Assignment 2

Theory Question

Shubham Kumar

2015098

~~A0~~

A1 = In RBF, we try to maximize the margin but also the smoothness of the kernel function effects the complexity of classifier. This effects ~~offer~~ overfitting. In RBF the hyperparameter plays the role for overfitting, size of margin. All this is dependent on hyper-parameter. If we choose small/a worst hyper parameter the RBF kernel would almost tends to be linear. whereas high value of hyperparameter may result in overfitting. Thus, in reality we choose best hyperparameters to avoid over-fitting yet mapping the data into higher dimensional space.

A2 = Given,

$$\left.\begin{aligned} y_1 &= \theta x_1 + b = 1 \\ y_2 &= \theta x_2 + b = -1 \end{aligned}\right] \rightarrow \text{two different classes}$$

for ~~max~~ maximising margin,

$$\max_\alpha L(\alpha) = \max \left(\sum_{n=1}^{N} \alpha_n - \frac{1}{2} \sum \sum a_n a_m t_n t_m x_n^T x_m \right)$$

$$N = 2, \quad \alpha_i \geq 0, \quad \sum_{n=1}^{N} \alpha_n t_n = 0.$$

$$\Rightarrow \alpha_1 - \alpha_2 = 0 \quad \text{—①} \qquad \alpha_1, \alpha_2 \geq 0.$$

$$\therefore \quad L = \alpha_1 + \alpha_2 - \frac{1}{2} \left(\alpha_1^2 t_1^2 x_1^T x_1 - \alpha_1 \alpha_2 t_1 t_2 x_1^T x_2 + \alpha_2 \alpha_1 t_2 t_1 x_2^T x_1 + \alpha_2^2 t^2 x_2^T x_2 \right)$$

$$= \alpha_1 + \alpha_2 - \frac{1}{2}\alpha_1^2 x_1^T x_1 + \alpha_1 \alpha_2 x_1^T x_2 - \frac{1}{2}\alpha_2^2 x_2^T x_2$$

Kernel, $k_{ij} = K(x_i, x_j) = x_i^T x_j$

$$\therefore L = \alpha_1 + \alpha_2 - \frac{1}{2}\alpha_1^2 k_{11} + \alpha_1 \alpha_2 k_{12} - \frac{1}{2}\alpha_2^2 k_{22}$$

$$\frac{\partial L}{\partial \alpha_1} = 1 - \alpha_1 k_{11} + \alpha_2 k_{12} = 0.$$

$$\frac{\partial L}{\partial \alpha_2} = 1 - \alpha_2 k_{22} + \alpha_1 k_{12} = 0$$

equate both,

$$\cancel{1} - \alpha_1 k_{11} + \alpha_2 k_{12} = \cancel{1} - \alpha_2 k_{22} + \alpha_1 k_{12}$$

$$\alpha_2 k_{12} - \alpha_1 k_{12} = \alpha_1 k_{11} - \alpha_2 k_{22}$$

$$k_{12} \underbrace{(\alpha_2 - \alpha_1)}_{\to 0} = \alpha_1 k_{11} - \alpha_2 k_{22}$$

$$\therefore \alpha_1 k_{11} = \alpha_2 k_{22}$$

put

$$\alpha_1 = \frac{\alpha_2 k_{22}}{k_{11}}$$

$$\therefore 1 - \alpha_2 k_{22} + \frac{\alpha_2 k_{22} k_{12}}{k_{11}} = 0$$

$$\alpha_2 = \frac{k_{11}}{k_{22}(k_{11} - k_{12})}$$

put it in α_1,

$$\alpha_1 = \alpha_2 \times \frac{k_{22}}{k_{11}} = \frac{k_{11}}{k_{22}(k_{11} - k_{12})} \frac{k_{22}}{k_{11}}$$

$$= \frac{1}{k_{11} - k_{12}}$$

$$\theta = \sum a_n t_n \phi(x_n) = \frac{1}{k_{11}-k_{11}} x_1 - \frac{1}{k_{22}} \frac{k_{11}}{k_{11}-k_{12}} x_2$$

$$\Rightarrow \boxed{w = \frac{1}{k_{11}-k_{12}} \left(x_1 - \frac{k_{11}}{k_{22}} x_2 \right)}$$

