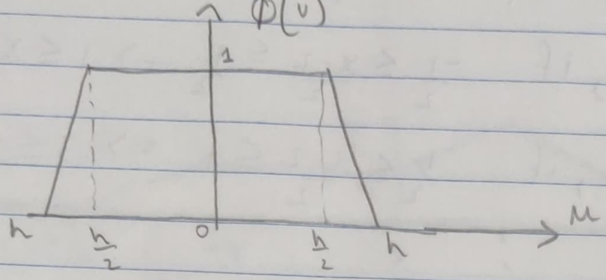
HW 6

↳ Hardik Prajapati (2678294168)

1.



$h = 1$

$D_{\omega_1} = \{0, 2, 5\}$

$D_{\omega_2} = \{4, 7\}$

$$\phi(u) = \begin{cases} 2u+2, & \text{if } -1 \le u \le -\frac{1}{2} \\ 1, & \text{if } -\frac{1}{2} \le u \le \frac{1}{2} \\ -2u+2, & \text{if } \frac{1}{2} \le u \le 1 \\ 0, & \text{else} \end{cases}$$

a) Now,

$$p_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{V_n} \phi\left(\frac{x - x_i}{h_n}\right)$$

$$V_n = \frac{1}{2}(1+2)(1) = \frac{3}{2}$$

$$h_n = 2(1) = 2.$$

$$p_n(x|\omega_1) = \frac{1}{3} \sum_{i=1}^{3} \frac{1}{\left(\frac{3}{2}\right)} \phi\left(\frac{x - x_i}{2}\right)$$

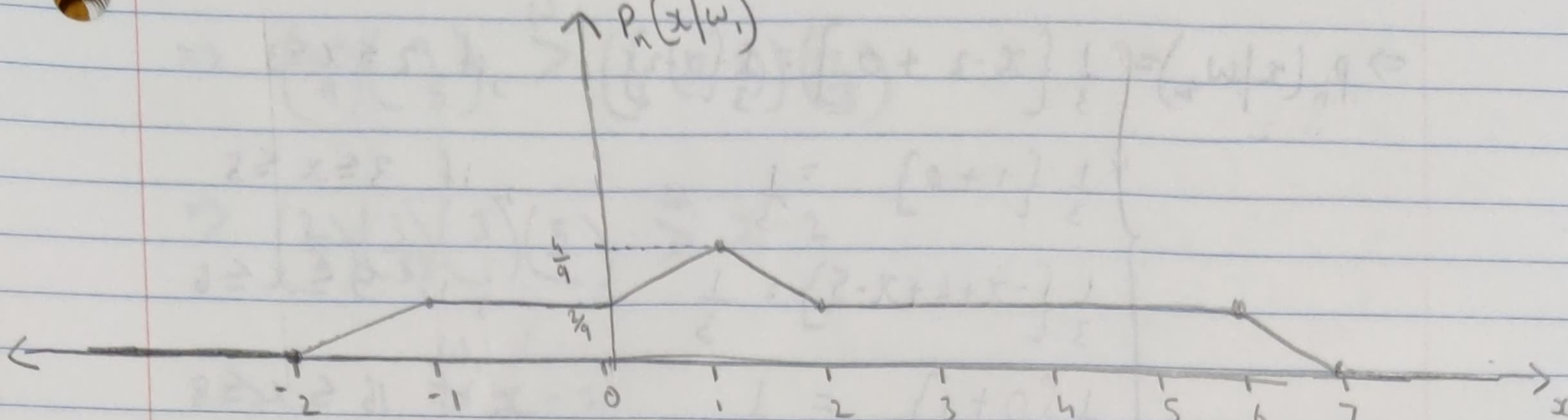
$$= \frac{2}{9}\left\{ \phi\left(\frac{x-0}{2}\right) + \phi\left(\frac{x-2}{2}\right) + \phi\left(\frac{x-5}{2}\right) \right\}$$

$$\phi\left(\frac{x}{2}\right) = \begin{cases} x+2 & \text{if } -1 \le \frac{x}{2} \le -\frac{1}{2} \Rightarrow -2 \le x \le -1 \\ 1 & \text{if } -\frac{1}{2} \le \frac{x}{2} \le \frac{1}{2} \Rightarrow -1 \le x \le 1 \\ -x+2 & \text{if } \frac{1}{2} \le \frac{x}{2} \le 1 \Rightarrow 1 \le x \le 2 \\ 0 & \text{else} \end{cases}$$

