229

Fonctions monotones, fonctions convexes, exemples et applications

[Rom] p 83

**Cadre** : On se placera sur $I$ un intervalle de $\mathbb{R}$

# I) FONCTIONS MONOTONES

## 1) Définitions et propriétés des fonctions monotones

**Def 1** Soit $f : I \to \mathbb{R}$ une fonction, $f$ est :
- croissante si $\forall x, y \in I$, $x \geq y \Rightarrow f(x) \geq f(y)$
- décroissante si $\forall x, y \in I$, $x \geq y \Rightarrow f(x) \leq f(y)$
- monotone sur $I$ si $f$ est croissante sur $I$ ou décroissante sur $I$

**NB** Si les inégalités sont strictes, la fonction est dite strictement croissante/décroissante/monotone

**Prop 2** Soient $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$
1. $f$ et $g$ ont même monotonie $\Rightarrow f + g$ aussi
2. Soit $\lambda \geq 0$, $f$ monotone $\Rightarrow \lambda f$ a la même monotonie
   $\Rightarrow -\lambda f$ de monotonie inverse
3. $f : J \to I$, $g : I \to \mathbb{R}$ monotones
   Si $f$ et $g$ ont même monotonie, $f \circ g$ est croissante
   Sinon $f \circ g$ est décroissante

**Ex 3** $\begin{array}{l}\mathbb{R} \to \mathbb{R} \\ x \mapsto x\end{array}$, $\begin{array}{l}\mathbb{R} \to \mathbb{R} \\ x \mapsto \lambda x\end{array}$, $\lambda \geq 0$

**C-Ex 4** $\begin{array}{l}\mathbb{R} \to \mathbb{R} \\ x \mapsto x\end{array}$ monotone mais pas $\begin{array}{l}\mathbb{R} \to \mathbb{R} \\ x \mapsto x^2\end{array}$

**Prop 5** Soit $f : I \to \mathbb{R}$, de classe $C^1$ sur $I$
$f'$ est de signe constant $\Leftrightarrow f$ est monotone

**Prop 6** Soit $f : I \to \mathbb{R}$ continue
$f$ injective $\Leftrightarrow f$ strictement monotone

**Thm 7**
1. Soit $f : I \to \mathbb{R}$ croissante
   $f$ continue $\Leftrightarrow f(I)$ est un intervalle
2. Soit $f : I \to J$
   $f$ est une bijection continue $\Leftrightarrow f^{-1}$ est continue

**Ex 8** $\exp : \mathbb{R} \to \mathbb{R}^{+*}$, $\exp(\mathbb{R}) = \mathbb{R}^{+*}$
$\exp^{-1} = \ln$ est continue sur $\mathbb{R}^{+*}$

## 2) Régularités

[Rom] p 43-44

**Thm 9** Soit $f : I \to \mathbb{R}$ croissante alors :
$f$ admet une limite à gauche ($f_g$) et à droite ($f_d$) en tout point
Précisément : $\boxed{\forall x \in I,\ f(x^-) = \sup_{a < t < x} f(t) \leq f(x) \leq f(x^+) = \inf_{x < t < b} f(t)}$

$\forall x < y \quad f(x^-) \leq f(y^-)$

**Def 10** Si $\alpha \in \mathring{I}$, si $f : I \to \mathbb{R}$ est discontinue en $\alpha$, avec des $f_d$ et $f_g$ en $\alpha$, alors $f$ a une discontinuité de $1^{ère}$ espèce en $\alpha$.

**Ex 11** $x \mapsto [x]$

**Cor 12** Une fonction monotone $f : I \to \mathbb{R}$ ne peut avoir que des discontinuités de 1ère espèce

**Thm - Cor 13** Soit $f : I \to \mathbb{R}$ monotone alors l'ensemble de ces points de discontinuité est dénombrable

[Rom] p 86 [Hau] p 200

**Thm 14** [ADMIS] [Lebesgue]
Une fonction monotone est dérivable presque partout

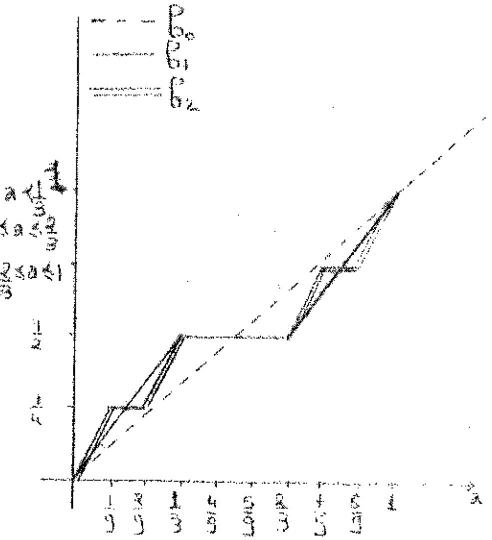
[Hau] p 193-196

**EX 15**

$f_n : [0,1] \to \mathbb{R}$

$f_0 : x \mapsto x$

$f_{n+1} : [0,1] \to \mathbb{R}$
$x \mapsto \begin{cases} \frac{1}{2} f_n(3x) & \text{si } 0 \leq x \leq \frac{1}{3} \\ \frac{1}{2} & \text{si } \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{1}{2} f_n(3x-2) + \frac{1}{2} & \text{si } \frac{2}{3} \leq x \leq 1 \end{cases}$



## 3) Monotonie et suite de fonctions

[FGN An2 p156] [Gou p231] [FGN An2 p156]

**Def 16** Soit $X$ un ensemble, $(E,d)$ un esp métrique, $(f_n)_n$ une suite de fonctions de $X$ dans $E$

- $(f_n)$ CV simplement sur $X$ vers $f : X \to E$ si :
$\forall \varepsilon > 0, \forall x \in X, \exists N \in \mathbb{N}, \forall n \geq N, d(f_n(x), f(x)) < \varepsilon$
- $(f_n)$ CV uniformément sur $X$ vers $f : X \to E$ si :
$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \forall x \in X, d(f_n(x), f(x)) < \varepsilon$

