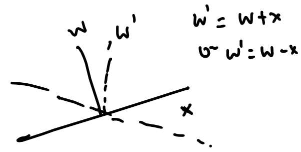
Optimization

Math for ML

→ Prob / Stat (Prob, Cond prob, hypothesisting, Est..)

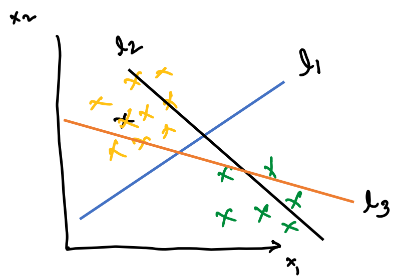
→ Linear Algebra / Coord Geometry - DB, Hyperplane, Dot. products, distances

→ Optimization { Calculus, Gradient descent }



Why optimization?

What is a DB, How do we define (Math) a best 'DB'?

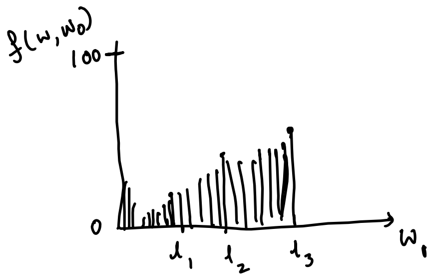


$$\max \uparrow f(w, w_0) = \frac{1}{n} \sum_{i=1}^{n} \frac{w^T \cdot x + w_0}{\|w\|} \cdot y_i$$

w, w_0 where $f(w, w_0)$ Max

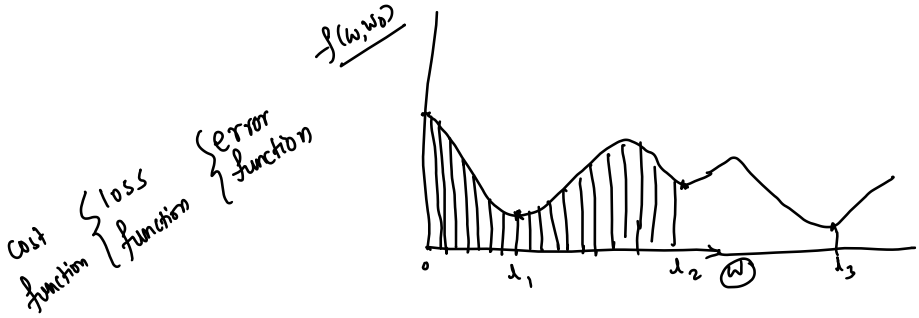
Which is the best 'DB'

$-f(w, w_0)$. Min? ✓



$w[w_1, w_2]$, w_0

Minimum Possible value



Data x_i y_i

(x_i, y_i) $\{x_i - R^d\}$

$x\{x_1, x_2\}$ $x_i - R^2$

Goal → $f(w, w_0)$ w, w_0 Minimum possible value.

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

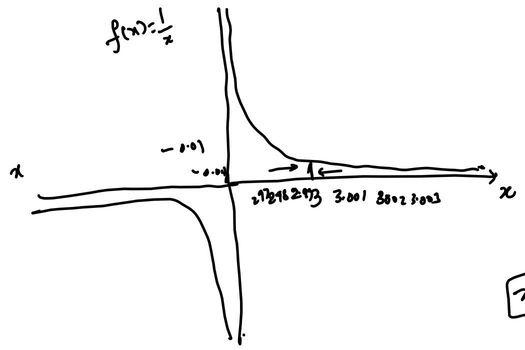
→ calculus → finding maxima or minima of a function.

w, w_0 at which $f(w, w_0)$ takes minimum possible value

Minima $\rightarrow$ derivative of a function $\rightarrow$ Limits and Continuity ?
(Partial derivative)

How do we find out if a function is Continious? $(x \in R)$

$f(x) = \frac{1}{x}$



$0.01 \rightarrow 100$
$0.001 \rightarrow 1000$
$0.0000\ldots \rightarrow \infty$

Mathematical

Limits

$\lim_{x \to k} f(x)$

$\boxed{x = 3}$ at $x = 0$? undefined

$\lim_{x \to 3} f(x) = \frac{1}{3}$

$f(x) = \frac{1}{x}$

$\lim_{x \to 0} f(x) \rightarrow$ Left hand Side limit / Right hand Side limit

$\left.\begin{array}{l} \lim_{x \to 3^+} f(x) \\ \lim_{x \to 3^-} f(x) \end{array}\right\}$

