March 2, 2023

DSML : Math for ML.

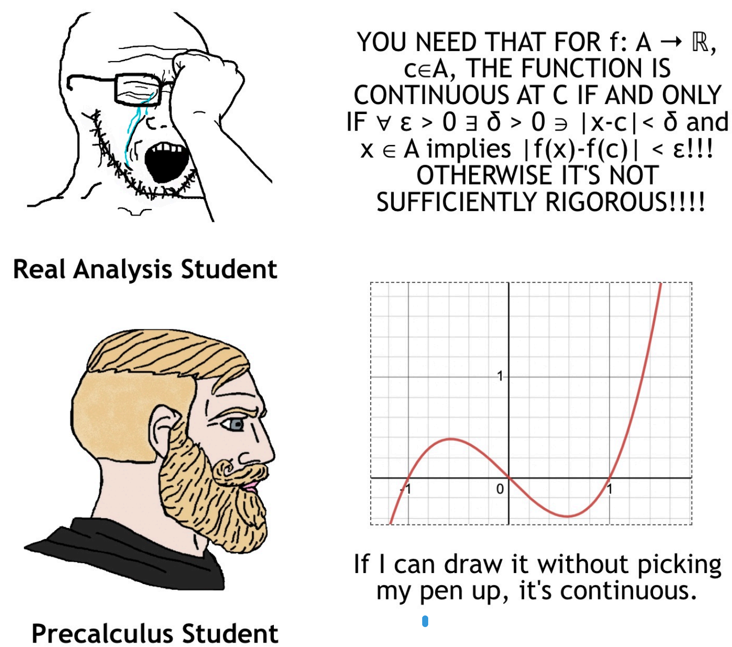
Optimization 2 : Towards Gradient Descent.

Agenda for today :

(a) Recap.

(b) Derivatives in a single dimension.

(c) Derivatives in multiple dimensions.



Detailed Recap :

$$\boxed{w_0^*, \bar{w}^* = \underset{\bar{w}, w_0}{\arg\max} \; \ell\left(D ; \bar{w}, w_0\right)}$$

The problem we want to solve.

→ Simple example : (1-D)

$$x^* = \underset{x}{\arg\max} \; -(x-2)^2$$

max value is 0

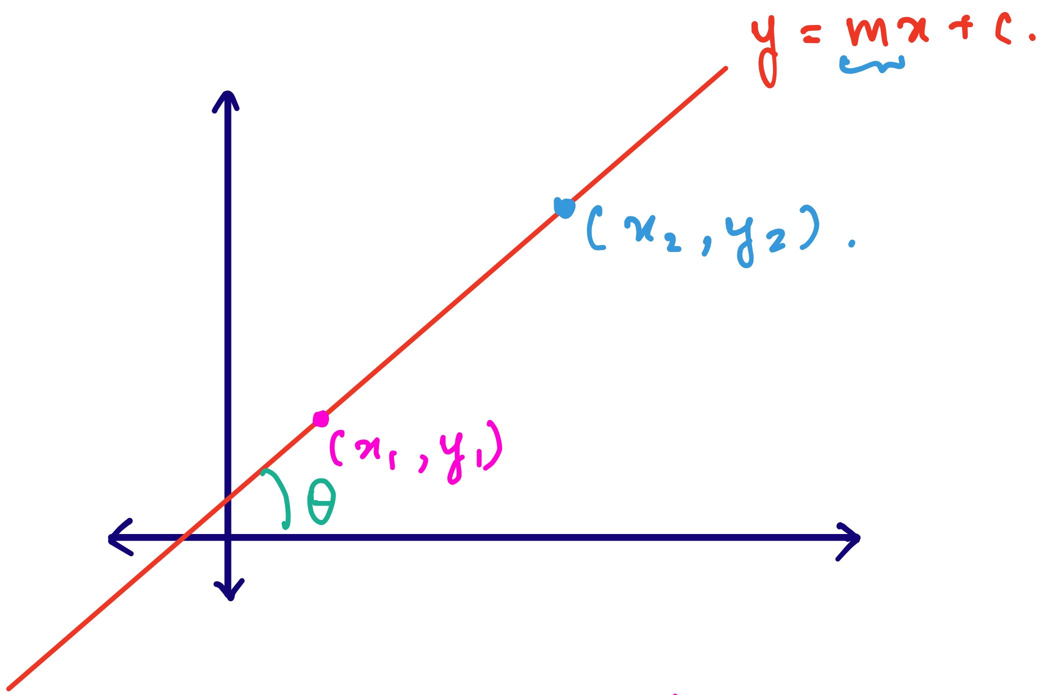
but, the x which gives us the max value is

$x = 2$.

