Thm 42 < Chain-rule >

Let $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ be open

$f : X \to \mathbb{R}^m$ and $g : Y \to \mathbb{R}^p$, $f(X) \subseteq Y$

Let x° be a point in X $y^\circ = f(x^\circ)$ and Define the composite function

$F = g \circ f$ by $F(x) = g(f(x))$, $\forall x \in X$

If f is differentiable at x° and g is differentiable at y°,

then F is differentiable at x° and $DF(x^\circ) = Dg(y^\circ) \cdot Df(x^\circ)$

Thm 43 < Mean Value Theorem >

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

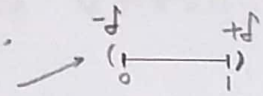
Let $x, y \in X$ be such that the line segment between x and y.

$$l(x,y) := \{ z \in \mathbb{R}^n \mid \exists \lambda \in [0,1] : z = \lambda x + (1-\lambda) y \} \text{ is contained in } X$$

Then, for each vector $a \in \mathbb{R}^m$, $\exists$ a vector $z \in l(x,y)$ such that

$$a \cdot [f(y) - f(x)] = a Df(z)(y-x)$$

proof) Let $h = y - x$. As X is open and contains $l(x,y)$,



$\exists \delta > 0$ such that $x + \lambda h \in X$ for $\lambda \in (-\delta, 1+\delta)$

Fix an arbitrary $a \in \mathbb{R}^m$ and define the real valued function ϕ_a

on the interval $(-\delta, 1+\delta)$ by $\phi_a(\lambda) = a \circ f(x + \lambda h) = \sum_{i=1}^{m} a_i f^i(x+\lambda h)$

By construction, ϕ_a is differentiable on $(-\delta, 1+\delta)$ and $\phi_a'(\lambda) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i f_j^i(x+\lambda h) \cdot h_j$

$= a \cdot Df(x + \lambda h) \cdot h$

The mean value theorem for univariate functions applied to ϕ_a implies that

$\exists \lambda^\circ \in (0,1)$ such that $\phi_a(1) - \phi_a(0) = \phi_a'(\lambda^\circ)$

Setting $z = x + \lambda^\circ h$, this is equivalent to $a[f(y) - f(x)] = a \cdot Df(z)(y-x)$. //

Thm 44

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

Let $x, y \in X$ be such $\ell(x, y) \subseteq X$.

Then $\exists$ a vector $z \in \ell(x, y)$ such that $\|f(y) - f(x)\| \leq \|Df(z)(y-x)\| \leq \|Df(z)\|_L \|y-x\|$

proof) ① $f(y) - f(x) = \underline{0}$, we are done.

② $f(y) - f(x) \neq \underline{0}$

By Theorem 43, $\forall a \in \mathbb{R}^m$, $\exists z \in \mathbb{R}^n$ such that $a \cdot [f(y) - f(x)] = a[Df(z)(y-z)]$

By Cauchy-Schwarz, $\underbrace{|a[f(y) - f(x)]|}_{\text{This is scalar.}} \leq \|a\| \|Df(z)(y-z)\|$

Setting $a = \dfrac{f(y) - f(x)}{\|f(y) - f(x)\|}$

$$\frac{\|f(y) - f(x)\|^2}{\|f(y) - f(x)\|} \leq \|Df(z)(y-x)\| \leq \|Df(x)\|_L \cdot \|y - x\|. \ \blacksquare$$

Thm 45

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

if $x^\circ \in X$ is a local maximizer (minimizer) of f, then $\nabla f(x^\circ) = \underline{0}$

proof) Suppose x° is local maximizer

$\forall t \in \mathbb{R}$ small enough, we have $f(x^\circ) \geq f(x + t \nabla f(x^\circ))$

Define $g : \mathbb{R} \to \mathbb{R}$ by $g(t) = f(x^\circ + t \nabla f(x^\circ))$ $\quad t \to 0,\ g(0) = f(x^\circ)$

$t = 0$ is a maximizer of g and the corresponding one-dimension result yields

$0 = g'(0) = \nabla f(x^\circ) \cdot \nabla f(x^\circ) = \|\nabla f(x^\circ)\|^2$

Hence $\nabla f(x^\circ) = \underline{0}$. //

Continuously differentiable

Ex) Derivatives needs not to be continuous.

