\ 0 & -3[3c-2] & 9c-6 \\ \hline 0 & 0 & 12c-3 \end{array}$$

<u>Conclusion</u>: given vectors are $\boxed{\text{linearly independent}}$ for $\boxed{c \neq 1/4}$

<u>Linearly Independent</u>	<u>Linearly Dependent</u>
→ If the given vectors are Linearly independent i.e $\boxed{c \neq 1/4}$	Then $c = 1/4$.
	Echelon form of x is. $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{bmatrix}$
Then we $\boxed{\text{can't find}}$ the unknown parameters $[w_0, w_1, w_2]$ $\boxed{\text{because } c \text{ is unknown}}$	→ There exists more than one solution
	→ $\boxed{N < (M+1)}$ under determined system
	→ $\boxed{\text{unknown parameters}}$ $\boxed{\text{can be found}}$ by $\boxed{\text{Legrangian Formulation}}$
	$\boxed{w = x^T (x x^T)^{-1} \cdot y}$

b] Find:

No of attributes = 2

No of data points = 3

$x_1 = (2, -1)$, $y_1 = 1$

$x_2 = (-1, 2)$ $y_2 = 1$

$x_3 = (3c, c)$ $y_3 = 1$

c] Find the hyper plane that generates the data using least square Regression, by applying direct method.

Sol: Linearly ~~Depen~~ Independent :

→ If the given vectors are Linearly independent i,e $\boxed{c \neq 1/4}$

→ As the $\boxed{c \text{ is unknown}}$

we $\boxed{\text{can't find}}$ the unknown parameters $[w_0, w_1, w_2]^T$

→ so $\boxed{\text{hyperplane is not possible}}$ if $[w_0, w_1, w_2]^T$ are unknown

b/c hyperplane $\boxed{y_i = w^T x_i + w_0}$

Linearly Dependent :

→ If the given vectors are $\boxed{\text{Linearly Dependent}}$

Then $\boxed{c = 1/4}$

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 1 & -1 & 2 & 1 \\ 1 & 3c & c & 1 \end{array}\right] \Rightarrow \begin{array}{l} R_2 = R_2 - R_1 \\ R_3 = R_3 - R_1 \end{array} \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -3 & 3 & 0 \\ 0 & 3c-2 & c+1 & 0 \end{array}\right]$$

Augumented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -3 & 3 & 0 \\ 0 & 0 & 12c-3 & 0 \end{array}\right] \leftarrow R_3 = 3R_3 + (3c-2)R_2$$

Linearly dependent $\quad 12c - 3 = 0 \Rightarrow c = 1/4.$

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$n+1 = 3 \qquad N = 2$

$(n+1) > N \Rightarrow$ $\boxed{\text{under determined system}}$

↳ solved by

Legrangian Formulation
(or)
Minimum Normed solution
i.e left pseudo inverse.

↳ under determined $\boxed{\text{Cannot be solved}}$ by $\boxed{\text{least square approximation}}$

Legrangian Formulation (or) Minimum Normed solution:

left pseudo inverse

$$\boxed{w = X^T (X X^T)^{-1} \cdot y}$$

$$X = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \end{bmatrix} \qquad y = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \end{bmatrix} \qquad y = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$w = X^T (XX^T)^{-1} \cdot y$$

$$XX^T = \underset{2\times 3}{\begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} 1 & 0 \\ 2 & -3 \\ -1 & 3 \end{bmatrix}}$$

$$= \begin{bmatrix} 6 & -9 \\ -9 & 18 \end{bmatrix}$$

$$(XX^T)^{-1} = \frac{1}{27}\begin{bmatrix} 18 & 9 \\ 9 & 6 \end{bmatrix}$$

$$X^T (XX^T)^{-1} = \frac{1}{27} \cdot \underset{3\times 2}{\begin{bmatrix} 1 & 0 \\ 2 & -3 \\ -1 & 3 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} 18 & 9 \\ 9 & 6 \end{bmatrix}}$$

$$X^T (XX^T)^{-1} = \frac{1}{27}\begin{bmatrix} 18 & 9 \\ 9 & 0 \\ 9 & 9 \end{bmatrix}$$

$$X^T (XX^T)^{-1} y = \frac{1}{27}\begin{bmatrix} 18 & 9 \\ 9 & 0 \\ 9 & 9 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \frac{1}{27}\begin{bmatrix} 18 \\ 9 \\ 9 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ 1/3 \end{bmatrix} = w$$

$$w_0 = 2/3 \qquad w_1 = 1/3 \qquad w_2 = 1/3$$

Hyper plane $\Rightarrow$ $y_i = w^T x_i + w_0$

$$y_i = w_1 \cdot x_1 + w_2 x_2 + w_0$$

Hyper plane $\boxed{y_i = 1/3\, x_1 + 1/3\, x_2 + 2/3}$ $\rightarrow$ using Legrangian Formulation

Q12 consider the following Regression problem

$\{(x_i, y_i)\}, i = 1, 2, 3, 4\}$

where $x_1 = (-1,1)^T$ $x_2 = (0,2)^T$ $x_3 = (1,1)^T$, $x_4 = (2,1)^T$

and $y_1 = -1$ $y_2 = 1$ $y_3 = 2$ $y_4 = 0$. let $xw = y$

If it is possible to find hyper plane that generates data

Sol:

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix}$$

$N = 4$ $(n+1) = 3$

$N > (n+1) \Rightarrow$ over determined system

↳ parameter can be found by Least square Regression

a]

$x(\text{domain}) = R^3$

$x(\text{Range}) = R^4$

$x: w \longrightarrow y$

$x: R^3 \longrightarrow R^4$

Right pseudo inverse $w = (x^T x)^{-1} \cdot x^T \cdot y$

Hyper plane $y_i = w^T x_i + w_0$

→ As $N > (n+1) \Rightarrow$ There Exsits some y's which doesn't have solution

$\boxed{x \text{ is not onto}}$

$S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\}$

→ As the given vectors uniquely Represented

$\boxed{x \text{ is one-one}}$

b] dimension of Range $(x) = n+1 = 3$

$\boxed{\begin{array}{l} x \text{ is uniquely Represented} \rightarrow x \text{ is one-one} \\ \text{Basis and dimension } R(x) = n+1 \end{array}}$

(iii) Basis of R (x) = n+1 = 3

$$S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\}$$

(iv) write the Expression to find the projection of y onto the Range of X

$$\boxed{y' = W X}$$

$N > (n+1) \Rightarrow$ overdetermined system

↳ least square Approx

Right pseudo inverse

$$\boxed{w = (X^T X)^{-1} \cdot X^T \cdot y}$$