MAT 2410: Optimisation

Robert Guénette

Département de Mathématiques et de Statistique

Chapitre 1
Minimisation d'une fonction numérique: existence et unicité

Références:

- Notes de cours: chapitre 1, sections 1.1 à 1.4.
- Livre de M. Delfour: chapitre 1.

Notations

- $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$
- Norme euclidienne: $\|x\| = \sqrt{\sum_{i=1}^d x_i^2}$
- Produit scalaire: $x \cdot y \equiv (x, y) \equiv x^t\, y = \sum_{i=1}^d x_i y_i$
- Distance entre 2 points: $d(x, y) = \|x - y\|$
- Boule ouverte: $B_r(a) = \{x \in \mathbb{R}^d \mid \|x - a\| < r\}$
- Boule fermée: $\overline{B_r(a)} = \{x \in \mathbb{R}^d \mid \|x - a\| \leq r\}$
- Ensemble ouvert: $U \subset \mathbb{R}^d$ est ouvert si

$$\forall x \in U \quad \exists B_r(x) \quad \text{tel que} \quad B_r(x) \subset U$$

- Ensemble fermé: $F \subset \mathbb{R}^d$ est fermé si $\mathbb{R}^d \setminus F$ est ouvert.

Notion de suite

Soit $\{x_n\}_{n=1}^{\infty}$ une suite de points de $\mathbb{R}^d$.

- On notera la convergence de cette suite par

$$x_n \to x \iff \lim_{n\to\infty} x_n = x \iff \lim_{n\to\infty} \|x_n - x\| = 0$$

- Critère de fermeture: $F \subset \mathbb{R}^d$ est fermé ssi

$$\forall x_n \to x \text{ et } x_n \in F \Longrightarrow x \in F$$

- Suite bornée: si $\exists M > 0$ tel que $\|x_n\| \leq M \qquad \forall n$
- Sous-suite: soit $\{n_k\}_{k=1}^{\infty}$ une suite strictement croissante d'indices

$$n_1 < n_2 < n_3 < \dots$$

La suite $\{x_{n_k}\}_{k=1}^{\infty}$ s'appelle une sous-suite de $\{x_n\}_{n=1}^{\infty}$.
Propriété: si $x_n \to x$, toutes les sous-suites convergent vers x.

Propriété fondamentale de $\mathbb{R}^d$

Théorème de Bolzano-Weierstrass

Toute suite bornée de $\mathbb{R}^d$ admet une sous-suite qui converge vers un point x de $\mathbb{R}^d$.

Remarque: si F est une partie fermée et bornée et $\{x_n\} \subset F$, alors le point limite de toute sous-suite convergente est dans F.

Infimum et suprémum

Soit $U \subset \mathbb{R}^d$ et $f : U \to \mathbb{R}$ une fonction numérique.

Par construction des nombres réels, l'ensemble $\{f(x) \,|\, x \in U\}$ admet un infimum $\alpha = \inf_{x \in U} f(x)$, i.e. la plus grande borne inférieure α vérifiant $\alpha \leq f(x) \quad \forall x \in U$.
Note: on peut avoir $\alpha = -\infty$.

De même, on peut définir le suprémum $\beta = \sup_{x \in U} f(x)$ comme la plus petite valeur β vérifiant $\beta \geq f(x) \quad \forall x \in U$.

Caractérisation de l'infimum

- $\alpha = \inf_{x \in U} f(x) > -\infty$

$$\forall \epsilon > 0 \quad \exists x \in U \text{ tel que } f(x) < \alpha + \epsilon$$

- $\alpha = \inf_{x \in U} f(x) = -\infty$

$$\forall N > 0 \quad \exists x \in U \text{ tel que } f(x) < -N$$

Existence des extréma: cas borné

Théorème d'existence des extréma (Weierstrass)

Soit $f : U \to \mathbb{R}$ une fonction numérique continue définie sur un domaine U fermé et borné de $\mathbb{R}^d$, alors

i) $\exists a \in U$ tel que $f(a) = \inf_{x \in U} f(x)$ ou encore $f(a) = \min_{x \in U} f(x)$

ii) $\exists b \in U$ tel que $f(b) = \sup_{x \in U} f(x)$ ou encore $f(b) = \max_{x \in U} f(x)$

Preuve: cas (i) seulement.
Supposons que $\alpha = \inf_{x \in U} f(x) > -\infty$.

$$\Longrightarrow \forall n \quad \exists x_n \in U \text{ tel que } f(x_n) < \alpha + 1/n$$

Or la suite $\{x_n\}$ est bornée car U est borné. Le théorème de Bolzano-Weierstrass implique l'existence d'une sous-suite convergente $x_{n_k} \to a$. Mais U est fermé, donc $a \in U$.

$$\begin{array}{rcl} f(x_{n_k}) & < & \alpha + 1/n_k \\ \downarrow & & \downarrow \\ \alpha \leq f(a) & \leq & \alpha + 0 \end{array}$$

d'où $f(a) = \alpha = \inf_{x \in U} f(x)$

Fonction à section bornée

Rappel: $\max_{x \in U} f(x) = -\min_{x \in U} -f(x)$

Par la suite, on se limitera uniquement au cas de la minimisation.

