Session-7

Optimization-2 Towards SGD

Feb 10, 2024

Agenda

1. Derivative
2. Finding derivative for any function
3. Checking if function is differentiable
4. Commonly used derivatives
5. Rules of differentiation
6. Using derivative for optimization

WHAT IS DERIVATIVE

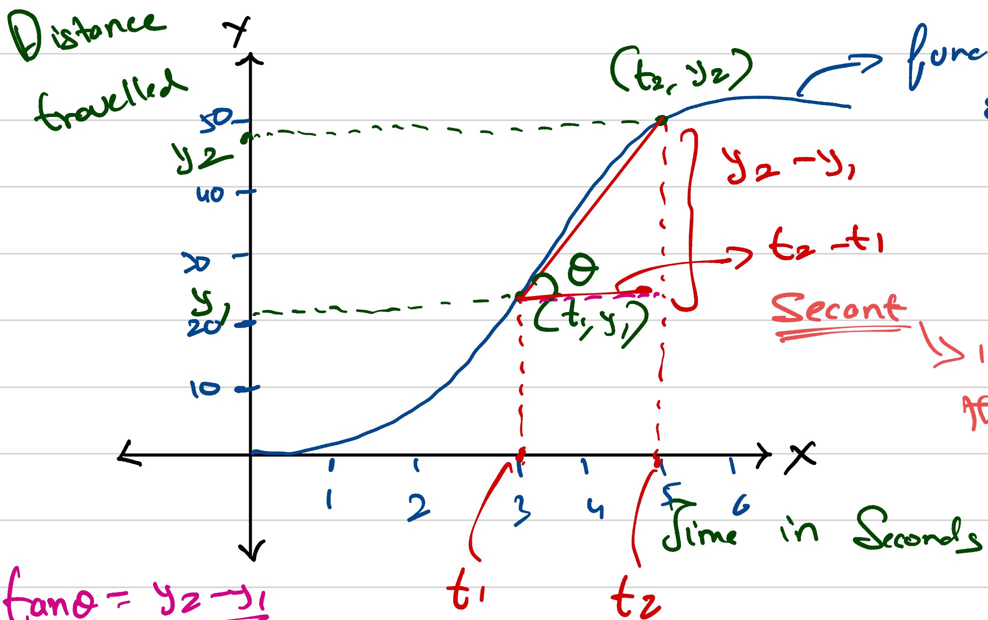
* Expression of $\frac{dy}{dx}$

$dy \rightarrow$ change in y

$dx \rightarrow$ change in x

$y = f(x)$

$\rightarrow$ Rate of change of y w.r.t. x



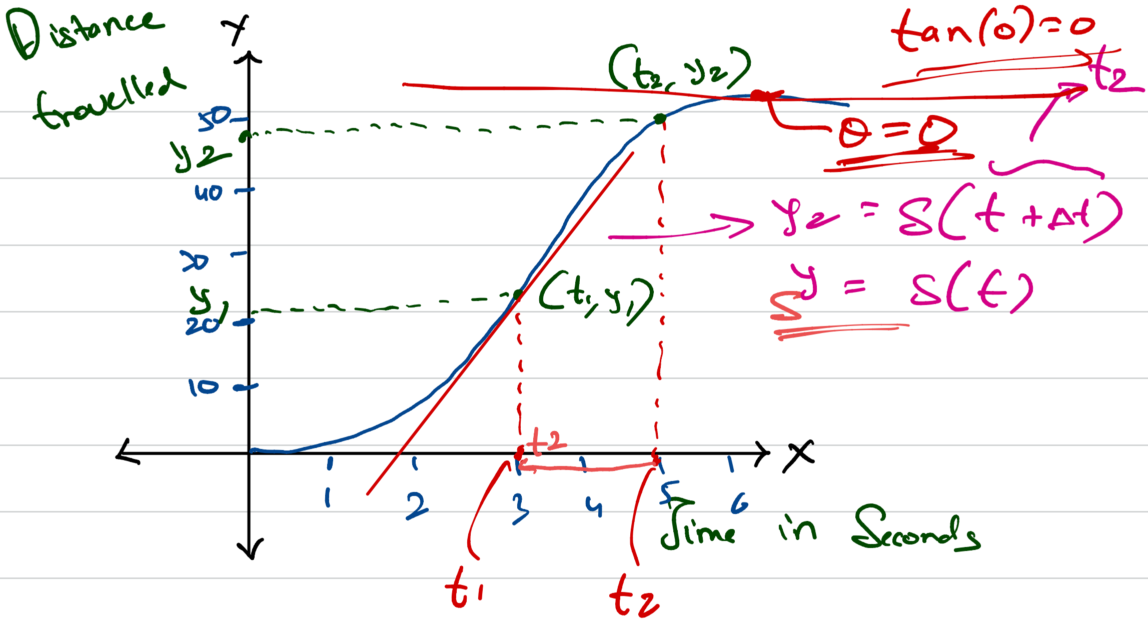
$\rightarrow$ function of speed w.r.t. time

