Date:
Page No.

Theory Ques^n

Q4 Linear SVM wont be able to solve it. But we can solve it by introducing kernels and thus making it linearly seperable after applying high-dimentional ~~problem~~ transformation. Or maybe one method could be to introduce a third input,

x_1	x_2	x_3	Y
0	0	0	1
0	1	0	0
1	0	0	0
1	1	1	1

→ now the pts become linearly seperable.

Q2 Assume a dataset ~~of~~ having only 2 datapts. a & b (x_1, x_2) where a (x_1) $\in C_0$ and b (x_2) $\in C_1$, we can say that max margin hyperplane can be given by : argmin $\|\vec{w}\|^2$. s.t : $w^T x_1 + w_0 = 1$
$w^T x_2 + w_0 = -1$

$$\Rightarrow \text{argmin } \|w\|^2 + \alpha_1(w^T x_1 + w_0 - 1) + \alpha_2(w^T x_2 + w_0 + 1)$$

where α_1 & α_2 → lagrange's ~~s~~ multipliers.
taking derivative,

$$w + \alpha_1 x_1 + \alpha_2 x_2 = 0 \Rightarrow$$

$$\text{and } \alpha_1 + \alpha_2 = 0 \Rightarrow \boxed{\alpha_1 = -\alpha_2}$$

$$\Rightarrow w = -\alpha_1(x_1 - x_2)$$

~~$\Rightarrow w_0 =$~~

taking derivative wrt w_0.

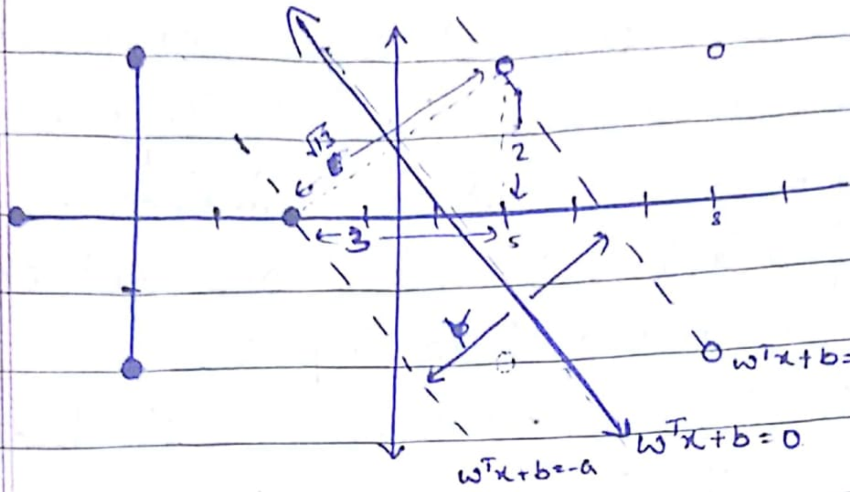
$$2w_0 + w^T x_1 + w^T x_2 = 0$$

$$\Rightarrow w^T(x_1 + x_2) = -2w_0$$

Scanned by CamScanner

since α_1 & α_2 remains unchanged, it shows us that ~~...~~ w_1 and w_2 remains undetermined.

Q3



Hard margin enforce that all pts lie outside the margin. → $\max \gamma = \frac{a}{\|w\|}$

Initial margin = 1.5

~~Fina~~ margin after shifting (ie removing X7) $= \sqrt{13}/2$

∴ Size of max margin ~~changes~~ increases on removing an SVM.

Q4 Since we deal with a very large dataset, so overfitting doesnt occurs, roughly overfitting occurs when complexity of model / training set size is too high, but practically this doesn't happen since normally data we consider is too ~~high~~ big

Q4 Continued:

Either we can use that or we can use kernelization to solve that.

$$K(x, x') = K(1 + x^T x_\delta')^2$$ ← polynomial kernel

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

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

$x' = \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}$

So kernel becomes.

$$[1, x_1^2, \sqrt{2}x_1x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2]^T$$

and same for x'.

where x' can be x'', x'^2, x'^3, x'

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

$$[\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4] \begin{bmatrix} 9 & -1 & -1 & 1 \\ -1 & 9 & 1 & -1 \\ -1 & 1 & 9 & -1 \\ 1 & -1 & -1 & 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

optimal value comes out to be $\frac{1}{8}$ for all α.

$$q(\alpha) \to \frac{1}{4} \Rightarrow \frac{1}{2}\|w\|^2 = \frac{1}{4}$$

$$\|w_0\| = \frac{1}{\sqrt{2}}$$

$$w_0 = \cancel{\frac{1}{8}\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix}} \ \frac{1}{8}[\phi(x_1) + \phi(x_2) + \phi(x_3) + \phi(x_4)]$$

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

Page No.
Date

$$[0, 0, 1/\sqrt{2}, 0, 0, 0] \begin{bmatrix} 1 \\ x^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix} = 0 \Rightarrow \boxed{x_1 x_2 = 0;}$$

$\cancel{\Rightarrow \text{if } x_1 = x_2 = -1}$

~~and~~ $\cancel{\Rightarrow x_1 = x_2 = 1}$

if $\left.\begin{aligned} x_1 = x_2 = -1 \\ x_1 = x_2 = 1 \end{aligned}\right\} y = -1$

if $\left.\begin{aligned} x_1 = -1 \quad x_2 = 1 \\ x_1 = 1 \quad x_2 = -1 \end{aligned}\right\} y = +1$

$\therefore$ Solved.