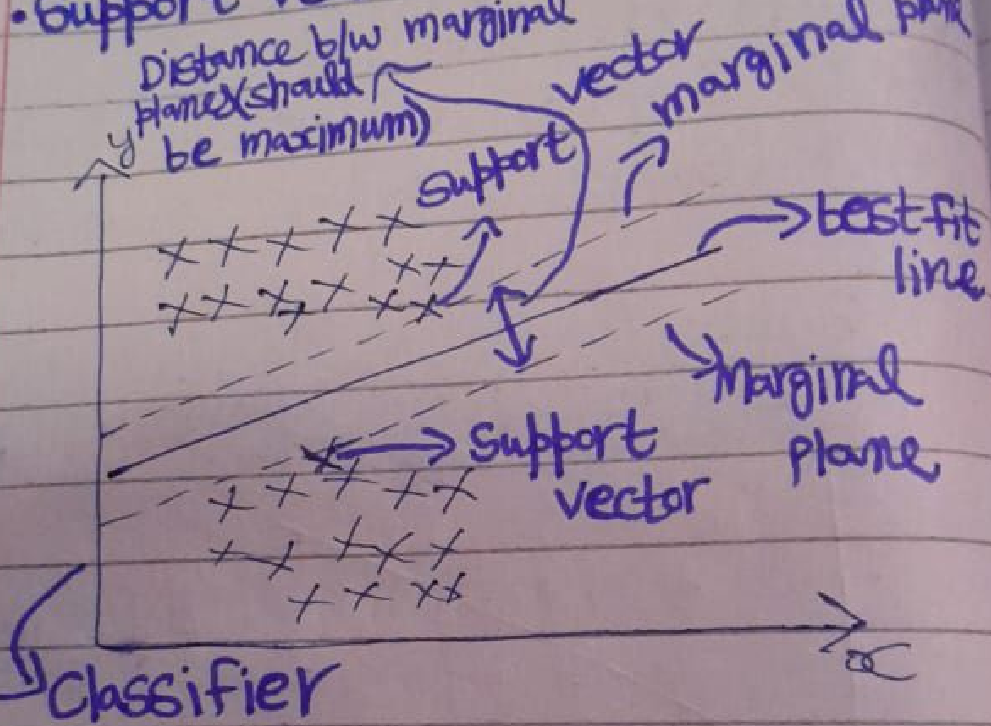
⇛ Support Vector Machine (SVM) Algorithm:

It is used for both classification and regression

1. SVC (Support Vector Classifier)
2. SVR (Support Vector Regressor)

• Support Vector Machine (SVC):



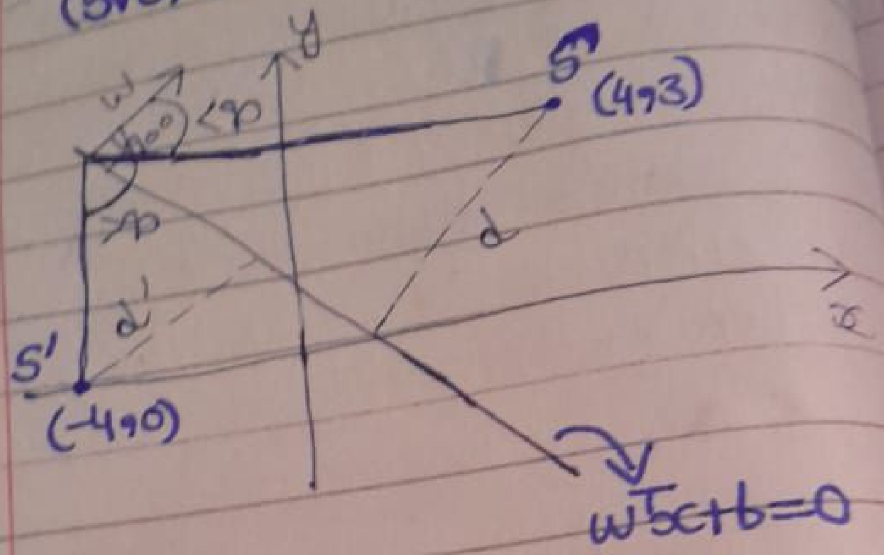
• Marginal plane pass through the nearest point from best fit line and the point called support vector

- Soft Margin and Hard Margin in SVM:
  - In **Soft Margin**, we get some amount of errors because points are overlapping.
  - In **Hard Margin**, there is no error the lines are clearly seperating points

Soft margin

- Hard Margin we drawn on previous page.

