Optimisation Convexe

Ivan Lejeune*

12 février 2024

1 Optimisation en dimension finie

Introduction. Méthode d'évaluation :

- CC noté en CM
- 2TPs notés
- Examen terminal à 40%

Notation. On considère un espace vectoriel normé de dimension n noté $(E, ||\cdot||)$ et U ouvert de E. On considère une fonction $f\colon U \to \mathbb{R}$. Dans la pratique, E sera égal à $\mathbb{R}^n$.
Soit $x \in U$, on note $f'(x)$ la différentielle (qu'on appellera plus simplement "dérivée") de f en x. On a donc, pour tout $h \in E$ tel que $||h||$ soit assez petit,

$$f(x+h) = f(x) + f'(x)\cdot h + ||h||\varepsilon(x,h)$$

avec $\varepsilon(x,h) \underset{h\to 0}{\to} 0$ et $f'(x) \in \mathscr{L}(E,\mathbb{R})$.
Avec cette notation, si f est dérivable en x, alors f admet des dérivées partielles en x dans toutes les directions, et si $(e_1,\dots,e_n)$ est une base de E, on note "$\partial_i f(x)$" ou encore $\frac{\partial f}{\partial x_i}(x)$" la dérivée partielle de f par rapport à la i-ème variable. On a alors

$$\partial f(x) = f'(x)\cdot e_i \qquad i = 1,\dots n$$

Ainsi, pour $h \in E$ tel que $h = \sum_{i=1}^n h_i e_i$; on a

$$\begin{aligned} f'(x)\cdot h &= f'(x)\cdot\left(\sum_{i=1}^n h_i e_i\right) \\ &= \sum_{i=1}^n h_i f'(x) e_i \\ &= \sum_{i=1}^n h_i \partial f(x) \end{aligned}$$

De même, si $x \mapsto f(x)$ est dérivable en x, on note $f''(x) \in \mathscr{L}(E;\mathscr{L}(E,\mathbb{R}))$ cette dérivée seconde et on considère $f''(x)$ comme une forme bilinéaire

$$f''(x) \in \mathscr{L}(E\times E,\mathbb{R})$$

Avec ces notations, le théorème fondamental de l'analyse (TTA) peut s'énoncer ainsi :

Théorème. Soit $f \in \mathscr{C}'(U,\mathbb{R})$. Alors pour tout $(x,y) \in U$ tel que $\forall t \in [0,1], x + t(y-x) \in U$, on a

$$f(y) = f(x) + \int_0^1 f'(x+t(y-x))\cdot(y-x)dt$$

*Cours inspiré de M. Marche

Formule de Taylor à l'ordre 2. Soit $f \in \mathscr{C}^2(U,\mathbb{R}), x \in U$. Alors il existe un voisinage ν de x tel que pour tout $h \in \nu$

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

Bien entendu, cette expression peut aussi se formuler ainsi :

$$\begin{aligned} f(x+h) = f(x) &+ \sum_{i=1}^{n} \partial_i f(x) h_i \\ &+ \frac{1}{2}\sum_{i,j=1}^{n} \partial^2_{i,j} f(x) h_i h_j \\ &+ \|h\|^2 \varepsilon(h) \end{aligned}$$

avec $\varepsilon(h) \underset{\|h\|\to 0}{\to} 0$, de manière a bien mettre en évidence la linéarité de la dérivée et la bilinéarité de la dérivée seconde.
En notant $\nabla f(x)$ le gradient de f évalué en x, et $\nabla^2 f(x)$ la matrice Hessienne de f évaluée en x, on a :

$$f(x+h) = f(x) + \nabla f(x) h + \frac{1}{2}{}^T h \nabla^2 f(x)\cdot h + o(\|h\|^2)$$

Définition. Soit $f : U \to \mathbb{R}$ de classe $\mathscr{C}^1$. Alors pour tout $a \in U$, il existe un unique vecteur, noté $\nabla f(a)$ tel que pour tout $h \in E$

$$f'(a)\cdot h = \langle \nabla f(a), h\rangle$$

où $\langle\cdot,\cdot\rangle$ désigne le produit scalaire canonique sur $\mathbb{R}^n$ (qui peut être noté aussi avec $\cdot$ mais dans ce cas attention aux confusions). C'est le **gradient** de f.

