continuous functions on metric spaces (last review of topology)

Def $f: X \to Y$ between MS is $\boxed{\text{ctn. at } x_0} \in X$ if
$\forall \varepsilon > 0\ \exists \delta > 0$ st. $d_X(x, x_0) < \delta \Rightarrow d_Y(f(x), f(x_0)) < \varepsilon$

i.e. $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$

ex

(1) f is ctn iff $\forall$ open $B \in Y$, $f^{-1}(B)$ is open in X.

(2) Let X be cpt. Y be any MS

$\Rightarrow f(X)$ is cpt.

(ctn function preserve compactness in MS.)

(c) $f: X \to \mathbb{R}$

if X cpt 且 f ctn $\Rightarrow$ f 有 max & min.

Def $f: X \to Y$ is $\boxed{\text{uniformly ctn.}}$ if
$\forall \varepsilon > 0\ \exists \delta > 0$ st. $d_X(x_1, x_2) < \delta \Rightarrow d_Y(f(x_1), f(x_2)) < \varepsilon$

Thm if $f: X \to Y$ ctn 且 X cpt.
$\Rightarrow f$ uni. ctn.

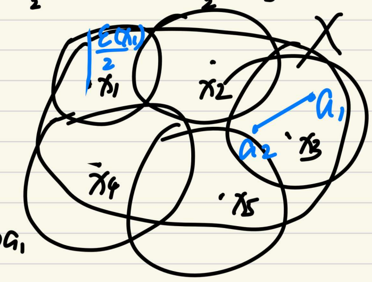
Pf let $\varepsilon > 0$

For each $x \in X$, can find $\delta(x) > 0$ st.
$f\left(B_{\frac{\delta(x)}{2}}(x_0)\right) \subseteq B_{\frac{\varepsilon}{2}}(f(x_0))$ by ctn.

So $\{B_{\delta(x)}(x_0)\}_{x \in X}$ is an open cover

$\Rightarrow \exists$ finite subcover $\{B_{\frac{\delta(x_1)}{2}}(x_0), \ldots, B_{\frac{\delta(x_k)}{2}}(x_k)\}$

let $\delta = \min_{i=1,\ldots,k} \frac{\delta(x_i)}{2}$



For each $a_1, a_2 \in X$ s.t.
$d(a_1, a_2) < \delta$
pick i st. $B_{\frac{\delta(x_i)}{2}}(x_i) \ni a_1$

Note: $d(a_2, x_i) \le d(a_2, a_1) + d(a_1, x_i)$
$\le \frac{\delta(x_i)}{2} + \frac{\delta(x_i)}{2} = \delta(x_i)$

So $d(f(a_1), f(a_2)) \le d(f(a_1), f(x_i)) + d(f(a_2), f(x_i))$
$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ □

ex $f: \mathbb{R} \to \mathbb{R}$
$x \mapsto x^2$ is not uni. ctn.

Differentiability

Def let $E \subseteq \mathbb{R}^n$ be open.
Say $f: E \to \mathbb{R}^m$ is differentiable at $x_0 \in E$
if $\exists$ a linear map $A: \mathbb{R}^n \to \mathbb{R}^m$ st.

$$\lim_{|h| \to 0} \frac{|f(x_0+h) - f(x_0) - Ah|}{|h|} = 0$$

Rmk we can use $|h| = \|h\|_2 = \sqrt{\sum_i h_i^2}$ here
But it does not matter if we use another norm.

Rmk Here $\lim_{|h| \to 0} G(h) = 0$ mean
$\forall \varepsilon > 0\ \exists \delta > 0$ st. $|h| < \delta \Rightarrow |G(h)| < \varepsilon$

ex when $n = m = 1$, A matrix is a #.

