Loss Minimization in Classification

Agenda

- Recap
  - Distance b/w origin and Line
  - Distance b/w point and Line
- Distance b/w 2 planes
- Putting it all together: Loss function
  - Gain function: Maths
  - Convert gain function to Loss function
- Weight update:
  - The perceptron Algo
  - Coding the Perceptron
- Non-Linear Decision Boundary: Circle

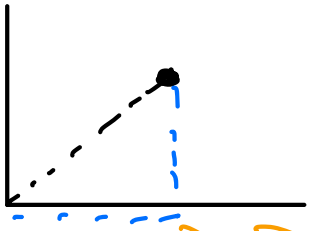
Recap

a) Vectors :

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad x = [\,1, 2, 3\,]$$

$$\|x\| \Rightarrow \sqrt{1^2+2^2+3^2}$$

b) Norm :



$\|x\| \Rightarrow$ L2 Norm $= \sqrt{\sum_{i=1}^{n} x_i^2}$

$\Rightarrow$ L1 Norm $\Rightarrow \sum_{i=1}^{n} |x_i|$

c) Dot product :

$x = [\,1, 2, 3\,]$

$y \Rightarrow [\,1, 2, 3\,]$

$x \cdot y$

$1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3$

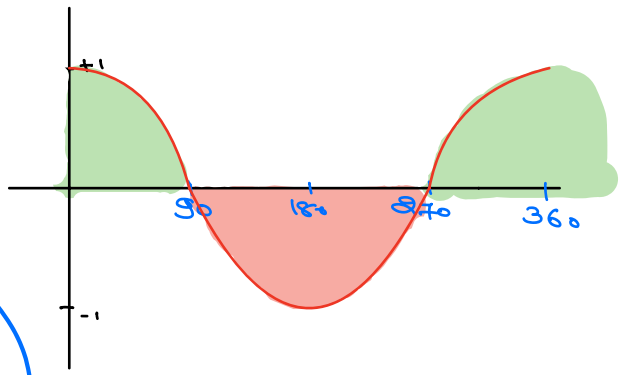
14

d) Angle b/w Vectors :

$\vec{x}$ and $\vec{y}$

$$\cos\theta = \frac{x^T \cdot y}{\|x\|\|y\|}$$

$$\theta = \cos^{-1}\left(\frac{x^T \cdot y}{\|x\|\|y\|}\right)$$



e) Eqn of Line and HalfSpaceeg

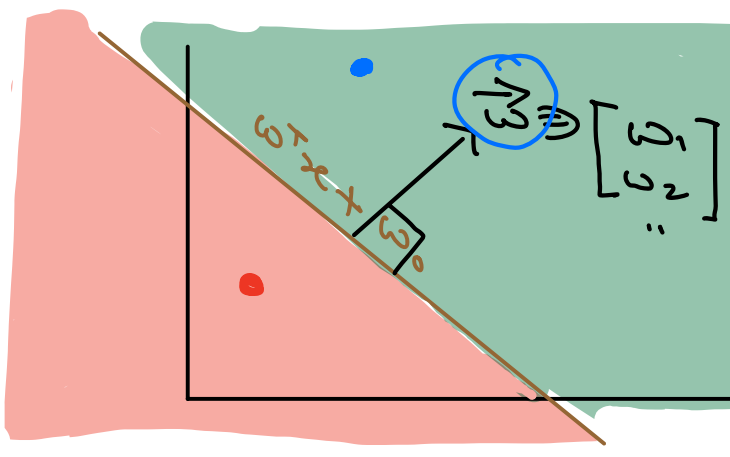
$y = mx + c$

$w_1 x + w_2 y + w_0 = 0$

$w_1 x_1 + w_2 x_2 + w_0 = 0$

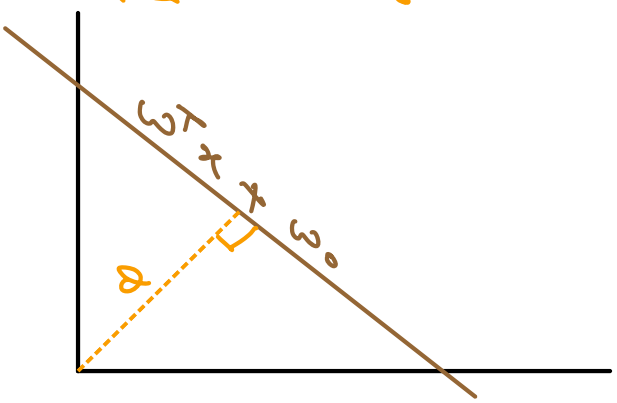
$w^T x + w_0 = 0$

$\vec{w} \Rightarrow \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$



c) Dist b/w origin and Line:

$$\frac{-\omega_0}{\|\omega\|}$$

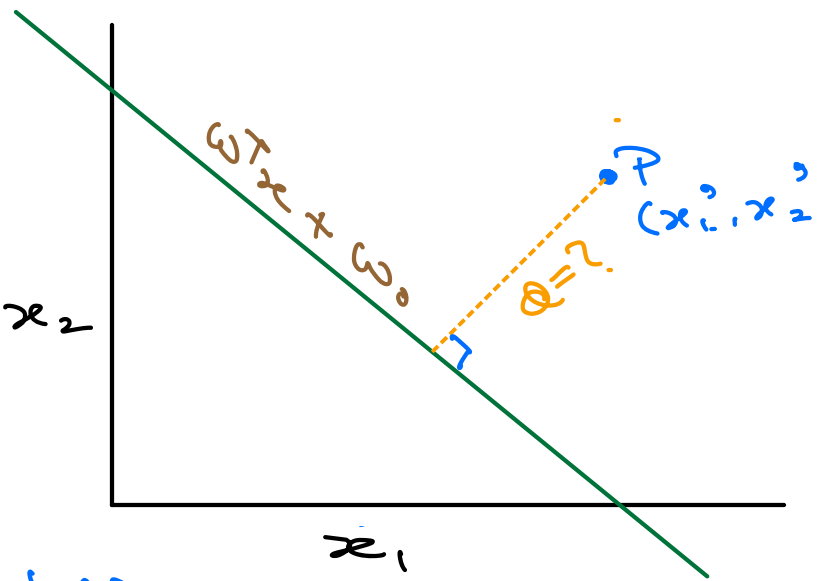


$2x + 3y - 6 \Rightarrow +(d)$

$2x + 3y + 6 \Rightarrow -(d)$

d) Dist b/w point and Line

$$d \Rightarrow \frac{\omega^T p + \omega_0}{\|\omega\|}$$



$$\Rightarrow \frac{\omega_1 x_1' + \omega_2 x_2' + \omega_0}{\|\omega\|}$$

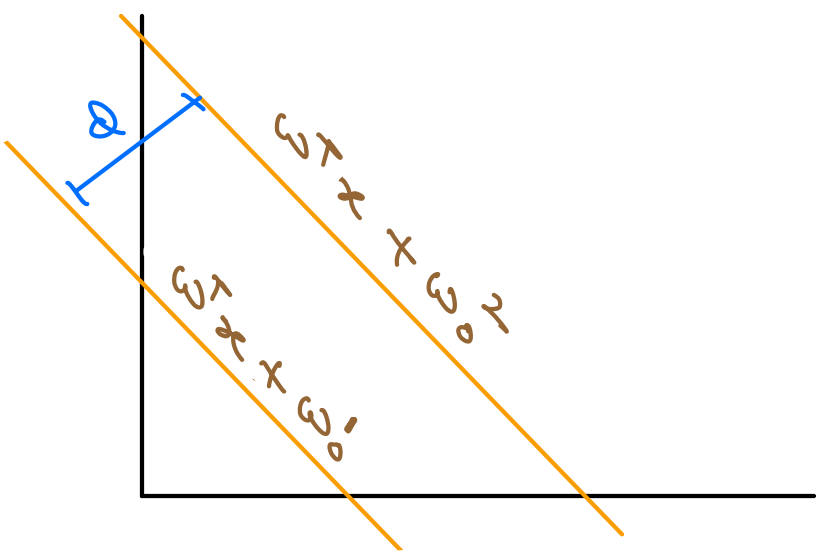
Q) Eq^n $\Rightarrow 2x - 3y + 6z = 2$ point = $(-6, 0, 0)$

ω ↓ plane, p ↓ 3D Coordi

$$\begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix} \quad \begin{bmatrix} -6 \\ 0 \\ 0 \end{bmatrix}$$

