Optimisation et optimisation numérique

Chapitre 1 : Premiers éléments d'optimisation

Lucie Le Briquer

16 janvier 2018

Table des matières

1 Introduction **2**

2 Théorème de projection **2**

3 Quelques structures intéressantes **3**

3.1 Fonctions convexes 4

3.2 Ellipcité 5

3.3 Fonctions semi-continue inférieurement (s.c.i.) 5

4 Conditions d'optimalité **7**

1 Introduction

De manière assez simple, l'optimisation consiste à minimiser $J(u)$ avec $u \in K \subset X$ où X est un espace topologique. $J\colon X \longrightarrow \mathbb{R},]-\infty, +\infty]$ est appelée fonctionnelle. On s'intéresse à l'existence, l'unicité et au calcul des solutions.
On peut par exemple s'intéresser à :

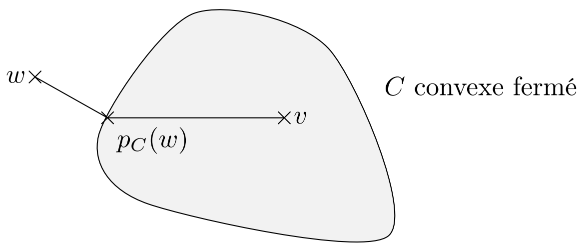
- J linéaire ($J(x) = Bx$) sur $K = \cap(C_i X \leqslant b_i)$, ce qui se ramène à de la programmation linéaire.
- J quadratique :
$$J(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$$
- J convexe (SMV Support Vector Machine)
- J non linéaire

Quelques références :

1. Ph. Ciailet, *Introduction à l'analyse numérique matricielle et à l'optimisation*
2. J.F. Bonnans, J.C. Gilbert, C. Lemaréchal et C.A. Sagistizabal, *Numerical Optimization*
3. J. Nocedal et S. Wright, *Numerical Optimization*
4. D.P. Bertseka,s *Non linear programming*
5. M. Nikola, *Optimization, Application in image processing* (cours MVA)

2 Théorème de projection

Définition 1 (convexe)
Un sous-ensemble $C \subset E$ e.v. est *convexe* si $\forall x, y \in C,\ \forall t \in [0,1] : tx + (1-t)y \in C$.



Théorème 1 (de projection sur les convexes fermés)
Soit V un Hilbert et C un *convexe fermé* de V. Alors pour tout $w \in V$, il existe un unique $p_C(w) \in C$ tel que :
$$|w - p_C(w)|_V = \inf_{v \in C} |w - v|_V$$
De plus, $\forall v \in C$ on a :
$$\langle w - p_C(w), v - p_C(w)\rangle_V \leqslant 0$$

Preuve.
Égalité du parallélogramme :

$$2|a|^2 + 2|b|^2 = |a-b|^2 + |a+b|^2$$

Soit (v_n) une suite minimisante. Soit $\varepsilon > 0$, $\exists n \geqslant 0\ \forall p \geqslant 0$:

$$|w - v_{n+p}|^2 \leqslant d^2 + \varepsilon$$

où $d = \inf_{v \in C} |w - v|$. Par l'égalité du parallélogramme :

$$2|w - v_{n+p}|^2 + 2|w - v_n|^2 = |v_{n+p} - v_n|^2 + 4\left|w - \frac{v_n + v_{n+p}}{2}\right|^2$$

Par suite :

$$\begin{aligned} |v_{n+p} - v_n|^2 &\leqslant 4d^2 + 4\varepsilon - 4\Big|w - \underbrace{\frac{v_n + v_{n+p}}{2}}_{\in C}\Big|^2 \\ &\leqslant 4d^2 + 4\varepsilon - 4d^2 = 4\varepsilon \end{aligned}$$

Par suite (v_n) est de Cauchy, V est complet. Si $p_C(w) = \lim v_n$ on a $p_C(w) \in C$ puisque C est fermé et $|w - p_C(w)|^2 \leqslant d^2$. L'unicité est laissée en exercice.
Enfin,

$$p_C(w) + t(v - p_C(w)) \in C \quad \text{si } v \in C \text{ et } t \in [0,1]$$

d'où :

$$|w - p_C(w)|^2 \leqslant \underbrace{|w - p_C(w) + t(v - p_C(w))|^2}_{\gamma(t)}$$

