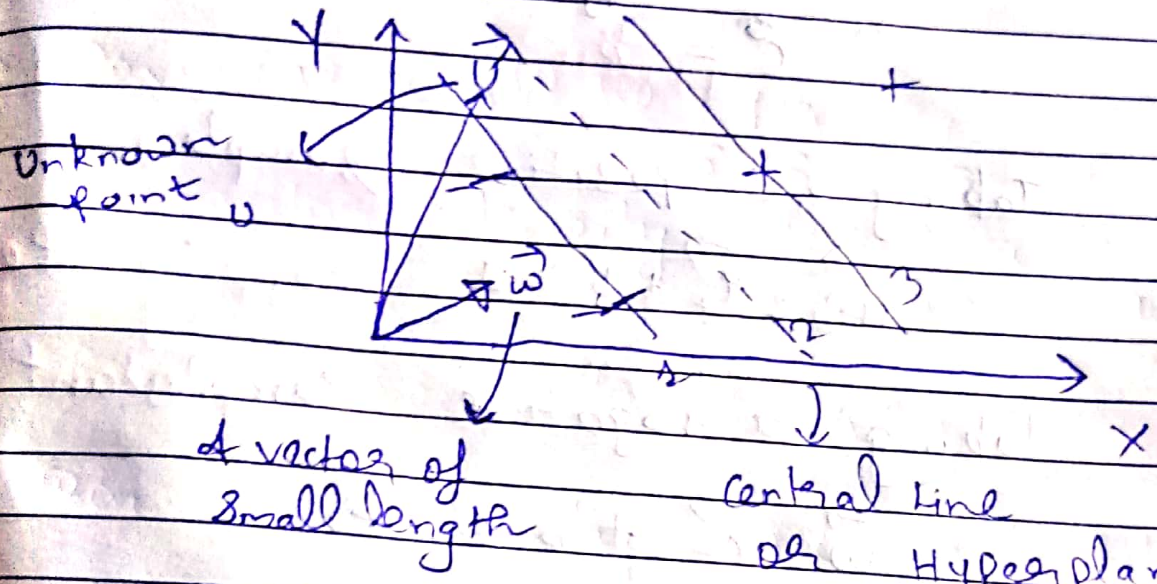
SVM classifier

Consider a system where we want to classify data points as positive and negative.



Taking dot product of w & u

$\vec{w} \cdot \vec{u} \geq c$ where c is a constant

Or we could say without the loss of generality

$$\boxed{\vec{w} \cdot \vec{u} + b \geq 0 \quad \text{Then } +}$$

Decision Rule — ①

②

$\vec{w}$ has to be perpendicular to the ~~median of the~~ decision boundary. But we do not know the length. Nor we know $\vec{b}$

$$c = -b$$

Taking a positive sample,

$$\vec{w} \cdot \vec{x}_{+} + b \geq 1$$

Taking a negative sample,

$$\vec{w} \cdot \vec{x}_{-} + b \leq -1$$

Let's introduce a variable y_i

y_i such that $y_i = +1$ for + samples and $y_i = -1$ for − samples.

For + samples

$$y_i(x_i \cdot \vec{w} + b) \geq 1$$

For − samples

$$y_i(\vec{x_i} \cdot \vec{w} + b) \geq 1$$

$$y_i(\vec{x_i}\cdot\vec{w}+\vec{b})-1 \geq 0$$

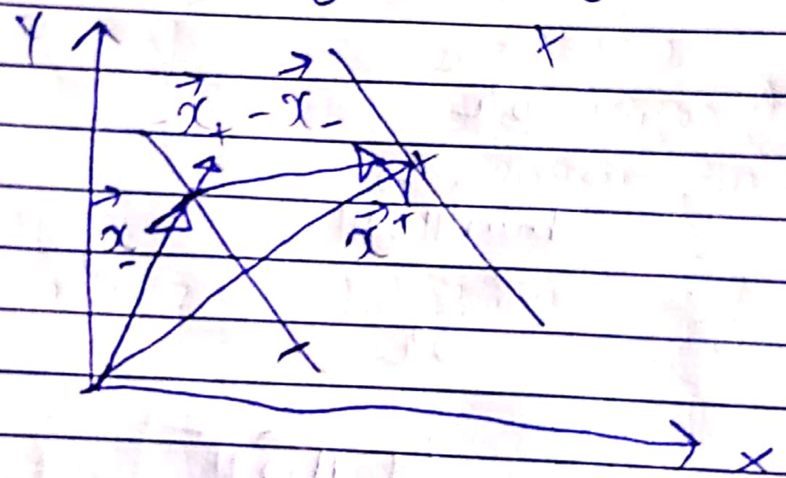
Now,

$$\boxed{y_i(\vec{x_i}\cdot\vec{w}+\vec{b})-1=0}$$

For x_i on lines 1 & 3 i.e. passing through the support vectors.

— ②

Redrawing the figure



$$\text{width} = (\vec{x}_+ - \vec{x}_-)\cdot\frac{\vec{w}}{\|\vec{w}\|}$$

$$\vec{x}_+\cdot\vec{w} = 1-b$$

$$\vec{x}_-\cdot\vec{w} = 1-b$$

from ②

$\therefore$ Width $= \frac{2}{||\vec{w}||}$ — (3)

We need to maximize the width

$$\max\left(\frac{2}{||\vec{w}||}\right)$$

or

$$\max\left(\frac{1}{||\vec{w}||}\right)$$ dropping the constant

or

$$\min(||\vec{w}||)$$

or

$$\min\left(\frac{1}{2}||\vec{w}||^2\right)$$

In order to maximize the width

$$L = \frac{1}{2}||\vec{w}||^2 - \sum \alpha_i \left[y_i(\vec{w}\cdot\vec{x}_i + b) - 1\right]$$

Taking Partial Derivative wrt $\vec{w}$ and setting it to 0

પ્રશ્ન
પેટા પ્રશ્ન
ક્રમાંક

$$\frac{\partial L}{\partial \vec{w}} = \vec{w} - \sum \alpha_i y_i x_i = 0$$

$$\Rightarrow \boxed{\vec{w} = \sum_i \alpha_i y_i \vec{x_i}}$$

Taking partial derivative wrt b

$$\frac{\partial L}{\partial b} = -\sum \alpha_i y_i = 0$$

$$\Rightarrow \boxed{\sum \alpha_i y_i = 0}$$

$$\therefore L = \frac{1}{2}\left(\sum \alpha_i y \vec{x_i}\right)\left(\sum \alpha_j y_j \vec{x_j}\right) - \sum \alpha_i y_i x_i \cdot \left(\sum \alpha_j y_j x_j\right) - \underbrace{\sum \alpha_i y_i b}_{0} + \sum \alpha_i$$

$$\boxed{L = \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i \cdot x_j}$$

we need to maximize L

(6)

$\sum \alpha_i y_i \vec{x_i} \cdot \vec{v} + b \geq 0$ Then +

else − samples

Here, $\vec{v}$ is the unknown point

Scanned with CamScanner