

# OPTIMISATION DÉTERMINISTE

## ING2 MATHS-INFO CYTECH

LOUIS GARRIGUE

Soit un ensemble  $K \subset \mathbb{R}^d$  une application  $f : K \rightarrow \mathbb{R}$  bornée inférieurement, c'est-à-dire qu'il existe  $c \in \mathbb{R}$  tel que  $f(x) \geq c$  pour tout  $x \in K$ . Optimiser  $f$ , c'est trouver le ou les minimums globaux de  $f$  sur  $K$ , c'est-à-dire trouver les solutions du problème

$$\inf_{x \in K} f(x),$$

s'il en existe. Ce problème très général a un ensemble indénombrable d'applications dans tous les domaines de l'ingénierie et des sciences, en intelligence artificielle, en statistiques, en finance, en physique, etc.

Nous ne nous concentrerons pas seulement sur des problèmes abstraits et théoriques, mais nous donnerons les concepts suffisants pour le travail d'un ingénieur, et nous présenterons les méthodes permettant de résoudre très concrètement ces problèmes. Les TPs associés auront de nombreuses applications numériques. Un bon compris entre théorie et application sera suivi pour une formation ingénieur.

Pour réviser et aller plus loin dans le cours, nous conseillons la lecture du livre

“Introduction à l'Optimisation”, Jean-Christophe Culioli,  
éditions Ellipses [1], Chapitres 1, 2 et 3

### CONTENTS

1. Rappels sur la différentiation	2
1.1. Dérivations d'ordre 1	2
1.2. Dérivation d'ordre 2	4
1.3. Développement de Taylor	4
1.4. Exemples de calculs	4
1.5. Rappels sur les matrices	6
2. Algorithme de descente générique	7
2.1. Descente	7
2.2. Algorithme générique	8
2.3. Test de convergence / arrêt de l'algorithme.	8
2.4. Convergence	9
2.5. Choix du pas	9
3. Vitesse de convergence	10
3.1. Cas $\gamma = \alpha = 1$	10
3.2. Remarque sur le cas $\gamma = 1$ et $\alpha < 1$	11
3.3. Remarque sur le cas $\gamma > 1$	11
3.4. Calculer les constantes en pratique	12

---

Date: September 30, 2025.

3.5.	Exemple de convergence	12
4.	Convexité	13
4.1.	Ensembles convexes	13
4.2.	Fonctions convexes	14
4.3.	Exemples	16
4.4.	Fonctions coercives	17
4.5.	Vocabulaire	18
5.	Les principales conditions d'optimalité	20
5.1.	$K$ ouvert	20
5.2.	$K$ fermé	23
5.3.	$K$ convexe	24
6.	Théorème KKT	25
	References	25

## 1. RAPPELS SUR LA DIFFÉRENTIATION

Nous aurons besoin de manipuler les dérivées. En effet, par exemple si  $f(x) = 3x^2 - x + 1$  et si on veut la minimiser sur  $\mathbb{R}$ , il faut d'abord trouver les points  $x \in \mathbb{R}$  tels que  $f'(x) = 0$ . Sur  $\mathbb{R}^d$ , la différentiation est plus délicate.

Généralement, on nomme “application” un élément qui à chaque élément d'un espace de départ, associe un élément d'un espace d'arrivée. Quand l'espace d'arrivée est  $\mathbb{R}$ , on parle aussi de “fonction”.

Dans toute cette section,  $\Omega \subset \mathbb{R}^d$  sera ouvert de  $\mathbb{R}^d$ .

### 1.1. Dérivations d'ordre 1.

1.1.1. *Gradient.* Rappelons que le gradient d'une application  $f : \Omega \rightarrow \mathbb{R}$ ,  $\mathcal{C}^1$  est l'application notée  $\nabla f : \Omega \rightarrow \mathbb{R}^d$  donnée par

$$\forall x = (x_1, \dots, x_d) \in \Omega, \quad (\nabla f)(x) := \begin{pmatrix} (\partial_{x_1} f)(x) \\ \vdots \\ (\partial_{x_d} f)(x) \end{pmatrix}. \quad (1)$$

Par exemple si  $d = 2$  et  $f(x) = x_1^2 + 3x_2$ , alors  $(\nabla f)(x) = \begin{pmatrix} 2x_1 \\ 3 \end{pmatrix}$ .

Nous noterons aussi  $\nabla f(x) := (\nabla f)(x)$  pour alléger le texte.

1.1.2. *Différentielle.* Prenons une fonction  $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ , donc plus générale que pour la définition du gradient. On rappelle la définition de la différentielle de  $f$ , notée  $df : \Omega \rightarrow \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ , où  $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$  est l'ensemble des applications linéaires allant de  $\mathbb{R}^d$  à  $\mathbb{R}^n$ .

**Définition 1.1** (Différentielle en  $x$ ). *Prenons  $x \in \Omega$ . Comme  $\Omega$  est ouvert, il existe  $r > 0$  tel que  $B(x, r) \subset \Omega$ . Supposons qu'il existe une application linéaire  $L : \mathbb{R}^d \rightarrow \mathbb{R}^n$  et  $c > 0$  et  $s \in [0, r]$  tels que pour tout  $y \in \mathbb{R}^d$  tel que  $\|y\| \leq s$ ,*

$$\|f(x + y) - (f(x) + Ly)\| \leq c \|y\|^2.$$

*Alors  $f$  est différentiable en  $x$  et on note  $d_x f = L$ .*

On dit que  $f$  est différentiable sur  $\Omega$  si elle l'est en tout point de  $\Omega$ .