**Thm 17 [Dini]**

Soit $(f_n)_n$ suite ~~croissante~~ de fonctions continues $f_n : [a,b] \to \mathbb{R}$ qui CVS vers une ~~f~~ continue $f$

- Si la suite $(f_n)_{n \geq 0}$ est croissante, la cvence est uniforme

Si les $f_n$ sont des fonctions croissantes, continue (Si $f_n$ CVS vers $f : I \to \mathbb{R}$ continue, elle y CVU) la cvence est uniforme

**Ex 18** $(f_n)_{n \geq 0}$ suite de fonctions $f_n : [0,1] \to \mathbb{R}$ :

$f_0 = 0, \quad f_{n+1}(x) = f_n(x) + \frac{1}{2}(x - f_n^2(x))$

[FGN An 2]

**Thm 19 [Helly]**

Soit $(f_n)_{n \geq 0}$ une suite de fonctions, $f_n : \mathbb{R} \to [-1,1]$ tq les $f_n$ soient croissantes sur $\mathbb{R}$.

Il existe une sous-suite de $(f_n)_{n \geq 0}$ qui converge simplement vers une fonction $f : \mathbb{R} \to [-1,1]$

|DEV 1|

## II) FONCTIONS CONVEXES

### 1) Définitions et propriétés

[Go p246-252-253]

**Def 20** $f : I \to \mathbb{R}$, $f$ est convexe sur $I$ si

$$\boxed{\forall x, y \in I, \forall \lambda \in [0,1], \quad f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)}$$

Dessin : Annexe 1 et 2.

**Rem 21** : Epigraphe $\Gamma(f)$ :

$$\Gamma(f) = \{(x,t) \in I \times \mathbb{R}, \ t \geq f(x)\}$$

La convexité de $f$ équivaut à celle de $\Gamma(f)$

convexité de $\Gamma(f)$ : si $M, N \in \Gamma(f)$, $[MN] \subset \Gamma(f)$

**Ex 22** $x \mapsto |x|$ convexe sur $\mathbb{R}$.

**Rem 23** Si $f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda) f(y)$, $f$ est concave

**Thm 24** Soit $f : I \to \mathbb{R}$, alors

En tout point $x$ de $I$, $f$ admet une dérivée à gauche (dàg) $f'_g(x)$ et une à droite (dàd) $f'_d(x)$.

$f'_d$ et $f'_g$ sont croissantes sur $I$ et si $x < y$ :

$$\boxed{f'_g(x) \leq f'_d(x) \leq \frac{f(y) - f(x)}{y - x} \leq f'_g(y) \leq f'_g(y)}$$

**Application 25** $I = ]a,b[$, $f : I \to \mathbb{R}$ convexe

alors $f$ est continue sur $]a,b[$

**Thm 26** $I = ]a,b[$, $f : I \to \mathbb{R}$ dérivable sur $I$, EQU :

1. $f$ est convexe
2. $f'$ est croissante sur $I$
3. la courbe représentative de $f$ se situe au-dessus de toutes ses tangentes ie

$$\forall x, y \in I, \ f(x) \geq f(y) + f'(y)(x-y)$$

**Ex 27** : exp.

**Rem 28** : Si $f$ est dérivable deux fois, $f$ convexe $\Leftrightarrow f'' \geq 0$

App$^{lem}$ 29 de (3) : $\forall x \in \mathbb{R},\ e^x \geq 1+x$, $\forall x \geq -1,\ \sqrt{1+x} \leq 1+\frac{x}{2}$

## 2) Dimension supérieure

**Déf 30** Si $E$ est un evn
$U$ est un ouvert convexe de $E$
$f : U \to \mathbb{R}$ convexe si $\forall x,y \in E, \forall \lambda \in [0,1]$

$$\boxed{f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)}$$

**Rem 31** $E \Leftrightarrow$ : $\forall x_1, \dots, x_n \in U$, $\forall \lambda_1, \dots, \lambda_n \in \mathbb{R}$ tq $\sum_{i=1}^n \lambda_i = 1$ :

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i)$$

**Ex 32** : Annexe 3

**Thm 32** Soit $U$ ouvert non vide convexe d'un $\mathbb{R}$-ev $E$ normé
$f : U \to \mathbb{R}$ convexe.

Admis

1) Si $f$ est majorée sur une partie ouvert non vide de $U$, elle est continue sur $U$
2) Si $\dim E < +\infty$, $f$ est continue sur $U$

**C-EX 34** En dimension infinie

$\Phi : E := \mathcal{C}^0([0,1], \mathbb{R}) \to \mathbb{R}$, $f \mapsto f(0)$ est convexe non continue sur $(E, \|\cdot\|_{L^1})$ mais continue sur $(E, \|\cdot\|_\infty)$

# III) APPLICATIONS

## 1) Inégalités de convexité

**Inégalité 35** [arithmético-géo]
$\forall x_1, \dots, x_n \in \mathbb{R}^+$

$$\sqrt[n]{x_1 \cdots x_n} \leq \frac{1}{n}(x_1 + \cdots + x_n)$$

**Inégalité 36** (Hölder)
Soient $(a_1, \dots, a_n) \in (\mathbb{R}^+)^n$, $(b_1, \dots, b_n) \in (\mathbb{R}^+)^n$, $p, q \in \mathbb{R}$, $p,q > 1$ tq $\frac{1}{p} + \frac{1}{q} = 1$

$$\sum_{i=1}^n a_i b_i \leq \left(\sum_{i=1}^n a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n b_i^q\right)^{\frac{1}{q}}$$

**Inégalité 37** (Minkowski)

$$\left(\sum_{i=1}^n (a_i + b_i)^p\right)^{1/p} \leq \left(\sum_{i=1}^n a_i^p\right)^{1/p} + \left(\sum_{i=1}^n b_i^p\right)^{1/p}$$

[Rom p 108-113]

**Inégalité 38** (Jensen)

$\Phi : I \to \mathbb{R}$ convexe continue
$f : [0,1] \to \mathbb{R}$ réglée

$$\int_0^1 \Phi \circ f \geq \Phi\left(\int_0^1 f\right)$$

**Rappel 39** : une fonction $f : [a,b] \to E$ est dite réglée si elle est limite uniforme sur $[a,b]$ de fonctions en escalier.

