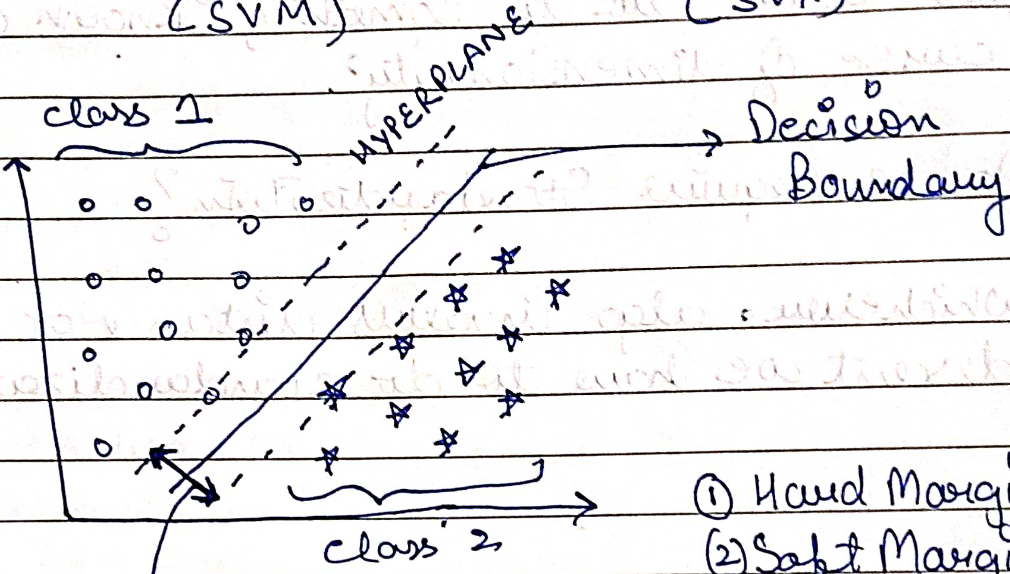
SUPPORT VECTOR MACHINE.

Classification (SVM) | Regression (SVR)

