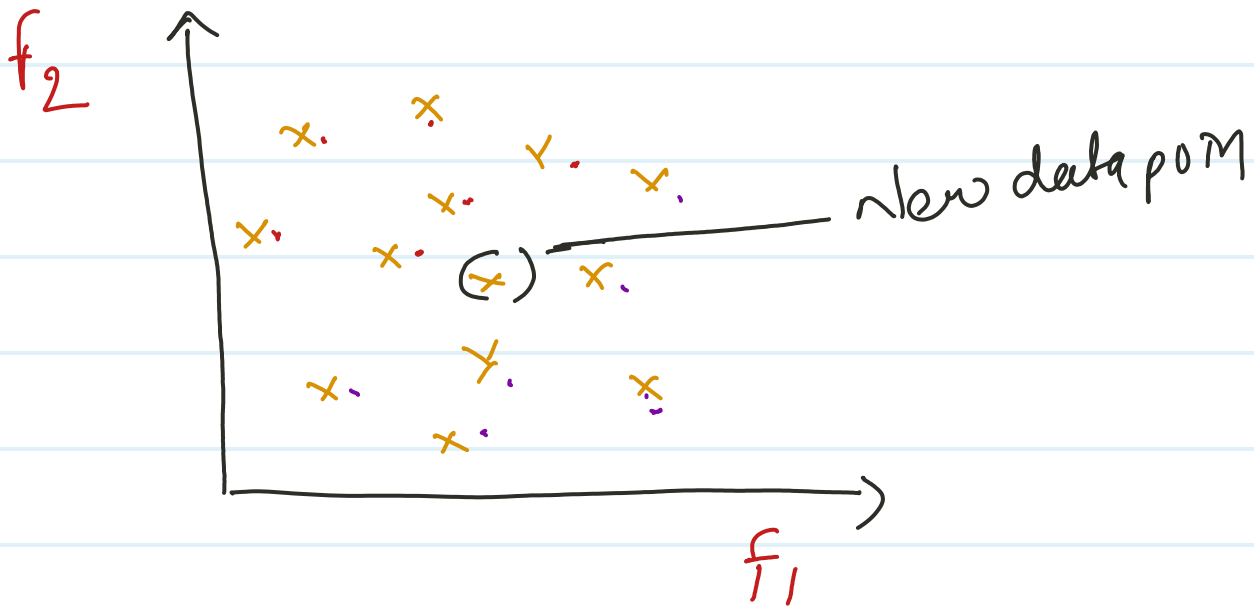
✡ KNN
(K- nearest neabour)

→ KNN works on distance based approach.

→ It is called lazy learner.

→ It is used to solve both classi. and Regression problem.

① classification -



k = 3, 5, 7, 9 = odd value.

k is hyperparameter.

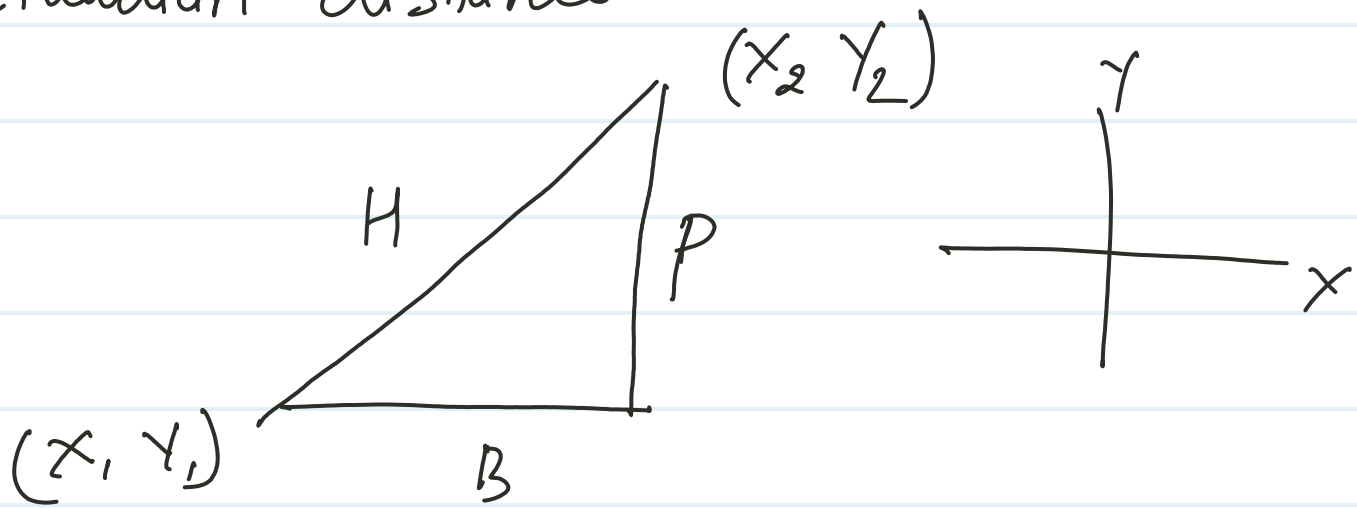
check distance from each point according to k value.

