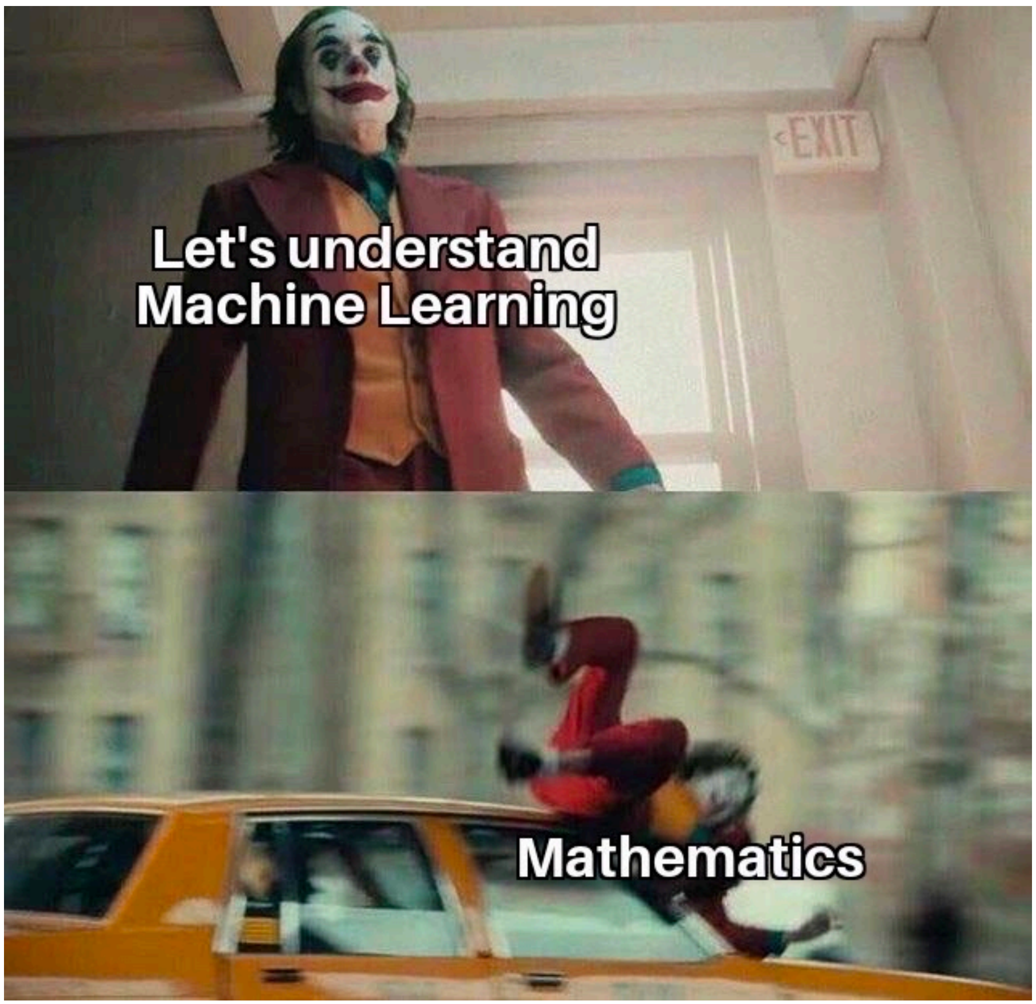
February 23, 2023

DSML : Math for ML.

Linear Algebra : Loss minimization in Classification



Recap:

(a) Vectors: $\bar{x} \in \mathbb{R}^d$, $\bar{x} = [x_1 \quad x_2 \quad \dots \quad x_d]$

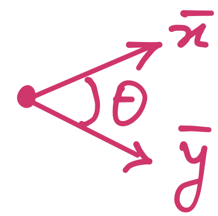
(b) Norm: $\|\bar{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_d^2}$

(c) Inner product: $\bar{x}, \bar{y} \in \mathbb{R}^d$

(Dot)

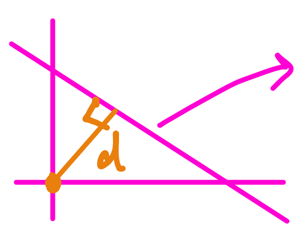
$$\bar{x} \cdot \bar{y} = \bar{x}^T \bar{y} = \sum_{i=1}^{d} x_i \cdot y_i$$

