# ☆ Calculus

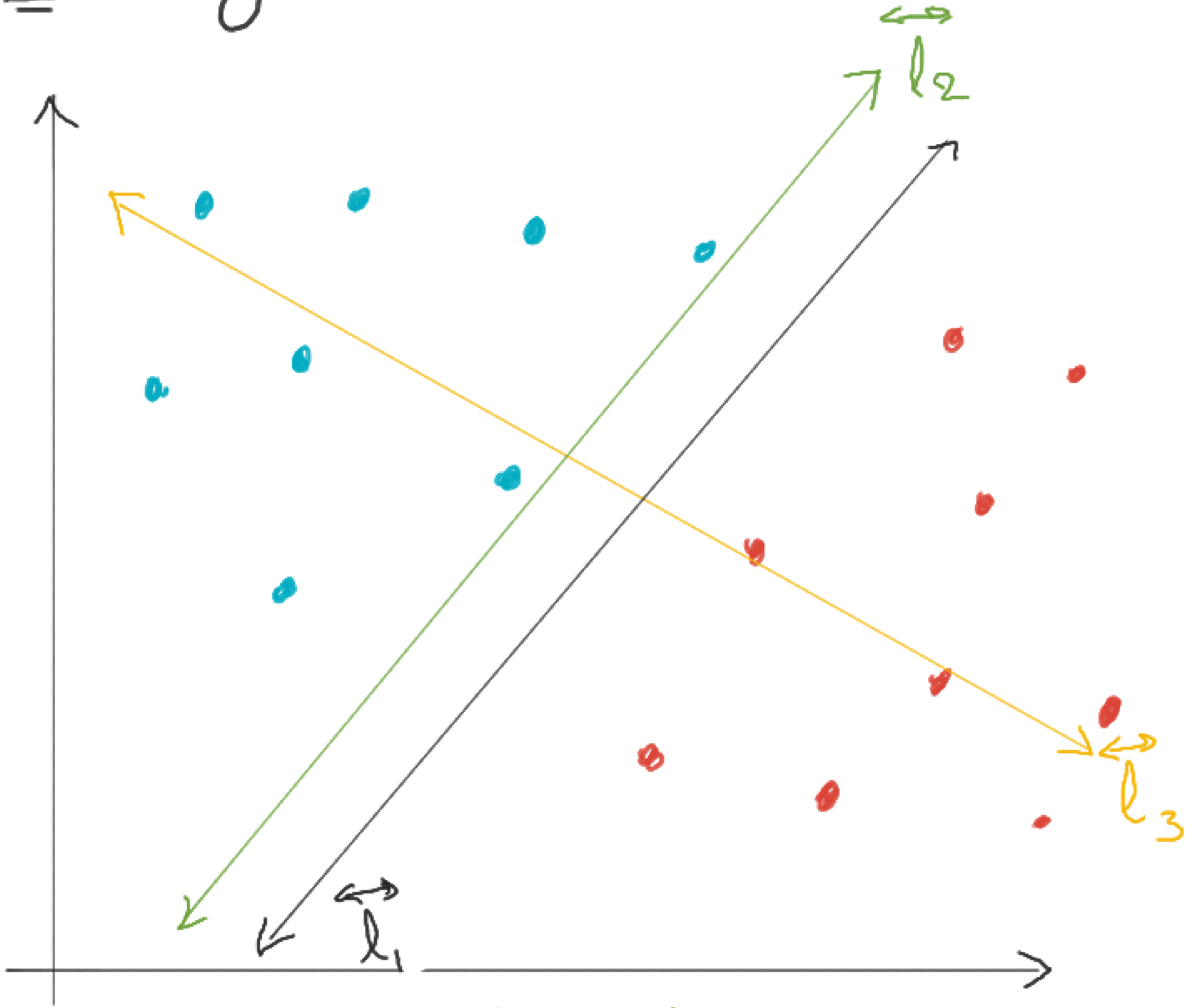
→ Why do we need calculus

→ Functions

→ Limits & Continuity

→ some important functions

## ☆ Why do we need calculus?




Accuracy: green line = 100%

black line = 100%

Why black line is better?

① Distance
② Future

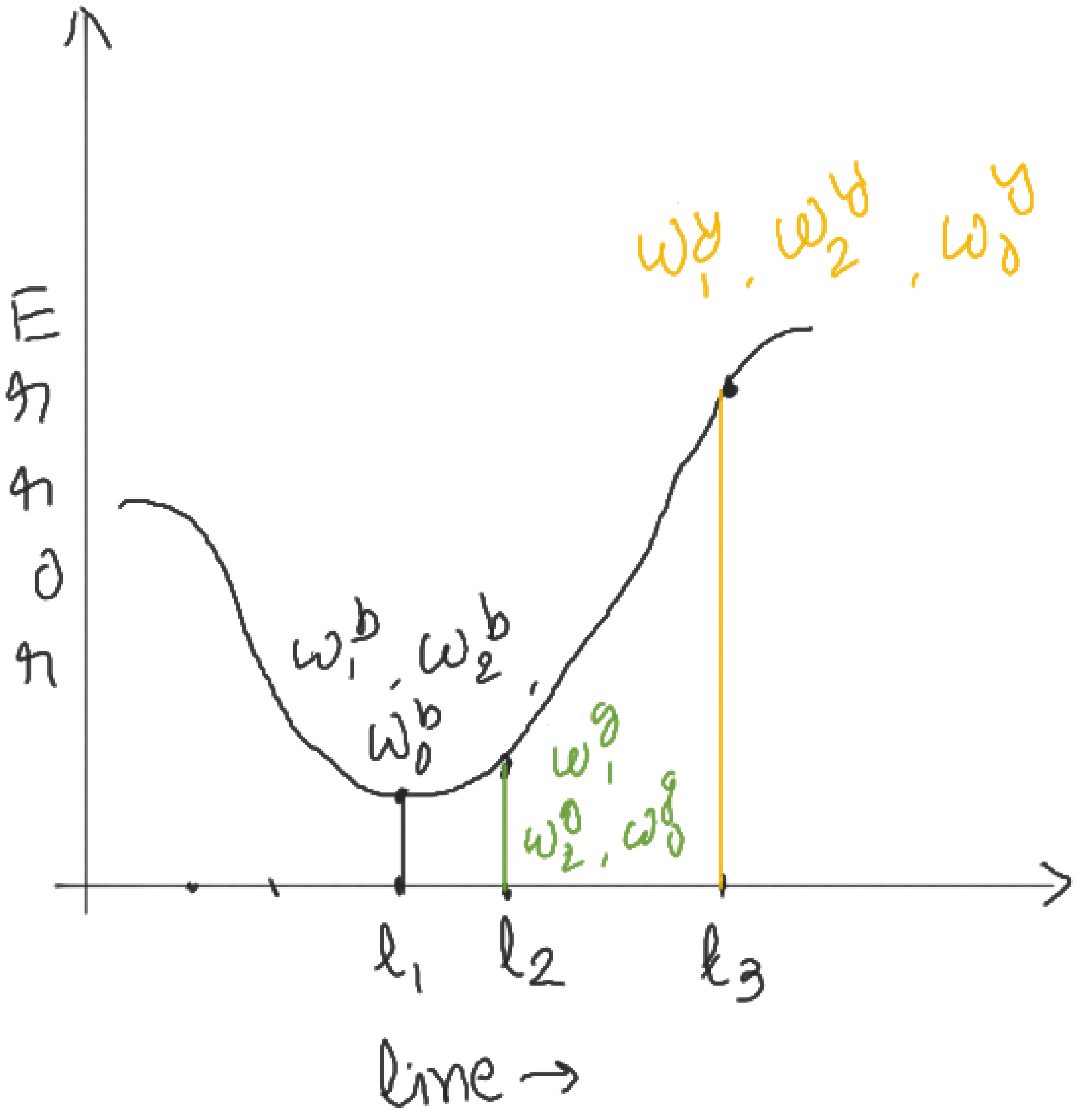
→ Gain Function

Loss Function → Error

Yellow → High

Green → Less

Black → Very less




We want to find the best-fit line.

