Name: Tejas Acharya

Classs: EE-559

Assignment: Homework 2

1. Num Classes $C = 3$

Method to be used, One v/s One

Discriminant Functions -> $g_{12}(\underline{x}) = -x_1 - x_2 + 5$

$g_{13}(\underline{x}) = -x_1 + 3$

$g_{23}(\underline{x}) = -x_1 + x_2 - 1$

$g_{ij}(\underline{x}) = -g_{ji}(\underline{x})$

Decision Rule ->

$\underline{x} \in \Gamma_k$ iff $g_{kj}(\underline{x}) > 0$ for all $j \neq k$

A We $g_{kj}(\underline{x}) = 0$ is Decision Boundary.

Boundary H_{12} -> $g_{12}(\underline{x}) = 0 \Rightarrow -x_1 - x_2 + 5 = 0$

$\Rightarrow \boxed{x_1 + x_2 = 5}$

$x_1 = 0 \Rightarrow x_2 = 5 \rightarrow (0,5)$
$x_2 = 0 \Rightarrow x_1 = 5 \rightarrow (5,0)$ } -> Line joining will be H_{12}

$H_{13} \rightarrow g_{13}(\underline{x}) = 0 \Rightarrow -x_1 + 3 \Rightarrow \boxed{x_1 = 3}$

$H_{23} \rightarrow g_{23}(\underline{x}) = 0 \Rightarrow -x_1 + x_2 - 1 = 0 \Rightarrow \boxed{x_2 - x_1 = 1}$

$x_1 = 0 \Rightarrow x_2 = 1 \rightarrow (0,1)$
$x_2 = 0 \Rightarrow x_1 = -1 \rightarrow (-1,0)$] -> Line joining will be H_{23}

