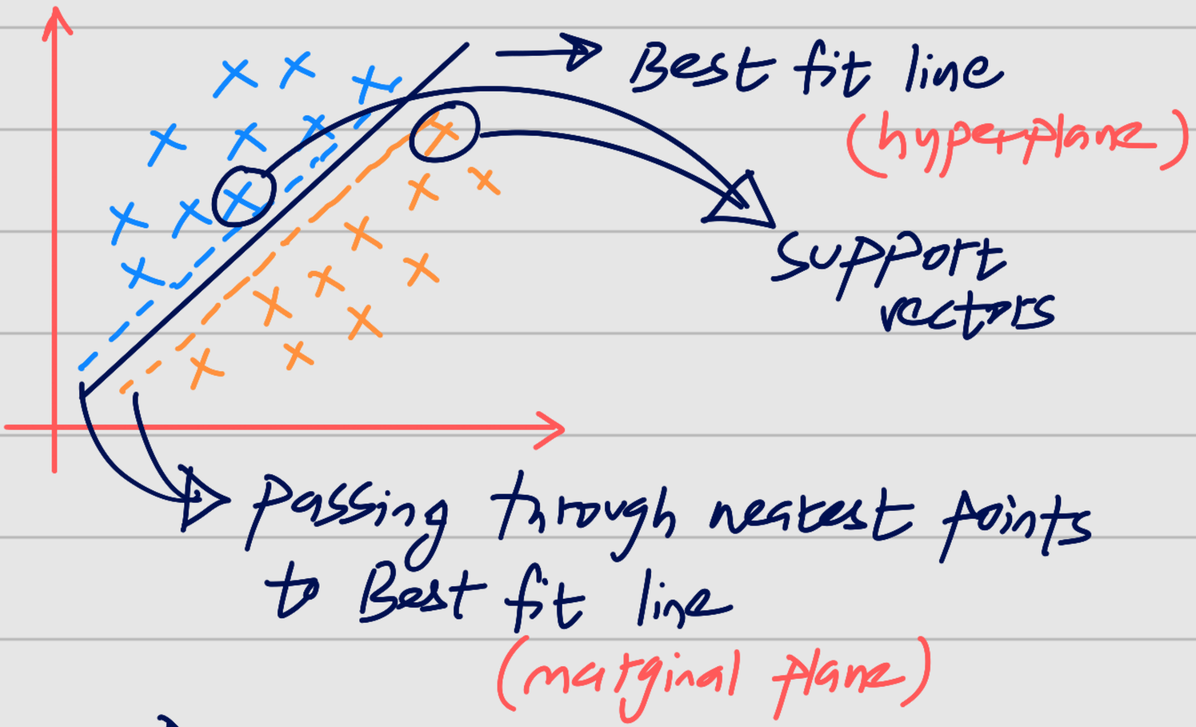
SVM Classification

- Supervised ML Algorithm
- Used for classification & Regression

→ Eg (Binary classification)