Suppose K = 5

1 + v = 3

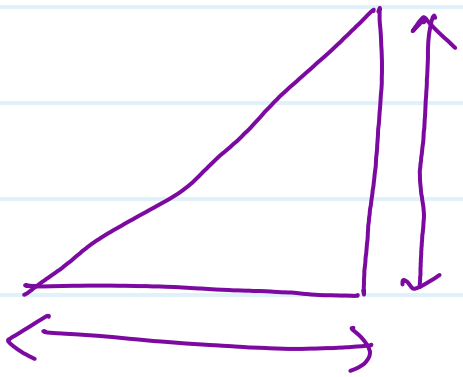
2 − v = 2

Eluadian distance



$$\text{Distance} = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

* Manhattan Distance =>

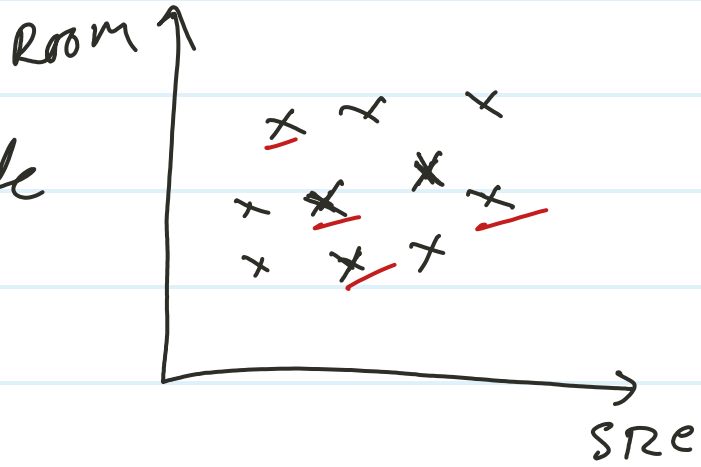


(2) Regression :-

location	Room	Area	Price
—	—	—	— Y

suppose k = 5

nearest point, calculate Avg.



limitation of KNN

(1) Cannot use for large dataset

(2) It highly affected with outlier.