$$b = \frac{1}{N_s} \sum \left(t_n - \sum a_m t_m k_{nm} \right)$$

$$= \frac{1}{2} \left(1 - (a_1 k_{11} - a_2 k_{12}) - 1 - (a_1 k_{21} - a_2 k_{22}) \right)$$

$$= \frac{1}{2} \left(\frac{1}{k_{22}} \frac{k_{11}}{k_{11}-k_{12}} \cdot k_{22} - \frac{1}{k_{11}-k_{12}} \cdot k_{11} \right)$$

$$\boxed{b = 0}$$

Location of maximum margin hyper plane is found with x_1, x_2 which are generalized and have generalize dimensionality

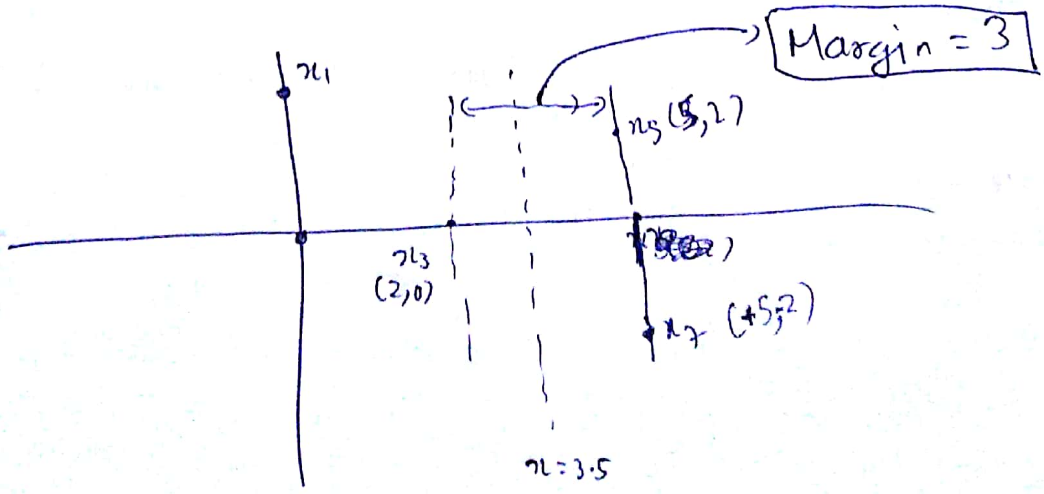
Hence Proved.

A.3:- from graph,

support vectors are $x = 2$ and $x = 5$

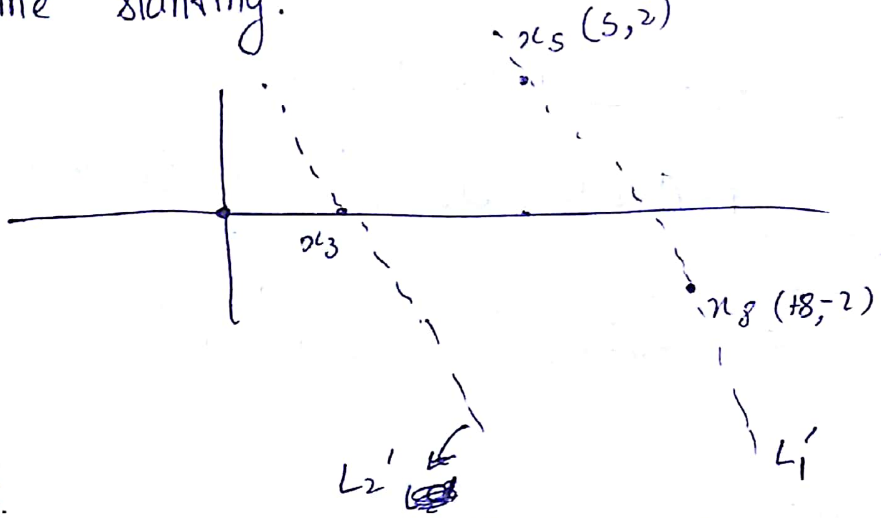
$\therefore$ Decision boundary $= \frac{2+5}{2} = 3.5$

$x = 3.5$



$\therefore$ Max margin = 3

If we remore x_7, the decision boundary will become slanting.