And we can get the usual def:

$$\lim_{t \to 0} \left| \frac{f(x_0+t) - f(x_0) - At}{t} \right| = 0 \Leftrightarrow \lim_{t \to 0} \left| \frac{f(x_0+t) - f(x_0)}{t} \right| = A$$

即 $f'(x_0)$

Rmk A is unique

Pf if A_1, A_2 both satisfies the def

$$\Rightarrow \frac{|A_1 h - A_2 h|}{|h|} \le \frac{|f(x_0+h) - f(x_0) - A_1 h|}{|h|} + \frac{|f(x_0+h) - f(x_0) - A_2 h|}{|h|}$$

So $\lim_{|h| \to 0} \frac{|A_1 h - A_2 h|}{|h|} = 0$

and **Rmk** $\boxed{\text{if } A \neq \text{zero matrix, 总是有 } \lim_{|h| \to 0} \frac{|Ah|}{|h|} \neq 0}$

$\Rightarrow A_1 = A_2$

Def Derivative

For $f: \mathbb{R}^n \to \mathbb{R}^m$ that is diff'ble at $x_0 \in \mathbb{R}^n$

称 A (s.t. $\lim_{|h| \to 0} \frac{|f(x_0+t) - f(x_0) - Ah|}{|h|} = 0$)

为 the derivative of f at x_0 并 denote A 为 $\underset{m \times n}{Df(x_0)}$

Rmk $Df(x_0)$ is the best linear approximation to the increment function $g: h \longmapsto f(x_0+h) - f(x_0)$

The remainder (error) is

$$r_{x_0}(h) = f(x_0+h) - f(x_0) - \underset{m\times n}{Df(x_0)} \cdot \underset{n\times 1}{h}$$

$r_{x_0}: \mathbb{R}^n \to \mathbb{R}^m$

因而 $r_{x_0}(h)$ 总是有 sublinear size (< linear)

即 $\left|\frac{r_{x_0}(h)}{h}\right| \to 0$ as $h \to 0$

或称 $r_{x_0}(h)$ is $o(|h|)$

Def Directional derivative

Let $A \subseteq \mathbb{R}^n$ be open

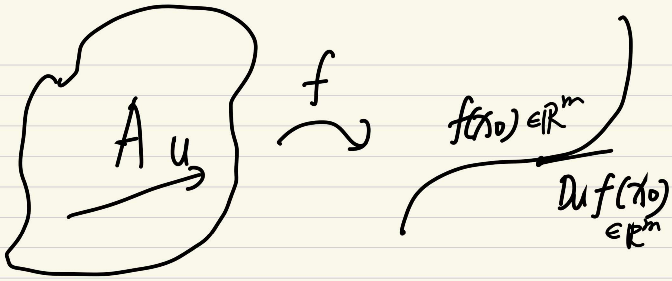
Suppose $x_0 \in A$ & $u \in \mathbb{R}^n$

We define the directional derivative $D_u f(x_0) \in \mathbb{R}^m$

as $\lim\limits_{t\to 0, t\in\mathbb{R}} \frac{f(x_0+tu) - f(x_0)}{t}$ (if it exists)

$D_u f: \mathbb{R}^n \to \mathbb{R}^m$

note: $f \in \mathbb{R}^{m\times n} = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$, $D_u f(x_0) = \begin{pmatrix} D_u f_1(x_0) \\ D_u f_2(x_0) \\ \vdots \\ D_u f_m(x_0) \end{pmatrix}$, where $D_u f_i(x_0) = \frac{d}{dt}(f_i(x_0) + tu)$



Note: $D_u f(x_0) = \frac{d}{dt}\Big|_{t=0} f(x_0 + tu)$

在 x_0 处，u方向上的微小变化量 $t\to 0$ ①

导致 f 的变化量 $f(x_0+tu) - f(x_0)$ ②

而 $D_u f(x_0)$ 则是二者比值

(u上微小变化量t引发的变化率)

ex $f: \mathbb{R}^2 \to \mathbb{R}$

$(x_1, x_2) \longmapsto \sin(x_1, x_2)$

$u = (1, 0)$

$$D_u f(x_1, x_2) = \frac{d}{dt}\Big|_{t=0} \sin((x_1+t)x_2) = x_2 \cos(x_1 x_2)$$