Let $f: \mathbb{R} \to \mathbb{R}$ defined as $f(x) = \begin{cases} x^2 \sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

f is differentiable $\forall x \neq 0$. $f'(x) = 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x})$

However, $f'(0) = \lim_{h\to 0} \dfrac{h^2 \sin(\frac{1}{h})}{h} = 0$

$\Rightarrow$ f' is not continuous since $\limsup_{x\to 0} f'(x) = \limsup_{x\to 0} \cos(\frac{1}{x}) = 1$

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

The function $f: X \longrightarrow \mathbb{R}^m$ is continuously differentiable on X if it is differentiable on X and the derivative Df is a continuous function from X to $L(\mathbb{R}^n, \mathbb{R}^m)$

The function $f: X \longrightarrow \mathbb{R}^m$ is of class $\mathcal{C}^k$ or $f \in \mathcal{C}^k(X)$, if the first k partial derivatives exist and are continuous on X

$\boxed{\text{Note}}$ The function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is continuously differentiable if and only if it is of class $\mathcal{C}^1$

Inverse and Implicit function theorems

Thm 46 < Inverse function theorem >

Let $X \subseteq \mathbb{R}^n$ be open and $f: X \to \mathbb{R}^n$ be continuously differentiable.

Suppose $Df(x^\circ)$ is invertible for some $x^\circ \in X$.

Then there is an open set V containing x° and an open set W containing $f(x^\circ)$

such that the inverse relation $f^{-1}: W \to V$ is a well-defined function.

Moreover, f^{-1} is continuously differentiable and $\forall y \in W$, its derivative is

given by

$$Df^{-1}(y) = [Df(f^{-1}(y))]^{-1}$$

proof)

Fix $x^\circ \in X$ and let $A := Df(x^\circ)$

By assumption, A is invertible and we define $\lambda := \frac{1}{2\|A^{-1}\|_L}$

Since X is open and f is continuously differentiable,

we can choose a $\delta > 0$ s.t. setting $U := B_\delta(x^\circ)$, $\bar{U} \subseteq X$ and

Closure of U

$$\|Df(x) - A\|_L \leq \lambda, \ \forall x \in \bar{U}. \quad (1)$$

Since A is invertible and $\forall x \in \bar{U}$, $\|Df(x) - A\| \leq \frac{1}{2\|A^{-1}\|_L}$

We can conclude by thm 30 that $Df(x)$ is invertible $\forall x \in \bar{U}$.

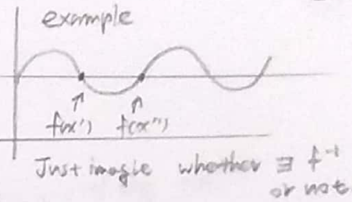
Claim 1 : f is one-to-one on $\bar{U}$

We define $\forall y \in \mathbb{R}^n$, $\phi_y(x) := x + A^{-1}(y - f(x)) \quad (2)$

This yields

$$D\phi_y(x) = I - A^{-1} \cdot Df(x) = A^{-1} \cdot [A - Df(x)]$$

Hence, for all $x \in \bar{U}$, $\underset{\substack{\uparrow \\ \text{By thm 27}}}{\|D\phi_y(x)\|_L \leq} \|A^{-1}\|_L \cdot \|A - Df(x)\|_L \underset{\substack{\uparrow \\ \text{By (1)}}}{\leq} \frac{1}{2}$



Together with theorem 44, this implies

$\forall x', x'' \in \bar{U}$, $\quad \|\phi_y(x') - \phi_y(x'')\| \leqq \frac{1}{2}\|x' - x''\|$

Suppose $\exists x', x'' \in \bar{U}$ such that $f(x') = f(x'') = y$ for some $y \in \mathbb{R}^n$

Then, $\|\phi_y(x') - \phi_y(x'')\| = \|x' - x''\| \leqq \frac{1}{2}\|x' - x''\|$ and hence $x' = x''$

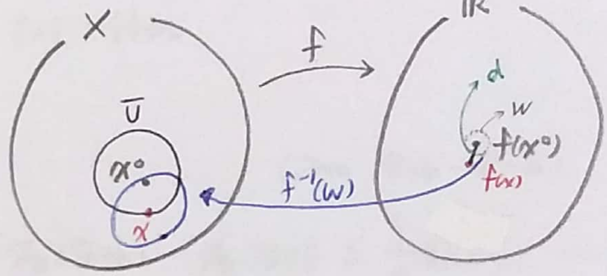
Therefore, f is one-to-one on $\bar{U}$.

Note that ∂U is closed and bounded $\Rightarrow$ therefore, compact (Heine Borel)

Since f and $\|\cdot\|$ are continuous, the Extreme Value theorem (Weierstrass)

implies that $d := \min_{x \in \partial U} \|f(x) - f(x^\circ)\|$ is well-defined.

