Département Sciences du Numérique

Calcul différentiel - Optimisation sans contraintes - Premiers algorithmes

O. Cots, J. Gergaud, S. Gratton, D. Ruiz et E. Simon

29 septembre 2020

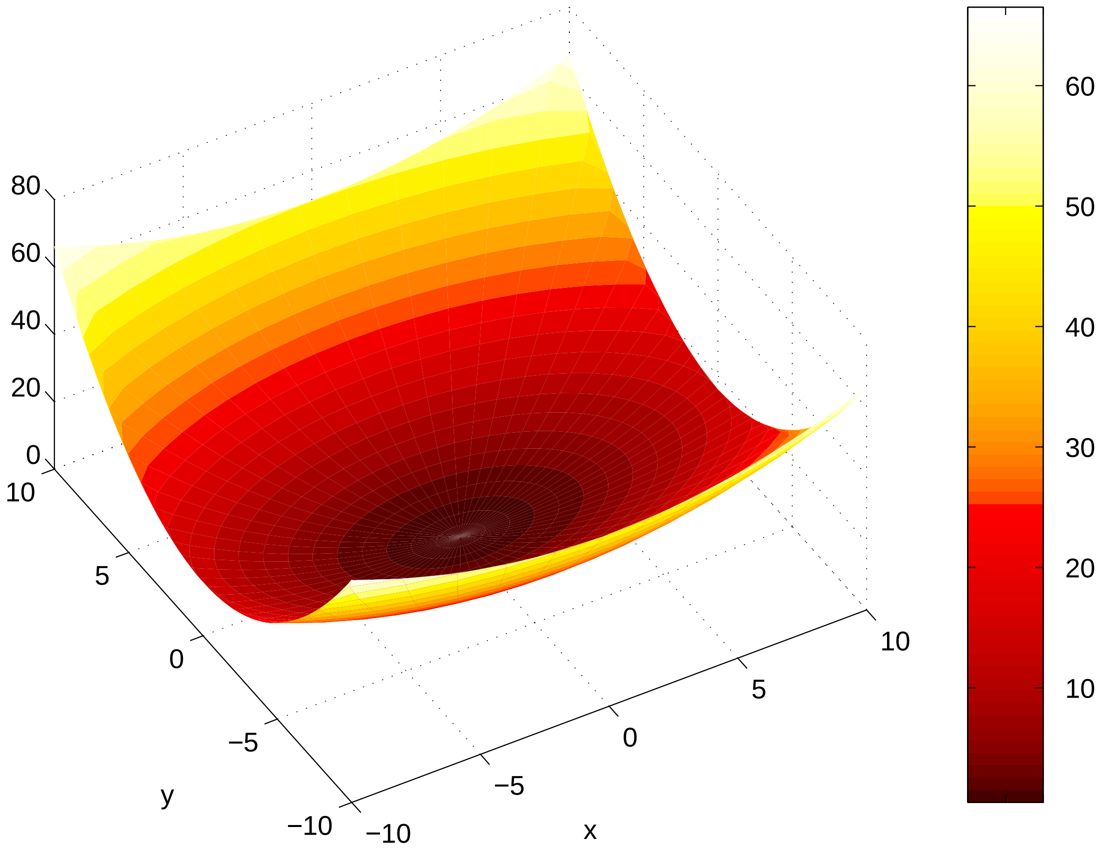

Table des matières

Chapitre 4

Existence de solution, unicité de solution

4.1 Introduction

Les problèmes d'optimisation où l'ensemble C est fini admettent toujours une solution, par contre, ceci n'est pas toujours le cas si C a un nombre infini d'éléments. Par exemple, le problème d'optimisation où la fonctionnelle à minimiser est $f(x) = 1/x$ et l'ensemble des contraintes est $C = \{x \in \mathbb{R}, x > 0\}$, n'admet pas de solution. En effet $f(x) > 0$ pour tout x dans C et pour tout $\varepsilon > 0$, il existe $x > 1/\varepsilon$ tel que $f(x) < \varepsilon$. Il est donc préférable, avant de vouloir calculer la solution, de s'assurer que le problème en admet une.

4.2 Existence de solution

4.2.1 Problèmes avec contraintes

Théorème 4.2.1

Soit (P) un problème d'optimisation avec contraintes $C \subset E$. Si f est continue et C est un compact non vide, alors le problème (P) admet une solution.

► C'est une application immédiate du théorème qui dit que l'image d'un compact par une application continue dans un espace séparé est un compact. ■

Remarque 4.2.1. On rappelle que, en dimension finie, un ensemble C est compact si et seulement si C est fermé et borné.