Par un développement de Taylor en $t = 0$ et γ ($\gamma'(0) \geqslant 0$), on obtient l'inégalité. □

3 Quelques structures intéressantes

Définition 2 (domaine d'épigraphe)

Soit $f\colon X \to]-\infty, +\infty]$,

$$\mathrm{dom}(f) = \{x \in X \mid f(x) < +\infty\} = (f < +\infty)$$

$$\mathrm{epi}(f) = \{(x,y) \in X \times \mathbb{R} \mid y \geqslant f(x)\}$$

3.1 Fonctions convexes

Définition 3 (fonction convexe et strictement convexe)

Soit $f\colon C \subset E \to \mathbb{R}$ où C est un convexe et E un e.v. On dit que f est *convexe* si :

$$\forall x, y \in C,\ \forall t \in [0,1],\ f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y)$$

On dit que f est strictement convexe si :

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

Exemple. Si $f(x) = \frac{1}{2}\langle Ax, x\rangle$ avec A symétrique positive, f est convexe. Si A est définie positive, f est strictement convexe. (à faire en exercice)

Remarque. Si $f\colon C \to \mathbb{R}$ est convexe, on peut considérer $\tilde{f}\colon E \to]-\infty, +\infty]$ définie par :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in C \\ +\infty & \text{sinon} \end{cases}$$

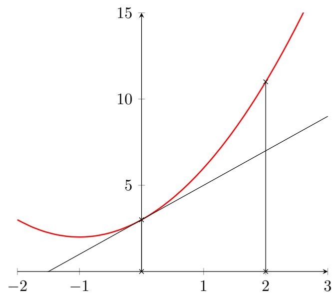
$\mathrm{dom}(\tilde{f}) = C$ et $\tilde{f}$ est convexe au sens étendu (convention $a + (+\infty) = +\infty$ si $a \in]-\infty, +\infty]$).

Propriété 1 (condition de convexité)

$f\colon U \subset E \to \mathbb{R}$, U ouvert, et dérivable en tout point d'un convexe C. Alors :

$$f \text{ est convexe } \Leftrightarrow\ f(y) \geqslant f(x) + f'(x)(y-x)\ \forall x, y \in C$$

$$f \text{ est strictement convexe } \Leftrightarrow\ f(y) > f(x) + f'(x)(y-x)\ \forall x \neq y \in C$$



Preuve. Voir TD $\square$

Exercice. Soit E un e.v.

1. Vérifier que $f\colon E \to]-\infty, +\infty]$ est convexe ssi $\mathrm{epi}(f)$ est convexe (en particulier $\mathrm{dom}(f)$ est convexe).
2. Si $(f_i)_{i\in I}$ est une famille de fonctions convexes alors $\sup_{i\in I} f_i$ est convexe.

3.2 Ellipcité

Définition 4 (fonction elliptique)

Une fonction $f\colon V \to \mathbb{R}$ où V est un Hilbert est dite elliptique si f est $\mathcal{C}^1$ et s'il existe $\alpha > 0$ tel que $\forall x, y \in V$ on a :

$$\langle \underbrace{\nabla f(y)}_{\in V} - \nabla f(x), y - x \rangle_V \geqslant \alpha |x - y|_V^2$$

Notation. $\langle \nabla f(x), h \rangle = f'(x)h = df(x)h$. $df(x) = f'(x) \in \mathcal{L}(V, \mathbb{R}) = V'$ le dual topologique.