$\lim_{x \to 3}$ RHL

$\frac{1}{3.01} = 0.33\ldots$

$\frac{1}{3.001} = 0.33\ldots$

$\frac{1}{3.0001} = 0.33\ldots$

$\lim_{x \to 3}$ LHL

$\frac{1}{2.90} = 0.32\ldots$

$\frac{1}{2.99} = 0.32\ldots$

$\frac{1}{2.9999} = 0.32\ldots$ ✓

$f(x)$

$x = 3$

$\frac{1}{3} = 0.33$

LHL = RHL = $f(a)$ ✓ $\underline{\underline{a = 3}}$

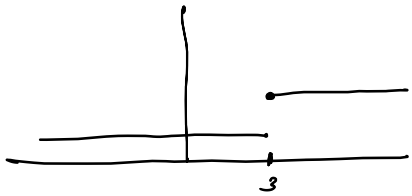
$\underline{\underline{a = 0}}$

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

LHL		RHL		$f(a)$
$\frac{1}{-0.01}$	-100	$\frac{1}{0.01}$	100	
$\frac{1}{-0.001}$	-1000	$\frac{1}{0.001}$	1000	undefined.
$\frac{1}{-0.00001}$	-100000	$\frac{1}{0.0001}$	10000	
$\vdots$	$-\infty$	$\vdots$	$+\infty$	

at '0' LHL = RHL = $f(a)$?

e.g $f(x) = \begin{cases} 2 & x < 3 \\ 5 & x \geq 3 \end{cases}$



$\lim_{x \to 2} f(x) = 2$ ✓

$f(3) = 5$?

$\lim_{x \to 3} f(x)$ < LHL, RHL

LHL	f(x)	f(x)	RHL
2.9 →	2	5	← 3.01
2.99 →	2	5	← 3.001
2.999 →	2	5	← 3.0001

LHL = 2, RHL = 5, $f(a) = 5$?

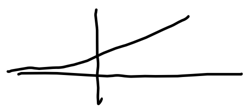
Continuity:-
a function $f(x)$ is Cont. at a if LHL = RHL = $f(a)$ ✓

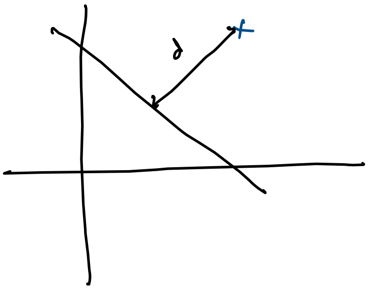
Cont. function:-
if it is Cont. for all values it Can take

all values a function can take is called 'domain'

(o exclu
[o inclusive)

function	Domain	Range (output)
$\frac{1}{x}$	$(-\infty, \infty)$ R	$(-\infty, \infty)$ R
e^x	R	$(0, \infty)$
$\log x$	$(0, \infty)$	R
$\frac{1}{1+e^{-x}}$ (Sigmoid)	R	$(0, 1)$ ← Prob





$d \rightarrow$ Sigmoid $\rightarrow (0-1)$

Examples

Here are a few example problems. Using the definition above, try to determine if they are continuous or not.

EXAMPLE

Is the function $f(x) = \begin{cases} 2x+1 & (x<3) \\ 3x-2 & (x \geq 3) \end{cases}$ continuous for all $x \in \mathbb{R}$?

We know that the graphs of $y = 2x+1$ and $y = 3x-2$ are continuous, so we only need to see if the function is continuous at $x = 3$. The procedure is simply using the definition above, as follows:

(i) Since $f(3) = 3 \times 3 - 2 = 7$, $f(3)$ exists.

(ii) In order to see whether the limit exists or not, we have to check the limit from both sides. The left-hand and right-hand limits are

$$\lim_{x \to 3^-} f(x) = \lim_{x \to 3^-} (2x+1) = 2 \times 3 + 1 = 7 \quad \text{and} \quad \lim_{x \to 3^+} f(x) = \lim_{x \to 3^+} (3x-2) = 3 \times 3 - 2 = 7,$$

respectively. Because the limits from both sides are equal, $\lim_{x \to 3} f(x)$ exists.