$s(t)$

Secant $\rightarrow$ it touches the curve at two points

$$\tan\theta = \frac{y_2 - y_1}{t_2 - t_1}$$

$$\frac{dy}{dx} = \frac{y_2 - y_1}{t_2 - t_1} \Longleftrightarrow \frac{s(t_2) - s(t_1)}{t_2 - t_1}$$



$$t_2 = t_1 + \Delta t \quad ; \quad \Delta t \to 0 \quad (\Delta t = 10^{-40})$$

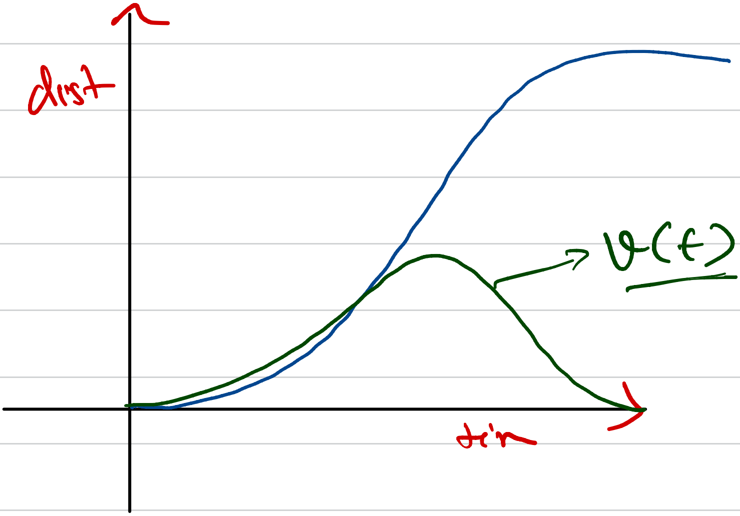
$$y_1 = s(t_1)$$

$$y_2 = s(t_1 + \Delta t) \quad ; \quad \Delta t \to 0$$

$$\frac{dy}{dx} = m = \frac{s(t_1 + \Delta) - s(t_1)}{(t_1 + \Delta t) - t_1}$$

$$\Rightarrow \frac{s(t_1 + \Delta t) - s(t_1)}{\Delta t}$$

$$\boxed{m = \frac{s(t + \Delta t) - s(t)}{\Delta t}}$$



How to find derivative of any given func

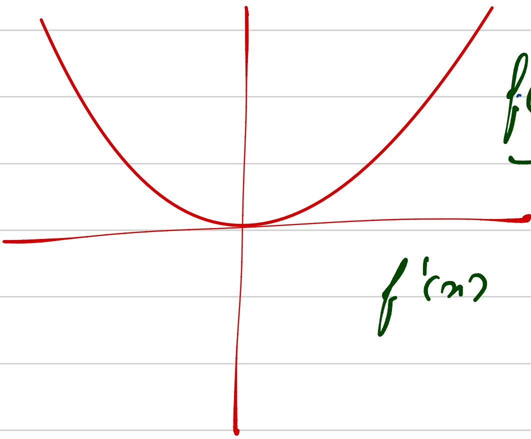
① Simplest way

$$\frac{d\,f(x)}{d(x)} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

→ m ; $\boxed{f'(x)}$

$$f'(x) = \frac{d\,f(x)}{dx}$$

② You can break down a func into smaller parts, use derivative of known function



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

