TD 1 : Introduction et rappels

Ex 1) ${}^t x Q x$ ou $\langle x, Qx \rangle$

$$f(x) = \frac{1}{2} x^T Q x + c^T x + b$$

$Q \in S_n(\mathbb{R})$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}$

* $\nabla f(x) = Qx + c$

* 1ère méthode : avec les dérivées partielles

$$f(x_1, \dots, x_n) = \frac{1}{2} \sum_{i,j} q_{ij} x_i x_j + \sum_i c_i x_i + b$$

$$\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = \sum_{j=1}^{n} q_{ij} x_j + c_i$$

$$\Rightarrow \nabla f(x) = Qx + c$$

* 2ème méthode : avec $Df(x)$

$$f(x+h) = \frac{1}{2}(x+h)^T Q (x+h) + c^T(x+h) + b$$

$$= \underbrace{\frac{1}{2} x^T Q x + c^T x + b}_{f(x)} + h^T Q(x) + c^T h + \underbrace{\frac{1}{2} h^T Q h}_{o(\|h\|)}$$

$\left(\leq \frac{1}{2} |||Q||| \, \|h\|^2 \right)$

$$h^T \underbrace{(Qx + c)}_{\nabla f(x)}$$

* $Hf(x)$?

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = q_{ij}$$

et $Hf(a) = Q$

(la formule de Taylor à l'ordre pour f est exacte).

Remarque : f est une fonction polynomiale de degré 2. Elle est convexe (resp. strictement) ssi $H \geq 0$ (resp. $H \gg 0$)

Ex2)

$$f(x_1, x_2) = \overbrace{100(x_2 - x_1^2)^2 + (1-x_1)^2}^{\geq 0}$$

modèle quadratique approché en $(1,1)$?

(approximation à l'ordre 2)

$\nabla f(1,1)$?

$Hf(1,1)$?

$$\nabla f(x_1, x_2) = \begin{pmatrix} 200(-2x_1)(x_2 - x_1^2) - 2(1-x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}$$

$$\rightarrow \nabla f(1,1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

(pt stationnaire)

$$Hf(x_1, x_2) = \begin{pmatrix} -400(x_2 - x_1^2) + 800x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$$

et

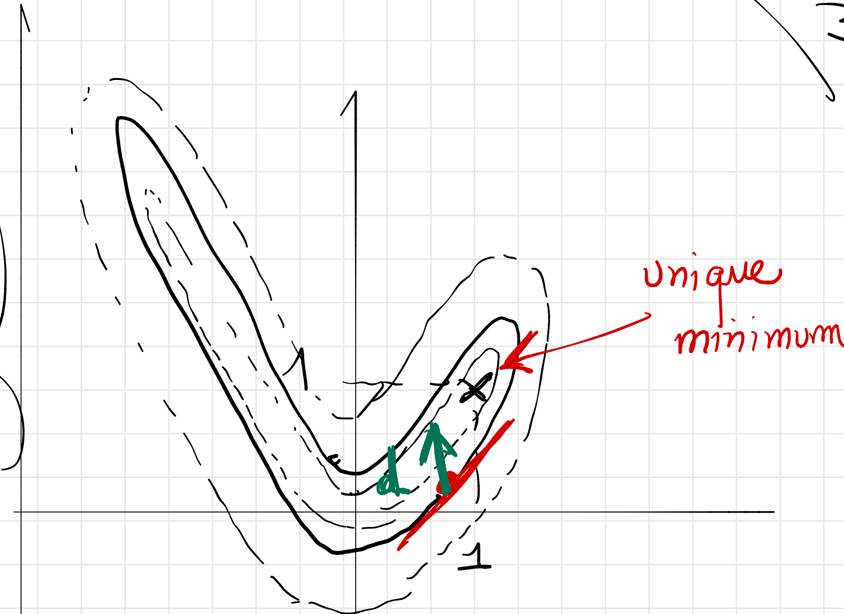
$$Hf(1,1) = \begin{pmatrix} 402 & -400 \\ -400 & 200 \end{pmatrix}$$

Autour de $(1,1)$, on a

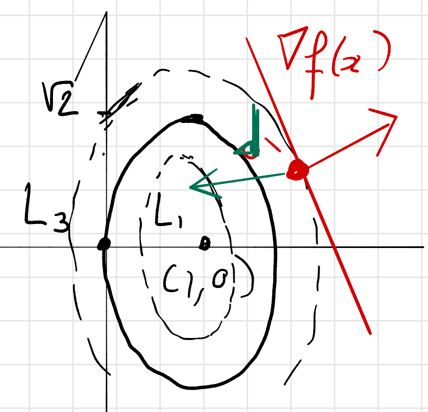
$$f((1,1)+(h_1,h_2)) \approx 0 + \langle 0, h \rangle + \frac{1}{2}(h_1,h_2)\begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

(soit $401h_1^2 - 400h_1h_2 + 100h_2^2$)

Remarque: f s'appelle la fonction de Rosenbrock. Elle possède 1 unique minimum en $(1,1)$. Ses lignes de niveaux ont une forme parabolique et définissent une "vallée" étroite autour du minimum.



