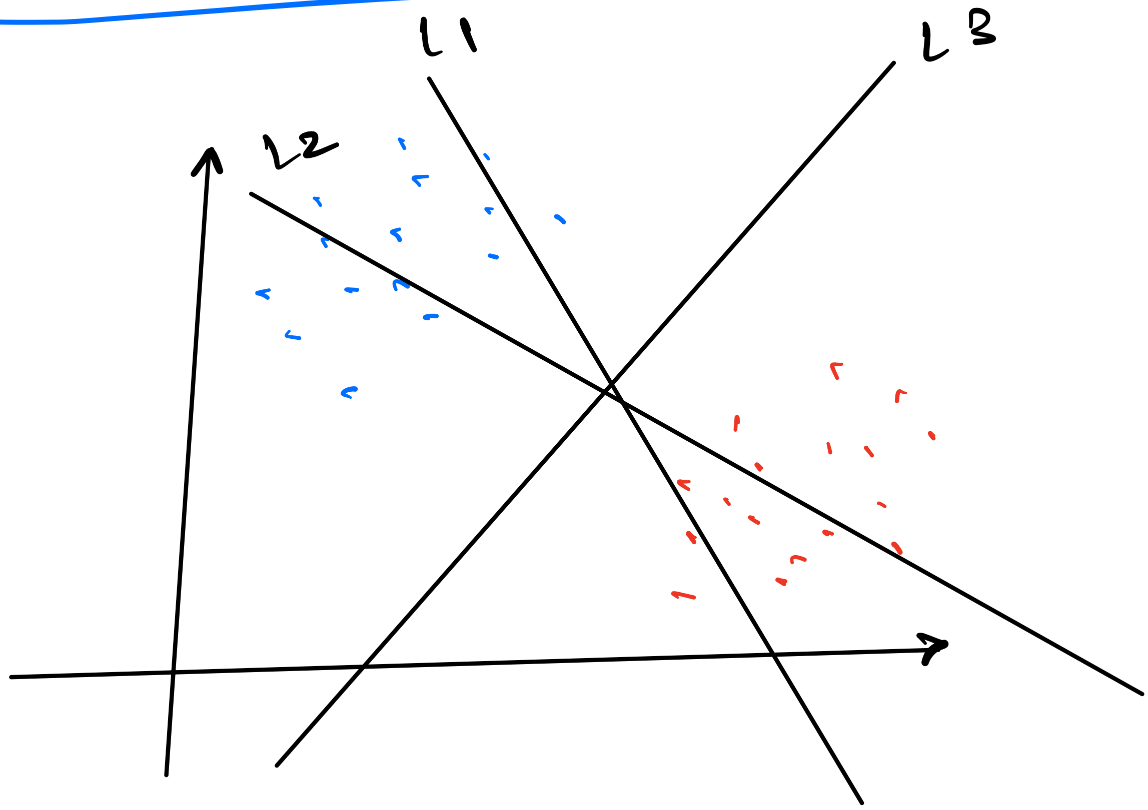
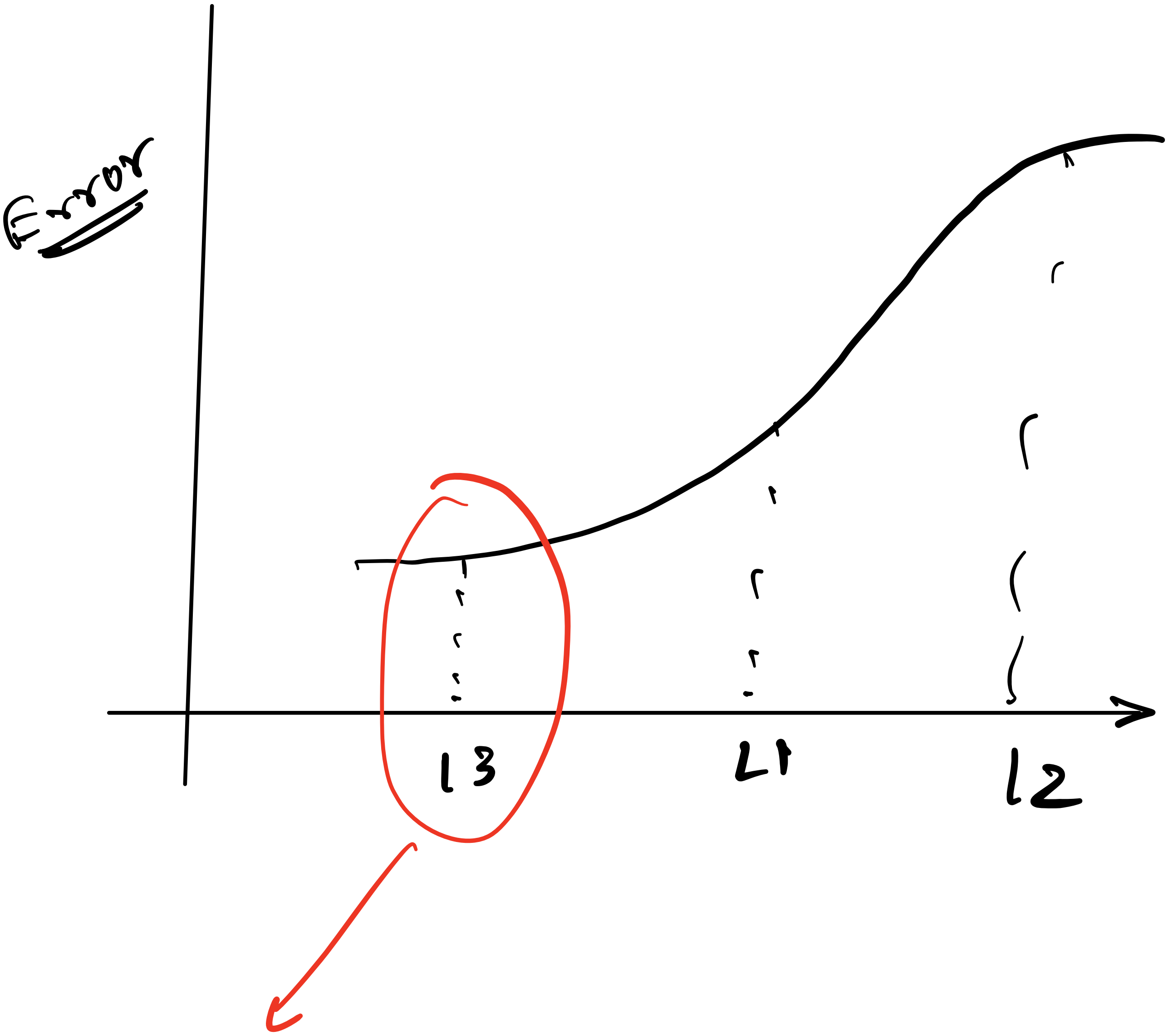
Optimization - 1