Rappel. Dans une base orthonormée, on a

$$\nabla f(a) = \sum_{i=1}^{n} \partial_i f(a) e_i$$

1.1 Résultats d'existence

Un outil fondamental à la compacité

Théorème 1.1.1. Soit K un compact de $\mathbb{R}^n$ et $f : K \to \mathbb{R}$ une fonction continue. Alors f est bornée et atteint ses bornes :

$$\sup_{x\in K} |f(x)| < +\infty$$

et il existe $\underline{x} \in K$ et $\overline{x} \in K$ tels que

$$\begin{aligned} f(\underline{x}) &= \inf_{x\in K} f(x) = \min_{x\in K} f(x) \\ f(\overline{x}) &= \sup_{x\in K} f(x) = \max_{x\in K} f(x) \end{aligned}$$

Preuve. Ce résultat à été démontré dans le cours de topologie / analyse fonctionnelle.
Puisque f est continue, $f(K)$ est une partie compacte de $\mathbb{R}$, c'est à dire une partie fermée et bornée de $\mathbb{R}$. Ainsi on a

$$-\infty < \inf f(K) \leq \sup f(k) < +\infty$$

et puisque $f(K)$ est fermée et que $\inf f(k)$ et $\sup f(k)$ sont adhérents à $f(K)$, on a

$$\inf f(K) = \min f(K) \in f(E)$$

et
$$\sup f(K) = \max f(K) \in f(E)$$

Definition 1.1.2. Soit $f : \mathbb{R}^n \to \mathbb{R}$. On dit que f est **coercive** si $f(x) \to +\infty$ lorsque $\|x\| \to +\infty$.

Théorème 1.1.3. Soit $f : \mathbb{R}^n \to \mathbb{R}$ continue et coercive. Alors f est minorée et atteint son minimum.

Preuve. Posons $A = f(0) + 1$.
Puisque f est coercive, il existe $\alpha > 0$ tel que
$$\forall x \in \mathbb{R}^n, \|x\| \geq \alpha \Longrightarrow f(x) \geq f(0) + 1$$
La boule $\overline{B}(0, \alpha)$ est un fermé borné de $\mathbb{R}^n$ donc un compact de $\mathbb{R}^n$ et $f|_{\overline{B}(0,\alpha)}$ est continue.
D'après le Théorème 1.1.1, f est minorée sur $\overline{B}(0, \alpha)$ et atteint son minimum en un certain $x_0 \quad (\forall x \in \overline{B}(0, \alpha), f(x) \geq f(x_0))$
Ainsi, soit $x \in \mathbb{R}^n$

a) si $x \in \overline{B}(0, \alpha)$, alors $f(x) \geq f(x_0)$
b) $x \notin \overline{B}(0, \alpha)$, alors $\|x\| > \alpha$ et donc $f(x) \geq f(0) + 1$ et $f(x) \geq f(x_0) + 1 > f(x_0)$ puisque $0 \in \overline{B}(0, \alpha)$

Ainsi $\forall x \in \mathbb{R}^n, f(x) \geq f(x_0)$ et x_0 est bien le minimum de f sur $\mathbb{R}^n$

Remarque 1.1.4. Ce dernier résultat peut être généralisé, sous les même hypothèses, au cas d'une fonction $f : K \to \mathbb{R}$ avec K fermé de $\mathbb{R}^n$.

1.2 Caractérisation des extremas sans contraintes

Un outil fondamental : le calcul différentiel.