∴ We are interested in finding $w_1^b, w_2^b$ & $w_0^b$. For simplicity, let's assume $w_1^b, w_2^b$ & $w_0^b \in [0, 20]$ and with step size of 0.1

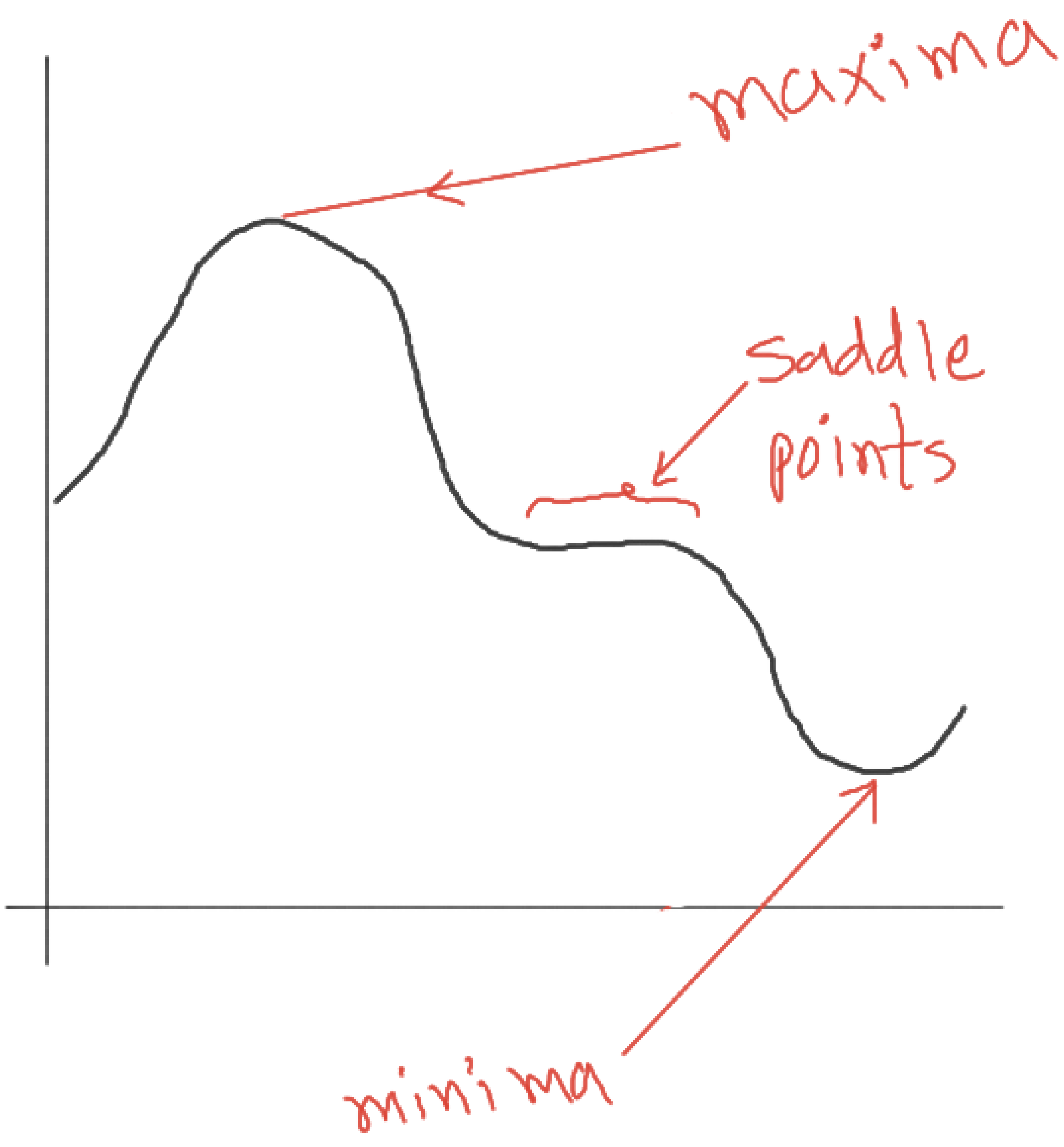
How many possible values $w_1, w_2$ & $w_0$ can take?

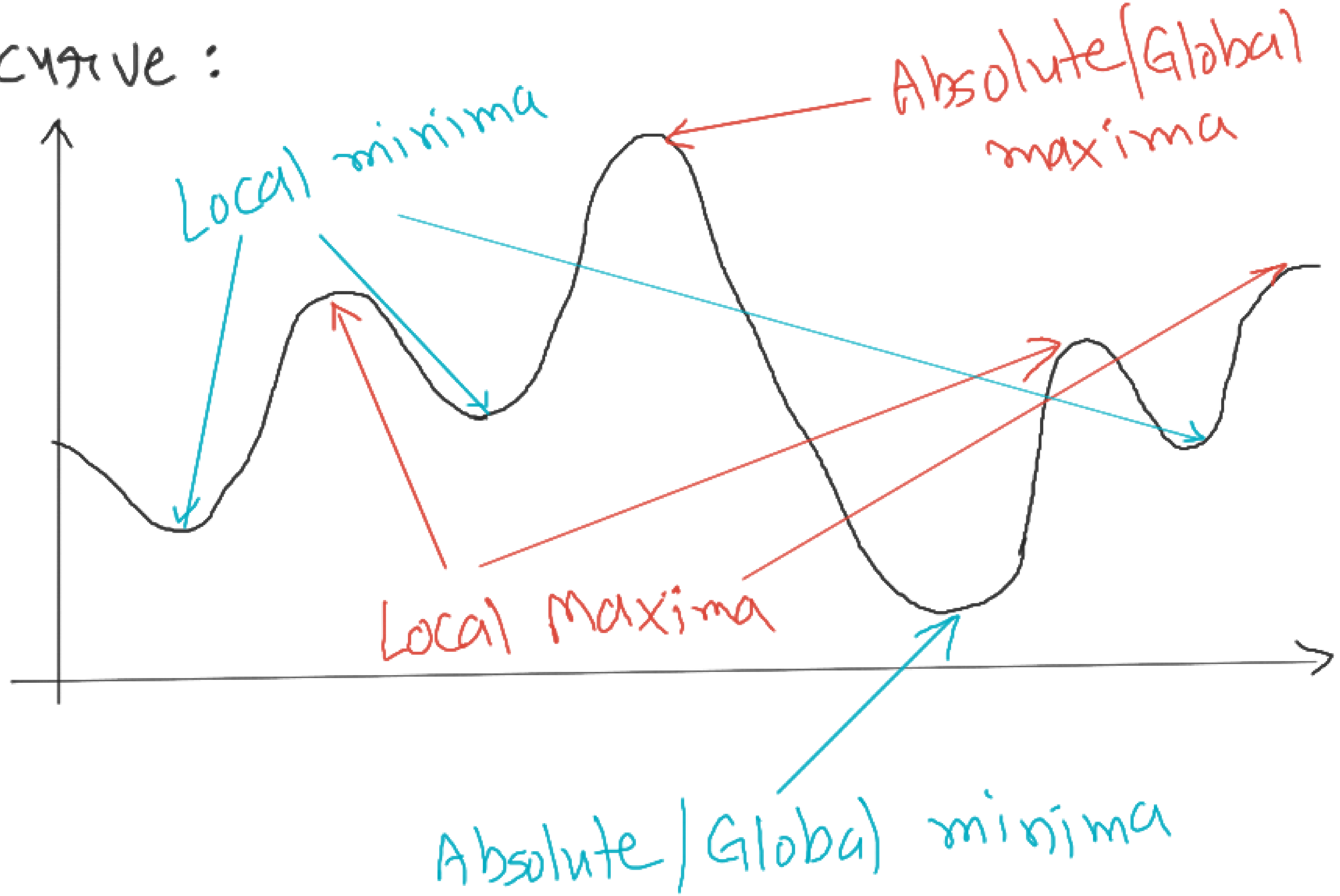
| Variable | $w_1$ | $w_2$ | $w_0$ |
|---|---|---|---|
| Possible Values | 201 | 201 | 201 |

∴ Total possible lines = $(201)^3$

∴ We need to find the best line from these $201^3$ lines. Try BruteForce

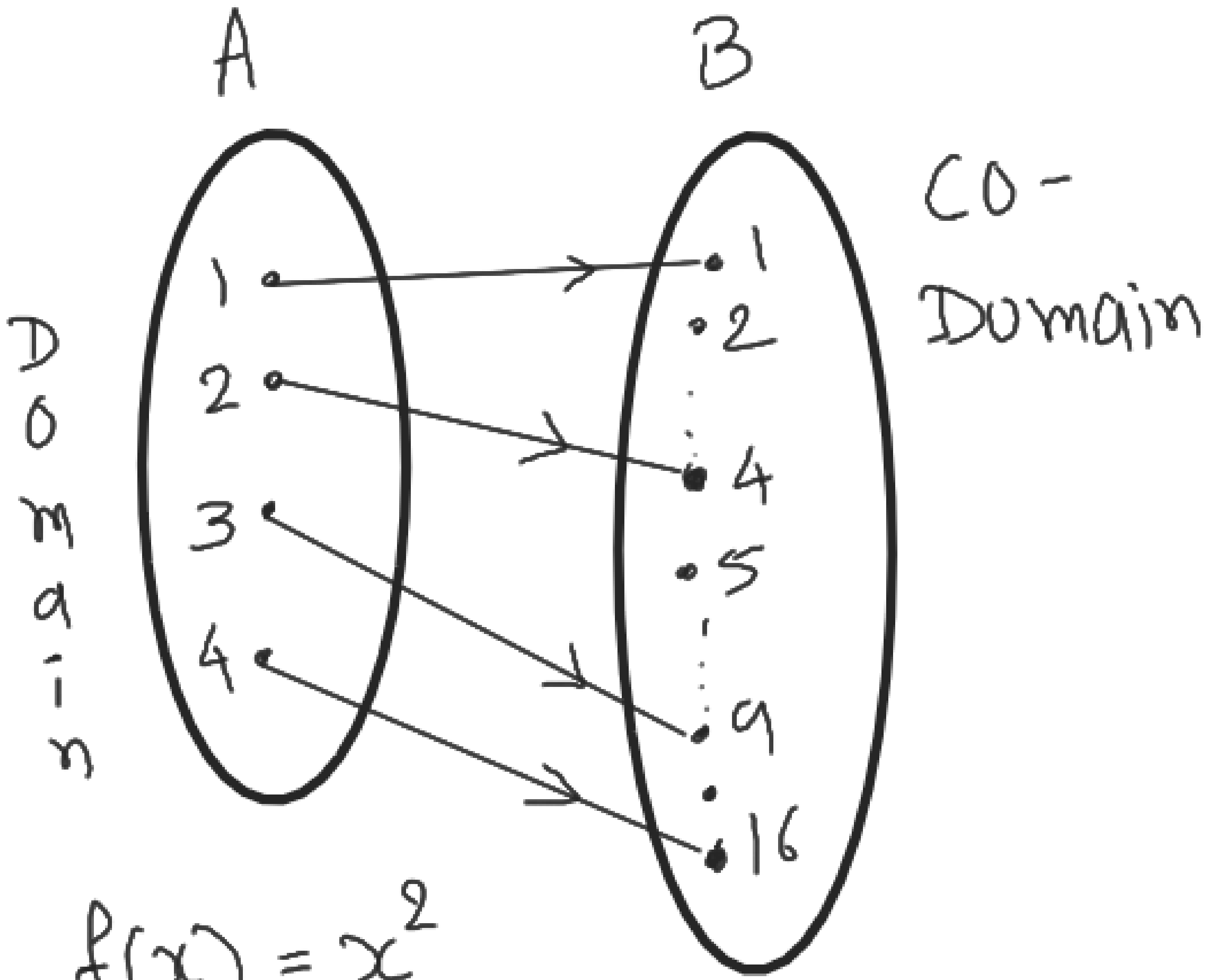
★ Components of a curve:



# ☆ Functions

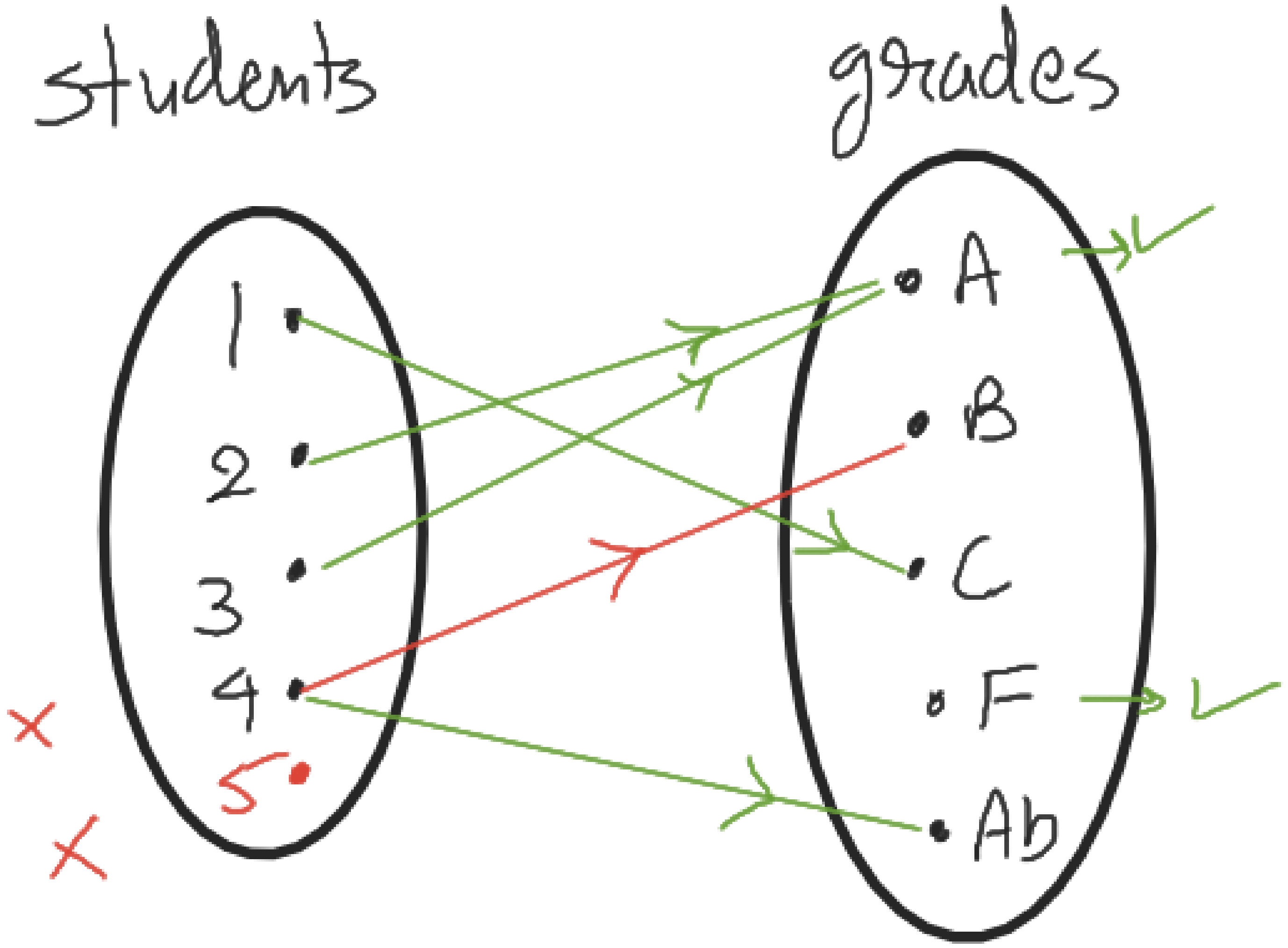
$$\vec{x} \longrightarrow \boxed{f(x)} \longrightarrow y$$



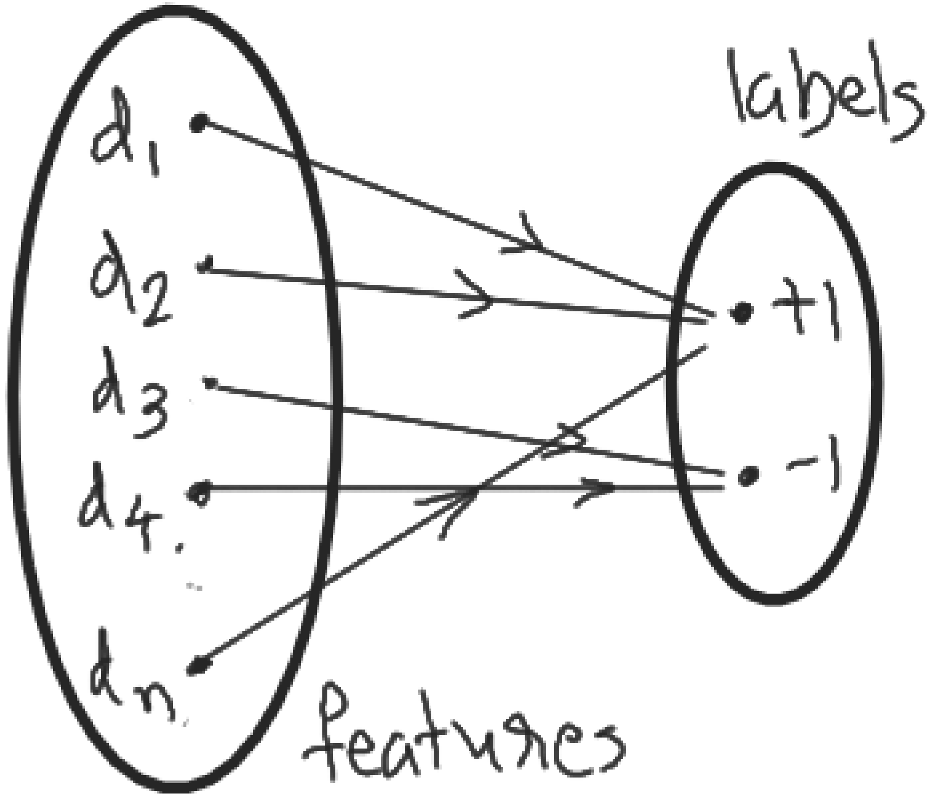

$f(x) = x^2$

$f : A \to B$

Range $= \{1, 4, 9, 16\}$




→ ML context:




$f(x_1, x_2) = \omega_1 x_1 + \omega_2 x_2 + \omega_0$

## ☆ Misc - Gain & Loss Functions

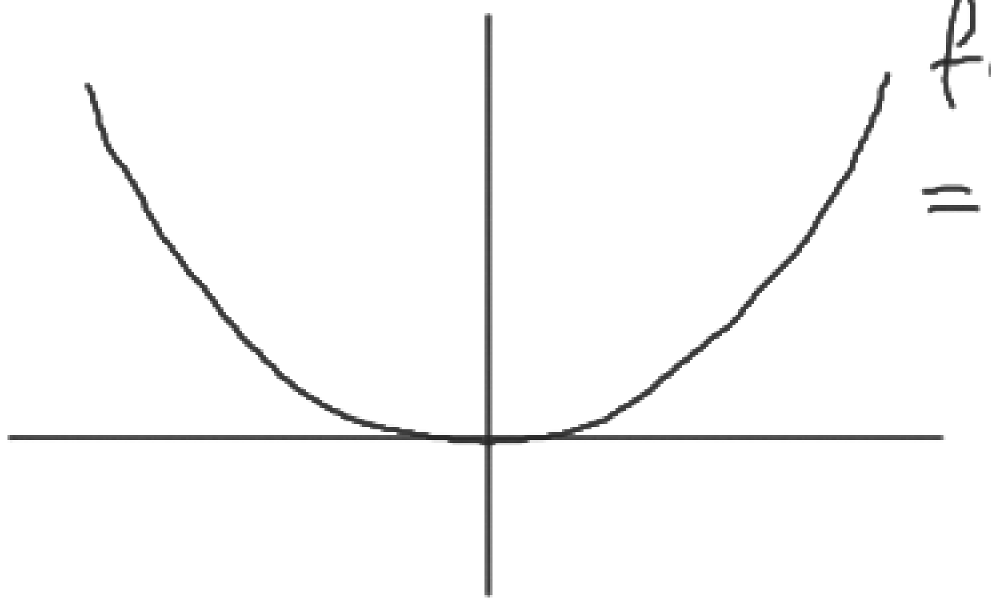
$$\underset{\vec{w},\, w_0}{\operatorname{argmax}} \ \sigma(\vec{w}, w_0, \vec{x}, \vec{y}) = \sum_{i=0}^{n} \frac{\vec{w}^T \cdot \vec{x}_i + w_0}{\|\vec{w}\|} \cdot y_i \quad \checkmark$$

$$= \prod_{i=0}^{n} \frac{\vec{w}^T \cdot \vec{x}_i + w_0}{\|\vec{w}\|} \cdot y_i \quad \times$$

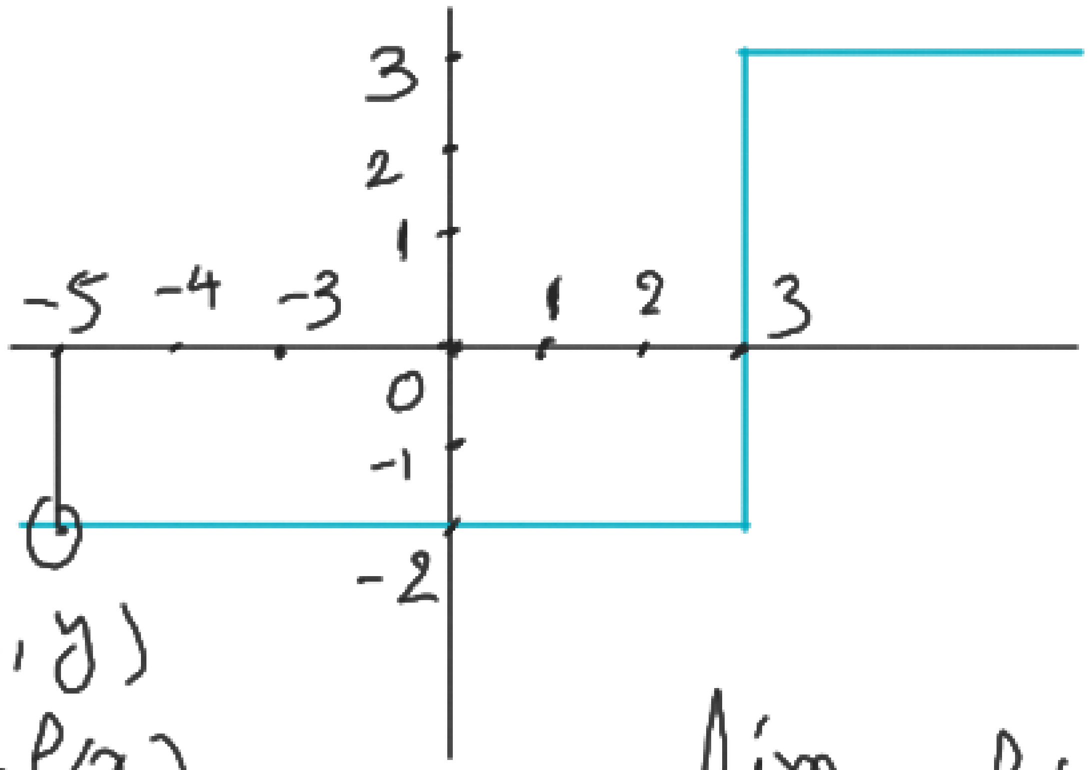
$$\sum_{i=1}^{4} a_i = a_1 + a_2 + \overset{0}{a_3} + a_4$$

$$\prod_{i=1}^{4} a_i = a_1 * a_2 * \overset{0}{a_3} * a_4$$

# ☆ Limits & Continuity



$f(x) = x^2$




$(x, y)$

$y = f(x)$

$f(x) = -2 \; ; \; x < 3$

$f(x) = 3 \; ; \; x > 3$

if $x = 3$, $f(x) = (?)$

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

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

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

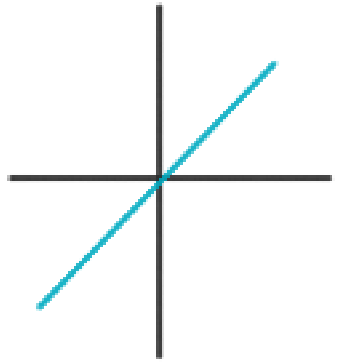
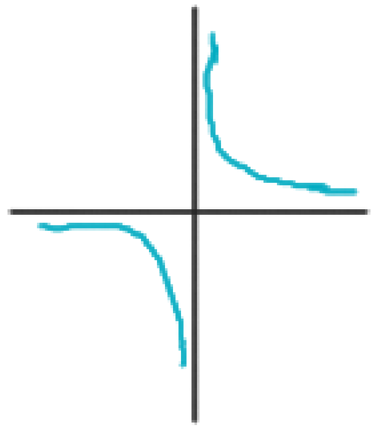
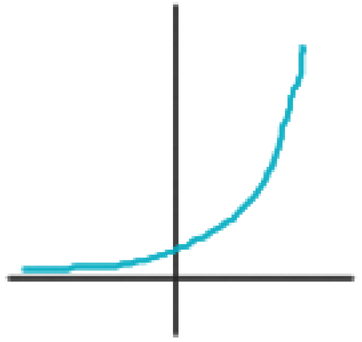
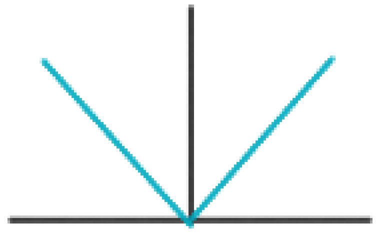
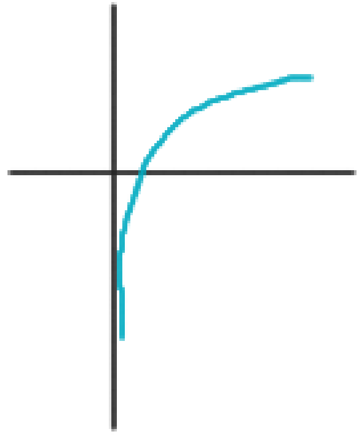
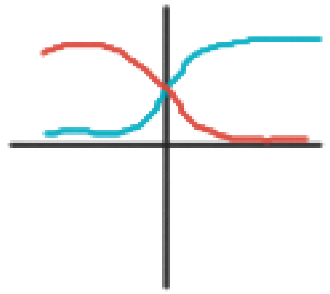
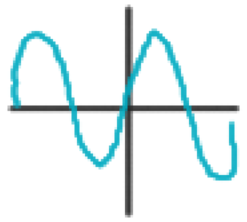
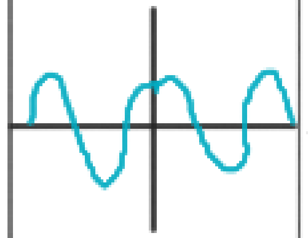
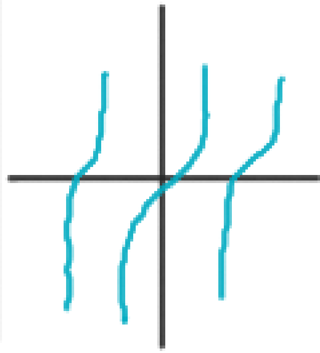
$\Rightarrow f(x)$ is continuous at $x = 0$