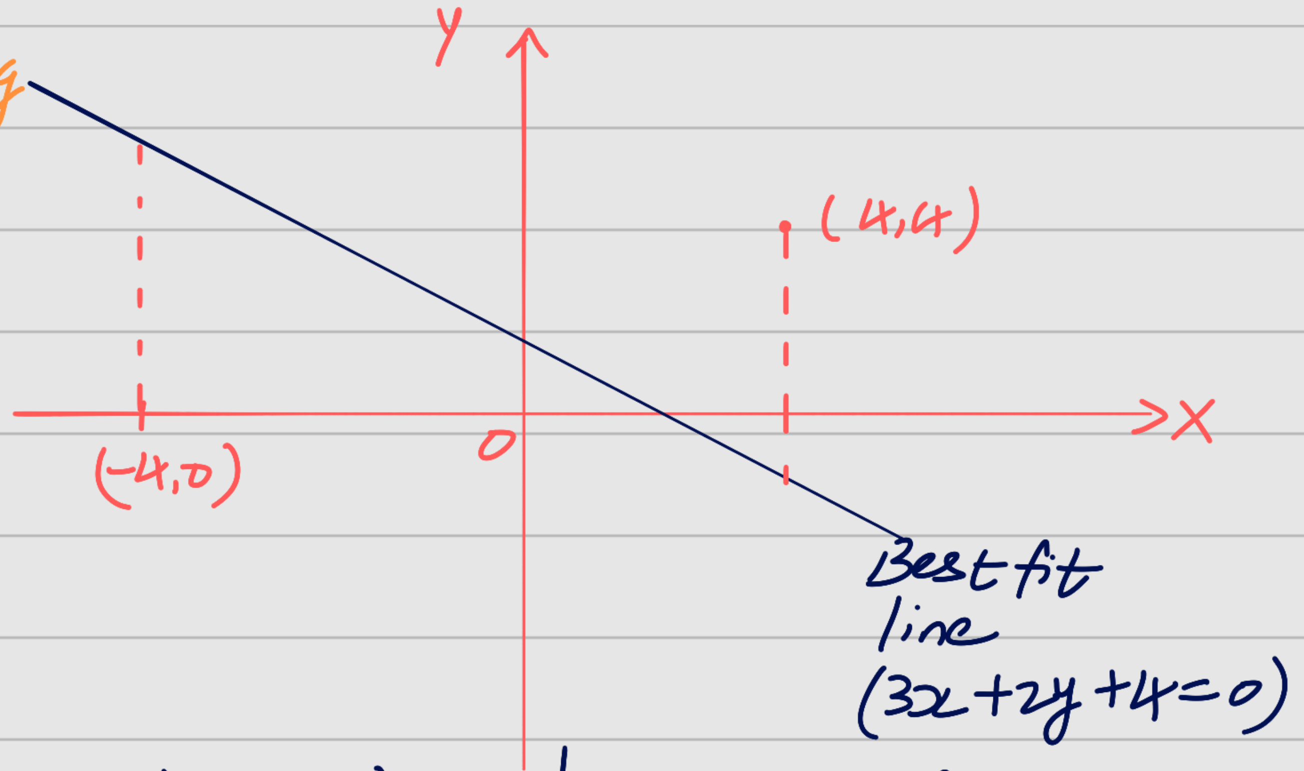


Distance b/w them should be MAX

→ Hard marginal Plane (All points belong to corresponding category) } impractical

Soft marginal plane (Some points of other category belongs here)