## 2) Optimisation

**Application 40** [Méthode du gradient à pas optimal]

[Cia]

Soit $f : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$, $A \in S_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$.

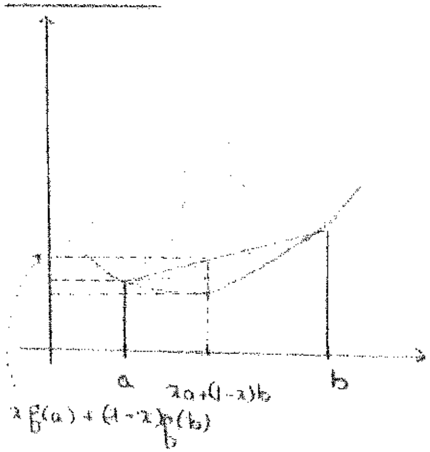
(P) : Minimiser $f(x)$, $x \in \mathbb{R}^n$

1) (P) a une et une seule solution $\bar{x}$ qui est l'unique solution de $\nabla f(x) = 0$
2) On proposera une méthode pour approcher $\bar{x}$

DEV2

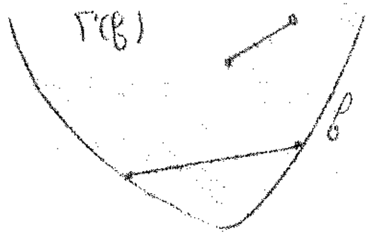
[Tes]

## Annexe 1

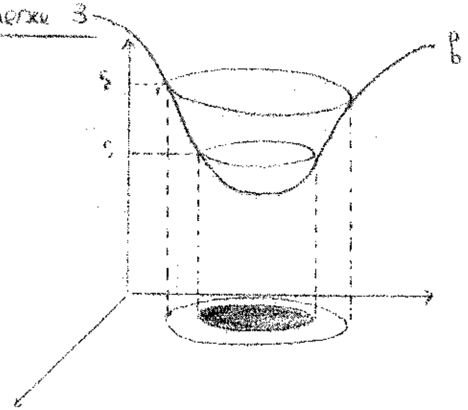



[Bec p27]

## Annexe 2




## Annexe 3




Références :

[Mar] : J-P Marco, Mathématiques Analyse L3
[Pom] : Pommelet, cours d'analyse

[Rom] : [Rombaldi] : Elts d'analyse réelle

[Hau] : Hauchecorne, les contre exemples en mathématiques, 2e ed.

[Gou] : Gourdon, les maths en tête, Analyse 2e ed

[FGN] : Oraux X-ENS

[Tes] : Testard, Analyse de l'implicite

Emily CLEMENT
Julien GABET

# Méthode du gradient à pas optimal

Emily Clement

3 octobre 2016

Référence : Optimisation et analyse convexe, Jean-Baptiste Hiriat-Urruty et Oraux X-ENS Algèbre 3, exercice 2.35

Prérequis : Inégalité de Kantorovitch, la (les...) démonstrations sont à la fin.

**Proposition .1.**

On pose la fonction $f : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c \end{array}$

où $A \in S_n^{++}(\mathbb{R}), b \in \mathbb{R}^n, c \in \mathbb{R}$

$$(\mathcal{P}) \quad \text{Minimiser } f(x), x \in \mathbb{R}^n$$

1. $(\mathcal{P})$ admet une **unique** solution $\overline{x}$, caractérisée comme étant l'unique solution de l'équation $\nabla f(x) = 0$
2. On proposera la méthode du gradient à pas optimal pour approcher $\overline{x}$ : On construit une suite $(x_k)$ par récurrence :

$$\begin{cases} x_0 \in \mathbb{R}^n & \text{quelconque} \\ x_{k+1} := x_k + t_k d_k & \text{quand } \nabla f(x) \neq 0 \end{cases}$$

où $d_k := -\nabla f(x_k)$ et $t_k \in \mathbb{R}^+$ est **l'unique réel positif minimisant** $t \mapsto f(x_k + td_k)$ sur $\mathbb{R}$.

**Remarque .1.** *La méthode converge rapidement pour $c_2(A)$ proche de 1, lentement pour $c_2(A)$ grand.*

*Démonstration.*

1. Montrons que $f$ admet un unique minimum, et caractérisons le.

**Différentiabilité de $f$ et calcul de $\nabla f(x)$, de $\nabla^2 f(x)$** $f$ est quadratique, polynomiale en $x$ donc différentiable deux fois sur $\mathbb{R}^n$

$$\begin{aligned} f(x+h) &= \frac{1}{2}\langle A(x+h), x+h\rangle + \langle b, x+h\rangle + c \\ &= \frac{1}{2}\langle Ax, x\rangle + \frac{2}{2}\langle Ax, h\rangle + \frac{1}{2}\langle Ah, h\rangle + \langle b, x\rangle + \langle b, h\rangle + c \\ &= f(x) + \langle Ax, h\rangle + \frac{1}{2}\langle Ah, h\rangle + \langle b, h\rangle \\ &= f(x) + \frac{1}{2}\langle Ah, h\rangle + \langle Ax+b, h\rangle \end{aligned}$$