Propriété 2 (CNS d'ellipticité au 2eme ordre)

Si f est $\mathcal{C}^2$ alors f est elliptique ssi $f''(x)(v, v) \geqslant \alpha |v|_V^2$

Notation. $f''(x) = d^2 f(x) \in \mathcal{L}(V, \mathcal{L}(V, \mathbb{R})) \equiv \mathcal{L}(V \otimes V, \mathbb{R})$

Preuve.
$\gamma(t) = f(x + tv),\ x, v \in V$

$$\begin{aligned} \gamma'(t) - \gamma'(0) &= df(x + tv)v - df(x)v \\ &= (f'(\underbrace{x + tv}_{y}) - f'(x))v \\ t(\gamma'(t) - \gamma'(0)) &= \langle \nabla f(y) - \nabla f(x), \underbrace{x - y}_{tv} \rangle \\ &\geqslant \alpha t^2 |v|_V^2 \end{aligned}$$

d'où

$$\frac{\gamma'(t) - \gamma'(0)}{t} \geqslant \alpha |v|_V^2$$

Puis passage à la limite $t \to 0$, $t \neq 0$. Autre sens en exercice. □

3.3 Fonctions semi-continue inférieurement (s.c.i.)

Définition 5

On dit que $f\colon E \to]-\infty, +\infty]$ est s.c.i. en $x \in E$ si $\forall \varepsilon > 0$, $\exists U$ $(x \in U)$ tel que $\inf_U f \geqslant f(x) - \varepsilon$.

Propriété 3

Soit $f\colon E \to]-\infty, +\infty]$. Sont équivalents :

1. f est s.c.i.
2. $\forall \lambda \in \mathbb{R},\ (f \leqslant \lambda)$ est fermé.
3. $\operatorname{epi}(f)$ est fermé.

Preuve. Voir TD $\square$

Corollaire 1

Si $(f_i)_{i\in I}$ est une famille de fonctions s.c.i. de $E \to]-\infty,+\infty]$, $\sup_I f_i$ est s.c.i.

Preuve.

$$\mathrm{epi}(\sup_I f_i) = \bigcap_{i\in I} \underbrace{\mathrm{epi}(f_i)}_{\text{fermé}} \text{ fermé}$$

$\square$

Propriété 4

Si $f\colon E \to]-\infty,+\infty]$ est s.c.i. alors $\mathrm{dom}(f)$ est fermé.

Preuve. En exercice. $\square$

Théorème 2

Si $f\colon E \to]-\infty,+\infty]$ est s.c.i. et $K \subset E$ compacte, alors l'infimimum de f sur K est atteint i.e. $\exists x_K \in K$ tel que $f(x_K) = \inf_K f$.

Preuve.
Soit $(x_n)_{n\geqslant 0}$ une suite minimisante, on en extrait une sous-suite convergente $(x_{n_k})_{k\geqslant 0}$.
Si $x_K = \lim_{k\to+\infty} x_{n_k} \in K$ alors $f(x_K) \leqslant \underline{\lim} f(x_{n_k}) = \inf_K f$ et $\inf_K f \leqslant f(x_K)$. $\square$

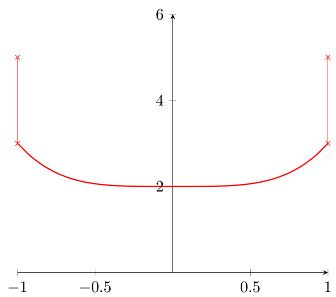
Remarque. Marge de progression utile en dimension infinie.

Théorème 3

Soit $f\colon V \to]-\infty,+\infty]$ avec V un Hilbert. Alors f est convexe s.c.i. ssi f est l'enveloppe supérieure de ses minorantes affines, i.e. :

$$f(x) = \sup_{(l,a)\in V'\times\mathbb{R},\ l+a\leqslant f} l(x) + a$$

Contre-exemple.



et $+\infty$ en dehors de $[-1,1]$. Cette fonction n'est pas s.c.i., on ne peut pas l'approcher par des droites aux points -1 et 1.

Preuve.
Si $\text{epi}(f) = \emptyset$ i.e. $f = +\infty$: ok. Sinon, soit $x \in \text{dom}(f)$ alors $\forall \lambda < f(x)$, $(x, \lambda) \notin \text{epi}(f)$. Or $\text{epi}(f)$ est un convexe fermé (f est convexe et s.c.i) d'où (projection sur les convexes fermés) il existe $(l, a) \in V' \times \mathbb{R}$ tel que $\forall y \in \text{dom}(f)$:

$$l(x) + a\lambda < l(y) + af(y)$$

Or pour $y = x$, on a $a\lambda < af(x)$, d'où $a > 0$. Par suite, quitte à diviser par a, on peut supposer que $a = 1$, ainsi :

$$l(x) + \lambda < l(y) + f(y) \text{ i.e. } f(y) > l(x - y) + \lambda$$

Vrai pour $y \in \text{dom}(f)$ mais aussi pour $y \notin \text{dom}(f)$.