La différentielle généralise les développements limités d'ordre 1, c'est-à-dire qu'on peut écrire que pour tout  $x \in \Omega, y \in \mathbb{R}^d$  tel que  $x + y \in \Omega$ ,

$$f(x + y) = f(x) + (d_x f) y + O(\|y\|^2), \quad (2)$$

quand  $\|y\| \rightarrow 0$ . Pour un  $x \in \Omega$  donné,  $d_x f$  est une application linéaire allant de  $\mathbb{R}^d$  dans  $\mathbb{R}^n$ . Si  $f$  et  $g$  sont deux applications composables et  $\mathcal{C}^1$ ,  $g \circ f$  est aussi  $\mathcal{C}^1$  et la règle de la chaîne donne

$$d_x(f \circ g)y = (d_{g(x)}f) \circ (d_x g)y. \quad (3)$$

**Exercice 1.2.** L'application exponentielle  $\exp : \mathcal{M}_d(\mathbb{R}) \rightarrow \mathcal{M}_d(\mathbb{R})$  est définie par  $\exp(A) := \sum_{k=0}^{+\infty} \frac{A^k}{k!}$ . Calculer sa différentielle en 0.

*Solution.* On prend  $A = 0$ . On a

$$\exp(A + H) = \exp H = 1 + H + \sum_{k=2}^{+\infty} \frac{H^k}{k!} = \exp A + H + \sum_{k=2}^{+\infty} \frac{H^k}{k!}.$$

On conjecture que  $(d_A \exp) H = H$ . On pose  $LH := H$ , qui est une application linéaire. On a

$$\begin{aligned} \|\exp(A + H) - ((\exp A) + LH)\| &= \left\| \sum_{k=2}^{+\infty} \frac{H^k}{k!} \right\| \leq \sum_{k=2}^{+\infty} \frac{\|H\|^k}{k!} = \|H\|^2 \sum_{k=0}^{+\infty} \frac{\|H\|^k}{(k+2)!} \\ &\leq \|H\|^2 \sum_{k=0}^{+\infty} \frac{\|H\|^k}{k!} = \|H\|^2 e^{\|H\|}. \end{aligned}$$

On peut donc écrire

$$\exp(A + H) = \exp(A) + LH + O_{\|H\| \rightarrow 0}(\|H\|^2),$$

donc  $\exp$  est différentiable en 0 et  $(d_0 \exp)H = LH = H$ , i.e.  $d_0 \exp = \mathbf{1}_{\mathcal{M}_d(\mathbb{R})}$ .  $\square$

1.1.3. *Dérivée.* Prenons un intervalle  $I$ . Pour les applications  $\xi : I \rightarrow \mathbb{R}^d$ , c'est-à-dire quand l'espace de départ est plongé dans  $\mathbb{R}$ , on définit la dérivée de  $\xi$  par

$$\xi'(t) := \frac{(d_t \xi)s}{s} \in \mathbb{R}^d$$

pour tout  $t \in I$  et tout  $s \in \mathbb{R}$ . Si  $\xi : I \rightarrow \mathbb{R}$ , la dérivée coïncide avec le gradient.

1.1.4. *Autre définition du gradient.* Reprenons une fonction  $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}$ . Sur tout point  $x \in \Omega$ , le gradient  $\nabla f(x)$  peut être en fait défini à partir de la différentielle par dualité (via le théorème de Riesz disant que les formes linéaires on un vecteur représentant) via

$$\forall y \in \mathbb{R}^d, \quad (d_x f)y = \langle \nabla f(x), y \rangle, \quad (4)$$

et on peut montrer qu'il est bien égal à celui donné en (1). La différentielle ne dépend pas du produit scalaire alors que le gradient si.

Lorsque ces quelques règles sont bien maîtrisées, on ne se trompe plus dans les calculs de dérivation de fonctions définies sur  $\mathbb{R}^d$ .

1.1.5. *Dérivée directionnelle.* La dérivée directionnelle généralise en dimension  $d$  la dérivée à gauche et à droite de la dimension 1. Dans cette définition,  $K$  n'est pas forcément un ensemble ouvert, c'est seulement un sous-ensemble de  $\mathbb{R}^d$ .

**Definition 1.3** (Dérivée directionnelle). Prenons  $K \subset \mathbb{R}^d$ ,  $x \in K$  et  $y \in \mathbb{R}^d$ . Si la limite

$$(\delta_x^+ f)(y) := \lim_{t \rightarrow 0^+} \frac{f(x + ty) - f(x)}{t}$$

existe, alors on dit que c'est la dérivée de  $f$  en  $x$  dans la direction  $y$ . On parle aussi de dérivée de Dini.

Par exemple pour  $f : [0, 1] \rightarrow \mathbb{R}$  définie par  $f(x) = x$ ,  $f$  n'est pas dérivable en 0 mais a une dérivée de Dini dans la direction 1 (c'est-à-dire à droite) en 0,  $(\delta_0^+ f)(1) = 1$ . Un des intérêts de la dérivée directionnelle est qu'elle peut être définie sur le bord  $\partial K$ . Si  $f$  est différentiable en  $x$ , alors pour tout  $y \in \mathbb{R}^d$ ,  $f$  a une dérivée de Dini en  $x$  dans la direction  $y$  et

$$(\delta_x^+ f)(y) = \langle \nabla f(x), y \rangle. \quad (5)$$

Mais si  $f$  a une dérivée en  $x$  dans toute direction  $y \in \mathbb{R}^d$ ,  $f$  n'est pas forcément différentiable, un contre-exemple est  $x \mapsto |x|$  en 0, qui a une dérivée à gauche et à droite mais n'est pas dérivable.