Exemple 4.2.1. Considèrons le problème suivant :

$$(P) \left\{ \begin{array}{l} \min \quad f(x) \\ x \in [0,1] \end{array} \right.$$

où f est la fonction suivante :

$$\begin{array}{rcl} f : [0,1] & \longrightarrow & \mathbb{R} \\ 0 & \longmapsto & 1 \\ x & \longmapsto & x \end{array}$$

Ce problème n'admet pas de solution. L'hypothèse du théorème (4.2.1) qui n'est pas vérifiée est la continuité de f. □

Exemple 4.2.2. Considérons le problème suivant :

$$(P) \left\{ \begin{array}{l} \min \quad f(x) = \frac{1}{x} \\ x \in [1,5] \end{array} \right.$$

(i) f est continue ;
(ii) $[1,5]$ est un fermé et borné, donc un compact de $\mathbb{R}$.

Par suite ce problème admet une solution. □

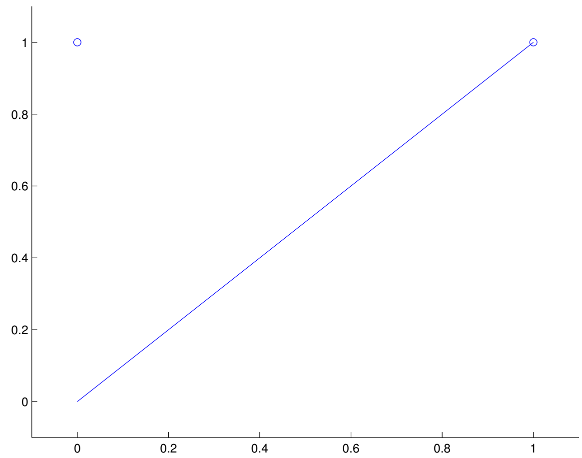

FIGURE 4.1 – *Exemple où f non continue*

Exemple 4.2.3. Considérons le problème suivant :

$$(P)\begin{cases} \min \quad f(x) = \frac{1}{x} \\ x \in]1,5] \end{cases}$$

Ce problème a une solution, mais les hypothèse du théorème ne sont pas vérifiées. □

Exemple 4.2.4. Considérons le problème suivant :

$$(P)\begin{cases} \min \quad f(x) = \frac{1}{x} \\ x \in [1,5[\end{cases}$$

Ce problème n'admet pas de solution, $C = [1, 5[$ n'est pas fermé. □

4.2.2 Problème sans contraintes

Définition 4.2.2

Une fonction $f : E \to \mathbb{R}$, E espace vectoriel normé, est dite $0-$coercive si et seulement si

$$f(x) \longrightarrow +\infty \text{ quand } \|x\| \longrightarrow +\infty. \tag{4.1}$$

Théorème 4.2.3

Soit (P) un problème d'optimisation avec contraintes où f est une fonction de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ et C est un fermé non vide. Si f est continue et 0-coercive, alors le problème admet une solution.

► Soit $(x_k)_{k\in\mathbb{N}^n}$ une suite minimisante de points de C, c'est-à-dire une suite de point de C telle que $\lim_{k\to+\infty} f(x_k) = \inf_{x\in\mathbb{R}^n} f(x) = \mu < +\infty$. Montrons que cette suite est bornée. Sinon il existe une sous-suite $(x_{n_k})_{n_k}$ telle que $||x_{n_k}||$ tende vers $+\infty$ lorsque n_k tend vers $+\infty$ et donc, comme f est 0-coercive, $\lim_{n_k\to+\infty} f(x_{n_k}) = +\infty$, ce qui est impossible.
Par suite il existe un réel $R > 0$ tel que la suite $(x_k)_{k\in\mathbb{N}}$ soit contenue dans $C \cap B_f(0, R)$ qui est un fermé borné de $\mathbb{R}^n$; c'est donc un compact dont on peut extraire une sous-suite qui converge vers x^*. Mais f est continue, et donc $f(x^*) = \mu$ et x^* est une solution du problème d'optimisation. ■

Remarque 4.2.2. Le théorème précédent s'applique si le problème d'optimisation est sans contraintes car dans ce cas $C = E$.

4.3 Cas convexe

Théorème 4.3.1

Si C est un convexe de E espace vectoriel normé et si f est une fonction de C à valeurs dans $\mathbb{R}$ convexe, alors l'ensemble des solutions est soit vide soit un ensemble convexe de E.