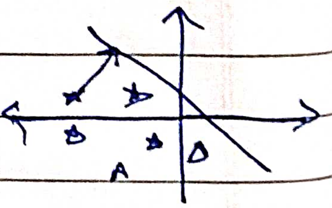


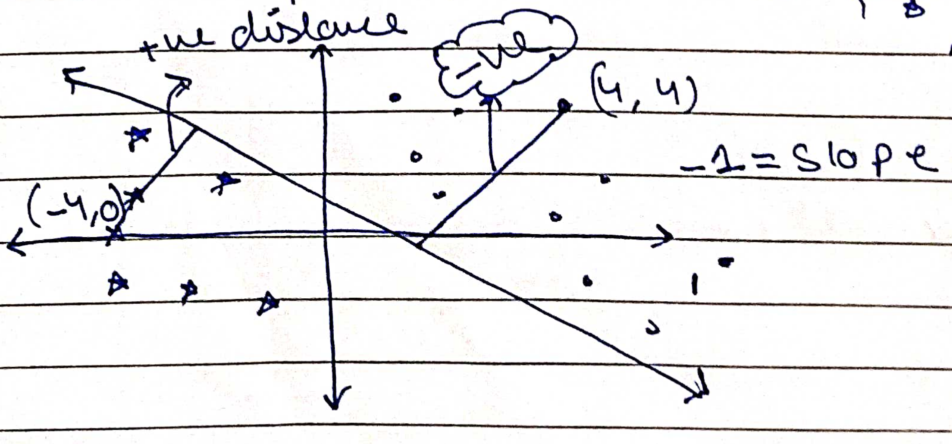
Main motto is to maximize the gap

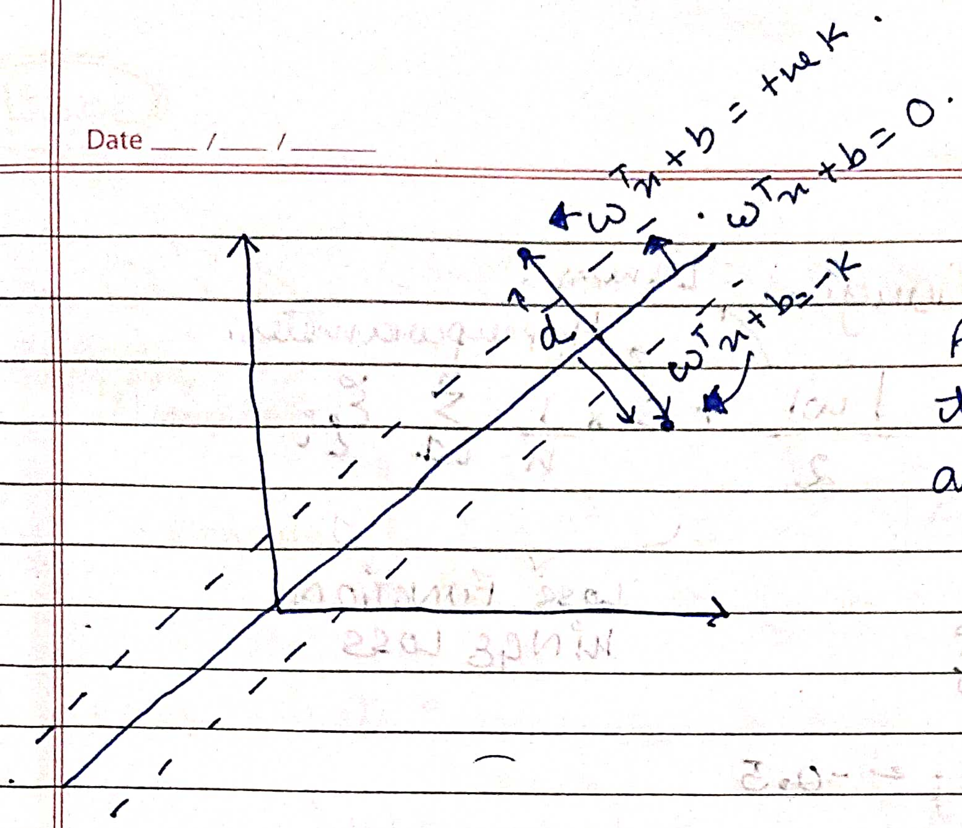
① Hard Margin SVM
② Soft Margin SVM

⇒ If your data is linearly separable [If the data can be classified into two classes using a st. line] SVM works perfectly fine

Distance from a pt to plane = $w^T x$







Aim is to calculate the distance 'd' and maximize

$$\therefore \quad w^T x_1 + \cancel{b} = +k$$
$$w^T x_2 + \cancel{b} = -k$$
$$- \qquad - \qquad +$$
$$w^T(x_1 - x_2) = 2k.$$

$$w^T(x_1 - x_2) = 2k \quad \text{—(i)}$$

Dividing (i) by $|w|$

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

Such that

$$y_i \times w^T x_i \geq 0.$$
$$\Downarrow$$

Can be applicable in Hard Margin or when data is linearly separable.

Let $k = 1$.

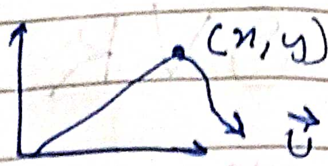
$$\left\{ y_i(\vec{w} \cdot \vec{x}_i + b) \geq 0 \right\}$$

$$\boxed{\max_{w,b} = \frac{2}{|w|}} \Longrightarrow \text{OPTIMIZATION FUNCTION}$$

$\Leftrightarrow$ Change w or b till the value is maximized.

Date ___ / ___ / ___

DECISION RULE.