→ The need for Calculus in ML.

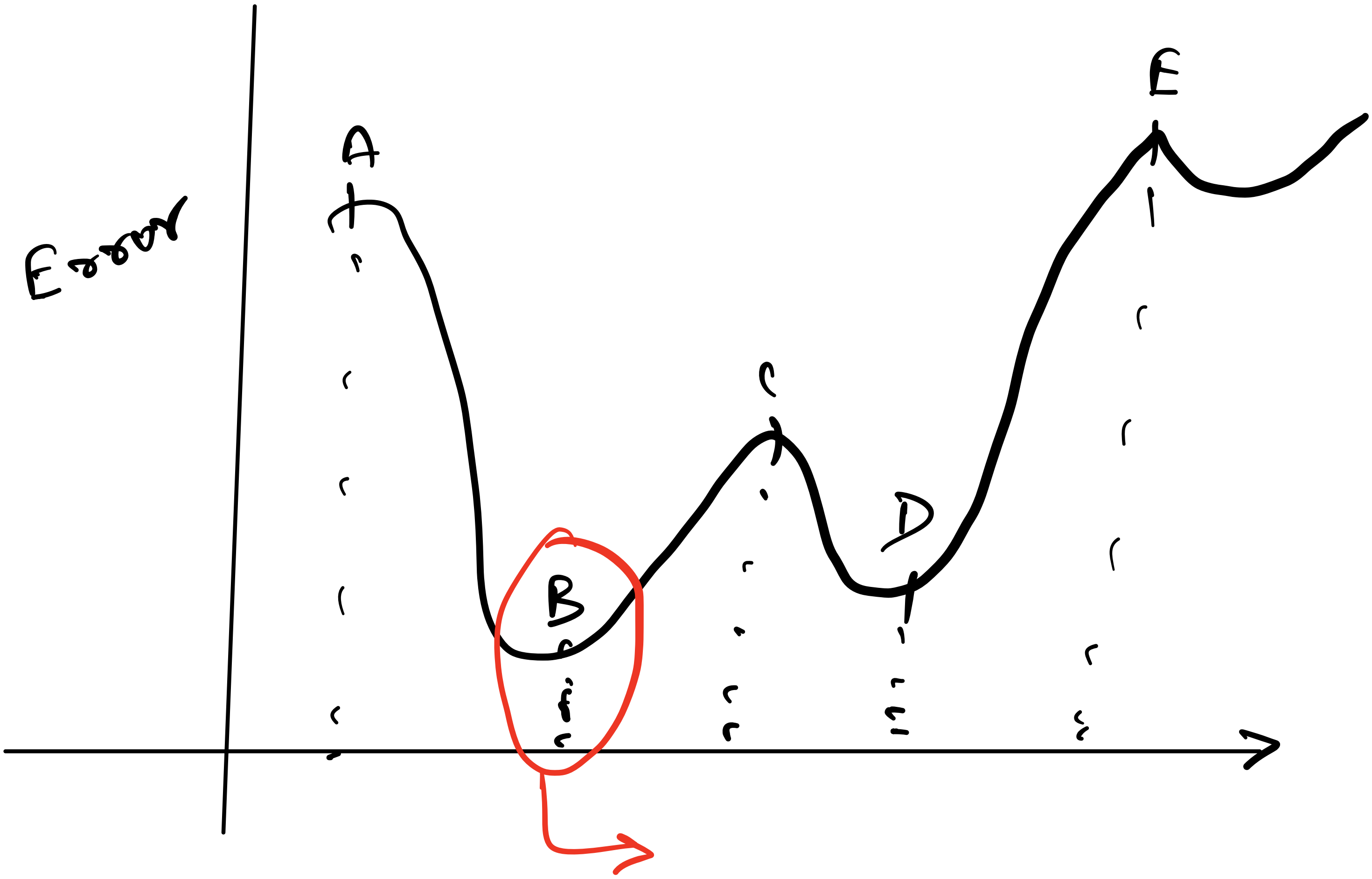
* Basic Intuition of a Classifier :