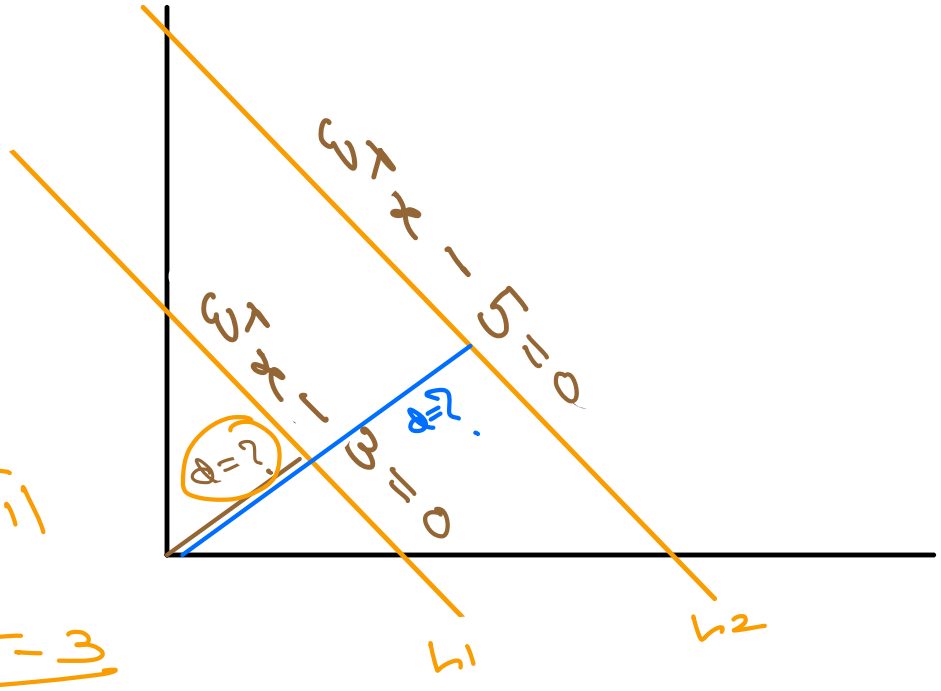
$$\frac{-12 + 0}{2^2 + 3^2 + 6^2} \Rightarrow \frac{-12 - 2}{7} \Rightarrow \frac{-14}{7} = -2$$

Distance b/w 2 Hyperplanes



$$\left| \frac{-w_0^2 + w_0^1}{\|w\|} \right|$$

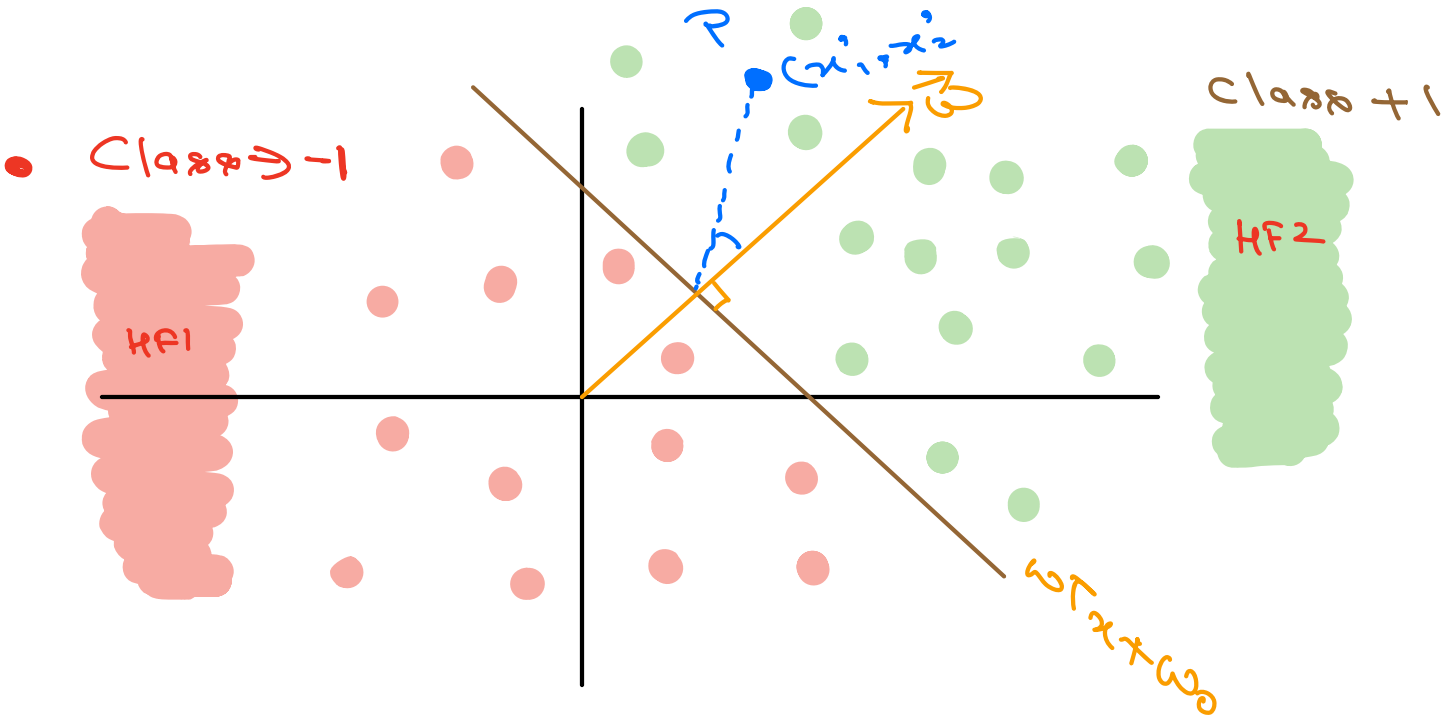
$$\left| \frac{w_0^1 - w_0^2}{\|w\|} \right|$$



