ANAND
2017218

ML - A2 : Theory Questions

Q.3 We can't ~~the~~ solve the problem using linear SVM.
Use the following Kernel : (Polynomial kernel) :

$$K(x, x') = K(1 + x^T x')^2 \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \; x' = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}$$

$$= 1 + x_1^2 x_1'^2 + 2x_1 x_2 x_1' x_2' + x_2^2 x_2'^2 + 2x_1 x_1' + 2x_2 x_2'$$

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

$$\phi(x') = [1, x_1'^2, \sqrt{2}x_1' x_2', x_2'^2, \sqrt{2}x_1', \sqrt{2}x_2']^T$$

★ Treating the problem as an optimal margin classifier problem (SVM :

	Iput vector		Label ($y^{(i)}$)
$x^{(1)}$	-1	-1	-1
$x^{(2)}$	-1	+1	+1
$x^{(3)}$	+1	-1	+1
$x^{(4)}$	+1	+1	-1

★ (K) Kernel Matrix : $K_{ij} = K(x^{(i)}, x^{(j)})$
There are 4 input features,
∴ K is 4x4 matrix.

(i) $\phi(x^{(1)}) = [1, 1, \sqrt{2}, 1, -\sqrt{2}, -\sqrt{2}]^T$
$x^{(1)} = (-1, -1)$

(ii) $\phi(x^{(2)}) = [1, 1, -\sqrt{2}, 1, -\sqrt{2}, \sqrt{2}]^T$
$x^{(2)} = (-1, +1)$

(iii) $\phi(x^{(3)}) = [1, 1, -\sqrt{2}, 1, \sqrt{2}, -\sqrt{2}]^T$
$x^{(3)} = (+1, -1)$

classmate

Date

Page

(ii) $\phi(x^{(4)}) = [1, 1, \sqrt{2}, 1, \sqrt{2}, \sqrt{2}]^T$

$x^{(4)} = (+1, +1)$

Using the formula $K_{ij} = \phi(x^{(i)})^T \cdot \phi(x^{(j)})$, we get:

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

* Minimized lagrange function $W(\alpha) = L(w, b, \alpha)$

$$= \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} y^{(i)} y^{(j)} \alpha_i \alpha_j K_{i,j}$$

$\therefore$ $W(\alpha)$ becomes:

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}(9\alpha_1^2 - 2\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + 2\alpha_1\alpha_4 + 9\alpha_2^2 + 2\alpha_2\alpha_3 - 2\alpha_2\alpha_4 + 9\alpha_3^2 - 2\alpha_3\alpha_4 + 9\alpha_4^2) \quad - (*)$$

Putting each $\frac{\partial W(\alpha)}{\partial \alpha_i} = 0$ we get the following equations:

$$9\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 = 1 \quad - (1)$$

$$-\alpha_1 + 9\alpha_2 + \alpha_3 - \alpha_4 = 1 \quad - (2)$$

$$-\alpha_1 + \alpha_2 + 9\alpha_3 - \alpha_4 = 1 \quad - (3)$$

$$\alpha_1 - \alpha_2 - \alpha_3 + 9\alpha_4 = 1. \quad - (4)$$

On solving (1), (2), (3), (4) using RREF, we get

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1/8 \quad - (5)$$

Putting (5) in (*) we get

$W(\alpha) = \frac{1}{4}$ which means $\frac{1}{2}\|w\|^2 = \frac{1}{4}$

$\therefore \|w\| = 1/\sqrt{2}$

$\therefore$ we get $w = \frac{1}{8}[\phi(x_1) + \phi(x_2) + \phi(x_3) + \phi(x_4)]$

$= [0, 0, -1/\sqrt{2}, 0, 0, 0]^T$.

Now, $w^T \phi(x) = 0$

$$\Rightarrow [0, 0, \frac{-1}{\sqrt{2}}, 0, 0, 0] \begin{bmatrix} 1 \\ x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix} = 0$$

$\Rightarrow \boxed{-x_1x_2 = 0} \rightarrow$ eq^n of hyperplane.