(d) Angle between vectors:



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

(e) Distance between origin and a line:

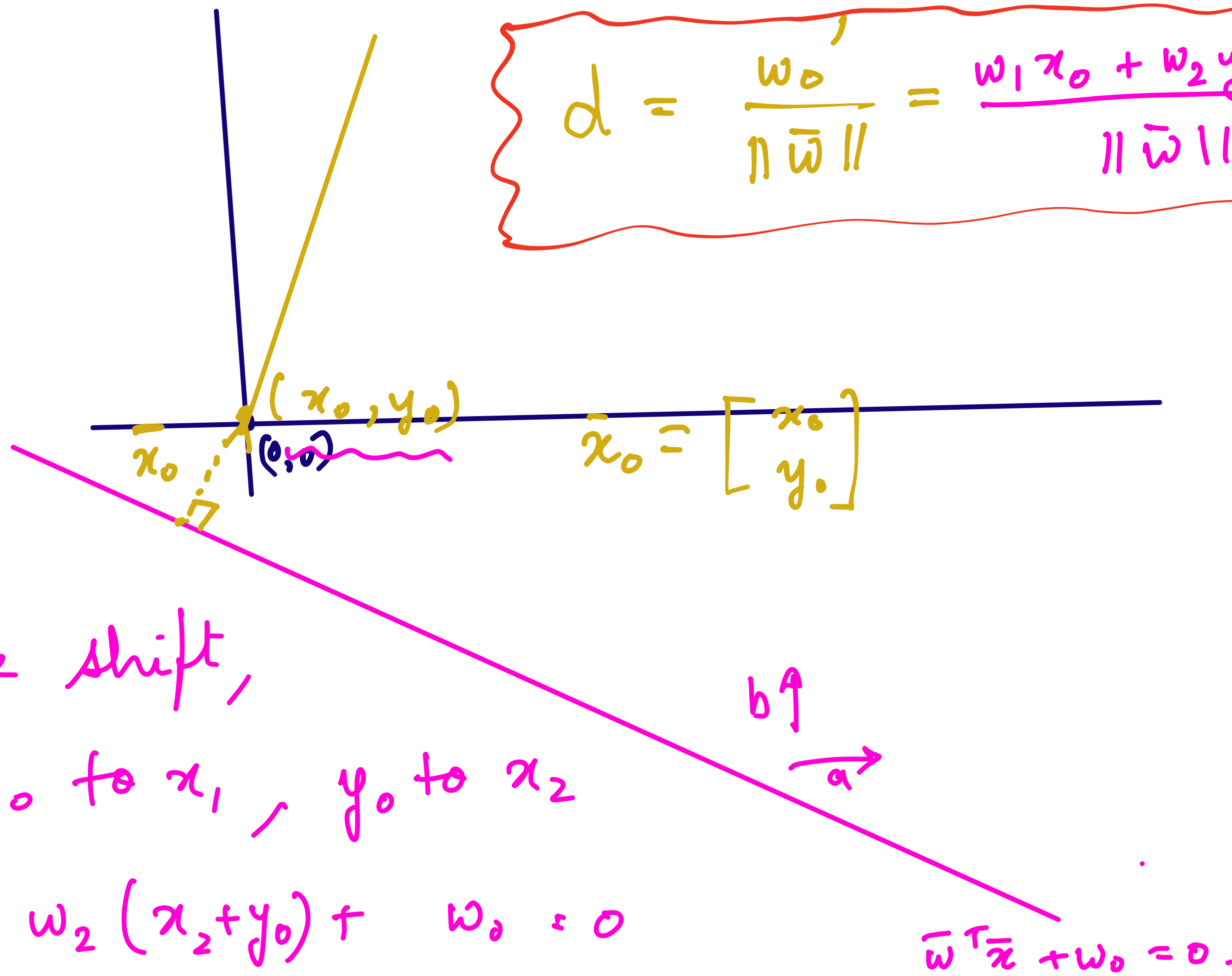


Line equation: $\bar{w}^T \bar{x} + w_0 = 0$

$$d = \frac{w_0}{\|\bar{w}\|}$$

Distance between a point and a line

$$d = \frac{w_0'}{\|\bar{w}\|} = \frac{w_1 x_0 + w_2 y_0 + w_0}{\|\bar{w}\|}$$