Eg

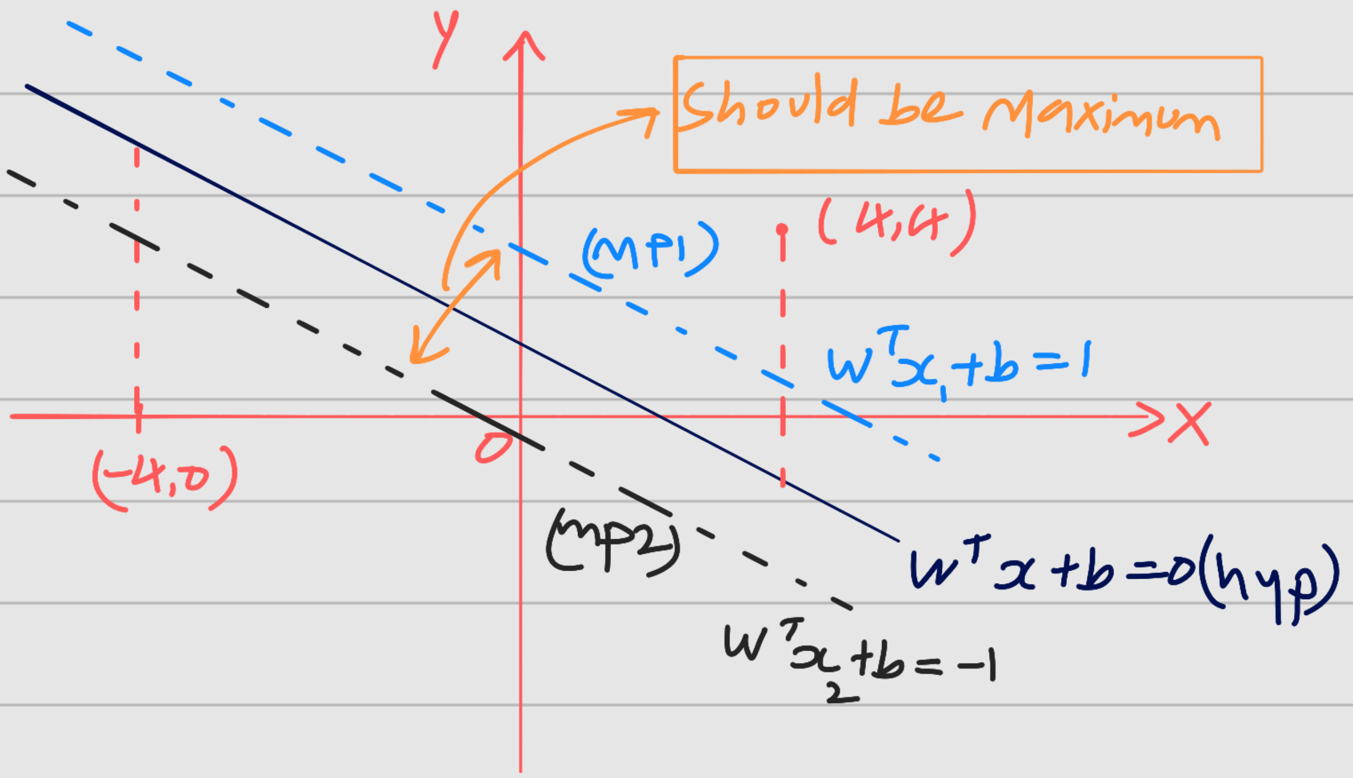


$3(-4)+2(0)+4$ | $3(4)+2(4)+4$

$-12+4=-8$ | $=24$

BELOW THE LINE = -ve Values

ABOVE THE LINE = +ve values



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

$w^T \rightarrow$ slope (magnitude & direction)

$\hookrightarrow$ dividing this by $|w|$ will give direction

$$\frac{w^T}{|w|}(x_1 - x_2) = \underbrace{\frac{2}{|w|}}$$ } we will get w vector

we have to maximize it

{final aim: maximize $\frac{2}{|w|}$ by altering (w, b) }

marginal plane distance

<u>Constraints</u>

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

(data point)

For all accurate datapoints...

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

with constraint

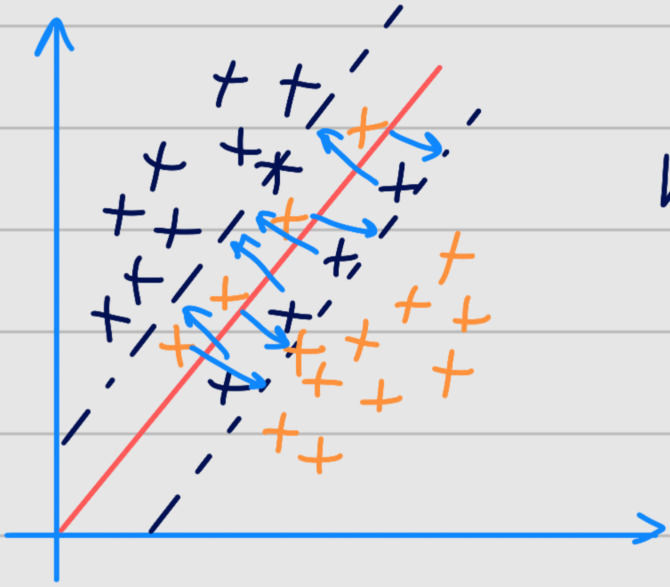
$\underset{(w,b)}{\text{Maximize}} \ \frac{2}{|W|}$ OR $\underset{(w,b)}{\text{Minimize}} = \frac{|W|}{2}$

loss f^n

$$\boxed{\text{cost } f^n = \min_{(w,b)} \left(\frac{|w|}{2}\right) + C_i \underbrace{\sum_{i=1}^{w} \eta_i}}$$

hyperparameter

points we want can avoid misclassification



here, c=8

$\eta_i \rightarrow$ Summation of distance of misclassified points from marginal plane (respective)

<u>Kernel f^n</u> $\rightarrow$ transforms data into a higher dimensional space to make it linearly separable. [1D $\rightarrow$ 2D $\rightarrow$ 3D $\rightarrow$ ...]