Thm Let $A \subseteq \mathbb{R}^n$ be open

$f: A \to \mathbb{R}^m$ be diff'ble at $x_0 \in A$

$\Rightarrow \forall u \in \mathbb{R}^n, \exists D_u f(x_0)$ 且 $D_u f(x_0) = \underset{\text{matrix}}{Df(x_0)} \cdot \underset{\text{vector}}{u}$

(In particular, $u \longmapsto D_u f(x_0)$ is linear)

Rmk

(1) $f \in \mathcal{F}(\mathbb{R}^n, \mathbb{R}^m)$

$Df(x_0) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \subseteq \mathcal{F}(\mathbb{R}^n, \mathbb{R}^m)$

note: 在一个 function space (of linear maps)

而特别地, $m=n=1$ 时 $\boxed{\mathrm{Hom}(\mathbb{R}, \mathbb{R}) = \mathbb{R}}$

因而 $f: A \to \mathbb{R}$ 在 $x_0 \in \mathbb{R}$ 的导数 $\frac{d}{dx}f(x_0)$

既是一个 linear map 又是一个 real number

而 $\frac{d}{dx}f: A \to \underbrace{\mathrm{Hom}(\mathbb{R}, \mathbb{R})}_{=\mathbb{R}}$

和 $f: A \to \mathbb{R}$ 来自同一个 function space.

$f \in \mathcal{F}(\mathbb{R}^n, \mathbb{R}^m)$, $f(x_0) \in \mathbb{R}^m$

$Df \in \mathcal{F}(A, \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m))$, $Df(x_0) \in \mathcal{F}(\mathbb{R}^n, \mathbb{R}^m)$

f 在某点的 derivative 是对整个 f 的一个 local approximation

因而 $Df(x_0)$ 和 f 来自同一个函数空间

而 $Df: A \to \mathbb{R}^{m\times n}$ 是 $A \subseteq \mathbb{R}^n$ 上的点到该点上的 f 的线性近似函数的映射，未必线性

而 $D: \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ 是线性的

$f \longmapsto Df$

(ex: $f: \mathbb{R} \to \mathbb{R}, x \mapsto x^3$

$Df = f': \mathbb{R} \to \mathbb{R}, x \mapsto 3x^2$)

(微分算子 $D \in \mathrm{Hom}(\mathcal{F}(\mathbb{R}^n, \mathbb{R}^n))$

是一个 $\mathcal{F}(\mathbb{R}^n, \mathbb{R}^n)$ 到自身的 linear operator)

If take $u \in \mathbb{R}^n$,

$D_u f(x_0) \in \mathbb{R}^n$

(In particular, 如果 $f: \mathbb{R}^n \to \mathbb{R}$, 则 $D_u f(x_0) \in \mathbb{R}$)

$D_u f(x_0)$ 即 $Df(x_0) \in \mathbb{R}^{m\times n}$ 在 $\vec{u} \in \mathbb{R}^n$ 方向上的作用

因而 (2) $D_u f(x_0) = Df(x_0) \cdot \vec{u}$ (if $Df(x_0)$ exists)

为什么 $Df(x_0)$ 不存在时 $D_u f(x_0)$ 仍可能存在?