$$f > \underbrace{-l + l(x) + \lambda}_{=\lambda \text{ en } x}$$

Comme λ est arbitraire, on a le résultat pour $x \in \text{dom}(f)$.
Si $x \notin \text{dom}(f)$, alors comme $\text{dom}(f)$ est un convexe fermé, on note $x_C = p_C(x)$ et $\lambda < f(x_C)$. $(x_C, \lambda) \cap \text{epi}(f) = \emptyset$. D'où il existe $(l, a) \in V' \times \mathbb{R}$ tel que :

$$l(x_C) + a\lambda < l(y) + af(y) \ \forall y \in \text{dom}(f)$$

De même en prenant $y = x_C$ on a $a > 0$ et donc on se ramène à :

$$l(x_C) + \lambda < l(y) + f(y) \ \forall \mathbf{y} \in \mathbf{V}$$

De plus comme $\langle x - x_C, y - x_C \rangle \leqslant 0 \ \forall y \in \text{dom}(f) = C$, on a

$$g_t(y) = l(x_C - y) + \lambda + t\underbrace{\langle x - x_C, y - x_C \rangle}_{\leqslant 0}$$

qui est une minorante affine affine de f. Donc :

$$g_t(x) = l(x_C - x) + \lambda + t|x - x_C|^2 \xrightarrow[t \to +\infty]{} +\infty$$

□

4 Conditions d'optimalité

Définition 6 (minimum local)

$J \colon X \longrightarrow \mathbb{R}$, X espace topologique.

- $x \in X$ est un minimum local de J si $\exists V$ un voisinage de X tel que $J(x) = \inf_V J$
- $x \in X$ est un minimum local strict de J si $\exists V$ un voisinage de X tel que :

$$J(x) < J(y) \ \forall y \in V \backslash \{x\}$$

- $x \in X$ est un minimum local de J par rapport à $U \subset X$ si $\exists V$ un voisinage de X tel que $J(x) = \inf_{V \cap U} J$

Théorème 4 (équation d'Euler)

Soit U un ouvert de E e.v. et $J \colon U \longrightarrow \mathbb{R}$ une fonction. Si $u \in U$ est un minimum local de J et J dérivable en u alors $J'(u) = 0$.

Remarque. C'est une condition nécessaire *du premier ordre* (CN1).

Preuve. Immédiat en considérant $\gamma(t) = J(u + tv)$ avec $v \in E$. $\gamma'(0) = J'(u)v = 0$ $\forall v \in E$. $\square$

Théorème 5 (CN2)

Si de plus J est deux fois dérivable en u alors $J''(u)(v, v) \geqslant 0$ $\forall v \in E$.

Preuve. Idem. $\square$

Théorème 6 (CS2)

$J : U \longrightarrow \mathbb{R}$ avec U ouvert de E e.v.n. On suppose que J est dérivable en u et $J'(u) = 0$.

1. Si $J''(x)$ existe et il existe $\alpha > 0$ tel que $J''(x)(u, u) \geqslant \alpha |v|^2$ $\forall v \in E$ alors x est un minimum local strict.
2. S'il existe B un ouvert contenant x tel que $\forall y \in B$ $J''(y)$ existe et $J''(y)(v, v) \geqslant 0$ $\forall v \in E$ alors x est un minimum local de J.

Preuve. cf. TD $\square$

Théorème 7

Soit $J : C \longrightarrow \mathbb{R}$ avec C un convexe de E et soit :

$$S = \{x \in C \mid J(x) = \inf J\}$$

l'ensemble des minima globaux sur C. On suppose qu'il existe $x_* \in C$ minimum local.

1. Si J est convexe alors S est convexe et $x_* \in S$ (i.e. un minimum local est global).
2. Si J est strictement convexe, alors $S = \{x_*\}$.

Preuve. cf. TD $\square$