Points

- $(2,4) \rightarrow$

$g_{12}([2,4]) = -2-4+5 = -1 \quad ; \quad g_{21}([2,4]) = 1$

$g_{13}([2,4]) = -2+3 = 1 \quad ; \quad g_{31}([2,4]) = -1$

$g_{23}([2,4]) = -2+4-1 = 1 \quad ; \quad g_{32}([2,4]) = -1$

For Γ_1, $g_{12}([2,4]) < 0$; $g_{13}([2,4]) > 0$

For Γ_2, $g_{21}([2,4]) > 0$; $g_{23}([2,4]) > 0$

For Γ_3, $g_{31}([2,4]) < 0$; $g_{32}([2,4]) < 0$

$\therefore g_{2j}([2,4]) > 0$ for $j \in \{1,3\}$

$\therefore [2,4] \in \Gamma_2$

- $(4,3) \rightarrow$

$g_{12}([4,3]) = -4-3+5 = -2 \quad ; \quad g_{21}([4,3]) = 2$

$g_{13}([4,3]) = -4+3 = -1 \quad ; \quad g_{31}([4,3]) = 1$

$g_{23}([4,3]) = -4+3-1 = -2 \quad ; \quad g_{32}([4,3]) = 2$

For Γ_1, $g_{12}(\underline{x}) < 0$, $g_{13}(\underline{x}) < 0$

For Γ_2, $g_{21}(\underline{x}) > 0$, $g_{23}(\underline{x}) < 0$

For Γ_3, $g_{31}(\underline{x}) > 0$; $g_{32}(\underline{x}) > 0$

$\therefore g_{3j}([4,3]) > 0$ for $j \in \{1,2\}$

$\therefore [4,3] \in \Gamma_3$

- $(1,2) \rightarrow$

$g_{12}([1,2]) = -1-2+5 = 2$; $g_{21}(\underline{x}) = -2$

$g_{13}([1,2]) = -1+3 = 2$; $g_{31}(\underline{x}) = -2$

$g_{23}([1,2]) = -1+2-1 = 0$; $g_{32}(\underline{x}) = 0$

For Γ_1, $g_{12}([1,2]) > 0$; $g_{13}([1,2]) > 0$

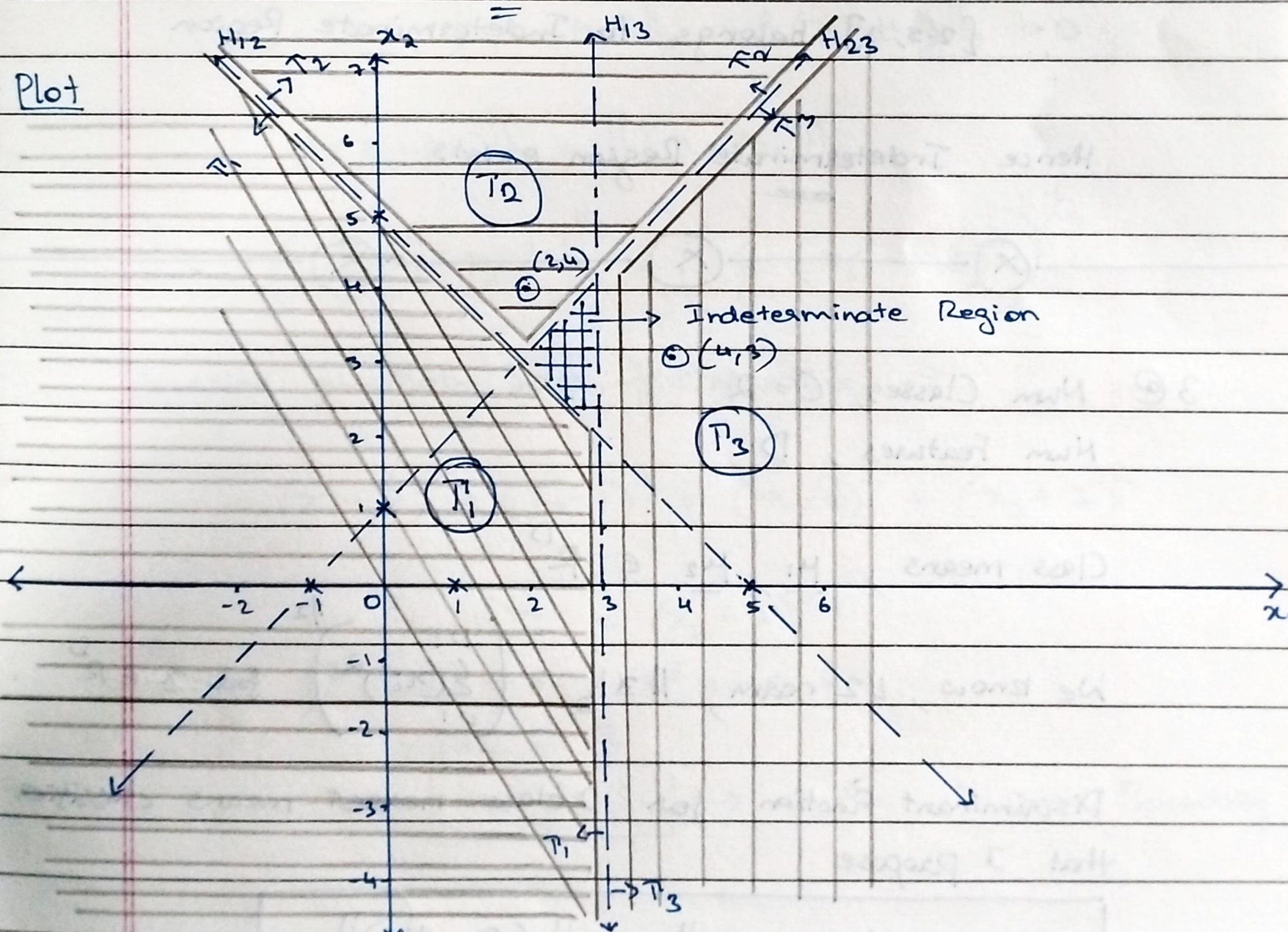
For Γ_2, $g_{21}([1,2]) < 0$; $g_{23}([1,2]) = 0$

For Γ_3, $g_{31}([1,2]) < 0$; $g_{32}([1,2]) = 0$

$\therefore$ $g_{1j}([1,2]) > 0$ for $j \in \{2,3\}$

$\therefore$ $\underline{[1,2] \in \Gamma_1}$

- Plot



~~(4,3)~~ (2.5, 3)