Cheat Sheet :

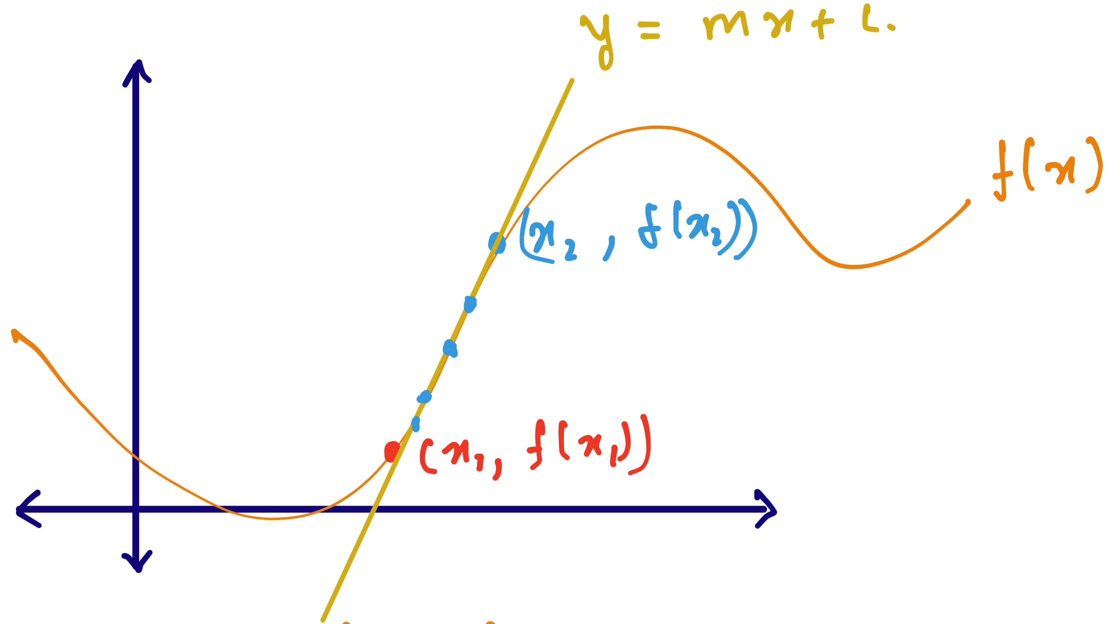
1. Left-hand limit (LHL) : $\lim_{x \to a^-} f(x)$.
2. Right-hand limit (RHL) : $\lim_{x \to a^+} f(x)$
3. Two-sided limit : $\lim_{x \to a} f(x)$
4. Domain : Set of all possible inputs.
5. Range : Set of all possible outputs.

Differentiation : Geometric Picture.



