19UCC023 Mohit Akhouri

IDS Assignment-2

Q1) - - - - - - - - - - - - - -

Soln: We can discuss this SVM with respect to p-dimensional space called the "hyperplane". In a 2 dimensional space, a hyperplane is a "straight line" & in 3D-space, hyperplane is 2D subspace.

For p value > 3, the imagination can be harder, but intuition remains the same:

$$\boxed{\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 0}$$

Equation of hyperplane in 2D-space. For a p-dimensional hyperplane:

equation is:

$$\boxed{\beta_0 + \beta_1 x_1 + \beta_2 x_2 - - + \beta_p x_p = 0}$$

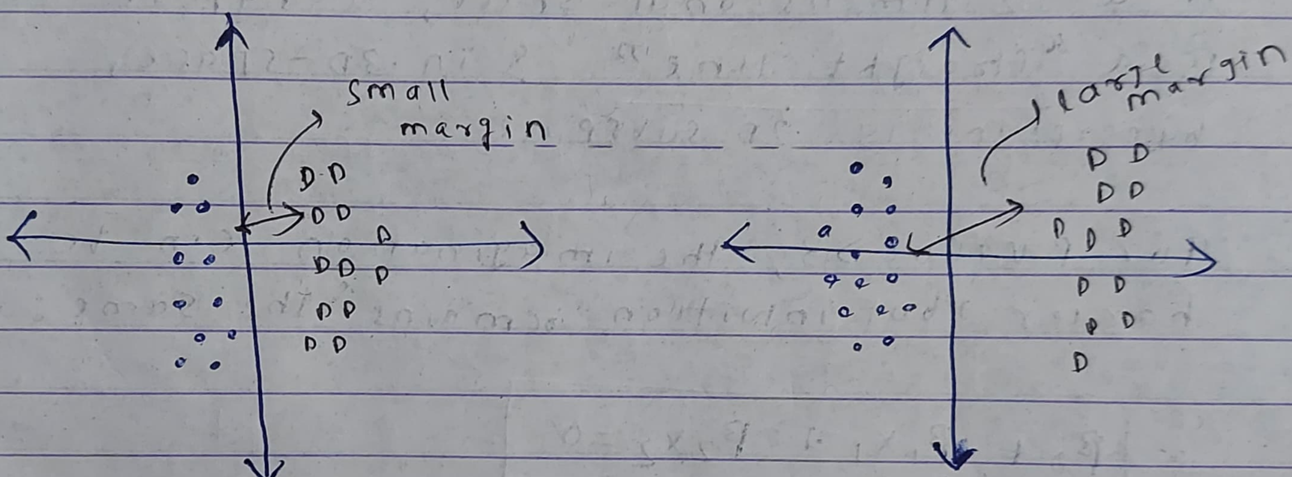
β_0 = intercept

β_1 = first axis

β_2 = second axis

19UCC023 - Mohit Akhouri

Now, we can remove noise by the help of svm. support vectors are data points that are closer to hyperplane and influence the position and orientation of hyperplane. we can use this support vectors to maximize the margin of the classifier.



So the conclusion is that with the help of svm, we can reduce the noise in the input sequence passed through maximal margin classifier.

19UCC023 – Mohit Akhouri

Q2) – – – – – – – – – – – – – – –

Soln: Hierarchial clustering is as follows:
we will find the matrix for different values of j:

for j = 0

	{P1}	{P2}	{P3}	{P4}	{P5}
{P1}	X	16	13	20	8
{P2}	16	X	11	4	24
{P3}	13	11	X	15	13
{P4}	20	4	15	X	18
{P5}	8	24	15	28	X

Formula for complete linkage:

$$\boxed{D(A,B) = \max_{\substack{a \in A \\ b \in B}} (a, b)}$$

$D(A,B) = \{P_4, P_5\}$ - in this example

Now, we can find the clusters as:

$$\boxed{\{\{P_1\}, \{P_2\}, \{P_3\}, \{P_4, P_5\}\}}$$

19UCC023 - Mohit Akhouri

Now we find matrix for j=1:

	{P1}	{P2}	{P3}	{P4, P5}
{P1}	X	16	13	20
{P2}	16	X	11	24
{P3}	13	11	X	15
{P4, P5}	20	24	15	X

Aplying formula for complete linkage, we get: $d(A,B) = \max_{a \in A, b \in B}(a,b) = \max(P_1 P_4, P_1 P_5)$