$d_1 = \frac{3}{\|w\|}$ $\quad d_2 = \frac{5}{\|w\|}$

$d_2 - d_1 \Rightarrow \frac{5-3}{\|w\|}$

HalfSpace a Point belongs to?



$P \Rightarrow [x_1', x_2']$?

① $\theta < 90$

$\cos\theta \Rightarrow$ +ve

② $\theta > 90$ and $\theta < 180$

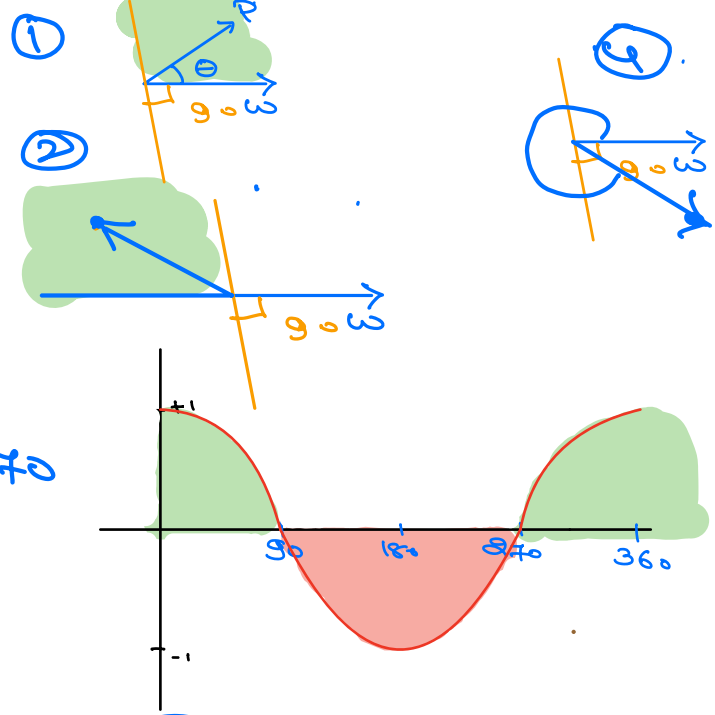
$\cos\theta \Rightarrow$ −ve

③ $\theta > 180$ and $\theta < 270$

$\Rightarrow \cos\theta \Rightarrow$ −ve

④ $\theta > 270$

$\cos\theta \Rightarrow$ +ve



$$\cos\theta \Rightarrow \frac{a \cdot b}{|a|\,|b|}$$

$$\cos\theta \Rightarrow \frac{w^T p}{|w|\,|p|}$$

$\cos\theta$ is +ve when $w^T p > 0$

$\cos\theta$ is −ve when $w^T p < 0$