$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

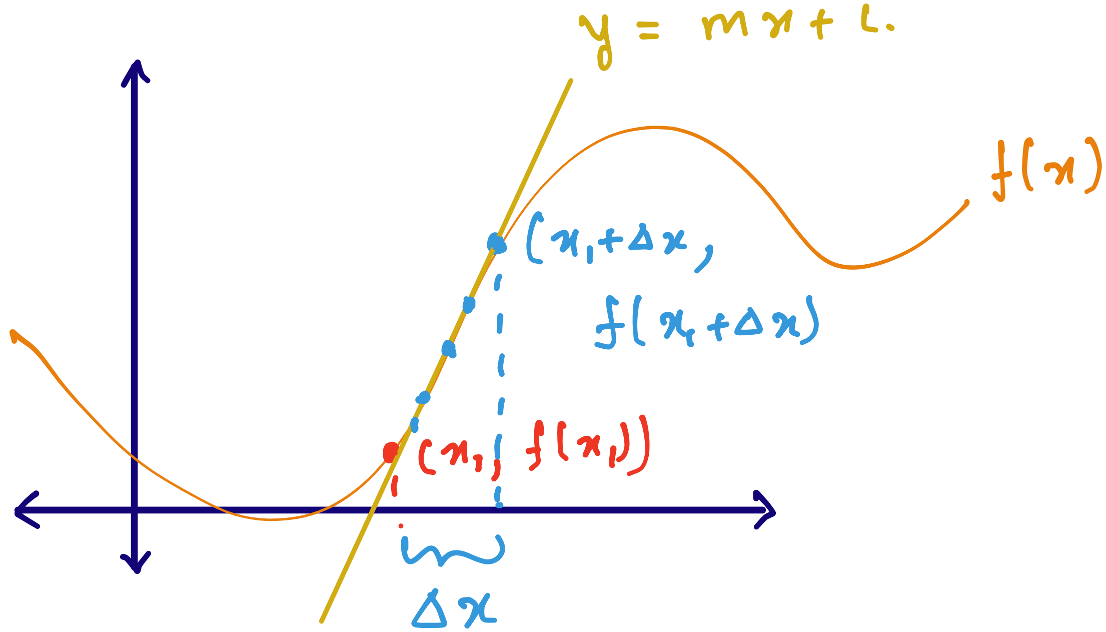
$$m = \tan(\theta).$$



$$y_1 = f(x_1)$$

$$y_2 = f(x_2)$$

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



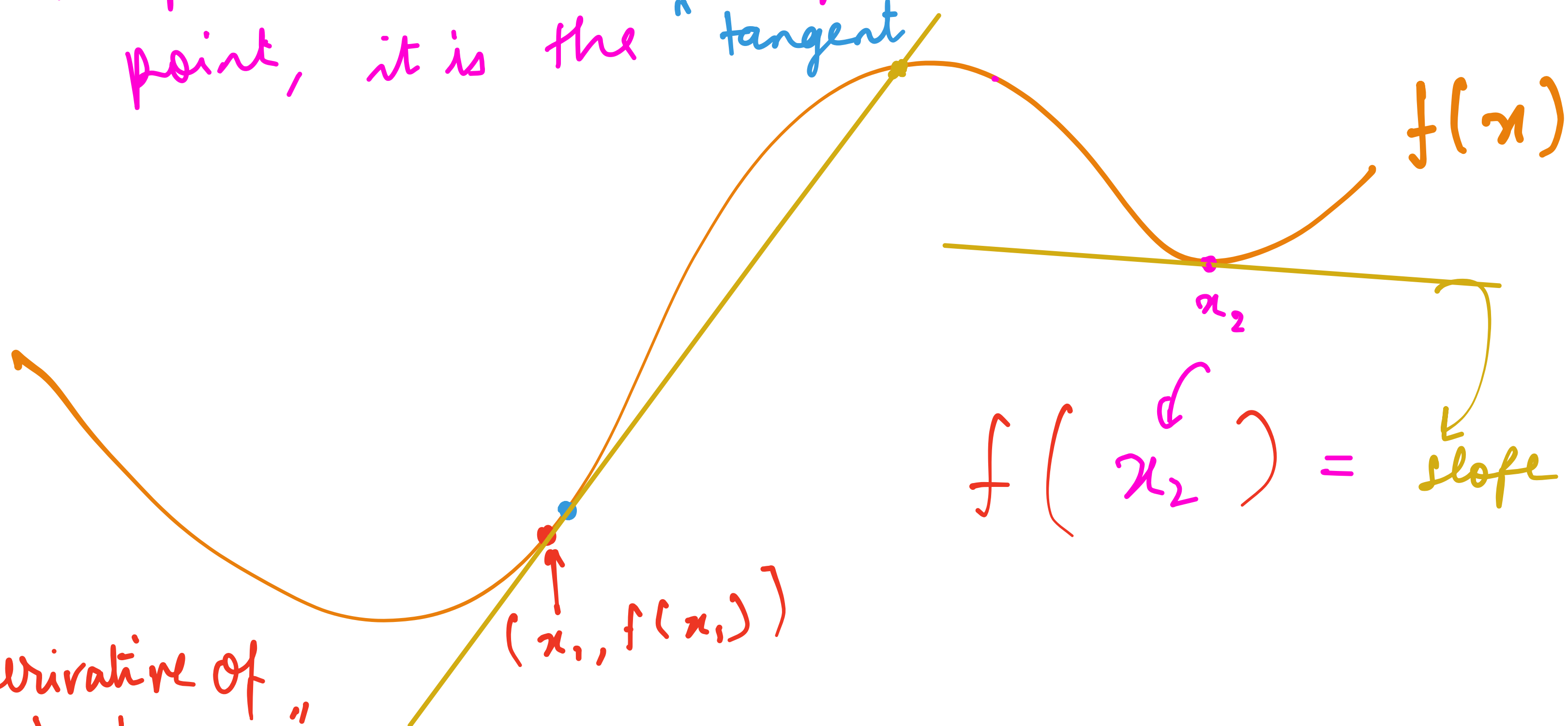
$$m = \frac{f(x_1 + \Delta x) - f(x_1)}{x_1 + \Delta x - x_1}$$