Théorème 1.2.1. Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \to \mathbb{R}$ de classe $\mathscr{C}^1$. Si $x_0 \in U$ est extremum local de f sur U alors on a $f'(x_0) = 0$ (ou $\nabla f(x_0) = 0$)

Preuve. Rappelons ce qu'il se passe pour une fonction $\varphi : I \subset \mathbb{R} \to \mathbb{R}$ qui admet par exemple un maximum local en $0 \in I$.
On a d'une part $\varphi'(0) = \lim\limits_{x \to 0^+} \frac{\varphi(x) - \varphi(0)}{x} \leq 0$ car $x > 0$ et $\varphi(x) - \varphi(0) \leq 0$
et d'autre part $\varphi'(0) = \lim\limits_{x \to 0^-} \frac{\varphi(x) - \varphi(0)}{x} \geq 0$ car $x < 0$ et $\varphi(x) - \varphi(0) \leq 0$
Dans le cas $E = \mathbb{R}^n$, supposons que f admet un maximum local en $x_0 \in U$. Soit e_i un vecteur de base.
On sait que $\partial_i f(x_0) = f'(x_0) \cdot e_i = \varphi'_{e_i}(0)$ avec $\varphi_{e_i}(t) = f(x_0 + t e_i)$
(où on remarque que $t \in]-\delta, \delta[$) puisque U est ouvert et $x_0 \in U$)
Puisque f admet un maximum local en x_0, il existe $r > 0$ tel que
$$\forall x \in B(x_0, r), \quad f(x) \leq f(x_0)$$
Soit $h \in \mathbb{R}^n$ tel que $\|h\| \leq r$ alors
$$f(x_0 + h) \leq f(x_0)$$
et en particulier
$$\begin{aligned} |t| \leq r &\Longrightarrow f(x_0 + t e_i) \leq f(x_0) \\ &\Longrightarrow \varphi'_{e_i}(0) = 0 \end{aligned}$$
Ainsi, toutes les dérivées partielles de f sont nulles en x_0 et donc $f'(x_0) = 0$

Remarque 1.2.2. Ce résultat est bien entendu faux si U n'est pas ouvert :

$$\begin{aligned} f : [0,1] &\to [0,1] \quad \text{est } \mathscr{C}^\infty \\ x &\to x \end{aligned}$$

0 est le minimum sur $[0,1]$, pourtant $f'(0) = 1$.

Definition 1.2.3. Soit $f : U \to \mathbb{R}$ et $a \in U$. a est un **point critique** si au moins une des conditions suivantes est satisfaite :

- $f'(a) = 0$
- $\nabla f(a) = 0$
- $\partial_i f(a) = 0$ pour $i \in [\![1, n]\!]$

Théorème 1.2.4. Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \to \mathbb{R}$ de classe $\mathscr{C}^2$.
Si $x_0 \in U$ est minimum local de f sur U, alors

$$f'(x_0) = 0 \text{ et } f''(x_0) \text{ est positive}$$

Au sens des formes bilinéaires symétriques cela donne

$$\forall \xi \in \mathbb{R}^n, f''(x_0) \cdot (\xi, \xi) \geq 0$$

Preuve. Soit $x_0 \in U$ et $h \in \mathbb{R}^n$. Pour tout $t \in \mathbb{R}$ assez petit, on a

$$f(x_0 + th) - f(x_0) \geq 0$$

Or, d'après Taylor, on a

$$f(x_0 + th) - f(x_0) = \frac{1}{2} f''(x_0) \cdot (th, th) + \underset{t \to 0}{o}(t^2)$$

d'où

$$\frac{1}{t^2}(f(x_0 + th) - f(x_0)) = \frac{1}{2} f''(x_0)(h, h) + \underset{t \to 0}{o}(1)$$

et donc

$$\frac{1}{t^2}(f(x_0 + th) - f(x_0)) \underset{t \to 0}{\to} \frac{1}{2} f''(x_0)(h, h)$$

Ainsi, on a bien

$$\forall h \in \mathbb{R}^n, f''(x_0) \cdot (h, h) \geq 0$$

Remarque 1.2.5. La réciproque est fausse :
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ telle que $f(x, y) = x^2 - y^4$. On a

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -4y^3 \end{pmatrix} \qquad \text{et} \qquad \nabla^2 f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -12y^2 \end{pmatrix}$$

On a alors

$$\nabla f(0, 0) = 0 \qquad \text{et} \qquad \nabla^2 f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Ainsi, $\nabla^2 f(0, 0) \geq 0$. Pourtant $(0, 0)$ n'est pas un extremum de f.

Remarque 1.2.6. Ces deux caractérisations sont des conditions nécessaires pour que x_0 soit un extremum (mais non suffisantes..)