因为 $D_u f(x_0)$ 是 f 在 x_0 处，对于某方向u上作用 $f_{u,x_0} = \frac{d}{dt} f(x_0+tu)$ 的线性近似

即使 f 在 x_0 处无法线性近似，它在方向u上的作用未必不可以

(3) note: remember $\mathbb{R}^n$ 为一个 vector space,

取n个 orthonormal 的 vector $\{\vec{u_1}, \dots, \vec{u_n}\}$,

$(D_{u_1} f, \dots, D_{u_n} f)$ 就可以完整描述 Df 的作用

不妨就取 $\vec{u_i} = \vec{e_i}$

if take $\vec{u_i} = \vec{e_i}$ then $D_{u_i} f(x_0) = Df(x_0) \cdot \vec{e_i}$

note that $\forall x \in \mathbb{R}^n$, $x = \sum_{i=1}^{n} a_i \vec{e_i}$ for some $a_1, \dots, a_n$

$$\Rightarrow \underset{\in\mathbb{R}^{m\times n}}{Df(x_0)} \cdot \underset{\in\mathbb{R}^n}{x} = Df(x_0) \cdot \left(\sum_i a_i \vec{e_i}\right)$$

$$= \sum_i a_i \left(Df(x_0) \vec{e_i}\right) = \sum_i a_i \underbrace{D_{e_i} f(x_0)}_{\frac{\partial f}{\partial x_i}(x_0)}$$

因而 $Df(x_0)$:

$$Df(x_0) = m\begin{bmatrix} | & & | \\ D_{e_1}f(x_0) & \cdots & D_{e_n}f(x_0) \\ | & & | \end{bmatrix}$$

We denote $D_{e_i}f(x_0)$ by $\frac{\partial f}{\partial x_i}(x_0)$, $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$

$$\Rightarrow D_{e_i}f(x_0) = \begin{pmatrix} D_{e_i}f_1(x_0) \\ \vdots \\ D_{e_i}f_m(x_0) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_i}(x_0) \\ \vdots \\ \frac{\partial f_m}{\partial x_i}(x_0) \end{pmatrix}$$

$$\Rightarrow \boxed{Df(x_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{bmatrix}}$$

the Jocobian

$$\text{So } Df(x_0)x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{bmatrix} \begin{bmatrix} a_1e_1 \\ \vdots \\ a_ne_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_i \frac{\partial f_1}{\partial x_i}(x_0) \\ \vdots \\ \sum_{i=1}^n a_i \frac{\partial f_m}{\partial x_i}(x_0) \end{bmatrix}$$

Rmk $f: \mathbb{R} \to \mathbb{R}$,

For every x, $f(x) \in \mathbb{R}$, $Df(x) \in \mathrm{Hom}(\mathbb{R}, \mathbb{R}) = \mathbb{R}$

因而 $f: \mathbb{R} \to \mathbb{R} \in \mathcal{F}(\mathbb{R}, \mathbb{R})$

$Df: \mathbb{R} \to \underbrace{\mathrm{Hom}(\mathbb{R}, \mathbb{R})}_{\mathbb{R}} \in \mathcal{F}(\mathbb{R}, \mathbb{R})$

来自同一函数空间

但是 for $f: \mathbb{R}^n \to \mathbb{R}^m$ where $n \neq 1$

$f: \mathbb{R}^n \to \mathbb{R}^m \in \mathcal{F}(\mathbb{R}^n, \mathbb{R}^m)$

$Df: \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \in \mathcal{F}(\mathbb{R}^n, \mathbb{R}^{m\times n})$

$\dim(\mathbb{R}^{m\times n}) = mn$

这是因为 $\mathbb{R} \to \mathbb{R}$ 的任何线性变换（甚至任何函数）对一个点的作用只有拉伸，因而 $\mathrm{Hom}(\mathbb{R}, \mathbb{R})$ 就是 $\mathbb{R}$ 上所有的 scalar，即 $\mathbb{R}$

而 $\mathbb{R}^n \to \mathbb{R}^m$ 的线性变换有各种形式（在多个维度上），如翻折，旋转等，$\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 的维度是 nm

因在 $f: \mathbb{R} \to \mathbb{R}$ 的 Df 和 f 来自同一空间

但是 $f: \mathbb{R}^n \to \mathbb{R}^m$ 的 Df 则并不和 f 来自同一空间

而是 $\forall x \in \mathbb{R}^n$, $Df(x)$ 和 f 来自同一空间.