Ex 4)

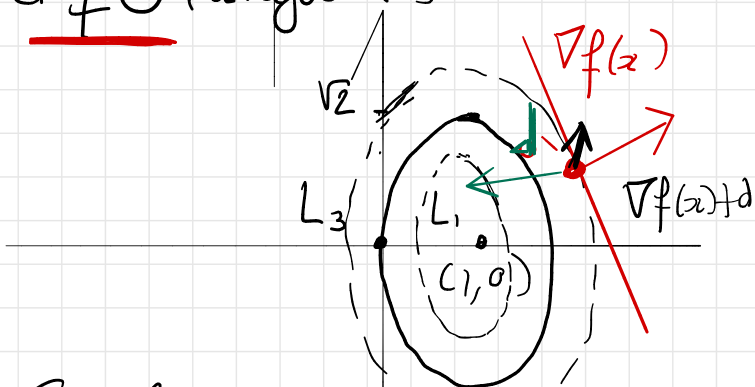


2) $f \in C^1(\mathbb{R}^n, \mathbb{R})$

$$\|\nabla f(x) + d\| \leq \|\nabla f(x)\|$$

d direction de descente en f ?

$d \neq 0$ (à rajouter)



On élève au carré :

$$\|\nabla f(x) + d\|^2 \leq \|\nabla f(x)\|^2$$

$$\Longleftrightarrow \quad \cancel{\|\nabla f(x)\|^2} + 2\langle d, \nabla f(x)\rangle + \|d\|^2 \leq \cancel{\|\nabla f(x)\|^2}$$

soit

$$2\langle d, \nabla f(x)\rangle \leq -\|d\|^2 < 0$$

et d est bien une direction de descente.

3) f C^1 et convexe tq

$f(y) < f(x) \Rightarrow y - x$ direction de descente ?

* f est C^1 et convexe

$$\Longrightarrow f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$$

(f au dessus de ses hyperplans tangents)

$$\Longrightarrow \langle \nabla f(x), y - x\rangle < 0$$

On note $d = y - x$

On a :

$$\underbrace{\langle \nabla f(x), d \rangle}_{<0} = \underbrace{\frac{f(x+sd) - f(x)}{s}} + \underbrace{o(1)}_{\to 0}$$

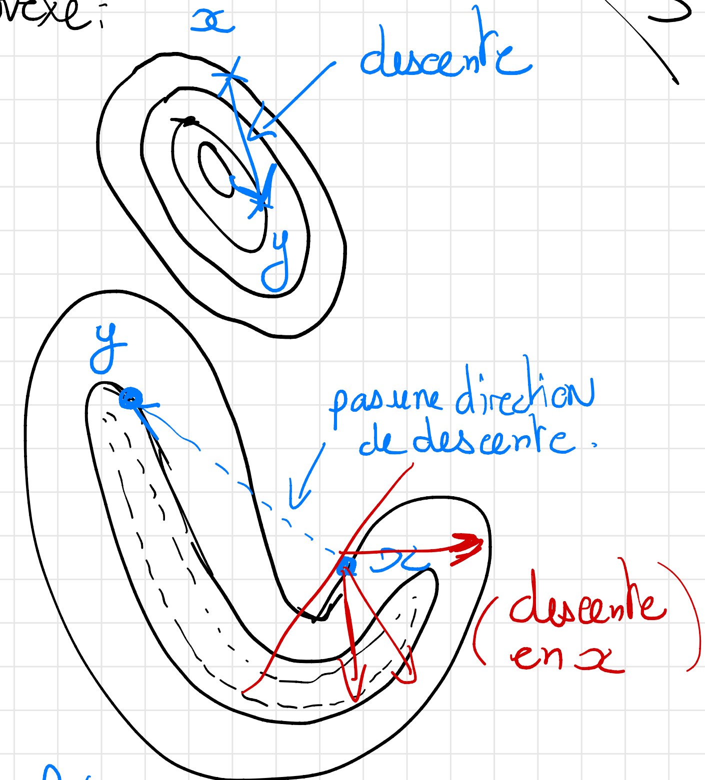
lorsque $s \to 0$, il existe $\varepsilon > 0$

tq $\dfrac{f(x+sd) - f(x)}{s} < 0$

si $s \in]0, \varepsilon]$

$d = y - x$ est bien une direction de descente.

* f convexe :



$f(y) < f(x)$