Recall: $f: \mathbb{R}^n \to \mathbb{R}^m$ is diffble at x_0 if

$\exists$ matrix $Df(x_0) \in \mathbb{R}^{m\times n}$ s.t. $\lim_{|h|\to 0} \frac{|f(x_0+h) - f(x_0) - Df(x_0)h|}{|h|} \doteq 0$

(can be locally linear approximated by a linear map)

And directional derivative: linear approximation $\boxed{D_uf(x_0) \in \mathbb{R}^m}$ of the effect of $f: \mathbb{R}^n \to \mathbb{R}^m$ on a direction $u \in \mathbb{R}^n$

$$D_uf(x_0) = \lim_{t\to 0} \frac{f(x_0+tu) - f(x_0)}{u}$$

Thm Let $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ diffble at x_0 (A open.)

$\Rightarrow \forall u \in \mathbb{R}^n$, $D_uf(x_0) = \underbrace{Df(x_0)}_{\text{matrix}} \underbrace{u}_{\text{vector}}$

Pf for $u = 0 \Rightarrow$ obviously $D_uf(x) = 0 \ \forall x$
$= Df(x) \cdot \underset{\sim}{u}_0$

for $u \neq 0$, WTS: $\lim_{t\to 0} \left| \frac{f(x_0+tu) - f(x_0) - Df(x_0)u}{t} \right| = 0$

equivalently $\lim_{t\to 0} \frac{|f(x_0+tu) - f(x_0) - Df(x_0)tu|}{|tu|} = 0$

(since u is const)

This is true, by taking $h = tu$, as the def of derivative.

Def Partial derivative

$f: A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$ open

记 $\frac{\partial f}{\partial x_j}(x_0) = D_{e_j} f(x_0) \in \mathbb{R}^m$ 为 the j^{th} partial derivative of f at $\underset{\in \mathbb{R}^n}{x_0}$

$= \frac{d}{dt}\Big|_{t=0} f(x_0 + te_j)$

ex when $m=1$

$$\frac{\partial f}{\partial x_j}((x_1, x_2, \dots, x_n)) = \frac{d}{dt}\Big|_{t=0} f(x_1, \dots, x_{j-1}, x_j + t, x_{j+1}, \dots, x_n)$$

$$= \frac{d}{ds}\Big|_{s=x_j} f(x_1, \dots, x_{j-1}, s, x_{j+1}, \dots, x_n)$$

$= \varphi'(x_j)$ where

$\varphi(s) = f(x_1, \dots, x_{j-1}, s, x_{j+1}, \dots, x_n)$

(treat all aboves vars as const except s)

If $m \neq 1$, we can write $f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$ for any $u \in \mathbb{R}^n$

then $D_uf(x_0) = \lim_{t\to 0} \frac{f(x_0+tu) - f(x_0)}{t}$

$$= \lim_{t\to 0} \begin{pmatrix} \frac{f_1(x_0+tu) - f_1(x_0)}{t} \\ \vdots \\ \frac{f_m(x_0+tu) - f_m(x_0)}{t} \end{pmatrix} = \begin{pmatrix} D_uf_1(x_0) \\ D_uf_2(x_0) \\ \vdots \\ D_uf_m(x_0) \end{pmatrix}$$

In particular, $\frac{\partial f}{\partial x_j}(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(x_0) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(x_0) \end{pmatrix}$

ex let $F: \mathbb{R}^2 \to \mathbb{R}^3$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2+y^2 \\ xy \\ \sin y \end{pmatrix}$

then $\frac{\partial F}{\partial x}(x,y) = \begin{pmatrix} \frac{\partial}{\partial x}(x^2+y^2) \\ \frac{\partial}{\partial x}(xy) \\ \frac{\partial}{\partial x}(\sin y) \end{pmatrix} = \begin{pmatrix} 2x \\ y \\ 0 \end{pmatrix}$

$\frac{\partial F}{\partial y}(x,y) = \begin{pmatrix} 2y \\ x \\ \cos y \end{pmatrix}$

if $u = (1,2)$ then $D_uF(x,y) = \frac{d}{dt}\Big|_{t=0} \begin{pmatrix} (x+t)^2 + (y+2t)^2 \\ (x+t)(y+2t) \\ \sin(y+2t) \end{pmatrix} = \begin{pmatrix} 2x+4y \\ y+2x \\ 2\cos(y) \end{pmatrix}$

In this case we can check

$D_{(1,2)}F(x,y) = 1 D_{e_1}F(x,y) + 2D_{e_2}F(x,y)$

which suggests (但并不证明) F being diffble

Thm Let $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ (A open). $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$

(a) If f is diffble at $x_0 \in \mathbb{R}^n$

则 $\boxed{Df(x_0)}$ 这个 $m\times n$ matrix 的 $\boxed{(i,j)^{th} \text{ entry 为 } \frac{\partial f_i}{\partial x_j}(x_0)}$

