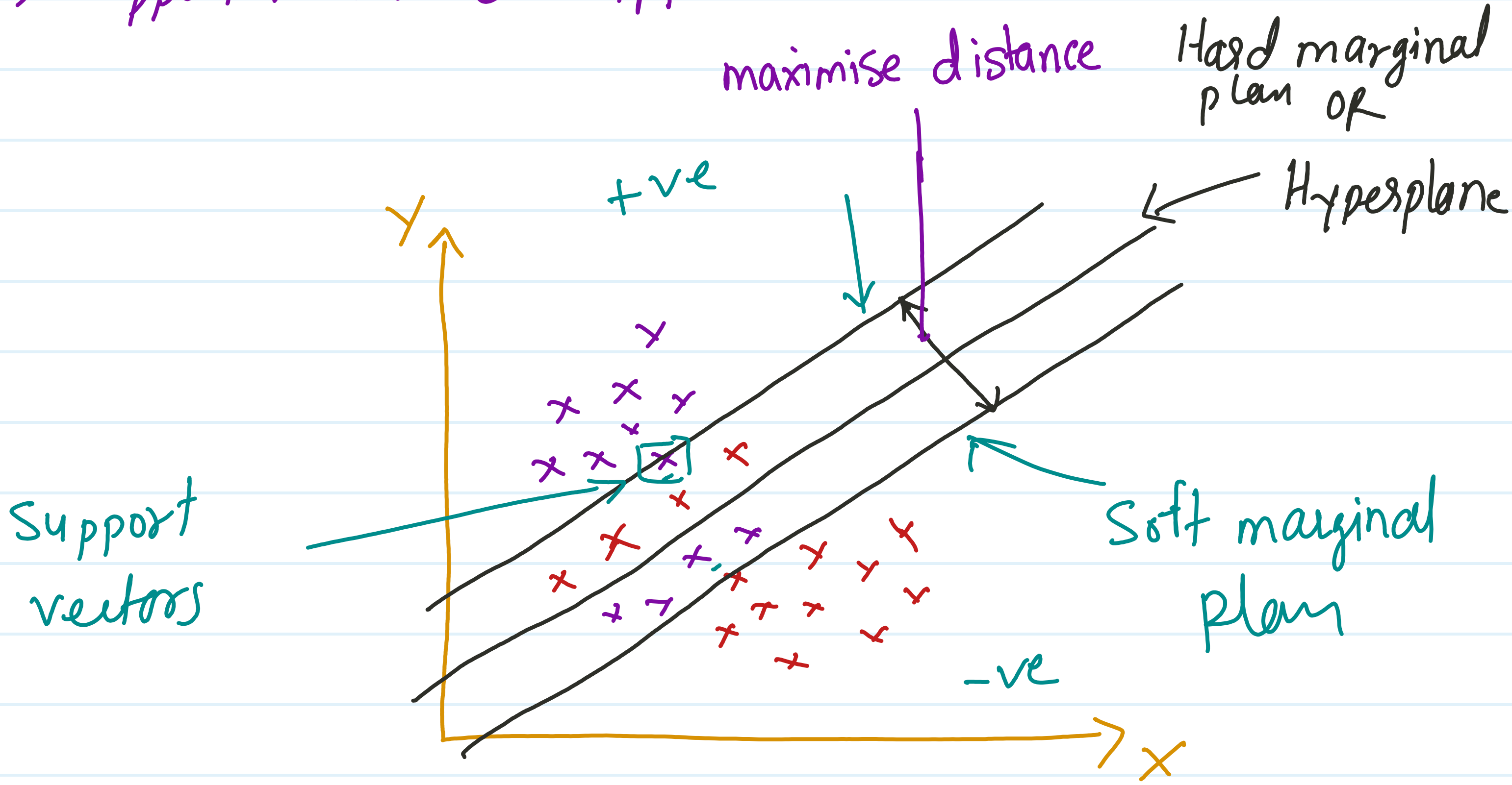
✡ Support vector Machine (SVM)

