Example of derivative of linear functions

✩ $\forall x \in X \quad DL|_x = L$

1 $\quad f: z \mapsto \bar{z} \qquad X = V = \mathbb{C}$, real vector space

$$\overline{z+h} = \bar{z} + \bar{h} \Rightarrow f(z+h) = f(z) + f(h) \Rightarrow Df|_z(h) = f(h) = \bar{h}$$

2.

$$f: \mathbb{R}^n \to \mathbb{R}^n \quad x \mapsto Ax, \quad A \in Mat(n \times n; \mathbb{R})$$

$$A(x+h) = Ax + Ah$$

3.

$$tr: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \qquad trA = \sum a_{ii}$$

$tr(A+H) = trA + trH$ ✩ the meaning of $H \to 0$?

Exercise $\qquad f(A) = A^3$

1. find $Df|_A(H)$

$$f(A+H) = (A+H)^3 = A^3 + A^2H + AHA + HA^2 + \underline{H^2A + HAH + AH^2 + H^3}$$

/how? triangle inequality

$$= A^3 + A^2H + AHA + HA^2 + o(H)$$

$$= f(A) + L_A(H) + o(H), \text{ where } L_A(H) = A^2H + AHA + HA^2.$$

2. Prove that $D(\cdot)^{-1}|_A H = -A^{-1}HA^{-1}$

① $(A+H)^{-1} = (A(I + A^{-1}H))^{-1} = (I + A^{-1}H)^{-1}A^{-1}$ — $(*)$

note that $(I+A^{-1}H)(I-A^{-1}H) = I - A^{-1}HA^{-1}H = I + o(H)$

Apply inverse on both side:

$$I - A^{-1}H = (I+A^{-1}H)^{-1}(I+o(H)) = (I+A^{-1}H)^{-1} + o(H)$$

$$\Rightarrow (I+A^{-1}H)^{-1} = I - A^{-1}H + o(H)$$

$$(*) = (I - A^{-1}H + o(H))A^{-1} = A^{-1} - \underbrace{A^{-1}HA^{-1}}_{D(\cdot)^{-1}|_A(H)} + o(H) \quad \divideontimes$$

Structure of derivative

If exists, $Df|_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \cong Mat(m\times n; \mathbb{R})$

$(Df|_x)_{ij} := \langle e_i, Df|_x e_j \rangle$ matrix element.

Partial Derivative

$\Omega \in \mathbb{R}^n$ $\quad f: \Omega \to \mathbb{R}$. $\quad \left.\frac{\partial f}{\partial x_j}\right|_x := \lim_{h\to 0} \frac{f(x+he_j) - f(x)}{h}$

$$= \lim_{h\to 0} \frac{f(x_1, \cdots, x_{j-1}, x_j + h, x_{j+1}, \cdots, x_n) - f(x)}{h}$$

$\Rightarrow (Df|_x)_{ij} = \frac{\partial f_i}{\partial x_j}$, or

$$Df|_x = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right) = \left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}\right|_x =: J_f(x)$$

Jacobian

If $Df|_x$ exists, $J_f(x)$ is its matrix representation. However, the converse is not true.

partial derivative "only examine some direction".

Theorem

1. all partial derivatives are bounded $\Rightarrow f$ is continuous
2. all partial derivatives are continuous $\Rightarrow f$ is continuously differentiable

$\Rightarrow$ a gap! only differentiable?