Soit $f : U \to \mathbb{R}$ une fonction numérique continue définie sur un domaine $U \subset \mathbb{R}^d$ fermé mais pas borné.

Fonction à section inférieure bornée: f est dite à section inférieure bornée si

$\exists k \in \mathbb{R}$ tel que $f^{-1}(-\infty, k] = \{x \in \mathbb{R}^d \mid f(x) \leq k\}$ soit borné et non vide

En notant $K = f^{-1}(-\infty, k]$, il est clair que nous avons la relation

$$\inf_{x \in U} f(x) = \inf_{x \in U \cap K} f(x)$$

Etant donné que K est fermé et borné, nous avons aussi que $U \cap K$ est fermé et borné. Ainsi le théorème d'existence des extréma s'applique.

Théorème d'existence: cas non borné

Soit $f : U \to \mathbb{R}$ une fonction numérique continue à section inférieure bornée et définie sur un domaine U fermé de $\mathbb{R}^d$, alors

$$\exists a \in U \text{ tel que } f(a) = \inf_{x \in U} f(x)$$

Fonction coercive

Soit $f : U \to \mathbb{R}$ une fonction numérique continue définie sur un domaine $U \subset \mathbb{R}^d$ fermé mais pas borné.

Fonction coercive: f est dite coercive si

$$\lim_{\substack{\|x\| \to \infty \\ x \in U}} f(x) = \infty$$

D'une manière formelle, ceci correspond à l'énoncé

$$\forall M > 0 \quad \exists N > 0 \text{ tel que } \|x\| > N \Longrightarrow f(x) > M.$$

Proposition: Toute fonction numérique continue et coercive est à section inférieure bornée.

Preuve: soit $k \in \mathbb{R}$ tel que $K = f^{-1}(-\infty, k]$ soit non vide. Montrons que K est borné.
Supposons le contraire.

$$\exists x_n \in K \text{ tel que } \|x_n\| \to \infty.$$

Ceci implique que $f(x_n) \to \infty$ car f est coercive. Mais $f(x_n) \leq k$ ce qui est impossible. Donc K est borné.

Ensemble convexe

Définition: $U \subset \mathbb{R}^n$ est dit convexe si

$$\forall x, y \in U \quad \forall \lambda \in [0,1] \Longrightarrow (1-\lambda)x + \lambda y \in U$$

Propriétés:

- Si $\{U_i\}_{i \in I}$ est une collection quelconque d'ensembles convexes, $\bigcap_{i \in I} U_i$ est convexe.
- La fermeture $\bar{U}$ d'un ensemble convexe est convexe.
- L'intérieur $int(U)$ d'un ensemble convexe est convexe ou vide.
- En général, la réunion d'ensembles convexes n'est pas convexe.

Exemples d'ensembles convexes:

- Hyperplan: $\{x \in \mathbb{R}^n \mid (x, a) = \alpha\}$ où $a \in \mathbb{R}^n$ $(a \neq 0)$ et $\alpha \in \mathbb{R}$
- Demi-espace: $\{x \in \mathbb{R}^n \mid (x, a) \leq \alpha\}$
- Intersection d'hyperplans: $\{x \in \mathbb{R}^n \mid Ax = b\}$ (si non vide) où A est une matrice de format $m \times n$ et $b \in \mathbb{R}^m$.
- Intersection de demi-espaces: $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ (si non vide)

Fonction convexe

Soit $f : U \longrightarrow \mathbb{R}$ une fonction numérique définie sur un ensemble convexe U.

Définition: f est dite convexe si

$$\forall x, y \in U \quad \forall \lambda \in [0,1] \Longrightarrow f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$$

Définition: f est dite concave si $-f$ est convexe.

Propriétés des fonctions convexes

Soient $f, g : U \longrightarrow \mathbb{R}$ des fonctions convexes.