1.2. **Dérivation d'ordre 2.** Soit  $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}$  deux fois différentiable. On note  $\mathcal{S}_d(\mathbb{R})$  l'ensemble des matrices carrées réelles symétriques de taille  $d \times d$ . On appelle Hessienne de  $f$  l'application  $\nabla^{\otimes 2} f : \Omega \rightarrow \mathcal{S}_d(\mathbb{R})$ , où

$$(\nabla^{\otimes 2} f)(x) := \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \leq i, j \leq d}.$$

**Lemma 1.4.** Pour tout  $x \in \Omega$ , on a bien  $\nabla^{\otimes 2} f(x) \in \mathcal{S}_d(\mathbb{R})$ .

*Proof.* Pour tout  $i, j \in \{1, \dots, d\}$ ,  $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ , par le théorème de Schwartz.  $\square$

Comme  $\nabla^{\otimes 2} f(x)$  est symétrique, on peut la diagonaliser. Il existe une matrice orthogonale  $P(x) \in O_d(\mathbb{R})$  (c'est-à-dire que  $P(x)P(x)^T = 1$ ) et  $D(x) \in \mathcal{M}_d(\mathbb{R})$  diagonale telles que

$$\nabla^{\otimes 2} f(x) = P(x)D(x)P(x)^{-1}. \quad (6)$$

1.3. **Développement de Taylor.** On a, pour  $\|y\| \rightarrow 0$ , le développement limité d'ordre 2 suivant, pour toute direction  $y \in \mathbb{R}^d$ ,

$$\boxed{f(x + y) = f(x) + \langle \nabla f(x), y \rangle + \frac{1}{2} \langle y, (\nabla^{\otimes 2} f) y \rangle + O_{\|y\| \rightarrow 0}(\|y\|^3)}. \quad (7)$$

1.4. **Exemples de calculs.** Donnons quelques exemples simples.

1.4.1.  $t \mapsto f(ta + b)$ . Supposons  $\Omega = \mathbb{R}^d$  et  $a, b \in \mathbb{R}^d$ , et prenons  $f : \mathbb{R} \rightarrow \mathbb{R}$ . On définit  $w : \mathbb{R} \rightarrow \mathbb{R}$  par

$$w(t) := f(ta + b)$$

et on veut calculer sa dérivée. On peut poser  $m : \mathbb{R} \rightarrow \mathbb{R}^d$  définie par  $m(t) := ta + b$ , on a alors  $w = f \circ m$ , donc la règle de la chaîne (3) donne, pour tout  $t, s \in \mathbb{R}$ ,

$$(d_t w) s = (d_{m(t)} f) \circ (d_t m) s.$$

Or, on peut calculer que  $(d_t m)s = as$  donc

$$\begin{aligned} w'(t) &= \frac{(d_{m(t)} f) \circ (d_t m) s}{s} = \frac{(d_{m(t)} f) \circ as}{s} = (d_{m(t)} f) \circ a = \langle \nabla f(m(t)), a \rangle \\ &= \langle \nabla f(ta + b), a \rangle. \end{aligned}$$

Le calcul est plus simple qu'en passant par (1). De même, on peut calculer

$$w''(t) = \langle a, ((\nabla^{\otimes 2} f)(ta + b)) a \rangle.$$

1.4.2. *Norme.* Calculons la différentielle de l'application  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  définie par  $f(x) := \|x\|^2$ .

On a  $f(x) = f(x_1, \dots, x_d) = \sum_{k=1}^d x_k^2$  donc

$$(\partial_{x_j} f)(x) = \sum_{k=1}^d \partial_{x_j} x_k^2 = \sum_{k=1}^d 2\delta_{k=j} x_k = 2x_j$$

et

$$\nabla f(x) = \begin{pmatrix} (\partial_{x_1} f)(x) \\ \vdots \\ (\partial_{x_d} f)(x) \end{pmatrix} = \begin{pmatrix} 2x_1 \\ \vdots \\ 2x_d \end{pmatrix} = 2x.$$

On calcule  $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \sum_{k=1}^d \frac{\partial^2}{\partial x_i \partial x_j} x_k^2 = 2\delta_{i=j}$  donc  $\nabla^{\otimes 2} f(x) = 2 \times \mathbb{1} = 2$  où  $\mathbb{1}$  est la matrice identité.

1.4.3. *Formes quadratiques.*

**Proposition 1.1: Gradient et Hessienne d'une forme quadratique**

Soit  $b \in \mathbb{R}^d$  et  $c \in \mathbb{R}$ ,  $A \in \mathcal{M}_n(\mathbb{R})$  et  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  telle que  $\forall x \in \mathbb{R}^d$ ,

$$f(x) := \langle Ax, x \rangle + \langle b, x \rangle + c. \quad (8)$$

On a alors

$$\nabla f(x) = (A^T + A)x + b, \quad \text{et} \quad \nabla^{\otimes 2} f(x) = A^T + A.$$

Sous cette forme,  $f$  est appelée une forme quadratique.

*Proof.* On fait un calcul direct et un calcul via la différentielle. On verra que le calcul utilisant la différentielle est beaucoup plus simple et rapide.

• Calcul direct. On appelle  $g(x) := \langle Ax, x \rangle$  et  $h(x) := \langle b, x \rangle$ . Alors  $g(x) = \sum_{i=1}^d (Ax)_i x_i = \sum_{1 \leq i, j \leq d} A_{ij} x_j x_i$  donc

$$\begin{aligned} (\partial_{x_k} g)(x) &= \sum_{1 \leq i, j \leq d} A_{ij} \partial_{x_k} x_j x_i = \sum_{1 \leq i, j \leq d} A_{ij} (2x_k \delta_{k=i=j} + x_i \delta_{k=j, i \neq k} + x_j \delta_{k=i, j \neq k}) \\ &= 2A_{kk}x_k + \sum_{\substack{1 \leq i \leq d \\ i \neq k}} A_{ik}x_i + \sum_{\substack{1 \leq j \leq d \\ j \neq k}} A_{kj}x_j = \sum_{i=1}^d A_{ik}x_i + \sum_{j=1}^d A_{kj}x_j \\ &= (A^t x)_k + (Ax)_k. \end{aligned}$$

De même,  $h(x) = \sum_{i=1}^d b_i x_i$  donc  $\partial_{x_k} h(x) = b_k$ .