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

$\Delta x \to 0$.

* If a line joins any 2 points on a function, it is called "secant"

* If a line touches a function at a single point, it is the "tangent"



"Derivative of $f(x)$ at x_1"

$$f'(x_1) = m = \lim_{\Delta x \to 0} \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x}$$

How to find the derivative function for any given function $f(x)$?

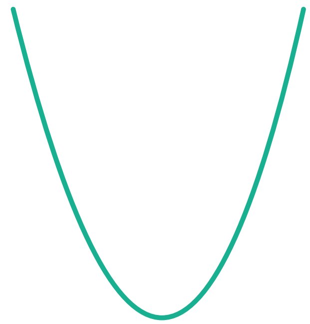
① Simplest way : Use the definition directly.

"Ab initio" method.

$$\frac{d f(x)}{dx} = f'(x_1) = \lim_{\Delta x \to 0} \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x}$$

② First, we remember some well known derivatives.

Then we use 4 rules to calculate it for unknown functions.



$\rightarrow$ parabola.

$f(x) = x^2$.

$$\frac{d}{dx} f(x) = \underline{\underline{2x}}$$

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

$$= \lim_{\Delta x \to 0} \frac{(x + \Delta x)^2 - (x)^2}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{\cancel{x^2} + (\Delta x)^2 + 2x\,\Delta x - \cancel{x^2}}{\Delta x}.$$

$$= \lim_{\Delta x \to 0} \frac{\cancel{\Delta x}\left(\Delta x + 2x\right)}{\cancel{\Delta x}}$$

$$= \lim_{\Delta x \to 0} \left(\Delta x + 2x\right) = \underline{\underline{2x}}$$

$$\boxed{\frac{d}{dx} x^n = n x^{n-1}}$$ → will be useful later, let's remember this.