$g_{12}([2.5, 3]) = -2.5 - 3 + 5 = -0.5$; $g_{21}(x) = 0.5$

$g_{13}([2.5, 3]) = -2.5 + 3 = 0.5$; $g_{31}(x) = -0.5$

$g_{23}([2.5, 3]) = -2.5 + 3 - 1 = -0.5$; $g_{32}(x) = 0.5$

For $\pi_1 \to g_{12}(x) < 0$; $g_{13}(x) > 0$

For $\pi_2 \to g_{21}(x) > 0$; $g_{23}(x) < 0$

For $\pi_3 \to g_{31}(x) < 0$; $g_{32}(x) > 0$

$\therefore$ ~~For~~ ~~none~~ [2.5, 3] doesn't satisfy the rule

$g_{kj}(x) > 0$ for all $j \neq k$

$\therefore$ [2.5, 3] belongs to Indeterminate Region

Hence Indeterminate Region exists.

3 (a) Num Classes, $C = 2$

Num Features, D

Class means, $\underline{\mu_1}, \underline{\mu_2} \in \mathbb{R}^D$

We know, L2 norm, $\|x\|_2 = \left(\sum_{i=1}^{D} (x_i)^2 \right)^{1/2}$ for $x \in \mathbb{R}^D$

Discriminant Function for 2 class nearest means classifier that I propose

$$\boxed{g(x) = \|(x - \underline{\mu_2})\|_2 - \|(x - \underline{\mu_1})\|_2}$$

with a Decision Rule,

$$g(x) \underset{\pi_2}{\overset{\pi_1}{\gtrless}} 0$$

Desion Boundary,

$$\|(x - \mu_1)\|_2 - \|(x - \mu_2)\|_2$$

It is a Linear Classifier.

(b) $\mu_1 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$, $\mu_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

To get Decision Boundary, $g(x) = 0$

$$g(x) = 0 \Rightarrow \|(x - \mu_2)\|_2 - \|(x - \mu_1)\|_2 = 0$$

$$\|(x - \mu_2)\|_2 = \|(x - \mu_1)\|_2$$

$$\left(\sum_{i=1}^{2} (x_i - \mu_{2i})^2 \right)^{1/2} = \left(\sum_{i=1}^{2} (x_i - \mu_{1i})^2 \right)^{1/2}$$

Squaring on Both Sides & Expanding $\Rightarrow$

$$\cancel{(x_1 - 0)^2} + (x_2 - 1)^2 = \cancel{(x_1 - 0)^2} + (x_2 + 2)^2$$

$$(x_2 - 1)^2 = (x_2 + 2)^2$$

$$\cancel{x_2^2} - 2x_2 + 1 = \cancel{x_2^2} + 4x_2 + 4$$

$$-6x_2 = 3$$

$\boxed{x_2 = -0.5}$ $\rightarrow$ is the Decision Boundary

• $\underline{(0,0) = x}$

$$g([0,0]) = \|x - \mu_2\|_2 - \|x - \mu_1\|_2$$

$$= \left((0-0)^2 + (0-1)^2\right)^{1/2} - \left((0-0)^2 + (0+2)^2\right)^{1/2}$$