$$f'(x) = \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

$$\Rightarrow \frac{(x+\Delta x)^2 - x^2}{\Delta x}$$

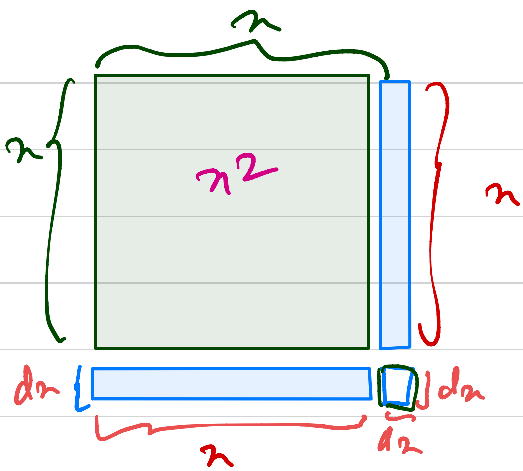
$(a+b)^2 = a^2 + 2ab + b^2$

$$\lim_{\Delta x \to 0} \Rightarrow \frac{\cancel{x^2} + \Delta x^2 + 2x \cdot \Delta x - \cancel{x^2}}{\Delta x}$$

$$\rightarrow \frac{\Delta x^2 + 2x \cdot \Delta x}{\Delta x}$$

$$\Rightarrow \frac{\cancel{\Delta x}(\Delta x + 2x)}{\cancel{\Delta x}} \Rightarrow \cancel{\Delta x}^{\,0} + 2x$$

$$\Rightarrow \underline{\underline{2x}}$$



$f(x) = x^2$, w.r.t. $\underline{dx}$

$f(x) = x^2$

$$f_{new}(x) = x^2 + x \cdot dx + x \cdot dx + dx^2$$

$$f_{new}(x) - f(x)$$

$$\cancel{x^2} + x \cdot dx + x \cdot dx + dx^2 - \cancel{x^2}$$

$$2x \cdot dx + \cancel{dx^2} \Rightarrow 0$$

$$\Rightarrow 2x \cdot dx$$

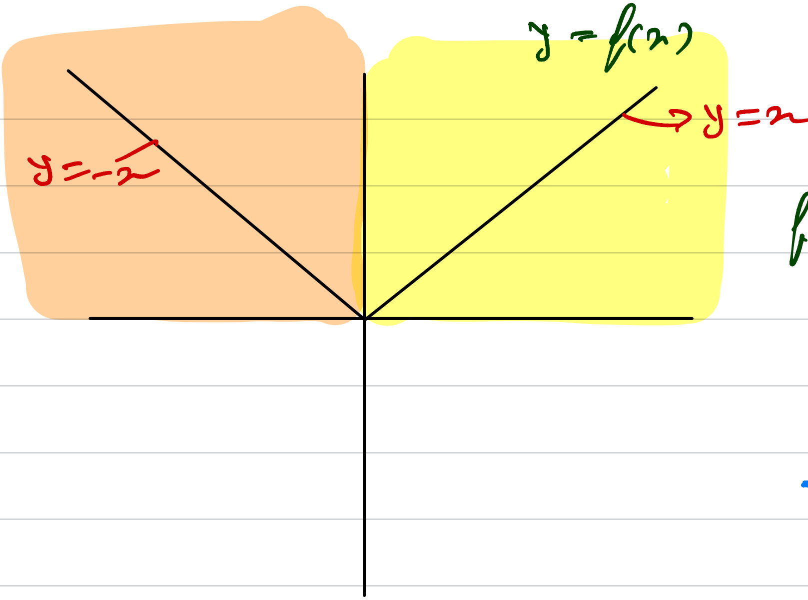
$$\frac{dy}{dx} = \frac{2x \cdot \cancel{dx}}{\cancel{dx}} \Rightarrow \underline{2x}$$

