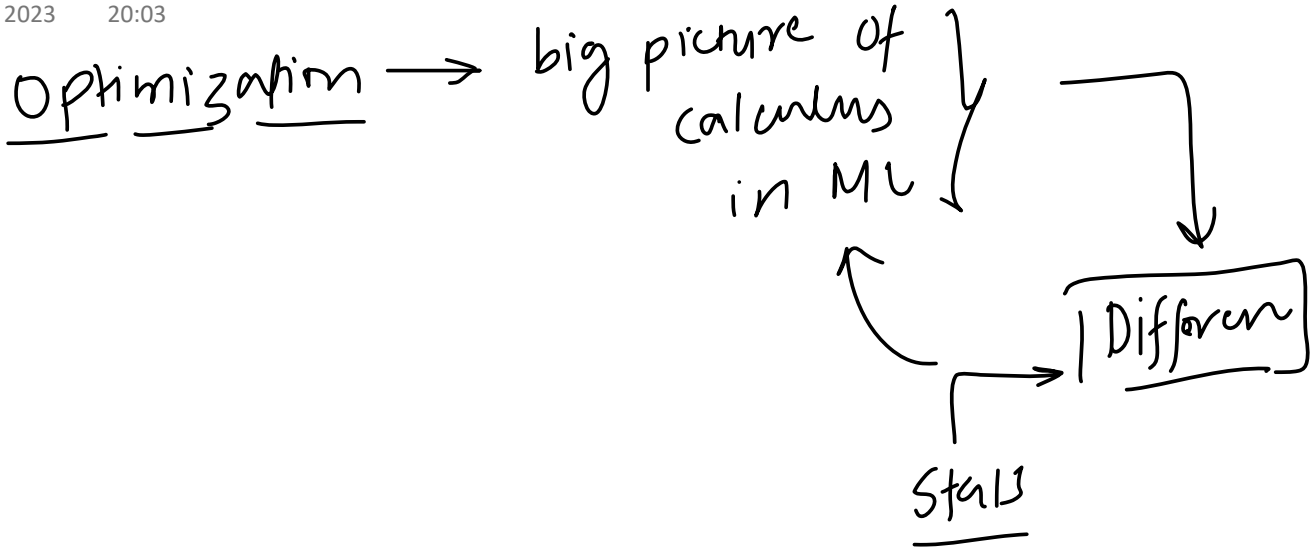
22 April 2023 20:03



What is differentiation

22 April 2023 08:29

Differentiation is the process of finding the derivative of a function. The derivative of a function represents the instantaneous rate of change of the function with respect to its variable, typically denoted as 'x'.

- What is differentiation?
- Why instantaneous?
- Relation with slope
- Maxima and Minima
- How to calculate derivative
- Intuition
- Derivative in ML

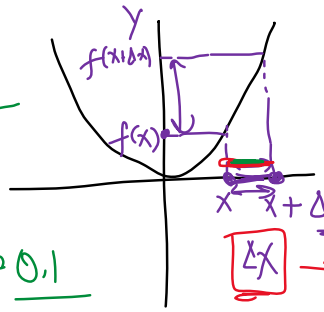
diff → slope

at a point

$\frac{dy}{dx} = 0$

maxima minima

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



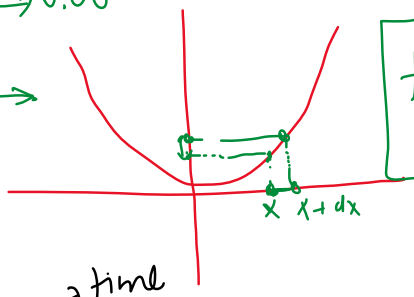
$$\text{rate of change} = \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

$\Delta x \to 0.1$

$dx \to 0.000000001$

$\Delta x \to 0.0000001$

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

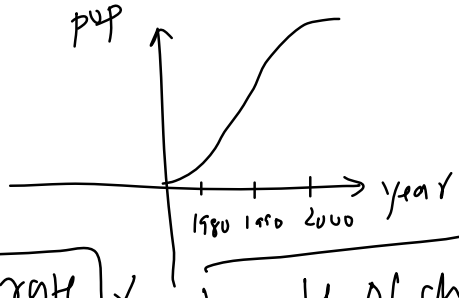


$$\frac{f(x+dx) - f(x)}{dx} = \frac{df(x)}{dx}$$ inst

$y = f(x) = x^3$ (time)

