* Differential Calulus

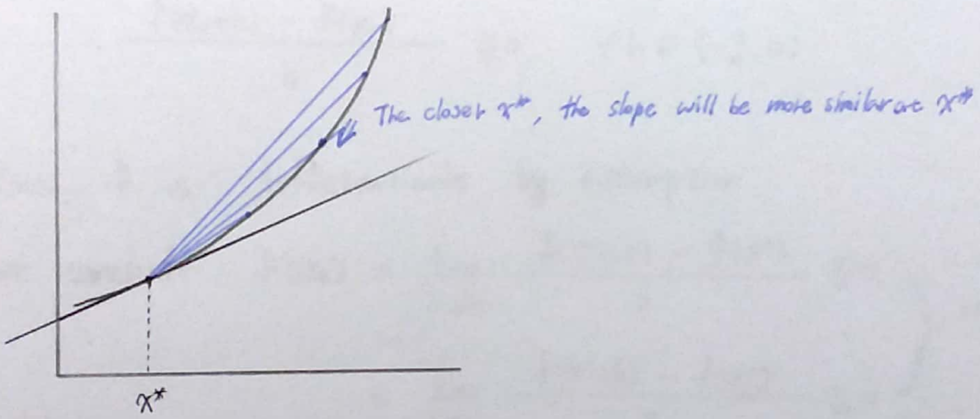
Recap for $\mathbb{R}$

Def : Let $(a,b) \subset \mathbb{R}$ and let $f : (a,b) \to \mathbb{R}$

- f is differentiable at $x \in (a,b)$

  if $f'(x) := \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$

- f is differentiable if it is differentiable at each point of its domain.



Thm 32 Let $f : (a,b) \to \mathbb{R}$
If f is differentiable at $x \in (a,b)$, then f is continuous at x.

proof) $\lim_{h\to 0} f(x+h) - f(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \cdot \lim_{h\to 0} h = f'(x) \cdot 0 = 0$

By thm 11, f is continuous at x.

Def Let (X,d) be a metric space. and $f : X \to \mathbb{R}$

$x_0 \in X$ is a local maximizer (minimizer) of f

if $\exists\, \delta > 0$ s.t. $\forall x \in B_\delta(x_0)$, $f(x) \leq f(x_0)$ $(f(x) \geq f(x_0))$

Thm 33 Let $f : (a,b) \to \mathbb{R}$ be a differentiable function.

and $x_0 \in (a,b)$ be a local maximizer (minimizer) of f.

Then, $f'(x_0) = 0$.

proof) (Maximizer case)

Suppose x_0 is a local maximizer.

Then $f(x_0+h) - f(x_0) \leq 0$, $\forall h \in \mathbb{R}$, satisfying $|h| < \delta$.

Hence, $\dfrac{f(x_0+h) - f(x_0)}{h} \leq 0$, $\forall h \in (0, \delta)$

$\dfrac{f(x_0+h) - f(x_0)}{h} \geq 0$, $\forall h \in (-\delta, 0)$

Since f is differentiable by assumption,

we conclude $f'(x_0) = \lim\limits_{\substack{h \downarrow 0 \\ h \to 0^+}} \dfrac{f(x_0+h) - f(x_0)}{h} \leq 0$

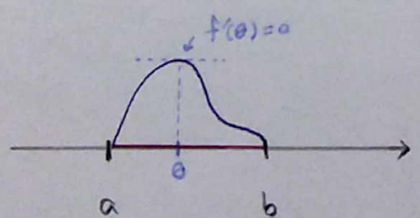
$= \lim\limits_{\substack{h \uparrow 0 \\ h \to 0^-}} \dfrac{f(x_0+h) - f(x_0)}{h} \geq 0$, which implies that $f'(x_0) = 0$.