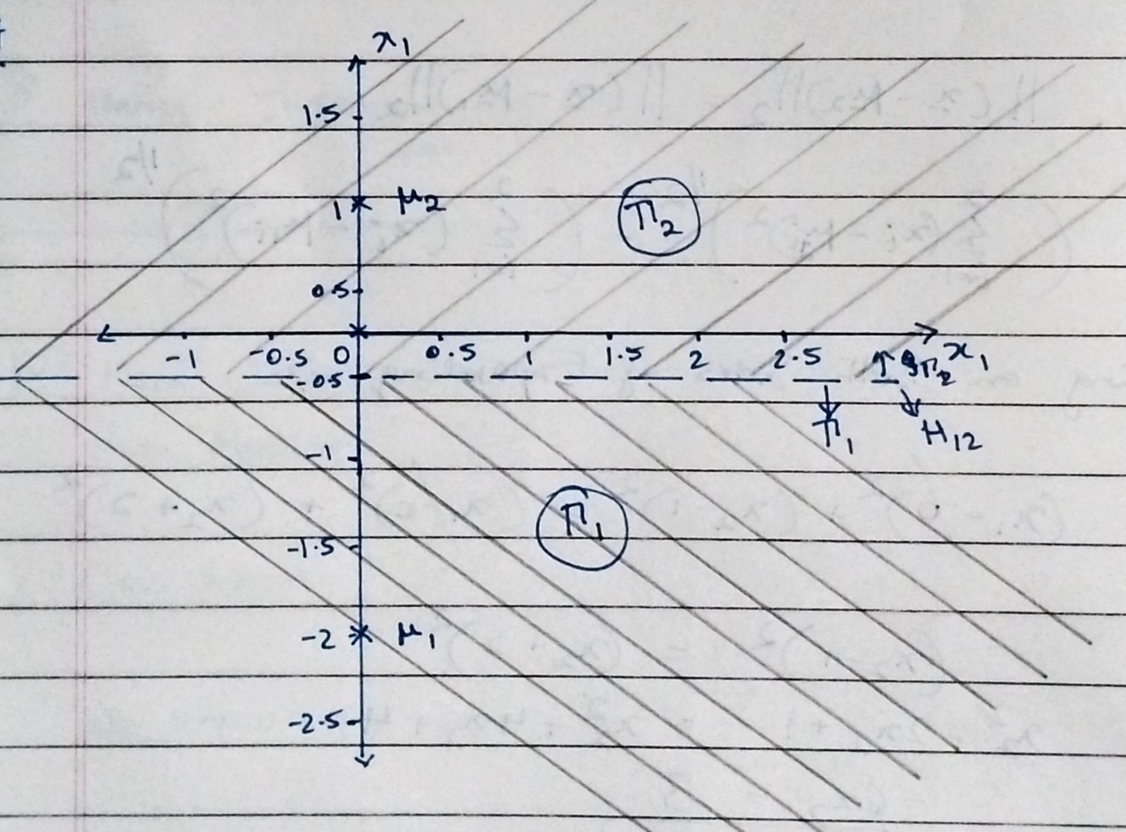
$$= \left(1^2\right)^{1/2} - \left(2^2\right)^{1/2}$$

$$= 1 - 2$$

$g([0,0]) = -1$ which is less than 0.

$\therefore [0,0] \in \Pi_2$

Plot



3 (c) Num Classes, $C=3$
Num Features, D

Method $\rightarrow$ Maximal Value Method

Class Means $\rightarrow$ $\underline{\mu}_1$, $\underline{\mu}_2$, $\underline{\mu}_3$

I choose the following discriminant functions,

$$g_1(\underline{x}) = -\|\underline{x} - \underline{\mu}_1\|_2$$
$$g_2(\underline{x}) = -\|\underline{x} - \underline{\mu}_2\|_2$$
$$g_3(\underline{x}) = -\|\underline{x} - \underline{\mu}_3\|_2$$

<u>Decision Rule</u>: $g_k(\underline{x}) > g_j(\underline{x}) \quad \forall\ j \neq k$
$\Rightarrow \underline{x} \in \Gamma_k$

Decision Boundary b/w Γ_k & Γ_j is

$$g_k(\underline{x}) = g_j(\underline{x})$$
$$\Rightarrow -\|\underline{x} - \underline{\mu}_k\|_2 = -\|\underline{x} - \underline{\mu}_j\|_2$$

$$\|\underline{x} - \underline{\mu}_k\|_2 = \|\underline{x} - \underline{\mu}_j\|_2$$

which is same as the $C=2$ class case.

Hence the classifier is Linear.