Example: ∂f are not all cont., but f is diff.

$$f(x,y)=\begin{cases}(x^2+y^2)\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) & (x,y)\neq(0,0)\\ 0\ , & (x,y)=(0,0)\end{cases}$$

along x, $\partial_x f$ is not cont:

$$\partial_x f|_{(0,0)}=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=0$$

$$\partial_x f|_{(x,y)\neq 0}=2x\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right)-\frac{x}{\sqrt{x^2+y^2}}\cos\left(\frac{1}{\sqrt{x^2+y^2}}\right)$$

Now, we see f is diff. at $(0,0)$. Let $h:=(h_1,h_2)$

$$\Rightarrow f(0+h)=\underbrace{f(0)}_{0}+\underbrace{Df|_{(0,0)}}_{0}h+\underbrace{(h_1^2+h_2^2)\sin\left(\frac{1}{\sqrt{h_1^2+h_2^2}}\right)}_{o(h)}\Rightarrow Df|_{(0,0)}\text{ exists.}$$

$$\lim_{h\to 0} \frac{\left|(h_1^2+h_2^2)\sin\left(\frac{1}{\sqrt{h_1^2+h_2^2}}\right)\right|}{\|h\|} = \lim_{h\to 0} \|h\| \left|\sin\left(\frac{1}{\sqrt{h_1^2+h_2^2}}\right)\right| = 0.$$

Product rule

$$D(f \odot g) = (Df)\odot g + f \odot (Dg)$$

at x

$$u \mapsto Df|_x u \odot g(x) + f(x) \odot Dg|_x u$$

Chain rule

$$D(f\circ g)|_x = Df|_{g(x)} \circ Dg|_x$$

Exercise

the derivative of $tr(AA^T)$ by $\begin{cases} 1. \text{ def} \\ 2. \text{ chain / product rule} \end{cases}$

1 $tr((A+H)(A+H)^T) = tr((A+H)(H^T+A^T)) = tr(AH^T + AA^T + HH^T + HA^T)$

$$= tr(AA^T) + \underline{tr(AH^T + HA^T)} + \underbrace{tr(HH^T)}_{o(H)}$$

2. $f(A) = tr(AA^T) = tr(\cdot) \circ g(A), \quad g(A) = AA^T$

$$Df|_x H = \underbrace{D\,tr(\cdot)|_{g(A)}}_{\text{linear}} \circ Dg|_A H$$

$$= tr(\cdot) \circ (AH^T + HA^T)$$

$$= tr(AH^T + HA^T)$$

$Dg|_A H$: $g(A+H) = (A+H)(A+H)^T$

$$= AA^T + AH^T + HA^T + HH^T$$

$$= g(A) + \underline{AH^T + HA^T} + o(H)$$

Integral

bounded : $\|f\|_\infty := \sup_{x\in I} \|f(x)\|_V < \infty$

all bounded function $f: I \to V$ is denoted by $\mathcal{L}^\infty(I,V)$

$(f_n)_{n\in\mathbb{N}}$ converges uniformly to f if

$$\|f_n - f\|_\infty := \sup_{x\in I} \|f_n(x) - f(x)\|_V \xrightarrow{n\to\infty} 0$$

a regulated function is the uniform limit of some step function.

the standard estimate

$$\left\| \int_a^b f(x)\,dx \right\|_V \le \int_a^b \|f(x)\|_V\,dx \le |b-a| \cdot \sup_{x\in[a,b]} \|f(x)\|_V .$$

$$\Rightarrow \int_a^b f(x)\,dx = \int_a^b \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix} dx = \begin{pmatrix} \int_a^b f_1(x)\,dx \\ \vdots \\ \int_a^b f_n(x)\,dx \end{pmatrix}$$

Mean Value Theorem.

$$f(x+y)-f(x)=\int_0^1 \underbrace{Df\big|_{x+ty}}_{\in V}\,y\,dt=\Big(\underbrace{\int_0^1 Df\big|_{x+ty}\,dt}_{\in \mathcal{L}(X,V)}\Big)\,y$$

Ex:

$$f(x_1,x_2)=x_1^2+x_2^2 \qquad f(x+y)-f(x):$$

① $f(x+y)-f(x)=2x_1y_1+2x_2y_2+y_1^2+y_2^2$

② $Df\big|_{x+ty}\,y=(2x_1+2ty_1 \quad 2x_2+2ty_2)\begin{pmatrix}y_1\\ y_2\end{pmatrix}=2x_1y_1+2ty_1^2+2x_2y_2+2ty_2^2$

