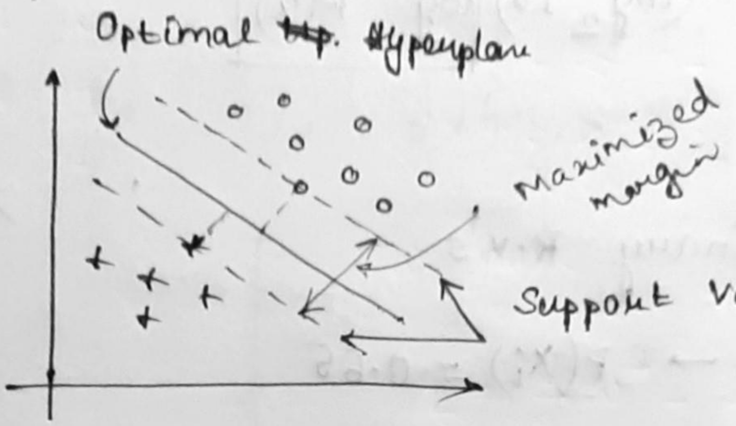
Support Vector Machines

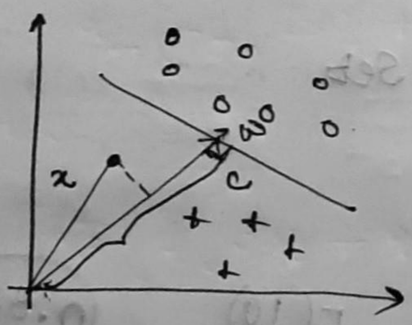
↳ A supervised learning machine where we try to find a hyper-plane that best seperates the two classes



Margin: Distance between the hyper-plane and observation closest to hyperplane / the support vectors

→ The best hyper-plane is the plane that has maximum distance from both classes

Mathematical Intuition



- Assume vector '$\vec{x}$' and a vector 'ω' perpendicular to hyperplane
- Distance of vector 'ω' from origin is 'c'

$\therefore$ $\vec{x} \cdot \vec{w} = c$ (Point on decision boundary)

$\vec{x} \cdot \vec{w} > c$ (+ve class)

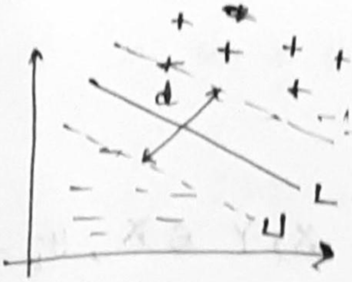
$\vec{x} \cdot \vec{w} < c$ (−ve class)

Decision Rule to classify

$$\vec{x} \cdot \vec{w} - c \geq 0$$

Putting $-c = b$, $\vec{x} \cdot \vec{w} + b \geq 0$

$$y = \begin{cases} +1, & \vec{x} \cdot \vec{w} + b \geq 0 \\ -1, & \vec{x} \cdot \vec{w} + b < 0 \end{cases}$$



Aim of SVM

Maximize the margin / maximize distance (d)

Constraint for Maximizing Margin

$$w_1 x_1 + w_2 x_2 + b = 0$$

$$x_2 = -\frac{b}{w_2} - \frac{w_1}{w_2} x_1$$

Intercept $= -b/w_2$

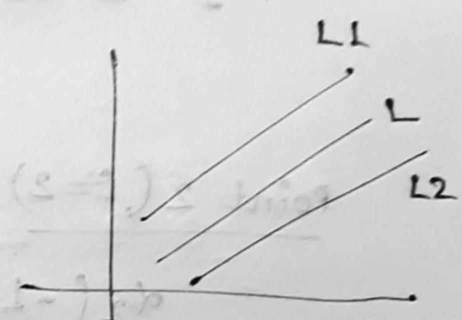
$\vec{w} \cdot \vec{x} + b = 0$

$w_1 x_1 + w_2 x_2 + b =$

L1 : $\vec{w} \cdot \vec{x} + b = 1$

Intercept $= \frac{-(1-b)}{w_2}$

L2: $\vec{w} \cdot \vec{x} + b = \frac{+(1-b)}{w_2}$.



L1 : $\vec{w} \cdot \vec{x} + b = 1$

L2 : $\vec{w} \cdot \vec{x} + b = -1$

} Considering -ve slope, $w_1, w_2 \rightarrow$ +ve

$$\therefore \; y_i(\vec{w} \cdot \vec{X} + b) \geq 1$$

$\vec{w} \cdot \vec{X} + b \leq -1$, when $y_i = -1$

$\vec{w} \cdot \vec{X} + b \geq 1$, when $y_i = +1$

$\underset{(w^*, b^*)}{\text{argmax}}$, $\frac{2}{\|w\|}$ such that $y_i(\vec{w} \cdot \vec{X} + b) \geq 1$

Numerical

Q)