$$\phi\left(\frac{x-2}{2}\right) = \begin{cases} x & , \text{if } -1 \le \frac{x-2}{2} \le -\frac{1}{2} \Rightarrow 0 \le x \le 1 \\ 1 & , \text{if } -\frac{1}{2} \le \frac{x-2}{2} \le \frac{1}{2} \Rightarrow 1 \le x \le 3 \\ -x+4 & , \text{if } \frac{1}{2} \le \frac{x-2}{2} \le 1 \Rightarrow 3 \le x \le 4 \end{cases}$$

$$\phi\left(\frac{x-5}{2}\right) = \begin{cases} x-3 & , \text{if } -1 \le \frac{x-5}{2} \le -\frac{1}{2} \Rightarrow 3 \le x \le 4 \\ 1 & , \text{if } -\frac{1}{2} \le \frac{x-5}{2} \le \frac{1}{2} \Rightarrow 4 \le x \le 6 \\ -x+7 & , \text{if } \frac{1}{2} \le \frac{x-5}{2} \le 1 \Rightarrow 6 \le x \le 7 \end{cases}$$

$$\Rightarrow p_n(x|w_1) = \begin{cases} \frac{2}{9}\{x+2+0+0\} = \frac{2}{9}\{x+2\} & , \text{if } -2 \le x \le -1 \\ \frac{2}{9}\{1+0+0\} = \frac{2}{9} & , \text{if } -1 \le x \le 0 \\ \frac{2}{9}\{1+x+0\} = \frac{2}{9}\{1+x\} & , \text{if } 0 \le x \le 1 \\ \frac{2}{9}\{-x+2+1+0\} = \frac{2}{9}\{-x+3\} & , \text{if } 1 \le x \le 2 \\ \frac{2}{9}\{0+1+0\} = \frac{2}{9} & , \text{if } 2 \le x \le 3 \\ \frac{2}{9}\{0-x+4+x-3\} = \frac{2}{9} & , \text{if } 3 \le x \le 4 \\ \frac{2}{9}\{0+0+1\} = \frac{2}{9} & , \text{if } 4 \le x \le 6 \\ \frac{2}{9}\{0+0-x+7\} = \frac{2}{9}\{-x+7\} & , \text{if } 6 \le x \le 7 \\ 0 & , \text{else} \end{cases}$$



Similarly,

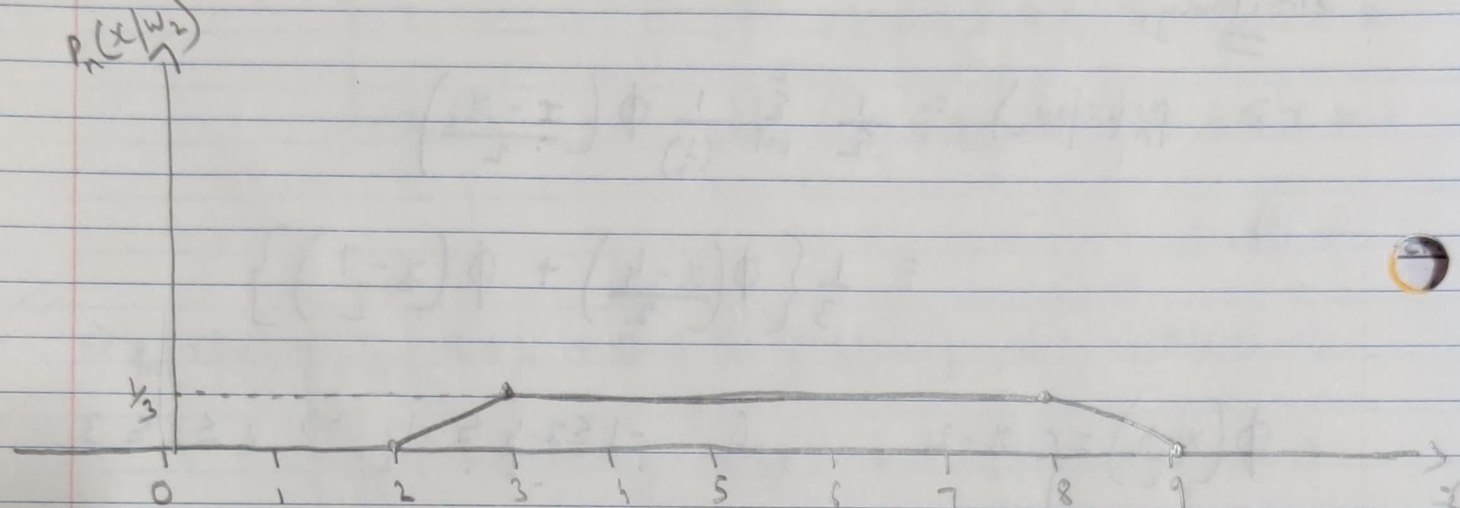
$$P_n(x|w_2) = \frac{1}{2}\sum_{i=1}^{2}\frac{1}{(\frac{3}{2})}\phi\left(\frac{x-x_i}{2}\right)$$