To make the shift,

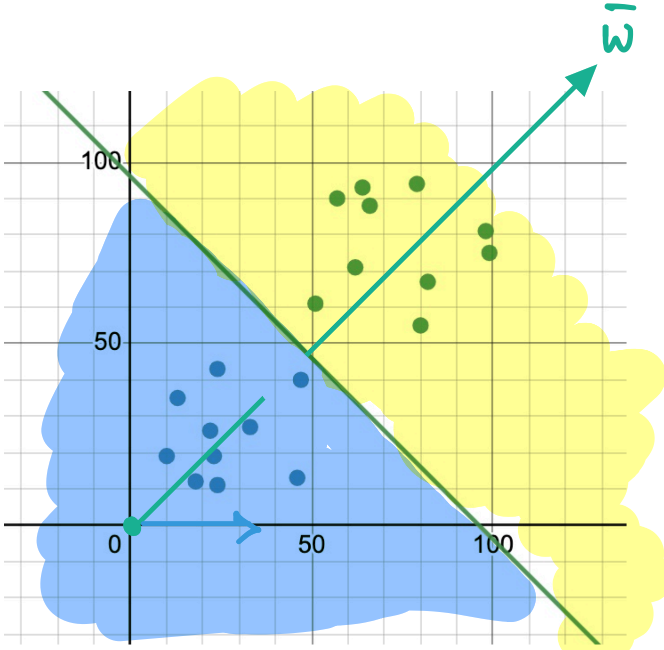
we add x_0 to x_1, y_0 to x_2

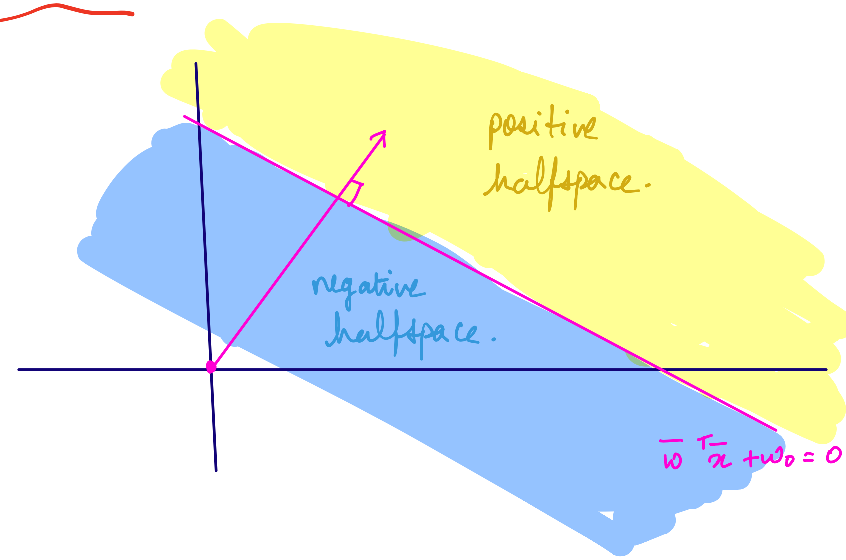
$$w_1(x_1 + x_0) + w_2(x_2 + y_0) + w_0 = 0$$

$$w_1 x_1 + w_1 x_0 + w_2 x_2 + w_2 y_0 + w_0 = 0.$$

$$w_1 x_1 + w_2 x_2 + \underbrace{(w_1 x_0 + w_2 y_0 + w_0)}_{w_0'} = 0.$$

How to find the halfspace a point belongs to.





