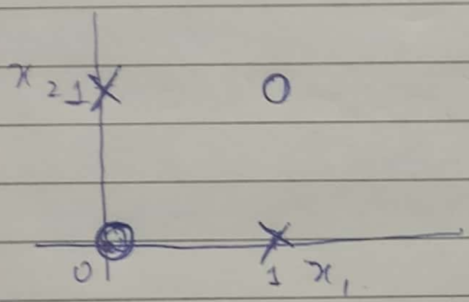
Ans 4. Model XOR using SVM?

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0



X - 1
O - 0

Since the above data is not linearly separable. We cannot use hard margin SVM with a linear kernel.

⇒ We need a kernel to move this into a higher dimension

Let the kernel be $K(x,z) = \phi^T(x)\,\phi(z)$

K $\phi(x) = (x^T z)^2$ ~~$\phi(x) = \{x_1, x_2\}$~~ ~~$K(x, z) = -x z$~~

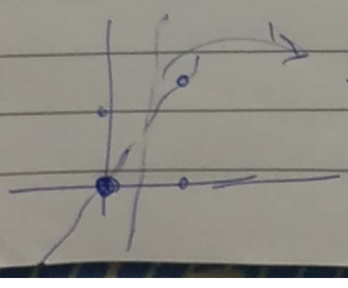
$\Rightarrow \phi(x) = \begin{Bmatrix} x_1^2 \\ x_2^2 \\ 2x_1x_2 \end{Bmatrix}$

~~If pick -1 to represent 0.~~

~~x_1~~	~~x_2~~	~~y~~
~~-1~~	~~-1~~	~~-1~~
~~-1~~	~~1~~	~~1~~
~~1~~	~~-1~~	~~1~~
~~1~~	~~1~~	~~-1~~

x		y
⇒ {0, 0, 0}	⇒	0
{0, 1, 0}	⇒	1
{1, 0, 0}	⇒	1
{1, 1, 2}	⇒	0

⇒ Datapts



→ separating hyperplane.

classmate

Date

Page

Ans 3

Hard margin SVM's separating hyperplane will be at mid-pt of x_3, x_1 & x_7.

$$\Rightarrow \quad x = 2 + \frac{(5-2)}{2}$$

$$= 3.5$$

Margin length $= \frac{2}{\|\theta\|}$ $2 \times 1.5 = 3$.

After removing x_7.

Support vector 1 : passes throug x_1 & x_3

2 : passes throug x_5 & x_7

Distance b/w support vectors $= 3.6$.

Ans 2

Let x_1 (label = 1), x_2 (label = -1) be the two data pts.

SVM. $\min \frac{1}{2}\|\theta\|^2$

s.t. $y(\theta^T x + b) \geq 1$

Applying lagrange multipliers.

$$\min \frac{\|\theta\|^2}{2} + \lambda_1 (y_1(\theta^T x_1 + b) - 1) + \lambda_2 (y_2(\theta^T x + b) - 1)$$

$$= \min \frac{\|\theta\|^2}{2} + \lambda_1 ((\theta^T x_1 + b) - 1) + \lambda_2 (-(\theta^T x_2 + b) - 1)$$

Scanned by CamScanner

Diff. wrt θ &

$$\vec{\theta} + \lambda_1(x_1) + \lambda_2(-x_2) = 0$$

Diff wrt b, we have

$$-\lambda_1 + \lambda_2 = 0$$

$$\Rightarrow \lambda_1 = \lambda_2 = \lambda \text{ (say)}$$

$$\vec{\theta} = \lambda(x_1 - x_2)$$

$\Rightarrow$ $\vec{\theta}$ is independent of dimensionality of space.

Ans! Overfitting depends on the c-factor and not on the space dimensionality.

Thus, gaussian kernel does not lead to overfitting.