→ Generic Formula

① $\dfrac{d\, x^n}{dx} = n \cdot x^{n-1}$

$$x^2 \Rightarrow 2 \cdot x^{2-1}$$

$$\Rightarrow \underline{\underline{2x}}$$



$y = |x|$

$$f(x) = \begin{cases} x & x > 0 \\ -x & x < 0 \\ 0 & x = 0 \end{cases}$$

$-(-2)$
$= 2$

Case 1: $x > 0$

$f(x) = x$

$$f'(x) = \lim_{\Delta x \to 0} \frac{(\cancel{x} + \Delta x) - \cancel{x}}{\Delta x}$$

$$\Rightarrow \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = 1$$

Case 2: $\underline{\underline{x < 0}}$

$f(x) = \underline{-x}$

derivative $= \underline{\underline{-1}}$

Case 3: $x = 0$

$\underline{f'(x) = 0}$

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

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

If $LHL \neq RHL$ func is not <u>differentiable</u>

$$\text{Speed} = \frac{x}{t}$$

$f(x)$

$$s = ut + \frac{1}{2}at^2$$

s → dist

u → Initial velocity

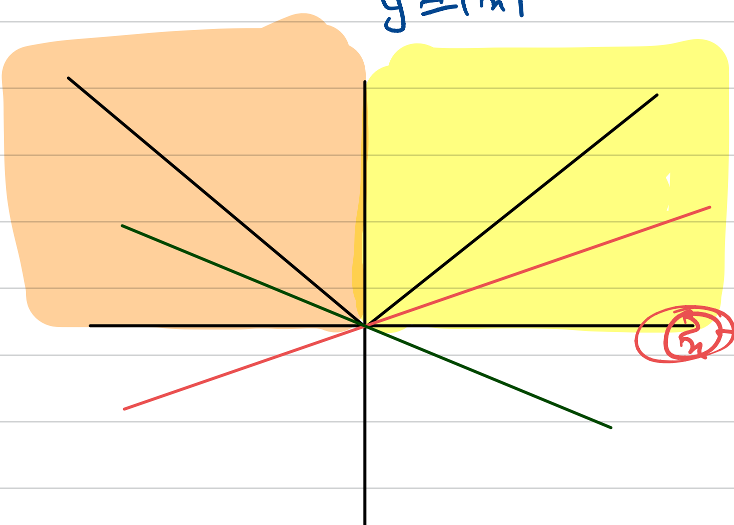
$$\underbrace{\frac{ds}{dt}}_{f'(x)} = u + \frac{1}{\cancel{2}}a \times \cancel{2}t = u + at$$

$$\boxed{v = u + at}$$

$$\underbrace{\frac{dv}{dt}}_{f''(x)} = a$$

$f(x)$ → once → $f'(x)$ → again → $f''(x)$

$y = |x|$



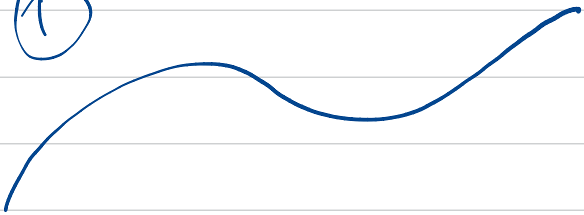
x — 1 tangent

more than one tangent → not differentiable

A func → continuous but not necessarily diff

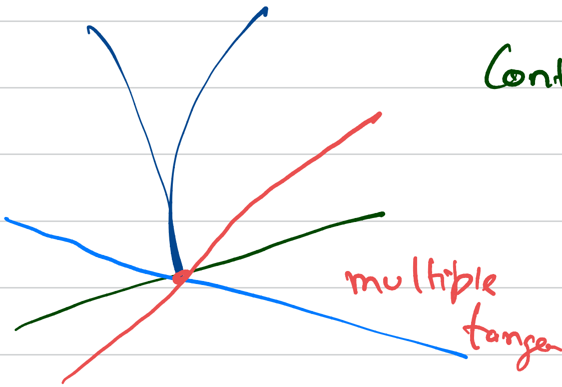
" → discontinuous can't be diff

1



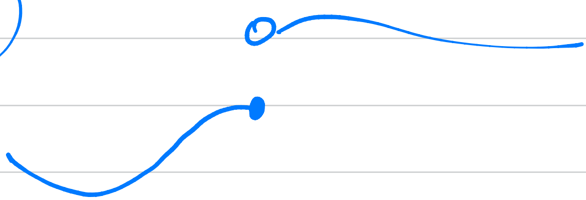
Both Cont & diff.