(d) $\underline{\mu}_1 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$, $\underline{\mu}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\underline{\mu}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

Decision Boundaries

$H_{12} \rightarrow \quad g_1(x) = g_2(x)$

$$\| \underline{x} - \mu_1 \|_2 = \| \underline{x} - \mu_2 \|_2$$

$$\left(\sum_{i=1}^{2} (x_i - \mu_{1i})^2 \right)^{1/2} = \left(\sum_{i=1}^{2} (x_i - \mu_{2i})^2 \right)^{1/2}$$

Squaring and Expanding

$$\left(\cancel{(x_1 - 0)^2} + (x_2 + 2)^2 \right)^{\cancel{1/2}} = \cancel{(x_1 - 0)^2} + (x_2 - 1)^2$$

$$\cancel{x_2^2} + 4x_2 + 4 = \cancel{x_2^2} - 2x_2 + 1$$

$$6x_2 = -3$$

$$\boxed{x_2 = -0.5} \rightarrow H_{12}$$

$H_{13} \rightarrow \quad g_1(x) = g_3(x)$

$$\| \underline{x} - \mu_1 \|_2 = \| \underline{x} - \mu_3 \|_2$$

Squaring and Expanding $\rightarrow$

$$\cancel{(x_1 - 0)^2} + (x_2 + 2)^2 = (x_1 - 2)^2 + \cancel{(x_2 - 0)^2}$$

$$\pm (x_2 + 2) = \pm (x_1 - 2)$$

$\Rightarrow x_2 + 2 = x_1 - 2$

$\boxed{x_2 - x_1 = -4}$

$x_2 = 0 \Rightarrow x_1 = 4 \Rightarrow (4, 0)$

$x_1 = 0 \Rightarrow x_2 = -4 \Rightarrow (0, -4)$

$x_2 + \cancel{2} = -x_1 + \cancel{2}$

$\boxed{x_2 = -x_1}$

$x_2 = 0 \Rightarrow x_1 = 0 \rightarrow (0, 0)$

$x_2 = 1 \Rightarrow x_1 = -1 \Rightarrow (-1, 1)$

$\underline{H_{23}} \rightarrow \; g_2(\underline{x}) = g_3(\underline{x})$

$\|\underline{x} - \underline{\mu}_2\|_2 = \|\underline{x} - \underline{\mu}_3\|_2$

Squaring & Expanding $\rightarrow$

$$(x_1 - 0)^2 + (x_2 - 1)^2 = (x_1 - 2)^2 + (x_2 - 0)^2$$

$$(x_2 - 1)^2 = (x_1 - 2)^2$$

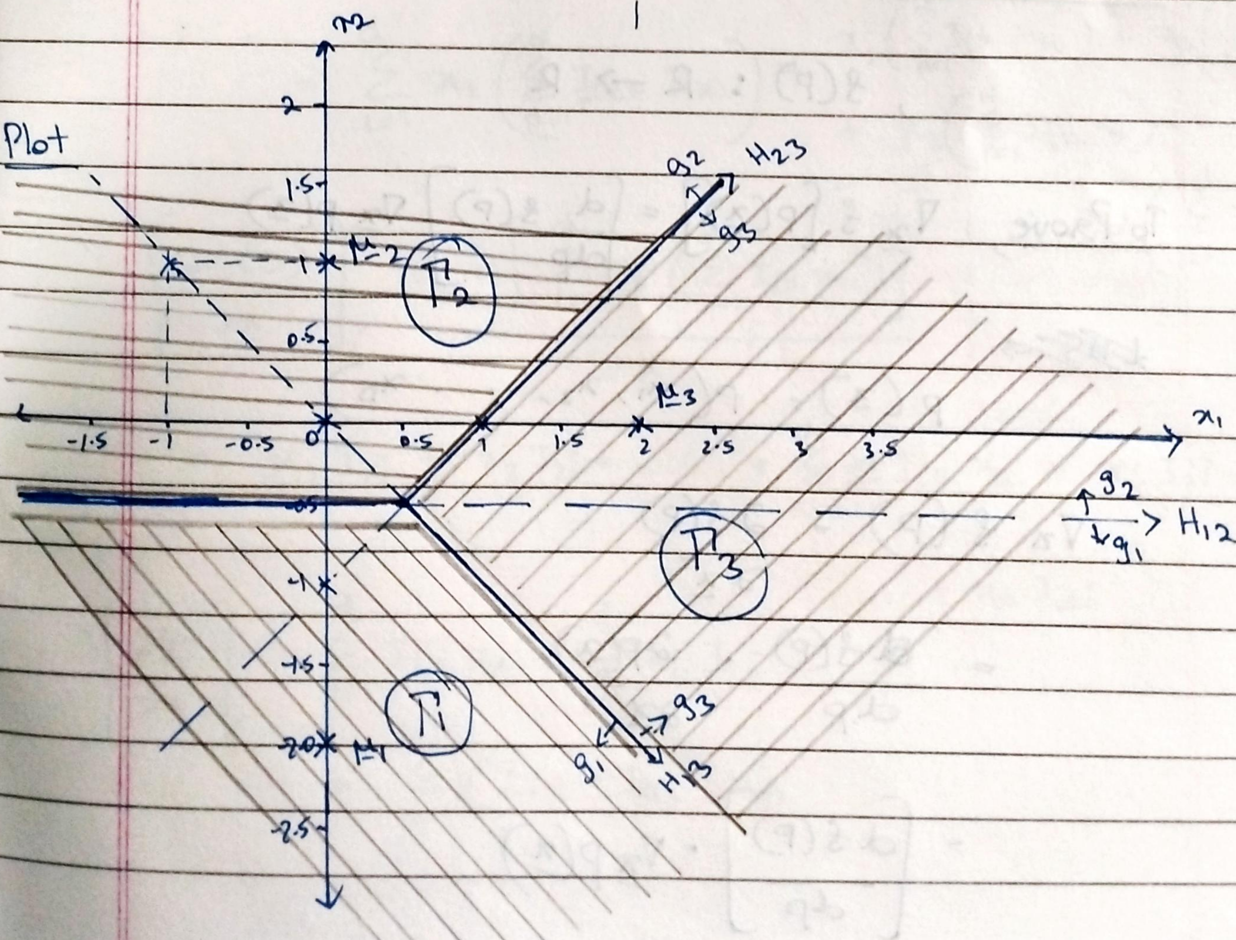
$\rightarrow \; x_2 - 1 = x_1 - 2$ | $x_2 - 1 = -x_1 + 2$