$\int_0^1 Df\big|_{x+ty}\,y\,dt=2x_1y_1+y_1^2+2x_2y_2+y_2^2$

③ $Df\big|_{x+ty}=(2x_1+2ty_1 \quad 2x_2+2ty_2)$

$\int_0^1 Df\big|_{x+ty}\,dt=\int_0^1(2x_1+2ty_1 \quad 2x_2+2ty_2)\,dt$

$=(2x_1+y_1 \quad 2x_2+y_2)$

$\int_0^1 Df\big|_{x+ty}\,dt\;y=(2x_1+y_1 \quad 2x_2+y_2)\begin{pmatrix}y_1\\ y_2\end{pmatrix}=2x_1y_1+y_1^2+2x_2y_2+y_2^2$

Derivative Estimate

$$\|f(x+y)-f(x)\|_V\le\|y\|_X\cdot\sup_{0\le t\le 1}\|Df\big|_{x+ty}\|$$

☆ Differentiating under an Integral

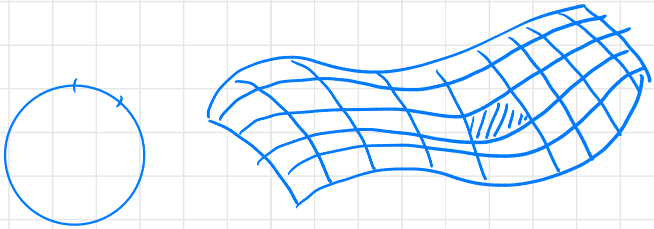
$f : I \times \Omega \to V$, Ω is open

$$g(x) = \int_a^b f(t,x)\,dt, \quad Dg(x) = \int_a^b Df(t,\cdot)\big|_x\,dt$$

☆ key : exchange the limit

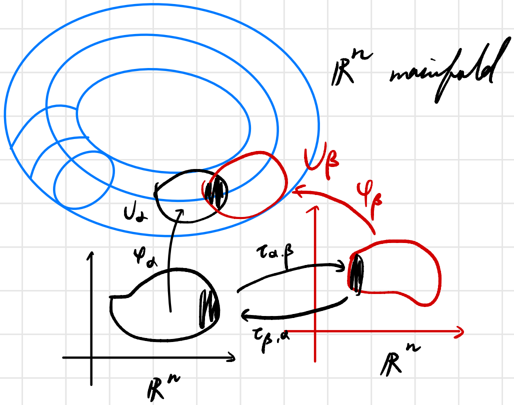
Curve

local :



manifold

atlas



Curve

A set $C \subset V$ for which there exists a continuous, surjective and locally injective map $\gamma : I \to C$

diff. $\to$ parametrization

If γ is globally injective γ, C is called simple curve.

For $I=(a,b)$, if $\exists$ (one is enough) $\lim_{t\to a}\gamma(t) = \lim_{t\to b}\gamma(t)$, C is closed.

if not, C is open, $\begin{cases} \lim_{t\to a}\gamma(t) \text{ is initial point} \\ \lim_{t\to b}\gamma(t) \text{ is final point} \end{cases}$

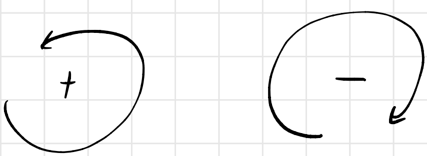
Re-parametrize

We have $\gamma: I \to C$, but we can also have

bijective: $\gamma' : J \to I$, $J, I \subset \mathbb{R}$ $\Rightarrow$ reparametrization

$\begin{cases} \nearrow : \text{preserving} \\ \searrow : \text{reversing} \end{cases}$

$\tilde{\gamma} := \gamma \circ \gamma' : J \to I \to C$



: only simple curve

why?