(3) Even affected with missing value

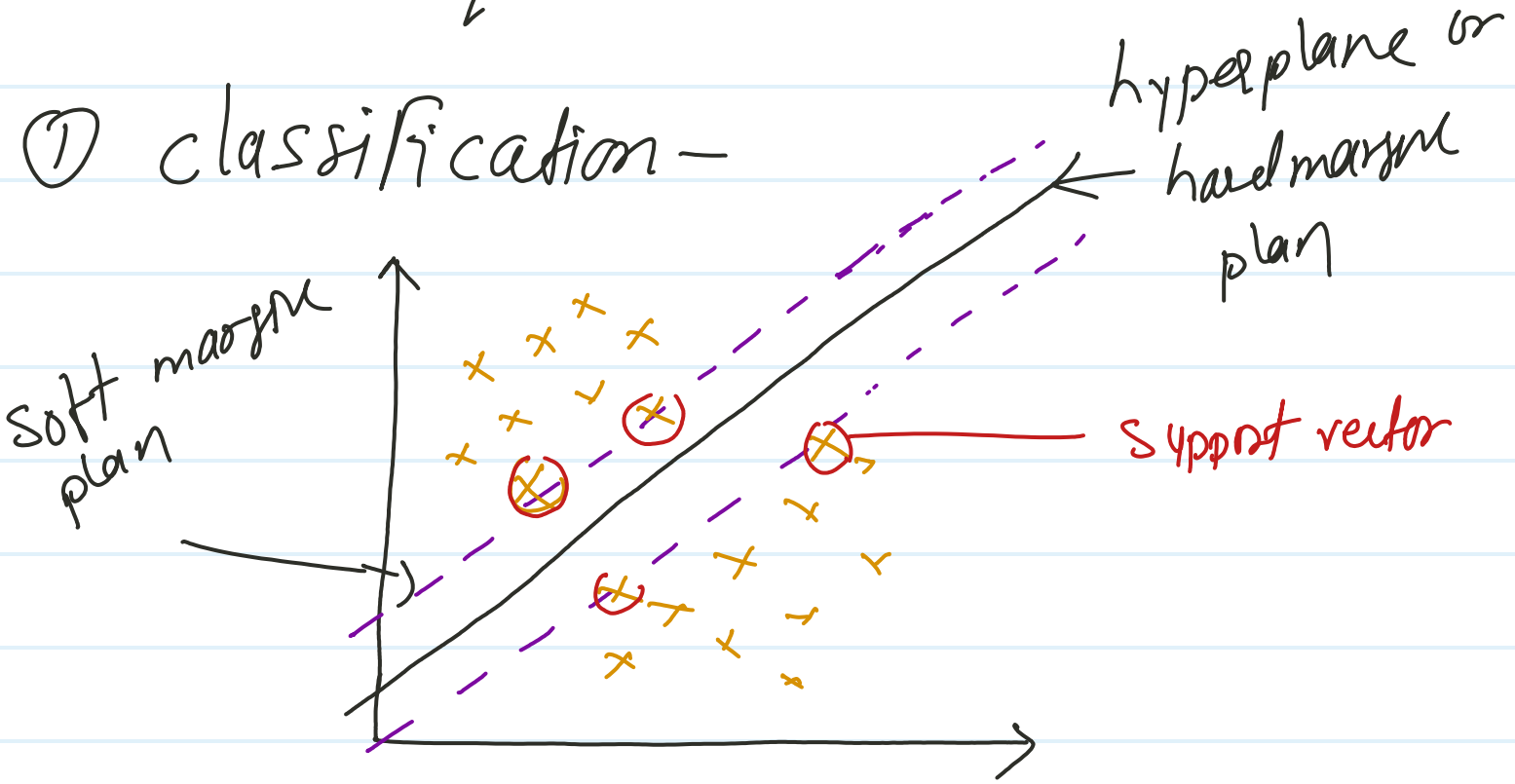
Application or Use

Feature selection

SVM
(Support vector machine)

SVC / SVR

① classification –



eqn

$$y = mx + c$$

or

$$y = \theta_0 + \theta_1 x$$

or

$$y = \beta_0 + \beta_1 x$$

OR

$$y = w_1 x + b$$

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

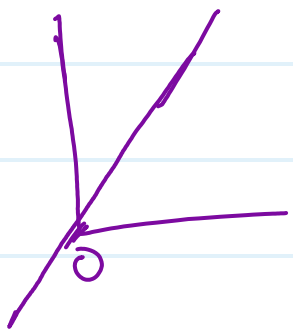
$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

