Rappel sur la différentiabilité

E, F e. normés

$f : E \to F$ une fonction

1) f est **Fréchet différentiable** si $\exists L \in \mathcal{L}(E, F)$ t.q.

$$f(x+h) = f(x) + Lh + o(\|h\|) \quad \left(\Leftrightarrow \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(x+h) - f(x) - Lh}{\|h\|} = 0 \right)$$

Dans ce cas, on note $L = f'(x)$ (c'est la dérivée)

2) **gradient de f en x** 2 conditions : . $F = \mathbb{R}$

. $\exists$ un prod. scalaire sur E $\langle \cdot, \cdot \rangle$

$\Rightarrow \exists !$ vecteur de E, noté $\nabla f(x) \in E$, tel que $\forall h \in E$

$$\boxed{f'(x).h = \langle \nabla f(x), h \rangle} \qquad (1)$$

indépendant du prod. scal — dépend du produit scalaire

Cas particulier : $E = \mathbb{R}^n$ et $\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i$ (p.s. euclidien)

$$\boxed{\nabla f(x) = \left(\frac{\partial f}{\partial x_i}(x) \right)_{i=1}^{n}} \qquad [\nabla f(x)]_i = (e^i)^T \nabla f(x) = f'(x).e^i$$

$$= \lim_{\substack{t \to 0 \\ \neq}} \frac{1}{t}\left[f(x+te_i) - f(x)\right]$$

$$= \frac{\partial f}{\partial x_i}(x)$$

$\Longrightarrow$ 2 manières de calculer $Df(x)$ dans ce cas

① $\underset{\in E}{\nabla f(x)} = \left(\dfrac{\partial f}{\partial x_i}\right)_{i=1}^{n}$ (pour le p.s. euclidien) $\longrightarrow$ $\nabla f(x)$ (pour un p.s. donné)

② Calculer $f'(x)$ et utiliser (1)

3) <u>Dérivées secondes</u> Si f est 2 fois différentiable en $x \in E$, et si $\exists\, L_1 \in \mathcal{L}(E,F)$, $\exists\, L_2 \in \mathcal{L}_2(E,F)$ et <u>symétrique</u> $\left(L_2(h,k) = L_2(k,h)\right)$ t.q.

$$f(x+h) = f(x) + L_1 h + \frac{1}{2} L_2(h,h) + o(\|h\|^2) \quad (1')$$

<u>alors</u> $f'(x) = L_1$, $f''(x) = L_2$

b) Hessien de f en x : 2 conditions :
- $E = \mathbb{R}$
- p. scalaire $\langle \cdot , \cdot \rangle$ sur E

$\Rightarrow$ $\exists!$ opérateur linéaire <u>auto-adjoint</u> $E \to E$ noté $\nabla^2 f(x)$ tel que $\forall h, k \in E$

$$\boxed{f''(x).(h,k) = \langle \nabla^2 f(x)\, h, k \rangle} \qquad (2)$$

indépendant du produit scalaire

hessien dépend du produit scalaire

Cas particulier où $E = \mathbb{R}^n$ et $\langle x, y \rangle = x^T$

$$\boxed{[\nabla^2 f(x)]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)} \qquad (3)$$

2 manières de calculer $\nabla^2 f(x)$ dans ce cas

$f(x) =$ fct des x_i $\longrightarrow$
- soit par (3) donne $\nabla_E^2 f(x)$ (pour le p.s. euclidien) $\leadsto$ $\nabla^2 f(x)$ (pour le p.s. donné)

pbe structurés $\longrightarrow$
- soit on calcule $f''(x)$ puis $\nabla^2 f(x)$ par (2)

$\boxed{\text{Ex } 1}$

$b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$ (non symétrique)

$x \in \mathbb{R}^n \longrightarrow f(x) = \frac{1}{2} x^T A x - b^T x \in \mathbb{R}$