SVM
- Support vector classifier
- Support vector Regressor

⇒ Support vector Classifier



$$Y = mx + c$$

or

$$Y = \theta_0 + \theta_1 X_1$$

or

$$Y = \beta_0 + \beta_1 X_1$$

Note :- we alway play with soft marginal plan for seperate point or

$$y = w_1 x_1 + b$$

$$y = w_1 x_1 + w_2 x_2 + w_3 x_3 + b$$

$$w^T = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

Transpose

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

*

$$ax + by + c = 0$$

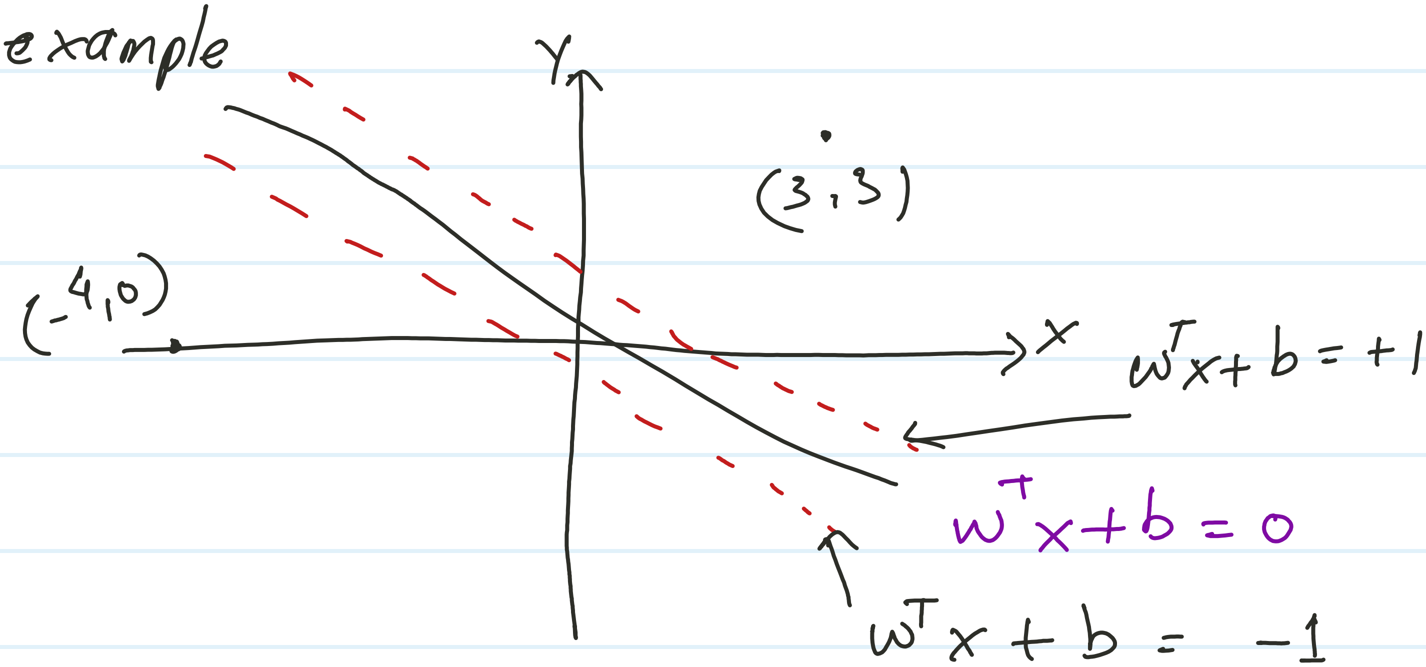
$$by = -ax - c$$

$$\boxed{y = -\frac{a}{b}(x) - \frac{c}{b}}$$

coefficient (slop) $m = -\dfrac{a}{b}$

Intercept $c = -\dfrac{c}{b}$

for example



we have a line eqn

$$3x + 2y + 4 = 0$$

for first point – $(-4, 0)$

$\rightarrow \quad 3 \times (-4) + 2 \times 0 + 4$

$\rightarrow \quad -12 + 4$

$\rightarrow \quad -8$ (–ve point)