Types of Kernels:

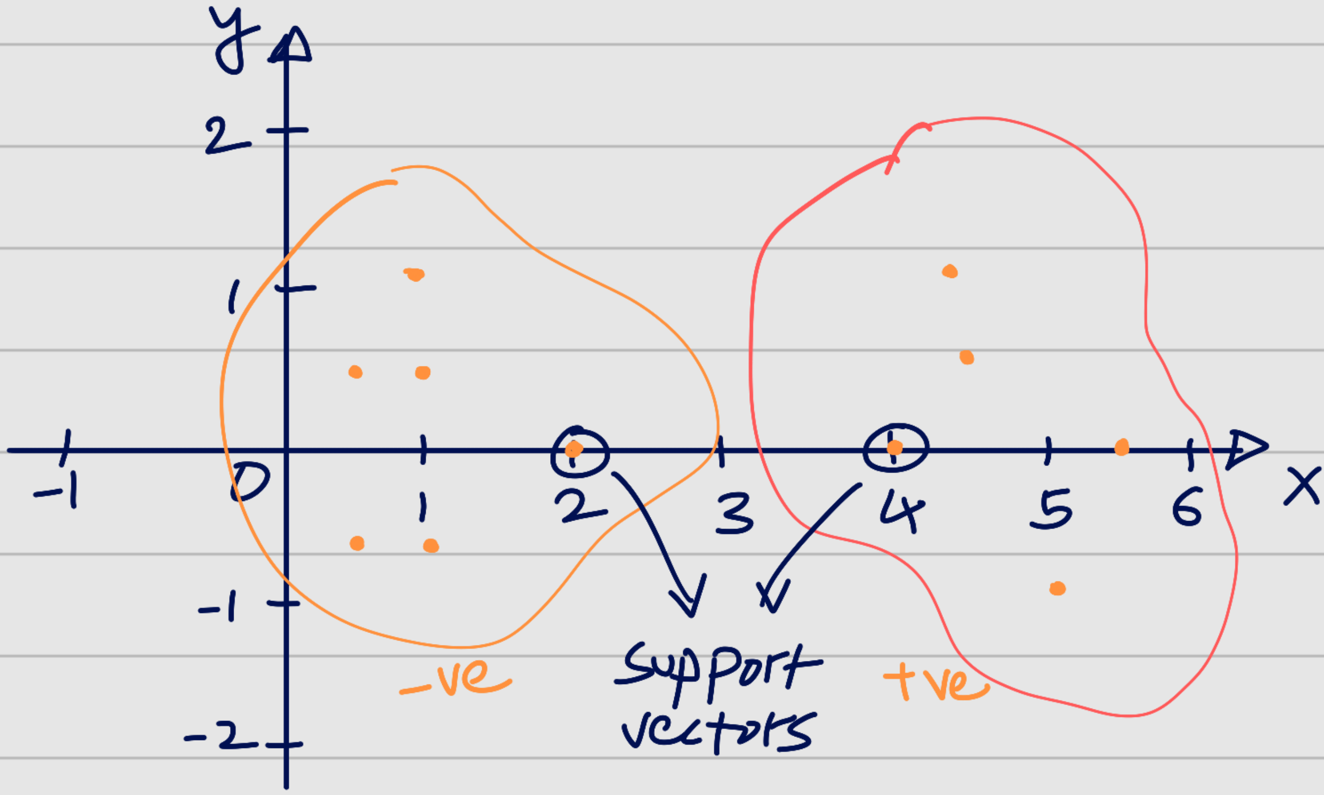
1. Linear Kernel → Suitable for linearly separable data
2. Polynomial Kernel → for curved boundries
3. Radial Basis f^n → captures complex relationships

Eg consider following set of data points given, find the optimal hyperplane for these data points.