(scalar value function 增长最快的方向)

$$Df(x_0) = \begin{pmatrix} | & | & & | \\ \frac{\partial f}{\partial x_1}(x_0) & \frac{\partial f}{\partial x_2}(x_0) & \cdots & \frac{\partial f}{\partial x_n}(x_0) \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} \text{---} \ \nabla f_1^T \ \text{---} \\ \text{---} \ \nabla f_2^T \ \text{---} \\ \vdots \\ \text{---} \ \nabla f_m^T \ \text{---} \end{pmatrix}$$

Def gradient

这是一个只面向标量值函数 $g: \mathbb{R}^n \to \mathbb{R}$ $(m=1)$ 的概念

$$\nabla g(x) = \begin{pmatrix} \frac{\partial g}{\partial x_1}(x) \\ \vdots \\ \frac{\partial g}{\partial x_n}(x) \end{pmatrix} = Dg(x)$$

(b) f is diffble at x_0 iff $\forall f_i$ is diffble at x_0

Pf (a) $\forall j$, $Df(x_0)e_j = D_{e_j}f(x_0) = \frac{\partial f}{\partial x_j}(x_0)$

(b) f is diffble at $x_0 \Leftrightarrow \exists$ matrix A s.t.

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

$\Leftrightarrow \forall i$, $\frac{|f_i(x_0+h) - f_i(x_0) - \overset{\text{row } i \text{ of } A}{A_i} h|}{|h|} \to 0$ as $|h| \to 0$

$\Leftrightarrow f_i$ is diffble at x_0 with $Df_i(x_0) = A_i$ (i^{th} row of A) □

ex let $f(x,y) = \begin{pmatrix} x^2+y^2 \\ xy \\ \sin y \end{pmatrix}$

if f is diffble then it derivative should be

$$Df(x,y) = \begin{pmatrix} \frac{\partial}{\partial x}(x^2+y^2) & \frac{\partial}{\partial y}(x^2+y^2) \\ \frac{\partial}{\partial x}(xy) & \frac{\partial}{\partial y}(xy) \\ \frac{\partial}{\partial x}(\sin y) & \frac{\partial}{\partial y}(\sin y) \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ y & x \\ 0 & \cos y \end{pmatrix}$$

Def this matrix for $f: \mathbb{R}^n \to \mathbb{R}^m$ with $\frac{\partial f_i}{\partial x_j}(x_0)$ is called the Jacobian at x_0

Rmk $\exists Df(x_0) \Longrightarrow$ Jacobian at x_0 exists 且 $= Df(x_0)$

但是 $\exists Df(x_0) \not\Longleftarrow$ Jacobian at x_0 exists

Rmk f diffble at $x_0 \Leftrightarrow f_1, \dots, f_m$ diffble at x_0

是一个自然且令人舒适的结论，因为这就表示了在处理 vector-value 的函数的导数时只需把它拆成 m 个 sclar-value 的函数即可

Thm Let $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ (A open)

对于 $x_0 \in A$，如果 $\exists$ some $B_\varepsilon(x_0)$ s.t. $\forall x \in B_\varepsilon(x_0)$, all $\frac{\partial f_i}{\partial x_j}$ exists (在整个 $B_\varepsilon(x_0)$ 上) 且 ctn. at x_0

$\Longrightarrow f$ diffble at x_0

(**Def** $f \in C^1$ if all $\frac{\partial f_i}{\partial x_j}$ exists 且 ctn on A (dom f)

(so the thm states: $\boxed{f \in C^1 \text{ around } x_0 \Rightarrow f \text{ diffble at } x_0}$

$\boxed{f \in C^1 \Rightarrow f \text{ diffble}}$))

Pf Since f is diffble at x_0 iff its components are, we may assume $m=1$ (that's why thm1(b) 重要 在考虑任何 $\mathbb{R}^n \to \mathbb{R}^m$ 函数的微分性质时只需考虑 $\mathbb{R}^n \to \mathbb{R}$ 时)

Let $r > 0$ s.t. $B_r(x_0) \subseteq A$

And $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ with $|h| < r$

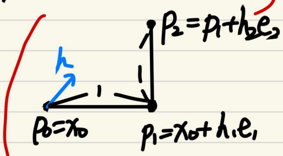
Set $p_0 = x_0$

$p_1 = p_0 + h_1 e_1$

$p_2 = p_1 + h_2 e_2$

$\vdots$

$p_n = p_{n-1} + h_n e_n = x_0 + h$



Let $\varphi_j(s) = f(p_{j-1} + s e_j)$

$\varphi_j : I_j \to \mathbb{R}$ where $I_j \supseteq B_{h_j}(0)$

$\Rightarrow \varphi_j'(s) = \frac{\partial}{\partial x_j} f(p_{j-1} + s e_j)$ so! $q_j \in B_h(x)$

在 $m=1$ 时，$D_u f(x_0) \in \mathbb{R}$ which means 可用thms in 一元分析

By MVT, $\exists c_j$ between 0 & h_j s.t.