W^T

$$\boxed{y = W^T x + b}$$

Line from origin $= W^T x = 0$

because $b = 0$



or

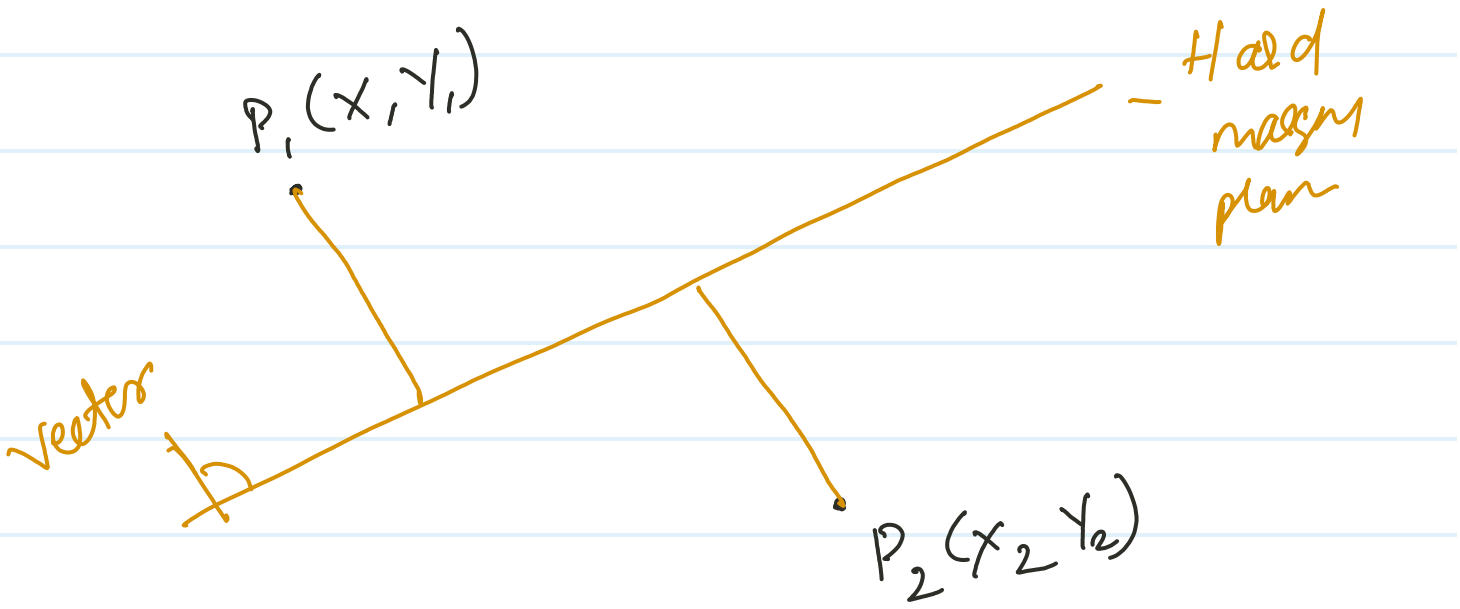
$$ax + by + c = 0$$

$$by = -ax - c$$

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

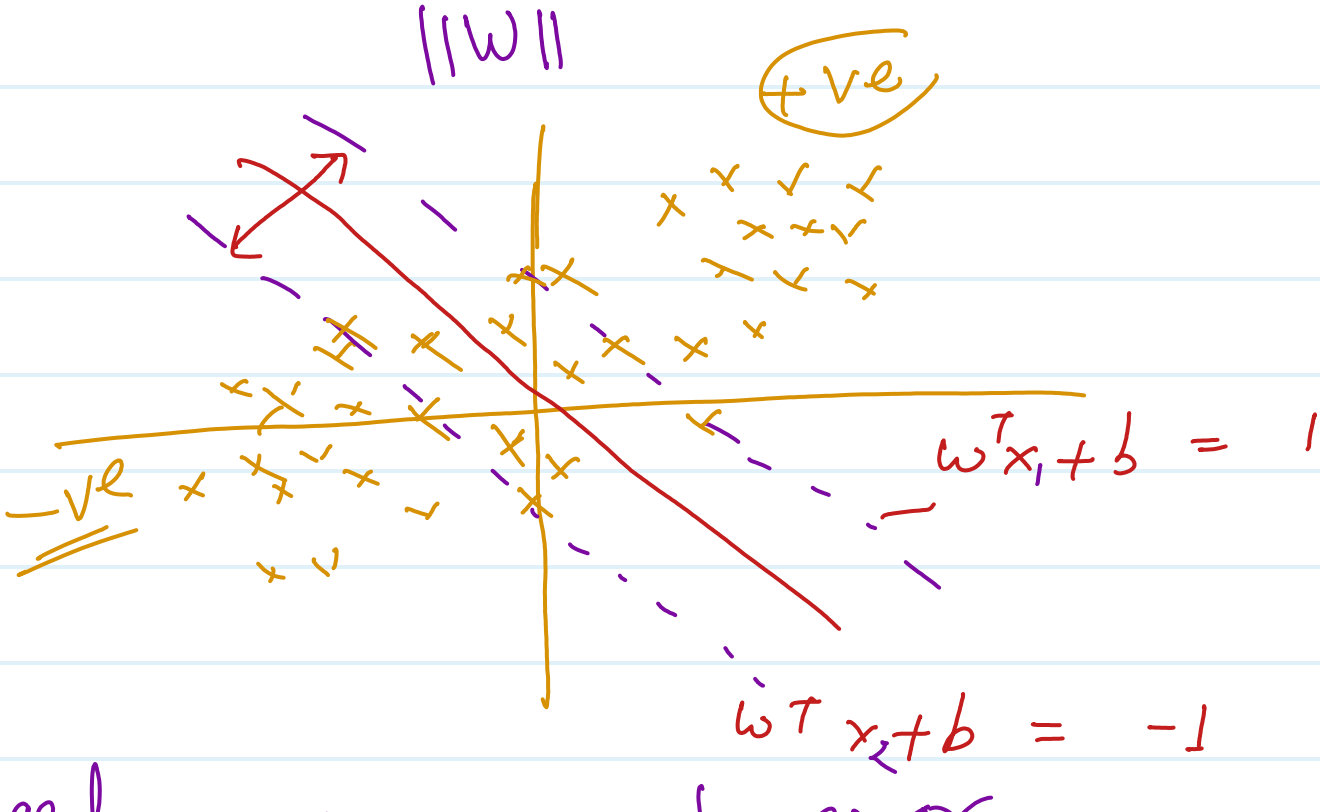
Coefficient $m = -\frac{a}{b}$

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



Distance of a point to a plane

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



Hard margine - n/o error

Soft margine = with some error

$$w^T x_1 + b = 1$$
$$w^T x_2 + b = -1$$