⇒ Support Vector Machine (SVC) Maths Intuition:



We know equation of straight line is

$ax+by+c=0$

can be written

$w_1x_1+w_2x_2+b=0$

$w^Tx+b=0$

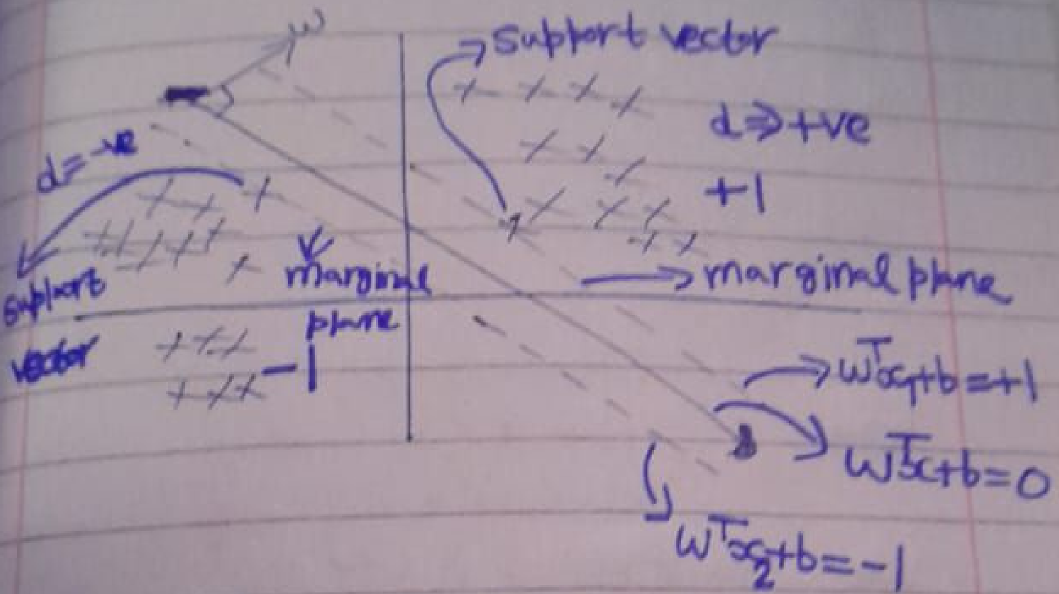
where $w=[w_1\ w_2]$ and $x=[x_1\ x_2]$

we took transpose for multiplication

- w will be perpendicular to line
- As below the plane angle b/w w and line is greater than 90 so distance will

be negative. And above the plane angle is less than 90 so distance will be positive.

$d' = -ve$ below the plane
$d = +ve$ above the plane



- We have to pickup the marginal planes with maximum distance b/w them

Distance will be given by subtracting equations.

$$w^T x_1 + \cancel{b} = 1$$
$$- \; w^T x_2 \; - \cancel{b} = + 1$$

$$\boxed{w^T(x_1 - x_2) = 2} \Rightarrow \text{Distance}$$

Unit vector (Magnitude of vector is 1):

$\hat{\omega} = \frac{\vec{w}}{||w||}$

So our distance in terms of unit vectors will become

$$\boxed{\frac{w^T}{||w||}(x_1 - x_2) = \frac{+2}{||w||}}$$

Cost Function:

$\frac{2}{||w||}$ ⇒ Distance b/w marginal planes (we have to maximize it)

- We will try to maximize this distance by changing the values of w and b.

Constraint such that

$$y_i = \begin{cases} +1 & w^Tx+b \geq 1 \\ -1 & w^Tx+b \leq -1 \end{cases}$$

⇒ This is for all correctly classified points

For all correct points

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

We can also write cost function as:

$$\underset{(w,b)}{\text{Maximize}} \frac{2}{||w||} \Rightarrow \boxed{\underset{(w,b)}{\text{Minimize}} \frac{||w||}{2}}$$

Cost function of SVM (SVC):

$$\underset{(w,b)}{\text{Minimize}} \frac{||w||}{2} + \boxed{C_i \sum_{i=1}^{n} \xi_i} \rightarrow \text{Hinge Loss}$$