$\boxed{x_2 - x_1 = -1}$ | $\boxed{x_2 + x_1 = 3}$

$x_2 = 0 \Rightarrow x_1 = 1 \rightarrow (1, 0)$ | $x_2 = 0 \Rightarrow x_1 = 3 \rightarrow (3, 0)$

$x_1 = 0 \Rightarrow x_2 = -1 \rightarrow (0, -1)$ | $x_1 = 0 \Rightarrow x_2 = 3 \rightarrow (0, 3)$

Plot



We chose points ~~for~~ $\{(0,0), (-1,1)\}$ for H_{13} and points $\{(1,0), (0,-1)\}$ for H_{23} ~~instead of~~ because they satisfied the discriminat functn of it's decision rule.

Points $\{(4,0), (0,-4)\}$ for H_{13} & $\{(3,0), (0,3)\}$ for H_{23} didn't satisfy the constraints.

---

4 (a) Given,

$$\underline{x} \in R^D$$

Scalar function, $P(\underline{x}) : R^D \to R$

, $f(P) : R \to R$

To Prove, $\nabla_{\underline{x}} f[P(\underline{x})] = \left[\frac{d}{dp} f(P)\right] \nabla_{\underline{x}} P(\underline{x})$

~~LHS~~ $\Rightarrow$

$$p(\underline{x}) = P(x_1, x_2, \ldots x_D)$$

$$\nabla_{\underline{x}} f(P) = \frac{\partial f(P)}{\partial \underline{x}}$$

$$= \frac{d f(P)}{dp} \cdot \frac{\partial P(\underline{x})}{\partial \underline{x}}$$

$$= \left[\frac{d f(P)}{dp}\right] \cdot \nabla_{\underline{x}} P(\underline{x})$$

(b) Find $\nabla_{\underline{x}} (\underline{x}^T \underline{x})$

$$f(\underline{x}) = \underline{x}^T \underline{x}$$
$$= \underline{x}^T \underline{\underline{I}}\, \underline{x}$$

$$\nabla_x f(\underline{x}) = \frac{\partial f(\underline{x})}{\partial \underline{x}}$$

$$= (\underline{\underline{I}} + \underline{\underline{I}}^T)\, \underline{x}$$

$$\nabla_x f(\underline{x}) = 2\underline{\underline{I}}\, \underline{x} = 2\underline{x}$$

