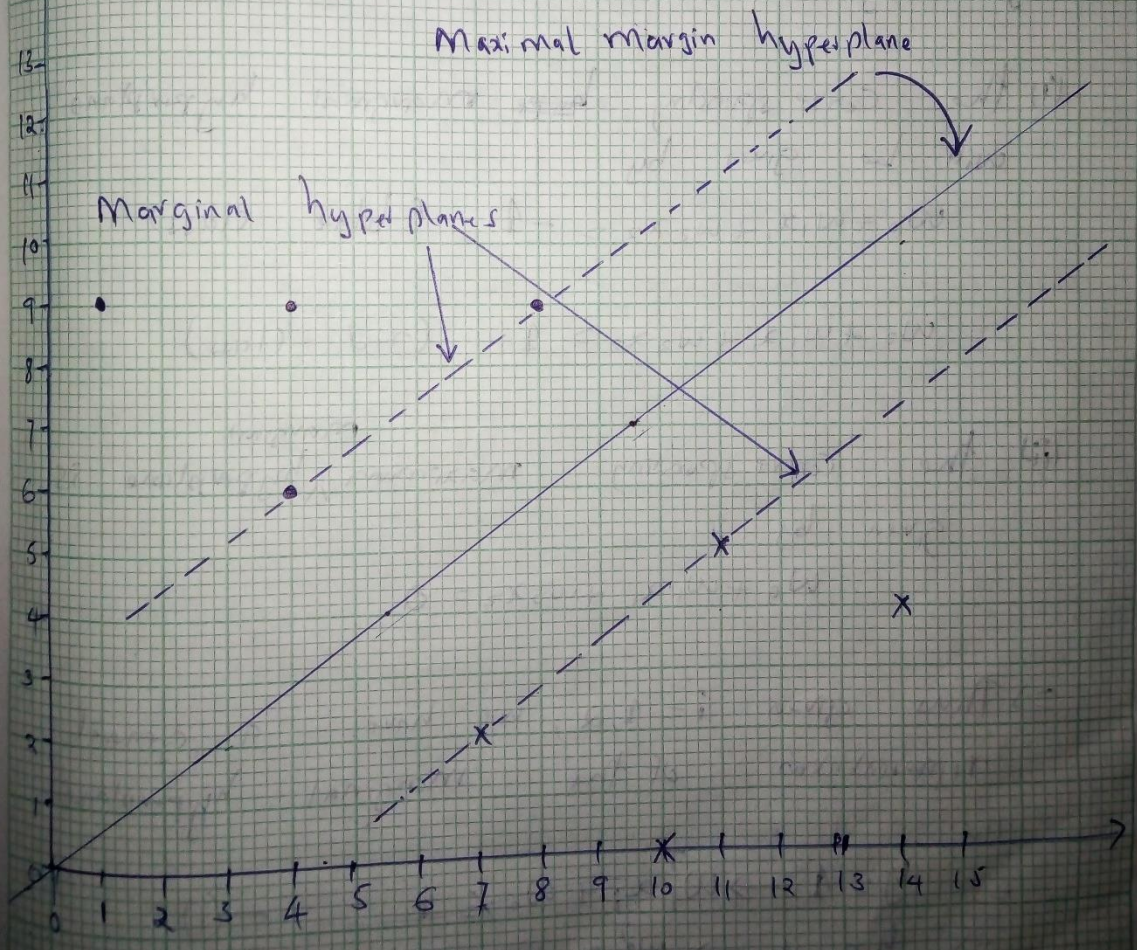
a) Draw Data points and sketch the maximum-margin hyperplane and marginal hyperplanes

Note
X - red class
● - Blue class



Write down the margin ρ

i.e

distance from the decision boundary to the margin boundary

Soln.

Thus we say that for a general datapoint given by (x, y) co-ordinates, then we have

(i) the corresponding ~~hyper~~ marginal hyperplane will be given by

$$w_0 + w_1 x_1 + w_2 x_2 = -1 \quad \text{(blue class)}$$

$$w_0 + w_1 x_1 + w_2 x_2 = 1 \quad \text{(red class)}$$

(ii) the corresponding maximum marginal hyperplane is given by

$$w_0 + w_1 x_1 + w_2 x_2 = 0.$$

=> Thus given $i = 1, 2$, we have a general representation of the marginal hyperplanes as

=> $y = 1 \rightarrow w_i^T x_i + w_0 \leq 1$ —(i)

=> $y = -1 \rightarrow w_i^T x + w_0 \geq 1$ —(ii)

=) Therefore, Combining equations (i) and (ii) we have

$$y(w_i^T x_i + w_0) \geq 1 \qquad \forall i = 1, 2$$

⇒ Finally upon formulation of this function we have the optimal minimization of the margin (ρ) function given as below

$$\text{Minimize}_{w_i, w_0} \quad \frac{1}{2} \|w\|^2$$

$$\text{S.t} \quad y(w_i^T x_i + w_0) \geq 1 \quad \forall (x, y)$$

with $i = 1, 2$

∴ Thus the above constraint will have to hold for all training data points (x, y) provided

(b) Choose 3 support vectors, write system of equations and solve for w_0, w_1, w_2.

Soln.