pop over time

pop



1990 ↔ 1980

$\frac{100c - 80c}{10}$

1990

pop → growth rate X

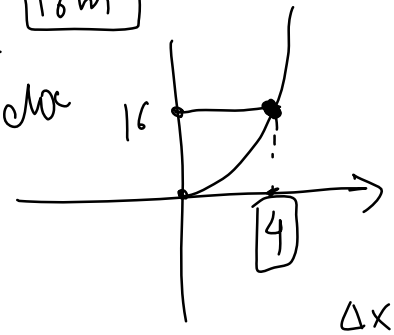
16m

rate of change

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

distance time

velocity



velocity / speed

at t = 4

0 → dist → 0

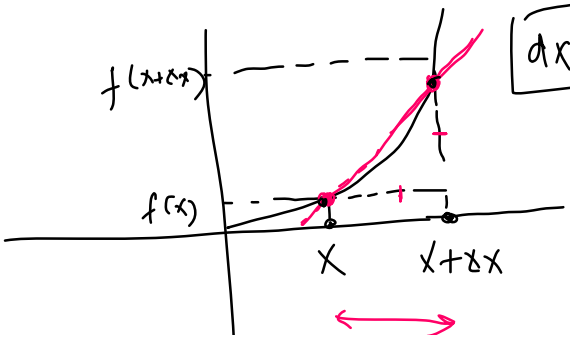
4 → dist → 16

$\frac{16-0}{4-0}$

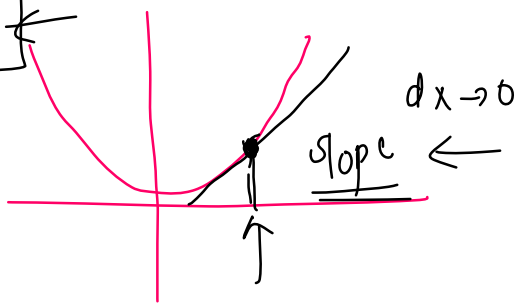
$\frac{16}{4} = 4$ m/s

$s = \frac{d}{t}$

$y = x^2$

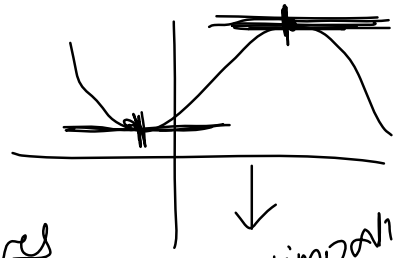


dx



$\frac{df}{dx}$ → slope at that point

x x+dx



diff = 0
derivative

derivatives

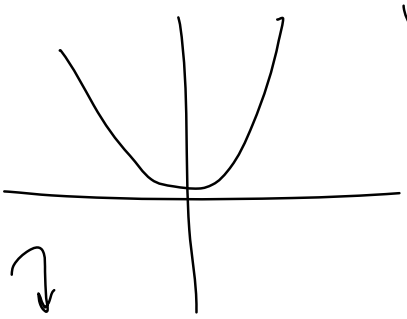
optimization

point

$y = x^2 + 2x$

$\frac{dy}{dx} = 2x + 2 = 0$ $x = -1$

$x = -\frac{2}{2} = -1$



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

$\frac{df(x)}{df} = \frac{f(x+dx) - f(x)}{dx}$

$dx \to 0$

$x^2 \to 2x$

$f(x+dx) = (x+dx)^2 = \frac{(x+dx)^2 - x^2}{dx}$

$f(x) = x^2$

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

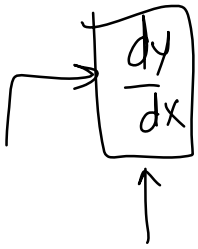
$\frac{df}{dx} = 2x$

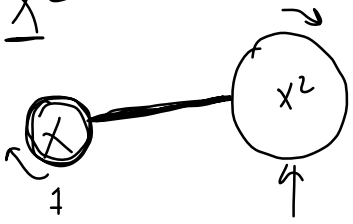
$\frac{(dx)^2 + 2x\,dx}{dx} = dx + 2x = 2x$