//

Thm 34 (Rolle's theorem)

Let $f : [a,b] \to \mathbb{R}$ be continuous and differentiable (a,b)

If $f(a) = f(b) = 0$, then $\exists\ \theta \in (a,b)$ s.t. $f'(\theta) = 0$.



proof) Since f is continuous and $[a,b]$ is compact,

Weierstrass' theorem implies that $\exists$ a local maximizer $x_M \in [a,b]$ and a local minizer $x_m \in [a,b]$.

- If $f(x_M) = f(x_m) = 0$, then $f(x) = 0$, $\forall x \in [a,b]$.

  This implies $f'(x) = 0$, $\forall x \in [a,b]$

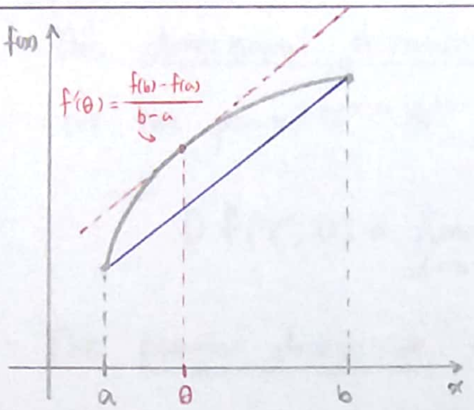
- If $f(x_M) > 0$, then $x_M \in (a,b)$. Since x_M is a local maximizer, we conclude $f'(x_M) = 0$ (by thm33)

Similar argument for $f(x_m) < 0$. //

Thm 35 ⟨ Mean Value theorem ⟩

Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function.

If $a, b \in \mathbb{R}$ satisfies $a<b$, then $\exists\ \theta \in (a,b)$ s.t. $f'(\theta) = \dfrac{f(b)-f(a)}{b-a}$ (slope)



proof) Define $\phi : [a,b] \longrightarrow \mathbb{R}$ by

$$\phi(x) = -\frac{f(b)-f(a)}{b-a}(x-a) + f(x) - f(a)$$

ϕ satisfies all assumption of Rolle's theorem.

Hence, $\exists\ \theta \in (a,b)$ such that

$$0 = \phi'(\theta) = -\frac{f(b)-f(a)}{b-a} + f'(\theta) \quad //$$

Thm 36 ⟨ Inverse function theorem on $\mathbb{R}$ ⟩

Let $f : \mathbb{R} \to \mathbb{R}$ be continuously differentiable with $f'(a) \neq 0$.

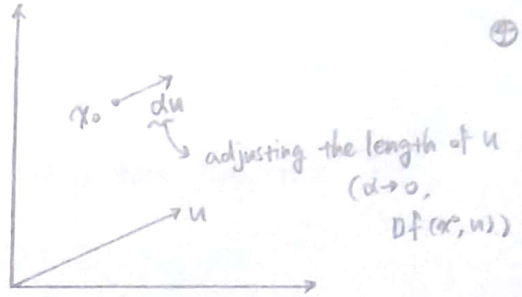
Then f is invertible in a neighborhood of a.

The inverse is countably differentiable and $\boxed{(f^{-1}(\cdot))' = \dfrac{1}{f'(a)}}$

$\Rightarrow$ This will be generalized in higher dimensions.

Partial and directional derivatives

(How do we get from $\mathbb{R}$ to $\mathbb{R}^n$?)



Def Let $X \subseteq \mathbb{R}^n$ be open.

- The <u>directional derivative</u> of $f: X \to \mathbb{R}$ in the direction $u \in \mathbb{R}^n$ at the point $x^\circ \in \mathbb{R}^n$ is defined by

$$Df(x^\circ, u) = \lim_{\alpha \to 0} \frac{f(x^\circ + \alpha u) - f(x^\circ)}{\alpha}$$, whenever the limit exists and is finite.