- $f + g$ est une fonction convexe.
- $k\,f$ est une fonction convexe avec $k > 0$.
- $\forall \alpha \in \mathbb{R},\ f^{-1}]-\infty, \alpha] = \{x \mid f(x) \leq \alpha\}$ est un ensemble convexe.
- On définit l'épigraphe de f par

  $$epi(f) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha\}$$

  On a que

  $$f \text{ est convexe} \Longleftrightarrow epi(f) \text{ est convexe.}$$

- Si $\{f_i\}_{i \in I}$ est une collection de fonctions convexes, on a que la fonction $\sup\limits_{i \in I} f_i$ définie par

  $$\left(\sup_{i \in I} f_i\right)(x) = \sup_{i \in I} f_i(x) \qquad \text{est convexe.}$$

Norme vectorielle

Une fonction $p : \mathbb{R}^n \longrightarrow \mathbb{R}$ est une norme vectorielle sur l'espace $\mathbb{R}^n$ si elle vérifie les conditions suivantes, $\forall x, y \in \mathbb{R}^n$,

- $p(x) \geq 0$,
- $p(x) = 0 \iff x = 0$,
- $p(\alpha\, x) = |\alpha|\, p(x) \qquad \forall \alpha \in \mathbb{R}$,
- $p(x + y) \leq p(x) + p(y)$. (inégalité du triangle)

Voici les exemples les plus importants:

1. $\|x\|_\infty = \max_i |x_i|$ (norme du maximum)
2. $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ (norme l_1)
3. $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$ (norme euclidienne)
4. $\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$ (norme $p > 1$)

Produit scalaire

Un produit scalaire sur l'espace vectoriel $\mathbb{R}^n$ est une application bilinéaire

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

vérifiant les conditions suivantes, $\forall x, y, z \in \mathbb{R}^n$,

- $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle, \qquad \forall \alpha, \beta \in \mathbb{R}$
- $\langle y, x \rangle = \langle x, y \rangle$, (symétrie)
- $\langle x, x \rangle \geq 0$ et $\langle x, x \rangle = 0 \iff x = 0.$ (positivité)

Propriétés:

1. $\|x\| \equiv \sqrt{\langle x, x \rangle}$ est une norme vectorielle.
2. $|\langle x, y \rangle| \leq \|x\| \, \|y\|$ (inégalité de Cauchy-Schwartz)

Exemple très important:

$\langle x, y \rangle \equiv (x, y)_A \equiv (Ax, y)$ si A est une matrice symétrique définie-positive

Exemples de fonctions convexes

1. Fonction linéaire: $f(x) = (a, x) + \alpha$ où $a \in \mathbb{R}^n$ et $\alpha \in \mathbb{R}$.
2. La fonction quadratique: $f(x) = x^2$.
3. $f(x) = \|x\|_2$ à cause de l'inégalité du triangle.
4. Si $g : I = [a, b] \to \mathbb{R}$ est une fonction convexe et croissante

   $$\alpha \leq \beta \Longrightarrow g(\alpha) \leq g(\beta)$$

   et $f : U \subset \mathbb{R}^n \to \mathbb{R}$ est convexe avec $f(U) \subset I$, alors on a que $g \circ f : U \to \mathbb{R}$ est convexe.
5. $f(x) = \|x\|_2^2 = \sum_{i=1}^n |x_i|^2$
   On applique 4. avec $f = g \circ h$ et $g(y) = y^2$ et $h(x) = \|x\|_2$.
6. Toutes les normes vectorielles.
7. La fonction quadratique: $f(x) = (Ax, x)$ avec A une matrice symétrique définie-positive.
8. $f(x) = (Ax, x) + (b, x) + c$ où $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ et idem pour A.

Fonction strictement convexe

Soit $f : U \longrightarrow \mathbb{R}$ une fonction définie sur l'ensemble convexe U.

Définition: f est dite strictement convexe si

$\forall \lambda \in]0,1[\quad \forall x, y \in U \quad x \neq y$

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y)$$

Géométriquement, le graphe d'une telle fonction doit avoir une certaine **rondeur**.

Exemples et contre-exemples:

1. La fonction linéaire $f(x) = (a, x) + \alpha$ est convexe mais pas strictement convexe.
2. $f(x) = \|x\|$ est strictement convexe.
3. $f(x) = \|x\|_\infty$ n'est pas strictement convexe.
   En effet, il suffit de prendre les points $x = (-1, 1)$ et $y = (1, 1)$ et de vérifier que
   $$1 = \|\frac{x+y}{2}\|_\infty < \frac{\|x\|_\infty + \|y\|_\infty}{2} = 1$$
4. De même pour $f(x) = \|x\|_1$
5. Si f est strictement convexe et g est convexe, $f + g$ est strictement convexe.