($dx \to 0$)

intuition

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





-2

$L(\beta_0)$

$\frac{dL}{d\beta_0}$

$L \to \frac{dL}{d\beta_0} \to 2$

β_0 change

$\beta_0 \to \uparrow\downarrow$

random

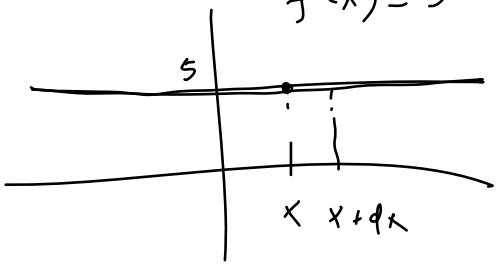
Derivative of a constant

22 April 2023 13:57

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

0

$y = 5$



$f(x) = 5$ $f(x+dx) = 5$

$\frac{dy}{dx} = \frac{f(x+dx) - f(x)}{dx}$

derivative of a constant = 0

$\frac{dy}{dx} = 0$ $5 - 5$

$y = \boxed{5}$

$\frac{dy}{dx} = 0$

Cheatsheet

22 April 2023 11:01

COMMON DERIVATIVES

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

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

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

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

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

$$\frac{d}{dx}(\csc x) = -\csc x \cot x$$

$$\frac{d}{dx}(\cot x) = -\csc^2 x$$

$$\frac{d}{dx}(\sin^{-1} x) = \frac{1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}(\cos^{-1} x) = -\frac{1}{\sqrt{1-x^2}}$$