(c) Prove $\nabla_x [\underline{x}^T \underline{x}] = 2\underline{x}$

$$f(\underline{x}) = \underline{x}^T \underline{x} = \underline{x}^T \underline{\underline{I}}\, \underline{x}$$

$$= \sum_{l=1}^{d} x_l \left(\sum_{k=1}^{d} I_{lk}\, x_k \right) = x_1 \left(\sum_{k=1}^{d} I_{1k}\, x_k \right) + x_2 \left(\sum_{k=1}^{d} I_{2k}\, x_k \right) + x_3 \left(\sum_{k=1}^{d} I_{3k}\, x_k \right) + \dots + x_d \left(\sum_{k=1}^{d} I_{dk}\, x_k \right)$$

$$\left[\frac{\partial f(\underline{x})}{\partial \underline{x}} \right]_i = \frac{\partial \left[\sum_{l=1}^{d} x_l \left(\sum_{k=1}^{d} I_{lk}\, x_k \right) \right]}{\partial x_i}$$

$$= x_1 I_{1i} + x_2 I_{2i} + \dots + \sum_{k=1}^{d} I_{ik}\, x_k + x_i I_{ii} + \dots + x_d I_{di}$$

$$\left[\frac{\partial f(\underline{x})}{\partial \underline{x}} \right]_i = \sum_{k=1}^{d} I_{ik}\, x_k + \sum_{l=1}^{d} x_l I_{li}$$

$$= \sum_{k=1}^{d} (I_{ik} + I_{ki})\, x_k$$

$$\left[\frac{\partial f(\underline{x})}{\partial \underline{x}}\right]_i = (I_{i1} + \cancel{I_{1i}})x_1 + (I_{i2} + \cancel{I_{2i}})x_2 + \dots + (I_{ii} + I_{ii})x_i + \dots (I_{id} + I_{di})x_d$$

$$= 2x_i$$

$$\nabla_{\underline{x}} f(\underline{x}) = \begin{bmatrix} \left(\frac{\partial f(\underline{x})}{\partial \underline{x}}\right)_1 \\ \left(\frac{\partial f(\underline{x})}{\partial \underline{x}}\right)_2 \\ \vdots \\ \vdots \\ \left(\frac{\partial f(\underline{x})}{\partial \underline{x}}\right)_d \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ \vdots \\ 2x_d \end{bmatrix} = 2\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$\boxed{\nabla_{\underline{x}} f(\underline{x}) = 2\underline{x}}$$

Hence Proved

(a) $\nabla_{\underline{x}}\left[(\underline{x}^T\underline{x})^3\right]$: Let $p(\underline{x}) = \underline{x}^T \cdot \underline{x}$

$\downarrow$

Let $f(p) = p^3$

$$\nabla_{\underline{x}}\left[f(p)\right] = \left[\frac{d f(p)}{dp}\right] \nabla_{\underline{x}}\, p(\underline{x})$$

$$\frac{d f(p)}{dp} = 3p^2$$

$$\nabla_{\underline{x}}\, p(\underline{x}) = \nabla_{\underline{x}}(\underline{x}^T \cdot \underline{x}) = 2\,\underline{x}$$

$$\therefore \nabla_{\underline{x}}\left[(\underline{x}^T \cdot \underline{x})^3\right] = 3(\underline{x}^T \cdot \underline{x})^2 \cdot 2\,\underline{x} = \underline{\underline{6(\underline{x}^T \cdot \underline{x})^2 \cdot \underline{x}}}$$