$$= \frac{1}{3}\left\{\phi\left(\frac{x-4}{2}\right) + \phi\left(\frac{x-7}{2}\right)\right\}$$

$$\phi\left(\frac{x-4}{2}\right) = \begin{cases} x-2 & \text{, if } -1 \le \frac{x-4}{2} \le -\frac{1}{2} \Rightarrow 2 \le x \le 3 \\ 1 & \text{, if } -\frac{1}{2} \le \frac{x-4}{2} \le \frac{1}{2} \Rightarrow 3 \le x \le 5 \\ -x+6 & \text{, if } \frac{1}{2} \le \frac{x-4}{2} \le 1 \Rightarrow 5 \le x \le 6 \end{cases}$$

$$\phi\left(\frac{x-7}{2}\right) = \begin{matrix} x-5 & \text{, if } -1 \le \frac{x-7}{2} \le -\frac{1}{2} \Rightarrow 5 \le x \le 6 \\ 1 & \text{, if } -\frac{1}{2} \le \frac{x-7}{2} \le \frac{1}{2} \Rightarrow 6 \le x \le 8 \\ -x+9 & \text{, if } \frac{1}{2} \le \frac{x-7}{2} \le 1 \Rightarrow 8 \le x \le 9 \end{matrix}$$

$$\Rightarrow p_n(x|\omega_2) = \begin{cases} \frac{1}{3}\{x-2+0\} = \frac{1}{3}\{x-2\} & , \text{if } 2 \le x \le 3 \\ \frac{1}{3}\{1+0\} = \frac{1}{3} & , \text{if } 3 \le x \le 5 \\ \frac{1}{3}\{-x+6+x-5\} = \frac{1}{3} & ; \text{if } 5 \le x \le 6 \\ \frac{1}{3}\{0+1\} = \frac{1}{3} & , \text{if } 6 \le x \le 8 \\ \frac{1}{3}\{0-x+9\} = \frac{1}{3}\{-x+9\} & , \text{if } 8 \le x \le 9 \end{cases}$$



Q1. (b) $P(\omega_1) = \frac{3}{5}$

$P(\omega_2) = \frac{2}{5}$

Q1. (c) Choose ω_1 if

$$P(\omega_1|x) > P(\omega_2|x)$$

$$\Rightarrow \quad p(x|\omega_1)P(\omega_1) > p(x|\omega_2)P(\omega_2)$$

$\Rightarrow \left(\frac{2}{9}\right)\left(\frac{3}{5}\right) \overset{\omega_1}{>} \left(\frac{1}{3}(x-2)\right)\left(\frac{2}{5}\right)$

$\Rightarrow \left(\frac{2}{9}\right)\left(\frac{3}{5}\right)(3)\left(\frac{5}{2}\right) \overset{\omega_1}{>} x-2$

$\Rightarrow 1 \overset{\omega_1}{>} x-2$

$\Rightarrow 3 \overset{\omega_1}{>} x$

$\Rightarrow \left(\frac{2}{9}(-x+7)\right)\left(\frac{3}{5}\right) \overset{\omega_1}{>} \left(\frac{1}{3}\right)\left(\frac{2}{5}\right)$

$\Rightarrow -x+7 \overset{\omega_1}{>} 1$

$\Rightarrow 6 \overset{\omega_1}{>} x$

$$\Rightarrow Y = \begin{cases} \omega_1, & \text{if } x < 3 \\ \omega_1 \text{ or } \omega_2, & \text{if } 3 \leq x < 6 \\ \omega_2, & \text{if } 6 \leq x \end{cases}$$

//

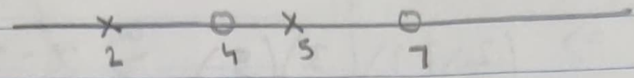
2.