$$\frac{d}{dx}(\tan^{-1} x) = \frac{1}{1+x^2}$$

$$\frac{d}{dx}(a^x) = a^x \ln(a)$$

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

$$\frac{d}{dx}(\ln(x)) = \frac{1}{x}, x > 0$$

$$\frac{d}{dx}(\ln|x|) = \frac{1}{x}$$

$$\frac{d}{dx}(\log_a(x)) = \frac{1}{x \ln(a)}$$

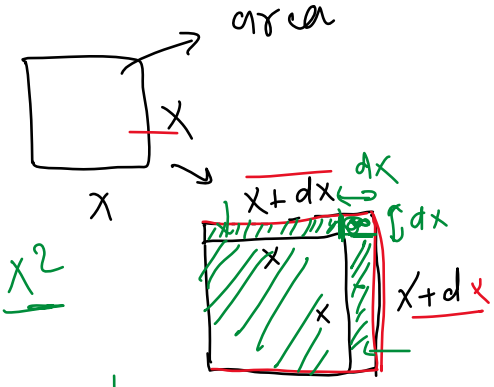
Power Rule

22 April 2023 10:59

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

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

$$\frac{dy}{dx} = \frac{f(x+dx) - f(x)}{dx}$$

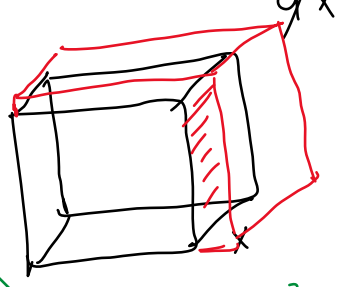


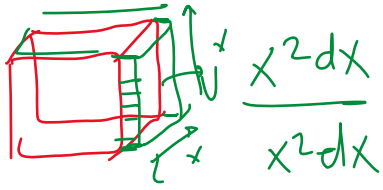
x^2

$f(x+dx) = x^2 + x\,dx + x\,dx + (dx)^2$

$= x^2 + 2x\,dx$

$$\frac{x^2 + 2x\,dx - x^2}{dx} = \frac{2x\,dx}{dx} = 2x$$





$y = f(x) = x^3$

$$\frac{dy}{dx} = \frac{f(x+dx) - f(x)}{dx}$$

$$\frac{x^3 + x^2dx + x^2dx + x^2dx - x^3}{dx}$$

$$\frac{3x^2\,dx}{dx} = 3x^2$$

$$\frac{dy}{dx} = 3x^2$$

power rule

$x^4 \rightarrow 4x^3$

$x^5 \rightarrow 5x^4$

$x^6 \rightarrow 6x^5$

$x^n \rightarrow nx^{n-1}$

$dx \rightarrow (0.000001)^2$

Sum Rule

22 April 2023 11:01

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

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

$$\boxed{(f(x) \pm g(x))' = f'(x) \pm g'(x)} \leftarrow$$

$$\frac{d(x^2 + \log x)}{dx} = \frac{dx^2}{dx} + \frac{d\log x}{dx}$$

$y = x$ $\quad h(x) = f(x) + g(x)$

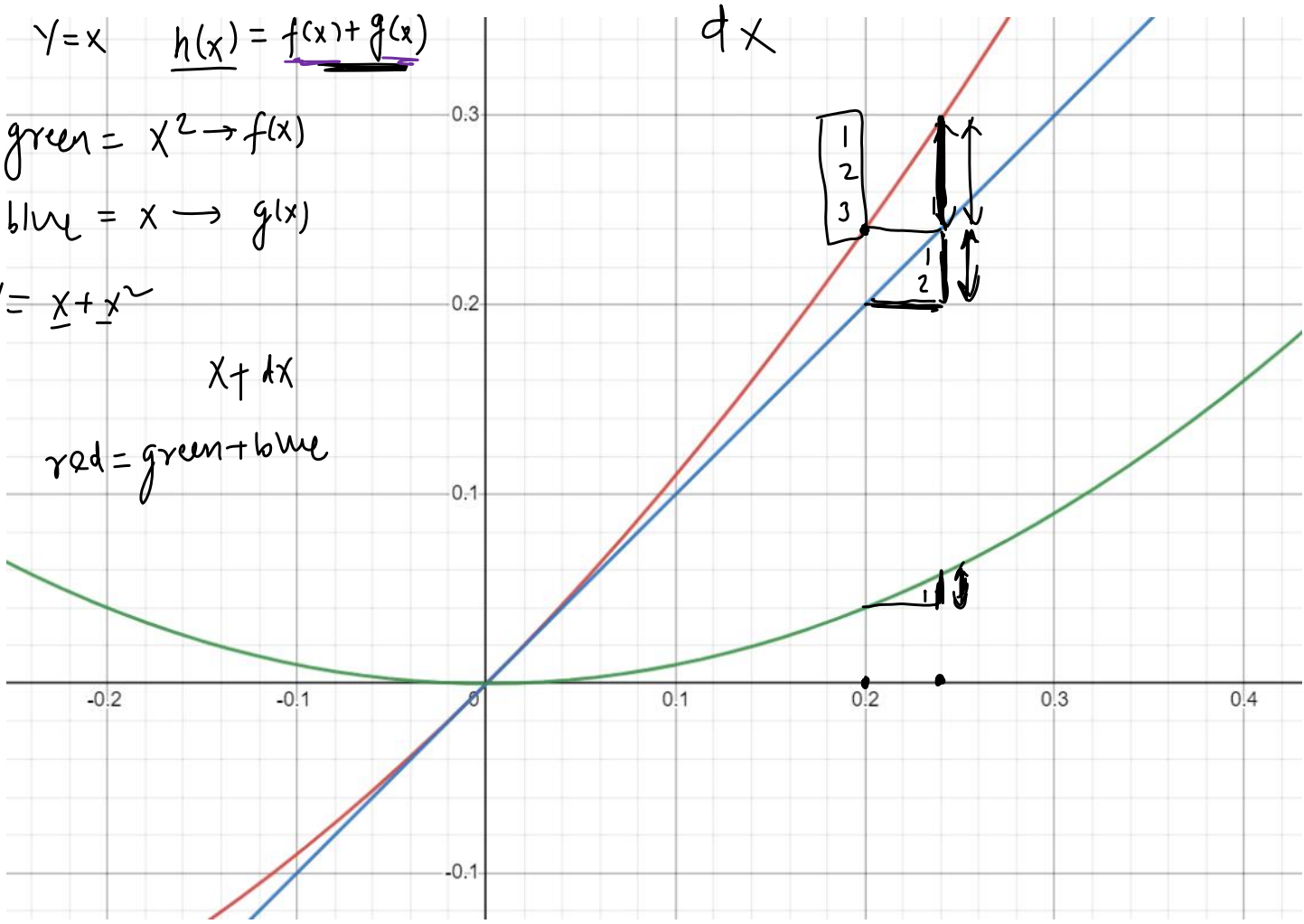
green $= x^2 \rightarrow f(x)$

blue $= x \rightarrow g(x)$

$y = x + x^2$

$x + dx$

red = green + blue



Product Rule

22 April 2023 11:01

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

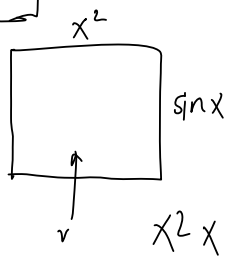