---

5 (a) Find $\nabla_{\underline{w}} \|\underline{w}\|_2$; We know

$$\|\underline{w}\|_2 = (\underline{w}^T \cdot \underline{w})^{1/2}$$

$$\nabla_{\underline{w}}\left[(\underline{w}^T \cdot \underline{w})^{1/2}\right] \quad ; \quad \text{Let } p(\underline{w}) = \underline{w}^T \cdot \underline{w}$$

$$\text{Let } f(p) = p^{1/2}$$

$$\nabla_{\underline{w}} f(p) = \left[\frac{d f(p)}{dp}\right] \cdot \nabla_{\underline{w}}\, p(\underline{w})$$

$$\frac{d f(p)}{dp} = \frac{1}{2} p^{(-1/2)} \quad ; \quad \nabla_{\underline{w}}(\underline{w}^T \cdot \underline{w}) = 2\underline{w}$$

$$\therefore \nabla_{\underline{w}} \|\underline{w}\|_2 = \frac{1}{2}(\underline{w}^T \cdot \underline{w})^{-1/2} \cdot 2\,\underline{w} = \underline{\underline{(\underline{w}^T \cdot \underline{w})^{-1/2}\, \underline{w}}}$$

(b) Find, $\nabla_{\underline{w}} \|\underline{M}\,\underline{w} - \underline{b}\|_2$; $\|\underline{w}\|_2 = (\underline{w}^T \cdot \underline{w})^{1/2}$

$$\Rightarrow \nabla_{\underline{w}}\left[(\underline{M}\,\underline{w} - \underline{b})^T \cdot (\underline{M}\,\underline{w} - \underline{b})\right]^{1/2}$$

Let $p(\underline{w}) = (\underline{\underline{M}}\,\underline{w} - \underline{b})^T \cdot (\underline{\underline{M}}\,\underline{w} - \underline{b})$

Let $f(p) = p^{1/2}$

$$\nabla_{\underline{w}} f(p) = \left[\frac{df(p)}{dp}\right] \cdot \nabla_{\underline{w}}\, p$$

$$\nabla_{\underline{w}}\, p = \cancel{\nabla_{\underline{w}}}\,\frac{\partial p(\underline{w})}{\partial \underline{w}} = \frac{\partial}{\partial \underline{w}}\left[\left\{(\underline{\underline{M}}\,\underline{w})^T - \underline{b}^T\right\}(\underline{\underline{M}}\,\underline{w} - \underline{b})\right]$$

$$= \frac{\partial}{\partial \underline{w}}\left[(\underline{\underline{M}}\,\underline{w})^T(\underline{\underline{M}}\,\underline{w}) - (\underline{\underline{M}}\,\underline{w})^T\underline{b} - \underline{b}^T\underline{\underline{M}}\,\underline{w} + \underline{b}^T\underline{b}\right]$$

$$= \frac{\partial}{\partial \underline{w}}\left[\underline{w}^T(\underline{\underline{M}}^T\underline{\underline{M}}\,\underline{w}) - \underline{w}^T(\underline{\underline{M}}^T\underline{b}) - (\underline{b}^T\underline{\underline{M}})\underline{w} + \underline{b}^T\underline{b}\right]$$

$$= \frac{\partial}{\partial \underline{w}}\left[\underline{w}^T(\underline{\underline{M}}^T\underline{\underline{M}}\,\underline{w}) - \underline{w}^T(\underline{\underline{M}}^T\underline{b}) - (\underline{\underline{M}}^T\underline{b})^T\underline{w} + \underline{b}^T\underline{b}\right]$$

We know if $f(\underline{x}) = \underline{y}^T\underline{x} = \underline{x}^T\underline{y} \rightarrow \dfrac{\partial f(\underline{x})}{\partial \underline{x}} = \underline{y}$

$$\rightarrow \nabla_{\underline{w}}\, p = \underline{\underline{M}}^T\underline{\underline{M}}\,\underline{w} - \underline{\underline{M}}^T\underline{b} - \underline{\underline{M}}^T\underline{b}$$

$$= \underline{\underline{M}}^T\underline{\underline{M}}\,\underline{w} - 2\underline{\underline{M}}^T\underline{b}$$

$$\boxed{\nabla_{\underline{w}} \|\underline{\underline{M}}\,\underline{w} - \underline{b}\|_2 = \left[(\underline{\underline{M}}\,\underline{w} - \underline{b})^T \cdot (\underline{\underline{M}}\,\underline{w} - \underline{b}) \right] \\ \left(\underline{\underline{M}}^T \underline{\underline{M}}\,\underline{w} - 2\,\underline{\underline{M}}^T \underline{b} \right)}$$