$$\boxed{L3 > L1 > L2}$$



Preferred among 3.



Calculus

$\vec{w}, w_0$

* Simple Searching Algo :

$\rightarrow$ Brute Force

eg:- $x_1, x_2 \longrightarrow w_1, w_2, w_0$

$w \longrightarrow [-10 \text{ to } +10]$

step-size = 0.1

$$\underbrace{[-10, -9.9, \dots -0.1]}_{100} \quad \underset{\downarrow \atop 1}{0} \quad \underbrace{[0.1, 0.2 \dots 10]}_{100}$$

$100 + 1 + 100 \longrightarrow \boxed{201}$

$$\left.\begin{array}{l} w_1 \longrightarrow 201 \\ w_2 \longrightarrow 201 \\ w_0 \longrightarrow 201 \end{array}\right]$$

Total Combinations :

$$\hookrightarrow (201) * (201) * (201)$$

$$= \underline{\underline{(201)^3}}$$

<u>Suppose</u> :-

$$1 \text{ op} \longrightarrow 10^{-6} \text{ sec}$$

$$(201)^3 \text{ op} \longrightarrow (201)^3 \times 10^{-6} \text{ sec}$$

$$\hookrightarrow \boxed{\underline{\underline{8.12 \text{ sec}}}}$$

3 features : x_1 x_2 x_3

↳

$(201)^4$ op → $(201)^4 * 10^{-6}$ sec

↳ $\boxed{27.2 \text{ mins}}$

w_1 w_2 w_3 w_0

↓ ↓ ↓ ↓

$201 \times 201 \times 201 \times 201$

4 features → $(201)^5$ op

↳ (91.3) hrs

How to solve this optimization Problem?

