ECE449
HW1
2180300806

t	$x_1 = \cos(2t)$	$x_2 = \cos(5t)$	y
1	-0.416	0.284	2.4
2	-0.654	-0.839	1.3
3	0.960	-0.7596	-0.5
4	-0.146	0.4081	1.6
5	-0.839	0.991	2.8

$X = \cancel{[x_1^T, x_2^T]}$ $[1, x_1^T, x_2^T]$, $\theta = [w_0, w_1, w_2]$, $Y = y^T$

Using mean squared error $\Rightarrow J(\theta) = \frac{1}{2}(X\theta - \vec{y})^T (X\theta - y)^T$

By minimizing $\nabla_\theta J(\theta)$. $\cancel{\Rightarrow \theta = X^T X}$ by setting $\nabla_\theta J(\theta) = 0$

$$\cancel{X = \begin{bmatrix} -0.416 & 0.284 \\ -0.654 & -0.839 \\ 0.960 & -0.7596 \\ -0.146 & 0.4081 \\ -0.839 & 0.991 \end{bmatrix}}$$

$\theta = [w_0, w_1, w_2]$ $\theta = (X^T X)^{-1} X^T Y$

$$Y = \begin{bmatrix} 2.4 \\ 1.3 \\ -0.5 \\ 1.6 \\ 2.8 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 5 & -1.0942 & 0.0842 \\ -1.0942 & 2.248 & -1.19 \\ 0.0842 & -1.19 & 2.511 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & -0.416 & 0.284 \\ 1 & -0.654 & -0.839 \\ 1 & 0.960 & -0.7596 \\ 1 & -0.146 & 0.4081 \\ 1 & -0.839 & 0.991 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 0.231 & 0.144 & 0.061 \\ 0.144 & 0.684 & 0.319 \\ 0.0607 & 0.3196 & 0.547 \end{bmatrix}$$

$$\theta = \begin{bmatrix} 0.231 & 0.144 & 0.061 \\ 0.144 & 0.684 & 0.319 \\ 0.0607 & 0.3196 & 0.547 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -0.416 & -0.654 & 0.960 & -0.146 & -0.839 \\ 0.284 & -0.839 & -0.7596 & 0.4081 & 0.991 \end{bmatrix} \begin{bmatrix} 2.4 \\ 1.3 \\ -0.5 \\ 1.6 \\ 2.8 \end{bmatrix}$$

$$\theta = \begin{bmatrix} 1.2495 \\ -1.776 \\ 0.753 \end{bmatrix} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix}$$

$\therefore y = 1.2495 - 1.776(\cos 2(t)) + 0.753 \cos(5(t))$

2

x_1	x_2	y
2	3	A
4	2	B
1	1	A
5	4	B
3	2	A
6	5	B
2	4	A
3	5	A
4	4	B
5	2	B

For $k=1$, Query ~~$(x_{1q}, x_{2q}) = (4,3)$~~

Query $= (4,3)$
$= (x_{1q}, x_{2q})$

x_1	x_2	distance	Included	label
2	3	2	×	A
4	2	1	✓	B
1	1	$\sqrt{13}$	×	A
5	4	$\sqrt{2}$	×	B
3	2	$\sqrt{2}$	×	A
6	5	$2\sqrt{2}$	×	B
2	4	$\sqrt{5}$	×	A
3	5	$\sqrt{5}$	×	A
4	4	1	✓	B
5	2	$\sqrt{2}$	×	B

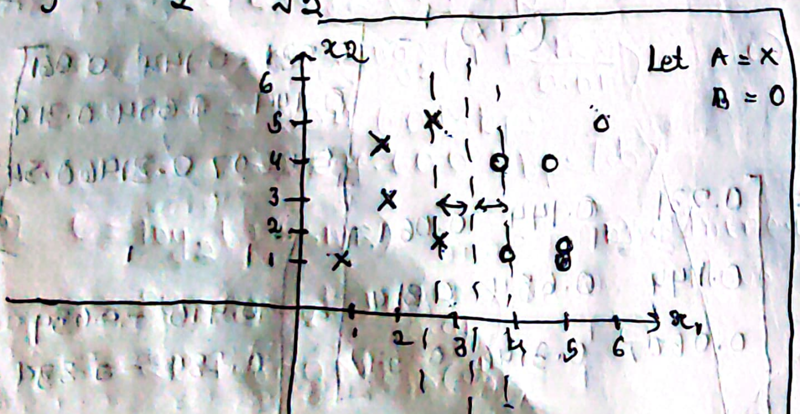
$$\text{distance} = \sqrt{(x_1 - x_{1q})^2 + (x_2 - x_{2q})^2}$$