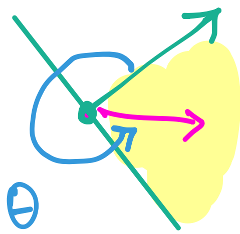
$$x + y - 98 = 0.$$

$$\bar{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$x = a \qquad \underbrace{1x + 0y}_{[1,\ 0]} - a = 0$$

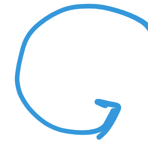
$$\Rightarrow \quad x - a$$

4]

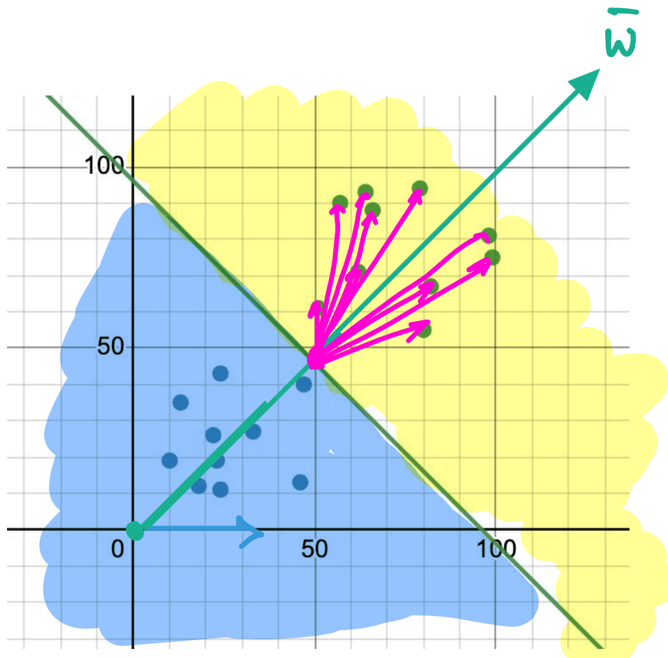


$270^\circ \le \theta \le 360^\circ$

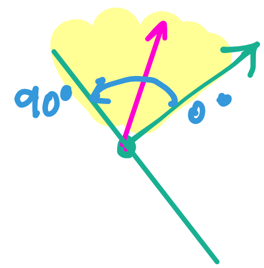
$\cos\theta \to +ve$



counter clockwise angles.



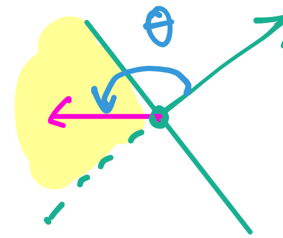
1]



$0^\circ \le \theta \le 90^\circ$

$\cos\theta \to +ve$

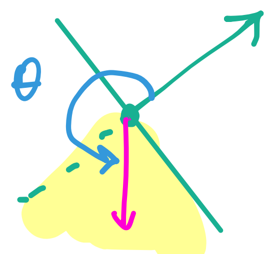
2]



$90^\circ \le \theta \le 180^\circ$

$\cos\theta \to -ve$

3]