► Supposons que l'ensemble des solutions ne soit pas vide. Soient x et y deux solutions alors $f(x) = f(y)$ car $(f(x) \leq f(y)$ et $f(y) \leq f(x))$. Par suite, pour tout $\alpha \in]0, 1[$, nous avons

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \leq \alpha f(x) + (1-\alpha)f(x) \leq f(x).$$

En conséquence $\alpha x + (1-\alpha)y$ est aussi une solution. ■

Théorème 4.3.2

Si C est un convexe de E espace vectoriel normé et si f est une fonction de C à valeurs dans $\mathbb{R}$ strictement convexe, alors il existe au plus un point x^* minimisant f sur C.

► Supposons qu'il existe deux solutions x_1 et x_2. Pour $\alpha \in]0, 1[$, on pose $x_\alpha = \alpha x_1 + (1-\alpha)x_2$, alors, puisque f est strictement convexe on a

$$f(x_\alpha) < \alpha f(x_1) + (1-\alpha)f(x_2) = f(x_1) = f(x_2),$$

ce qui est impossible. ■

Théorème 4.3.3

Si C est un convexe de E espace vectoriel normé et si f est une fonction de C à valeurs dans $\mathbb{R}$ convexe, alors tout minimum local x^* de f sur C est un minimum global de f sur C.

► Soit x^* un minimum local de f sur C. Il existe donc $\eta > 0$ tel que pour tout $x \in C \cap B(x^*, \eta)$, $f(x^*) \leq f(x)$. Supposons maintenant qu'il existe dans C un point y tel que $f(y) < f(x^*)$. Alors, puisque f est convexe, on a pour tout $\alpha \in]0, 1[$

$$\begin{aligned} f(x^* + \alpha(y - x^*)) = f((1-\alpha)x^* + \alpha y) &\leq (1-\alpha)f(x^*) + \alpha f(y) \\ &< (1-\alpha)f(x^*) + \alpha f(x^*) = f(x^*). \end{aligned}$$

Mais pour α suffisamment proche de 0, $x^* + \alpha(y - x) \in B(x^*, \eta)$, d'où la contradiction. ■

Chapitre 5

Condition nécessaire, condition suffisante de solution Cas sans contraintes et cas de contraintes convexes

5.1 Condition du premier ordre

5.1.1 Cas sans contraintes

Théorème 5.1.1

Soient Ω un ouvert d'un espace vectoriel normé E et f une application de Ω à valeurs dans $\mathbb{R}$. Si f admet un minimum local en x^* et si f est dérivable en x^* alors on a l'équation parfois appelée équation d'Euler

$$f'(x^*) = 0. \tag{5.1}$$

► Soit $h \in E$, comme Ω est ouvert, il existe $\eta > 0$ tel que la fonction

$$\begin{array}{rcl} \varphi :]-\eta, \eta[& \longrightarrow & \mathbb{R} \\ t & \longmapsto & \varphi(t) = f(x^* + th) \end{array}$$

soit bien définie. φ est dérivable en 0 et $\varphi'(0) = f'(x^*) \cdot h$. Mais x^* est un minimum local de f, donc 0 est un minimum local de φ, par suite on a

$$0 \geq \lim_{t \to 0^-} \frac{\varphi(t) - \varphi(0)}{t} = \varphi'(0) = \lim_{t \to 0^+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0.$$

Ainsi, pour tout $h, \varphi'(0) = f'(x^*) \cdot h = 0$. ■

Définition 5.1.2 – Point critique

Un point qui vérifie $f'(x) = 0$ est dit un point critique et sa valeur en f, $f(x)$ une valeur critique.

5.1.2 Cas de contraintes convexes

Théorème 5.1.3

Soient Ω un ouvert d'un espace vectoriel normé E et $f : \Omega$ à valeur dans $\mathbb{R}$. Soit $C \subset \Omega$ convexe. Si f admet un minimum local en x^* sur C et si f est dérivable en x^* alors on a l'inéquation d'Euler

$$\forall y \in C,\ f'(x^*) \cdot (y - x^*) \geq 0. \tag{5.2}$$

► Soit $y \in C$, alors la fonction

$$\begin{array}{rcl} \varphi : [0,1] & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \varphi(t) = f(x^* + t(y - x^*)) \end{array}$$

est bien définie et admet une dérivé à droite en 0 $\varphi'^{+}(0) = f'(x^*) \cdot (y - x^*)$. Mais 0 est un minimum local de φ et donc $\varphi(0) \leq \varphi(t)$ pour t suffisamment proche de 0. Par suite

$$\varphi'^{+}(0) = \lim_{t \to 0^+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0.$$

■

Remarque 5.1.1. (i) Si C est un sous espace affine ($C = x_0 + V$, avec V sous-espace vectoriel de E) alors l'inéquation d'Euler (5.2) devient

$$\forall h \in V, f'(x^*) \cdot h = 0$$

(ii) Si $C = E$ alors l'inéquation d'Euler (5.2) devient l'équation d'Euler (5.1).

