For XOR data set

x_1	x_2	Y	
0	0	0	-1
0	1	1	1
1	0	1	1
1	1	0	-1

Here we are using the Kernal function

for $x = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ $x' = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

$$K(x, x') = (x^T x' + 1)^2$$

$$(u_1 v_1 + u_2 v_2 + 1)^2$$

$$= 1 + (u_1 v_1 + u_2 v_2)^2 + 2(u_1 v_1 + u_2 v_2)$$

$$= 1 + u_1^2 v_1^2 + u_2^2 v_2^2 + 2 u_1 v_1 u_2 v_2 + 2u_1 v_1 + 2u_2 v_2$$

So the equivalent transformation

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

for $(x_1, x_2) = (0,0)$

$$\phi[0,0] = [1, 0, 0, 0, 0, 0]^T$$

$$\phi[0,1] = [1, 0, \sqrt{2}, 0, 1, 0]^T$$

$$\phi[1,0] = [1, \sqrt{2}, 0, 1, 0, 0]^T$$

$$\phi[1,1] = [1, \sqrt{2}, \sqrt{2}, 1, 1, \sqrt{2}]^T$$

$$L(\alpha) = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\sum_{i=1}^{4}\sum_{j=1}^{4} \alpha_i \alpha_j y_i y_j K_{ij}$$

S.T $\alpha_i y_i = 0 \Rightarrow \alpha_1 + \alpha_4 = \alpha_2 + \alpha_3$

$\alpha_i \geq 0$

Scanned by CamScanner

inner product = K
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 1 & 4 \\ 1 & 1 & 4 & 4 \\ 1 & 4 & 4 & 9 \end{bmatrix}$$

Our lagrange equation =

$a_1 + a_2 + a_3 + a_4 - \frac{1}{2}\Big(a_1^2 + 4a_2^2 + 4a_3^2 + 9a_4^2 +$

$- a_1a_2 - a_1a_3 + a_1a_4$

$- a_1a_2 - a_1a_3 + a_1a_4$

$+ a_2a_3 + a_2a_3 - 4a_2a_4 - 4a_4a_2 - 4a_3a_4 - 4a_4a_3\Big)$

$= a_1 + a_2 + a_3 + a_4 + \frac{a_1^2}{2} + 2a_2^2 + 2a_3^2 + 9a_4^2 - a_1a_2$

$- a_1a_3 + a_1a_4 + a_2a_3 - 4a_2a_4 - 4a_3a_4$

$\frac{\partial L}{\partial a_1} = 0 \Rightarrow a_1 - a_2 - a_3 + a_4 = 1 \quad (i)$

$\frac{\partial L}{\partial a_2} = 0 \Rightarrow 4a_2 - a_1 + a_3 - 4a_4 = 1 \quad (ii)$

$\frac{\partial L}{a_3} = 0 \Rightarrow 4a_3 - a_1 + a_2 - 4a_4 = 1 \quad (iii)$

$\frac{\partial L}{a_4} = 0 \Rightarrow 9a_4 + a_1 - 4a_2 - 4a_3 = 1 \quad (iv)$

from the following equations

$a_1 = \frac{13}{3} \quad a_2 = \frac{8}{3} \quad a_3 = \frac{8}{3} \quad a_4 = 2$

as we know

$$\omega = \sum_{i=1}^{4} \alpha_i y_i \Phi(x_i)$$

$$-\frac{13}{3}\begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix} + \frac{8}{3}\begin{bmatrix}1\\0\\\sqrt{2}\\0\\1\\0\end{bmatrix} + \frac{8}{3}\begin{bmatrix}1\\\sqrt{2}\\0\\1\\0\\0\end{bmatrix} - 2\begin{bmatrix}1\\\sqrt{2}\\\sqrt{2}\\1\\1\\\sqrt{2}\end{bmatrix}$$

$$= \begin{bmatrix}-1\\\frac{2\sqrt{2}}{3}\\\frac{2\sqrt{2}}{3}\\-2\sqrt{2}\\\frac{2}{3}\\\frac{2}{3}\end{bmatrix}$$

$$g(x_1, x_2) = \omega^T \Phi(x)$$

$$= \begin{bmatrix}-1 & \frac{2\sqrt{2}}{3} & \frac{2\sqrt{2}}{3} & \cancel{-2\sqrt{2}\ \frac{2}{3}\ \frac{2}{3}} & \frac{2}{3} & \frac{2}{3} & -2\sqrt{2}\end{bmatrix}\begin{bmatrix}1\\\sqrt{2}x_1\\\sqrt{2}x_2\\x_1^2\\x_2^2\\\sqrt{2}x_1x_2\end{bmatrix}$$

∴ decision boundary

$$= -1 + \frac{4}{3}x_1 + \frac{4}{3}x_2 + \frac{2}{3}x_1^2 + \frac{2}{3}x_2^2 - 4x_1x_2 = 0$$

Q. Bayes classifiers means classification of datapoints into 2 or more classes based on their probability of occurrence of their attributes in class1 or class2.