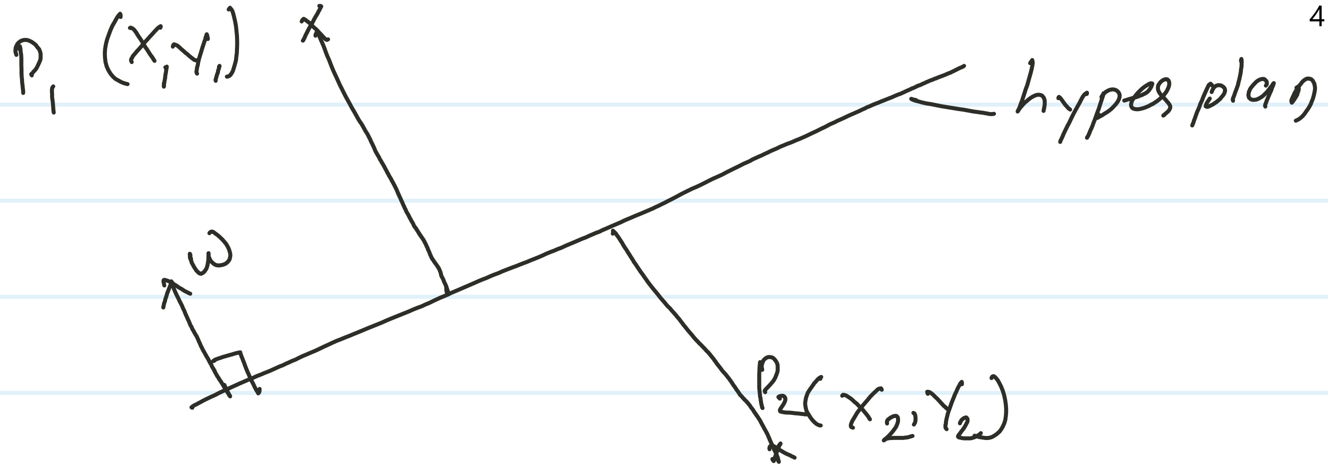
for second point – $(3, 3)$

$\rightarrow \quad 3 \times 3 + 2 \times 3 + 4$

$\rightarrow \quad 9 + 6 + 4$

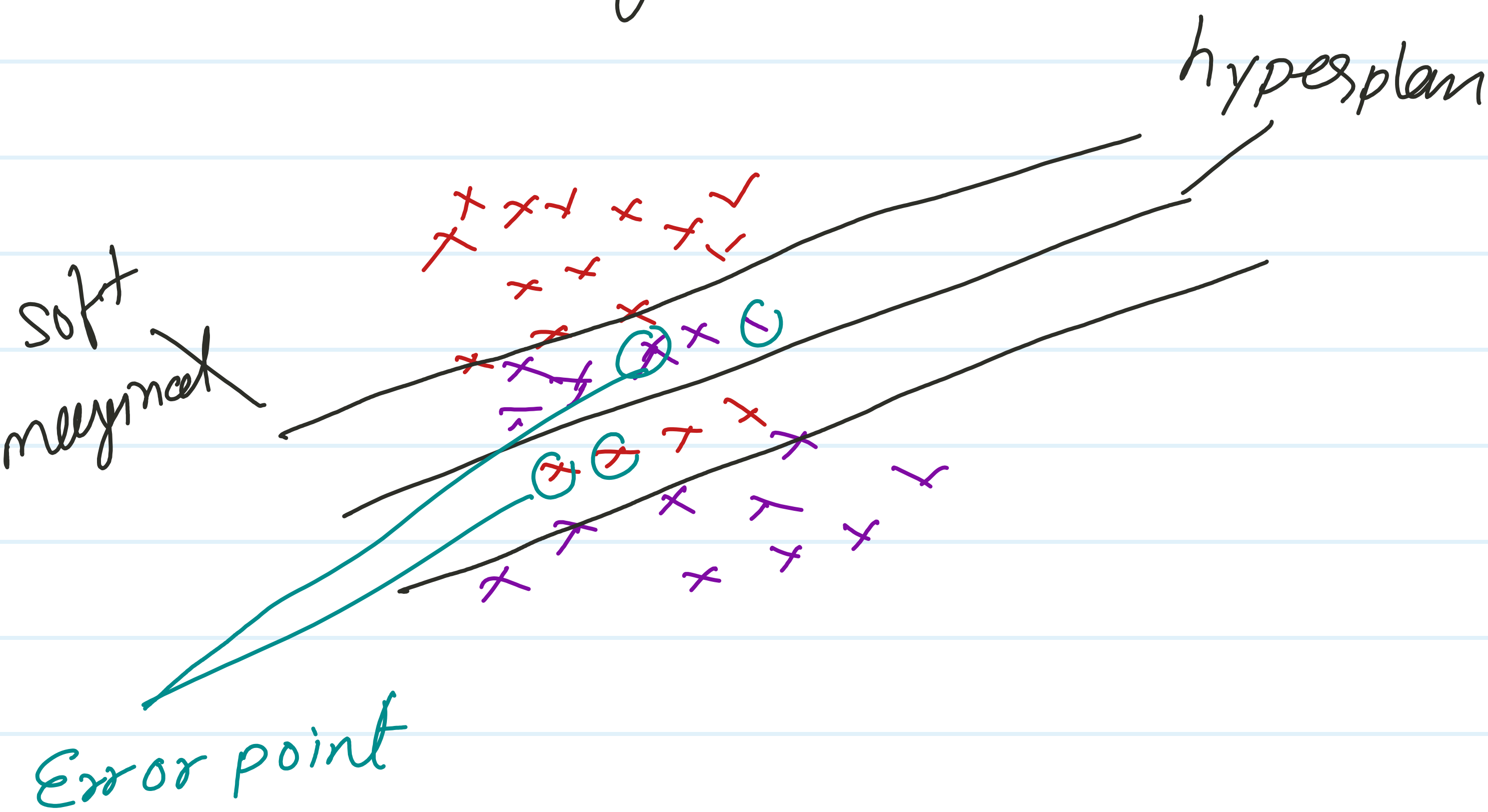
$\rightarrow \quad 19$ (+ve point)