Donc comme, si $f$ est deux fois différentiable en $x$, on a le développement de Taylor-Young d'ordre deux suivant :

$$f(x+h) = f(x) + \langle \nabla f(x), h\rangle + \frac{1}{2}\left\langle \nabla^2 f(x)h, h\right\rangle + \|h\|^2 \varepsilon(h)$$

où $\varepsilon(h) \xrightarrow[h\to 0]{} 0$

on a bien ici que $\nabla^2 f(x) = A$ et $\nabla f(x) = Ax + b$

**Convexité de $f$** : $\forall x, y \in \mathbb{R}^n, \forall \lambda \in [0,1]$ :

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \frac{\lambda^2}{2}\langle Ax, x\rangle + \lambda(1-\lambda)\langle Ax, y\rangle + \frac{(1-\lambda)^2}{2}\langle Ay, y\rangle \\ &\quad + \lambda(\langle b, x\rangle + c) + (1-\lambda)(\langle b, y\rangle + c) \\ &= \frac{\lambda}{2}\langle Ax, x\rangle - \frac{\lambda(1-\lambda)}{2}\langle Ax, x\rangle \\ &\quad + \frac{1-\lambda}{2}\langle Ay, y\rangle - \frac{\lambda(1-\lambda)}{2}\langle Ay, y\rangle \\ &\quad + \lambda(1-\lambda)\langle Ax, y\rangle + \lambda(\langle b, x\rangle + c) + (1-\lambda)(\langle b, y\rangle + c) \\ &= \lambda f(x) + (1-\lambda)f(y) - \frac{\lambda(1-\lambda)}{2}\langle A(x-y), x-y\rangle \\ &\leqslant \lambda f(x) + (1-\lambda)f(y) \end{aligned}$$

car $A \in S_n^{++}$ et $0 < \lambda < 1$,
d'où

$$\boxed{f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)}$$

Ce qui montre la convexité de $f$.

**Coercivité de** $f$ : Soit $x \in \mathbb{R}^n$, alors :

$$
\begin{aligned}
|f(x)| &= \left|\frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c\right| \\
&\geqslant \left|\frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle\right| - |c| \\
&\geqslant \frac{1}{2}|\langle Ax, x\rangle| - |\langle b, x\rangle| - |c| \\
&\geqslant \frac{1}{2}|\langle Ax, x\rangle| - \|b\|\|x\| - |c| \\
&\geqslant \frac{1}{2}\lambda_{\min}\|x\|^2 - \|b\|\|x\| - |c| \rightarrow +\infty \text{ quand } \|x\| \rightarrow +\infty
\end{aligned}
$$

Il existe un et un seul $\overline{x} \in \mathbb{R}^n$ minimisant $f$ sur $\mathbb{R}^n$.
Comme $f$ est différentielle et convexe, $\overline{x}$ est l'unique solution de l'équation $\nabla f(x) = 0$
La solution exacte de $(\mathcal{P})$ est donc $x = A^{-1}b$ car $\nabla f(x) = Ax + b$ ($A$ est symétrique définie positive donc inversible).
N.B : $f(\overline{x}) = -\frac{1}{2}\langle A^{-1}b, b\rangle + c$

2. En ce qui concerne la méthode du gradient à pas optimal pour approcher la solution, on va tout d'abord montrer que :
   (a)
   $$f(x_{k+1}) - f(\overline{x}) = [f(x_k) - f(\overline{x})]\left[1 - \frac{\|d_k\|^4}{\langle Ad_k, d_k\rangle\langle A^{-1}d_k, d_k\rangle}\right]$$
   (b) Puis montrer que, en posant $(\lambda_i)_{i\in[\![1,n]\!]}$ les valeurs propres de $A$, rangées dans l'ordre décroissant et $c_2(A) = \frac{\lambda_1}{\lambda_n}$ le *conditionnement* de $A$, que :
   $$f(x_k) - f(\overline{x}) \leqslant [f(x_0) - f(\overline{x})]\left[\frac{c_2(A) - 1}{c_2(A) + 1}\right]^{2k}$$
   $$\|x_k - \overline{x}\| \leqslant \left[\frac{2f(x_0) - f(\overline{x})}{\lambda_n}\right]^{\frac{1}{2}}\left[\frac{c_2(A) - 1}{c_2(A) + 1}\right]^{k}$$

   Et ainsi conclure que la rapidité de convergence de la méthode.
   (a) On sait que $d_k := -\nabla f(x_k)$, comme $x_{k+1} = x_k + t_k d_k$.
   **Caractérisons** $t_k$ **:**
   $$
   \begin{aligned}
   f(x_k + td_k) &= \frac{1}{2}\langle Ax_k + tAd_k, x_k + td_k\rangle + \langle b, x_k + td_k\rangle + c \\
   &= \frac{1}{2}t^2\langle Ad_k, d_k\rangle + t\langle Ax_k + b, d_k\rangle + f(x_k)
   \end{aligned}
   $$

Cette fonction, qui est un polynôme d'ordre deux en $t$, admet un minimum qui est :

$$\boxed{t_k = \frac{\|d_k\|^2}{\langle Ad_k, d_k \rangle}} > 0$$

Trouvons une relation entre $d_{k+1}$ et les indices précédents :

$$\begin{aligned} d_{k+1} &= -\nabla f(x_{k+1}) \\ &= -A(x_k + t_k d_k) - b \\ &= -(Ax_k + b) - t_k A d_k \\ &= d_k - t_k A d_k \end{aligned}$$

$$\boxed{d_{k+1} = d_k - t_k A d_k}$$

Donc on a immédiatement :

$$\langle d_{k+1}, d_k \rangle = \|d_k\|^2 - \langle d_k, Ad_k \rangle \frac{\|d_k\|^2}{\langle Ad_k, d_k \rangle} = 0$$

$$\boxed{\langle d_{k+1}, d_k \rangle = 0}$$

Combinons tout cela dans $f$ :

$$\begin{aligned} f(x_{k+1}) &= \frac{1}{2} t_k^2 \langle Ad_k, d_k \rangle + t_k \underbrace{\langle Ax_k + b, d_k \rangle}_{=\|d_k\|^2} + f(x_k) \\ &= \frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle} - \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle} + f(x_k) \\ &= f(x_k) - \frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle} \end{aligned}$$

$$\boxed{f(x_{k+1}) = -\frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle} + f(x_k)}$$

d'où le résultat cherché.

