

# Optimisation Convexe

Ivan Lejeune\*

2 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  $f: U \rightarrow \mathbb{R}$ . 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  $\varepsilon(x, h) \xrightarrow{h \rightarrow 0} 0$  et  $f'(x) \in \mathcal{L}(E, \mathbb{R})$ .

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

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

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

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

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

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

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

**Théorème.** Soit  $f \in \mathcal{C}^1(U, \mathbb{R})$ . 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. Charlier et M. Marche

**Formule de Taylor à l'ordre 2.** Soit  $f \in \mathcal{C}^2(U, \mathbb{R})$ ,  $x \in U$ . Alors il existe un voisinage  $\nu$  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 :

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

avec  $\varepsilon(h) \xrightarrow{\|h\| \rightarrow 0} 0$ , de manière à 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 :

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

**Definition.** Soit  $f: U \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$ . 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  $f: K \rightarrow \mathbb{R}$  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  $\bar{x} \in K$  tels que

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

**Preuve.** Ce résultat a 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

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

et

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

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

**Théorème 1.1.3.** Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  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

$$\forall x \in \mathbb{R}^n, \|x\| \geq \alpha \implies f(x) \geq f(0) + 1$$

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  $x_0$  ( $\forall x \in \overline{B}(0, \alpha), f(x) \geq f(x_0)$ )

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 mêmes hypothèses, au cas d'une fonction  $f: K \rightarrow \mathbb{R}$  avec  $K$  fermé de  $\mathbb{R}^n$ .

## 1.2 Caractérisation des extremas sans contraintes

Un outil fondamental au calcul différentiel.

**Théorème 1.2.1.** Soit  $U \subset \mathbb{R}^n$  ouvert et  $f: U \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$ . 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  $\varphi: I \subset \mathbb{R} \rightarrow \mathbb{R}$  qui admet par exemple un maximum local en  $0 \in I$ .

On a d'une part  $\varphi'(0) = \lim_{x \rightarrow 0^+} \frac{\varphi(x) - \varphi(0)}{x} \leq 0$  car  $x > 0$  et  $\varphi(x) - \varphi(0) \leq 0$

et d'autre part  $\varphi'(0) = \lim_{x \rightarrow 0^-} \frac{\varphi(x) - \varphi(0)}{x} \geq 0$  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  $\varphi_{e_i}(t) = f(x_0 + te_i)$

(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

$$\begin{aligned} |t| \leq r &\implies f(x_0 + te_i) \leq f(x_0) \\ &\implies \varphi'_{e_i}(0) = 0 \end{aligned}$$

Ainsi, toutes les dérivées partielles de  $f$  sont nulles en  $x_0$  et donc  $f'(x_0) = 0$