$\cos\theta$ is 0 when $w^T p = 0$

$a \Rightarrow$ +ve the point is Classified as Class +1

$a \Rightarrow$ −ve the point is Classified as Class −1

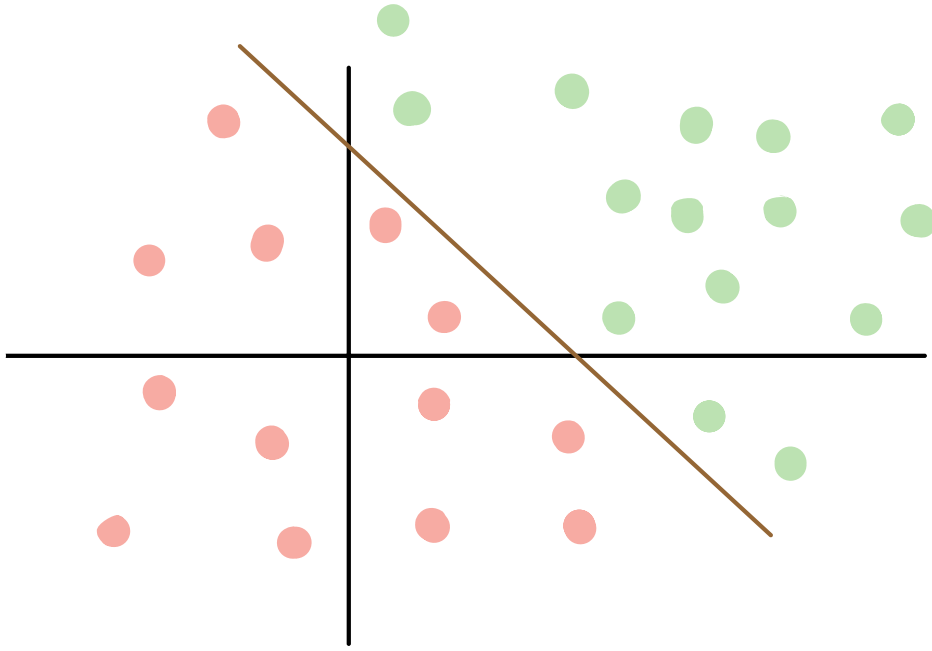
← ML Model

② Eqn of line $\Rightarrow 3x_1 + 4x_2 + 1 = 0$

Feature Vector = X ⇒

	x_1	x_2	Pred-Class
①	1	4	+ve +1
②	−2	−1	−ve −1
③	3	−1	+ve +1

Putting it all Together
Loss Function



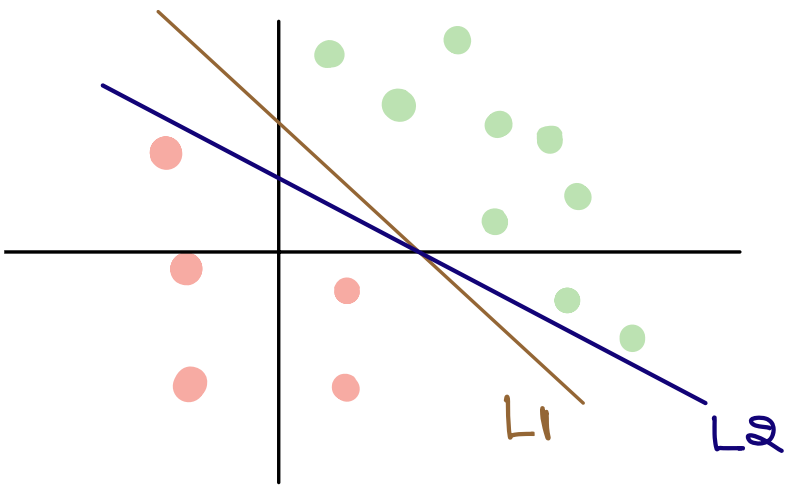
① How are labels assigned to New Datapoint

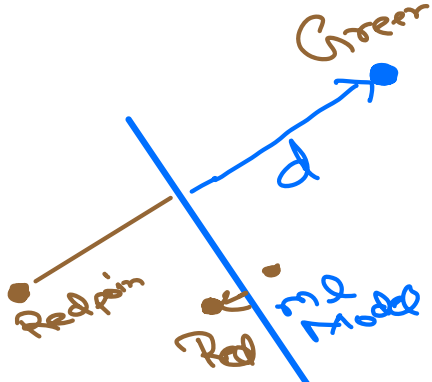
$$d = \frac{w^T x + w_0}{\|w\|} \rightarrow \begin{matrix} +ve \\ -ve \end{matrix}$$