↓

★ Gradient Descent Algo

① Functions

② Limits

③ Continuity

④ Differentiability

→ Slope, tangent
→ Single variable
→ multi variable

→ minima, maxima

→ next

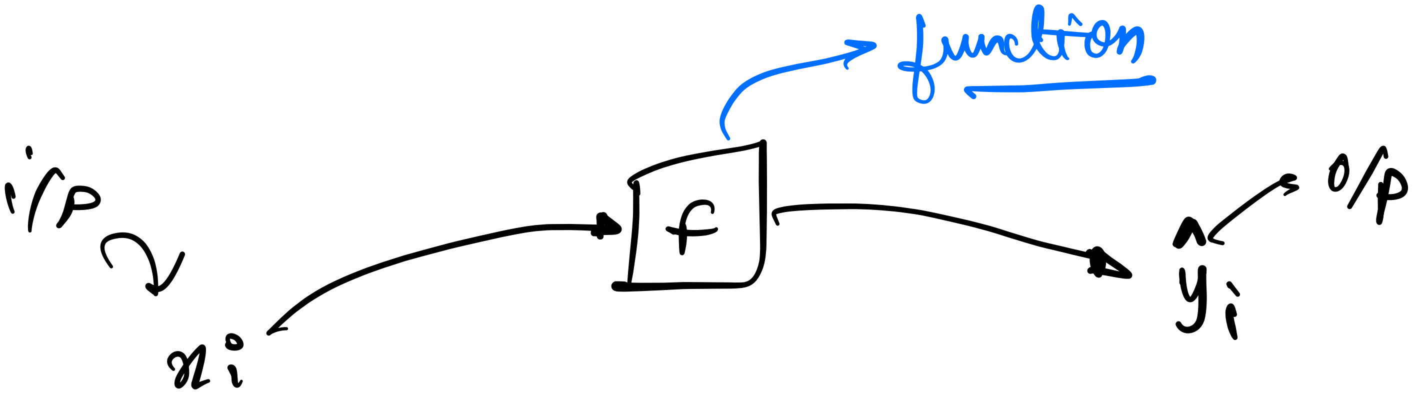
* Classifier in Mathematical Terms :-

$$D : \{ (\overset{\text{feature}}{x_i}, \underset{\text{label}}{y_i})$$

$$y \in \{+1, -1\}$$

Goal : Minimize the error

(i) function



$$\boxed{f(x_i) = \hat{y}_i}$$

predicted value

i/p

$$\boxed{\begin{array}{l} f \rightarrow x^2 \\ f(2) = 4 \end{array}}$$

ideally: $$\boxed{\hat{y}_i \approx y_i}$$

actual - label

predicted-label

eg

YOE	Salary
1	20L
3	60L
2	40L

$$\underline{\text{Salary}} = f(\text{YOE}) = \boxed{\underline{20L * \text{YOE}}}$$

$\text{YOE} = 4 \longrightarrow$ Predicted value

$$\hat{y}_i = \underline{80L}$$

$$y_i = \underline{80.1L}$$

$$G(D, \vec{w}, w_0)$$

gain function

This is the function that we need to optimize.

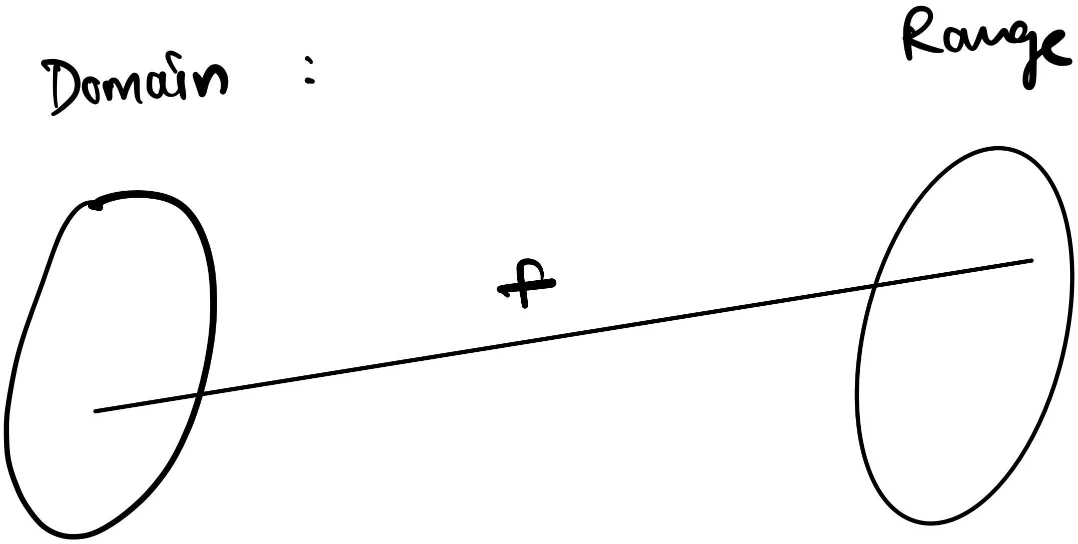
max

* Domain & Range

$f(x) = y$

↙ i/p

↳ o/p



Domain : Collection of all possible inputs that the function can take.

Range : Collection of all possible outputs.

eg :

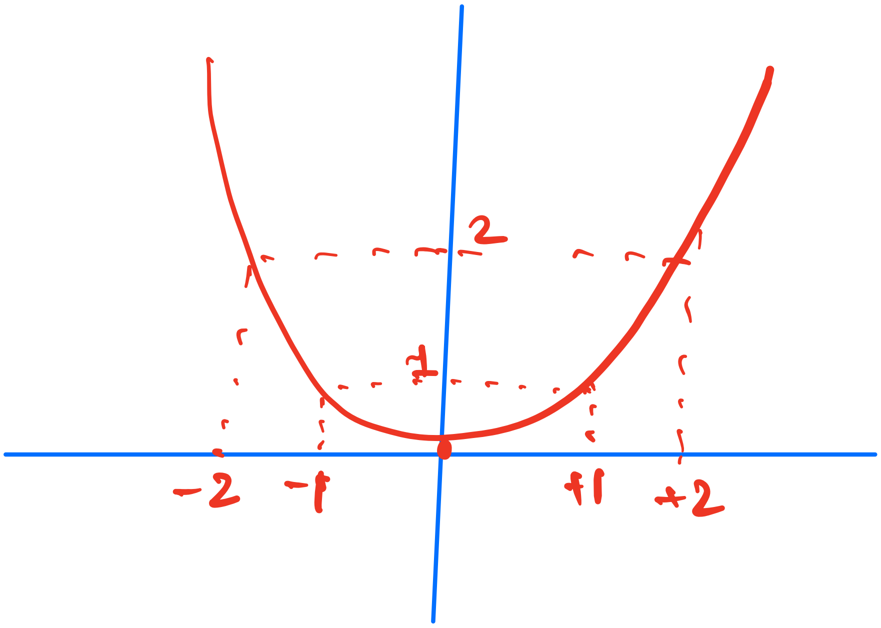
$$y = f(x) = x^2$$

$$x = 2 \longrightarrow y = x^2 = (2)^2 = 4$$

$$x = -4 \longrightarrow y = x^2 = (-4)^2 = 16$$

Domain : $-\infty$ to ∞ (i/p)