- The <u>partial derivative</u> of $f: X \to \mathbb{R}$ with respect to its i-th argument x_i at a point $x^\circ = (x_i^\circ, x_{-i}^\circ)$ is defined as

$$D_{x_i} f(x^\circ) = \lim_{\alpha \to 0} \frac{f(x_i^\circ + \alpha, x_{-i}^\circ) - f(x^\circ)}{\alpha}$$, whenever the limit exists and is finite.

Equivalently, $D_{x_i} f(x^\circ) = Df(x^\circ, e_i)$ ← e_i unit vector $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ ← ith

- The notations $D_i f(x^\circ)$, $f_i(x^\circ)$ or $\frac{\partial}{\partial x_i} f(x^\circ)$ are synonyms

- The <u>Gradient</u> of $f: X \to \mathbb{R}$ at x° is given by

$$\nabla f(x^\circ) = (f_1(x^\circ), f_2(x^\circ), \dots, f_n(x^\circ))$$

- The <u>partial derivative of the function $f_i(x)$</u> with respect to the K-th argument is defined by $f_{iK}(x)$ or $\frac{\partial}{\partial x_K} f_i(x)$

- For 1-dimension (one) functions differentiablity in X implies continuity in X
  This is not true for higher-dimensional functions.

Ex: $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \begin{cases} 1 & \text{if } x^2 = y,\ x \neq 0 \\ 0 & \text{else} \end{cases}$ at $(0,0)$ all directional derivatives exist and are 0.

Differentiability

If $f: \mathbb{R}^n \to \mathbb{R}^m$, we treat f as a vector of component functions f^i

$$f = \begin{pmatrix} f^1 \\ \vdots \\ f^m \end{pmatrix}, \quad \text{where } f^i : \mathbb{R}^n \to \mathbb{R}.$$

" differentiability $\approx$ approaching locally by linear function. "

In $\mathbb{R}$: f is differentiable in x^0 if $\exists\, a \in \mathbb{R}$ s.t.

$$\lim_{h \to 0} \frac{f(x^0+h) - [f(x^0) + ah]}{h} = 0$$

Def Let $X \subseteq \mathbb{R}^n$ be open and let $f : X \to \mathbb{R}^m$

- f is differentiable at a point $x \in X$ if $\exists\, A_x \in \mathbb{R}^{m \times n}$ s.t

$$\lim_{h \to 0} \frac{\| f(x+h) - f(x) - A_x h \|}{\|h\|} = 0 \qquad \underset{n\times 1}{h} = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \\ \vdots \\ \frac{n}{k} \end{pmatrix} \Rightarrow \lim_{k \to \infty} h = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- f is differentiable if it is differentiable in every point in its domain.
- The derivative of f at x is given by $Df : X \to \mathbb{R}^{m \times n}$ satisfying $x \longmapsto A_x$
- The derivative of f at x, df_x is given by the linear mapping induced by $Df(x)$; $df_x : \mathbb{R}^m \to \mathbb{R}^n$, $h \longmapsto A_x h = Df(x) \cdot h$

Thm 39 Let $X \subseteq \mathbb{R}^n$ be open and $f : X \to \mathbb{R}^m$ be differentiable at $x \in X$.
Then, f is continuous at x

Sketch) $\lim_{h\to 0} \| f(x+h) - f(x) \| \leq \lim_{h\to 0} (\| f(x+h) - f(x) - A_x h \| + \| A_x h \|)$

$$= \lim_{h\to 0} \frac{\| f(x+h) - f(x) - A_x h \|}{\|h\|} \cdot \lim_{h\to 0} \|h\| + \lim_{h\to 0} \| A_x h \| = 0 \quad \|$$

(Thm 37 : Taylor
Thm 38 : $f_{ik} = f_{ki}$)

Thm 40 Let $X \subseteq \mathbb{R}^n$ be open and $f : X \to \mathbb{R}^m$.
f is differentiable iff each of component functions is differentiable at x.
Moreover, if f is differentiable at x, then the partial derivative of the component functions exist at x, and the derivative of f at x is the matrix

$$Df(x) = \begin{bmatrix} Df^1(x) \\ \vdots \\ Df^m(x) \end{bmatrix} = \begin{bmatrix} \nabla f^1(x) \\ \vdots \\ \nabla f^m(x) \end{bmatrix} = \begin{bmatrix} f^1_1(x) & \cdots & f^1_n(x) \\ \vdots & \ddots & \vdots \\ f^m_1(x) & \cdots & f^m_n(x) \end{bmatrix}$$

$f^j_i(x)$ ← j-th component, ← i-th partial

We call this matrix of partial derivatives of component functions the <u>Jacobian</u> of f at x.