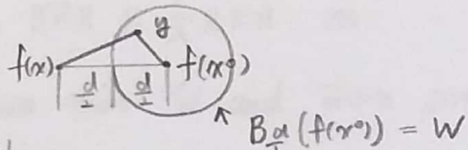
Since f is one-to-one on $\bar{U}$, $d > 0$



We define $W := \{y \in \mathbb{R}^n \mid \|y - f(x^\circ)\| < \frac{d}{2}\}$

By triangle inequality, $\forall y \in W$ and $\forall x \in \partial U$

$$\|y - f(x^\circ)\| < \frac{d}{2} \leqq \|y - f(x)\| \quad (3)$$



We set $V := U \cap f^{-1}(W)$ and note that

V is open and non-empty, since it contains x°

<u>Claim 2 : f is onto</u>

Fix arbitrary $y \in W$, we are going to show that $y \in f(V)$

Trick: Define $h : \bar{U} \to \mathbb{R}$ by

$$h(x) := \|y - f(x)\|^2 = (y - f(x)) \cdot (y - f(x))$$

By the extreme value theorem, h attains its minimum on $\bar{U}$

The minimum is not attained on ∂U, because by (3) already

x° attains a lower value than any $x \in \partial U$

Hence, the minimum is attained for some $\bar{x} \in U$.

By thm 45, we get $\nabla h(\bar{x}) = \underline{0}$ or $Df(\bar{x}) \cdot (y - f(\bar{x})) = 0$

As observed above $Df(\bar{x})$ is invertible, hence $\ker(Df(\bar{x})) = \{\underline{0}\}$ (By this, $y = f(\bar{x})$ b/c $Df(\bar{x}) \neq 0$)

Hence $y = f(\bar{x})$. Clearly, $\bar{x} \in f^{-1}(W)$ and hence $\bar{x} \in V$

Claim 3: f^{-1} is continuously differentiable and $Df^{-1}(y) = [Df(f^{-1}(y))]^{-1}$

We are going to show that
$$\lim_{k \to \underline{0}} \frac{\| f^{-1}(y+k) - f^{-1}(y) - [Df(f^{-1}(y))]^{-1} k \|}{\|k\|} = 0$$

Take arbitrary $y,\ y+k \in W$. Then $\exists\ \tilde{x},\ \tilde{x}+k \in V$ such that

$f(\tilde{x}) = y$ and $f(\tilde{x}+k) = y+k$.

Fixing an arbitrary $z \in W$, we get from (2) that

$$\phi_z(\tilde{x}+h) - \phi_z(\tilde{x}) = h - A^{-1}k$$

(Since $\tilde{x}+h-\tilde{x} = h$)

Moreover. $\|h\| - \|A^{-1}k\| \leq \|h - A^{-1}k\| = \|\phi_z(\tilde{x}+h) - \phi_z(\tilde{x})\| \leq \frac{1}{2}\|h\|$

This implies $\|h\| \leq 2\|A^{-1}k\| \leq 2\|A^{-1}\|_L \cdot \|k\| \leq \frac{1}{\lambda}\|k\|$ (4)

This implies that if y and $y+k$ are close that $\tilde{x}$ and $\tilde{x}+h$ are close

and hence f^{-1} is continuous.

Set $T = [Df(\tilde{x})]^{-1}$. Then $f^{-1}(y+k) - f^{-1}(y) - Tk = TT^{-1}h - Tk$

$$= \underset{\substack{\uparrow \\ T^{-1}}}{I}h - Tk = T[Df(\tilde{x}) \cdot h - \underbrace{(y+k-y)}_{f(\tilde{x}+h) - f(\tilde{x})}] = -T[f(\tilde{x}+h) - f(\tilde{x}) - Df(\tilde{x}) \cdot h]$$

$k \geq \lambda\|h\|$

By (4), this implies
$$0 \leq \frac{\|f^{-1}(y+k) - f^{-1}(y) - Tk\|}{\|k\|} \leq \|T\|_L \frac{\|f(\tilde{x}+h) - f(\tilde{x}) - Df(\tilde{x}) \cdot h\|}{\lambda\|h\|} \quad (5)$$

Inequality (4) also implies that $h \to \underline{0}$ as $k \to \underline{0}$. Since f is differentiable and $\|T\|_L$ is a constant, the right-hand-side of (5) goes to 0 as $k \to \underline{0}$

We conclude that $\lim_{k \to \underline{0}} \frac{\| f^{-1}(y+k) - f^{-1}(y) - Tk \|}{\| k \|} = 0$

and hence f^{-1} is differentiable on W and $Df^{-1}(y) = (Df(f^{-1}(y)))^{-1}$

$(Df(f^{-1}(y)))^{-1}$ is therefore the composition of three differentiable functions

f^{-1} is differentiable and hence continuous. $Df(x)$ is continuous by assumption

and $(\cdot)^{-1}$ is continuous by thm 31. $\blacksquare$

↑

the function that maps each invertible linear operator into its inverse.