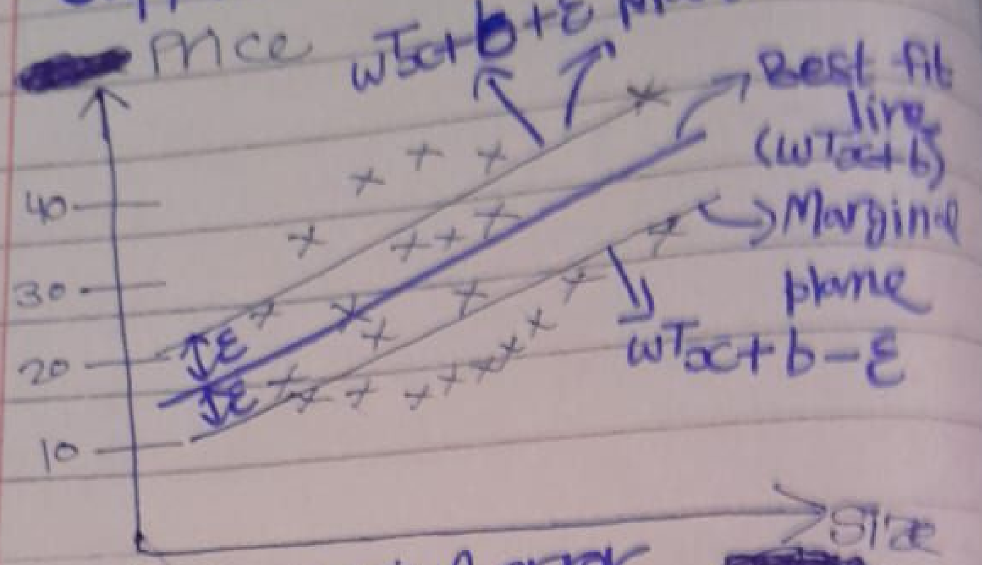
- We have added these Hyperparameters to our cost function because there is soft margin (in real scenarios) where points overlap and errors exist.

$C_i \Rightarrow$ How many points we can consider or bear to be missclassified (means how many errors are ignorable)

$\sum_{i=1}^{n} \xi_i$ = Summation of the distance of incorrect data points from the marginal plane

eta

- Support Vector Regression



- ε is the marginal error

Cost Function:

$$\underset{w,b}{\text{Minimize}} \ \frac{\|w\|}{2} + \boxed{C_i \sum_{i=1}^{n} \xi_i} \Rightarrow \text{Hinge loss}$$

Constraint:

$$\boxed{|y_i - w_i x_i| \le \varepsilon + \xi_i} \Rightarrow \text{Loss function}$$

$y_i \Rightarrow$ True points

$w_i x_i \Rightarrow$ Predicted points

- Most of points should fall within marginal plane.

$\xi \Rightarrow$ ~~It is distance of incorrect~~ means point outside margin plane from that margin plane