5.1.3 Problèmes convexes

Théorème 5.1.4

Soit $f : \Omega \subset E \to \mathbb{R}$, Ω ouvert d'un espace vectoriel normé E et soit $C \subset \Omega$ convexe. On suppose que f est convexe sur C et dérivable en tout point de C, alors les conditions suivantes sont équivalentes.

(i) x^* est un minimum global de f sur C.
(ii) x^* est un minimum local de f sur C.
(iii) Pour tout $y \in C$, $f'(x^*) \cdot (y - x^*) \geq 0$.

► (i)⇒ (ii) est évident.
(ii)⇒ (iii) est le théorème (5.1.2) précédent.
(iii)⇒ (i) ?
f est convexe, par suite (cf. le théorème 3.2.2) nous avons grâce à (*iii*)

$$f(y) \geq f(x^*) + f'(x^*) \cdot (y - x^*) \geq f(x^*).$$

■

Remarque 5.1.2. Si C est un ouvert convexe (iii) est équivalent à $f'(x^*) = 0$. L'équation d'Euler est donc dans ce cas une condition nécessaire et suffisante de solution.

Corollaire 5.1.5

On considère le problème au moindres carrés linéaire

$$(P) \begin{cases} \min \ f(\beta) = \frac{1}{2}\|y - X\beta\|^2 \\ \beta \in \mathbb{R}^p. \end{cases}$$

Alors β^* est une solution de (P) si et seulement si ce point vérifie les équations normales

$$X^T X\beta = X^T y. \tag{5.3}$$

► le problème est un problème convexe et on a $\nabla f(\beta) = X^T X\beta - X^T y$. ■

5.2 Conditions du deuxième ordre

5.2.1 Condition nécessaire

Théorème 5.2.1 – Condition nécessaire du deuxième ordre

Soit Ω un ouvert d'un espace vectoriel normé E et $f : \Omega \to \mathbb{R}$. Si x^* est un minimum local de f et si f est deux fois dérivables en x^* alors $f''(x^*)$ est semi-définie positive.

► Soit $h \neq 0$ un vecteur quelconque de E. x^* est un minimum local de f, donc $f'(x^*) = 0$ et il existe $\theta_0 > 0$ tel que pour tout $0 < \theta < \theta_0$ on ait $f(x^*) \leq f(x^* + \theta h)$. f étant deux fois dérivable en x^* on a par Taylor-Young

$$\begin{aligned} f(x^*+\theta h) - f(x^*) &= f'(x^*)\cdot h + \frac{\theta^2}{2} f''(x^*)\cdot(h,h) + \|\theta h\|^2\varepsilon(\theta h)) \\ &= \frac{\theta^2}{2}(f''(x^*)\cdot(h,h) + 2\|h\|^2\varepsilon(\theta h)) \geq 0. \end{aligned}$$

En divisant par θ^2 et en passant à la limite dans le membre de droite on en déduit que $f''(x^*)\cdot(h,h) \geq 0$. Ceci étant vrai pour tout h, on obtient le résultat. ■

Remarque 5.2.1. Il ne s'agit bien que d'une condition nécessaire (prendre $f(x) = x^3$).

5.2.2 Condition suffisante

Définition 5.2.2

Soit $B \in \mathcal{L}_2(E, \mathbb{R})$ une forme bilinéaire symétrique définie sur E, espace vectoriel normé.

(i) B est dite semi-définie positive si et seulement si pour tout $h \in E$

$$B(h,h) \geq 0.$$

(ii) B est dite définie positive si et seulement si pour tout $h \in E, h \neq 0$

$$B(h,h) > 0.$$

(iii) B est uniformément définie positive ou elliptique si et seulement si il existe $c > 0$ tel que pour tout $h \in E$

$$B(h,h) \geq c\|h\|^2.$$

Remarque 5.2.2. Si E est un espace vectoriel de dimension fini il y a équivalence entre la définie positivité et l'ellipticité.

Théorème 5.2.3 – Condition suffisante du deuxième ordre

Soit Ω un ouvert d'un espace vectoriel normé E et $f : \Omega \to \mathbb{R}$ dérivable sur Ω.