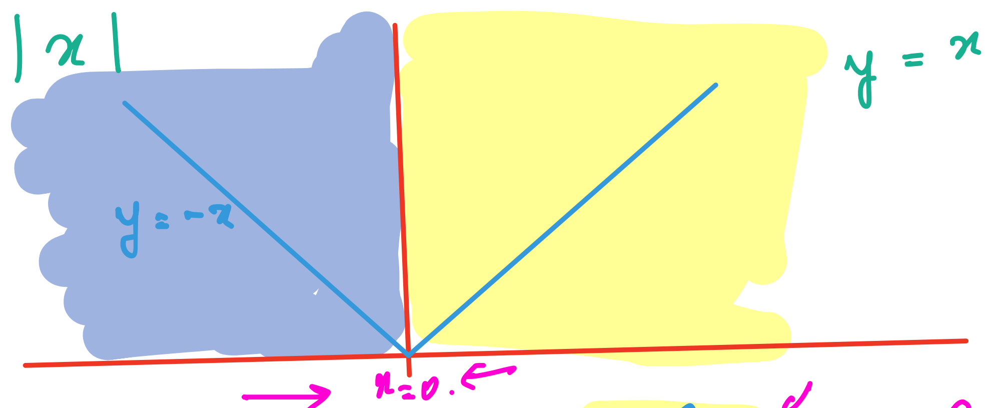
$$f(x) = x$$

$$f'(x) = 1$$

$$f(x) = -x$$

$$f'(x) = -1.$$

$f(x) = |x|$



Case 1 : $x > 0$

$f(x) = x$

$f'(x) = 1$

Case 2 : $x < 0$

$f(x) = -x$

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

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

$$= \lim_{\Delta x \to 0} \frac{\cancel{-x} - \cancel{\Delta x} + \cancel{x}}{\cancel{\Delta x}}$$

$$= \lim_{\Delta x \to 0} -\underline{1} = \underline{-1}$$

Case 3 : $x = 0$

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

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

Is $f'(x)$ a continuous function?

NO

How do we check if a function is differentiable?

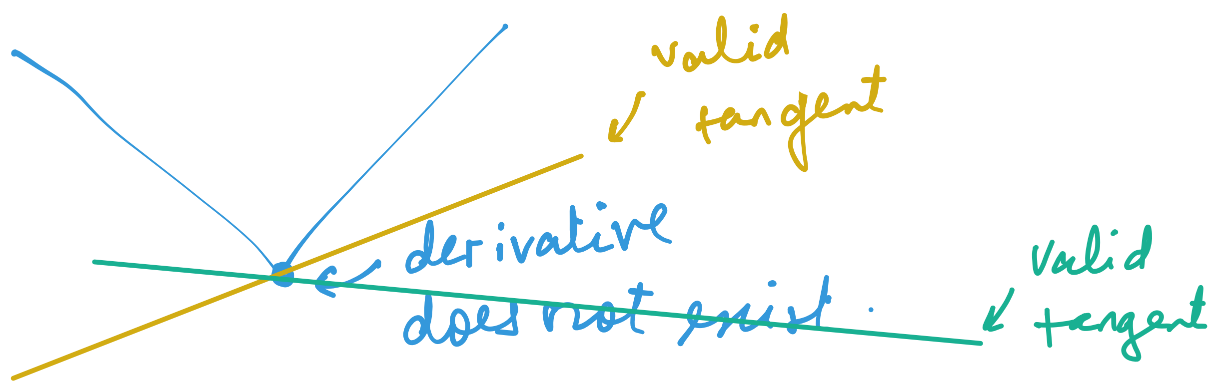
Step 1 : Get $f'(x)$.

Step 2 : If $f'(x)$ is discontinuous, then we say $f(x)$ is not differentiable!!

How to do this visually?

* If we have some sharp points, where we can draw multiple tangents, then we know that the function is non-differentiable.

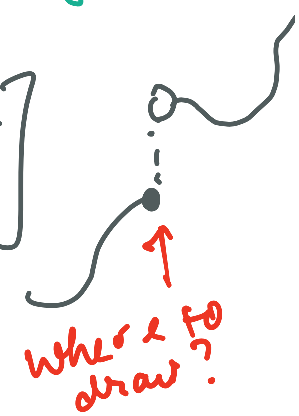
Eg 1 :



I]

II]

III]



Rules of Differentiation:

Some common derivatives to remember:

(a) $\frac{d}{dx} x^n = n x^{n-1}$

(b) $\frac{d}{dx} \log(x) = \frac{1}{x}$

(c) $\frac{d}{dx} e^x = e^x$

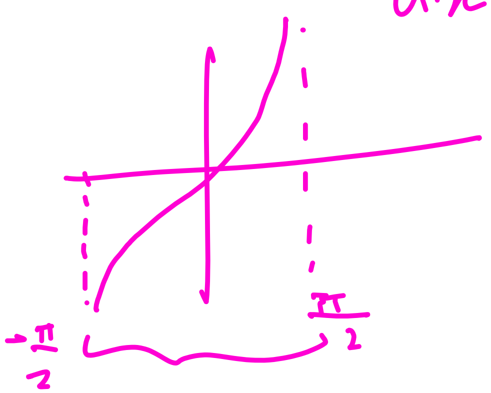
(d) $\frac{d}{dx} (c) = 0$

(e) $\frac{d}{dx} \sin(x) = \cos(x)$

(f) $\frac{d}{dx} \cos(x) = -\sin(x)$

(g) $\frac{d}{dx} \tan(x) = \sec^2 x$

$= \frac{1}{\cos^2(x)}$



$\tan(x)$

(1) Linearity rule : (a) Let $h(x) = f(x) + g(x)$

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

Eg : $f(x) = x^3 + \log(x)$

$f'(x) = 3x^2 + \frac{1}{x}$

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

$c \in \mathbb{R}$.

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

Eg : $f(x) = 3 \sin(x)$

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

② Product Rule : Let $h(x) = f(x) \cdot g(x)$.

Then $h'(x) = \underbrace{f'(x)} \cdot g(x) + f(x) \cdot g'(x)$

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

$= x \cdot x$.

$= 1 \cdot x + x \cdot 1$

$= x + x$

$= 2x$

$f(x) = \underbrace{\sin(x)} \cdot \underbrace{\cos(x)}$.

$= \cos(x) \cdot \cos(x) + \sin(x)(-\sin(x))$

$= \cos^2(x) - \sin^2(x) = \cos(2x)$.

③ Quotient Rule:

Let $h(x) = \dfrac{f(x)}{g(x)}$

Then $\dfrac{d}{dx}[h(x)] = \dfrac{d}{dx}\left[\dfrac{f(x)}{g(x)}\right]$

$$= \frac{g(x)\,f'(x) - g'(x)\,.\,f(x)}{[g(x)]^2}.$$

Eg: $f(x) = \dfrac{\sin(x)}{\cos(x)}$

$$= \frac{\cos(x)\,.\,\cos(x) - (-\sin(x))\,.\,\sin(x)}{(\cos(x))^2}$$

$$= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \underline{\underline{\sec^2 x}}$$

④ Chain Rule : Let $h(x) = f(g(x))$

Then,

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

Eg: ① $f(x) = \log(x^2)$

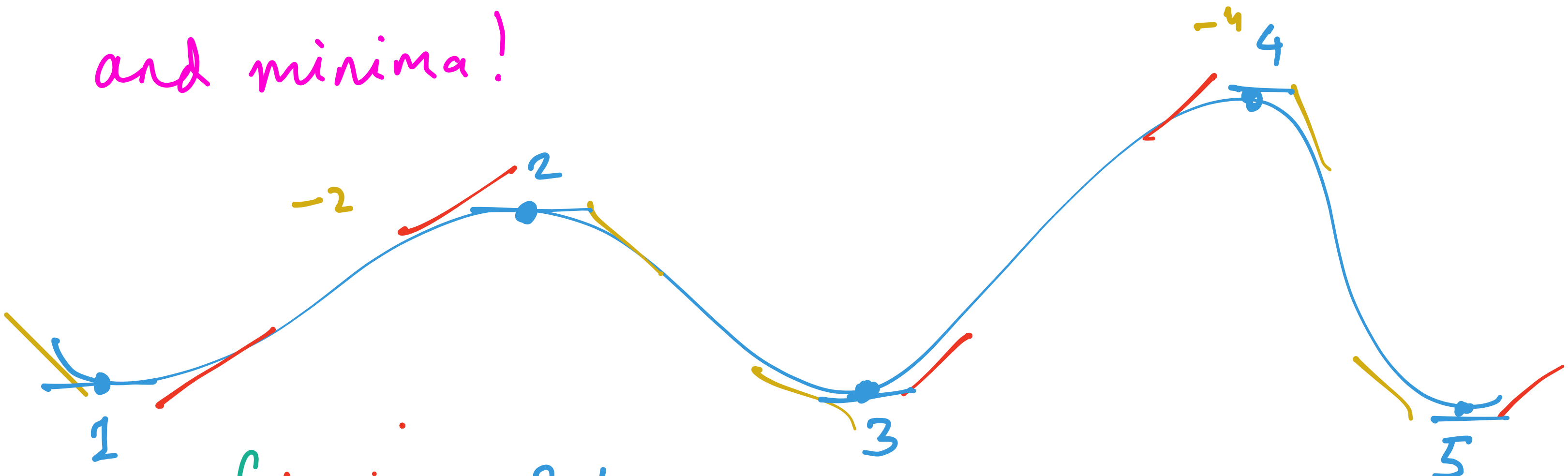
$$= \frac{1}{x^{\cancel{2}}} \cdot 2\cancel{x} = \frac{2}{x}.$$