$k = 2$

No. of feature $= 1 = d$

$D_{\omega_1} = \{2, 5\}$

$D_{\omega_2} = \{4, 7\}$



$$\hat{P}_{kNN}(x|\omega_1) = \frac{k}{n} \frac{1}{V_d R_k^d(x)} \quad , \quad V_d = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$$

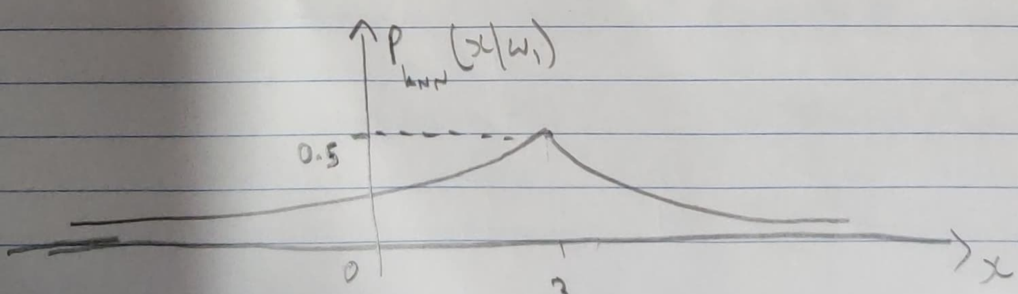
R_k = Radius of x from its k^{th} Nearest neighbour.

$$\Rightarrow V_1 = \frac{\pi^{1/2}}{\Gamma\left(\frac{1}{2}+1\right)} = \frac{\pi^{1/2}}{\Gamma\left(\frac{3}{2}\right)} = \frac{\sqrt{\pi}}{\frac{\sqrt{\pi}}{2}} = 2.$$

$$R_k = \begin{cases} |x-5| , & \text{if } x \leq 3 \\ |x-2| , & \text{else} \end{cases}$$

$$\Rightarrow \hat{P}_{kNN}(x|\omega_1) = \frac{2}{2} \times \left\{ \frac{1}{2R_k} \right\}$$

$$\Rightarrow \hat{P}_{kNN}(x|\omega_1) = \begin{cases} \dfrac{1}{2|x-5|} & \text{if } x \leq 3 \\[2ex] \dfrac{1}{2|x-2|} & \text{else} \end{cases}$$

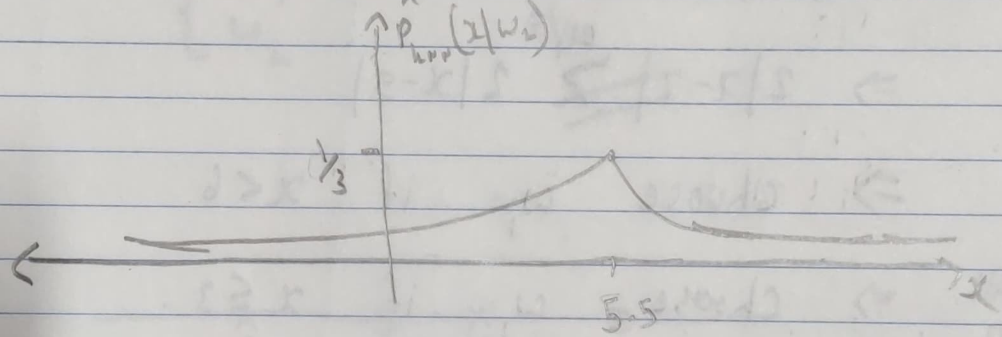


//

$$\Rightarrow \hat{P}_{kNN}(x|\omega_2) = \frac{2}{2}\left\{\frac{1}{2R_k}\right\}$$

$$R_k = \begin{cases} |x-7| & \text{, if } x \leq 6 \\ |x-4| & \text{, else} \end{cases}$$

$$\Rightarrow \hat{P}_{kNN}(x|\omega_2) = \begin{cases} \frac{1}{2|x-7|} & \text{, if } x < 5.5 \\ \frac{1}{2|x-4|} & \text{, else} \end{cases}$$