$\boxed{d(A,B) = 24}$

Now the clusters ~~would~~ would be : {{P1}, {P3}, {P2, P4, P5}}

Now we find matrix for j=2:

	{P1}	{P3}	{P2, P4, P5}
{P1}	X	13	20
{P3}	13	X	15
{P2, P4 P5}	20	15	X

Now, $\boxed{d(A,B) = 20}$, clusters become = {{P1 P2 P4 P5}, {P3}}

19UCC023 - Mohit Akhouri

For $j = 3$:

	$\{P_3\}$	$\{P_1, P_2, P_4, P_5\}$
$\{P_3\}$	X	15
$\{P_1, P_2, P_4, P_5\}$	15	X

Now, the $\boxed{\text{maximum weight} = 15}$

$$\boxed{d(A, B) = 15}$$ = Ans

Now, the clusters become: $\boxed{\{P_1, P_2, P_3, P_4, P_5\}}$

As now there are no more clusters, we will stop our process of "hierarchical clustering" and our final cluster becomes:

$$\boxed{\{P_1, P_2, P_3, P_4, P_5\}}$$ = Ans

19UCC023 - Mohit Akhouri

Q3) - - - - - - - - - - - - - -

Soln: Minimum distance classifier, is a supervised classifier. In this we have to calculate the distance of any given pattern x_0 from each of the classes. Lastly, we will assign x_0 to those classes where distance is minimum.

MDC in mathematical terms:

$$\boxed{D(x_0, C_i) \leq D(x_0, C_j)} \quad \forall j \neq i$$

Now, we will derive MDC from Baye's classifier.

MDC is a special case of Baye's classifier: Conditions that need to be satisfied are:

1. Prior probabilities of each class is same.

$$P_1 = P_2 = - - - - = P_M = 1/M$$

2. Probability density function follows Gaussian distribution.

3. dispersion matrix of each class are same and equal to I.

$$\Sigma_1 = \Sigma_2 = - - - - = \Sigma_m = I_{D \times D}$$

19UCC023 – Mohit Akhouri

Proof: $x \in R^D$

Baye's rule:

$$P_i\, p_i(x) \geq P_j\, p_j(x) \quad \forall j \neq i$$

$$= \frac{1}{M} \frac{1}{(\sqrt{2\pi})^D (\Sigma_i)^{1/2}} \exp\{-1/2\,(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)\}$$

$$\geq \frac{1}{M} \frac{1}{(\sqrt{2\pi})^D |\Sigma_j|^{1/2}} \exp\{-\frac{1}{2}(x - \mu_j)^T \Sigma_j^{-1} (x - \mu_j)\}$$

$$\boxed{\Sigma_i = \Sigma_j = \Sigma_i^{-1} = \Sigma_j^{-1} = \text{all equal to } I}$$

$$\boxed{|I| = 1, \quad I^{-1} = I}$$

from ~~...~~ properties

$$= \exp\{-\frac{1}{2}(x - \mu_i)^T I^{-1} (x - \mu_i)\}$$

$$\geq \exp\{-1/2\,(x - \mu_j)^T I^{-1} (x - \mu_j)\}$$

$$= \boxed{(x - \mu_i)^T I (x - \mu_i) = (x - \mu_i)^T (x - \mu_i)}$$

Taking logarithm both sides:

$$-1/2\,(x - \mu_i)^T (x - \mu_i) \geq -1/2\,(x - \mu_j)^T (x - \mu_j)$$

$$\forall i \neq j$$

19UCC023 - Mohit Akhouri

$$\Rightarrow \boxed{(x-\mu_i)^T (x-\mu_i) \le (x-\mu_j)^T (x-\mu_j)}$$

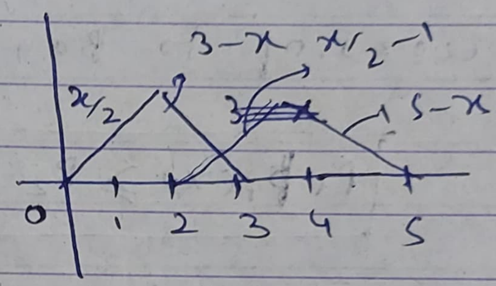
We have derived the ~~MDC~~ condition for minimum distance classifier (MDC).