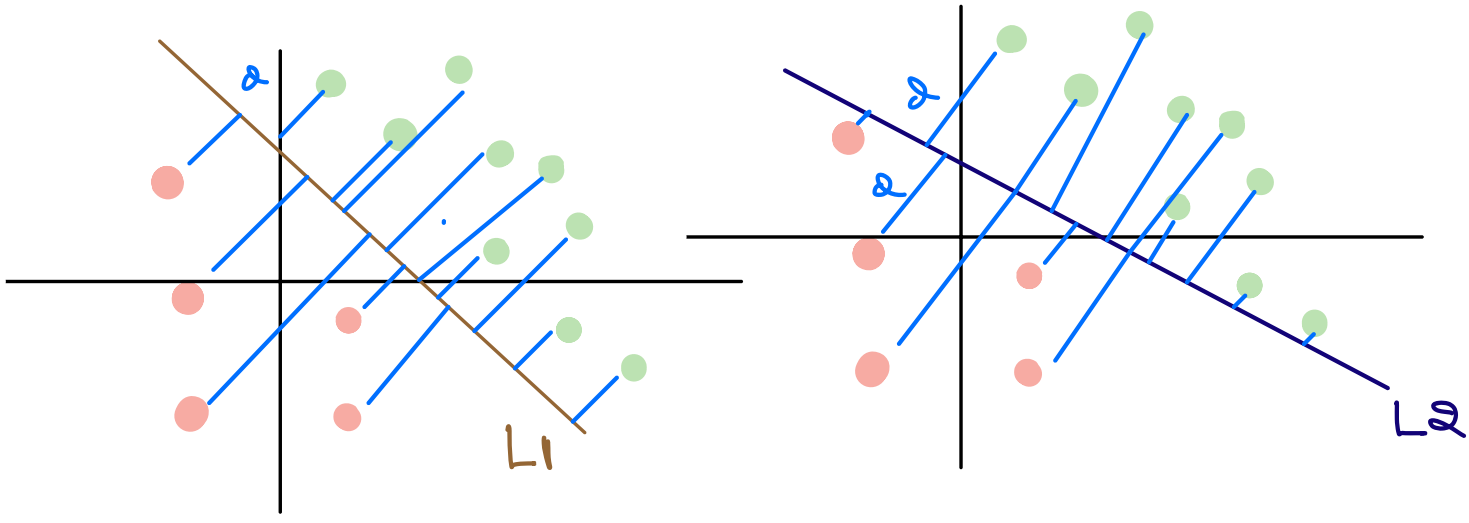
if $w^T x + w_0 > 0$:
+ve

else
-ve :

② Which line is better: L1 or L2?





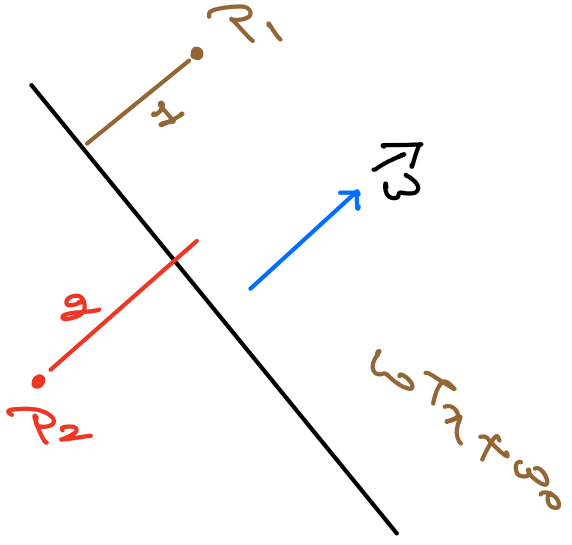


→ Sum of all distances

Sum(L1) Sum(L2)

L1 is better if Sum(L1) > Sum(L2)

Sum of Dist $= \sum_{i=1}^{M} \left(\frac{\vec{w}^T x_i + w_0}{\|w\|} \right)$



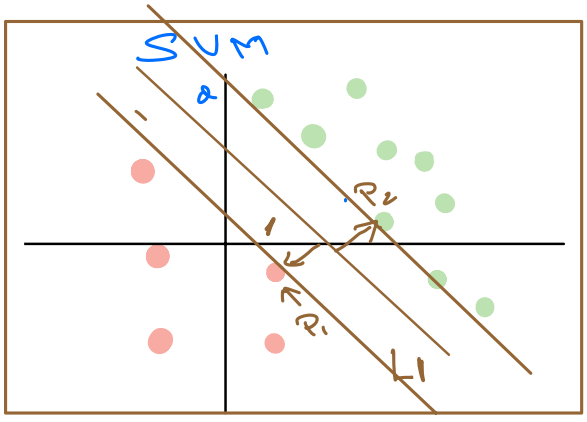
$G(x, \vec{w}, w_0) = \sum_{i=1}^{M} \left(\frac{\vec{w}^T x_i + w_0}{\|w\|} \right)$