$$\varphi_j(h_j) - \varphi_j(0) = \varphi_j'(c_j) \cdot h_j$$

$$f(p_{j-1} + h_j e_j) - f(p_{j-1}) = \frac{\partial f}{\partial x_j}(\underbrace{p_{j-1} + c_j e_j}_{q_j})\, h_j$$

$$f(x_0+h) - f(x_0) = \sum_{j=1}^n \left(f(p_j) - f(p_{j-1})\right) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(q_j) h_j$$

$$f(x_0+h) - f(x_0) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(p_0) \cdot h_j = \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(p_0)\right) h_j$$

If $A = \left(\frac{\partial f}{\partial x_1}(p_0), \dots, \frac{\partial f}{\partial x_n}(p_0)\right)$

$$\Rightarrow |f(x_0+h) - f(x_0) - Ah| \le |h| \sum_{j=1}^n \left|\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(p_0)\right|$$

$$\Rightarrow \frac{|f(x_0+h) - f(x_0) - Ah|}{|h|} \le \sum \left|\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(p_0)\right|$$

continuity of partials gives this $\to 0$ as $|h| \to 0$ □

Intuition: We want to change one variable at a time

Rmk The converse of this thm is false

counterex $f: \mathbb{R} \to \mathbb{R}$

$x \mapsto x^2 \sin(\frac{1}{x})$

diffble but f' not ctn. ($\notin C^1$)

Pf 复盘

let $r > 0$ s.t. $B_r(x) \subseteq A$, and all partials are defined and ctn. at A.

Let $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$ s.t. $\|h\| < r$

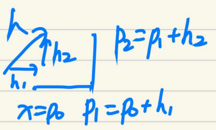
Set $p_0 = x$

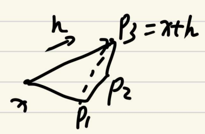
$p_1 = p_0 + h_1 e_1$

$p_2 = p_1 + h_2 e_2$

$\vdots$

$p_n = p_{n-1} + h_n e_n = x + h$





Then $f(x+h) - f(x) = f(p_n) - f(p_0)$

$= \sum_i f(p_i) - f(p_{i-1})$

Define: $\varphi_i(s) = f(p_{i-1} + s e_i)$ ($\varphi_i : [0, h_i] \to \mathbb{R}$)

$\underline{ex}$ $\varphi_1(s) = f(p_0 + se_1)$, $\varphi_1(h_1) = f(p_1)$

$\varphi_2(s) = f(p_1 + se_2)$, $\varphi_2(h_2) = f(p_2)$

$\vdots$

$\varphi_i : s \longmapsto f(p_{i-1}+se_i)$

e_i 是这个方向，s 是沿这个方向的变化量

note: φ_i 表示 f 的值随 h 的第 i 分量的变化而变化的量

$f(p_{i-1}+se_i)$ 是 f 随 x 在这个方向上的变化量产生的变化量

note: $\varphi_i'(s) = \dfrac{\partial}{\partial x_i} f(\underbrace{p_{i-1}+se_i}) \in B_r(x_0)$

$\varphi_i'(s) = \lim\limits_{s_0 \to 0} \dfrac{\varphi_i(s+s_0)-\varphi_i(s)}{s_0}$

即 $x = p_{i-1}+se_i$ 处，s 加上微小量 s_0 后 f 相对 x 的微小变化率

因而 $\varphi_i'(s)$ ∃ 且 ctn.

by MVT: $\underline{\varphi_i(h_i) - \varphi_i(0)} = h_i\, \varphi_i'(c_i)$ for some $c_i \in (0, h_i)$

$= f(p_i) - f(p_{i-1})$

因而 φ 把 x 在 e_i 上的变化量映射于 $f(x)$ 的相应变化量

$$\begin{aligned}
\Longrightarrow f(x+h) - f(x) &= \sum_i \left(f(p_i) - f(p_{i-1})\right) \\
&= \sum_i \left(\varphi_i(h_i) - \varphi_i(0)\right) \\
&= \sum_i h_i\, \varphi_i'(c_i) \\
&= \sum_i h_i \frac{\partial}{\partial x_i} f(p_{i-1} + c_i e_i)
\end{aligned}$$

因而 φ' 是偏导数 $\dfrac{\partial}{\partial x_i}$

$\varphi'(s)$ 即 f 随（x 在 e_i 上的变化）从 0 到 h_i 过程中在 $s \in [0, h_i]$ 时的变化速率

$\underline{Rmk}$ my thought on directional derivative

完全可以令 $m=1$ 来看

$$D_u f(x) = \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t}$$

t 是一个和 h 一样的虚拟变量，表示在 u 方向上的（微小）变化量

因而 $D_u f(x)$ 即：在 u 方向上的微小变化量引起的 f 在 x 处的变化量有多大

特别地，$u = e_i$ 时，表示 $(x_1, \dots, x_n)$ 中 x_i 的微小变化量引起的 f 在 x 处的变化量有多大

因而 x 是一个初始(const)量，$D_u f(x) = \left.\dfrac{d}{dt}\right|_{t=0} f(x+ut)$

note: $D_u f(x)$ 对于 u 是 linear 的

即 $\boxed{D_{a_1\vec{u}_1 + a_2\vec{u}_2} f(x) = a_1 D_{\vec{u}_1} f(x) + a_2 D_{\vec{u}_2} f(x)}$