Smooth

If $\exists\, \gamma$: (i) γ is continuously differentiable on int I

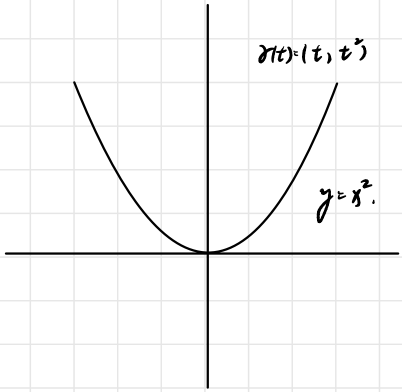
(ii) $D\gamma\big|_t \neq 0$ for all $t \in \text{int } I$.

smooth reparametrization : γ' has (i), (ii)

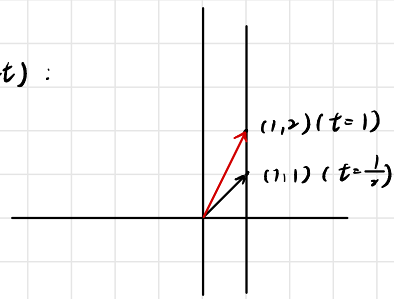
Tangent vector

C^* : oriented smooth curve and $p \in C^*$

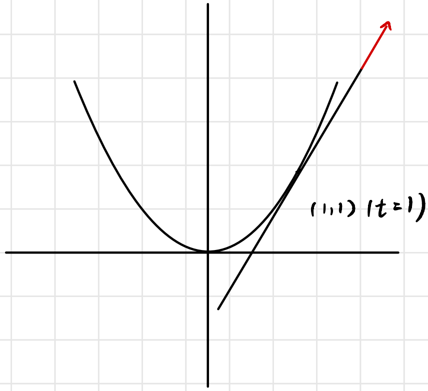
$p = \gamma(t)$, $\quad T\circ\gamma(t) = \dfrac{\gamma'(t)}{\|\gamma'(t)\|}$



$\gamma'(t) = (1, 2t)$:

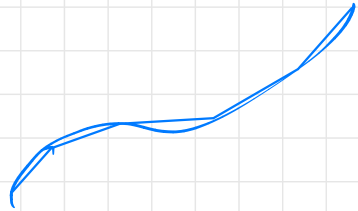


? dealing with vector



Curve Length

$$\ell(C) := \sup_{P} \ell_P(C)$$



However, in practice

$$\ell(C) = \int_a^b \|\gamma'(t)\|\, dt$$

$\Rightarrow \ell^{-1}$ is a natural parametrization

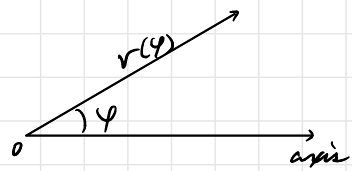
$$\gamma = \ell^{-1} : I \to C \qquad \operatorname{int} I = (0, \ell(C))$$

$$\| $$

$$\gamma \circ (\ell \circ \gamma)^{-1}$$

In polar coordinate, $r = r(\varphi)$, $0 < \varphi < 2\pi$.

$$\ell = \int_0^{2\pi} \sqrt{r^2(\varphi) + r'^2(\varphi)}\, d\varphi.$$



$$\gamma : \varphi \longmapsto \begin{pmatrix} r(\varphi)\cos\varphi \\ r(\varphi)\sin\varphi \end{pmatrix}$$

$$\gamma'(\varphi) = \begin{pmatrix} r'(\varphi)\cos\varphi - r(\varphi)\sin\varphi \\ r'(\varphi)\sin\varphi + r(\varphi)\cos\varphi \end{pmatrix}$$

$$\Rightarrow \|\gamma'(\varphi)\| = \sqrt{r'^2(\varphi) + r^2(\varphi)} \;\Rightarrow\; \ell(C) = \int_0^{2\pi} \sqrt{r'^2(\varphi) + r^2(\varphi)}\, d\varphi \;*$$

t is not anyway unique or special!