↓

Gain function

(find value of $\vec{w}$ and w_0

To maximize SOD



L → G

In Maths :

① Dataset : $X = \{(\vec{x}_i, y_i)\}_{i=1}^{n}$ (features, labels)

② Classifier : $\vec{\omega}^T \vec{x} + \omega_0 = 0$
or
Model

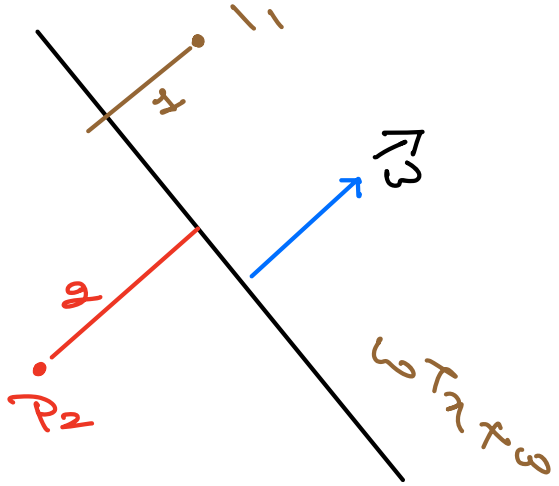
③ Loss function : – Gain Function

– (Sum of distances b/w CL and $\vec{x}$)

$$\text{Loss fn} = -\sum_{i=1}^{n} \left(\frac{\vec{\omega}^T \vec{x}_i + \omega_0}{\|\vec{\omega}\|} \right) \times y_i$$

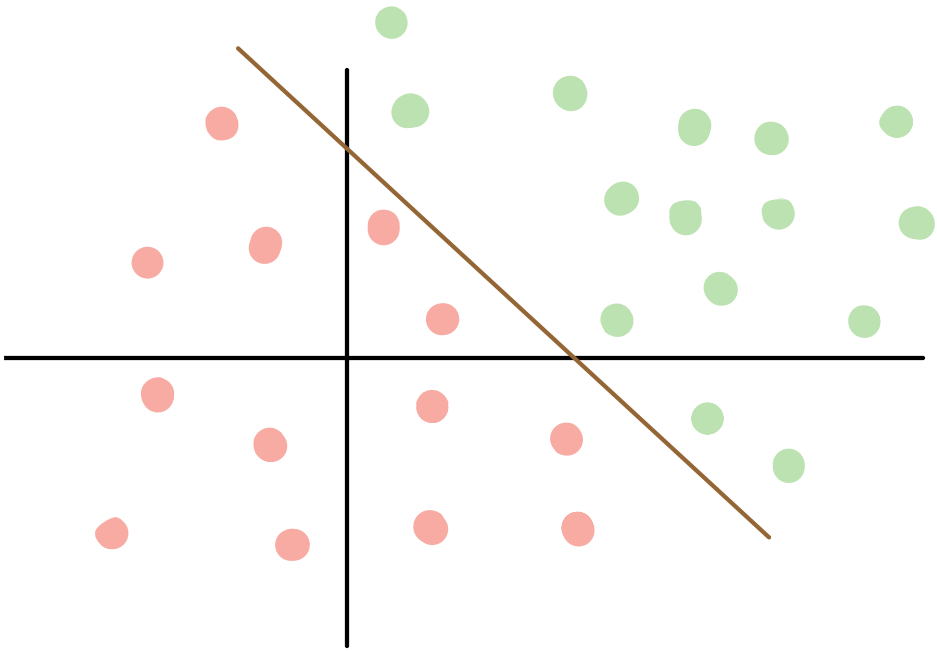
y_i : actual P_0

What is significance of this



How do we reduce Misclassification?