Range : 0 to ∞ (o/p)



* For every x, there must be exactly one y value

$$G(\vec{x}_i, y_i, \vec{w}, w_0)$$

$$= \left(\frac{w^T x + w_0}{\| \vec{w} \|} \right) * y_i$$

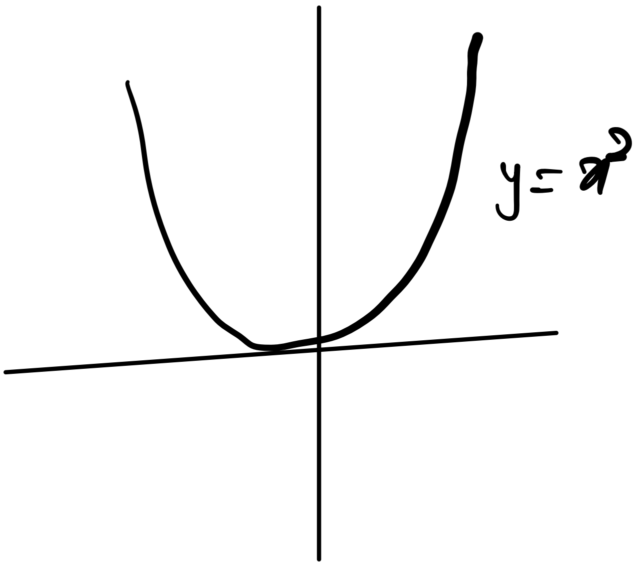
$$\Rightarrow \underset{\rightarrow \boxed{\vec{w}, w_0}}{\text{argmax}} \; \underbrace{G(\vec{x}, \vec{y}, \vec{w}, w_0)}$$

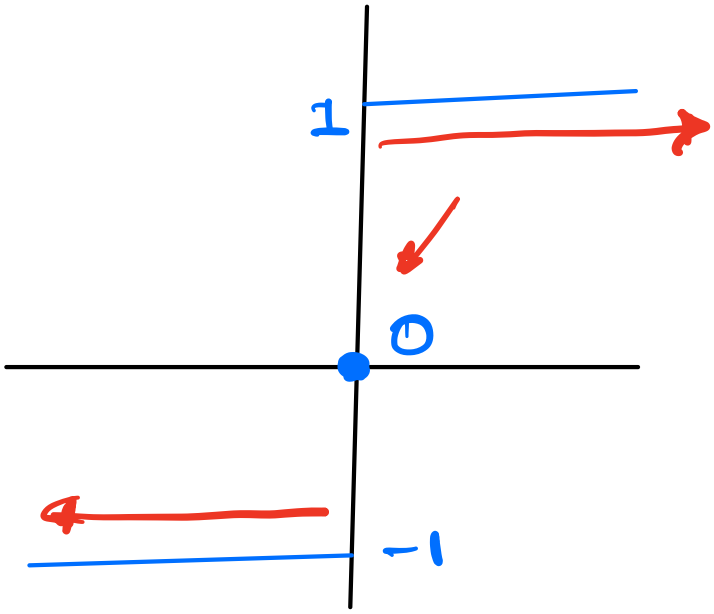
→ The best possible classifier is defined by $\vec{w}, w_0$ which will give me the max value of the gain function.