Remarque 1.2.7. Il existe bien entendu un résultat "symétrique" du Théorème 1.2.4 :

Théorème. Si x_0 est un **maximum local** alors

$$f''(x_0) \cdot (h,h) \leq 0, \quad \forall h \in \mathbb{R}^n$$

Théorème 1.2.8. Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \to \mathbb{R}$ de classe $\mathscr{C}^2$.
Soit $x_0 \in U$ tel que $f'(x_0) = 0$, alors :

(i) Si $f''(x_0)$ est **définie positive**, alors x_0 est **minimum local strict** de f.

(ii) Si $f''(x_0)$ est **définie négative**, alors x_0 est **maximum local strict** de f.

(iii) si $f''(x_0)$ a des valeurs propres non nulles de signe différent alors x_0 n'est pas un extremum. On est alors en présence d'un **point-selle** (ou "col").

(iv) si certaines valeurs propres sont nulles on ne peut pas conclure sans effectuer une étude locale au voisinage des points critiques.

Remarque 1.2.9. La terminologie "définie positive" signifie, au choix, l'une des définitions équivalents suivantes

(i) $\forall h \in \mathscr{B}(0,r), f''(x_0) \cdot (h,h) > 0$

(ii) f'' est positive et non dégénérée.

(iii) $\mathsf{SP}_{\mathbb{R}}(\nabla^2 f(x_0)) \subset \mathbb{R}_+^*$ (on rappelle que $\mathsf{SP}_{\mathbb{R}}$ correspond au spectre réel)

La terminologie "définie négative" se déduit immédiatement.

Preuve. Reprenons la formule de Taylor Young à l'ordre 2 pour un point critique :

$$\forall h \in \mathscr{B}(0,r), f(x_0+h) - f(x_0) = \frac{1}{2} f''(x_0) \cdot (h,h) + o(||h||^2)$$

(i) Puisque $f''(x_0)$ est définie positive, on a

$$\alpha ||h||^2 \leq f''(x_0) \cdot (h,h) \leq \beta ||h||^2$$

où

$$\alpha = \min\{\lambda\}_{i=1}^d > 0 \qquad \text{et} \qquad \beta = \max\{\lambda\}_{i=1}^d > 0$$

Ainsi, il existe $r > 0$ tel que

$$\forall h \in \mathscr{B}(0,r) \setminus \{0\}, f(x_0+h) - f(x_0) > 0$$

et x_0 est donc minimum local strict.

(ii) On peut appliquer la même méthode pour $f''(x_0)$ définie négative.

(iii) On suppose que $f''(x_0) \cdot (h,h)$ prend des valeurs strictement positives et strictement négatives. Ainsi, x_0 n'est pas un extremum local. On est en présence d'un "point selle".

(iv) Si certaines valeurs propres sont nulles, on ne peut pas conclure sur la nature de x_0 sans effectuer d'étude locale.

Remarque 1.2.10. Ces caractérisations peuvent s'étendre sans problème aux dérivées d'ordre supérieures, lorsqu'elles existent.

Remarque 1.2.11. Méthode de recherche d'extremas libres
Voir TD2 et 3.

1.3 Convexité : définitions et caractérisations

Definition 1.3.1. Soit U une partie de $\mathbb{R}^n$. On dit que U est **convexe** si

$$\forall (x,y) \in U \times U, \forall t \in [0,1],$$
$$tx + (1-t)y \in U$$

Definition 1.3.2. Soit U une partie convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$.
On dit que f est **convexe** si

$$\forall (x,y) \in, U \times U, \forall t \in [0,1],$$
$$f(tx + (1-t)y) \leq tf(x = +(1-t)f(y))$$

On dit que f est **strictement convexe** si

$$\forall (x,y) \in, U \times U, \forall t \in [0,1],$$
$$f(tx + (1-t)y) < tf(x = +(1-t)f(y))$$