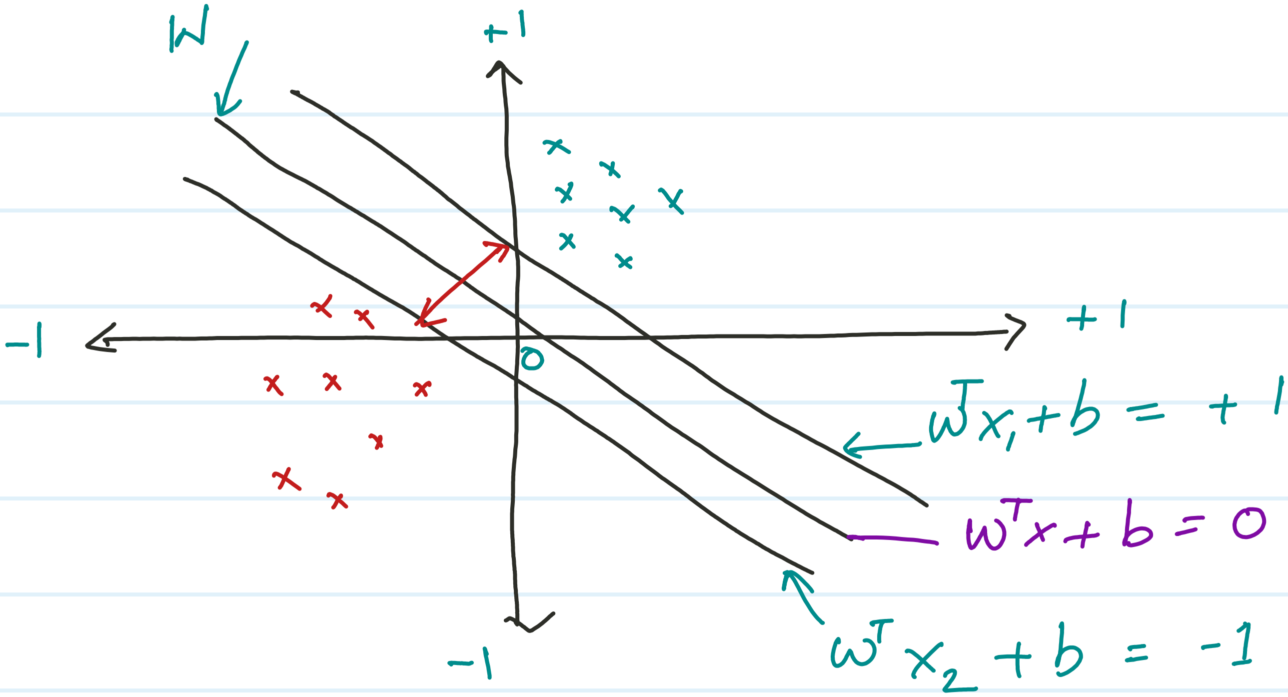
From this example if point is above line point will be positive or if point is below line point will be negative.