(i) Calculer $\nabla f(x)$ et $\nabla^2 f(x)$ pour le produit scalaire euclidien $(x,y) \longmapsto \sum_{i=1}^n x_i y_i$

$$f(x+h) = f(x) + \overbrace{(x^T A - b^T)}^{\nabla f(x)^T h} h + \frac{1}{2} h^T A h$$

$x^T A^T h$

$$\Rightarrow \nabla f(x) = A x - b$$

voir (i')

$$f(x+h) = f(x) + \overbrace{\left(\frac{1}{2} x^T A h + \frac{1}{2} h^T A x - b^T h\right)}^{\nabla f(x)^T h} + \frac{1}{2} h^T A h$$

$$f'(x)\, h = \nabla f(x)^T h = \Big[\underbrace{\tfrac{1}{2} x^T A + \tfrac{1}{2} x^T A^T - b^T}_{g}\Big] h$$

$$f''(x) \cdot (h,h) = h^T \nabla^2 f(x)\, h = h^T A h$$

- $\forall h \in \mathbb{R}^n : \nabla f(x)^T h = g^T h \Rightarrow \nabla f(x) = g$

$\forall h \in \mathbb{B}, \langle u, h \rangle = \langle v, h \rangle \overset{?}{\Rightarrow} u = v$

oui car $\quad$ ——— $\quad \langle u - v, h \rangle = 0$

$h = u - v \Rightarrow \| u - v \|^2 = 0 \Rightarrow u = v$

$$\rightsquigarrow \boxed{\nabla f(x) = \frac{1}{2}(A^T + A)x - b}$$

- $\forall h : h^T \nabla^2 f(x) h = h^T A h \quad \overset{?}{\Rightarrow} \quad \nabla^2 f(x) = A$

non !

On suppose que $h^T B h = h^T A h, \quad \forall h$

$$(h+k)^T B (h+k) = (h+k)^T A (h+k), \quad \forall h, k$$

$$\cancel{h^T B h} + k^T B h + h^T B k + \cancel{k^T B k}$$

$$= \cancel{h^T A h} + k^T A h + h^T A k + \cancel{k^T A k}$$

$$h^T(B+B^T)h = h^T(A+A^T)h$$

$$\Rightarrow \quad h^T(B+B^T) = h^T(A+A^T), \ \forall h$$

$$\Rightarrow \quad B+B^T = A+A^T$$

$$\underbrace{\nabla^2 f(x) + \nabla^2 f(x)^T}_{\substack{\parallel \\ 2\,\nabla^2 f(x)}} = A+A^T$$

$$\rightsquigarrow \quad \boxed{\nabla^2 f(x) = \frac{A+A^T}{2}}$$

$$A \text{ sym} \Rightarrow \nabla^2 f(x) = A$$

On suppose toujours que A est symétrique dans une fonction quadratique

$$f(x) = \frac{1}{2} x^T A x - b^T x$$

$$A = A_s + A_2$$

$$A_s := \frac{A + A^T}{2}$$

$$A_2 := \frac{A - A^T}{2}$$

$$\frac{1}{2} x^T A x = \frac{1}{2} x^T A_s x + \cancel{\frac{1}{2} x^T A_2 x}$$

$$= 0$$

en effet $(x^T A_2 x)^T = x^T A_2 x$

$$\| \quad x^T A_2^T x$$

$$\| \quad - x^T A_2 x$$

$$\Rightarrow x^T A_2 x = 0$$

À faire

1.2 et 1.3

f convexe $\Longleftrightarrow$ $A \succeq 0$ $\Longleftrightarrow$ $v^T A v \geq 0 \quad \forall v$