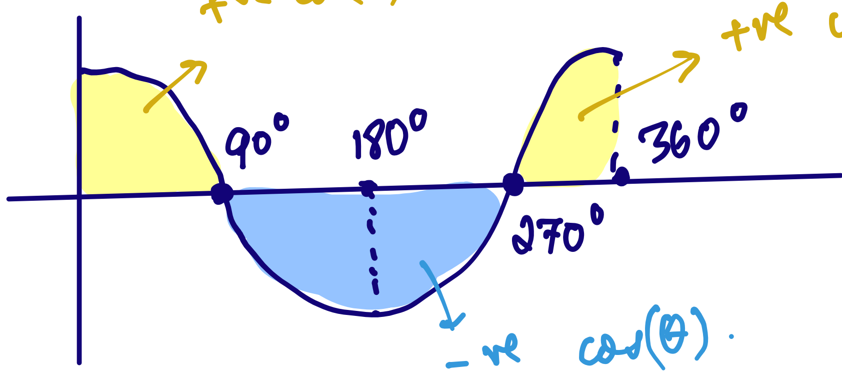


$180^\circ \le \theta \le 270^\circ$

$\cos\theta \to -ve$

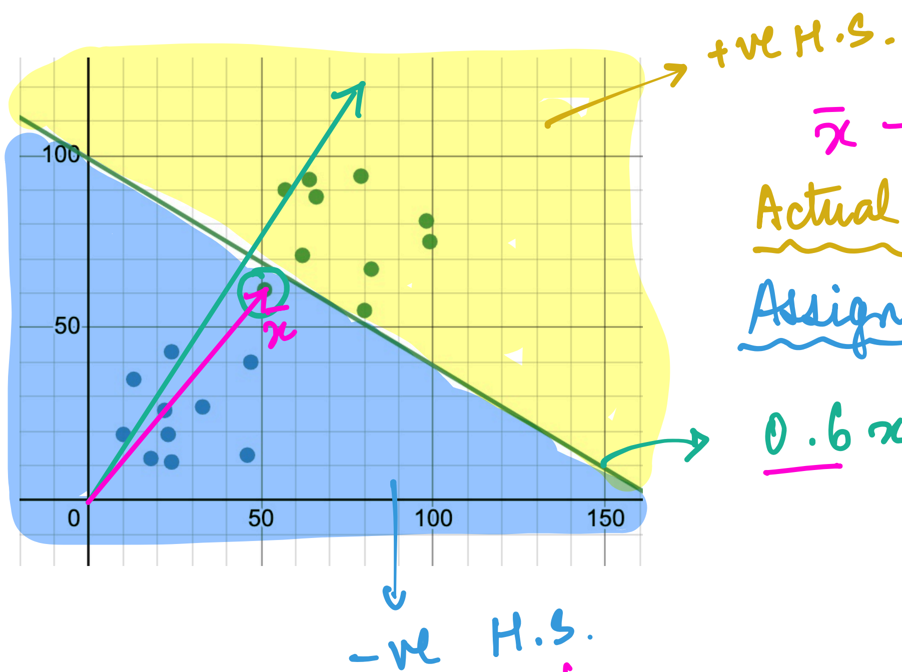
+ve cos(θ)

+ve cosθ



-ve cos(θ).

Putting it all together: Loss function.



$\bar{x} \rightarrow$ misclassified point.

Actual label : +1

Assigned label : − 1.

$0.6x + y - 99 = 0$

Q] How does a classifier assign labels?

$\bar{x} \rightarrow$ what is the label of this?

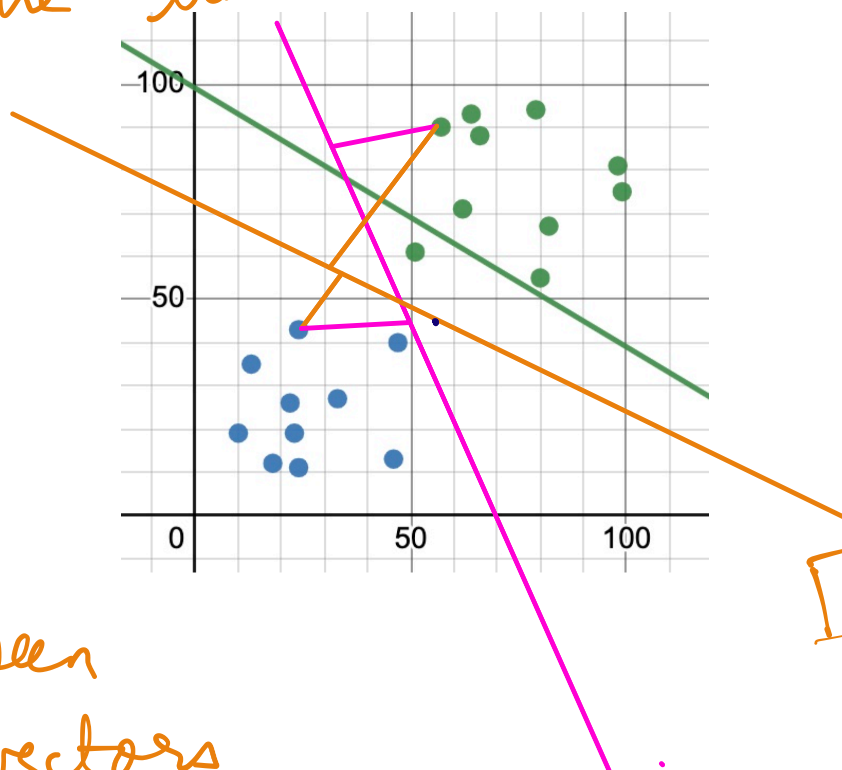
Ans. We first calculate $d = \dfrac{\bar{w}^T\bar{x} + w_0}{\|\bar{w}\|}$

If $d > 0 \rightarrow$ label $= +1$

If $d < 0 \rightarrow$ label $= -1$

Q2] Which line in the figure below is a better classifier?