$$d = \frac{\omega^T p}{\|\omega\|}$$

✱ point above the line will be positive and below the line will be negative.





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

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

$$\overline{\hspace{6cm}}$$

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

magnitude of w

The eqn dived by $||w||$, To get unit vector.

$$\frac{w^T (x_1 - x_2)}{||w||} = \left.\frac{2}{||w||}\right\} \text{ maximize}$$

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

for all the correct point

constant $= y_i \times (w^T x + b) \geq 1$

$$\boxed{\max_{(w,b)} = \frac{2}{\|w\|}}$$

marginal plan distance

Revers the eqn to min. magnitude

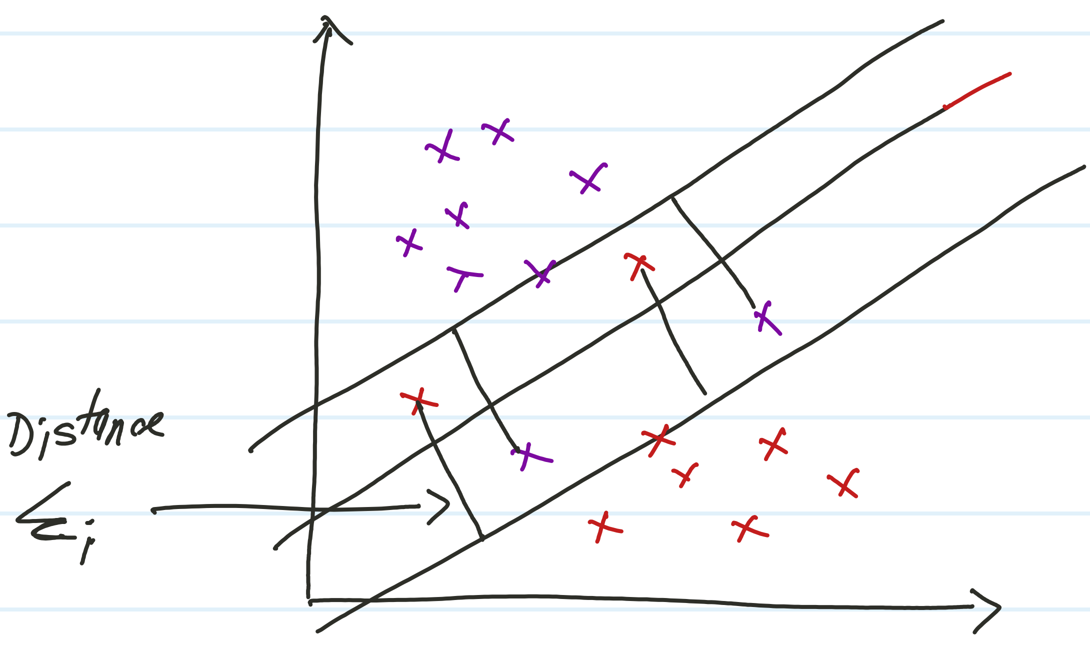
$$\min_{w,b} = \frac{\|w\|}{2}$$

* Cost function

$$\min_{(w,b)} = \frac{\|w\|}{2} + \boxed{C_i \sum_{i=1}^{w} \xi_i}$$

hingeloss

To avoid point for misclassification



Error point

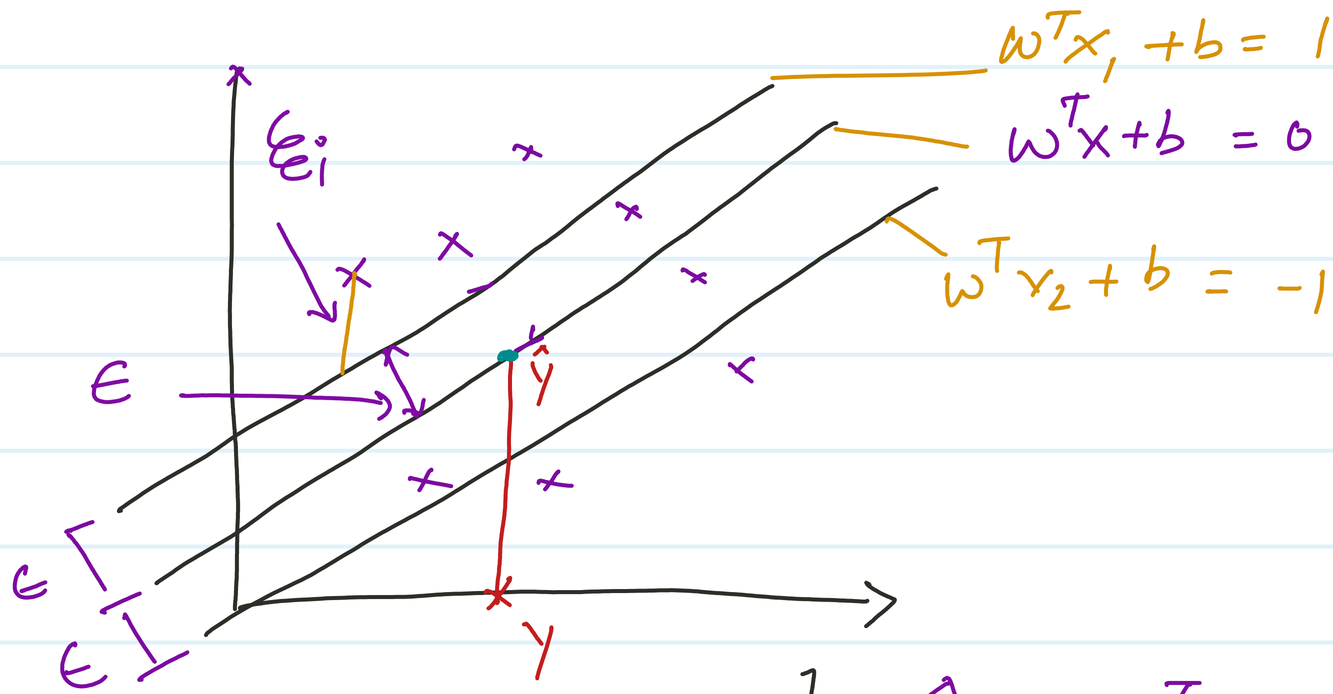
$C_i = 4$

C_i = How many point we can ignore for misclassification

ϵ_i = Sum of the distance of misclassfied point from marginal plan / hyperplan

Hyperparameter.

✩ SVR (Support Vector Regressor)



$w^T x + b = 0$

$w^T x - \epsilon$

$\hat{y} = w^T x_i$

cost function

$$\min_{(w,b)} \frac{\|w\|}{2} + \boxed{C_i \sum_{i=1}^{n} |\xi_i|} \text{ — Hinge loss}$$