Date
Page

Q3 Let N be equal to the total number of data points and N_1 be the data points for ~~test~~ Class 1 and N_2 for ~~test~~ Class 2

We label t as $\frac{N}{N_1}$ for C_1 and $-\frac{N}{N_2}$ for C_2

Then the sum of square error function

$$E = \frac{1}{2} \sum_{n=1}^{N} (w^T x_n + w_0 - t_n)^2$$

$$\frac{\partial E}{\partial w} = 0 \Rightarrow \sum_{n=1}^{N} (w^T x_n + w_0 - t_n) x_n = 0 \quad (1)$$

$$\frac{\partial E}{\partial w_0} = 0 \Rightarrow \sum_{n=1}^{N} (w^T x_n + w_0 - t_n) = 0 \quad (11)$$

from 11 we can compute

$$w_0 = -w^T \bar{m}$$

here $\bar{m} = \frac{1}{N} \sum_{n=1}^{N} x_n$

ie $\bar{m} = \frac{1}{N} (N_1 \bar{m}_1 + N_2 \bar{m}_2)$

Here $\bar{m}_1$ and $\bar{m}_2$ are corresponding mean of dataset belong to class C_1 and C_2 and we can see

$$\sum_{n=1}^{N} t_n = N_1 \frac{N}{N_1} - N_2 \frac{N}{N_2} = 0$$

equation (4') can also be written as

$$\left(S_W + \frac{N_1 N_2}{N} S_B\right) \omega = N(\bar{m}_1 - \bar{m}_2)$$

Here $S_W = \sum_{n \in C_1} (x_n - \bar{m}_1)(x_n - \bar{m}_1)^T + \sum_{n \in C_2} (x_n - \bar{m}_2)(x_n - m_2)^T$

$$S_B = (\bar{m}_2 - \bar{m}_1)(\bar{m}_2 - \bar{m}_1)^T$$

To get the direction of ω
we can see $S_B \cdot \omega$ is always in the direction of $(\bar{m}_2 - \bar{m}_1)$

$$\omega \propto S_\omega^{-1} (\bar{m}_2 - \bar{m}_1)$$

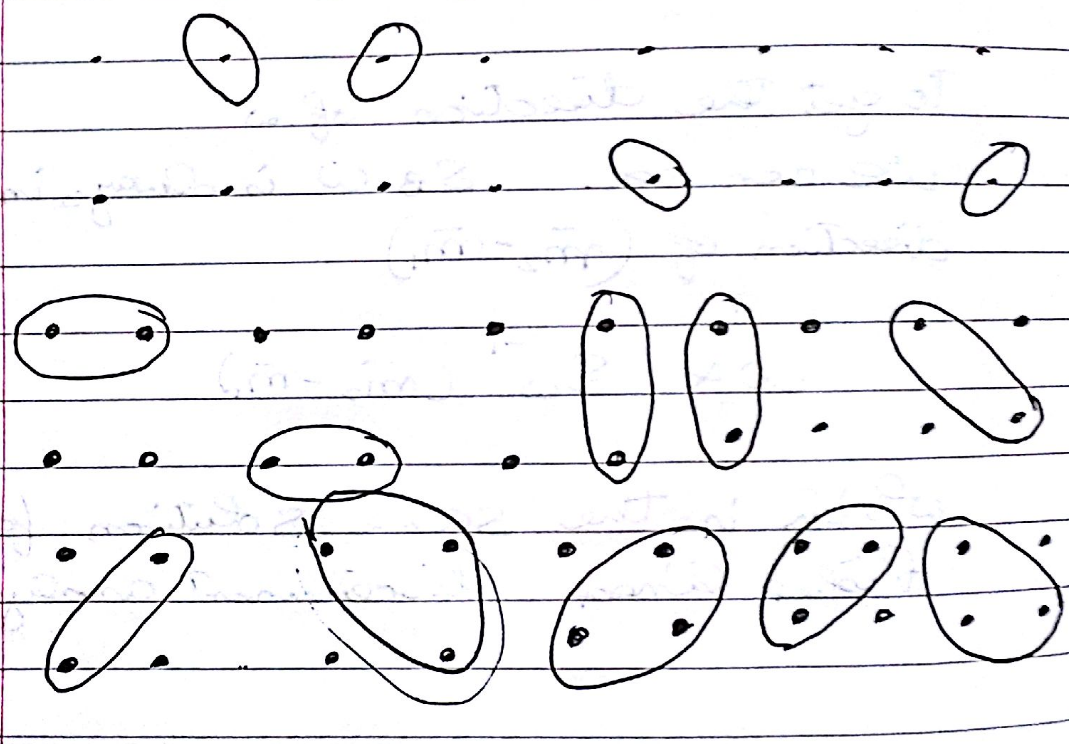
which is the same solution for fisher linear discriminant analysis

Scanned by CamScanner

Q5 Data points $x_1\ (-1,-1)$ $x_2\ (-1,+1)$
$x_3\ (+1,-1)$ $x_4\ (+1,+1)$

a) Linear kernel cannot shatter this

b) Polynomial kernel of degree 2 can shatter this. we take here ~~a possible~~ an ellipse to show shattering



C Gaussian kernel can ~~at~~ shatter it