Q2. b) $\hat{P}(\omega_1) = \frac{2}{4} = \frac{1}{2}$

$\hat{P}(\omega_2) = \frac{2}{4} = \frac{1}{2}$

Q2.(c) Choose ω_1 if

$$P(\omega_1|x) > P(\omega_2|x)$$

$$\Rightarrow P(x|\omega_1)P(\omega_1) > P(x|\omega_2)P(\omega_2)$$

$$\text{As } P(\omega_1) = P(\omega_2)$$

$$\Rightarrow P(x|\omega_1) \overset{\omega_1}{>} P(x|\omega_2)$$

Case:- $x \leq 3$

$$\frac{1}{2|x-5|} \overset{\omega_1}{>} \frac{1}{2|x-7|}$$

$$\Rightarrow 2|x-7| \overset{\omega_1}{>} 2|x-5|$$

$\Rightarrow$ choose ω_1 if $x < 6$

$\Rightarrow$ Choose ω_1 if $x \leq 3$.

Case:- $x > 3$ & $x < 5.5$

$$\Rightarrow \frac{1}{2|x-2|} \overset{\omega_1}{>} \frac{1}{2|x-7|}$$

$$\Rightarrow |x-7| \overset{\omega_1}{>} |x-2|$$

$$\Rightarrow x < \frac{9}{2}$$

$\Rightarrow$ Choose ω_1 if $3 < x < 4.5$

Choose ω_2 if $4.5 < x < 5.5$

Case:- $x > 5.5$

$$\Rightarrow \frac{1}{2|x-2|} \overset{\omega_1}{>} \frac{1}{2|x-4|}$$

$$\Rightarrow |x-4| \overset{\omega_1}{>} |x-2|$$

$$\Rightarrow x < 3$$

$\Rightarrow$ choose ω_2 if $x > 5.5$

$$\Rightarrow Y = \begin{cases} \omega_1 & \text{if } x < 4.5 \\ \omega_2 & \text{else} \end{cases}$$

Q2. (d).

$Y = \omega_2 \quad \forall x \in \mathbb{R}$, based on $k=2$.

$\Rightarrow$ This is a bad estimator as it will always estimate ω_2 because of the upper edge in tie-Breaker Decision.

Q3.

$x_1 = [1, 0]^T$, $z_1 = -1$
$x_2 = [0, 1]^T$, $z_2 = -1$
$x_3 = [0, -1]^T$, $z_3 = -1$
$x_4 = [-1, 0]^T$, $z_4 = 1$

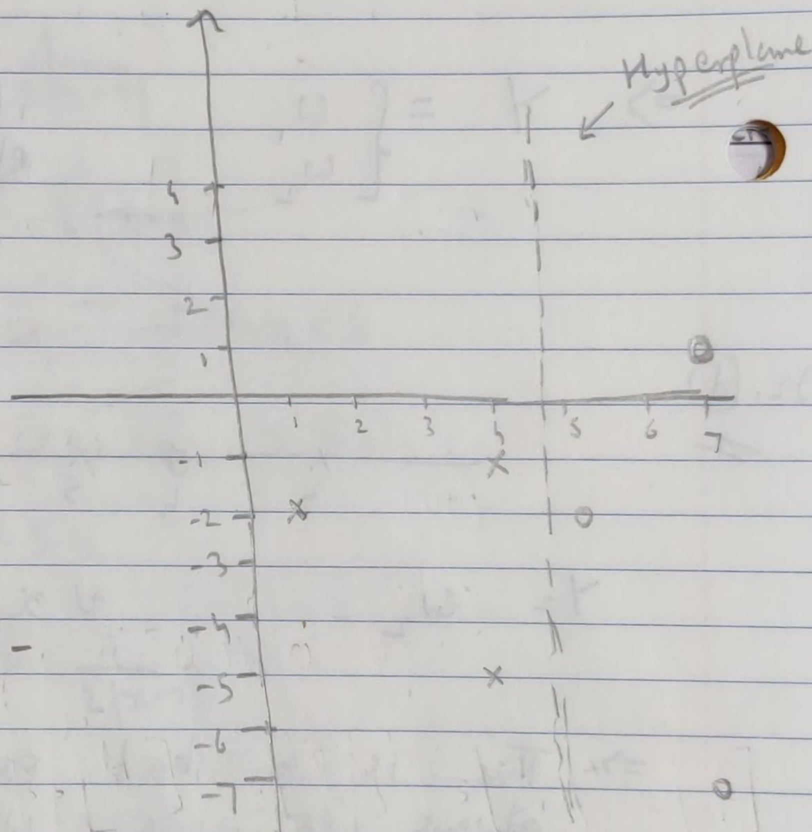
$x_5 = [0, 2]^T$, $z_5 = 1$
$x_6 = [0, -2]^T$, $z_6 = 1$
$x_7 = [-2, 0]^T$, $z_7 = 1$