Proposition 1.3.3. Soit U un ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ de classe $\mathscr{C}^1$.
On a f convexe si et seulement si

$$\forall (u,v) \in U \times U, f(v) \geq f(u) + f'(u) \cdot (v-u)$$

De plus, on a f strictement convexe si et seulement si

$$\forall (u,v) \in U \times U, f(v) > f(u) + f'(u) \cdot (v-u)$$

Preuve. Supposons f convexe sur U. Soit $(x,y) \in U \times U$ et $t \in]0,1[$. On a

$$\begin{aligned} f(ty + (1-t)x) &= f(x + t(y-x)) \\ &\leq tf(y) + (1-t)f(x) \\ &\leq f(x) + t(f(y) - f(x)) \end{aligned}$$

d'où

$$\forall t \in]0,1[, \frac{1}{t}(f(x + t(y-x)) - f(x)) \leq f(y) - f(x)$$

En faisant tendre t vers 0, on obtient

$$f'(x) \cdot (y-x) \leq f(y) - f(x)$$

Réciproquement, supposons que

$$\forall (x,y) \in U \times U, f(y) \geq f(x) + f'(x) \cdot (y-x)$$

Soient $(u,v) \in U \times U$ et $t \in]0,1[$.
On pose $x = tu + (1-t)v$ et $y = u$. On a alors

$$f(u) \geq f(tu + (1-t)v + f'(tu + (1-t)v) \cdot ((1-t)(u-v)) \tag{1}$$

On pose $x = tu + (1-t)v$ et $y = v$. On a alors

$$f(u) \geq f(tu + (1-t)v + f'(tu + (1-t)v) \cdot (t(v-u)) \tag{2}$$

La combinaison linéaire $t(1) + (1-t)(2)$ donne alors

$$tf(u) + (1-t)f(v) \geq f(tu + (1-t)v)$$

Donc f est convexe.
En appliquant le même raisonnement avec des inégalités strictes on peut montrer que si

$$f(y) > f(x) + f'(x) \cdot (y - x)$$

alors f est strictement convexe.
Enfin, supposons f strictement convexe.
On considère

$$x, y \in U, x \neq y, t \in]0,1[, w = tx + (1-t)y$$

On a f convexe par hypothèse donc, d'après ce qui précède, on a

$$f(w) - f(x) \geq f'(x) \cdot (w - x)$$

Or

$$f(w) < tf(x) + (1-t)f(y)$$
$$\text{et } w - x = (1-t)(y - x)$$

Ainsi, on obtient :

$$tf(x) + (1-t)f(y) - f(x) > f'(x) \cdot ((1-t)(y-x))$$

d'où

$$(1-t)(f(y) - f(x)) > (1-t)f'(x) \cdot (y - x)$$

et donc

$$f(y) > f(x) + f'(x) \cdot (y - x)$$

Remarque 1.3.4. En analyse fonctionnelle, un opérateur

$$T : X \to X^*$$

sur un espace vectoriel topologique X est dit **opérateur monotone** si

$$\forall u, v \in X, (T_u - T_v, u - v) \geq 0$$

Proposition 1.3.5. Soit U ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ de classe $\mathscr{C}^1$. Alors f est convexe si et seulement si f' est monotone sur U, c'est à dire :

$$\forall (u, v) \in U, (f'(v) - f'(u)) \cdot (v - u) \geq 0$$

De plus, f est strictement convexe si et seulement si f est strictement monotone.

Preuve. Si f est convexe, alors d'après la proposition 1.3.1, on a

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

On applique cette inégalité au couple (u, v) puis au couple (v, u) :

$$f(u) \geq f(v) + f'(v) \cdot (u - v)$$
$$f(v) \geq f(u) + f'(u) \cdot (v - u)$$

La somme de ces expressions donne

$$0 \geq f'(v) \cdot (u - v) - f'(u) \cdot (u - v)$$

d'où

$$0 \leq (f'(v) - f'(u)) \cdot (v - u)$$

Réciproquement, supposant f' monotone, on veut montrer que

$$\forall (x,y) \in U, f(y) - f(x) - f'(x)\cdot(y-x) \geq 0$$

Le TFA permet de reformuler $f(y) - f(x)$ en

$$\begin{aligned} f(y) - f(x) - f'(x)\cdot(y-x) &= \int_0^1 f'(x+t(y-x))\cdot(y-x)dt - f'(x)\cdot(y-x) \\ &= \int_0^1 (f'(x+t(y-x))\cdot(y-x) - f'(x)\cdot(y-x))dt \end{aligned}$$

Or, par monotonie, $\forall t \in \,]0,1[$, on a

$$(f'(x+t(y-x)) - f'(x))\cdot(t(y-x)) \geq 0$$

d'où

$$(f'(x+t(y-x)) - f'(x))\cdot(y-x) \geq 0$$

et donc

$$f(y) - f(x) - f'(x)\cdot(y-x) \geq 0$$

f est donc convexe.