$(cf(x))' = c(f'(x))$

$\frac{dy}{dx} = x^2 \sin x$

$\frac{d}{dx} x^2 \sin x + x^2 \frac{d}{dx} \sin x$

$x^3 \log x$

$y = x^2 \sin x$



$\frac{dy}{dx} = \frac{f(x+dx) - f(x)}{dx}$

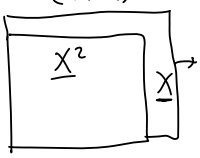
$y = x^2$

$(x+dx)^2 (x+dx)$

$(x+dx)^2$ area $-$ area / dx

$x^2 + x$

$(x+dx)^2$



$(x+dx)$

$+ + \rightarrow$ red area

$a \; b$

$y \uparrow$

$(x^2 + (dx)^2 + 2x\,dx)(x+dx)$

$x^2 x + x(dx)^2 + 2x \cdot x\,dx + x^2 dx + (dx)^3 + 2x(dx)^2 - x^2 x$

dx

$x\,dx + 2x^2 + x^2 + 2x\,dx$

$2x^2 + x^2 + 3x\,dx$

$dx = 0$

$x^2 + x$

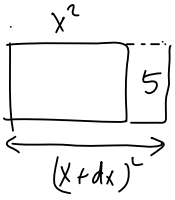
$2x \times + x^2$

$2x^2 + x^2 = 3x^2$

$x^2 x \rightarrow x^3$

$3x^2$

$y = 5x^2 = 5\frac{d}{dx}(x^2) = 5\,2x = 10x$



$5x^2$

$\frac{(x+dx)^2 5 - 5x^2}{dx}$

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

$\frac{5x^2 + 5(dx)^2 + 5(2x)dx - 5x^2}{dx}$

$\frac{5(dx)^2 + 5(2x)dx}{dx} = 5dx + 5(2x)$

$5(2x)$

Quotient Rule

22 April 2023 13:53

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$$

$$\frac{d}{dx}\frac{x^2}{\sin x} =$$

$$x^2(\sin x)^{-1}$$

$a \quad b$

$$\frac{\frac{d}{dx}x^2 \sin x - x^2\frac{d}{dx}\sin x}{(\sin x)^2}$$

$$\frac{dy}{dx} = \frac{2x\sin x - x^2\cos x}{(\sin x)^2}$$

Chain Rule

22 April 2023 11:01

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

$f(x)$ $g(x)$

$y = f(g(x)) = \sin(x^2)$

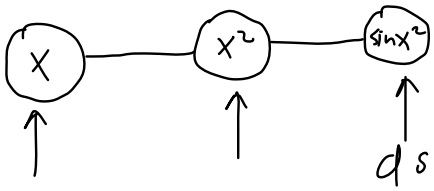
deep learning

$g(x) = x^2$

$f(g(x)) = \sin(x^2)$

$y = f(g(x))$

$$\frac{dy}{dx} = \frac{df}{dg}\frac{dg}{dx}$$



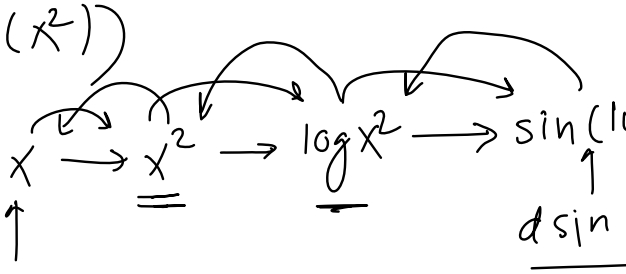
$$\frac{d\sin(x^2)}{dx^2} \times \frac{dx^2}{dx}$$