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

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

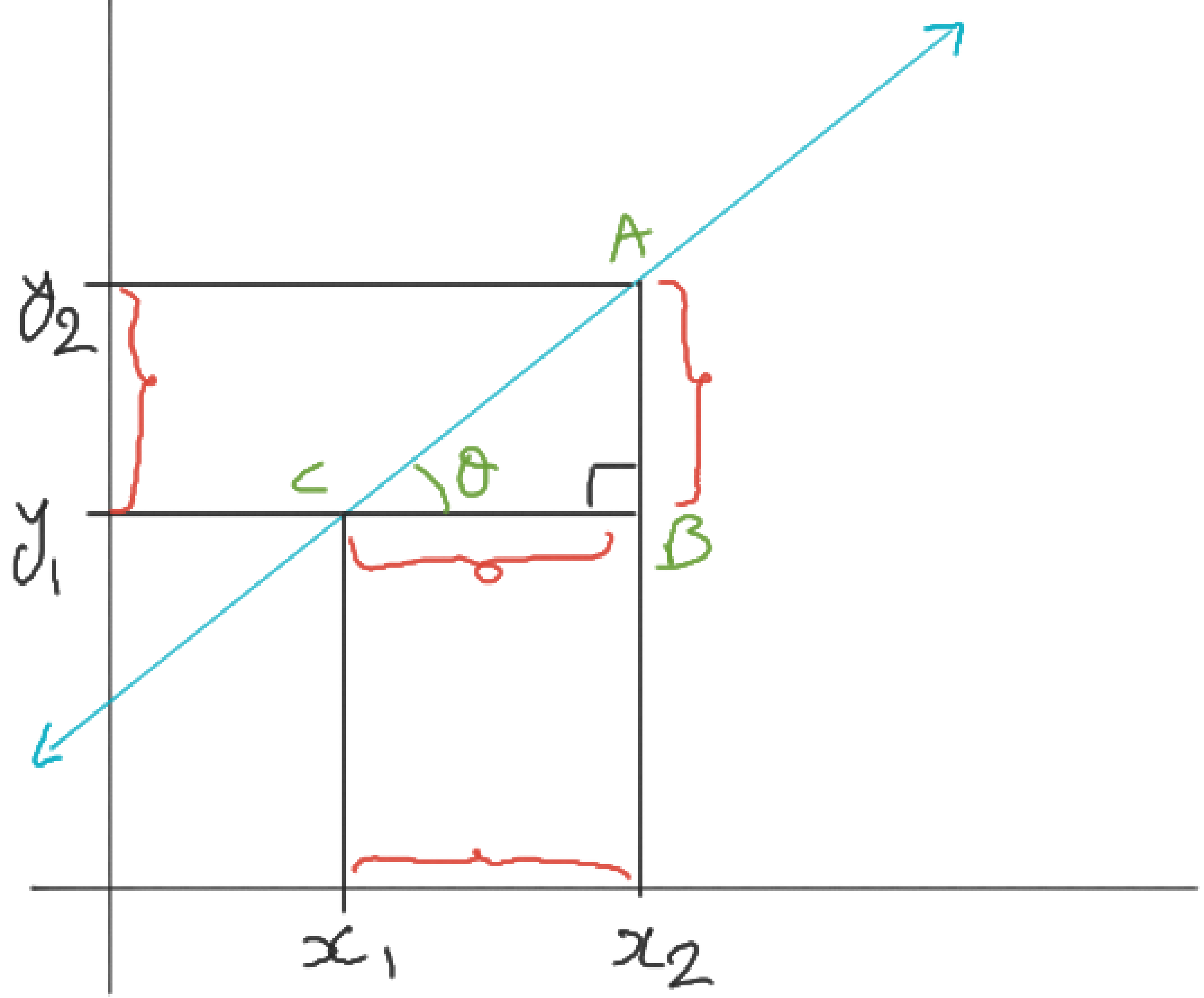
$\therefore f(x)$ is a discontinuous function

# ☆ Some important / most used functions:

| Sr. No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| function | $y = x$ | $y = \frac{1}{x}$ | $y = e^{x}$ | $y = \lvert x \rvert$ | $y = \ln(x)$ | $y = \frac{1}{1 + e^{-x}}$ | $y = \sin(x)$ | $y = \cos(x)$ | $y = \tan(x)$ |
| Domain | $[-\infty, +\infty]$ | $R - \{0\}$ | $R$ OR $[-\infty, +\infty]$ | $[-\infty, +\infty]$ | $R^{+}$ | $R$ | $R$ | $R$ | $R$ |
| Range | $[-\infty, +\infty]$ | $R - \{0\}$ | $R^{+}$ | $R^{+}$ | $R$ | $(0, 1)$ | $[-1, 1]$ | $[-1, 1]$ | $R$ |
| Continuous? | Y | N | Y | Y | Y | Y | Y | Y | N |
| Plot |  |  |  |  |  |  |  |  |  |