constant $[y_i - w^T x_i] \leq \boxed{\epsilon + |\xi_i|}$

hyperparameter

Limitation –

① Impact by outlier

② Required standarized datapoint x and y

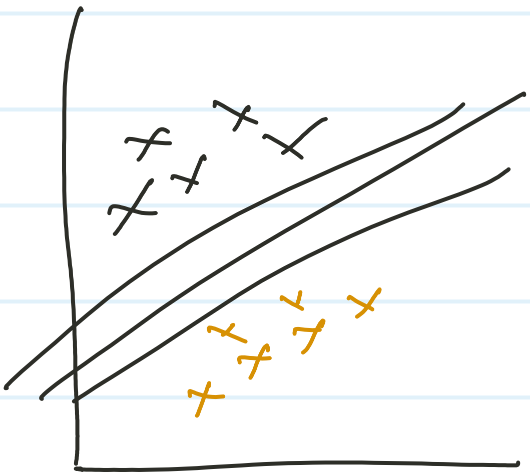
Advantages –

① SVM is used solve multi class classification problem

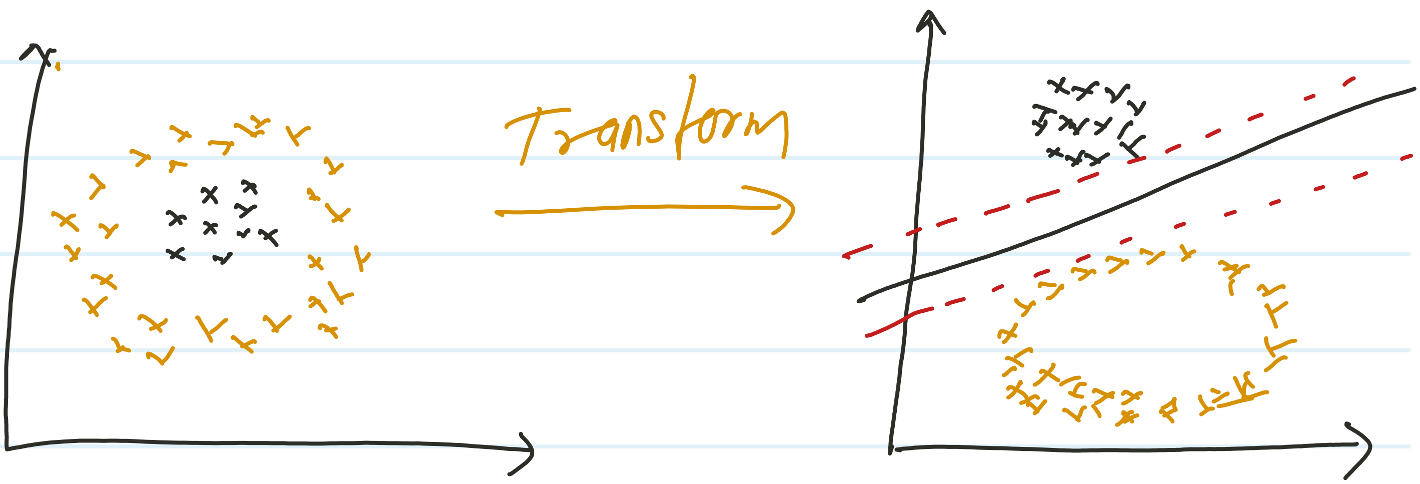
② Even it does not required Linear data like logistic and linear Regression for model building.