• En utilisant  $d_x f$ . On a

$$\begin{aligned} f(x+y) &= \langle A(x+y), x+y \rangle + \langle b, x+y \rangle + c \\ &= f(x) + \langle Ax, y \rangle + \langle Ay, x \rangle + \langle Ay, y \rangle + \langle b, y \rangle \\ &= f(x) + \langle (A + A^T)x + b, y \rangle + O(\|y\|^2). \end{aligned}$$

Donc pour que la définition (2) soit respectée, on déduit la différentielle  $(d_x f)y = \langle (A + A^T)x + b, y \rangle$ . Or, le gradient est défini via (4) donc  $\nabla f(x) = (A + A^T)x + b$ .

• Hessienne. On calcule

$$\frac{\partial^2 g}{\partial x_i \partial x_j}(x) = \sum_{1 \leq k, p \leq d} A_{kp} \frac{\partial^2}{\partial x_i \partial x_j} x_k x_p = \sum_{1 \leq k, p \leq d} A_{kp} \delta_{k=i, p=j} = A_{ij} + A_{ji} = (A + A^T)_{ij}.$$

donc  $\nabla^{\otimes 2} f(x) = A + A^T$ . □

1.4.4. *Forme quadratique particulière.* Si  $f(x_1, x_2) := x_1^2 + \lambda x_2^2$  où  $\lambda \in \mathbb{R}$ , alors  $\nabla f(x) = 2 \begin{pmatrix} x_1 \\ \lambda x_2 \end{pmatrix}$  and  $\nabla^{\otimes 2} f(x) = 2 \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$ .

## 1.5. Rappels sur les matrices.

**Définition 1.5** (Matrices symétriques). Prenons  $M \in \mathcal{M}_d(\mathbb{R})$ .

- Si  $M^T = M$ , on dit que  $M$  est symétrique et on note  $M \in \mathcal{S}_d(\mathbb{R})$
- Si  $M \in \mathcal{S}_d(\mathbb{R})$  et si toutes les valeurs propres de  $M$  sont positives, on dit que  $M$  est positive et on note  $M \in \mathcal{S}_d^+(\mathbb{R})$  ou  $M \geq 0$ , c'est équivalent à ce que  $\forall y \in \mathbb{R}^d, \langle y, My \rangle \geq 0$ .
- Si  $M \in \mathcal{S}_d(\mathbb{R})$  et si toutes les valeurs propres de  $M$  sont strictement positives, on dit que  $M$  est définie positive et on note  $M \in \mathcal{S}_d^{++}(\mathbb{R})$  ou  $M > 0$ , c'est équivalent à ce que  $\forall y \in \mathbb{R}^d, \langle y, My \rangle > 0$ .

Pour  $M$  symétrique, on a  $M = P^T D P$  où  $D = \text{diag}(\lambda_1, \dots, \lambda_d)$  et  $P$  est orthogonale, donc pour tout  $y \in \mathbb{R}^d$ ,

$$\langle y, My \rangle = \langle y, P^T D P y \rangle = \langle P y, D P y \rangle = \sum_{j=1}^d \lambda_j ((P y)_j)^2. \quad (9)$$

*Preuve dans la définition 1.5.* Montrons que  $M \in \mathcal{S}_d^+(\mathbb{R})$  est équivalent à  $\forall y \in \mathbb{R}^d, \langle y, My \rangle \geq 0$ .

Supposons que toutes les valeurs propres  $\lambda_1, \dots, \lambda_d$  sont positives. Soit  $y \in \mathbb{R}^d$ , alors on voit par (9) que  $\langle y, My \rangle \geq 0$ .

Supposons que pour tout  $y \in \mathbb{R}^d, \langle y, My \rangle \geq 0$ . Soit  $e_1, \dots, e_d \in \mathbb{R}^d$  la base canonique de  $\mathbb{R}^d$ . Soit  $j \in \{1, \dots, d\}$ , alors on voit par (9) que  $\langle P^{-1}e_j, MP^{-1}e_j \rangle = \lambda_j$  et donc  $\lambda_j \geq 0$ .  $\square$

## Partie algorithmique

### 2. ALGORITHME DE DESCENTE GÉNÉRIQUE

Numériquement, minimiser  $f$  sur  $K$  consiste souvent, par un algorithme, à démarrer d'un élément  $x_0 \in K$  et à générer une suite  $(x_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$  telle que

$$f(x_{n+1}) \leq f(x_n).$$

On espère que cette suite converge vers un minimum global.

**2.1. Descente.** L'intérêt principal de la définition de la dérivée de Dini 1.3 est que

- Si  $(\delta_x^+ f)(y) > 0$  alors  $f$  est croissante dans la direction  $y$ ,
- Si  $(\delta_x^+ f)(y) < 0$  alors  $f$  est décroissante dans la direction  $y$ .

Une notion encore plus faible est la direction de descente.

**Définition 2.1** (Direction de descente). Soient  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  et  $x \in \mathbb{R}^d$ . Le vecteur  $y \in \mathbb{R}^d$  est une direction de descente pour  $f$  à partir du point  $x$  si  $t \mapsto f(x + ty)$  est décroissante en  $t = 0$ , c'est-à-dire s'il existe  $\eta > 0$  tel que  $\forall t \in ]0, \eta]$ ,  $f(x + ty) < f(x)$ .

