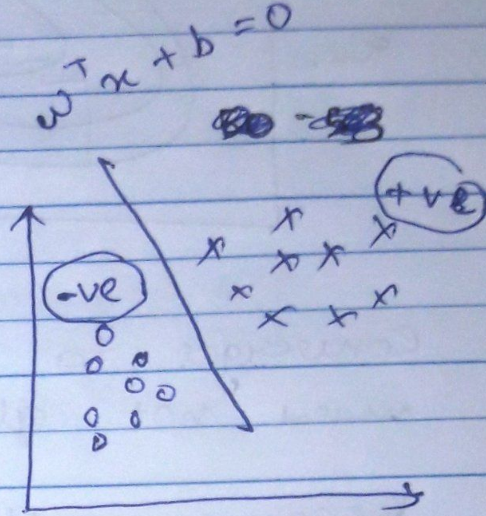
(02) SVM - Formulating Objective :

Binary classification



Find optimal hyperplane : which seperate data (best)

$x_i \in R^n$ ← n dimension

Dataset :

feature

$X = \{ x_1, x_2, \dots x_m \}$

m examples

Labels :

m labels

$Y = \{ y_1, y_2 \dots y_m \}$

$y_{(i)} \in \{-1, +1\}$

Key idea of SVM :

- Seperate data with maximum margin
- Hyperplane's dimension $= (n-1)$

$$= w^T x + b = 0$$

w = vector

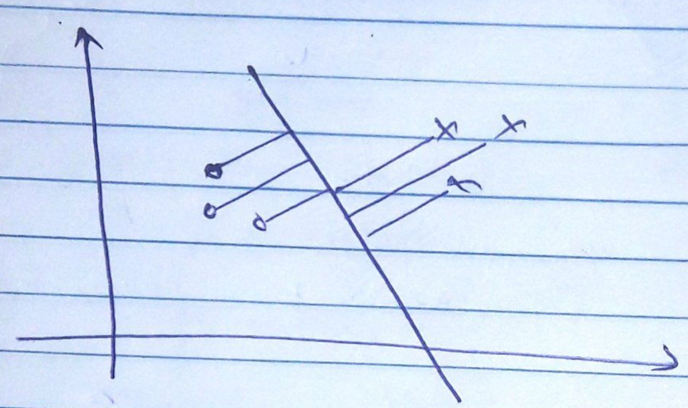
b = bias or intercept term

$$w^T x + b = 0$$

$$= [\underbrace{\theta_0}_{b} + \theta_1 x_1^{(1)} + \theta_2 x_1^{(2)} + \theta_3 x_1^{(3)} \cdots]$$

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- <u>Optimal</u> <u>Hyperplane</u> :



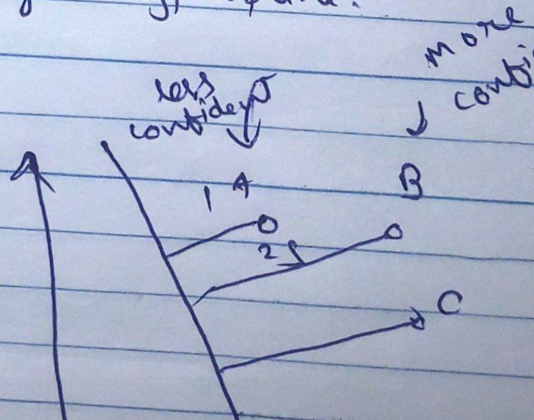
(a) $w^T x + b > 0$ if $x^{(i)} \in$ +ve class

(b) $w^T x + b < 0$ if $x^{(i)} \in$ -ve class

Then it is called Seperating hyperplane.

- It is should be Correct
- And Confident

Confident hyperplane



$$d_1 = w^T x_{(A)} + b > 0$$

$$d_2 = w^T x_{(B)} + b >> 0$$

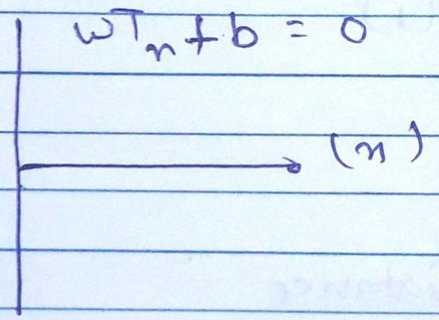
$$\Rightarrow \quad y_{pred} = g(w^T x + b)$$

where $g(z) = +1$ if $z \geq 0$

$g(z) = -1$ if $z < 0$

No Probability prediction.

- Distance of point from line:

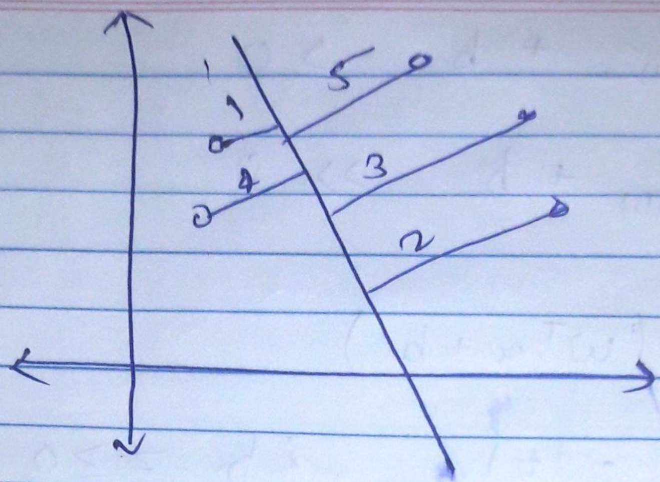


$$= \frac{w^T x + b}{\sqrt{w_1^2 + \cdots w_n^2}}$$

$$= \frac{w^T x + b}{\|w\|_2}$$

L-2 norm

<u>Goal:</u>
To maximize the minimum distance of point from the hyperplane.



$$\gamma^{(i)} = \frac{w^T . x^{(i)} + b}{\|w\|_2}$$

Minimum Distance:

$$\gamma = \min_{i=1 \ldots m} \gamma^{(i)}$$

Goal:
(a) Maximize minimum distance

$w^T x + b > 0$ if x +ve (Class)

$w^T x + b < 0$ if -ve

All the points should have atleast γ (gamma) distance

<u>Optimisation / SVM objective</u>

$\max_{\gamma, w b} \gamma$ (maximize γ)

such that

$$\underbrace{y^{(i)}\left(w^T x^{(i)} + b\right)}_{\text{Absolute distance}} \geq \gamma$$

for all $i = 1 \ldots m$

- Hard to solve this problem.