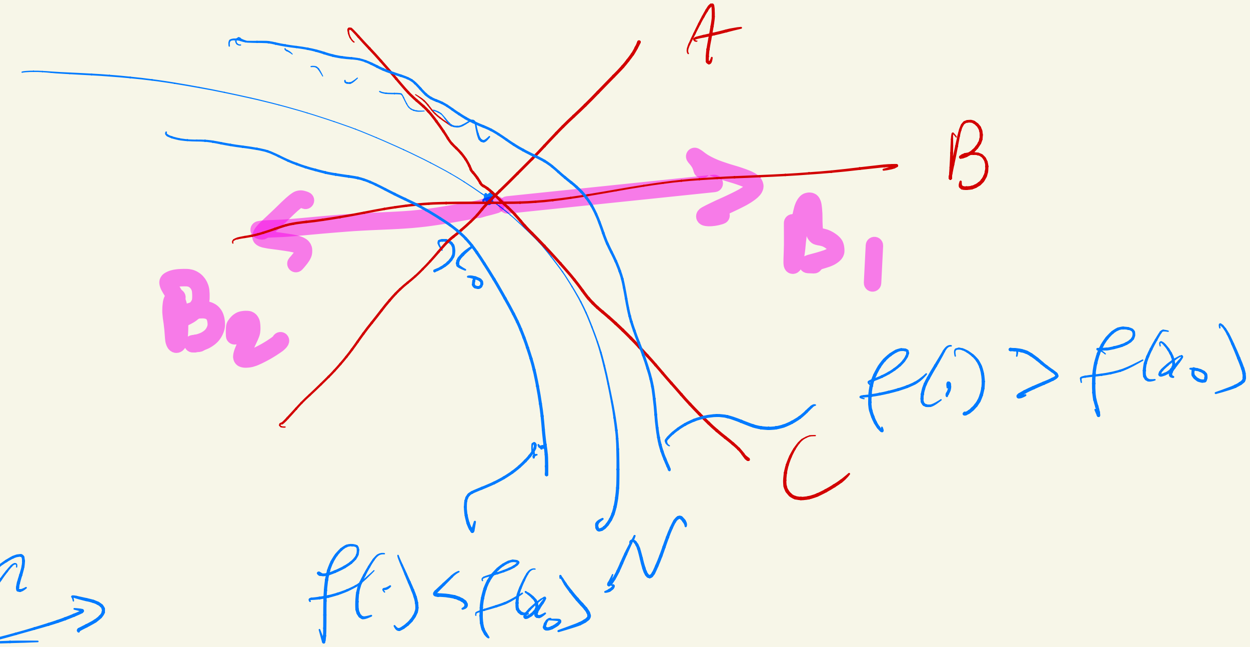
f strict convexe $\Longleftrightarrow$ $A \succ 0$ $\Longleftrightarrow$ $v^T A v > 0 \quad \forall v \neq 0$

$\boxed{\text{Ex 3}}$

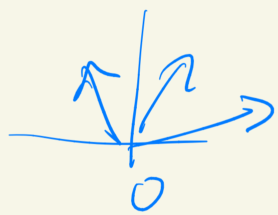
$f : \mathbb{R}^n \to \mathbb{R}$ $\qquad x_0 \in \mathbb{R}^n$ $\qquad Df(x_0) \neq 0$