x_5, x_8 are new support vectors.

eq^n of L_1',

$$\frac{y-2}{\cancel{x}} = \overbrace{-\frac{4}{3}}^{\text{slope}}(x-5)$$

$$3y - 6 = -4x + 20$$

$$y = \underbrace{-\frac{4}{3}}_{\text{slope}} x + \underbrace{\frac{26}{3}}_{\text{intercept.}}$$

eq^n of L_2',

$y = mx + c$ $\quad$ ($m = -\frac{4}{3}$ because $L_1' \parallel L_2'$)

$$\frac{y-0}{x-2} = -\cancel{\frac{4}{3}} \quad m = -\frac{4}{3}$$

$$y = -\frac{4}{3}x + \frac{1}{3}$$

Distance b/w lines L_1' and $L_2^* = \dfrac{26/3 - 8/3}{\sqrt{1+(-4/3)^2}}$

$\cong 3.6$

$\therefore$ Margin increases from 3 to 3.6 when we removed x_7.

A4 = XOR

x_1	x_2	z
0	0	0
0	1	1
1	0	1
1	1	0

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

$x = [x_1, x_2]^T$

$x_i = [x_{i1}, x_{i2}]^2$

$\therefore K(x, x_i) = 1 + x_1^2 x_{i1}^2 + 2x_1 x_2 x_{i1} x_{i2} + x_2^2 x_{i2}^2 + 2x_1 x_{i1} + 2x_2 x_{i2}$

$\therefore \phi(x_i) = [1, x_i^2, \sqrt{2}\, x_1 x_2, x_2^2, \sqrt{2} x_1, \sqrt{2} x_2]^T$

We need to maximize,

$L = \Sigma \alpha_i - \frac{1}{2} \alpha^T H \alpha$

$\alpha = [\alpha_1 . \alpha_2\ \alpha_3\ \alpha_4]^T$

$$H = \begin{bmatrix} 9 & 1 & 1 & 1 \\ 1 & 9 & 1 & 1 \\ 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 9 \end{bmatrix}$$

Objective function,

$Q(\alpha) = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}(9\alpha_1^2 - 2\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + \cancel{2\alpha_1\alpha_4} 2\alpha_1\alpha_4 + 9\alpha_2^2 + 2\alpha_2\alpha_3 - 2\alpha_2\alpha_4 + 9\alpha_3^2 - 2\alpha_3\alpha_4 + 9\alpha_4^2)$

Optimizing $Q(\alpha)$

$$9\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4 = 1$$
$$-\alpha_1 + 9\alpha_2 + \alpha_3 - \alpha_4 = 1$$
$$-\alpha_1 + \alpha_2 + 9\alpha_3 - \alpha_4 = 1$$
$$\alpha_1 - \alpha_2 - \alpha_3 + 9\alpha_4 = 1$$

$$\therefore \alpha_{o1} = \alpha_{o2} = \alpha_{o4} = \frac{1}{8}$$

All input vectors are support vector,

$$Q_o(\alpha) = \frac{1}{4}$$

$$\Rightarrow \frac{1}{2}\|w_o\|^2 = \frac{1}{4}$$

~~Optimum~~ $\quad \|w_o\| = \frac{1}{\sqrt{2}}$

Optimum weight vector,

$$w_o = \frac{1}{8}\left[-\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ -\sqrt{2} \\ -\sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ -\sqrt{2} \\ \sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ \sqrt{2} \\ -\sqrt{2} \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix} \right]$$

$$= \begin{bmatrix} 0 \\ 0 \\ -1/\sqrt{2} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$w_o^T \phi(x) = \begin{bmatrix} 0 & 0 & -1/\sqrt{2} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix} = 0$$

$$= 0 + 0 - x_1x_2 + 0 + 0 + 0$$
$$= -x_1x_2 = 0$$

This is the polynomial form of SVM for XOR prob,

~~$x_1 \cdot x_2 = 1$~~

~~$x_1 \cdot x_2 = 0$~~

$x_1 = x_2 = -1 \Rightarrow y = -1$

$x_1 = x_2 = 1 \Rightarrow y = -1$

$$\left.\begin{array}{l} x_1 = -1, 1 \\ x_2 = 1, -1 \end{array}\right\} y = 1$$