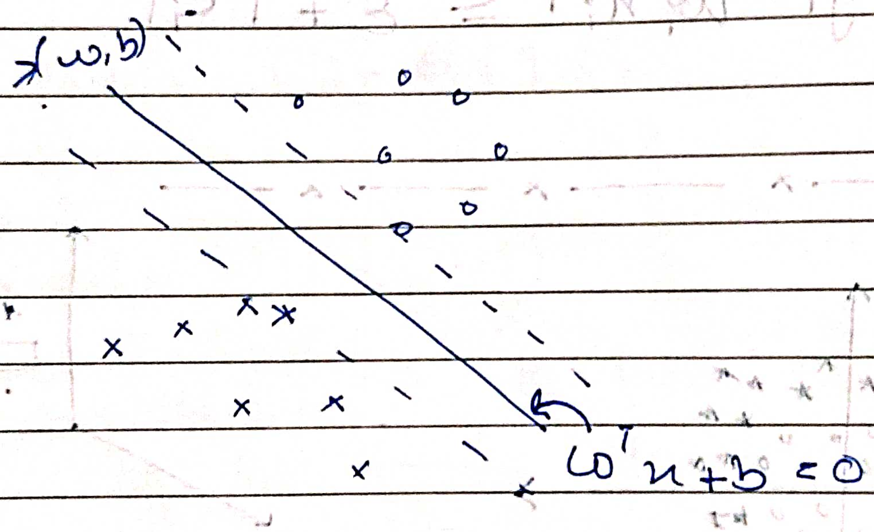
$\vec{w} \cdot \vec{U} \geq c$ → Scalar.

$\vec{w} \cdot \vec{U} - c \geq 0$

$\boxed{\vec{w} \cdot \vec{U} + b \geq 0}$ → Decision Rule.

if $\vec{U} \geq 0$ pt is +ve

$\vec{U} \leq 0$ pt is -ve.



for any $\overrightarrow{x_i}$

$$\hat{y} \text{ (Pred)} = \begin{cases} +1 & \text{if } \vec{w} \cdot \vec{x_i} + b \geq 0 \\ -1 & \text{if } \vec{w} \cdot \vec{x_i} + b < 0 \end{cases}$$

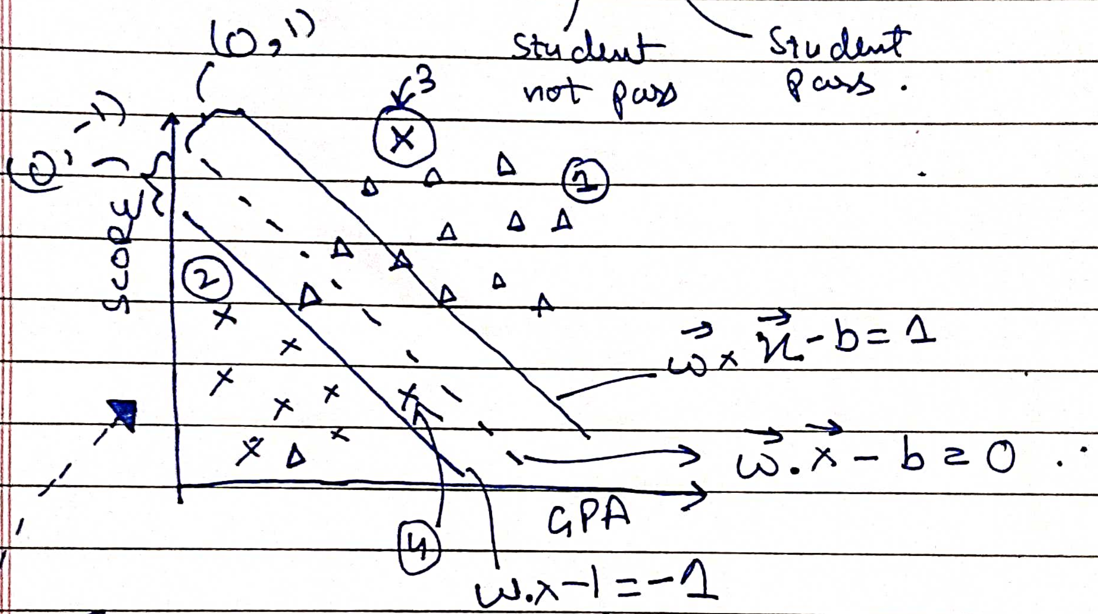
SOFT MARGIN SVM.

$$\begin{bmatrix} \vec{X}_i \\ \vdots \\ \vec{X}_n \end{bmatrix} \quad \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