(1, 0.5) (1, 1) (1, -0.5) (-0.5, 0.5) (0.5, 0.5) (2, 0)
(4, 0) (4.5, 1) (4.5, 0.5) (5, -1) (5.5, 0)

Sol^n Steps

(i) Draw all points on plane
(ii) Find Support vectors
(iii) Add bias value in vectors & convert 2D to 3D
(iv) Consider parameters for every support vectors like $\alpha 1, \alpha 2$ etc.
(v) Find linear equations & solve them to find parameter values
(vi) Final eqn of hyperplane ($\bar{w} = \sum_{i=1}^{n} \alpha_i S_i$)



$$\overset{s_1}{\begin{pmatrix}2\\0\end{pmatrix}} \quad \overset{s_2}{\begin{pmatrix}4\\0\end{pmatrix}} \xrightarrow{\text{add bias}} \overset{s_1}{\begin{pmatrix}2\\0\\1\end{pmatrix}} \quad \overset{s_2}{\begin{pmatrix}4\\0\\1\end{pmatrix}}$$

α_1 $\qquad$ α_2

$$(\alpha_1 s_1)\cdot s_1 + (\alpha_2 s_2)\cdot s_1 = -1$$
$$(\alpha_1 s_1) s_2 + (\alpha_2 s_2) s_2 = 1$$

$$\alpha_1 \begin{pmatrix}2\\0\\1\end{pmatrix}\begin{pmatrix}2\\0\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}4\\0\\1\end{pmatrix}\begin{pmatrix}2\\0\\1\end{pmatrix} = -1$$

$$5\alpha_1 + 9\alpha_2 = -1 \rightarrow ①$$

$$\alpha_1 \begin{pmatrix}2\\0\\1\end{pmatrix}\begin{pmatrix}4\\0\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}4\\0\\1\end{pmatrix}\begin{pmatrix}4\\0\\1\end{pmatrix} = 1$$

$$9\alpha_1 + 17\alpha_2 = 1 \rightarrow (2)$$

<u>from (1) & (2)</u>

$$\hookrightarrow \alpha_1 = -6.5$$
$$\alpha_2 = 3.5$$

$$\bar{w} = \alpha_1 s_1 + \alpha_2 s_2$$

$$\bar{w} = -6.5\begin{pmatrix}2\\0\\1\end{pmatrix} + 3.5\begin{pmatrix}4\\0\\1\end{pmatrix}$$

$$\bar{w} = \begin{pmatrix}-13\\0\\-6.5\end{pmatrix} + \begin{pmatrix}14\\0\\3.5\end{pmatrix}$$

$$\bar{w} = \begin{pmatrix}1\\0\\-3\end{pmatrix}$$

$\begin{matrix}1\\0\end{matrix}\Big\} \rightarrow$ orientation of line

- $\rightarrow \begin{pmatrix}1\\0\end{pmatrix}$ vertical
- $\rightarrow \begin{pmatrix}0\\1\end{pmatrix}$ horizontal
- $\rightarrow \begin{pmatrix}1\\1\end{pmatrix}$ 45°