$N = \{ x \in \mathbb{R}^n : f(x) = f(x_0) \}$

$(n = 2)$



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



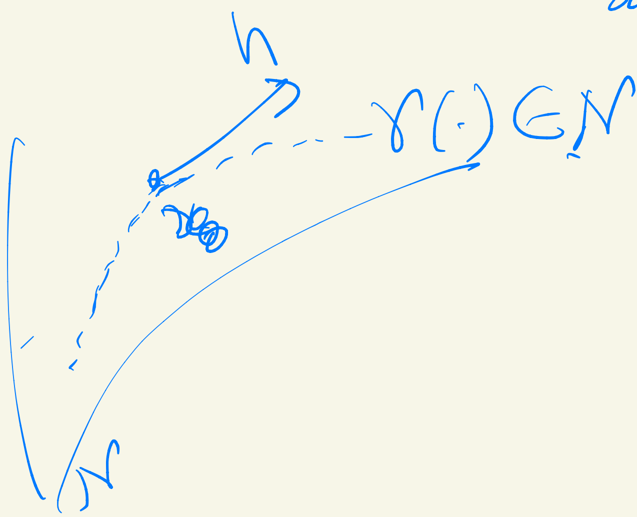
Question 1

le $Df(x_0)$ est aligné sur

A	~~||||~~ ||||
B	|
C	||

$$\boxed{Df(x_0) \perp h, \quad \forall h \in T_{x_0} N}$$

$h \in T_{x_0} N \iff \exists\, \gamma :]-1,1[\longmapsto N$



$\gamma(0) = x_0$

$\gamma'(0) = h$

$$\underbrace{f(\gamma(t))} \geq f(x_0), \quad \forall t \in]-1,1[$$

dérivée p. t en $t = 0$

$(f \circ \gamma)(t)$

$$f'(\gamma(0)) \cdot \gamma'(0) = 0$$

$$\langle Df(x_0), h \rangle = f'(x_0) \cdot h = 0, \quad \forall h \in T_{x_0} N$$

$\boxed{\text{Question 2}}$ $Df(x_0)$ dirigé suivant B_1 ou $\cancel{B_2}$

$$f(x_0 + t\, Df(x_0)) \overset{?}{\geq} f(x_0)$$ tout $t > 0$ petit

$$\frac{f(x_0+t\nabla f(x_0))-f(x_0)}{t} \geq 0 \quad \forall t>0 \text{ petit}$$

$$\lim_{t\downarrow 0}\left(\quad '' \quad \right) = f'(x_0)\cdot \nabla f(x_0)$$

$$= \langle \nabla f(x_0), \nabla f(x_0) \rangle$$

$$= \| \nabla f(x_0) \|^2$$

$$> 0$$

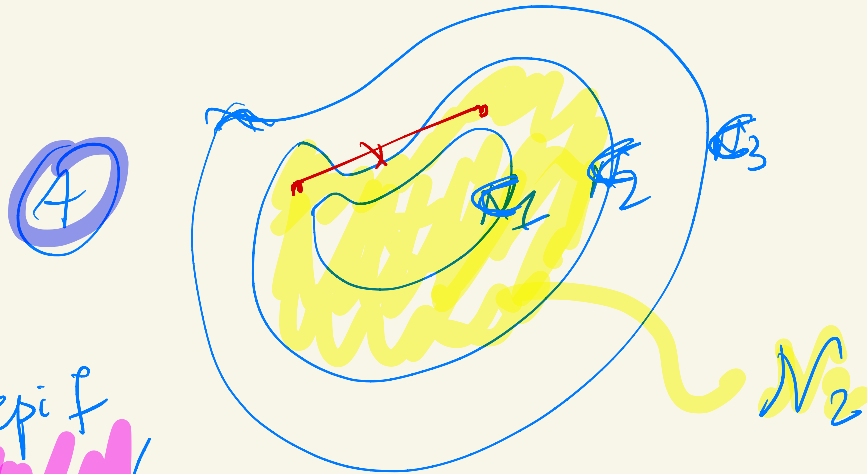
<u>Conclusion:</u> f croît dans la direction du gradient

$\boxed{\text{Ex 5}}$

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

$C_i := \{ x \in \mathbb{R}^2 : f(x) = i \}$

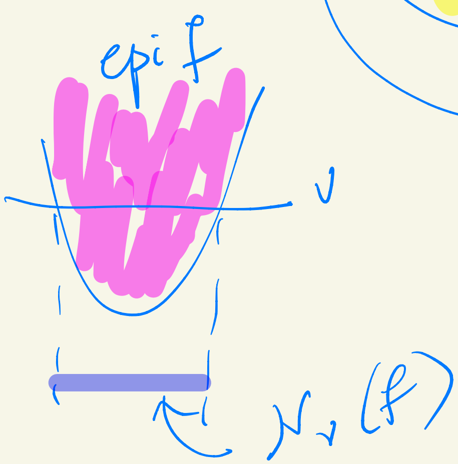
Peut-on trouver 1 fonction convexe ayant les courbes de niveau suivantes



$\mathrm{epi}\, f = \{ (x, \alpha) \in E \times \mathbb{R} : f(x) \leq \alpha \}$