19UCC023 – Mohit Akbouri

Q4) – – – – – – – – – – – – – –

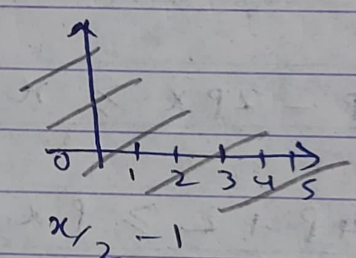
Soln: $M = 2$

$P_1 = P, \quad P_2 = 1 - P$



$$p_1(x) = \begin{cases} x/2, & 0 < x < 2 \\ 3 - x & 2 \le x \le 3 \\ 0, & \text{o.w.} \end{cases}$$

$$p_2(x) = \begin{cases} (x-2)/2 & , 2 < x < 4 \\ 5 - x & , 4 \le x \le 5 \\ 0, & \text{o.w.} \end{cases}$$



Using Baye's classifier:

Case 1: $0 < x < 2$

$P_2(x) = 0, \quad P_1(x) = x/2 \Rightarrow$ pattern x_0 in class 1

case 2: $4 \le x \le 5$

$P_1(x) = 0, \quad P_2(x) = 5 - x \Rightarrow$ Pattern x_0 in class 2

case 3: $3 < x < 4$

$P_1(x) = 0$ & $P_2(x) = (x-1)/2 \Rightarrow$ Pattern x_0 in class 2

19UCC023 – Mohit Akhouri

Case 4: $2 \le x \le 3$

Let us assume pattern x in class 1

According to Baye's classifier:

$$\boxed{P_1 p_1(x) \geq P_2 p_2(x)}$$

$$2P(3-x) \geq (1-P)(x-2)$$

$$6P - 2Px \geq x - 2 - Px + 2P$$

$$4P + 2 \geq Px + x$$

$$x(P+1) \leq 4P + 2$$

$$\boxed{x \leq \frac{4P+2}{1+P}} \text{ - For class 1}$$

$$\Rightarrow \boxed{x > \frac{4P+2}{1+P}} \text{ - For class 2}$$

19UCC023 - Mohit Akhouri

Now, we will find the probability of misclassification:

$$\cancel{E = \sum_{i=1}^{M} \int}$$

Prior probability

$$\boxed{E = \sum_{i=1}^{M} P_i \int_{\Omega_i^c} p_i(x)\,dx}$$

Ω = Total domain of consideration

$\Omega_1 \quad \Omega_2 \quad \ldots \quad \Omega_m$

domain of class 1 — domain of class 'm'

$\Omega_1 = 0$ to $\frac{4P+2}{1+P}$, $\Omega_2 = \frac{4P+2}{1+P}$ to 5

$\Omega_1 = [0, \frac{4P+2}{1+P}]$, $\Omega_2 = (\frac{4P+2}{1+P}, 5]$

$$\cancel{E = P \int_{\frac{4P+2}{1+P}}^{5} P_1(x)\,dx}$$

$$\boxed{\begin{aligned} \Omega_1^c &= \left(\frac{4P+2}{1+P}, 5\right] \\ \Omega_2^c &= \left[0, \frac{4P+2}{1+P}\right] \end{aligned}}$$

19UCC023 - Mohit Alkhouri

Therefore,

$$E = P\int_{\frac{4P+2}{1+P}}^{5} p_1(x)\,dx + (1-P)\int_{0}^{\frac{4P+2}{1+P}} p_2(x)\,dx$$

For $3 \le x < 5$, $p_1(x) = 0$, for $0 \le x < 2$, $p_2(x) = 0$

$$E = P\int_{\frac{4P+2}{1+P}}^{3} (3-x)\,dx + (1-P)\int_{2}^{\frac{4P+2}{1+P}} \left(\frac{x-2}{2}\right)dx$$

$$\cancel{= P\cdot\left[\frac{1}{2}\cdot(3-x)^2\right]}$$

$$= P\cdot\left[3x - \frac{1}{2}x^2\right]_{\frac{4P+2}{1+P}}^{3} + \frac{(1-P)}{2}\cdot\left[\frac{1}{2}x^2 - 2x\right]_{2}^{\frac{4P+2}{1+P}}$$

$$= P\left[\left(9 - \frac{9}{2}\right) - \left(\frac{3(4P+2)}{P+1} - \frac{(4P+2)^2}{2(P+1)^2}\right)\right]$$

$$+ \frac{(1-P)}{2}\left[\frac{1}{2}\left(\frac{4P+2}{1+P}\right)^2 - 2\left(\frac{4P+2}{1+P}\right) - \frac{4}{2} + 4\right]$$