We choose $s_1 = \begin{pmatrix} 8 \\ 9 \end{pmatrix}$

$\Rightarrow \tilde{s}_1 = \begin{pmatrix} \\ \end{pmatrix} = ?$

So we are going to augment each vector with a bias input of 1 as below

$\Rightarrow \tilde{s}_1 = \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix}$

$\Rightarrow s_2 = \begin{pmatrix} 7 \\ 2 \end{pmatrix} \Rightarrow \tilde{s}_2 = \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}$

$\Rightarrow s_3 = \begin{pmatrix} 11 \\ 5 \end{pmatrix} \Rightarrow \tilde{s}_3 = \begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix}$

Forming the system of equations

(i) $\rightarrow w_0 \tilde{s}_1 \tilde{s}_1 + w_1 \tilde{s}_2 \tilde{s}_1 + w_2 \tilde{s}_3 \tilde{s}_1 = -1$ (for negative class)

(ii) $\rightarrow w_0 \tilde{s}_1 \tilde{s}_2 + w_1 \tilde{s}_2 \tilde{s}_2 + w_2 \tilde{s}_3 \tilde{s}_2 = +1$ (positive class)

(iii) $\rightarrow w_0 \tilde{s}_1 \tilde{s}_2 + w_1 \tilde{s}_2 \tilde{s}_3 + w_2 \tilde{s}_3 \tilde{s}_3 = +1$ (positive class

Substituting $\bar{s}_1$, $\bar{s}_2$ and $\bar{s}_3$ into equations (i), (ii) and (iii) we have

$$w_0 \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix}\begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix} + w_1 \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}\begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix} + w_2 \begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix}\begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix} = -1 \quad \text{—(i)}$$

$$w_0 \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix}\begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix} + w_1 \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}\begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix} + w_2 \begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix}\begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix} = 1 \quad \text{—(ii)}$$

$$w_0 \begin{pmatrix} 8 \\ 9 \\ 1 \end{pmatrix}\begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix} + w_1 \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}\begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix} + w_2 \begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix}\begin{pmatrix} 11 \\ 5 \\ 1 \end{pmatrix} = 1 \quad \text{—(iii)}$$

Simplifying (i), (ii) and (iii) we have that

$$146 w_0 + 75 w_1 + 134 w_2 = -1$$

$$75 w_0 + 54 w_1 + 88 w_2 = 1$$

$$134 w_0 + 88 w_1 + 147 w_2 = 1$$

Solving for w_0, w_1 and w_2 we have

In matrix form we have that

$$\left(\begin{array}{ccc|c} 146 & 75 & 134 & -1 \\ 75 & 54 & 88 & 1 \\ 134 & 88 & 147 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 146 & 75 & 134 & -1 \\ 0 & -2259 & -2798 & -221 \\ 0 & -2798 & -3506 & -280 \end{array}\right)$$

$$\Rightarrow \left(\begin{array}{ccc|c} 146 & 75 & 134 & -1 \\ 0 & -2259 & -2798 & -221 \\ 0 & 0 & 91250 & 14162 \end{array}\right)$$

$$\Rightarrow 91250 w_2 = 14162$$

$$\Rightarrow w_2 = 0.1552$$

$$\Rightarrow -2259 w_1 - 2798(0.1552) = -221$$

$$w_1 = -0.0944$$

$$146 w_0 + 75(-0.0944) + 134(0.1552) = -1$$

$$w_0 = -0.1008$$

(c) (i) Removing $X_1 = (7,2)$

Ans

Yes, the solution will change since the point $(7,2)$ is used as s_2 in the part (b) to find w_0, w_1, w_2,

⇒ This means that we will have to alter s_2 to be given by $(4,6)$ which will change the solution.

(ii) $X_7 = (4,9)$

Ans

Removing $(4,9)$ will not change the solution since the point is not considered as any of the support vector

⇒ Thus, its removal will have no impact on the out come of the solution i.e w_0, w_1 and w_2

(iii) $X_1 = (7,2)$ and $X_2 (10,0)$

Ans

Removing both X_1 and X_2 will change the solution since X_1 has been used as one of the 3 support vectors

=) Thus, removing X_1 implies that we also alter the choice of S_2 which changes the corresponding output respectively

(iv) $X_1 = (7,2)$ and $X_3 = (11,5)$

Ans

This definately changes the solution since both points X_1 and X_3 has been used as support vectors, thus removing them results into a corresponding changes in S_2 and S_3 which drastically changes solution of w_0, w_1, w_2

(v) $X_3 = (11,5)$ and $X_6 = (4,6)$

Ans

Removing X_3 and X_6 will change the solution since X_3 is used as a support vector thus changing it, results into a change in the solution of w_0, w_1, w_2

(d) Adding point(s) to the data set

(i) (5, 0) with class = red

Ans

This will not change the solution since this point will be classified below the lower marginal hyperplanes and thus will not have any impact in chosing the support vectors

(ii) (2, 0) with class = red

Ans

Yes, this changes the solution (w_0, w_1, w_2) since the decision of the marginal hyperplanes will change to accomodate the inclusion of the point (2, 0),

=> This corresponds to a change in choice of the support vectors, resulting to change in solution.

(iii) (0,3) with class = blue

ANS

No, this will have no change in the solution since the point lies within the upper marginal hyperplane and thus has no impact on the choice of the support vectors S_1, S_2, S_3

(iv) (9,3) with class = red and (5,8) with class = blue

ANS

No, this will not change the solution since the points will be clustered within the frames of both the marginal hyperplanes, thus no impact on the support vectors

(v) (8,4) = red, (5,6) => blue

ANS

Yes, this changes the solution because both points will alter the decision of the marginal hyperplane, corresponding to change in the choice of the support vectors

=> This results to a corresponding change in w_0, w_1, w_2

THANK YOU!!!