$u = [\phi_1(x), \phi_2(x)]^T$: $\phi_1(x) = x_2^2 - 2x_1 + 3$
$\phi_2(x) = x_1^2 - 2x_2 - 3$

$u_1 = [1, -2]$

$u_2 = [4, -5]$

$u_3 = [4, -1]$
$u_4 = [5, -2]$
$u_5 = [7, -7]$
$u_6 = [7, 1]$
$u_7 = [7, 1]$



Hyperplane equation :-

$$\boxed{x - 4.5 = 0}$$

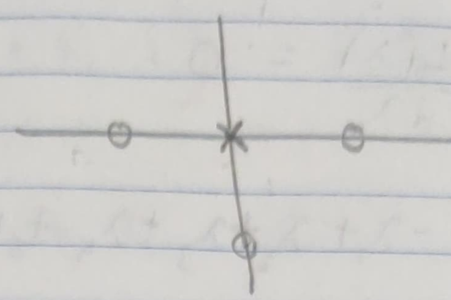
Q4.

$x_1 = [0,0]^T \quad , z_1 = -1$

$x_2 = [1,0]^T \quad , z_2 = 1$

$x_3 = [0,-1]^T \quad , z_3 = 1$

$x_4 = [-1,0]^T \quad , z_4 = 1.$



$$K(x_i, x_j) = (x_i^T x_j + 1)^2$$

a) $$L_D(\underline{\lambda}) = -\frac{1}{2}\sum_{i=1}^{4}\sum_{j=1}^{4} \lambda_i \lambda_j z_i z_j K(x_i, x_j) + \sum_{i=1}^{4} \lambda_i$$

$$= -\frac{1}{2}\sum_{i=1}^{4}\sum_{j=1}^{4} \lambda_i \lambda_j z_i z_j (x_i^T x_j + 1)^2 + \sum_{j=1}^{4} \lambda_i$$

$$= -\frac{1}{2}\Big\{ \lambda_1\lambda_1(-1)(-1)\left([0\ 0]\begin{bmatrix}0\\0\end{bmatrix}+1\right)^2 + 2\lambda_1\lambda_2(-1)(1)\left([0\ 0]\begin{bmatrix}1\\0\end{bmatrix}+1\right)^2$$

$$+2\lambda_1\lambda_3(-1)(1)\left([0\ 0]\begin{bmatrix}0\\-1\end{bmatrix}+1\right)^2 + 2\lambda_1\lambda_4(-1)(1)\left([0\ 0]\begin{bmatrix}-1\\0\end{bmatrix}+1\right)^2$$

$$+\lambda_2\lambda_2(1)(1)\left\{[1\ 0]\begin{bmatrix}1\\0\end{bmatrix}+1\right\}^2 + 2\lambda_2\lambda_3(1)(1)\left\{[1\ 0]\begin{bmatrix}0\\-1\end{bmatrix}+1\right\}^2$$

$$+2\lambda_2\lambda_4(1)(1)\left\{[1\ 0]\begin{bmatrix}-1\\0\end{bmatrix}+1\right\}^2 + \lambda_3\lambda_3(1)(1)\left\{[0\ -1]\begin{bmatrix}0\\-1\end{bmatrix}+1\right\}^2$$

$$+2\lambda_3\lambda_4(1)(1)\left\{[0\ -1]\begin{bmatrix}-1\\0\end{bmatrix}+1\right\}^2 + \lambda_4\lambda_4(1)(1)\left\{[-1\ 0]\begin{bmatrix}-1\\0\end{bmatrix}+1\right\}^2\Big\}$$

$$+\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$$

$$\Rightarrow L_D(\underline{\lambda}) = -\frac{1}{2}\Big\{\lambda_1^2 - 2\lambda_1\lambda_2 - 2\lambda_1\lambda_3 - 2\lambda_1\lambda_4 + 4\lambda_2^2 + 2\lambda_2\lambda_3 + 0$$

$$+ 4\lambda_3^2 + 2\lambda_3\lambda_4 + 4\lambda_4^2\Big\} + \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$$