For $k = 3$,

x_1	x_2	dist	Included	label
2	3	2	✓	A
4	2	1	✓	B
1	1	$\sqrt{13}$	×	A
5	4	$\sqrt{2}$	✓	B
3	2	$\sqrt{2}$	✓	A
6	5	$2\sqrt{2}$	×	B
2	4	$\sqrt{5}$	×	A
3	5	$\sqrt{5}$	×	A
4	4	1	✓	B
5	2	$\sqrt{2}$	✓	B

For $k=1$, prediction = B
For $k=3$, prediction = B
(d)

3



	x_1	x_2	y
	2	3	1
α_2	$\boxed{4}$	$\boxed{2}$	−1
	1	1	1
	5	4	−1
α_5	$\boxed{3}$	$\boxed{2}$	1
	6	5	−1
	2	4	1
α_8	$\boxed{3}$	$\boxed{5}$	1
α_9	$\boxed{4}$	$\boxed{4}$	−1
	5	2	−1

* By inspection x^2, x^5, x^8, x^9 are support vectors.

• Define the Lagrangian SVM ⇒

$$\mathcal{L}(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{n} \alpha_i y_i x_i^T w + b\left(\sum_{i=1}^{n} \alpha_i y_i\right) + \sum_{i=1}^{n} \alpha_i$$

$$= \frac{1}{2}[w_1, w_2]\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} - \Big(\alpha_1 \cdot 1 \cdot [2,3] \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \alpha_2 \cdot (-1) \cdot [4,2]\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} +$$
$$\alpha_3 \cdot (1) \cdot [1,1] \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \alpha_4 \cdot (-1) \cdot [5,4] \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \alpha_5 (1) [3,2] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} +$$
$$\alpha_6 \cdot (-1) \cdot [6,5] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \alpha_7 \cdot (1) \cdot [2,4] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \alpha_8 (1) [3,5] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} +$$
$$\alpha_9 (-1) \cdot [4,4] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \alpha_{10} (-1) [5,2] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 + \alpha_7 + \alpha_8$$
$$+ \alpha_9 + \alpha_{10} + b(\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 + \alpha_5 - \alpha_6 + \alpha_7 + \alpha_8 - \alpha_9 - \alpha_{10}\Big]$$

- By setting the Lagrangian derivative to:

$$\frac{\partial L}{\partial w} = 0 \rightarrow w^* = \sum_{i=1}^{n} \alpha_i y_i x^i$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{n} y_i \alpha_i = 0$$

- To solve the dual problem $\Rightarrow$

$$\alpha_1^*, \dots \alpha_{10}^* = \max_{\alpha_1, \alpha_2 \dots \alpha_{10}} -\frac{1}{2} \sum_{i=1}^{10} \sum_{j=1}^{10} \alpha_i \alpha_j y_i y_j x^{i^T} x^j + \sum_{i=1}^{10} \alpha_i$$

- By calculating the derivate of L with respect to α_i $i \in \{1, 10\}$ and set 0. we get that only $\alpha_2, \alpha_5, \alpha_8, \alpha_9 \geq 0$ which proves the inputs x^2, x^5, x^8, x^9 are the support vectors with 2 from each class.

- Then solving $\frac{\partial L}{\partial w} = 0$ and $\frac{\partial L}{\partial b} = 0$

$$w^* = \begin{bmatrix} -1.7 \\ 0.1 \end{bmatrix} \quad b = 5.6 \quad \text{margin width} = \frac{2}{\|w^*\|} = 1.1744$$

- Decision Rule $\Rightarrow$ $-1.7x_1 + 0.1x_2 + 5.6 \geq 1 = $ 'A' (a)

$-1.7x_1 + 0.1x_2 + 5.6 \leq -1 = $ 'B'