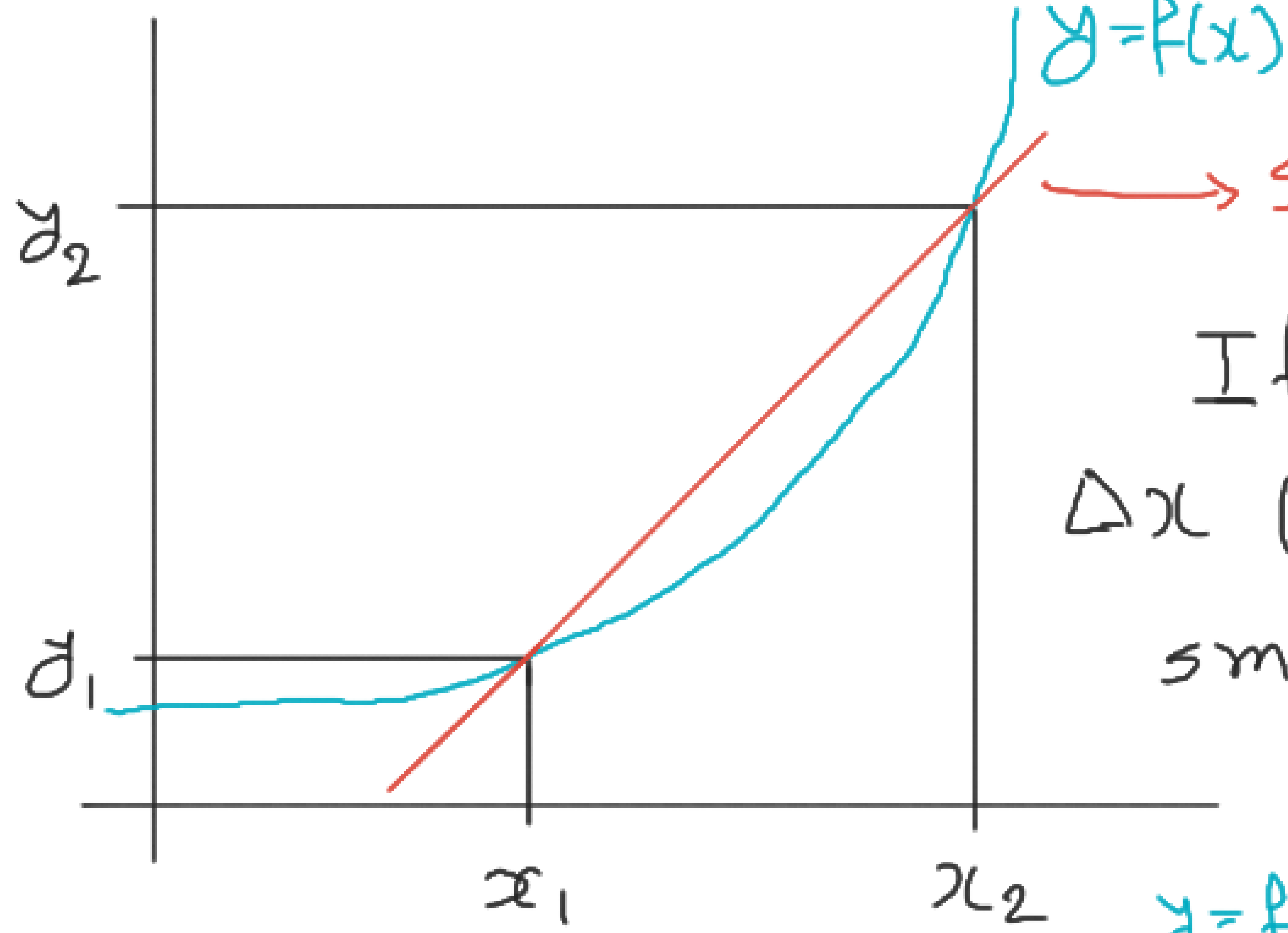
# ☆ Differentiation

slope = m = tan θ

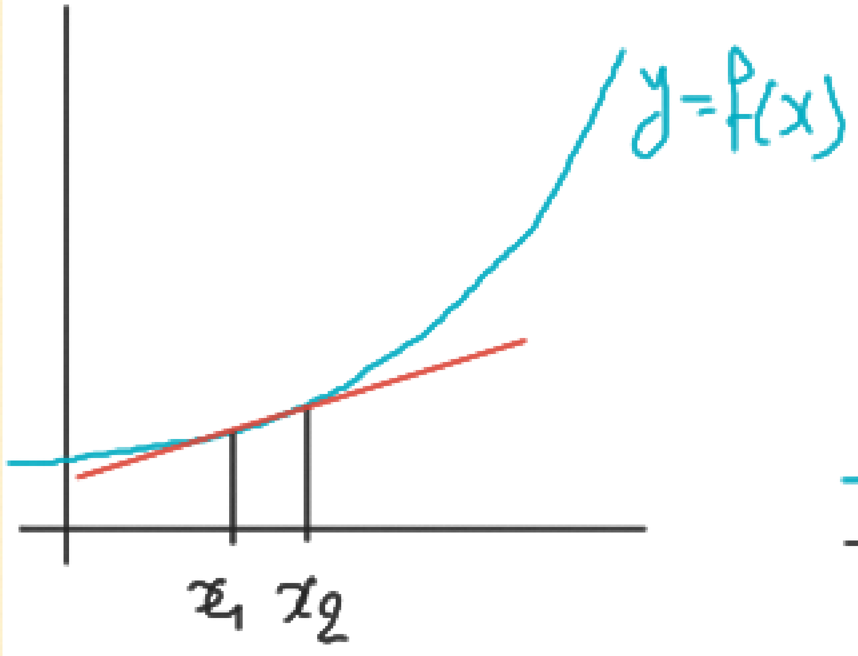


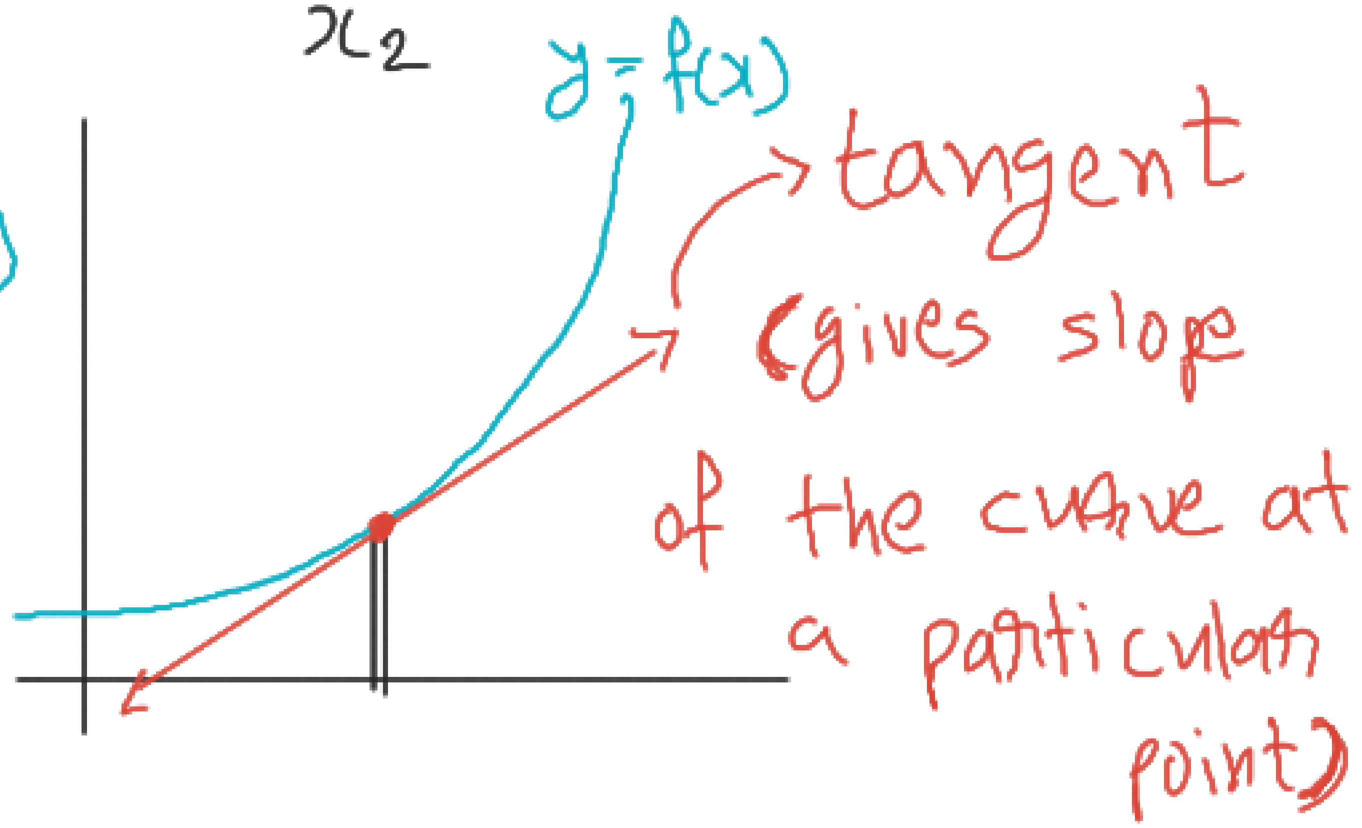

$y = mx + c$

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\text{Rise}}{\text{Run}} = \frac{\Delta y}{\Delta x}$$

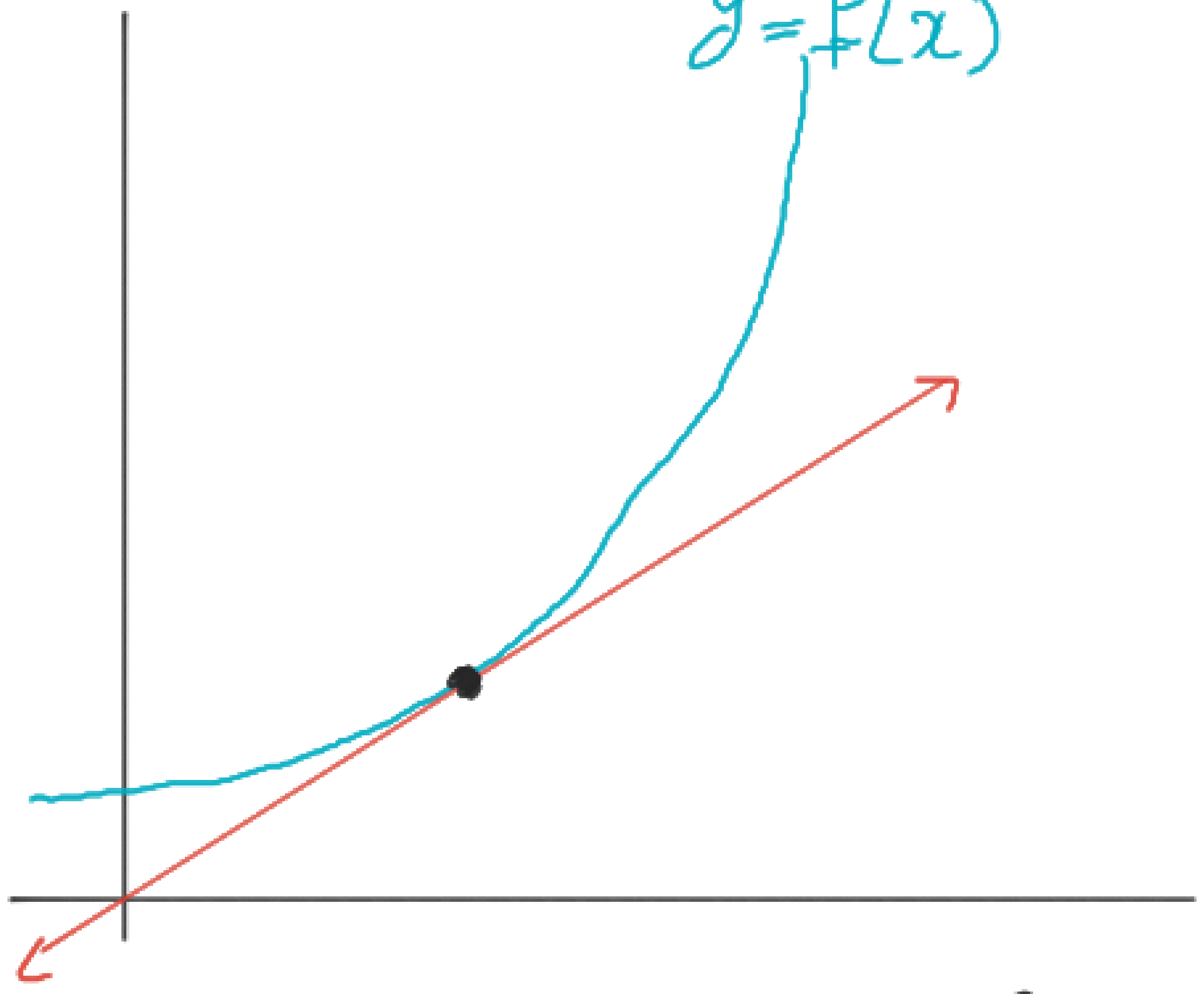



If we make $\Delta x$ $(x_2 - x_1)$ as small as possible then:




Now on, if $\Delta x$ is almost 0, we will call it $dx$ & $\Delta y$ as $dy$ $\therefore$ slope $= dy/dx$




Example: Let $f(x) = x^2$

let $x_1 = 3 \therefore y_1 = 9 = f(x_1)$

if $x_2 = 5 \Rightarrow y_2 = 25 = f(x_2)$

Let $x_2 - x_1 = h \Rightarrow x_2 = (x_1 + h)$

$\therefore f(x_2) = f(x_1 + h)$

slope of the curve m:

$$\frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$$

$$= \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

$$= \frac{f(x_1 + h) - f(x_1)}{x_1 + h - x_1}$$

$$\frac{\Delta y}{\Delta x} = \frac{f(x_1 + h) - f(x_1)}{h}$$

To find slope at a point, we need $\Delta x$ to be very small

$\therefore x_2 - x_1 \to 0 \therefore h \to 0$

$$\boxed{\therefore \frac{dy}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}}$$