(iii) Now from (i) and (ii), we have $\lim_{x \to 3} f(x) = f(3) = 7$, so the function is continuous at $x = 3$. □

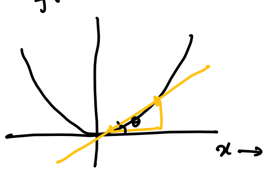
Calculus

derivative

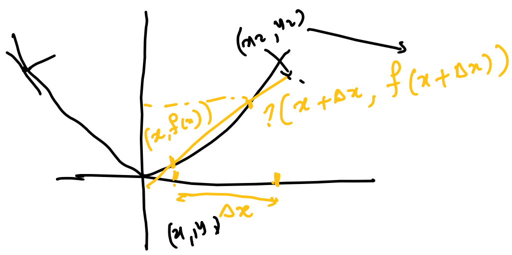
derivatives

$f(x) \rightarrow \{ f'(x)$ or $\frac{d f(x)}{dx} \}$

$f(x) = x^2$



$y = mx + c$, m - slope



$m? \rightarrow \frac{y_2 - y_1}{x_2 - x_1}$

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

if $\Delta x \to 0$ then the line touches the graph at only one point, is called 'Tangent'

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

0.

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

$$\lim_{\Delta x \to 0} \Rightarrow 2x + \Delta x = 2x + 0 = \underline{\underline{2x}}$$

derivative using first Principle

derivative of x^2 : $2x$

derivate of $= x^n = n \cdot \underline{x^{n-1}}$

→ Constant rule $\frac{d}{dx} C = 0$

→ power rule $\frac{d}{dx} x^n = n \cdot x^{n-1}$ (numpy)

→ Product rule $\frac{d}{dx}(f(x) \cdot g(x)) = f'(x) \cdot g(x) + g'(x) \cdot f(x)$

→ [Chain rule → $\frac{d}{dx}(g(f(x))) = g'(f(x)) \cdot f'(x)$] ← -[(dot · products)]

$\partial [E(f(x)) : (\frac{\partial E}{\partial w} \cdot \frac{\partial w}{\partial x})]$

$x^1 \to 1 \cdot x^0$

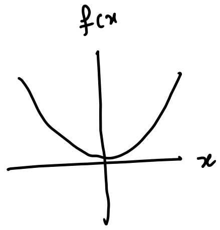
$f(x) = x^2 - x + 2$ $\quad f'(x) = \underline{\underline{2x - 1}}$

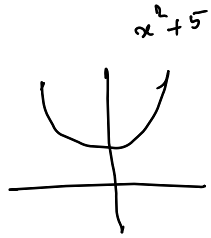
$f((a)+(b)) = f'(a) + f'(b)$ $\quad f(x) = g(x) + h(x) = g'(x) + h'(x)$

→ How to find Maxima or minima?

$f'(x) = \underline{2x-1} = 0 \quad 2x - 1 = 0 = \boxed{x = \frac{1}{2}}$

→ Minimum , minima Same! $f(x) = x^2$





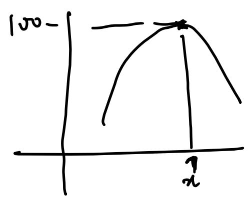
minimum → f(x) can take = 0

minima → a point at which f(x) takes minimum value = 0

f(x) minimum = 5

minima = 0

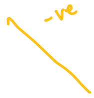
Maximum , Maxima



$f(x)$ max. = 100

maxima = 7

-ve



decreasing

+ve 0 -ve

$f''(x) < 0$

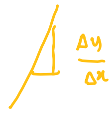
Maxima

Increasing

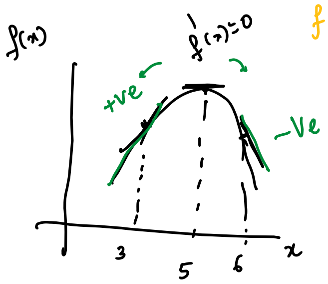
-ve , 0 , +ve

$f''(x) > 0$

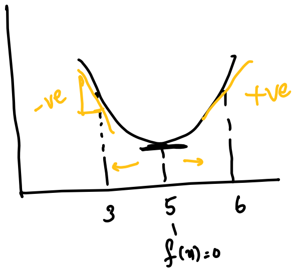
Minima



Maxima



$f''(x)$ -ve

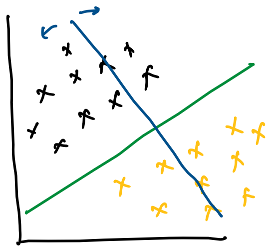


$f''(x) > 0$

$f'(x) = 0$ $\quad f''(x) > 0$ then minima

$f'(x) = 0$ $\quad f''(x) < 0$ then maxima

→ Optimization

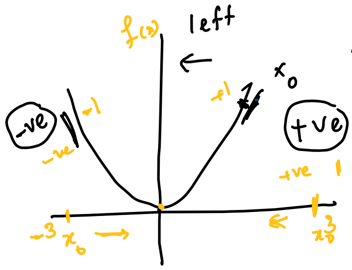


function $f(w, w_0)$

$$\rightarrow \left[-\frac{1}{n} \sum_{i=1}^{n} \frac{w^T \cdot x + w_0}{\|w\|} \cdot y_i \right]$$ Minimum