Ans

$$\boxed{E = P\left[9 - \frac{9}{2} - \frac{3(4P+2)}{P+1} - \frac{(4P+2)^2}{2(P+1)^2}\right] + \frac{(1-P)}{2}\left[\frac{1}{2}\left(\frac{4P+2}{1+P}\right)^2 - 2\left(\frac{4P+2}{1+P}\right) + 2\right]}$$

19UCC023 - Mohit Akhouri

Q5) — — — — — — — — — — — — —

Soln: the single perceptron to perform the task of logical AND gate with 3 inputs are as follows:

x_1	x_2	x_3	O/P	$w^T x$	Predicted o/p ($\theta=3$)
0	0	0	0	0	0
0	0	1	0	0	0
0	1	0	0	0	0
0	1	1	0	0	0
1	0	0	0	0	0
1	0	1	0	0	0
1	1	0	0	0	0
1	1	1	1	0	0

$$\boxed{w_i(t+1) = w_i(t) + \eta \delta_i x_i}$$

Let $\eta = 0.5$

$w_1(t+1) = 0 + 0.5\,(1-0)\times 1 = 0.5$

$w_2(t+1) = 0 + 0.5 \times (1-0) \times 1 = 0.5$

$w_3(t+1) = 0 + 0.5 \times (1-0) \times 1 = 0.5$

$$\Rightarrow \boxed{w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}}$$

19UCC023 - Mohit Akhouri

Now the working of Predicted O/P is as follows:

x_1	x_2	x_3	O/P	$w^T X$	Predicted O/P
1	1	1	1	1.5	0

After compairing "1.5 < 3"

Now,

$w_1(t+2) = 0.5 + 0.5(1-0) * 1 = 1$

$w_2(t+2) = 0.5 + 0.5(1-0) \times 1 = 1$

$w_3(t+3) = 0.5 + 0.5(1-0) \times 1 = 1$

$$\boxed{w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}$$

Now the working of Predicted O/P is as follows:

x_1	x_2	x_3	O/P	$w^T X$	Predicted O/P
1	1	1	1	3	1

after comparing 3 = 3, here $\boxed{\text{error} = 0}$

19UCC023 - Mohit Akhouri

As we can see, we cannot find ~~$w_T x$~~ other inputs for which $w^T x$ can exceed max. value 2.

The values for which $w^T x$ can be at most 2 are not misclassified.

We came to the conclusion that with the random values: $\theta = 3$ & $\eta = 0.5$, we can correctly classify logical AND gate with 3 inputs in 3 iterations of single perceptron.

Q6)

19UCC023 - Mohit Akhouri

Q6) - - - - - - - - - - - - -

(a) - - - - - - - - - - - - -

Soln:

directly density reachable = Let p & q be two points in S & P. Let them be directly density reachable from q w.r.t. Eps and minPts iff:

① $p \in Neps(q)$
② $|Neps(q)| \geq minpts$

Density reachable = P is density reachable from q wrt Eps and minpts if there exists a class of points $P_1, P_2 -- P_n$ where $P_1 = q$ & $P_n = P$ and P_{i+1} is density reachable from P_i.

Density connected = P is density connected from q if there exists a point o from where p & q are directly reachable.

19UCC023 - Mohit Akhouri

(b) - - - - - - - - - - - - -

Soln: The Algorithm for DBSCAN clustering technique is as follows:

Take a pattern P from dataset and find it's neighbourhood:

Now check the condition:

if $\boxed{|N_{eps}(P)| \geq minpts}$, then P is in a dense region so it is a core point and it must be in a cluster according to DBSCAN algorithm, otherwise pattern P is in the noise point (border point).

(c) - - - - - - - - - - - - - -

Soln: Advantages of DBSCAN algorithm are:

(i) DBSCAN is not susceptible to noise
(ii) We can obtain non-convex clusters
(iii) DBSCAN calculates the clusters by itself.

19UCC023 - Mohit Akhouri

Disadvantages of DBSCAN algorithm ~~algorithm~~ are:

(i) Quality of DBSCAN algo. depends on distance measure.

(ii) If dataset is too sparse or density varies, DBSCAN fails to identify cluster.

(iii) DBSCAN is sensitive to clustering parameters "minpts" and "EPS".