* Limits & Continuity :-

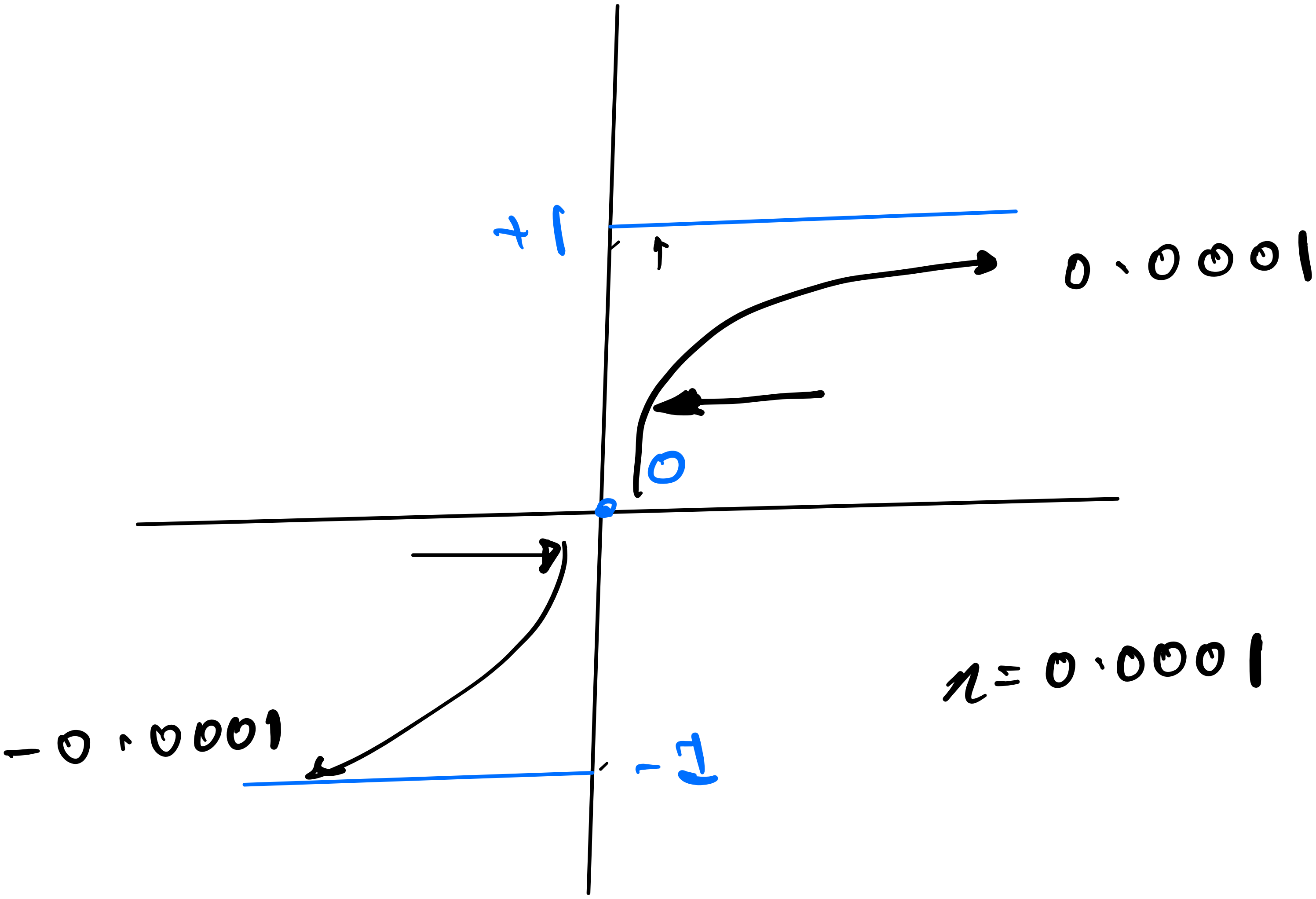
Continuous function

Dis-continuous function





✓ $$y = f(x) = \begin{cases} +1, & x > 0 \\ -1, & x < 0 \\ 0, & x = 0 \end{cases}$$



① Right hand Limit (RHL)