Q4(b) $\dfrac{dL_D(\lambda)}{d\lambda_1} = 0$

$$\Rightarrow -\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + 1 = 0 \qquad -(i)$$

$$\frac{dL_D(\lambda)}{d\lambda_2} = 0$$

$$\Rightarrow \lambda_1 - 4\lambda_2 - \lambda_3 + 1 = 0 \qquad -(ii)$$

$$\frac{dL_D(\lambda)}{d\lambda_3} = 0$$

$$\Rightarrow \lambda_1 - \lambda_2 - 4\lambda_3 - \lambda_4 + 1 = 0 \qquad -(iii)$$

$$\frac{dL_D(\lambda)}{d\lambda_4} = 0$$

$$\Rightarrow \lambda_1 - \lambda_3 - 4\lambda_4 + 1 = 0 \qquad -(iv)$$

$$\Rightarrow \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -4 & -3 & 0 \\ 1 & -1 & -4 & -1 \\ 1 & 0 & -1 & -4 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}$$

$$\Rightarrow \lambda_1 = 3, \quad \lambda_2 = 1/2, \quad \lambda_3 = 2/3, \quad \lambda_4 = 5/6$$

Q4. (c) $L_p = \frac{\|w\|^2}{2} - \sum_{i=1}^{n} \lambda_i \left(z_i (w x_i + b) - 1 \right)$ ← Primal form.

$$\frac{dL_p}{dw} = 0 \Rightarrow w = \sum_{i=1}^{n} \lambda_i z_i x_i$$

$$\frac{dL_p}{db} = 0 \Rightarrow \sum_{i=1}^{n} \lambda_i z_i = 0$$

Decision Boundary :-

$$\delta_D(x) = w x + w_0$$

$$\Rightarrow \delta_D(x) = \sum_{i=1}^{n} \lambda_i z_i x_i x + w_0$$

Using Kernel trick,

$$\Rightarrow \delta_D(x) = \sum_{i=1}^{n} \lambda_i z_i K(x_i, x) + w_0$$

where, $K(x_i, x)$ = Kernel

Q4 (d) For $\lambda_j \neq 0$,

$$z_j (w^T u_j + w_0) - 1 = 0$$

$$\Rightarrow w^T u_j + w_0 = \frac{1}{z_j}$$

$$\Rightarrow \sum_{i=1}^{n} \lambda_i z_i u_i u_j + w_0 = \frac{1}{z_j}$$

Using Kernel Trick,

$$\Rightarrow w_0 = \frac{1}{z_j} - \sum_{i=1}^{n} \lambda_i z_i K(x_i, x_j)$$

Q4.(e) $w_0^{(1)} = \frac{1}{z_j} - \sum_{i=1}^{n} \lambda_i z_i K(x_i, x_j)$

$$= \frac{1}{(-1)} - \sum_{i=1}^{4} \lambda_i z_i (x_i^T x_j + 1)^2$$

$$= -1 - \left\{ (3)(-1)\left([0\ 0]\begin{bmatrix}0\\0\end{bmatrix} + 1 \right)^2 + \left(\tfrac{1}{2}\right)(1)\left\{ [1\ 0]\begin{bmatrix}0\\0\end{bmatrix} + 1 \right\}^2 + \left(\tfrac{2}{3}\right)(1)\left\{ [0\ -1]\begin{bmatrix}0\\0\end{bmatrix} + 1 \right\}^2 + \left(\tfrac{5}{6}\right)(1)\left\{ [-1\ 0]\begin{bmatrix}0\\0\end{bmatrix} + 1 \right\}^2 \right\}$$