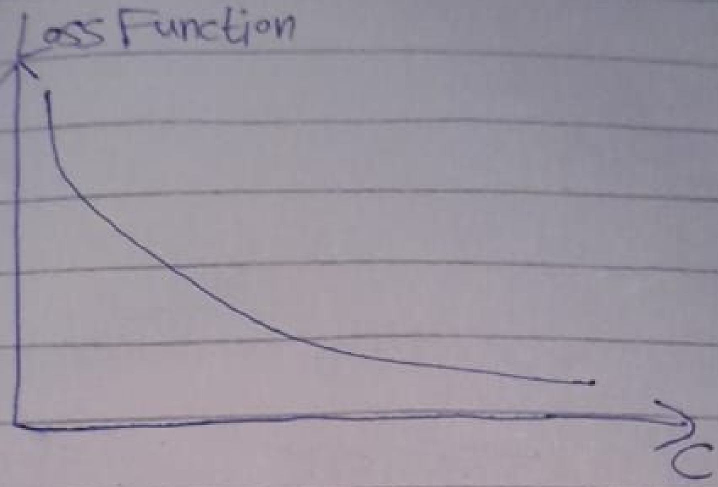
↓

eta

$\varepsilon \rightarrow$ Margin error

$\xi \rightarrow$ Error above the margin

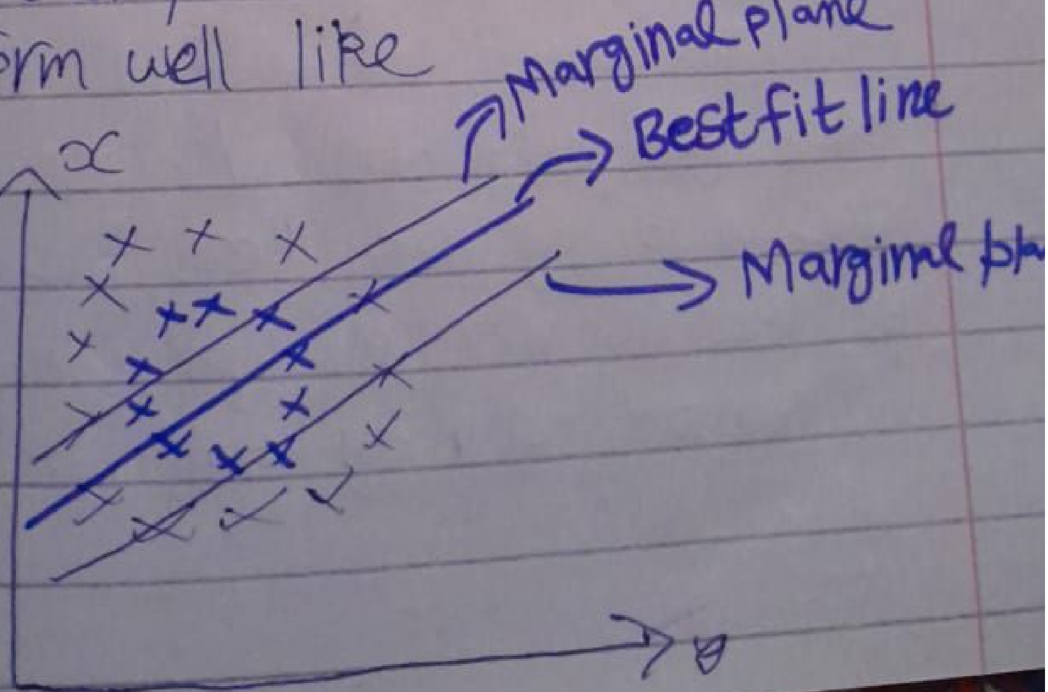
- **Relationship b/w loss function and c**



- As c increases loss function decreases c is number of points allowed outside marginal error

- **SVM Kernels**

When datapoints are not in linear shape SVC linear lines doesn't perform well like

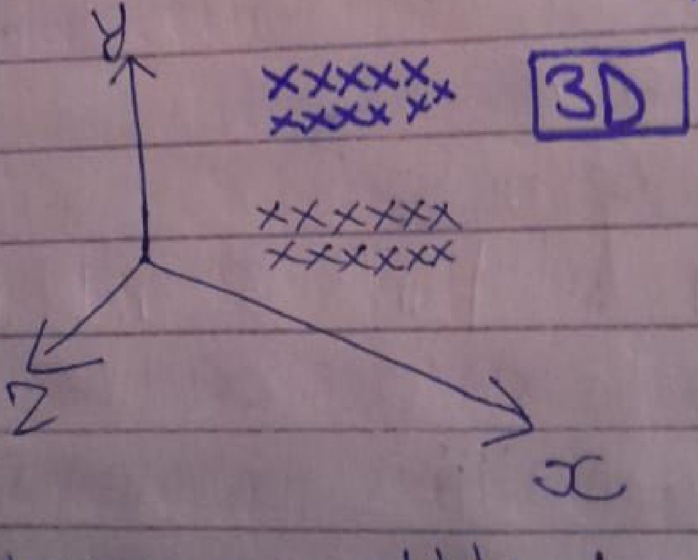


To solve this problem we use **SVM Kernel**

⇓ Apply on dataset

Transformations ⇒ (Apply mathematical formula)

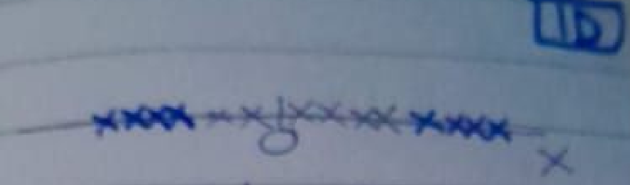
⇓



Now you can apply Linear SVC on it by draw best fit plane and marginal planes and our accuracy will increase.

Example:

Let 1 D dataset



We cannot solve it using Linear SVC so we will apply Transformation

Transformation $\Rightarrow y = x^2 \Rightarrow$ SVM Kernel

Now dataset will become

- Now we can apply linear SVC with good accuracy

$\Rightarrow$ Types of SVM Kernel:

① Polynomial Kernel
② RBF Kernel
③ Sigmoid Kernel