② $f(x) = e^{-x}$ Inner : $-x$
Outer : e^{x}

$$-1 \cdot e^{-x} = -e^{-x}$$

How to use derivatives for Optimization!

(a) Optimization is nothing but finding maxima and minima!

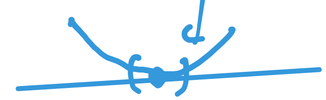


Maxima: 2, 4

Minima: 1, 3, 5.

critical points.

sign(m) = 0



3

1

sign(m) = -1

2

sign(m) = +1

1(b) How to find candidate points for minima & maxima? Given $f(x)$, check if

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

"If derivative is 0, then function is approximately flat at the point"

(c) Once we have the candidate points, how to check if it is a maxima or minima?

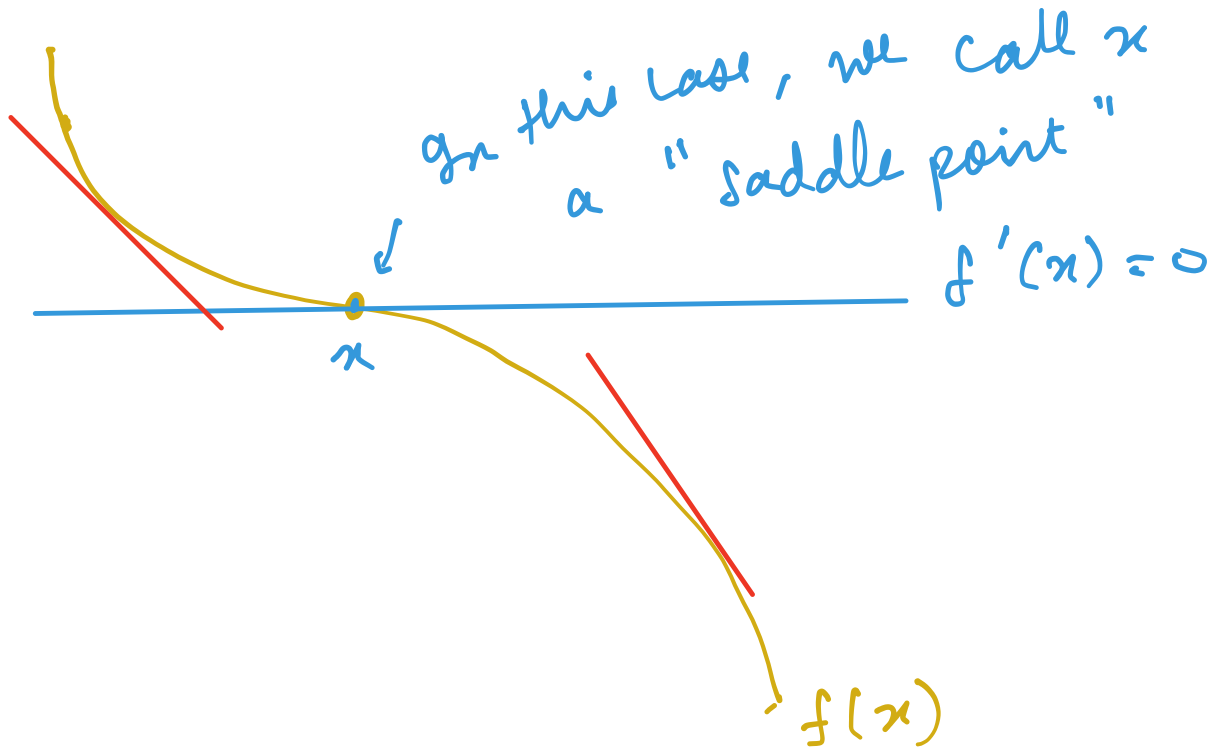
Given $f(x)$ and candidate points, calculate $\boxed{f''(x)}$

$f''(x) > 0$ <u>Minima</u>

$f''(x) < 0$ <u>Maxima</u>

$f''(x) = 0$. saddle point

rate of change of the slope.