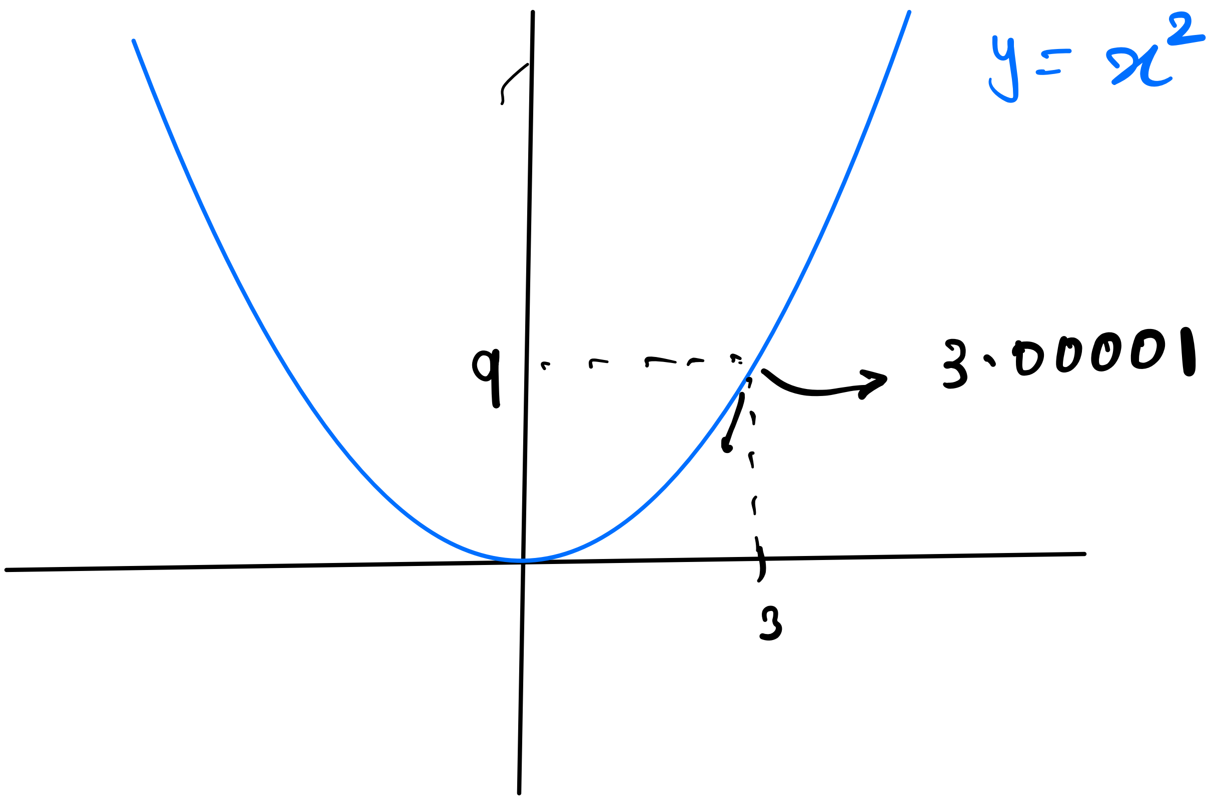
$$\lim_{x \to 0^+} f(x) = 1$$

② Left Hand Limit (LHL)

$$\lim_{x \to 0^-} f(x) = -1$$

$$\boxed{f(x) \text{ is Discontinuous at } x = 0}$$

* Continuous function



RHL :-

$$\lim_{x \to 3^+} f(x) = (3.00001)^2$$

$$= 9.0000000001$$

$$\approx 9$$

LHL

2.999999

$$\underset{x \to 3^-}{} f(x) = (2.99999)^2$$

$$\approx 9$$

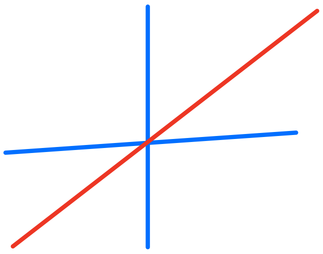
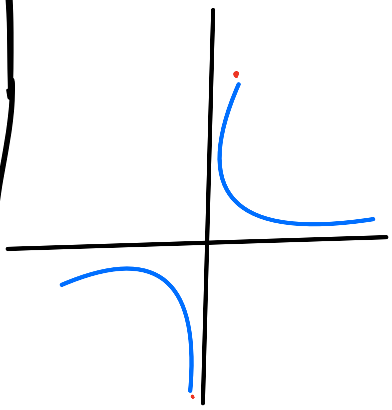
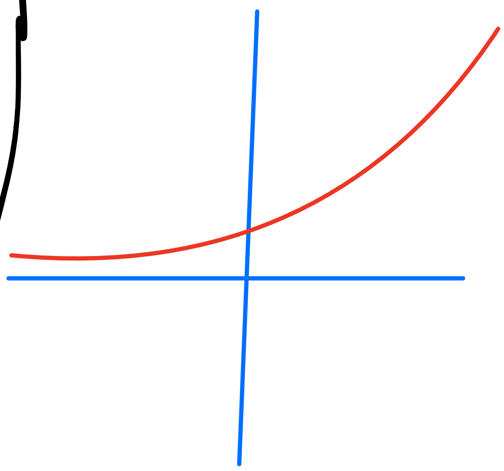
- mathematically,