$$= -1 - \left\{ -3 + \frac{1}{2} + \frac{2}{3} + \frac{5}{6} \right\}$$

$$= 0.$$

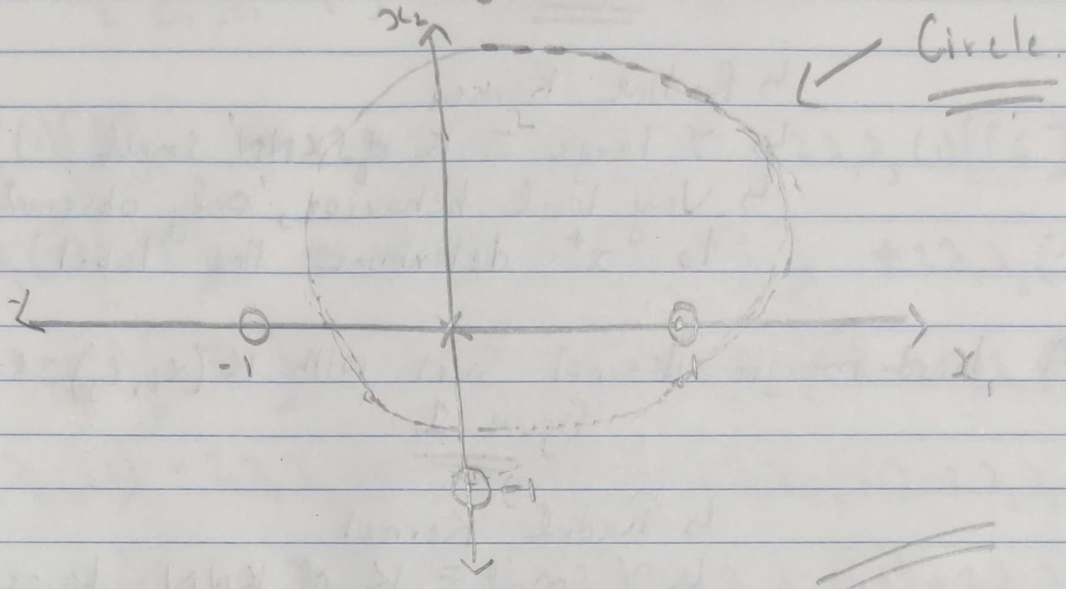
$$g_d(x) = \sum_{i=1}^{4} \lambda_i z_i K(x_i, x) + w_0^{(1)}$$

$$= (3)(-1)\left([0\ 0]\begin{bmatrix}x_1\\x_2\end{bmatrix} + 1 \right)^2 + \left(\tfrac{1}{2}\right)(1)\left\{ [1\ 0]\begin{bmatrix}x_1\\x_2\end{bmatrix} + 1 \right\}^2 + \left(\tfrac{2}{3}\right)(1)\left([0\ -1]\begin{bmatrix}x_1\\x_2\end{bmatrix} + 1 \right)^2 + \left(\tfrac{5}{6}\right)(1)\left\{ [-1\ 0]\begin{bmatrix}x_1\\x_2\end{bmatrix} + 1 \right\}^2 + 0$$

$$= -3 + \frac{1}{2}(x_1 + 1)^2 + \frac{2}{3}(-x_2 + 1)^2 + \left(\frac{5}{6}\right)(-x_1 + 1)^2$$

$$= \frac{-18 + 3x_1^2 + 3 + 6x_1 + 4x_2^2 + 4 - 8x_2 + 5x_1^2 + 5 - 10x_1}{6}$$

$$\Rightarrow \delta_d(x) = \frac{8x_1^2 + 4x_2^2 - 4x_1 - 8x_2 - 6}{6}$$



QS. a) A soft-margin linear SVM with $C = 0.02$:- Figure 4.

↳ Cost function is low.

↳ That means allowing more violations

↳ Margin is wider.

b) A soft-margin linear SVM with $C = 20$:- Figure 3

↳ Cost function is High

↳ Less violations

↳ Margin is narrow

c) A hard-margin kernel SVM with $K(x_i, x_j) = x_i^T x_j + (x_i^T x_j)^2$:- Figure 5

↳ Kernel fun is symmetric in x_i, x_j

↳ Polynomial + Linear factors

Q5. d. A hard-margin Kernel SVM with $K(x_i, x_j) = \exp(-5\|x_i - x_j\|^2)$:-

Figure 6.

↳ Radial Kernel
↳ γ large ≡ K of KNN Small
↳ Very local behavior, only observations very close to x^* determine the label.

e) A hard-margin kernel SVM with $K(x_i, x_j) = \exp(-\frac{1}{5}\|x_i - x_j\|^2)$:-

Figure 1

↳ Radial Kernel
↳ γ Small ≡ K of KNN Large
↳ low variance, high Bias