$$\boxed{f(x_{k+1}) - f(\overline{x}) = [f(x_k) - f(\overline{x})] \left[ 1 - \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle 2 (f(x_k) - f(\overline{x}))} \right]}$$

Montrons l'égalité finale :

$$\begin{aligned} \langle A^{-1} d_k, d_k \rangle &= \langle A^{-1}(Ax_k + b), Ax_k + b \rangle \\ &= 2 \left[ \frac{1}{2} \langle Ax_k, x_k \rangle + \langle b, x_k \rangle + \frac{1}{2} \langle A^{-1} b, b \rangle \right] \\ &= 2 [f(x_k) - f(\overline{x})] \end{aligned}$$

$$\boxed{f\left(x_{k+1}\right)-f\left(\overline{x}\right)=\left[f\left(x_k\right)-f\left(\overline{x}\right)\right]\left[1-\frac{\|d_k\|^4}{\langle Ad_k,d_k\rangle\langle A^{-1}d_k,d_k\rangle}\right]}$$

(b) On pose $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ les vp de $A$, $c_2 := \dfrac{\lambda_1}{\lambda_n}$ (conditionnement de $A$).

**Lemme .1** (Inégalité de Kantorovitch).

Soit $A$ une matrice symétrique définie positive, de valeurs propres $\lambda_1 \geqslant \cdots \geqslant \lambda_n$, alors : $\forall x \in \mathbb{R}^n$ :

$$\|x\|^2 \leqslant \langle Ax,x\rangle\left\langle A^{-1}x,x\right\rangle \leqslant \frac{1}{4}\left[\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right]^2\|x\|^4$$

$$\begin{aligned}
\frac{\|d_k\|^4}{\langle Ad_k,d_k\rangle\langle A^{-1}d_k,d_k\rangle} &\geqslant \frac{\|d_k\|^4}{\frac{1}{4}\left[\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right]^2\|d_k\|^4}\\
&\geqslant \frac{4}{\left[\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right]^2}\\
&\geqslant \frac{\frac{\lambda_1}{\lambda_n}}{\left(\frac{\lambda_1}{\lambda_n}+1\right)^2}\\
&\geqslant \frac{c_2\left(A\right)}{\left(c_2\left(A\right)+1\right)^2}
\end{aligned}$$

En injectant dans la dernière égalite obtenue :

$$\begin{aligned}
f\left(x_{k+1}\right)-f\left(\overline{x}\right) &\leqslant \left[f\left(x_k\right)-f\left(\overline{x}\right)\right]\left[1-4\frac{c_2\left(A\right)}{\left(c_2\left(A\right)+1\right)^2}\right]\\
&\leqslant \left[f\left(x_k\right)-f\left(\overline{x}\right)\right]\left[\frac{c_2\left(A\right)-1}{\left(c_2\left(A\right)+1\right)}\right]^2
\end{aligned}$$

Donc :

$$\boxed{f\left(x_k\right)-f\left(\overline{x}\right)\leqslant\left[f\left(x_0\right)-f\left(\overline{x}\right)\right]\left[\frac{c_2\left(A\right)-1}{\left(c_2\left(A\right)+1\right)}\right]^{2k}}$$

Lien avec $\|x_k - \overline{x}\|$ ?

mal rédigé

$$\begin{cases} f(x_k) - f(\overline{x}) = \frac{1}{2}\langle Ax_k, x_k\rangle + \langle b, x_k\rangle + c - f(\overline{x}) \\ = \frac{1}{2}\langle Ax_k, x_k\rangle + \frac{2}{2}\langle(-A\overline{x}), x_k\rangle + c - \left(-\frac{1}{2}\langle A^{-1}b, b\rangle + c\right) \\ = \frac{1}{2}\left[\langle A(x_k - \overline{x}), x_k\rangle + \langle A(-\overline{x}), x_k\rangle + \langle A^{-1}b, b\rangle\right] \\ = \frac{1}{2}\langle A(x_k - \overline{x}), x_k - \overline{x}\rangle \end{cases}$$

car, comme $-b = A\overline{x}$ :

$$\langle A^{-1}b, b\rangle = \langle \overline{x}, A\overline{x}\rangle$$

D'où :

$$\boxed{f(x_k) - f(\overline{x}) \geqslant \frac{1}{2}\lambda_n\|x_k - \overline{x}\|}$$

□

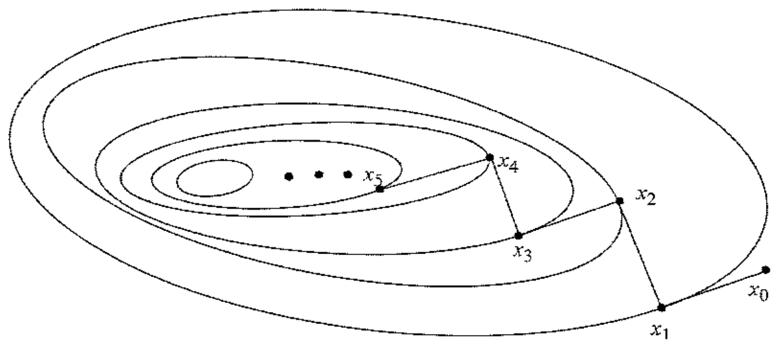
**Remarque .2** (Un mot sur la convergence). *On voudrait :*

$$\frac{f(x_k) - f(\overline{x})}{f(x_0) - f(\overline{x})} < \varepsilon (\textit{ une convergence rapide en fait})$$

