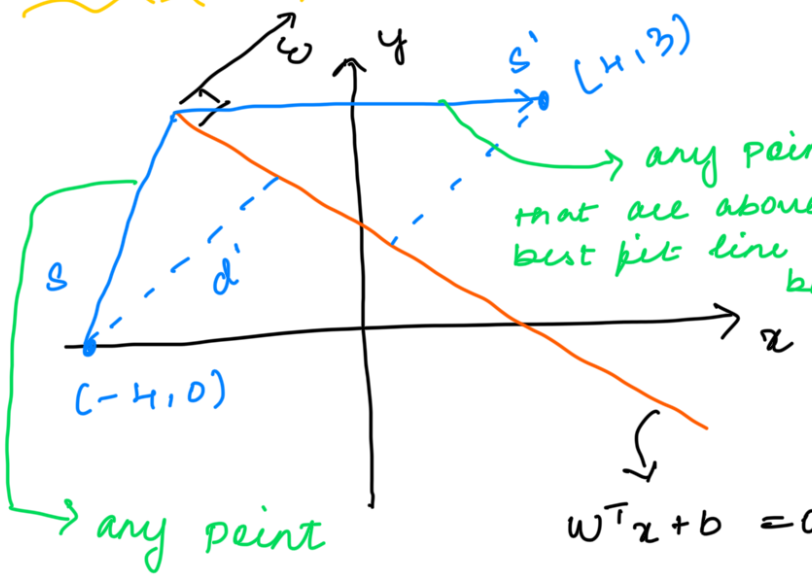
Support Vector Machine (SVC) Math Intuition



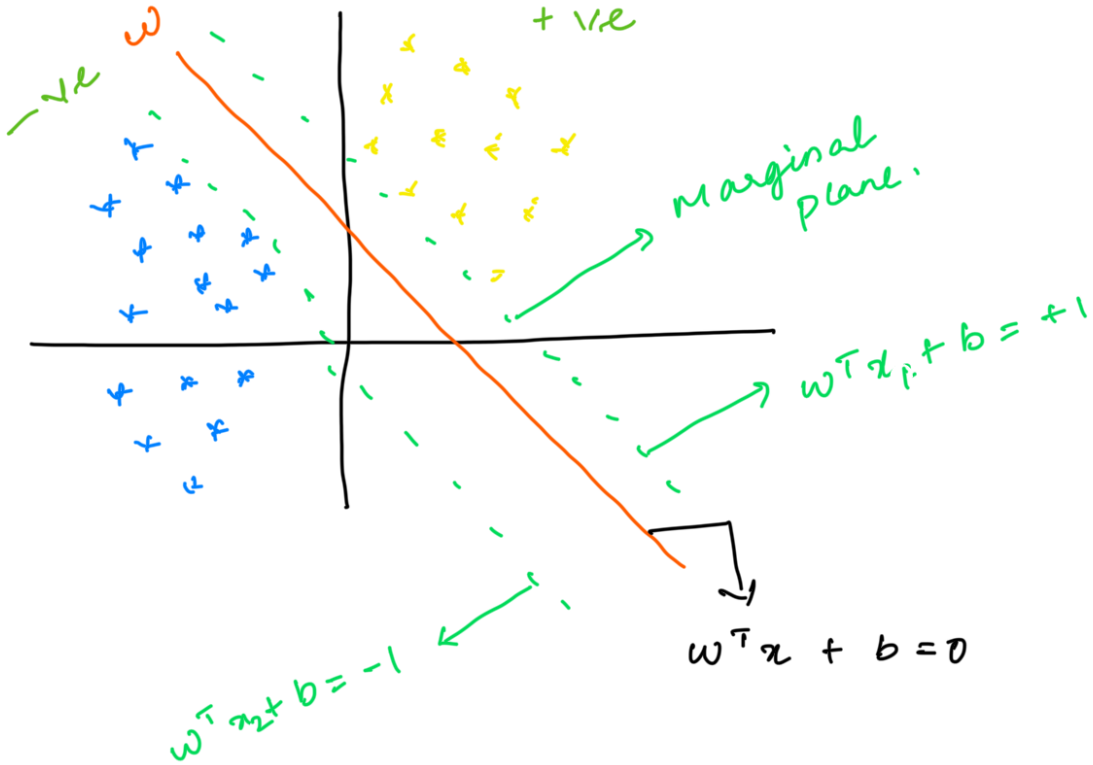
equation of a straight line

$$ax + by + c = 0$$

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

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

* distance will be -ve below the plane
* distance will be +ve above the plane.



Therefore,

$$w^T x_1 + b = 1$$

$$w^T x_2 + b = -1$$

$$(-)\quad (-)\quad (+)$$

$$w^T (x_1 - x_2) = +2$$

(solving linear equation)

Unit Vector

Magnitude of vector is 1

cost function

Maximize $\frac{2}{\|w\|}$ $\Rightarrow$ Distance between marginal plane

w, b

constraint such that

$$y_i \begin{cases} +1 & \text{if } w^T x + b \geq 1 \\ -1 & \text{if } w^T x + b \leq -1 \end{cases}$$

dealing with +ve points

dealing with -ve points

$\Downarrow$

These are applicable only for correctly classified points

For all correct points

constraint $\longrightarrow$ $\boxed{y_i (w^T x + b) \geq 1}$

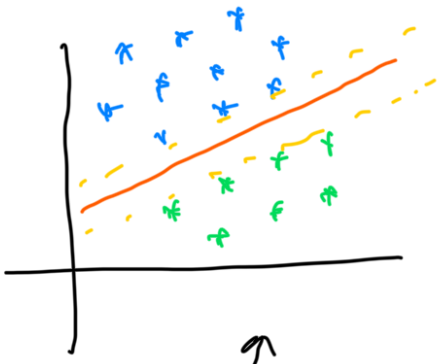
maximize $\frac{2}{\|w\|}$ = minimize $\frac{\|w\|}{2}$

$\min_{w,b} \|w\|$ $\min_{w,b}$

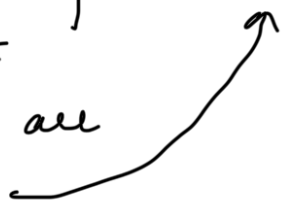
$\min_{w,b}$

cost function of SVM (SVC)

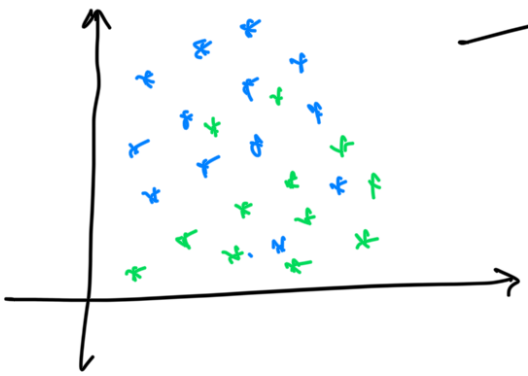
$$\min_{w,b} \frac{\|w\|}{2}$$



I can use this as my cost function, if my data points are clearly seperable



however, in the real world that is not the case,



→ there will be data over lapping. hence we will have tune it with hyper parameter.

In this case my cost function will be,

$$\min_{w,b} \frac{\|w\|}{2} + \boxed{C_i \sum_{i=1}^{n} h_i} \Rightarrow \text{Hinge Loss}$$

hyper parameter → $C_i \Rightarrow$ how many point we want to avoid misclassification.

$h_i \Rightarrow$ Summation of the distance of the incorrect data points from

the marginal plane.

* This cost function is related to soft margin.