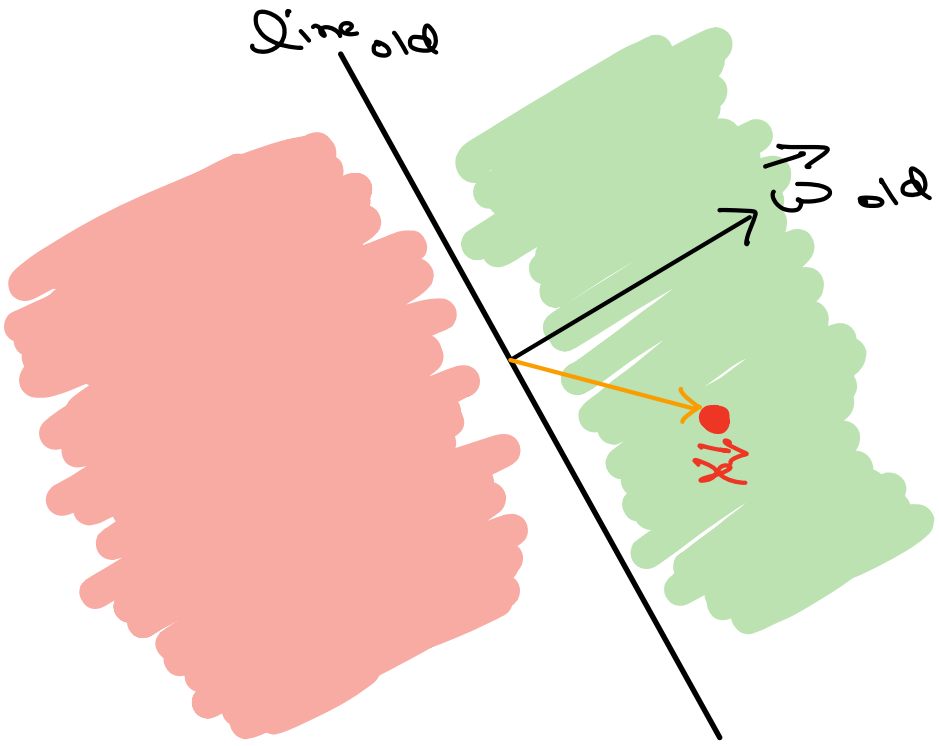
Perceptron Learning Algorithm



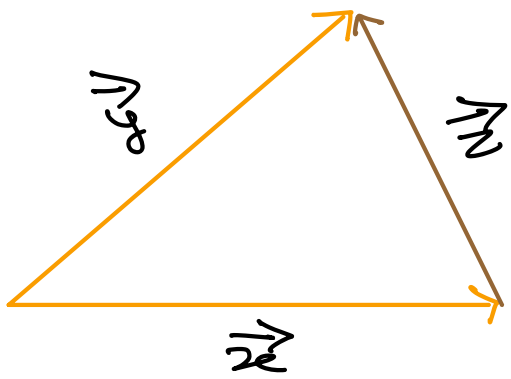
x_1	x_2	$\hat{y}$

Q: How do we identify that a point is misclassified?

Q2: How do we reduce misclassification?



Vector addition and Subtraction

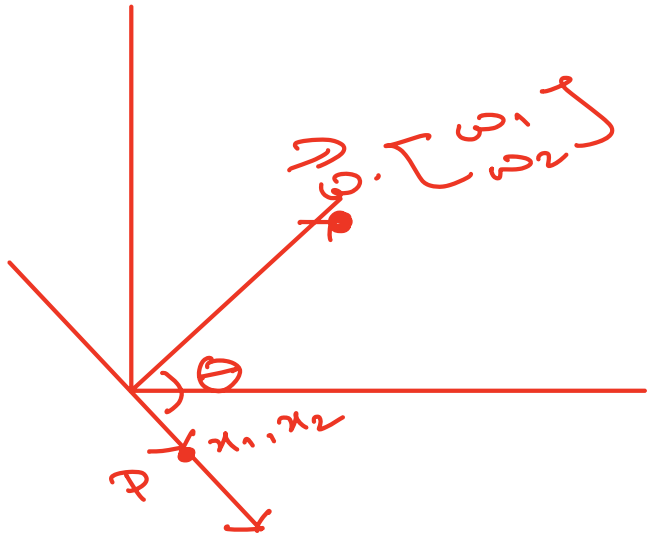


① $y_{actual} = -1$

$y_{label} = +1$

② $y_{actual} \Rightarrow -1$

$y_{label} \Rightarrow +1$



$\omega^T x = 0$

$\omega^T p = 0$

$$\cos\theta \Rightarrow \frac{\omega^T \cdot d}{\|d\| \|\omega\| \|p\|}$$

$$\cos\theta \Rightarrow \frac{0}{\|d\| \|\omega\| \|p\|}$$

$$\cos\theta = 0$$

$$\theta \Rightarrow \cos^{-1} 0 \Rightarrow 90^\circ$$