Si  $f$  est dérivable en  $x \in K$  dans la direction  $y \in \mathbb{R}^d$  et que  $(\delta_x^+ f)(y) < 0$ , alors  $y$  est une direction de descente. Mais  $y$  peut être une direction de descente sans que la dérivée de  $f$  en  $x$  dans la direction  $y$  soit définie. Penser par exemple à une fonction discontinue,  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(t) := 0$  si  $t \neq 0$  et  $f(0) := 1$ , alors  $f$  n'est pas dérivable en 0 à droite ou à gauche mais  $-1$  et  $1$  sont des directions de descente au point  $t = 0$ .

La dérivée directionnelle permet de savoir si une direction est de descente.

**Proposition 2.2.** Soient  $f : K \rightarrow \mathbb{R}$  dérivable en  $x \in K$  dans la direction  $y \in \mathbb{R}^d$ . Si  $(\delta_x^+ f)(y) < 0$  alors  $y$  est une direction de descente. Si  $y$  est une direction de descente alors  $(\delta_x^+ f)(y) \leq 0$ .

*Proof.* On a le développement

$$f(x + ty) = f(x) + t (\delta_x^+ f)(y) + O_{t \rightarrow 0}(t^2).$$

Si  $(\delta_x^+ f)(y) < 0$ , alors  $(\delta_x^+ f)(y) \neq 0$  et pour  $t > 0$  petit,  $f(x + ty) - f(x)$  a le signe de  $(\delta_x^+ f)(y)$ . Si  $y$  est une direction de descente, alors pour  $t > 0$  assez petit  $f(x + ty) - f(x) \leq 0$  et on déduit que  $(\delta_x^+ f)(y) \leq 0$ .  $\square$

En particulier, grâce à (5), le gradient permet de savoir si on est sur une direction de descente. Mais il permet encore mieux, il permet de donner la direction de plus forte descente. L'ensemble des directions possibles est donné par  $\mathbb{S} := \{y \in \mathbb{R}^d, \|y\| = 1\}$ , on ne retient pas la norme car c'est la direction qui importe.

**Proposition 2.3** (Direction de plus forte descente). *Soit  $f : K \rightarrow \mathbb{R}$  une fonction différentiable sur  $\overset{\circ}{K}$  et  $x \in \overset{\circ}{K}$ . Alors le problème*

$$\min_{\substack{y \in \mathbb{R}^d \\ \|y\|=1}} \langle \nabla f(x), y \rangle \quad (10)$$

*est minimisé pour  $y_* = -\frac{\nabla f(x)}{\|\nabla f(x)\|}$ , qui est la direction de plus forte descente.*

*Proof.* Le problème (10) est posé sur l'ensemble compact  $\mathbb{S}$  et la fonction  $y \mapsto \langle \nabla f(x), y \rangle$  est continue car linéaire, donc le problème admet un minimiseur global.

Pour tout  $y \in \mathbb{R}^d$  tel que  $\|y\| = 1$ , on a

$$\langle \nabla f(x), y_* \rangle = -\|\nabla f(x)\| = -\|\nabla f(x)\| \|y\| \underset{\text{Cauchy-Schwartz}}{\leq} \langle \nabla f(x), y \rangle.$$

□

**2.2. Algorithme générique.** Afin de générer une suite  $(x_n)_{n \in \mathbb{N}}$  telle que  $f(x_{n+1}) \leq f(x_n)$ , on va chercher une direction de descente et itérer. Le principe suivant sera suivi par les algorithmes que nous utiliserons. On considère qu'on dispose de fonctions indépendantes permettant d'évaluer  $f$  et  $\nabla f$  en n'importe quel point  $x \in K$ .

---

**Algorithm 1** Algorithme générique de descente

---

```

1: procedure DESCENT( $f, \nabla f, x_0$ )
2:    $k \leftarrow 0$ 
3:   while test de convergence non satisfait do
4:     Trouver une direction  $y_k \in \mathbb{R}^d$  telle que  $\langle \nabla f(x_k), y_k \rangle < 0$  (i.e. direction
       de descente)
5:     Choisir un pas  $s_k > 0$  tel que  $f(x_k + s_k y_k) < f(x_k)$ 
6:      $x_{k+1} = x_k + s_k y_k$ 
7:      $k \leftarrow k + 1$ 
8:   end while
9:   return  $x$ 
10: end procedure

```

---

Le choix du pas peut ne pas nécessiter de faire appel à  $f$ , dans ce cas l'algorithme demande seulement des appels à  $\nabla f$  et la fonction sera de la forme **descent**( $\nabla f, x_0$ ).

**2.3. Test de convergence / arrêt de l'algorithme.** Pour tout  $x, y \in \mathbb{R}^d$ , définissons leur distance relative

$$D(x, y) := \frac{\|x - y\|}{\frac{1}{2}(\|x\| + \|y\|)}$$



Soit  $x_*$  un point de minimum local de  $f$ . Le test d'arrêt idéal serait  $x_k = x_*$  mais bien sûr on ne connaît pas  $x_*$ . En pratique, on choisit une tolérance  $\varepsilon \in \mathbb{R}_+^*$  et on adopte un ou plusieurs critères d'arrêt parmi les suivants

- Optimalité :  $\|\nabla f(x_k)\| < \varepsilon$ ,
- Stagnation de la solution :  $D(x_{k+1}, x_k) < \varepsilon$ ,
- Stagnation de la valeur :  $D(f(x_{k+1}), f(x_k)) < \varepsilon$ ,
- Nombre d'itérations  $k$  dépassant un seuil fixé.

On privilégie les erreurs relatives aux erreurs absolues.