Etude de la fonction quadratique

Soit A une matrice symétrique définie-positive de format $n \times n$.

Proposition: La fonction quadratique

$$f(x) = (Ax, x)$$

est strictement convexe.
Preuve: Posons $\|x\|_A^2 = (Ax, x) = f(x)$. Nous savons que $(x, y)_A = (Ax, y)$ défini un produit scalaire et que $\|x\|_A$ est la norme associée.

$$\begin{array}{rcl} \|(1-\lambda)x + \lambda y\|_A^2 & = & (1-\lambda)^2 \|x\|_A^2 + \lambda^2 \|y\|_A^2 + 2\lambda(1-\lambda)(x, y)_A \\ & < & (1-\lambda)^2 \|x\|_A^2 + \lambda^2 \|y\|_A^2 + 2\lambda(1-\lambda) \|x\|_A \|y\|_A \\ & & \text{par Cauchy-Schwartz} \\ & = & ((1-\lambda)\|x\|_A + \lambda\|y\|_A)^2 \\ & < & (1-\lambda)\|x\|_A^2 + \lambda\|y\|_A^2 \quad \text{car } g(t) = t^2 \text{ strict. convexe} \end{array}$$

À la ligne (2), on a égalité seulement si $x \parallel y \iff y = \alpha x$ et $\alpha > 0$. À la ligne (5), on a égalité seulement si $\|x\|_A = \|y\|_A$.

$$y = \alpha x, \quad \|x\|_A = \|y\|_A \Longrightarrow \alpha \|x\|_A = \|x\|_A \Longrightarrow \alpha = 1$$

donc $x = y$ ce qui prouve la stricte convexité de f.

Théorème d'unicité

Théorème d'unicité:
Soit $f : U \longrightarrow \mathbb{R}$ une fonction numérique strictement convexe définie sur un ensemble convexe U.
Si f admet un point minimisant, i.e. $f(a) = \inf_{x \in U} f(x)$, alors ce point est unique.
Preuve: Soient a_1 et a_2 deux points minimisants. Supposons que $a_1 \neq a_2$ avec

$$f(a_1) = f(a_2) = \inf_{x \in U} f(x) = \alpha.$$

Posons $a = \dfrac{a_1 + a_2}{2} \in U$ car U est convexe.

$$\Longrightarrow f(a) < \frac{f(a_1) + f(a_2)}{2} = \alpha$$

car f est strictement convexe.

$$\Longrightarrow \alpha = \inf_{x \in U} f(x) \leq f(a) < \alpha$$

ce qui est impossible. Donc $a_1 = a_2$.

Fonction semi-continue inférieure

Notation:

$$\bar{\mathbb{R}} = \mathbb{R} \cup \infty$$

On accepte les opérations suivantes:

- $\infty + \alpha = \infty \qquad \forall \alpha \in \mathbb{R}$
- $\alpha \times \infty = \infty \qquad \forall \alpha > 0$

Définition:
$f : U \longrightarrow \bar{\mathbb{R}}$ est dite semi-continue inférieure si une des conditions équivalentes est satisfaite

- $\forall k \in \mathbb{R} \qquad f^{-1}(-\infty, k] \subset \mathbb{R}^n$ est fermé.
- $epi(f) \subset \mathbb{R}^n \times \mathbb{R}$ est fermé.
- $\forall x_n \to a \quad \Longrightarrow \quad f(a) \leq \underline{\lim}\, f(x_n)$
  $\underline{\lim}\, f(x_n)$ est la plus petite limite parmi toutes les sous-suites convergentes.

Fonction semi-continue inférieure

Généralisation du théorème de Weierstrass: cas borné
Soit $f : U \longrightarrow \mathbb{R}$ une fonction propre semi-continue inférieurement définie sur un domaine U fermé et borné, alors il existe $a \in U$ tel que

$$f(a) = \inf_{x \in U} f(x)$$

De plus, si f est une fonction strictement convexe, on a unicité du point minimisant.

Fonction semi-continue inférieure

Généralisation du théorème de Weierstrass: cas non borné
Soit $f : U \longrightarrow \mathbb{R}$ une fonction propre semi-continue inférieurement, à section inférieure bornée définie sur un domaine U fermé et non borné, alors il existe $a \in U$ tel que

$$f(a) = \inf_{x \in U} f(x)$$

De plus, si f est une fonction strictement convexe, on a unicité du point minimisant.