$N_v(f) = \{ x \in \mathbb{R}^2 : f(x) \leq v \}$

$\boxed{f \text{ convexe}} \Rightarrow \boxed{N_v(f) \text{ convexe } \forall v \in \mathbb{R}}$



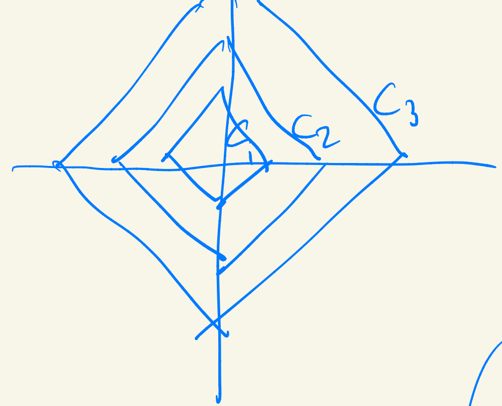
$x_0, x_1 \in N_v(f)$, $t \in [0,1]$

$f((1-t)x_0 + t x_1) \overset{?}{\leq} v$

$\leq (1-t) \underbrace{f(x_0)}_{\leq v} + t \underbrace{f(x_1)}_{\leq v} \leq v$

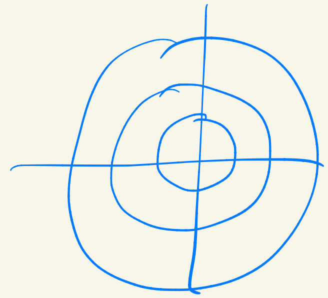
$\Rightarrow (1-t)x_0 + t x_1 \in N_v(f)$

B



Oui, par exemple.

$$f(x) = \|x\|_1 = \sum_{i=1}^{n} |x_i|$$



$\|\cdot\|_2$

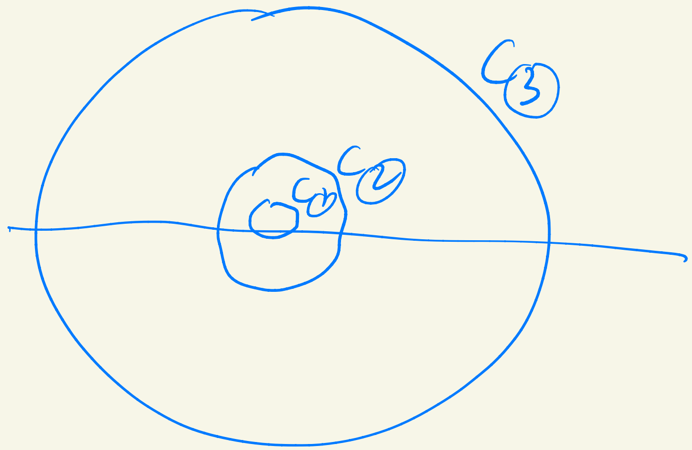
$$\boxed{\text{Une norme est convexe}}$$

$$\|(1-t)x_0 + t x_1\|$$

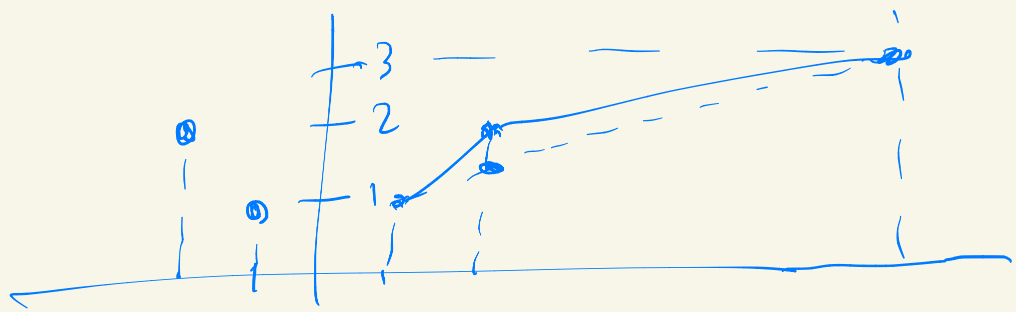
$$\leq \|\underbrace{(1-t)}_{\geq 0} x_0\| + \|\underbrace{t}_{\geq 0} x_1\|$$