☆ Let's verify this formula! Let $y = x^2 \Rightarrow \boxed{\dfrac{dy}{dx} = 2x}$

| $x$ | $y$ | $dx$ | $x+h$ | $f(x+h)$ | $dy$ | $dy/dx$ | |
|---|---|---|---|---|---|---|---|
| 2 | 4 | 0.00001 | 2.00001 | 4.00004 | 0.00004 | 4 | $= 2x$ ? |
| 3 | 9 | 0.00001 | 3.00001 | 9.00006 | 0.00006 | 6 | $= 2x$ ? |
| 4 | 16 | 0.00001 | 4.00001 | 16.00008 | 0.00008 | 8 | $= 2x$ ? |

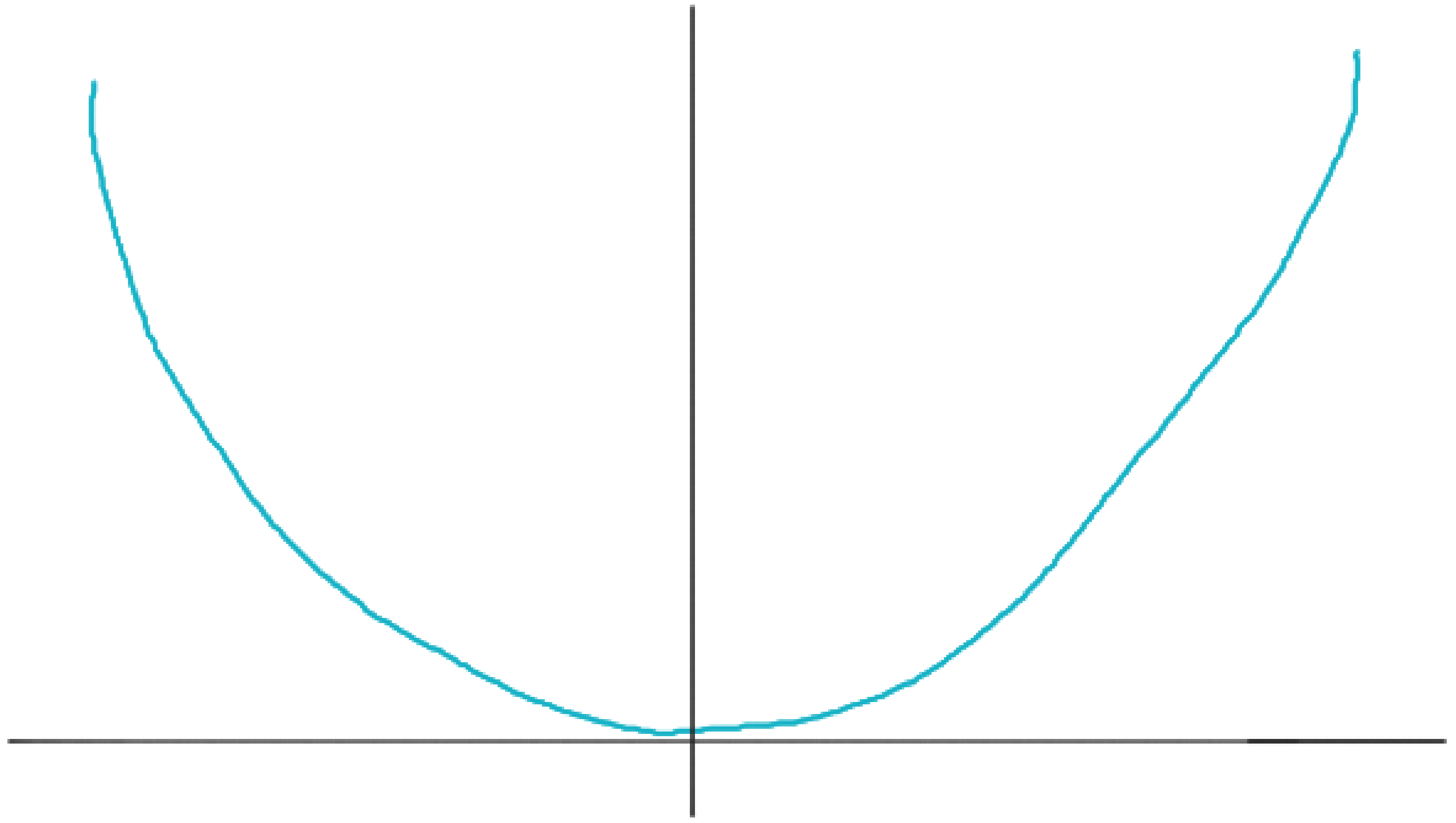
$$\frac{dy}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

☆ A function is continuous at a point 'a' if & only if :

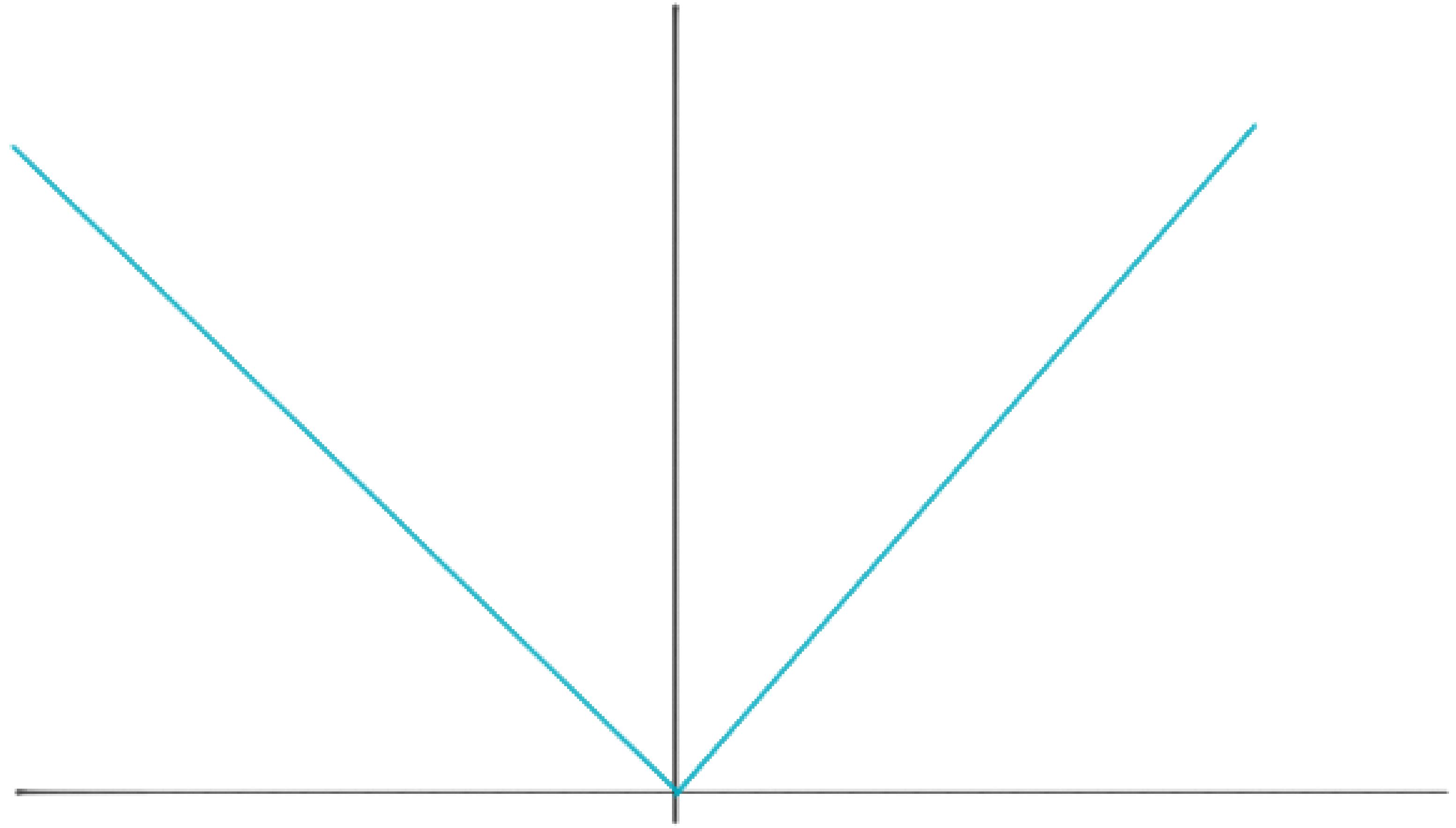
$$\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x) = \lim_{x \to a} f(x)$$

Practically, ① If we can draw the function curve without lifting our pen & ② If the curve doesn't have sharp corners