$$\text{(-)} \quad \text{(+)}$$
$$w^T (x_1 - x_2) = 2$$

both value div. by $\|w\|$

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

constant $y \begin{cases} +1, & w^T x + b \geq 1 \\ -1, & w^T x + b \leq -1 \end{cases}$

for all the correct point

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

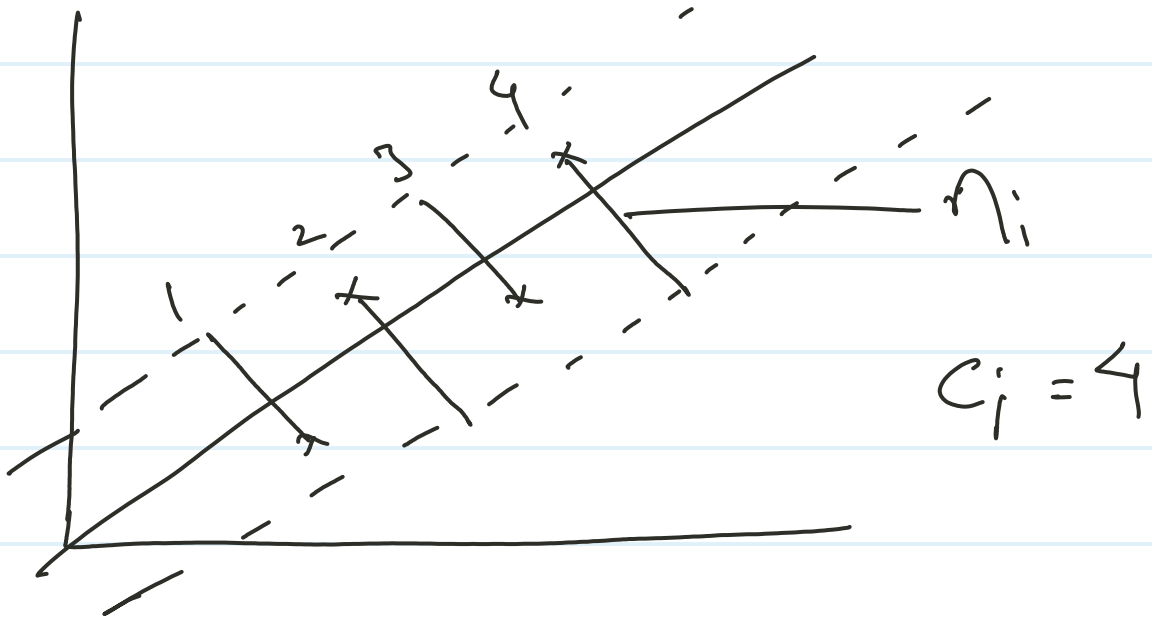
$$\max_{(w,b)} = \frac{2}{\|w\|}$$

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

* cost function of SVC -

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

η = eta


1
2
3
4
η_i
$C_i = 4$