$$\leq (1-t)\|x_0\| + t\|x_1\|$$

C

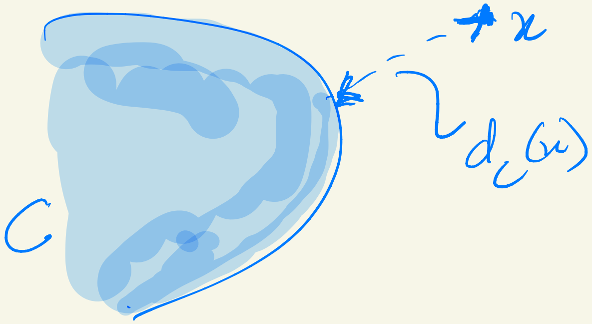


$\nexists$ fct convexe avec les courbes de niveau



$\boxed{\text{Ex } 7}$

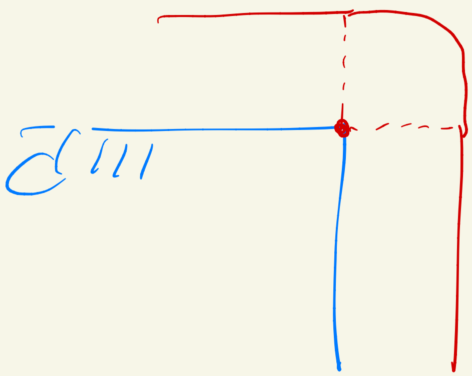
C un ensemble convexe de E e. normé



$d_C : E \to \mathbb{R} : x \longmapsto d_C(x)$

$$\boxed{d_C(x) := \inf_{z\in C} \|x-z\|}$$

distance à 1 ensemble convexe



$\forall x,y \in E \,, \forall t \in [0,1]$

$$d_C((1-t)x+ty) \overset{?}{\leq} (1-t)\, d_C(x) + t\, d_C(y)$$

$z \in C \iff z = (1-t)z_1 + t z_2$ avec $z_i \in C$

$$= \inf_{z_i \in C} \|(1-t)x+ty - [(1-t)z_1 + t z_2]\|$$

$$\underbrace{\|(1-t)(x-z_1) + t(y-z_2)\|}$$

$$\leq \inf_{z_i \in C} \left[(1-t)\|x-z_1\| + t\|y-z_2\|\right]$$

$$\geq \inf_{z\in C} (1-t)\|x-z\| + \inf_{z\in C} t\|y-z\|$$

car $\boxed{\inf_{x\in X}\left[f_1(x)+f_2(x)\right] \geq \inf_{x\in X} f_1(x) + \inf_{x\in X} f_2(x)}$

$$= (1-t)\|x-z_x\| + t\|y-z_y\|$$ avec $z_x, z_y \in C$ (n'∃ pas nécessairement!)

$$\geq \underbrace{(1-t)}_{\geq 0}\inf_{z\in C}\|x-z\| + \underbrace{t}_{\geq 0}\inf_{z\in C}\|y-z\|$$

$$\inf_{(z_1,z_2)\in C} \otimes = \inf_{z_1\in C}\left[\inf_{z_2\in C}\otimes\right]$$

$$= \inf_{z_1\in C}(1-t)\|x-z_1\| + t\inf_{z_2\in C}\|y-z_2\|$$

$$= (1-t)\,d_C(x) + t\,d_C(y)$$

$\boxed{C \subset (1-t)C + tC}$ toujours vrai q.q. soit C

$x = (1-t)x + tx \quad \forall x\in C$

$\boxed{C = (1-t)C + tC}$ vrai si C est convexe

$C \ni (1-t)x + ty \quad \forall x,y\in C$