## 2.4. Convergence.

**Definition 2.4** (Algorithme globalement convergent). *Soit un algorithme itératif qui génère une suite  $(x_k)_{k \in \mathbb{N}}$  pour résoudre  $\min_{x \in K} f(x)$ , où  $f$  est  $\mathcal{C}^1(K)$ . L'algorithme est dit globalement convergent si, quel que soit le point initial  $x_0 \in K$ ,  $\|\nabla f(x_k)\| \xrightarrow[k \rightarrow +\infty]{} 0$ .*

Attention, cette notion ne garantit pas que la limite soit un minimum, même local. L'algorithme peut converger vers un point critique qui n'est pas un minimum, un point selle par exemple.

**2.5. Choix du pas.** Considérons que nous avons une méthode pour choisir la direction de descente  $y_k$ . Il reste maintenant à définir une stratégie de recherche linéaire pour le calcul du pas  $s_k$ . Nous donnons ici deux manières naturelles.

**2.5.1. Pas constant.** Il consiste simplement à prendre  $s_k = s$  constant pour tout  $k \in \mathbb{N}$ . La question est alors comment choisir un pas qui garantisse la convergence de l'algorithme ? Un pas trop petit aidera à la convergence mais il faudra beaucoup d'itérations alors qu'un pas trop grand ne fera pas converger.

**2.5.2. Pas optimal.** On peut choisir un pas qui rende la fonction à minimiser la plus petite possible dans cette direction. Cette méthode est appelée méthode à pas optimal. Il reste donc, à chaque itération de l'algorithme 1, à résoudre le problème

$$\inf_{t \in \mathbb{R}} f(x_k + sy_k),$$

et on appelle  $s_k = s$  un minimiseur global s'il existe.

**Lemma 2.5** (Pas optimal pour les formes quadratiques). *Soit  $A \in \mathcal{S}_d^{++}(\mathbb{R})$  et  $b \in \mathbb{R}^d$ , on considère la forme quadratique  $Q(x) := \frac{1}{2} \langle x, Ax \rangle - \langle b, x \rangle$ . Imaginons qu'on itère avec une formule  $x^{k+1} = x^k + sy^k$ , c'est-à-dire dans la direction  $y^k$ . Alors le pas optimal  $s^k$  est*

$$s^k = \frac{\langle b - Ax^k, y^k \rangle}{\langle y^k, Ay^k \rangle}. \quad (11)$$

*Proof.* On définit

$$g(s) := Q(x^k + sy^k) = Q(x^k) + s \langle x^k, Ay^k \rangle - s \langle b, y^k \rangle + \frac{1}{2} s^2 \langle y^k, Ay^k \rangle.$$

On a

$$g'(s) = \langle x^k, Ay^k \rangle - \langle b, y^k \rangle + s \langle y^k, Ay^k \rangle,$$

et  $g''(s) = \langle y^k, Ay^k \rangle > 0$ .  $g$  est donc fortement convexe et a donc un unique minimiseur global. On le trouve en résolvant  $g'(s^k) = 0$ , qui donne le résultat.  $\square$

La résolution de ce problème de minimisation unidimensionnel peut coûter cher en temps de calcul si on doit le trouver numériquement. Pour cette raison, on peut lui préférer parfois l'algorithme de gradient à pas constant.

### 3. VITESSE DE CONVERGENCE

Soit  $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^N$  une suite convergente donnée par un algorithme, on note  $x_* := \lim_{k \rightarrow +\infty} x_k$ .

Dans la plupart des cas rencontrés usuellement, les vitesses de convergence des algorithmes sont telles que

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^\gamma} \xrightarrow{k \rightarrow +\infty} \alpha.$$

pour  $\alpha > 0$  et  $\gamma \geq 1$ . Nous ne considérerons pas d'autres cas.

**Definition 3.1** (Vitesse de convergence). *Supposons qu'il existe  $\alpha > 0$  et  $\gamma \geq 1$  tels que*

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^\gamma} \xrightarrow{k \rightarrow +\infty} \alpha.$$

- Si  $\gamma = 1$  et  $\alpha < 1$ , on dit que la convergence est linéaire et  $\alpha$  est appelé taux de convergence
- Si  $\gamma = \alpha = 1$ , nous référons à la Section 3.1
- Si  $\gamma > 1$  on dit que la convergence est superlinéaire d'ordre  $\gamma$ . Si  $\gamma = 2$ , on dit que la convergence est quadratique.

Nous ne regarderons  $\gamma = 1$  et  $\alpha > 1$  car alors  $\|x_k - x_*\| \rightarrow +\infty$ .

**3.1. Cas  $\gamma = \alpha = 1$ .** Dans le cas où  $\gamma = \alpha = 1$ , on ne sait même pas *a priori* si  $x_k$  converge, par exemple  $x_k = k$  et  $x_k = 1/k$  sont dans cette situation. Nous pourrions essayer de conjecturer quelque chose de plus précis, comme dans la définition suivante.

**Definition 3.2** (Conjecture fréquente lorsque  $\gamma = \alpha = 1$ ). *Si  $\gamma = \alpha = 1$ , nous aurons souvent*

$$\|x_k - x_*\| \underset{k \rightarrow +\infty}{\sim} \frac{C}{k^p}, \tag{12}$$

où  $C > 0$  et  $p > 0$ .

Si on a (12), on voit que

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = \left(1 + \frac{1}{k}\right)^{-p} \xrightarrow{k \rightarrow +\infty} 1,$$

ce qui correspond à la grandeur de la Définition 3.1 avec  $\gamma = \alpha = 1$ .

### 3.2. Remarque sur le cas $\gamma = 1$ et $\alpha < 1$ .

#### Proposition 3.1: Convergence dans le cas $\gamma = 1$ et $\alpha < 1$

Supposons que

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \xrightarrow{k \rightarrow +\infty} \alpha.$$

où  $\alpha < 1$ , alors il existe  $\beta \in [\alpha, 1[$  et  $C > 1$  tel que pour  $k$  assez grand,

$$\|x_k - x_*\| \leq C\beta^k.$$

Si en plus  $\alpha > 0$ , alors  $\|x_k - x_*\| \geq C'(\beta')^k$  pour  $C' > 0$ ,  $\beta' \in ]0, \alpha]$  et  $k$  assez grand.