2

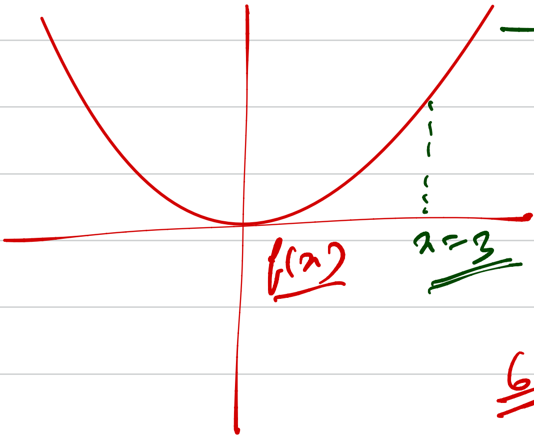


Cont but Not dif

3

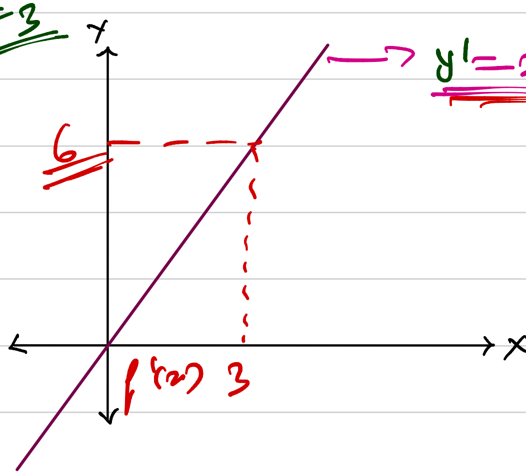


Neither



$$y = x^2$$

$$y' = 2x \quad / \quad y' = 6 \text{ at } x = 2$$



$$y' = 2x$$

→ Some most common derivatives

1. $\dfrac{d}{dx} x^n = n \cdot x^{n-1}$

2. $\dfrac{d \log(x)}{dx} = \dfrac{1}{x}$

3. $\dfrac{d\, e^x}{dx} = e^x$

const

4. $\dfrac{d\, c}{dx} = 0$

5. $\dfrac{d \sin(x)}{dx} = \cos x$

6. $\dfrac{d \cos(x)}{dx} = -\sin(x)$

7. $\dfrac{d \tan(x)}{dx} = \sec^2(x)$

$\sin x = \frac{P}{h}$ / $\cos(x) = \frac{b}{h}$ / $\tan x = \frac{P}{b}$

$\operatorname{cosec}(x) = \frac{1}{\sin x} = \frac{h}{P}$ / $\sec(x) = \frac{1}{\cos(x)} = \frac{h}{b}$ / $\cot(x) = \frac{1}{\tan x} = \frac{b}{P}$