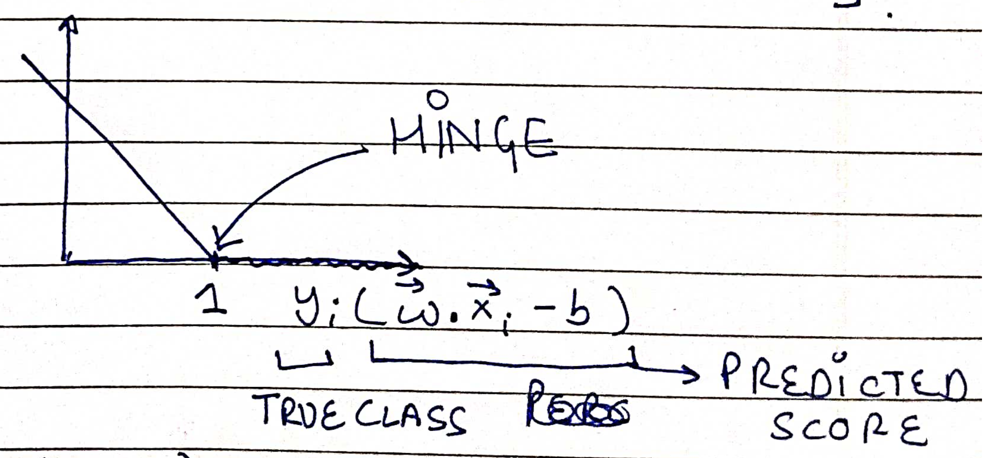
$y \rightarrow \{-1, 1\}$.

-1: Student not pass

1: Student pass.



$$\boxed{HINGE\ LOSS = \max(0, 1 - y_i(\vec{\omega} \times \vec{x}_i - b))}$$



1) $\max(0, 1 - 1(>1)) = 0$
2) $\max(0, 1 + 1(<-1)) = 0$
3) $\max(0, 1 + 1(>1)) = >1 \rightarrow$ LOSS >1
4) $\max(0, 1 + 1($ b/w -1 and $0)$

Soft Margin $\xrightarrow{}$ L1 norm.

Hyperparameter.

- $\underset{w,b}{\text{argmin}} \quad \frac{|w|}{2} + \underbrace{C \times \frac{1}{n} \sum_{i=1}^{n} \xi_i}_{\text{LOSS FUNCTION.}}$

HINGE LOSS

Q What is ξ

Pt one

Eg: $y_i \times w^T x_i = -0.5$

or

$y_i \times w^T x_i = 1 - (1.5)$

Pt two

$y_i \times w^T x = 1 - (2.5)$

$\boxed{y_i \times w^T x_i = 1 - \xi_i}$

ξ is the distance of the misclassified point from its actual plane.

SVR (support vector Regression)

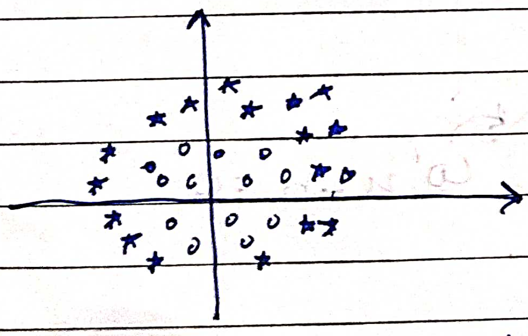
Minimize.

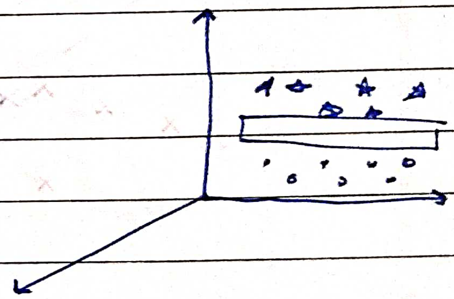
$$\text{Argmin } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}|\xi_i|$$

Constraints:

$$|y_i - w_i x_i| \le \varepsilon + |\xi_i|$$

— x — x — x —





Uses KNN

We use kernals here [Rbf, polynoical,

Default