Let's take the distances between our feature vectors and the line



We want to maximize the distances between our feature vectors and the line.

option 3

100% ✓

option2

100% ✓

In Math:

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

↳ feature vectors

$$y_i = \begin{cases} +1 & - \text{ +ve H.S.} \\ -1 & - \text{ -ve H.S.} \end{cases}$$

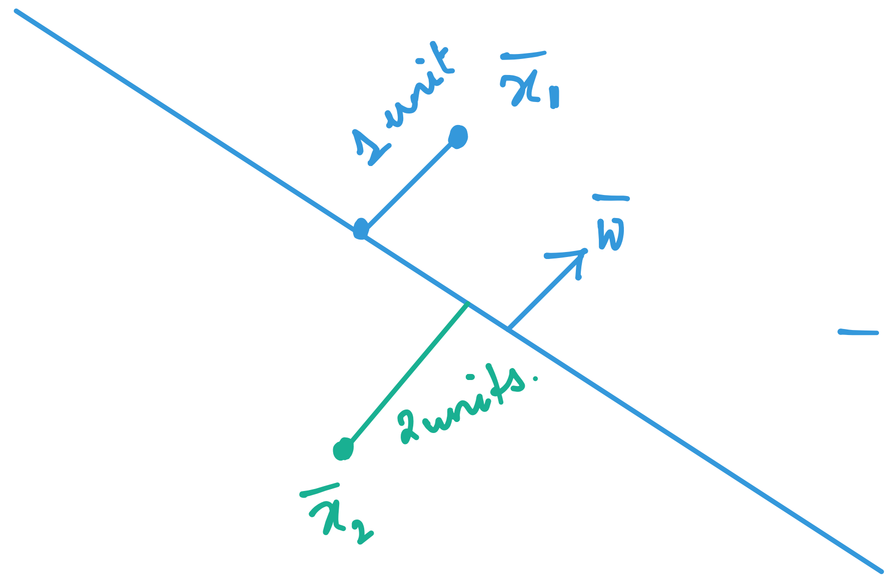
2. Classifier : $\bar{w}^T \bar{x} + w_0 = 0$.

3. Sum of all distances between classifier and feature vectors.

$$L(X; \bar{w}, w_0) = \sum_{i=1}^{n} \left(\frac{\bar{w}^T \bar{x}_i + w_0}{\|\bar{w}\|} \right) y_i$$

-ve

-ve



3 ← what we want.

−1 ← what we got!!

$$g(x; \bar{w}, w_0) = \cancel{3}. \quad +1 \quad + \quad -2.$$

$$= -1$$

1. ~~Take abs. value of distances.~~ $f(x) = |x|$. ← not differentiable
2. ~~Take square~~ ← Taking derivative of this is more complex.
3. Multiply by the labels. ✓

To automate the process of finding the best classifier, we must solve the following problem:

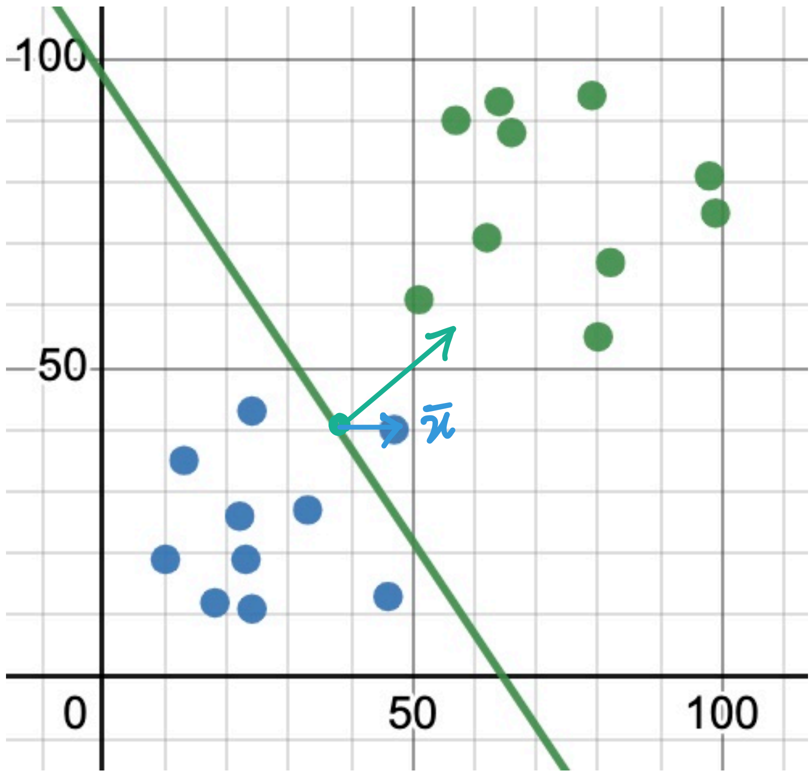
$$\max_{\bar{w}, w_0} g(X; \bar{w}, w_0)$$

$$= \max_{\bar{w}, w_0} \sum_{i=1}^{n} \left(\frac{\bar{w}^T \bar{x}_i + w_0}{\| \bar{w} \|} \right) y_i$$

How to get loss function from Gain function?

$$l(x; \bar{w}, w_0) = -g(x; \bar{w}, w_0)$$

Perceptron learning algorithm:



$x + y - 98 = 0.$

Q] How to detect if $\bar{x}$ is misclassified?

if $y_i == \text{sign}\left(\frac{\bar{w}^T \bar{x} + w_0}{\|\bar{w}\|}\right)$:
  print("correct")
else:
  print("Incorrect").

Q] How to update $\bar{w}$ to fix this?

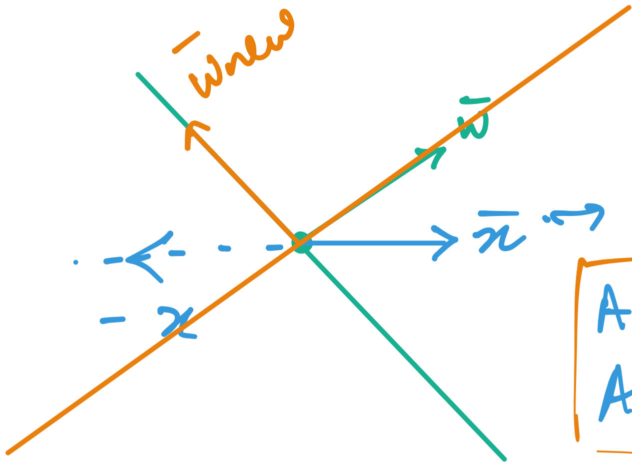
① $\boxed{\bar{w}_{new} = \bar{w} + \bar{x}}$

$\bar{x}$ is misclassified if, $\bar{x}$'s actual label : +1
and assigned label : −1

② $\boxed{\bar{w}_{new} = \bar{w} - \bar{x}}$

$\bar{x}$ is misclassified if, $\bar{x}$'s actual label : −1
and assigned label : +1

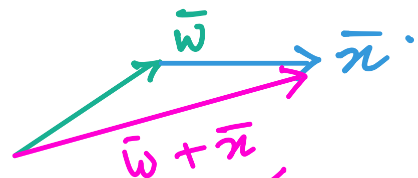
stop when $\bar{w}$ is not updated. →



$\bar{x} \rightarrow$ misclassified.

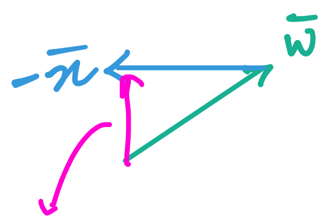
Actual label $\rightarrow -1$
Assigned label $\rightarrow +1$.