① $f(x) = x^2$ is differentiable



② $f(x) = |x|$ is not differentiable



**☆ Some Imp. Derivatives:**

① $\frac{d}{dx} x^n = n x^{n-1}$

② $\frac{d}{dx} c = 0$ ; $c$ = constant

③ $\frac{d}{dx} \sin x = \cos x$

④ $\frac{d}{dx} \cos x = -\sin x$

⑤ $\frac{d}{dx} \log x = \frac{1}{x}$

⑥ $\frac{d}{dx} e^x = e^x$

## ☆ Imp Rules Of Differentiation:

① Sum Rule: $$\frac{d}{dx}(f(x)+g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$$

② Product Rule: $$\frac{d}{dx}(f(x).g(x)) = g(x)\frac{d}{dx}f(x) + f(x)\frac{d}{dx}g(x)$$

③ Division Rule (Quatient Rule): $$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{g(x)\frac{d}{dx}f(x) - f(x)\frac{d}{dx}g(x)}{(g(x))^2}$$

④ Chain Rule: $$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x)$$

$\boxed{\text{Try these:}}$

① $\frac{d}{dx} x^5 = 5x^4$

② $\frac{d}{dx}(3x^4) = 12x^3$

③ $\frac{d}{dx}(2x^3 + 3x^2)$

$= \frac{d}{dx} 2x^3 + \frac{d}{dx} 3x^2$

$= 6x^2 + 6x$

④ $\frac{d}{dx}(5x^2 + 18)$

$= \frac{d}{dx} 5x^2 + \frac{d}{dx} 18$

$= 10x + 0$

⑤ $\frac{d}{dx}(x \cdot \log x)$

$= \log x \cdot \frac{d}{dx} x + x \cdot \frac{d}{dx} \log x$

$= \log x + x \cdot \frac{1}{x}$

$= \log x + 1$

⑥ $\frac{d}{dx} \frac{\log x}{x}$

$= \dfrac{x \frac{d}{dx} \log x - \log x \frac{d}{dx} x}{x^2}$

$= \dfrac{\cancel{x} \cdot \frac{1}{\cancel{x}} - \log x \cdot 1}{x^2}$

$= \dfrac{1 - \log x}{x^2}$

$\boxed{\text{Try these!}}$

⑦ $\frac{d}{dx}\sin(3x^4)$

$= \cos(3x^4)\cdot\frac{d}{dx}3x^4$

$= \cos(3x^4)\cdot 12x^3$

$= 12x^3\cos(3x^4)$

⑧ $\frac{d}{dx}\cos(5x^2+2x)$

$= -\sin(5x^2+2x)\,\frac{d}{dx}(5x^2+2x)$

$= -\sin(5x^2+2x)\cdot(10x+2)$

⑨ $\frac{d}{dx}e^{5x^3-3x^2}$

$= e^{5x^3-3x^2}\cdot\frac{d}{dx}(5x^3-3x^2)$

$= e^{5x^3-3x^2}\cdot(15x^2-6x)$

⑩ $\frac{d}{dx}e^{\log(x)\cdot\sin(x)}$

$= e^{\log x\cdot\sin x}\cdot\frac{d}{dx}(\log x\cdot\sin x)$

$= e^{\log x\cdot\sin x}(\sin x\cdot\frac{1}{x}+\log x\cos x)$

☆ $y = \dfrac{1}{1+e^{-x}}$

$$\frac{dy}{dx} = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{1}{(1+e^{-x})} \cdot \frac{e^{-x}}{(1+e^{-x})} = \frac{1}{(1+e^{-x})} \cdot \left(1 - \frac{1}{1+e^{-x}}\right)$$

$$\boxed{\frac{d}{dx} f(x) = f(x)\,(1 - f(x)) \quad \text{for } f(x) = \frac{1}{1+e^{-x}}}$$

## ☆ Towards gradient descent - What is argmax/argmin?

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

(i) Maximum possible value of $f(x) \rightarrow \max(f(x))$

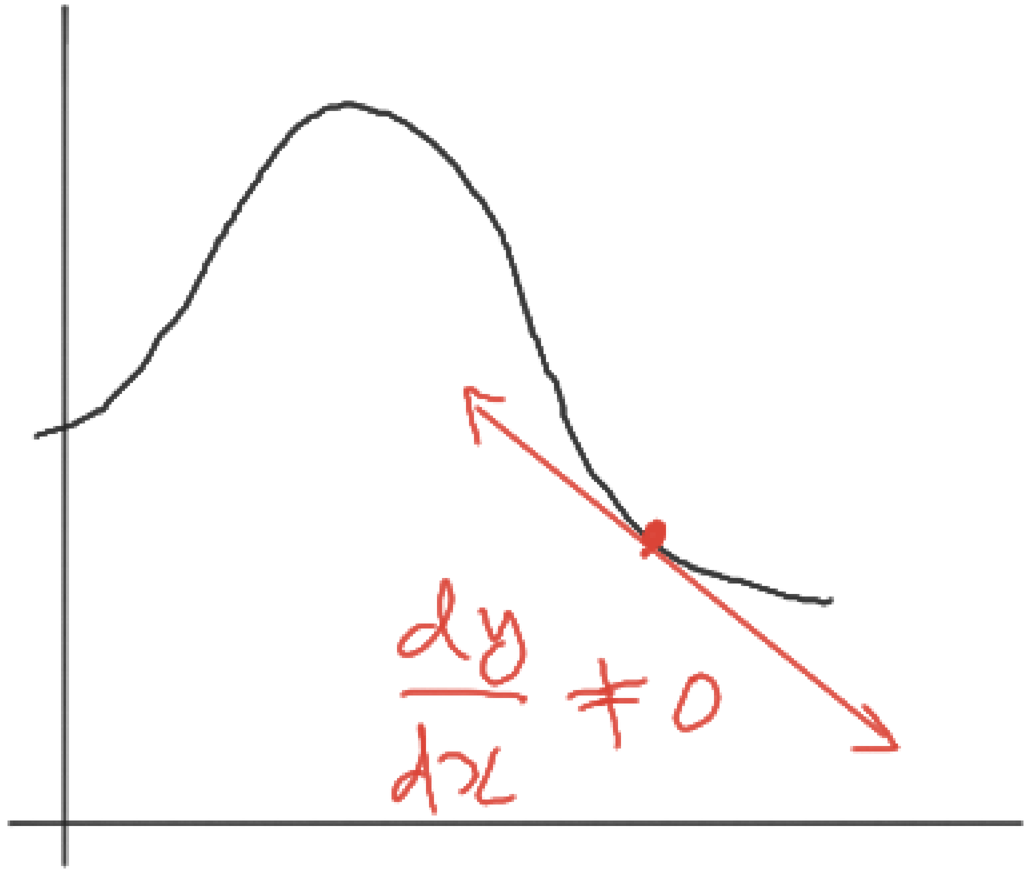
(ii) Value of $x$ when $f(x)$ is maximum. $\rightarrow \underset{x}{\operatorname{argmax}} f(x)$

$$\underset{\vec{w},\, w_0}{\operatorname{argmax}} \quad \sigma(\vec{w}, w_0, x, y) = \sum_{i=1}^{n} \frac{\vec{w} \cdot \vec{x} + w_0}{\|\vec{w}\|} \cdot y_i$$ means

value of $\vec{w}$ & $w_0$ when $\sigma$ is maximum

☆ Differentiation in context of ML (gradiend descent):

Let's say graph (or curve) of our gain function is as below:




suppose profit of an organization can be given as a function of sales ($x$) as:

$$f(x) = 41 - 72x - 18x^2$$

$$\therefore f'(x) = -72 - 36x \Rightarrow f''(x) = -36$$

☆ If $f''(x) < 0$ then the function is concave downwards ( $\cap$ ) at that point.

But if $f''(x) > 0$ then the function is concave upwards ( $\cup$ ) at that point

**☆ Steps to reach to optima (either minima or maxima):**

① Compute $f'(x)$

② Equate $f'(x)$ to $0$ & find value of $x$ (at $x = c$)

③ Calculate value of $f''(x)$ at each value of $x$.

④ If $f''(x) > 0 \Rightarrow x = c$ is a minima

If $f''(x) < 0 \Rightarrow x = c$ is a maxima

If $f''(x) = 0 \Rightarrow x = c$ is neithe maxima nor minima but it is a saddle point

☆ Examle : ① $f(x) = 41 - 72x - 18x^2$ ② $f(x) = 41 - 32x - 72x^2 - 18x^3$

① step-1 : $f'(x) = -72 - 36x$ , $f'(x) = -32 - 144x - 54x^2$

② step-2 : Let $f'(x) = 0 \Rightarrow -36x = 72 \Rightarrow \boxed{x = -2}$

$54x^2 + 144x + 32 = 0 \Rightarrow x = -0.245$ or $x = -2.422$

In this type of case we may get more than one values of $x$.

③ step-3 : $f''(x) = -36$ for $x = -2$ (or, for any value of $x$)

$f''(x) = -144 - 108x$ $\therefore$ For $x = -0.245$, $f''(x) = -117.54$

& for $x = -2.422$, $f''(x) = 117.58$

④ $f''(x) < 0 \Rightarrow x = -2$ is a maxima

$f''(-0.245) < 0 \Rightarrow x = -0.245$ is a maxima while $f''(-2.422) > 0 \Rightarrow x = -2.422$ is a minima