Lesson A 03:

Multivariate Calculus.

Yu Liang

Topic 1: Definition of Derivative

Topic 2: Calculus Rules.

- sum Rule
- power Rule
- product Rule
- Chain Rule
- Total Derivative

Topic 3: Derivative of Named function.

- $1/x$, $\sin(x)$, $\cos(x)$, e^x.

Topic 4: Derivative structure. ∇, ∇^2.

Topic 5: Taylor Series.

Topic 6: Optimization.

Definition of a Derivative.

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

e.g: $\sin'(x) = \frac{d \sin(x)}{dx} = \lim_{\Delta x \to 0} \frac{\sin(x+\Delta x) - \sin(x)}{\Delta x}$

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

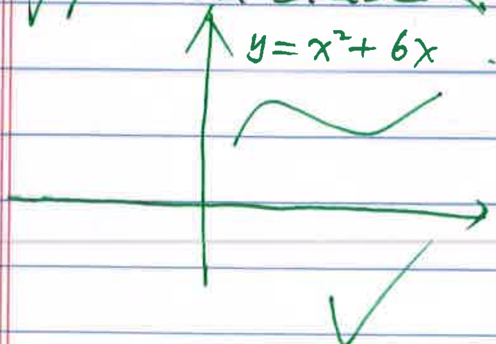
$$= \lim_{\Delta x \to 0} \cos(x) \cdot \frac{\sin(\Delta x)}{\Delta x}$$

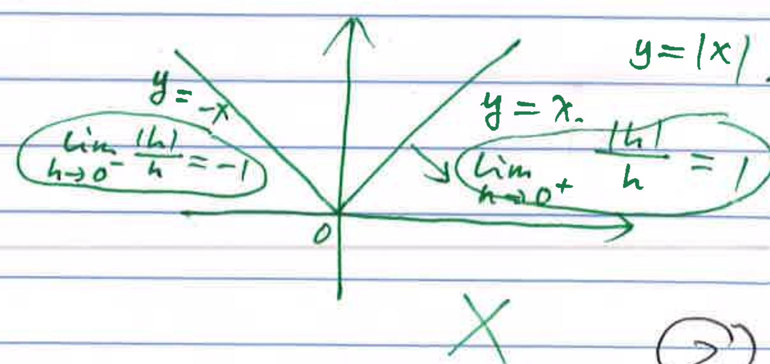
$$= \cos x$$

e.g: $\frac{d}{dx}\left(\frac{1}{x}\right) = \lim_{h \to 0} \frac{\left[\frac{1}{x+h} - \frac{1}{x}\right]}{h} = \lim_{h \to 0}\left[\frac{-h}{h(x^2+xh)}\right]$

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

Differentiable $\rightleftarrows$ derivative exist.





Rules of Calculus.

① Sum Rule.

$$\frac{d}{dx}\left(f(x)+g(x)\right) = \frac{df(x)}{dx} + \frac{dg(x)}{dx}$$

② Power Rule.

$$\frac{d}{dx}\left(ax^{b}\right) = ab\,x^{b-1}$$

③ Product Rule.

$$\frac{d}{dx}\left(f(x)\cdot g(x)\right) = f'(x)\cdot g(x) + f(x)\cdot g'(x).$$

④ Chain Rule.

$$z = f(y), \quad y = g(x).$$

$$\frac{dz}{dx} = \frac{dz}{dy}\cdot\frac{dy}{dx}$$

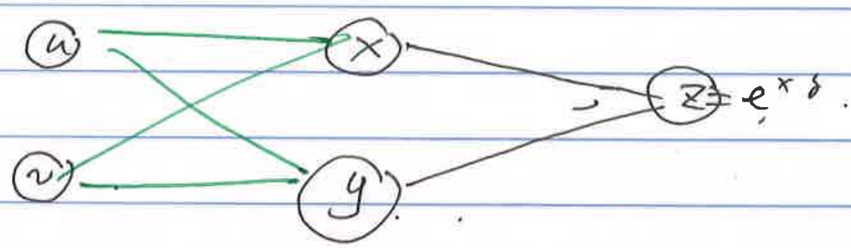
Application of Chain Rule.

① $f(x) = \ln(\underbrace{x^2-1}_{g(x)})$.

solution: $\dfrac{df(g)}{dx} = \dfrac{df(g)}{d(g)} \cdot \dfrac{d(g)}{dx}$

$= \dfrac{1}{g(x)} \cdot \dfrac{d(x^2-1)}{dx} = \dfrac{2x}{x^2-1}$

② $z = e^{xy}$. $x = 2u+v$. $y = \dfrac{u}{v}$.



Solution: $\dfrac{\partial z}{\partial u} = \underbrace{\dfrac{\partial z}{\partial x}}_{e^{xy}\cdot y} \cdot \underbrace{\dfrac{\partial x}{\partial u}}_{2} + \underbrace{\dfrac{\partial z}{\partial y}}_{e^{xy}\cdot x} \underbrace{\dfrac{\partial y}{\partial u}}_{1/v}$.

$\dfrac{\partial z}{\partial v}$. ?

Example of Derivative of Sigmoid func.

$$\sigma(x) = \frac{1}{1+e^{-z}}.$$

$$\frac{d\sigma(x)}{dx} = \frac{-1}{(1+e^{-z})^2} \cdot \frac{d(1+e^{-z})}{dz}$$

$$= \frac{e^{-z}}{(1+e^{-z})^2}$$

$$= \frac{1+e^{-z} - 1}{(1+e^{-z})^2}$$

$$= \sigma(x)\,(1-\sigma(x)).$$

Exercise:

$$\sigma(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

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

Derivative Structure

⓪ Jacobian J of function $f: R^n \to R^m$: $J(f) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$

① Given function : $f(\vec{x}): R^n \to R$

$$\nabla f(\vec{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Gradient

Hessian.

$$H(\vec{x}) = \nabla^2 f(\vec{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ & \ddots & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

② Taylor series

univariate $(f: R \to R)$

$$f(x) = f(c) + f'(c) \cdot (x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots$$

$$= \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n \quad (\text{e.g. } e^x = 1 + x + \frac{x^2}{2} \cdots)$$

multivariate: $(f: R^n \to R)$

$$f(x) = f(c) + \nabla f(\vec{x})^T (\vec{x} - \vec{c}) + \frac{1}{2!}(\vec{x} - \vec{c})^T H(\vec{c})(\vec{x} - \vec{c})$$

$$y = f(\vec{x})$$

Optimization and Vector Calculus

Analytical Solution ✓	Iterative Numerical Solution
	gradient descent (min)
	$\vec{x}^{(n+1)} = \vec{x}^{(n)} - \eta \cdot \nabla f(\vec{x}^n)$
$\nabla f(x) = 0$	
	gradient ascent (max)
	$\vec{x}^{(n+1)} = \vec{x}^{(n)} + \eta \cdot \nabla f(\vec{x}^n)$
e.g.: $\begin{cases} y = x^2 + 1 \\ \underset{x \in R}{\operatorname{argmin}}\ y \end{cases}$	η: Learn Rate
	where
	$\nabla f(\vec{x}^n)$ = gradient
$\begin{cases} z = (x_1+1)^2 + (x_2-2)^2 + 5 \\ \underset{(x_1, x_2) \in R^2}{\operatorname{argmin}}\ z \end{cases}$	