Thm 41 Let $X \subseteq \mathbb{R}^n$ be open and $f : X \to \mathbb{R}^m$
If the partial derivative of the component functions exist and are continuous on X,
Then, f is differentiable on X.

proof) We show that f^j is differentiable (j-th component function)
Fix some $x \in X$ and $\varepsilon > 0$.
Since X is open and the partial derivatives are continuous,
$\exists\, r > 0$ s.t. $B_r(x) \subseteq X$ and $\forall x+h \in B_r(x)$
and $i = 1, \dots, n$ $\underline{|f^j_i(x+h) - f^j_i(x)| < \frac{\varepsilon}{n}}$ $\cdots (*)$

↓

WTS: $$\lim_{h\to 0} \frac{|f_i^j(x+h) - f_i^j(x) - \nabla f^j(x)\cdot h|}{\|h\|} = 0$$

Suppose $\|h\| < r$. and define $V_0 = \underline{0}$, $V_1 = (h_1, 0, \ldots, 0)$

$V_2 = (h_1, h_2, 0, \ldots, 0)$

$\vdots$

$V_k = (h_1, h_2, \ldots, h_k, 0, \ldots, 0)$

Then, $\underline{f^j(x+h) - f^j(x) = \sum_{i=1}^{n} [f^j(x+V_i) - f^j(x+V_{i-1})]}$ $\cdots$ (**)

Note that $\|V_i\| \leq \|h\| < r$ by our construction.

$f'(\theta) = \frac{f(b)-f(a)}{b-a} \Rightarrow$ Imagine b, a in here

$b = x+V_i$
$a = x+V_{i-1}$

The one-dimensional mean value theorem implies for $i=1,\ldots,n$ that there exists $\theta_i \in (0,1)$ s.t. $\underline{f^j(x+V_i) - f^j(x+V_{i-1}) = h_i \cdot f_i^j(\underbrace{x+V_{i-1}}_{a} + \underbrace{\theta_i h_i e_i}_{\text{direction vector}})}$ $\cdots$ (***)

Putting everything (*), (**), (***) together, we get

$$|f_i^j(x+h) - f_i^j(x) - \nabla f^j(x)\cdot h| \overset{(**)}{=} \left|\sum_{i=1}^{n}[f^j(x+V_i) - f^j(x+V_{i-1})] - \sum_{i=1}^{n} h_i \cdot f_i^j(x)\right|$$

$$\overset{(***)}{=} \left|\sum_{i=1}^{n} h_i [f_i^j(x+V_{i-1} + \theta_i h_i e_i) - f_i^j(x)]\right|$$

"Cauch-Schwartz"

$$\leq \sum_{i=1}^{n} |h_i| \cdot |f_i^j(x+V_{i-1} + \theta_i h_i e_i) - f_i^j(x)|$$

$$\overset{(*)}{<} \sum_{i=1}^{n} |h_i| \cdot \frac{\varepsilon}{n} \leq \varepsilon \cdot \|h\|$$

We conclude that $\forall h \neq \underline{0}$ satisfying $\|h\| < r$,

$$\frac{|f^j(x+h) - f^j(x) - \nabla f^j(x)\cdot h|}{\|h\|} < \varepsilon.$$

Hence, f^j is differentiable $j=1,\ldots,m$.

Then, Thm 40 implies that f is differentiable. //