*Proof.* Posons  $S_k := \|x_k - x_*\|$ . Dans ce cas, il existe  $N \in \mathbb{N}$  tel que pour tout  $k \geq N$ ,  $\alpha - \varepsilon \leq \frac{S_{k+1}}{S_k} \leq \alpha + \varepsilon$ , où  $\alpha + \varepsilon < 1$ . Pour tout  $k \geq N$  on a alors

$$(\alpha - \varepsilon)^{k-N} S_N \leq S_k \leq (\alpha + \varepsilon)^{k-N} S_N$$

et donc en posant  $C_0 := \alpha^{-N} S_N$ ,  $C_- := (1 - \varepsilon/\alpha)^{-N}$  et  $C_+ := (1 + \varepsilon/\alpha)^{-N}$ , on a

$$C_- \left(1 - \frac{\varepsilon}{\alpha}\right)^k \leq \frac{S_k}{C_0 \alpha^k} \leq C_+ \left(1 + \frac{\varepsilon}{\alpha}\right)^k,$$

On ne peut pas conclure que  $\lim_{k \rightarrow +\infty} \frac{S_k}{C_0 \alpha^k} = 1$  sans ajouter d'autres conditions, mais on peut intuitivement penser que  $S_k \simeq C_0 \alpha^k$  (pas au sens des équivalents). En fait en définissant  $\beta := \alpha + \varepsilon < 1$  et  $C := C_0 C_+$ , on a  $S_k \leq C\beta^k$ . On peut faire pareil pour la borne inférieure.  $\square$

Si on a exactement  $\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = \alpha$ , alors  $S_k = S_0 \alpha^k$ . Dans tous les cas, on a une convergence exponentielle de  $\|x_k - x_*\|$  vers 0, même si on parle de vitesse de convergence linéaire !

**3.3. Remarque sur le cas  $\gamma > 1$ .** Dans le cas où  $\gamma > 1$ , en supposant que  $\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^\gamma} = \alpha$  exactement, posons  $S_k := \|x_k - x_*\|$ . Nous avons alors

$$\begin{aligned} S_k &= \alpha S_{k-1}^\gamma = \alpha^{1+\gamma} S_{k-2}^{\gamma^2} = \alpha^{1+\gamma+\gamma^2} S_{k-3}^{\gamma^3} = \dots = \alpha^{1+\gamma+\gamma^2+\dots+\gamma^{k-1}} S_0^{(\gamma^k)} = \alpha^{\frac{\gamma^k-1}{\gamma-1}} S_0^{(\gamma^k)} \\ &= e^{\gamma^k((\gamma-1)^{-1} \ln \alpha + \ln S_0) + (\gamma-1)^{-1} \ln \alpha}. \end{aligned}$$

On voit que  $S_k \rightarrow 0$  est équivalent à  $(\gamma - 1)^{-1} \ln \alpha + \ln S_0 < 0$ , lui-même équivalent à  $S_0 < \alpha^{\frac{1}{1-\gamma}}$ . Par exemple si  $\alpha = \gamma = 2$ , il faut que  $S_0 < 1/4$ . Ceci montre qu'intuitivement, il faudra partir avec  $u_0$  pas trop loin de la solution  $u$  pour que l'algorithme converge, sinon il divergera. Dans le cas de convergence vers 0, c'est une convergence super-exponentielle de  $\|x_k - x_*\|$  vers 0.

**Proposition 3.2: Convergence dans le cas  $\gamma > 1$** 

Supposons que

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^\gamma} \xrightarrow{k \rightarrow +\infty} \alpha.$$

où  $\gamma > 1$ , alors il existe  $\beta > 0$  et  $C > 0$  tels que pour  $k$  assez grand,

$$\|x_k - x_*\| \leq C\beta^{\gamma^k}.$$

Attention ce résultat ne précise pas si  $\beta < 1$ , qui est nécessaire pour avoir la convergence vers 0.

*Proof.* En reprenant le calcul précédent, on a, pour  $N$  assez grand,  $S_{k+1}/S_k^\gamma \leq \alpha + \varepsilon$  où  $\alpha + \varepsilon < 1$  (on considère que c'est vrai pour  $N \geq 0$  sans perte de généralité),

$$S_k \leq (\alpha + \varepsilon)S_{k-1}^\gamma \leq \dots \leq (\alpha + \varepsilon)^{\frac{\gamma^k - 1}{\gamma - 1}} S_0^{(\gamma^k)} = C(\alpha + \varepsilon)^{a\gamma^k}$$

où  $C := (\alpha + \varepsilon)^{(\gamma-1)^{-1}}$  et  $a := (\gamma - 1)^{-1} + \ln S_0$ , il reste à choisir  $\beta := (\alpha + \varepsilon)^a$ .  $\square$

**3.4. Calculer les constantes en pratique.** Pour évaluer le type de vitesse de convergence donné en Définition 3.2, on pourra commencer par calculer numériquement  $\gamma$  puis  $\alpha$ . En définissant  $S_k := \|x_k - x_*\|$  et  $J_k := S_{k+1}/S_k$ , on pourra évaluer numériquement les limites

$$\lim_{k \rightarrow +\infty} \frac{\ln J_{k+1}}{\ln J_k} = \gamma, \quad \text{puis} \quad \lim_{k \rightarrow +\infty} \frac{S_{k+1}}{S_k^\gamma} = \alpha.$$