Randomly initiate $[x, \; f(x) = x^2]$

Randomly took 'x' → $f'(x) \rightarrow x_0$



$x_0 \rightarrow f'(x_0)$

$x = x_0 - f'(x)$

how to move towards Minimum?

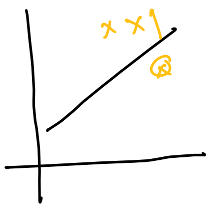
$$\boxed{x = x_0 - \text{gradient}}$$

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

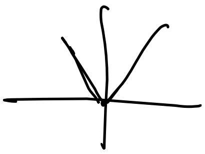
$x = x_0(-\text{gradient})$

$= 3 - 1$

$= 2$

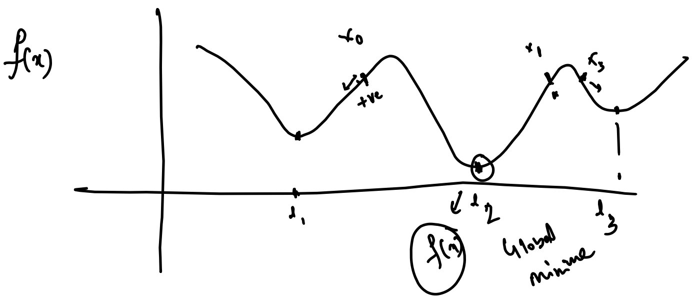


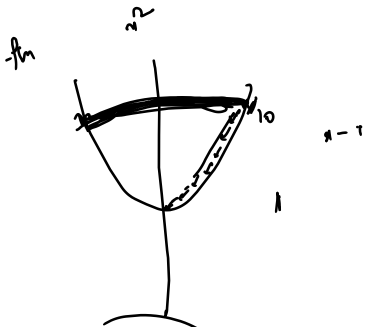
Non-differentiable



(abrupt turn point) - Non derivable

Restart : 10:15 pm



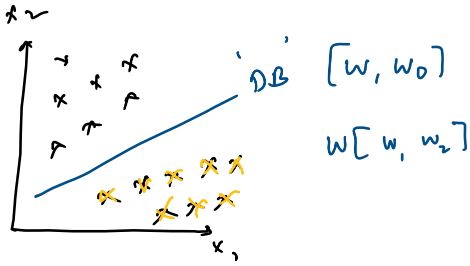


$\Rightarrow$ (X − derivative)

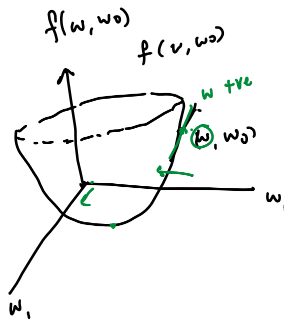
learning rate)

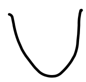
X − (0.1) x derivative

derivatives $f(x)$, x − minima where $f(x)$ takes minimum value.



Gradient (multi dimensional)





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

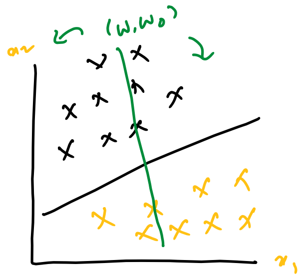
$f'(x)$ $f'(y)$ Partial derivatives

when deriving for x, treat y as Constant

$$f'(x) = \frac{\partial(x^2+y^2)}{\partial x} = 2x$$

$$f'(y) = \frac{\partial(x^2+y^2)}{\partial y} = 2y$$

$x=2,\ y=4$



$$\nabla f(x,y)_{(2,4)} = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \rightarrow \begin{bmatrix} 4 \\ 8 \end{bmatrix} \rightarrow$$

(Gradient descent)

$w = w_0 - (l.R).$ derivate

$$\boxed{w = w_0 - (l.R).\ \text{gradient}}$$

Gradient = (vector)

$$\boxed{w_i = w_i^0 - l.R.\ \text{gradient}}$$