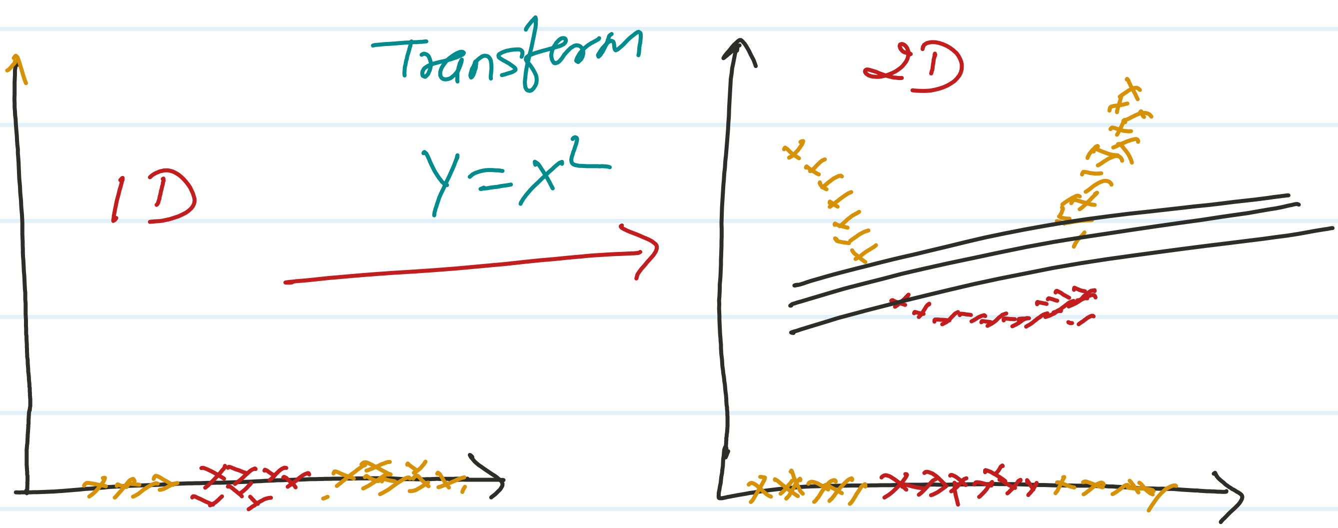
SVM kernel

SVM kernel only used for classification problem



Linear seprable





① polynomial kernal
② RBF kernel
③ sigmoid kernel

① polynomial kernal

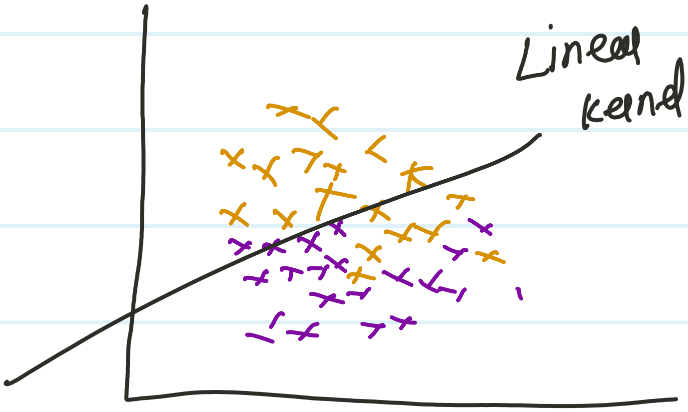
we convert Datapoint from 2D to 3D

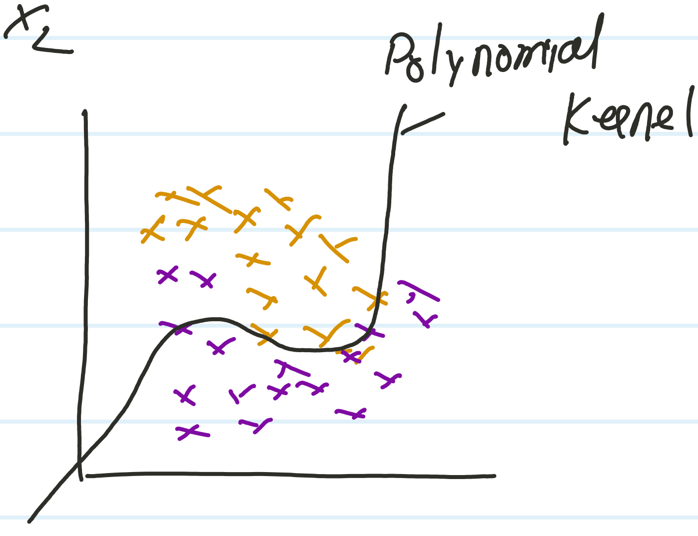
$$f(x_1, x_2) = (x_1^T x_2 + 1)^d$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{bmatrix}$$

x_1 x_2 $x_1 x_2$ $x_1 x_2$



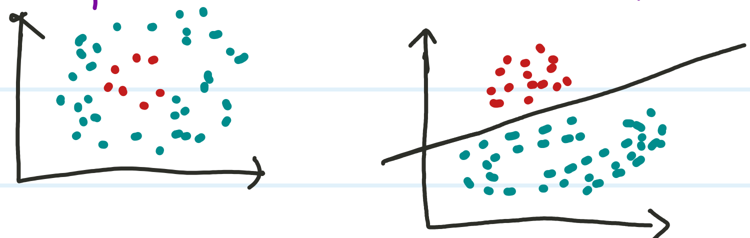


2. RBF (Radial Basis kernel)

We change dimension in it.

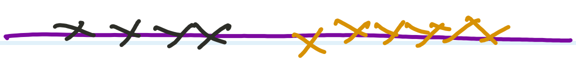
$$K(\vec{x}, l^i) = \frac{e^{-\|\vec{x} - l^i\|^2}}{2\sigma^2}$$

it is used to separate circular data point.



③ Sigmoid kernel ⇒

1D data point



$$\Rightarrow \frac{1}{1+e^{-z}}$$