*Il faudrait :*

$$k \geqslant \frac{\ln \varepsilon}{2\ln\left(\frac{c_2(A)-1}{c_2(A)+1}\right)} \sim c_2(A)/4\ln\left(\frac{1}{\varepsilon}\right)$$

**Exemple .1** (Cas $n = 2$). *Pour le cas $n = 2$, les courbes de $f$ ressemblent à des ellipses très effilées*




*Convergence de $(x_k)_k$ : lente, en zigzag.*

**Lemme .2** (Inégalité de Kantorovitch).

Soit $A$ une matrice symétrique définie positive, de valeurs propres $\lambda_1 \geqslant \cdots \geqslant \lambda_n$, alors : $\forall x \in \mathbb{R}^n$ :

$$\|x\|^2 \leqslant \langle Ax, x\rangle \langle A^{-1}x, x\rangle \leqslant \frac{1}{4}\left[\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right]^2 \|x\|^4$$

*Démonstration du lemme de Kantorovich via l'étude de fonction.*
C'est la manière la plus courte de procéder.

**Tronc commun aux deux preuves** Tout d'abord, on se ramène au cas $\|x\| = 1$ en remarquant que :

$$\left\langle A\frac{x}{\|x\|}, \frac{x}{\|x\|}\right\rangle \left\langle A^{-1}\frac{x}{\|x\|}, \frac{x}{\|x\|}\right\rangle = \frac{1}{\|x\|^4}\langle Ax, x\rangle \langle A^{-1}x, x\rangle$$

On doit donc montrer que, pour $x \in \mathbb{R}^n$ tel que $\|x\| = 1$ :

$$1 \leqslant \langle Ax, x\rangle \langle A^{-1}x, x\rangle \leqslant \frac{1}{4}\left[\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right]^2$$

**Inégalité de gauche : Minoration** On écrit les coordonnées de $x = (x_1, \cdots, x_n)^t$ dans une b.o.n. de vecteurs propres ($A \in S_n(\mathbb{R})$ donc on applique le théorème Spectral), on obtient alors :

$$\begin{aligned}
(*)\ \langle Ax, x\rangle \langle A^{-1}x, x\rangle &= \left(\sum_{i=1}^n \lambda_i x_i^2\right)\left(\sum_{i=1}^n \frac{1}{\lambda_i} x_i^2\right) \\
&= \sum_{i=1}^n \left(\sqrt{\lambda_i} x_i\right)^2 \sum_{i=1}^n \left(\frac{x_i}{\sqrt{\lambda_i}}\right)^2 \quad \text{[Cauchy-Schwarz] :} \\
&\geqslant \left(\sum_{i=1}^n \left(\sqrt{\lambda_i} x_i\right)\left(\frac{x_i}{\sqrt{\lambda_i}}\right)\right)^2 \\
&\geqslant \left(\sum_{i=1}^n x_i^2\right)^2 = 1
\end{aligned}$$

**Inégalité de droite : Majoration**

Outil : $\forall a, b \geqslant 0$, $\sqrt{ab} \leqslant \frac{1}{2}(a+b)$, en factorisant l'expression $(*)$ par

$\frac{\lambda_1}{\lambda_n}$, le bouzin devient :

$$\begin{aligned}
\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} &= \sqrt{\frac{\lambda_1}{\lambda_n}\left(\sum_{i=1}^{n}\frac{\lambda_i}{\lambda_1}x_i^2\right)\left(\sum_{i=1}^{n}\frac{\lambda_n}{\lambda_i}x_i^2\right)} \\
&\text{[Application de l'outil] :} \\
&\leqslant \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}\left(\sum_{i=1}^{n}\frac{\lambda_i}{\lambda_1}x_i^2 + \left(\sum_{i=1}^{n}\frac{\lambda_n}{\lambda_i}x_i^2\right)\right) \\
&= \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}\left(\sum_{i=1}^{n}\left(\frac{\lambda_i}{\lambda_1}+\frac{\lambda_n}{\lambda_i}\right)x_i^2\right)
\end{aligned}$$

Ensuite, on fait une gentille étude de la fonction bien posée :

$$f : \begin{array}{ccc} [\lambda_1, \lambda_n] & \to & \mathbb{R} \\ x & \mapsto & \frac{x}{\lambda_1} + \frac{\lambda_n}{x} \end{array}$$

— $f$ croissante sur $\left[\lambda_1, \sqrt{\lambda_1, \lambda_n}\right]$

— $f$ décroissante sur $\left[\sqrt{\lambda_1\lambda_n}, \lambda_n\right]$

Maximum de $f$ : en $\lambda_1$ ou en $\lambda_n$ (Au tableau : faire un tableau de variation, mais le lecteur comprendra que sous LaTeX, non.) or $f(\lambda_1) = f(\lambda_n) = 1 + \frac{\lambda_1}{\lambda_n}$ donc égalité, match nul, on retourne à la majoration, on peut faire sortir tous les $\lambda_i$ de la somme, donc on le fait :

$$\begin{aligned}
\sqrt{\langle Ax, x\rangle\langle A^{-1}x, x\rangle} &= \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}\left(1+\frac{\lambda_n}{\lambda_1}\right)\left(\sum_{i=1}^{n}x_i^2\right) \\
&\leqslant \frac{1}{2}\left(\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right)\sum_{i=1}^{n}x_i^2 \\
&\leqslant \frac{1}{2}\left(\sqrt{\frac{\lambda_1}{\lambda_n}}+\sqrt{\frac{\lambda_n}{\lambda_1}}\right)
\end{aligned}$$

C'est donc ce que l'on voulait.