$y = \sin x^2$

$\frac{dy}{dx}$

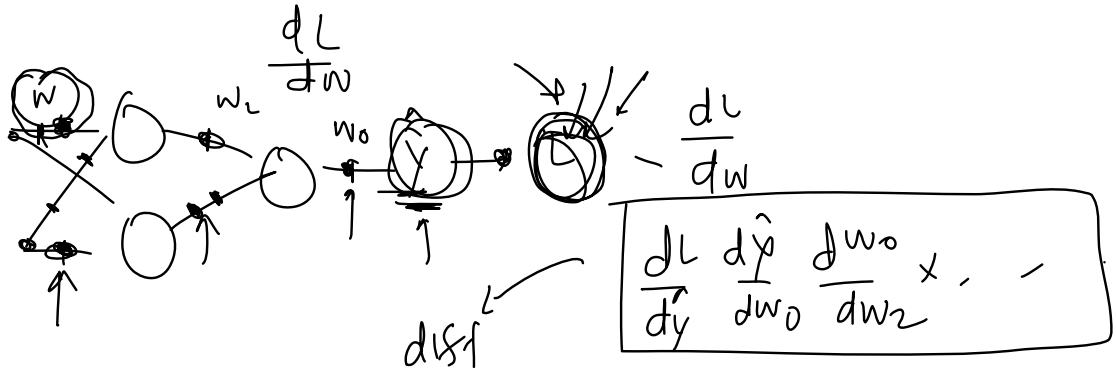
$\cos x^2 \; 2x = \boxed{2x\cos x^2}$

$\sin(\log(x^2))$



$$\frac{d\sin(\log x^2)}{d\log x} \quad \frac{d\log x^2}{dx^2} \quad \frac{dx^2}{dx}$$

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

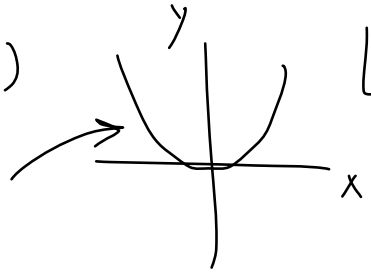


Partial Differentiation

22 April 2023 13:50

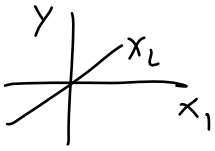
$y = f(x)$ ↑ single variable function

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



2d

$y = f(x_1, x_2)$



surface

$f(x_1, x_2 \dots x_n)$

$z = f(x,y) = x^2 + y^2$ ← 3d parabola

$\frac{dz}{dz}$

$z = x^2 + y^2$

compute

$\frac{\partial z}{\partial x} = \frac{2x+0}{2(1)}$

$\frac{\partial z}{\partial y} = 2y$

$2(2)$

$= 4$

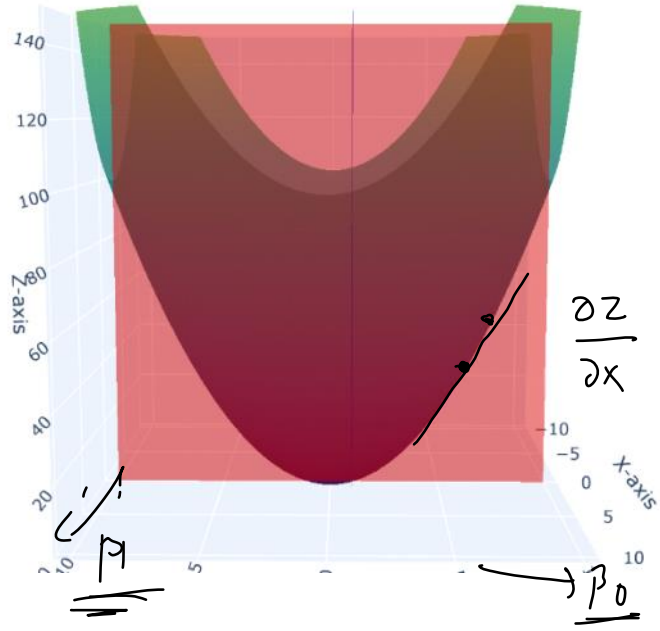
$= 2$

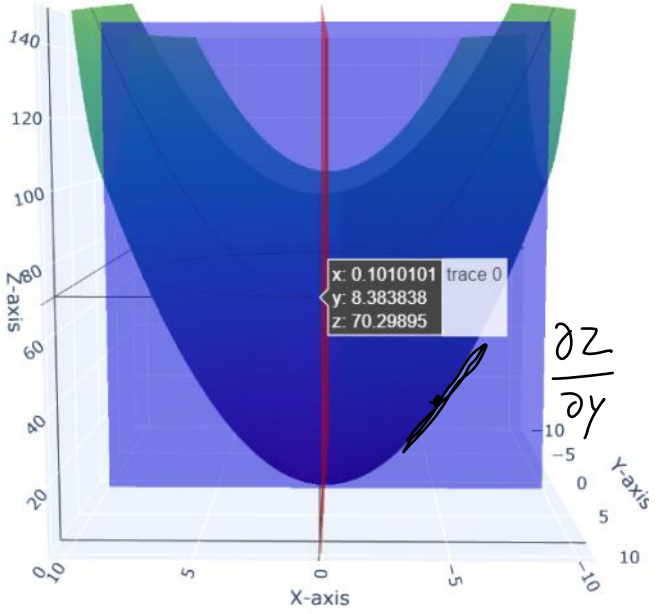
$\frac{\partial z}{\partial x}$ $\frac{\partial z}{\partial y}$ ← partial derivative