Proposition 1.3.6. Soit U ouvert convexe de $\mathbb{R}^n$ et $f: U \to \mathbb{R}$ de classe $\mathscr{C}^2$. Alors f est convexe si et seulement si

$$\forall u \in U, f''(u) \text{ est positive}$$

De plus, si pour tout $u \in U, f''(u)$ est définie positive, alors f est strictement convexe (la réciproque est fausse)

Preuve. Supposons f convexe. D'après la proposition 1.3.5, f' est monotone donc

$$\forall (u,v) \in U, (f'(v) - f'(u))\cdot(v-u) \geq 0$$

Soient $x \in U, \xi \in \mathbb{R}^n$ et $t > 0$. On applique la formule de monotonie avec $u = x$ et $v = x + t\xi$:

$$(f'(x+t\xi) - f'(x))(t\xi) \geq 0$$

On divise cette inégalité par t^2 et on fait tendre t vers 0, on obtient que

$$f''(x)\cdot(\xi,\xi) \geq 0$$

Réciproquement, si f'' est positive sur U alors d'après le TFA, on a :

$$(f'(u) - f'(v))\cdot(v-u) = \int_0^1 f''(u+t(u-v))\cdot(v-u, v-u)$$

D'où

$$(f'(u) - f'(v))\cdot(v-u) \geq 0$$

et donc f convexe.
Si maintenant f'' est définie positive, on obtient de la même manière que f' est strictement monotone et donc f strictement convexe.

Definition 1.3.7. Soit $f: \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$. On dit que f est **elliptique** s'il existe $\alpha > 0$ tel que

$$\forall (u,v) \in \mathbb{R}^n, (f'(v) - f'(u))\cdot(v-u) \geq \alpha \|v-u\|^2$$

Proposition 1.3.8. Soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$ et elliptique. Alors f est strictement convexe et coercive. Elle vérifie la propriété :

$$\forall (x,y) \in \mathbb{R}^n, \quad f(y) - f(x) - f'(x)\cdot(y-x) \geq \frac{\alpha}{2}\|y-x\|^2$$

De plus, si f est $\mathscr{C}^2$, alors

$$\forall u \in \mathbb{R}^n, \forall w \in \mathbb{R}^n, \quad f''(u)\cdot(w,w) \geq \alpha\|w\|^2$$

Preuve. f' est strictement monotone donc f est strictement convexe.
Posons

$$A = f(y) - f(x) - f(x)\cdot(y-x)$$

On a

$$A = \int_0^1 (f'(x+t(y-x)) - f'(x))\cdot(y-x)de$$

par application du TFA, et par ellipticité de f, on a

$$(f'(x+t(y-x)) - f'(x))\cdot(t(y-x)) \geq \alpha t^2\|y-x\|^2$$

Ainsi,

$$A \geq \int_0^1 \alpha t\|y-x\|^2 dt = \frac{\alpha}{2}\|y-x\|^2$$

Appliquant cette inégalité avec $x = 0$, on a

$$f(y) - f(0) - f'(0)\cdot y \geq \frac{\alpha}{2}\|y\|^2$$

et donc

$$\forall y \in \mathbb{R}^n, f(y) \geq \underbrace{\frac{\alpha}{2}\|y\|^2 - \|f'(0)\|\|y\| + f(0)}_{\substack{\text{polynôme de degré 2 en } y \\ \text{de coefficient dominant strictement} \\ \text{positif, noté } p(\|y\|)}}$$