Rules for differentiation

1. Linearity

$$h(x) = g(x) + f(x)$$

$$h'(x) = g'(x) + f'(x)$$

Ex: $x^3 + \log(x)$

$x^3 \to 3 \cdot x^{3-1} \to 3x^2$

$\log(x) \to \frac{1}{x}$

$\Rightarrow 3x^2 + \frac{1}{x}$

$$h(x) = c \cdot f(x)$$

$$h'(x) = c \cdot f'(x)$$

Ex $f(x) = 3 \cdot \sin(x)$

$$f'(x) = 3 \cdot \cos(x)$$

2. Product Rule

$$h(x) = f(x) \cdot g(x)$$

$$h'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

Ex: $f(x) = x \cdot \sin(x)$

$$f'(x) = \frac{d(x)}{dx} \cdot \sin(x) + \frac{d(\sin(x))}{dx} \cdot x$$

$$\Rightarrow 1 \cdot \sin(x) + \underline{\cos(x) \cdot x}$$

③ Quotient Rule

$$h(x) = \frac{f(x)}{g(x)}$$

$$h'(x) = \frac{g(x) \cdot f(x)' - g'(x) f(x)}{\{g(x)\}^2}$$

Ex: $\frac{\sin x}{\cos(x)}$ — $f(x)$, — $g(x)$

$$\frac{\cos(x) \cdot \cos(x) + (\sin x) \sin x}{\cos^2 x}$$

$$\Rightarrow \frac{\cos^2 x + \sin^2 x}{\cos^2 x}$$

$\sin^2 x + \cos^2 x = 1$

$$\Rightarrow \frac{1}{\cos^2 x} = \sec^2 x$$

4 Chain Rule

$$h(x) = f(g(x))$$

$$h'(x) = f'(g(x)) \cdot g'(x)$$

Ex. 1 $\log(x^2) = f(g(x))$

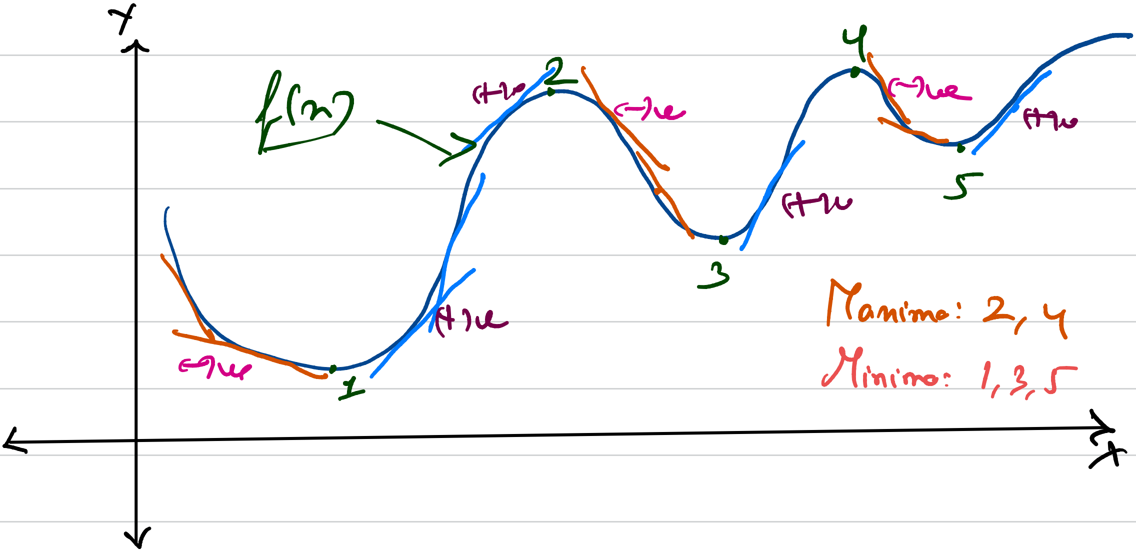
$$\frac{1}{\cancel{x^2}} \times 2 \cdot \cancel{x} \Rightarrow \frac{2}{x}$$

2 $e^{-x} \Rightarrow e^{-x} \cdot (-1) \Rightarrow -e^{-x}$

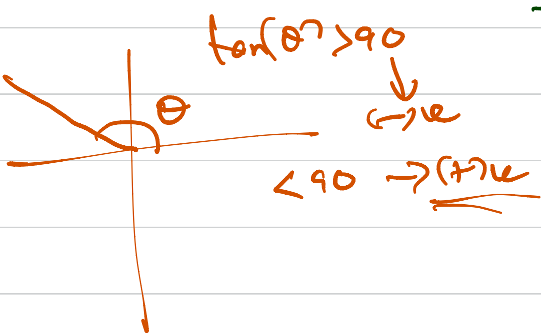
$\frac{4}{5}$

How To Use Derivatives For Optimisation.