$\therefore$ for $x^{(1)} = (-1, -1)$, $-x_1x_2 = -1 \Rightarrow y = -1$

for $x^{(2)} = (-1, 1)$, $-x_1x_2 = +1 \Rightarrow y = 1$

for $x^{(3)} = (1, -1)$, $-x_1x_2 = +1 \Rightarrow y = 1$

for $x^{(4)} = (+1, +1)$, $-x_1x_2 = -1 \Rightarrow y = -1$

$\therefore$ Solved XOR problem using SVM.

Ans 4. Let us assume $\phi(x)$ is p-dimensional $(\mathbb{R}^2 \rightarrow \mathbb{R}^p)$.

$\therefore \phi(x) = [\phi_1(x), \phi_2(x), \phi_3(x), \ldots, \phi_p(x)]^T$

& $\phi(x') = [\phi_1(x'), \phi_2(x'), \phi_3(x'), \ldots, \phi_p(x')]^T$

Now, $\phi(x, x') = (x^T x' + 1)^2$

$= \cancel{\phi}\ (x^T x')^2 + 2x^T x' + 1$

~~Can be written as~~:

classmate

Date

Page

$$= \left(\sum_{i=1}^{2} x_i x'_i\right)^2 + \left(2\sum_{i=1}^{2} x_i x'_i\right) + 1$$

$$= \sum_{i,j=1}^{2} x_i . x_j . x'_i . x'_j + \sum_{i=1}^{2} (\sqrt{2} x_i)(\sqrt{2} x'_i) + 1$$

$$= (x_1 x_1)(x'_1 x'_1) + (x_1 x_2)(x'_1 x'_2) + (x_2 x_1) + (x'_2 x'_1) + (x_2 x_2)(x'_2 x'_2)$$

$$+ (\sqrt{2} x_1)(\sqrt{2} x'_1) + (\sqrt{2} x_2)(\sqrt{2} x'_2) + 1$$

$$\therefore K(x, x') = [x_1^2, x_1 x_2, x_2 x_1, x_2 x_2, \sqrt{2} x_1, \sqrt{2} x_2, 1] \begin{bmatrix} x_1'^2 \\ x'_1 x'_2 \\ x'_2 x'_1 \\ x_2'^2 \\ \sqrt{2} x'_1 \\ \sqrt{2} x_2^2 \\ 1 \end{bmatrix}$$

$$= [x_1^2, \sqrt{2} x_1 x_2, x_2^2, \sqrt{2} x_1 \sqrt{2} x_2\ 1] \begin{bmatrix} x_1'^2 \\ \sqrt{2} x'_1 x'_2 \\ x_2'^2 \\ \sqrt{2} x'_1 \\ \sqrt{2} x'_2 \\ 1 \end{bmatrix}$$

Comparing with

$$K(x, x') = (\phi(x))^T \phi(x')$$

we get

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

$\therefore$ it contains 6 entries (6-dimension feature space)

Scanned by CamScanner

classmate
Date
Page

Que. 5

(a) What is b for SVM?

★ We know that $w^* = \sum_{i=1}^{m} \alpha_i \, y^{(i)} x^{(i)}$, (m = 10 here)

$$= \sum_{i=1}^{10} \alpha_i y^{(i)} \begin{bmatrix} x_1^{(i)} \\ y_1^{(i)} \end{bmatrix}$$

$$= 0.414 \begin{bmatrix} 4 \\ 2.9 \end{bmatrix} + 0 + 0 + -1.18 \begin{bmatrix} 2.5 \\ 1 \end{bmatrix} + 0 + 0$$

$$+ 1.18 \begin{bmatrix} 3.5 \\ 4 \end{bmatrix} + 0 + (-0.414) \begin{bmatrix} 2 \\ 2.1 \end{bmatrix} + 0$$

$$\Rightarrow w^* = \begin{bmatrix} 2.008 \\ 3.8712 \end{bmatrix}$$

★ $b^* = \left(\left(\max_{i = y^{(i)} = -1} w^{*T} x^{(i)} + \min_{i = y^{(i)} = 1} w^{*T} x^{(i)} \right) / 2 \right) * -1$

The following data points have y = -1 :
$[x^{(3)}, x^{(4)}, x^{(6)}, x^{(8)}, x^{(9)}]$