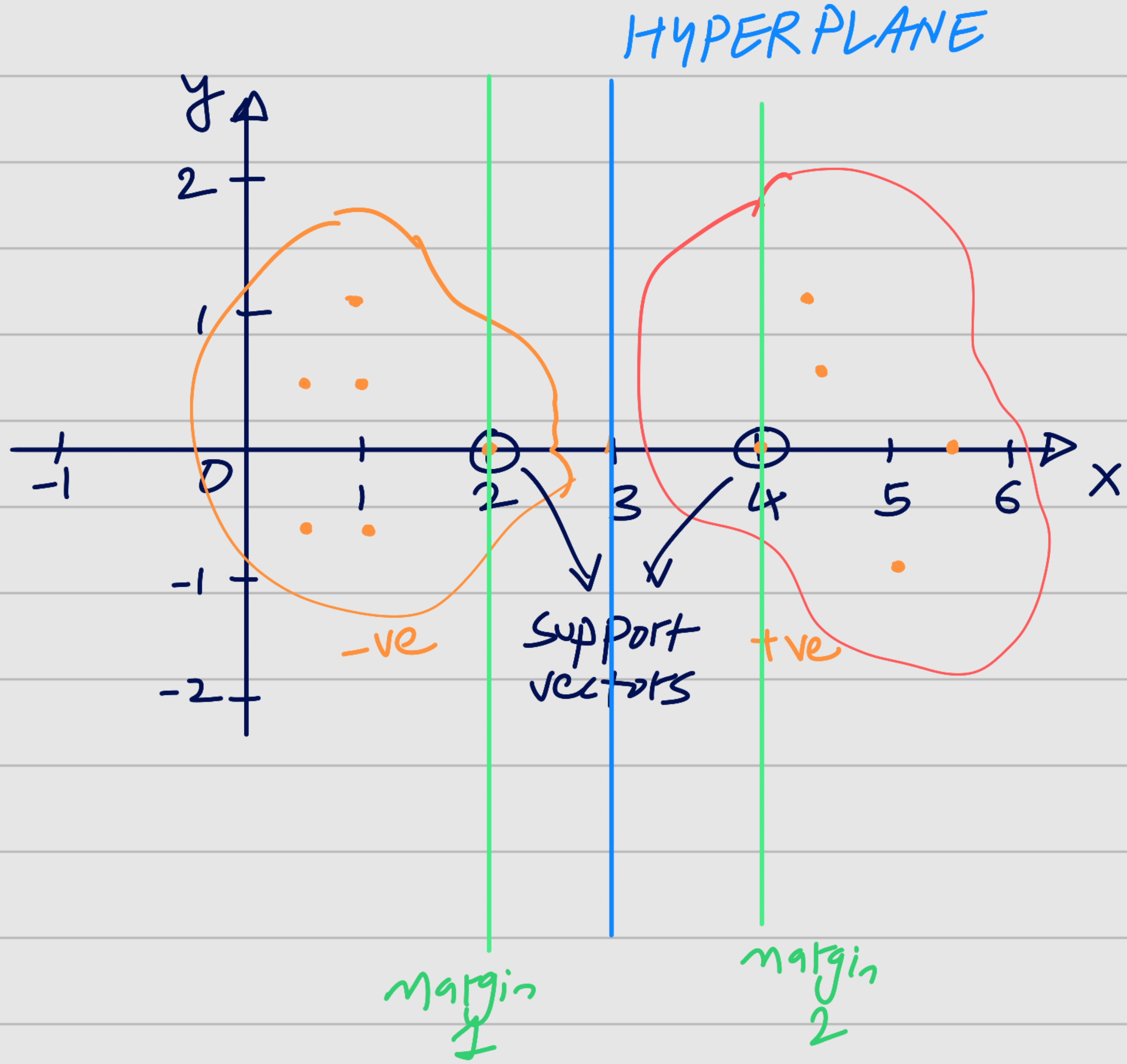
$-3 \rightarrow$ bias

$b = -3$

$b + \underline{3} = 0$

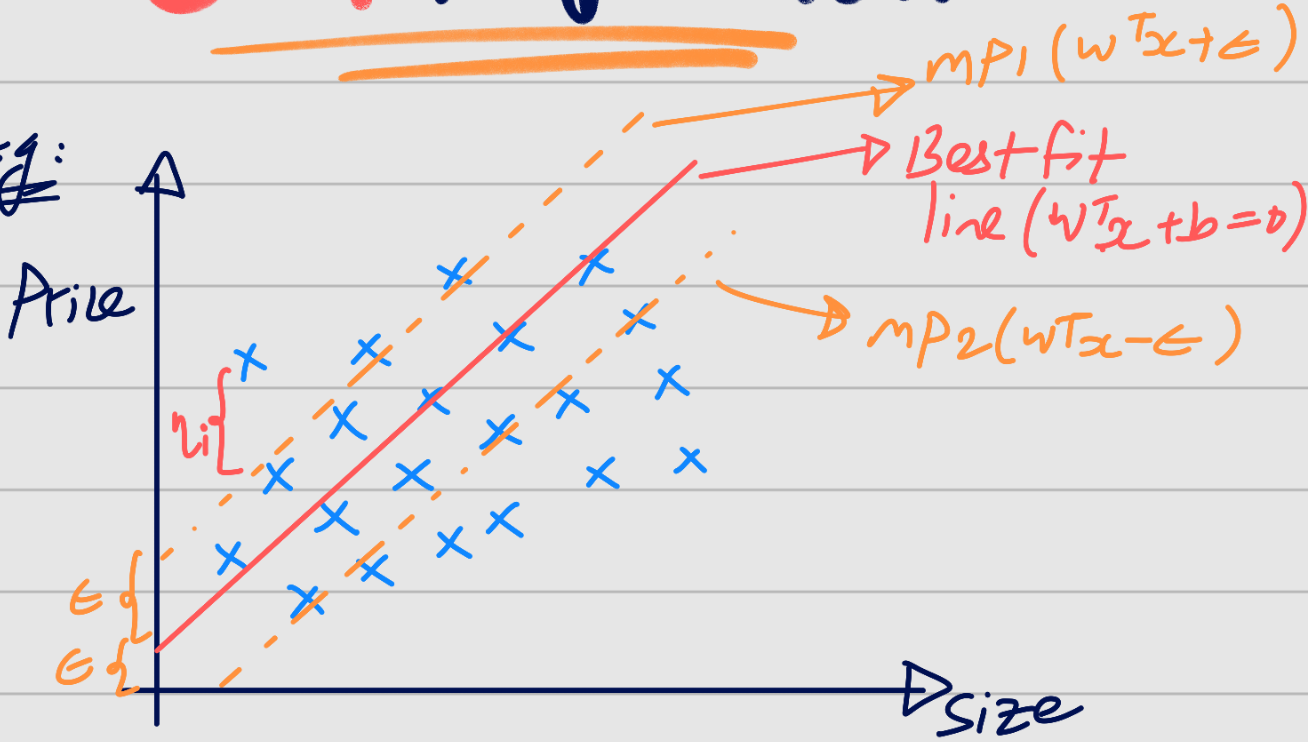
$\downarrow$

(x coordinate)


HYPERPLANE
y
2
1
-1
-2
-1
0
1
2
3
4
5
6
X
-ve
support vectors
+ve
margin 1
margin 2

SVM Regression

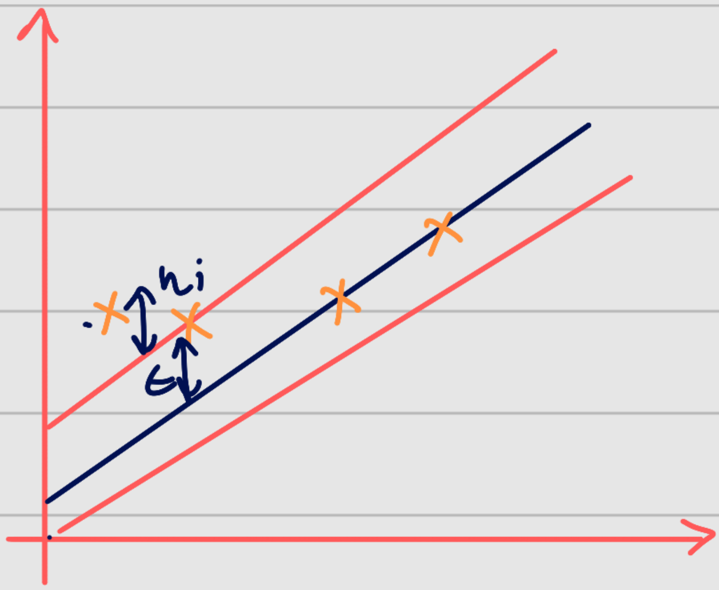
Eg:



First we create a Best fit line

$$\text{constraint: } \left|y_i - \underbrace{w^T x_i}_{\hat{y}}\right| \le \epsilon + |\eta_i| \quad \Big\}\ \text{good thing}$$

$$\text{cost } f^n : \underset{(w,b)}{\text{minimize}} \ \frac{|w|}{2} + \underbrace{C\sum_{i=1}^{n}|\eta_i|}_{\text{hyperparameter}}$$



h_i
ε