$-3, -2, -1, 0, -1, -2, -3, -4.$

If $f''(x) = 0$, the derivative test is inconclusive.

There are four possibilities, the first two cases where *c* is an extremum, the second two where *c* is a (local) saddle point: $\frac{d^6 f(x)}{dx^6}$

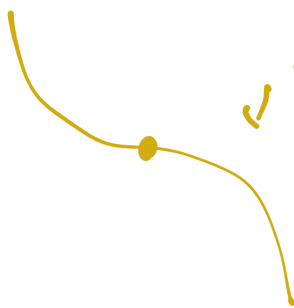
Eg: If $\frac{d^6 f(x)}{dx^6} < 0$.

5

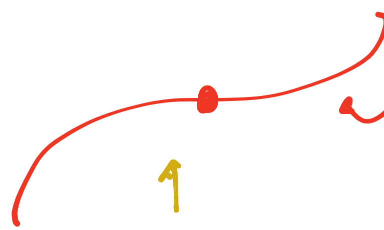
- If *n* is odd and $f^{(n+1)}(c) < 0$, then *c* is a local maximum.
- If *n* is odd and $f^{(n+1)}(c) > 0$, then *c* is a local minimum.
- If *n* is even and $f^{(n+1)}(c) < 0$, then *c* is a strictly decreasing point of inflection.
- If *n* is even and $f^{(n+1)}(c) > 0$, then *c* is a strictly increasing point of inflection.

If $\frac{d^5 f(x)}{dx^5} < 0$

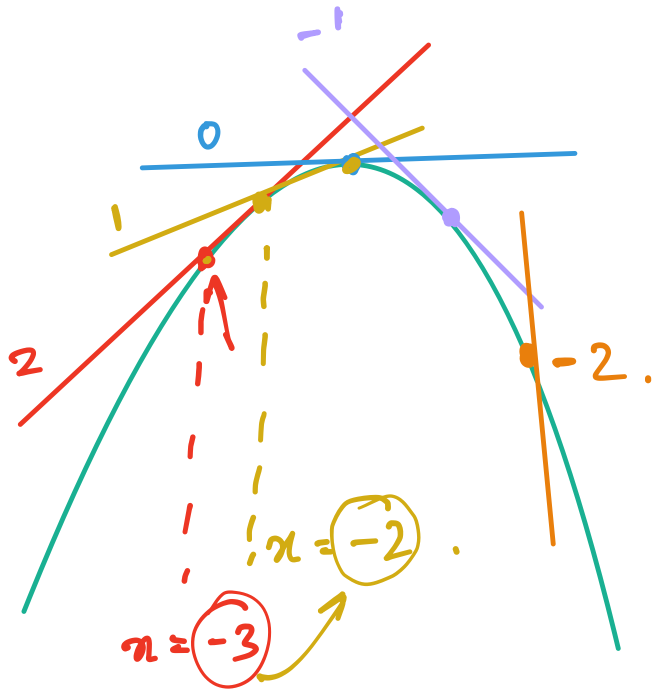
Since *n* must be either odd or even, this analytical test classifies any stationary point of *f*, so long as a nonzero derivative shows up eventually.



strictly decreasing point of inflection.



a strictly increasing point of inflection.



If $f''(x) < 0$

then we have a maxima.

$f'(x)$.

2 1 0 -1 -2

-1 -1 -1 -1 } $f''(x)$.

$$2x - 3y + 6z - 2 = 0$$

Q 5] (Day 83)

- 70kmph.

- 50kmph.

-70.