$w^{*T} x^{(3)} = 11.686$
$w^{*T} (x^{(4)}) = 8.8912$
$w^{*T} (x^{(6)}) = 11.1702$
$w^{*T} x^{(8)} = 6.81$
$w^{*T} x^{(9)} = \underline{12.145}$

$\therefore \max_{i = y^{(i)} = -1} w^{*T} x^{(i)} = 12.145$

& The following data points have y = +1 :
$[x^{(1)}, x^{(2)}, x^{(5)}, x^{(7)}, x^{(10)}]$

$w^{*T} x^{(1)} = 19.258$
$w^{*T} x^{(2)} = 23.512$
$w^{*T} x^{(5)} = 19.517$
$w^{*T} x^{(7)} = 22.508$
$w^{*T} x^{(10)} = \underline{18.714}$

$$\therefore \min_{i=y(i)=1} w^{*T} x^{(i)} = \underline{\underline{18.714}}.$$

$$\therefore b^* = \frac{-(12.145 + 18.714)}{2} = 15.4295$$

$\therefore$ **Ans**: $\boxed{b = 15.4295}$

(b) Identify the support vectors:

Support vectors are those for which $\alpha_i \neq 0$

$\therefore$ S.V. are:

$x^{(1)}, x^{(4)}, x^{(7)}, x^{(9)}$

$= [(4, 2.9), (2.5, 1), (3.5, 4), (2, 2.1)]$

(c) classify point (3,3)

Eq^n of hyperplane $= w^T x + b$

$$= \begin{bmatrix} +2.008 & 3.8712 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix} + 15.4295$$

$$= 6.024 + 11.613 + 15.4295$$

$$= 33.067$$

Since $w^T x + b \geq 0$

$\therefore g(w^T x + b) = 1 = h_{w,b}(x)$

$\therefore$ (3,3) is classified as $\boxed{y = 1}$.

classmate
Date
Page

Que. 6 $K(u,v) = e^{-\|u-v\|^2}$ (taking $\gamma = 1$)

$$= e^{-(u-v)^2} = e^{-u^2 - v^2 + 2uv}$$
$$= \left(e^{-u^2-v^2}\right) \cdot \left(e^{2uv}\right)$$

Using taylor series for e^{2uv}

$$e^{2uv} = 1 + 2uv + \frac{(2uv)^2}{2!} + \frac{(2uv)^3}{3!} + \dots + \infty$$

$$\therefore K = \left(e^{-u^2-v^2}\right)\left(1 + 2uv + \frac{(2uv)^2}{2!} + \frac{(2uv)^3}{3!} + \dots \infty\right)$$

$$= e^{-u^2-v^2}\left(1 + \sqrt{2}u\sqrt{2}v + \left(\sqrt{\frac{2^2}{2!}}u^2 \cdot \sqrt{\frac{2^2}{2!}}v^2\right) + \left(\frac{\sqrt{2^3}}{\sqrt{3!}}u^3 \sqrt{\frac{2^3}{3!}}v^3\right) + \dots + \infty\right)$$

$$= e^{-u^2}e^{-v^2} + \sqrt{2}ue^{-u^2}\sqrt{2}ve^{-v^2} + \sqrt{\frac{2^2}{2!}}u^2e^{-u^2} \cdot \sqrt{\frac{2^2}{2!}}v^2e^{-v^2} + \dots \infty$$

Comparing this equation with $\phi(u)^T\phi(v)$, we get

$$\phi(u) = \lim_{P\to\infty} e^{-u^2}\begin{bmatrix} 1 \\ \sqrt{2}u \\ \sqrt{\frac{2^2}{2!}}u^2 \\ \sqrt{\frac{2^3}{3!}}u^3 \\ \vdots \\ \sqrt{\frac{2^P}{P!}}u^P \end{bmatrix}$$

Scanned by CamScanner

* What can be said about the coefficients of higher order terms of u?

for u^p, the coefficient term is $e^{-u^2}\sqrt{\frac{2^p}{p!}}$

$$\lim_{p\to\infty} \frac{2^p}{p!} = 0.$$

$$\therefore \quad e^{-u^2}\sqrt{\frac{2^p}{p!}} \to 0 \text{ as } p \to \infty.$$

$\therefore$ coefficient of higher order terms are close to 0.