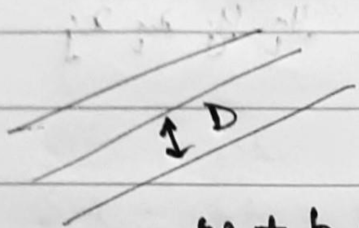
x	class
1	X
2	X
4	Y

unknown: w; b; α_i, $J(w)$

$$wx + b \le -1, \quad x \in X$$

$$wx + b \ge 1, \quad x \in Y$$

$$\boxed{wx + b = 0}$$ (decision boundary)



$$D = \frac{\|w\|^2}{2}$$

$$w + b + 1 = 0$$

$$2w + b + 1 = 0$$

$$\cancel{J(w)} \quad L = \frac{1}{2} w^2 - \sum_i \alpha_i \left(y_i (wx + b) - 1 \right)$$

$$y_i = \{-1, +1\}$$

$$L = \frac{w^2}{2} - \alpha_1 \left(-1(w + b + 1) \right) - \alpha_2 \left(-1(2w + b + 1) \right) - \alpha_3 \left(+1(4w + b - 1) \right)$$

$\frac{1}{2} w^2$

$$L = \frac{w^2}{2} - \alpha_1(-w - b - 1) - \alpha_2(-2w - b - 1) - \alpha_3(4w + b - 1)$$

$$\frac{\delta L}{\delta w} = 0, \quad \frac{\delta L}{\delta b} = 0, \quad \frac{\delta L}{\delta \alpha_i} = 0$$

$$\frac{\partial L}{\delta \alpha_1} = w + b + 1, \quad \frac{\delta L}{\delta \alpha_2} = 2w + b + 1$$

$$\frac{\delta L}{\delta \alpha_3} = -4w - b + 1, \quad \frac{\delta L}{\partial b} = \alpha_1 + \alpha_2 - \alpha_3$$

$$\frac{\delta L}{\delta w} = w + \alpha_1 + 2\alpha_2 - 4\alpha_3$$

$w + \alpha_1 + 2\alpha_2 - 4\alpha_3 = 0$ — ①

$\alpha_1 + \alpha_2 - \alpha_3 = 0$ — ②

$w + b + 1 = 0$ — ③ $\qquad w = 0, \; b = -1$

$2w + b + 1 = 0$ — ④ $\qquad$ Does to satisfy 5

$-4w - b + 1 = 0$ — ⑤ $\qquad w = 1, \; b = -3$

$\cancel{-b + 3 + 1} \qquad \cancel{-2w + 2 = 0}$ $\qquad$ Does to satisfy ②

$\alpha_1 = 0, \quad \alpha_2 = \alpha_3, \quad w = 2\alpha_3$

$w = 1, \; b = -3, \quad \alpha_3 = 1/2 = \alpha_2, \quad \alpha_1 = 0$

Two-Feature Problem of SVM

ex)

X	Y	Class
1	3	X
2	4	X
3	1	Y

$$w_1 x_1 + w_2 x_2 + b = 0$$

$$w_1 x + w_2 y + b \leq -1 \quad x_1, x_2 \in X$$

$$w_1 x + w_2 y + b \geq +1 \quad x_1, x_2 \in Y$$

Y_i

$\downarrow$

$$-1(w_1 + 3w_2 + b + 1)\, \alpha_1$$
$$-1(2w_1 + 4w_2 + b + 1)\, \alpha_2$$
$$+1(3w_1 + w_2 + b - 1)\, \alpha_3$$

$$L = \frac{1}{2}w^2 - \alpha_1(w_1 + 3w_2 + b + 1) - \alpha_2(2w_1 + 4w_2 + b + 1) + \alpha_3(3w_1 + w_2 + b - 1)$$

$$\boxed{\frac{\delta L}{\delta w} = w - \alpha_1(1,3)^T - \alpha_2(2,4)^T + \alpha_3(3,1)^T}$$

$$\frac{\delta L}{\delta \alpha_1} = -(w_1 + 3w_2 + b + 1)$$

$$\frac{\delta L}{\delta b} = (-\alpha_1 - \alpha_2 + \alpha_3)$$

$$\frac{\delta L}{\delta \alpha_2} = \cancel{2w_1} = -2w_1 - 4w_2 - b - 1$$

$\{ w^T w = w^2 \}$

$$\frac{\delta L}{\delta w} = w - \alpha_1 (1,3)^T - \alpha_2 (2,4)^T + \alpha_3 (3,1)^T = 0$$

$$w_1 = -\alpha_1 - 2\alpha_2 + 3\alpha_3$$

$$w_2 = -3\alpha_1 - 4\alpha_2 + \alpha_3$$

$$\frac{\delta L}{\delta \alpha_3} = 3w_1 + w_2 + b - 1 = 0$$

$$\frac{\delta L}{\delta b} = -\alpha_1 - \alpha_2 + \alpha_3 = 0$$

$$\frac{\delta L}{\delta \alpha_1} = w_1 + 3w_2 + b + 1 = 0$$

$$\frac{\delta L}{\delta \alpha_2} = 2w_1 + 4w_2 + b + 1 = 0$$

$$2w_1 + 4w_2 + b + 1 = 0$$
$$3w_1 + w_2 + b - 1 = 0$$

$\cancel{\alpha_2 = \alpha_3}$ $\qquad$ $\cancel{-2\alpha_2 + 3\alpha_3 = w_1}$

$\cancel{-4\alpha_2 + \alpha_3 = w_2}$

$w = \frac{1}{2}, \; w_2 = -1, \; b = 0$