① $\bar{w}_{new} = \bar{w} + \bar{x}$

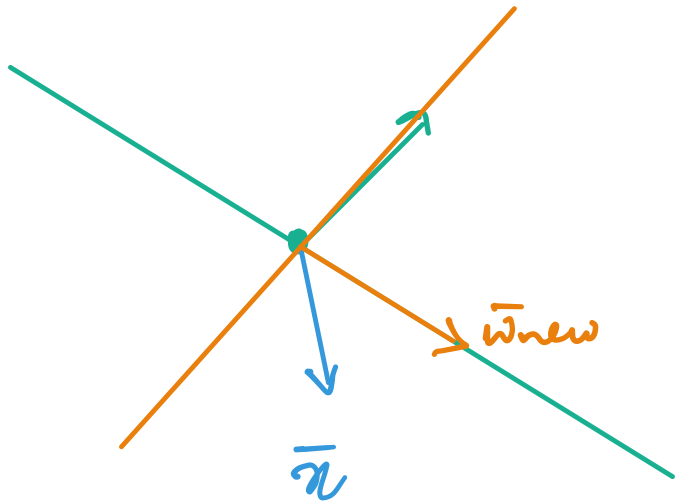


rotating clockwise.

② $\boxed{\bar{w}_{new} = \bar{w} - \bar{x}}$ ✓



$\bar{w} - \bar{x} \rightarrow$ rotating counter-clockwise.



Actual label : +1

Assigned label : -1

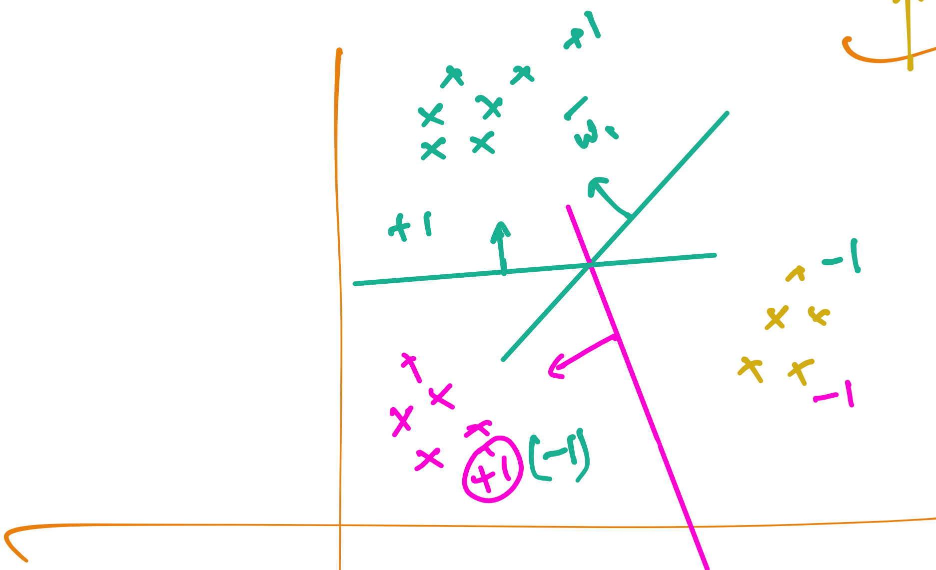
$$\bar{w}_{new} = \bar{w} + \bar{x}$$

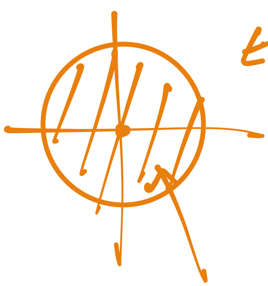
Perceptron learning code (approx.)

```
def assigned_label(x, w, w0):
    d = ( w.dot(x) + w0 ) / np.linalg.norm(w)
    if d > 0:
        return +1
    else:
        return -1

def learn(data, w, w0):
    x = data[0]
    y = data[1]

    count_updates = 1
    while count_updates > 0:
        count_updates = 0
        for (i, xi) in enumerate(x):
            if y[i] != assigned_label(xi, w, w0):
                if y[i] > 0:
                    w = w + x
                    count_updates += 1
                else:
                    w = w - x
                    count_updates += 1
    return (w, w0)
```

How to calculate y_i

$$(x-a)^2 + (y-b)^2 - r^2 = 0.$$





outside +ve.

Inside will be -ve