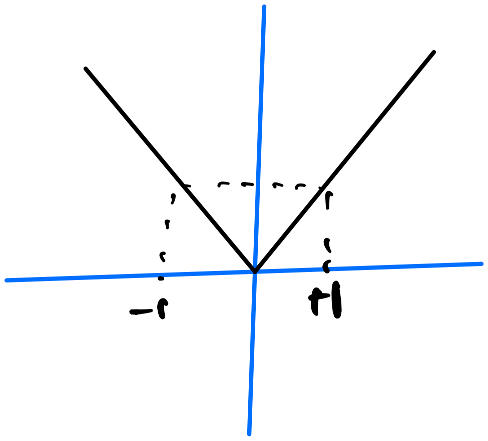
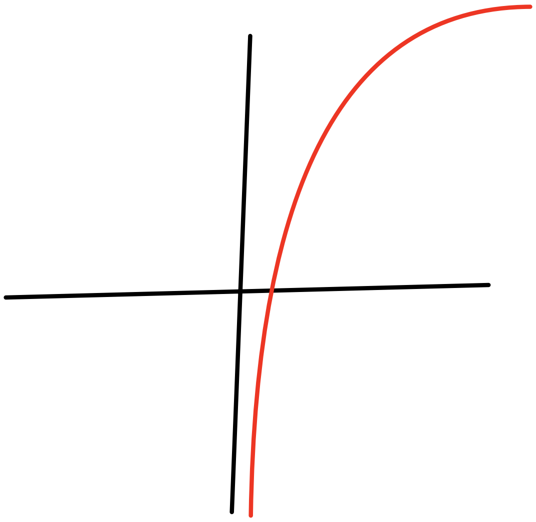
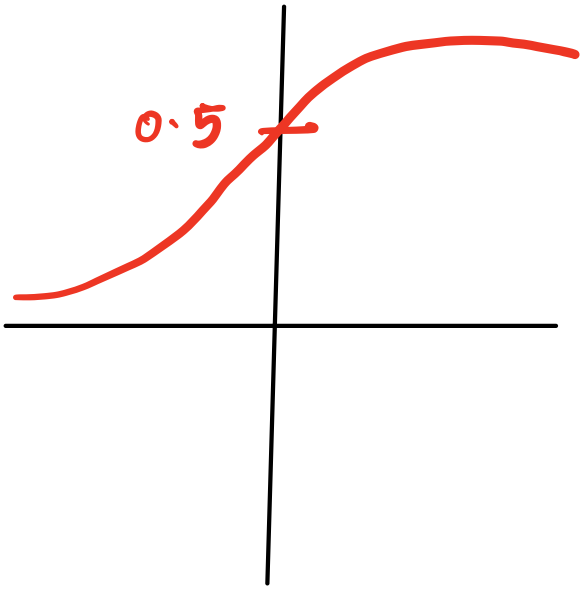
$$\boxed{\underline{\text{LHL} = \text{RHL} = f(x) \text{ at } x = a}} \checkmark$$

$$\Downarrow$$

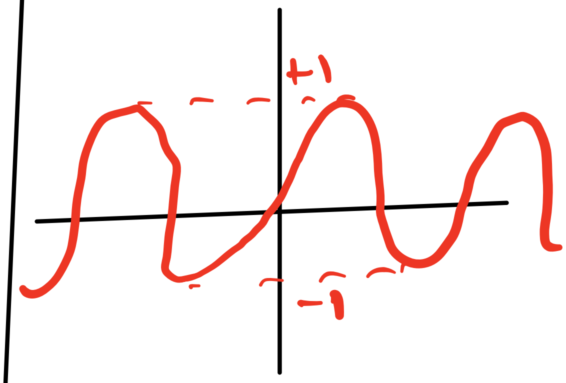
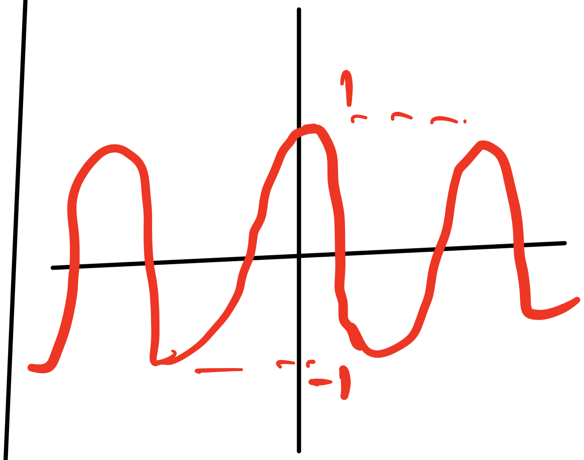
$$\boxed{\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x) = f(a)}$$

* Some important functions :

func	Domain	Range	Continuous	Graph ✓
① $y = x$	$(-\infty, \infty)$	$(-\infty, \infty)$	Conti	
② ✓$y = \frac{1}{x}$	$(-\infty, \infty)$	$(-\infty, \infty)$	Dis-Conti	✓ 
③ $y = e^{x}$	$(-\infty, \infty)$	$(0, \infty)$	Conti	

④ $y = \lvert x \rvert$	$(-\infty,\infty)$	$(0,\infty)$	Conti	
⑤ $y = \ln(x)$	$(0,\infty)$	$(-\infty,\infty)$	Conti	
⑥ $y = \dfrac{1}{1+e^{-x}}$	$(-\infty,\infty)$	$(0,1)$	Conti	

Sigmoid

$y = \sin\theta$	$(-\infty, \infty)$	$[-1, 1]$	Conti	
$y = \cos\theta$	$(-\infty, \infty)$	$[-1, 1]$	Conti	
$y = \tan\theta$	$(-\infty, \infty)$	$(-\infty, \infty)$	Disconti