On a

$$p(\|y\|) \underset{\|y\|\to+\infty}{\to} +\infty$$

donc f est convexe.
Supposons enfin que f est $\mathscr{C}^2$. Soit $u \in \mathbb{R}^n, \xi \in \mathbb{R}^n, t > 0$, on a

$$(f'(u+t\xi) - f'(u))\cdot(t\xi) \geq \alpha t^2\|\xi\|^2$$

par ellipticité de f. D'où, en divisant par t^2 et faisant tendre t vers 0, on a

$$f''(u)\cdot(\xi,\xi) \geq \alpha\|\xi\|^2$$

1.4 Optimisation des fonctions convexes

Théorème 1.4.1. Soit U un ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une fonction convexe, $x_0 \in U$, alors

(i) si x_0 est un minimum local de f c'est un minimum global

(ii) si f est strictement convexe, alors elle admet au plus un minimum et c'est un minimum strict

(iii) si en outre f est de classe $\mathscr{C}^1$ alors x_0 est un minimum de f si et seulement si

$$\forall y \in U, f'(x_0)\cdot(y-x_0) \geq 0$$

Preuve.

(i) Supposons que x_0 soit un minimum local, alors

$$\exists \eta > 0, \forall z \in B(x_0, \eta), f(z) \geq f(x_0)$$

Soit $y \in U$. On pose $z = tx_0 + (1-t)y$, avec $t \in]0,1[$
Alors $\|z - x_0\| = (1-t)\|y - x_0\|$ et il est possible de choisir t suffisamment proche de 1 pour avoir $\|z - x_0\| < \eta$ et donc

$$f(x_0) \leq f(z) = f(tx_0 + (1-t)y)$$

ou encore

$$f(x_0) \leq tf(x_0) + (1-t)f(y)$$

puisque f est convexe. Cette dernière inégalité donne en divisant par $(1-t)$

$$\forall y \in U, f(x_0) \leq f(y)$$

et donc x_0 est un minimum global sur U.

(ii) supposons f strictement convexe et que le min de f soit atteint en x_0 et y_0. Alors

$$f\left(\frac{x_0 + y_0}{2}\right) < \frac{1}{2}f(x_0) + \frac{1}{2}f(y_0) = \min_U f$$

par stricte convexité, ce qui est absurde.

(iii) Supposons en outre f de classe $\mathscr{C}^1$ et notons $f(x_0) = \min_U f$. Soit $y \in U$. On a

$$\forall t \in [0,1], \quad f(x_0) \leq f(x_0 + t(y - x_0))$$

et donc

$$\forall t \in]0,1[, \quad 0 \leq \frac{f(x_0 + t(y - x_0)) - f(x_0)}{t}$$

et à la limite en 0

$$0 \leq f'(x_0) \cdot (y - x_0)$$

Notez que cette implication est vrai même si f n'est pas convexe.
Supposons maintenant que $\forall y \in U$

$$f'(x_0) \cdot (y - x_0) \geq 0$$

pour un certain x_0 donné. On a

$$f(y) \geq f(x_0) + f'(x_0) \cdot (y - x_0)$$

par convexité, et donc

$$f(y) \geq f(x_0)$$

Ainsi x_0 réalise le minimum de f sur U.

Proposition 1.4.2. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathscr{C}^1$ et elliptique (constante d'ellipticité α).
Alors f est minorée, atteint son minimum en un unique point x_0, caractérisé par

$$f'(x_0) = 0$$

Preuve. La preuve est immédiate en vertu de ce qui précède

- f est $\mathscr{C}^0$ et coercive (Proposition 1.3.4) donc f est minorée et atteint son min (Théorème

1.1.2)

- f est strictement convexe (Proposition 1.3.4) donc le min est unique (Théorème 1.4.1)
- si x_0 est le min de f sur $\mathbb{R}^n$, comme $\mathbb{R}^n$ est un ouvert, alors $f'(x_0) = 0$ (Théorème 1.2.1). Si $f'(x_0) = 0$, puisque f est convexe, on a

$$\forall y \in \mathbb{R}^n, f(y) - f(x_0) \geq f'(x_0) \cdot (y - x_0) = 0$$

Par la proposition 1.3.1. Et x_0 est donc le min de f