(1,2)

$x = 1$ $y = 2$

β_0 β_1





Higher Order Derivatives

22 April 2023 14:01

$y = x^3$ $\quad \frac{d}{dx}\left(\frac{dy}{dx}\right) = 3x^2$ $\quad = \boxed{6x}$ $\quad \frac{d^2y}{dx^2} \leftarrow \frac{d}{dx}\left(\frac{dy}{dx}\right)$

$s \rightarrow v \rightarrow a$

$\frac{d^3y}{dx^3}$

instant rate of change of slope

slope = 0 $\quad \frac{dy}{dx} \rightarrow$ maxima / minima

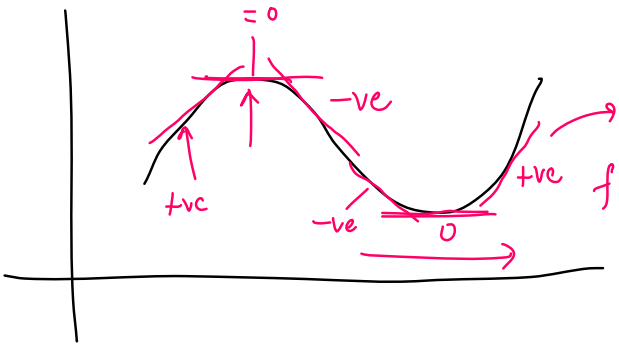
derivative

$\frac{d^2y}{dx^2}$ maxima

$\boxed{\frac{d^2y}{dx^2}} \rightarrow$ rate of change of slope

$\boxed{\frac{dy}{dx}} \leftarrow$ 1st order $\rightarrow$ Newton-hessian matrix

$\boxed{-ve}$ maxima $\boxed{-ve}$



$f''(x) \geq 0$

$\frac{d^2L}{dx^2}$

$f''(x)$

$\boxed{f''(x)} >$

-ve mov

$f''(x)$

$f'(x)$

$\frac{d}{dx}(Ax)$ ← $\frac{d}{dx}(Cx) = C$ (constant) $\quad \frac{d}{dx}5x = 5$

$\frac{d}{dx}Ax \qquad A^T = A$

matrix

$\frac{d}{dx}Ax \rightarrow A$

$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$ $f_1(x_1, x_2)$, $f_2(x_1, x_2)$

$\rightarrow \begin{bmatrix} \frac{df_1}{dx_1} & \frac{df_1}{dx_2} \\ \frac{df_2}{dx_1} & \frac{df_2}{dx_2} \end{bmatrix}$

$\rightarrow \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \rightarrow A \qquad \frac{d}{dx}Ax = A$

$y = x^T A x \qquad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad A^T = A \qquad \begin{bmatrix} a_{11} & a_x \\ a_x & a_{22} \end{bmatrix}^T$

$\frac{dy}{dx} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_x \\ a_x & a_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\qquad \begin{bmatrix} a_{11} & a_x \\ a_x & a_{22} \end{bmatrix}$ ←

$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11}x_1 + a_x x_2 \\ a_x x_1 + a_{22}x_2 \end{bmatrix}$ $f(x_1, x_2)$

$1 \times 2 \qquad 2 \times 1$

$a_{11}x_1^2 + a_x x_2^2 + a_x x_1^2 + a_{11}x_2^2$

$\begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \end{bmatrix}$

$\begin{bmatrix} 2a_{11}x_1 + 2a_x x_1 \\ 2a_x x_2 + 2a_{22}x_2 \end{bmatrix} = 2 \begin{bmatrix} a_{11}x_1 + a_x x_1 \\ a_x x_2 + a_{22}x_2 \end{bmatrix}$ ←

↓

$2 \begin{bmatrix} a_{11} & a_x \\ a_x & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

{matrix calculus}

$\rightarrow (2Ax)^T \qquad 2x^T A^T = 2x^T A$

$\rightarrow 2 \times 2$

$\rightarrow 4x$ ←