(i) Si x^* est un point de Ω tel que $f'(x^*) = 0$, f deux fois déribable en x^* et $f''(x^*)$ elliptique, alors x^* est un minimum local de f.

(ii) Si f est deux fois dérivable sur Ω et s'il existe une boule $B(x^*, \eta) \subset \Omega$ telle que, pour tout $x \in B(x^*, \eta), f''(x)$ est semi-définie positive et si $f'(x^*) = 0$, alors x^* est un minimum local de f.

► (i) La formule de Taylor-Young nous permet d'écrire pour tout h suffisamment petit

$$\begin{aligned} f(x^*+h) - f(x^*) &= \frac{1}{2} f''(x^*)\cdot(h,h) + \|h\|^2\varepsilon(h) \\ &\geq \frac{1}{2}(c + 2\varepsilon(h))\|h\|^2. \end{aligned}$$

Par suite il existe $\eta > 0$ tel que pour tout h tel que $\|h\| < \eta$ on ait

$$f(x^*+h) - f(x^*) > 0.$$

(ii) Si pour tout $x \in B(x^*, \eta), f''(x)$ est semi-postive, f est convexe sur l'ouvert $B(x^*, \eta)$. Par suite, comme $f'(x^*) = 0$, x^* est un minimum local de f.

■

Remarque 5.2.3. Pour les conditions nécessaires de solution du premier ordre et du deuxième ordre, on exploite le développement de Taylor-Young le long de toute direction h, combiné avec le fait que l'on a un minimum dans cette direction. Pour la condition suffisante du deuxième ordre, cette seule information, pourtant relative à toute direction h donnée, est insuffisante (cf. l'exercice 5.3.1, extrait de [5] page 50).

5.3 Exercices

✍ **Exercice 5.3.1**. On considère la fonction

$$\begin{array}{rcl} f : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) = 3x_1^4 - 4x_1^2 x_2 + x_2^2. \end{array}$$

1. Montrer que l'unique point critique de f est $\bar{x} = \begin{pmatrix} 0 & 0 \end{pmatrix}$.

2. La condition nécessaire de solution du deuxième ordre est-elle vérifiée ?

3. La condition suffisante de solution du deuxième ordre est-elle vérifiée ?

4. On fixe maintenant $d \in \mathbb{R}^2$ et on considèrela fonction

$$\begin{array}{rcl} \varphi : \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \varphi(t) = f(\bar{x} + td). \end{array}$$

Montrer que $t = 0$ est un minimum local de φ.

5. Calculer $f(x_1, 2x_1^2)$. Conclusion (faire le lien avec la remarque 5.2.3). □

✍ **Exercice 5.3.2**. Attention à l'intuition dans $\mathbb{R}$, cf. [5] page 52 !

1. On considère une fonction f de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ dérivable en tout point. On suppose que f admette un minimum local en $\bar{x}$ et que $\bar{x}$ est l'unique point critique de f. Démontrer que $\bar{x}$ est un minimum global de f.

2. On considère maintenant la fonction

$$\begin{array}{rcl} f : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) = 2x_1^3 + 3e^{2x_2} - 6x_1 e^{x_2} \end{array}$$

Montrer que $\bar{x} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ est l'unique point critique de f, que $\bar{x}$ est un minimum local de f, mais que f n'admet par de minimum global. □

✍ **Exercice 5.3.3**. Donner un exemple de fonction f (de $\mathbb{R}$ dans $\mathbb{R}$), deux fois dérivable, ayant un minimum strict en un point $\overline{x}$ et telle que dans toute boule $\mathcal{B}(\overline{x}, \rho)$ il existe un point $x \in \mathcal{B}(\overline{x}, \rho)$ vérifiant $f''(x) < 0$.
□

Bibliographie

[1] Carpentier. Cours de 3ième année enseeiht, filière informatique et mathématiques appliquées. INPT-ENSEEIHT, 1983. ↩ 12.

[2] Gerard Cornuejols and Reha Tütüncü. *Optimization Methods in Finance*. Cambridge University Press, 2007. ↩ 11.

[3] J.E. Dennis and Jr. Robert B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. SIAM, 1996. ↩ 56.

[4] J.M. Devos. Fermat "le premier homme du monde". France 3, IREM de Toulouse, CRDP Midi-Pyrénées, 1995. Casette vidéo. ↩ 4.

[5] Jean-Baptiste Hiriart-Urruty. *L'Optimisation*. Que sais-je. Presses Universitaires de France, 1996. ISBN : 2 13 047981 2. ↩ 50.