□

*Démonstration du lemme de Kantorovich via étude des barycentres.*
L'intérêt de cette preuve est là, certes, mais si le lecteur veut le mettre en

développement, celui du haut est plus gérable. Par contre celui là utilisant les barycentre non improvisable, il vaut mieux l'avoir en tête/avoir la vision géométrique n'est pas du luxe.
Comme dans la preuve précédente : on reste avec $\|x\| = 1$ et on exprimera nos vecteur dans la b.o.n. de vecteurs propres : $\exists P \in \mathcal{O}_n(\mathbb{R})$, $\Delta := \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ :

$$A =^t P\Delta P, \text{ donc } A^{-1} =^t P\Delta^{-1}P$$

On obtient :

$$\langle Ax, x\rangle = \langle {}^tP\Delta Px, x\rangle = \langle \Delta(Px), Px\rangle, \langle A^{-1}x, x\rangle = \langle \Delta^{-1}(Px), Px\rangle$$

Posons :

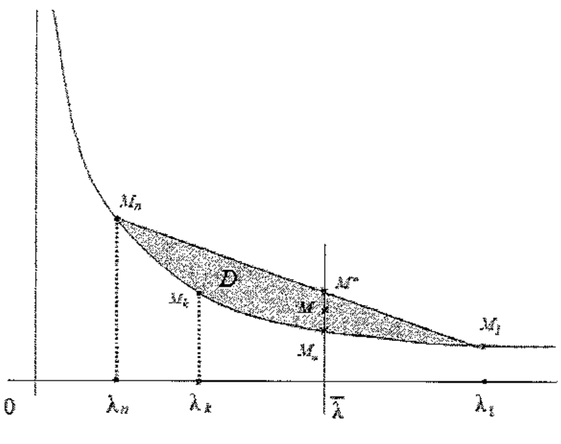
$$P : \begin{array}{ccc} = \{x, \|x\| = 1\} & \to & \\ x & \mapsto & y = Px \end{array}$$

Cela revient à ce qu'on a fait au début de la première preuve...

$P$ est une bijection de sur elle même (le vérifier), donc on peut étudier que $\forall y = (y_1, \cdots y_n) \in$ :

$$1 \leqslant \langle \Delta y, y\rangle \cdot \langle \Delta^{-1}y, y\rangle \leqslant \frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2$$

On pose la fonction $\delta : \theta \mapsto \dfrac{1}{\theta}$, son ~~épi~~graphe :




$M_k = \left(\lambda_k, \dfrac{1}{\lambda_k}\right) = (\lambda_k, \theta(\lambda_k))$ pour $1 \leqslant k \leqslant n$.
On pose $\alpha_i := y_i^2$, on a la combinaison convexe des $\lambda_i$ et $\lambda_i^{-1}$ :

$$\langle \Delta y, y\rangle = \sum_{i=1}^{n} y_i^2 \lambda_i, \langle \Delta^{-1}y, y\rangle = \sum_{i=1}^{n} y_i^2 \lambda_i^{-1}$$

On a posé $M = (a, b)$ le barycentre des $(M_k, \alpha_i)$,$M$ et $M_1M_n$ sont dans $D$, le domaine convexe.
On pose le point $M_* = (a, \theta(a))$ le point du graphe $\theta$ de même abscisse et $M^*$ l'intersection de $M_1M_*$ et de $M_1M_n$.
Ainsi $b = \sum_{i=1}^{n} \alpha_i \frac{1}{\lambda_i}$ et $\theta(a) = \frac{1}{\sum_{i=1}^{n} \alpha_i \lambda_i} =: \frac{1}{\overline{\lambda}}$
On peut voir l'ordonnée de $M^*$ comme un point de la droite $M_1M_n$ ce qui donne :

$$\theta(a) = \frac{1}{\lambda_1} + \frac{1}{\lambda_n} - \frac{\overline{\lambda}}{\lambda_1\lambda_n}$$

On a par le dessin/convexité/construction que l'ordonnée de $M$ est supérieure à celle de $M_*$ :

$$\sum_{i=1}^{n} \alpha_i \frac{1}{\lambda_i} \geqslant \frac{1}{\overline{\lambda}}$$

Ce qui montre en somme la minoration.
L'ordonne de $M^*$ est supérieure à celle de $M$ donc :

$$\sum_{i=1}^{n} \alpha_i \frac{1}{\lambda_i} \leqslant \frac{1}{\lambda_1} + \frac{1}{\lambda_n} - \frac{\overline{\lambda}}{\lambda_1\lambda_n}$$

Donc :

$$\left(\sum_{i=1}^{n} \alpha_i \lambda_i\right)\left(\sum_{i=1}^{n} \alpha_i \frac{1}{\lambda_i}\right) \overline{\lambda}\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_n} - \frac{\overline{\lambda}}{\lambda_1\lambda_n}\right) = \frac{\overline{\lambda}\left(\lambda_1 + \lambda - \overline{\lambda}\right)}{\lambda_1\lambda_n}$$

Idem après une étude de la fonction

$$u \quad \mapsto \quad \frac{u\left(\lambda_1 + \lambda_n - u\right)}{\lambda_1\lambda_n}$$

qui a son maximum en $\frac{\lambda_1 + \lambda_n}{2}$ d'où la majoration souhaitée en remplaçant $\overline{\lambda}$ par cette valeur. □

*Que retenir de ces deux preuves ?* La première est bien plus rapide, l'autre à le mérite de faire un dessin, mais mieux vaut expliquer l'idée car les calculs sont plus laborieux, et Cauchy-Schwarz simplifie toute la vie plus haut dans la première version...(mais s'il vous manque un 'oraux X-ENS Algèbre 3' le développement de la deuxième version se trouve dans le Urruty donc cela peut 'sauver'...