$f''(x) = ??$



* The points of a function where derivative = 0 ; those points are called
  → Points of inflection
  or
  → Critical points



- At critical points slope change from -ve to +ve or vice versa

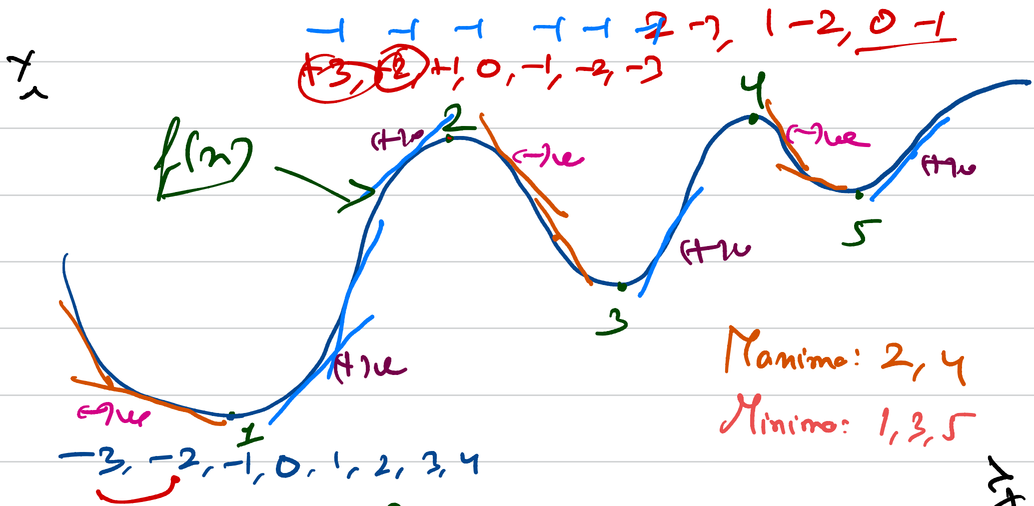
→ At 1, 2, 3, 4, 5, $f'(x) = 0$

→ How to identify critical points of $f(x)$

$f'(x) = 0$ } all the pts that satisfy ↓ critical point

→ How to tell if given critical point is maxima or minima?

- take 2 points to left and right of crt. point compare slope.



Maxima: 2, 4

Minima: 1, 3, 5

$(-2) - (-3) = +1$

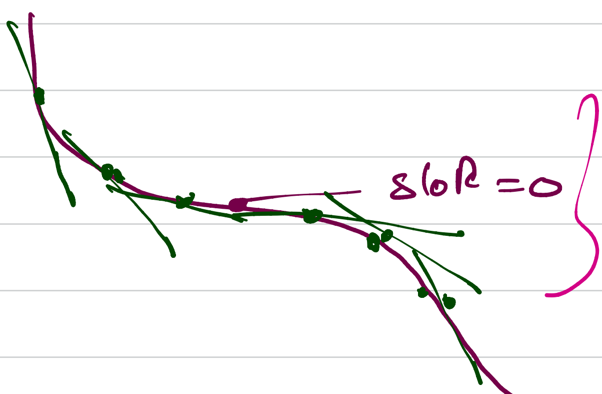
$f'(x)$ = Rate at which $f(x)$ is changing w.r.t. x.

$-1 - (-2) = +1$

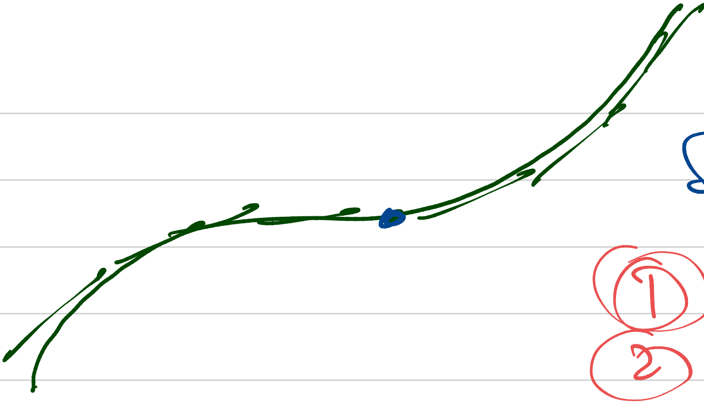
$f''(x)$ = Rate at which my slope is changing w.r.t. x.

$f''(x) \rightarrow (+)$ve (minima)

$f''(x) \rightarrow (-)$ve (maxima)



saddle point



So a critical Point

1 → Minima

2 → Maxima

3 → Saddle Point