# Théorème de sélection de Helly

Emily Clement

Master 2 MEEF

**Bibliographie :**
- *Oraux X-ENS Analyse 2*, Serge Francinou, Hervé Gianella, Serge Nicolas
- *Analyse mathématique, La maîtrise de l'implicite*, Frédéric Testard.

**Théorème .1** (Théorème de sélection de Helly).

Soit $(f_n)$ suite croissante de $\mathbb{R}$ dans $[-1,1]$.
Montrer qu'il existe une sous-suite de $(f_n)$ qui CVS vers $f : \mathbb{R} \to [-1,1]$

**Étape** 1 : Soit $E \subset \mathbb{R}$ dénombrable.
Soit $(f_n)_n : E \to \mathbb{R}$ tel que :

$$|f_n(x)| \leqslant 1, \forall n \in \mathbb{N}, \forall x \in E$$

Montrer qu'il existe une sous-suite de $(f_n)$ qui converge simplement vers une fonction $f : E \to \mathbb{R}$ .

**Étape 2 : Appliquer à $E = \mathbb{Q}$**

**Étape 3 : $\mathbb{R}\backslash\mathbb{Q}$** :

1. On prend $x \in \mathbb{R}$ (supposé pas dans $\mathbb{Q}$ sinon le travail est trivial) et on montre que si $\lim\limits_{x_-} f = \lim\limits_{x_+} f$ alors $(f_n(x))_n$ converge également vers $l$.
2. On montre que l'ensemble des points qui ne vérifient pas cette propriétés est dénombrable, et on réapplique l'étape 1.

*Démonstration.*

**Étape** 1

$E$ est dénombrable, donc on a une suite $(r_k)_{k\geqslant 0}$ tel que

$$E = \bigcup_{k\in\mathbb{N}} r_k$$

Comme $f_n$ est bornée dans $E$, $f_n(r_0)$ est bornée sur $\mathbb{R}$ donc on peut en extraire une sous-suite $f_{\varphi_0(n)}(r_0)$ convergente, de même pour la suite bornée $f_{\varphi_0(n)}(r_1)$ : $f_{\varphi_0\circ\varphi_1(n)}(r_1)$, et évidemment $f_{\varphi_0\circ\varphi_1(n)}(r_0)$ converge également.

Au final en itérant ce procédé :

$$\underline{\forall p \in \mathbb{N}, \forall k \in [0,p]} \quad f_{\varphi_0\circ\varphi_1\circ\cdots\circ\varphi_n(p)}(r_k)$$

converge.

On pose donc la suite $g_n := f_{\varphi_0\circ\varphi_1\circ\cdots\circ\varphi_n(n)}$, $g_N$ est une suite extraite de $f_n$ qui converge simplement vers une fonction $f$ de $E$ dans $[-1,1]$ (les inégalités larges passent à la limite).

**Étape** 2 :

[$\mathbb{Q}$ dénombrable], [Étape 1] : $(f_n)$ CVS sur $\mathbb{Q}$ vers une fonction $f$ : $\mathbb{Q} \rightarrow [-1,1]$

**Étape** 3 : Soit $x \in \mathbb{R}$, on a deux cas :

1. $\lim\limits_{x_-} f = \lim\limits_{x_+} f$

   On le traduit par : $\forall\varepsilon > 0, \exists\eta > 0, \forall t \in [x-\eta, x+\eta] \cap \mathbb{Q}$ : $|f(t) - l| \leqslant \varepsilon$. (1)

   On prend deux rationnels : un à gauche $\alpha \in [x-\eta, x]$ et un à droite ~~$\beta$in~~ $\beta \in$ $[x, x+\eta]$, par croissance :

   $$f_n(\alpha) \leqslant f_n(x) \leqslant f_n(\beta)$$

   $[\alpha, \beta \in \mathbb{Q}, (f_n)_n$ CVS dans $\mathbb{Q}]$ :
   $\exists N_0 \geqslant 0, \forall n \geqslant N_0$ :

   $$f_n(\beta) \leqslant f(\beta) + \varepsilon, f_n(a) \geqslant f(\alpha) - \varepsilon$$

   Donc : [(1)]

   $$l - 2\varepsilon \leqslant f(\alpha) - \varepsilon \leqslant f_n(x) \leqslant f(\beta) + \varepsilon \leqslant l + 2\varepsilon$$

   Donc $(f_n)_n$ converge vers $l$.

2. $\lim\limits_{x_-} f \neq \lim\limits_{x_+} f$, donc par croissance : $\lim\limits_{x_-} f < \lim\limits_{x_+} f$.

   Posons :$D := \left\{ x \in E, \lim\limits_{x_-} f < \lim\limits_{x_+} f \right\}$.

   $D$ est au plus dénombrable, en effet :

   [$\mathbb{Q}$ est dense dans $\mathbb{R}$] donc $\exists q_x \in \mathbb{Q}$ tel que :

$$\lim_{x_-} f \leqslant q_x \leqslant \lim_{x_+} f$$

   si on pose :

$$q : x \mapsto q_x$$

(*) [$f$ croissante ] : $q$ injective.

   $D$ est donc dénombrable.

   On ré-applique la question 1 : $(f_n)$ admet une suis-suite qui converge sur $D$, qui évidemment converge encore sur $\mathbb{R} \backslash D$.

□

(*) Si $\begin{cases} x < y \\ q_x = q_y \end{cases}$ alors

$$\lim_{x^-} f \leq \underset{\substack{\| \\ q_y}}{q_x} \leq \lim_{x^+} f$$

$$f \nearrow \quad \overline{\lim_{x^+} f \leqslant \underbrace{\lim_{y^-}}_{} f \leq q_y \leq \lim_{y^+} f}$$