Si  $\alpha = 1$ , on pourra conjecturer un comportement asymptotique de type Définition 3.2. On pourra alors numériquement calculer

$$p = - \lim_{k \rightarrow +\infty} (\ln(1 + k^{-1}))^{-1} \ln J_k = - \lim_{k \rightarrow +\infty} k \ln J_k, \quad (13)$$

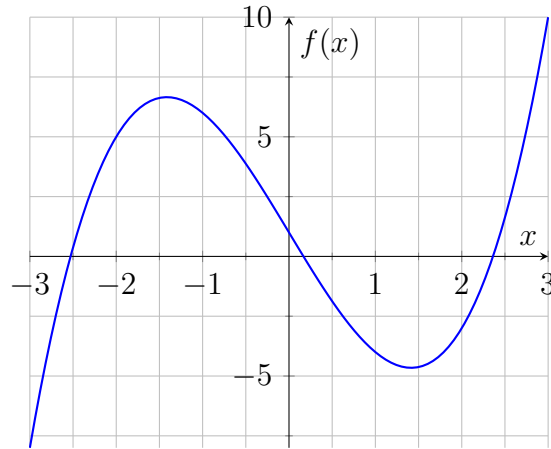
et enfin

$$C = \lim_{k \rightarrow +\infty} \frac{S_k}{k^p}.$$

**3.5. Exemple de convergence.** Considérons  $f(x) = x^3 - 6x + 1$ , tracée en Figure 1. L'unique minimum local est en  $x_* = \sqrt{2}$ . Partant de  $x_0 = 2$ , on compare plusieurs algorithmes

- $x_{k+1} = x_k - \mu(x_k^2 - 2)$ . Pour  $0 < \mu < 1/\sqrt{2}$ , convergence linéaire avec taux  $\alpha = |2\mu\sqrt{2} - 1|$ . Si  $\mu = (2\sqrt{2})^{-1}$ , la convergence est superlinéaire.
- $x_{k+1} = \frac{1}{2}(x_k + \frac{2}{x_k})$ . Convergence quadratique : en théorie 4 itérations suffisent pour obtenir 5 chiffres exacts. En pratique, on a 11 chiffres significatifs après 4 itérations.

## Partie théorique

FIGURE 1. Graph de la fonction  $f(x) = x^3 - 6x + 1$ 

## 4. CONVEXITÉ

## 4.1. Ensembles convexes.

**Définition 4.1** (Intervalle). *Un intervalle  $I \subset \mathbb{R}$  est un sous-ensemble connexe de  $\mathbb{R}$  qui n'est pas réduit à un singleton. Il est donc de la forme  $[b, +\infty[$  ou  $]b, +\infty[$  ou  $[a, b]$  ou  $[a, b[$  ou  $]a, b]$  ou  $]a, b[$  ou  $] - \infty, a]$  ou  $] - \infty, a[$ , où  $a, b \in \mathbb{R}$  et  $a < b$ .*

**Définition 4.2** (Ensemble convexe). *Un ensemble  $K \subset \mathbb{R}^d$  est dit convexe si*

$$\forall x, y \in K, \forall t \in [0, 1], \quad tx + (1 - t)y \in K.$$

Autrement dit, un ensemble est convexe si quand on prend deux points lui appartenant, le segment reliant ces deux points y est également. En dimension  $d = 1$ , les ensembles convexes sont les intervalles et les singletons. En figure 2, en dimension  $d = 2$ , nous dessinons deux exemples.

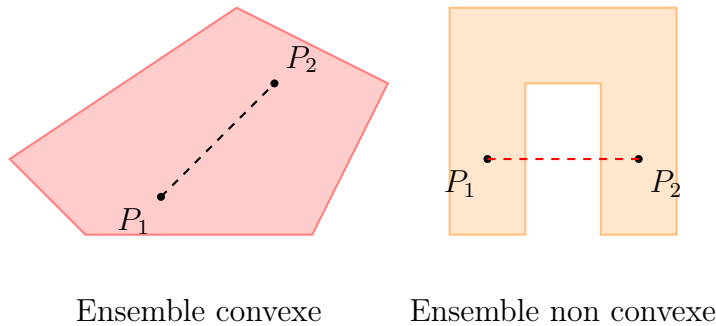


FIGURE 2. Exemples sur la convexité en dimension 2

## 4.1.1. Exemples.

- $\mathbb{R}^d$  est convexe
- la boule ouverte de centre  $a \in \mathbb{R}^d$  et de rayon  $R > 0$ ,  $B(a, R) \subset \mathbb{R}^d$ , est convexe

- n'importe quel sous-espace affine de  $\mathbb{R}^d$  est convexe
- pour  $d = 1$ , un ensemble  $C \subset \mathbb{R}$  est convexe ssi c'est un intervalle
- les demi-espaces affines

$$C := \{x \in \mathbb{R}^d \mid \langle a, x \rangle \leq \lambda\},$$

où  $a \in \mathbb{R}^d \setminus \{0\}$  et  $\lambda \in \mathbb{R}$ , sont convexes.

#### 4.2. Fonctions convexes.

**Définition 4.3** (Convexité des fonctions). Soit  $K \subset \mathbb{R}^d$  et une fonction  $f : K \rightarrow \mathbb{R}$  une fonction. On dit que  $f$  est

- *convexe* si

$$\forall x, y \in K, \forall t \in [0, 1], \quad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

- *strictement convexe* si

$$\forall x, y \in K \text{ tels que } x \neq y, \forall t \in ]0, 1[, \quad f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

- *fortement convexe* s'il existe  $\alpha > 0$  tel que

$$\forall x, y \in K, \forall t \in [0, 1], \quad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x - y\|^2$$

- *concave* si  $-f$  est convexe.

On a que fortement convexe  $\implies$  strictement convexe  $\implies$  convexe. En figure 3 nous donnons quelques exemples. Les fonctions affines sont les seules fonctions convexes et concaves. Visuellement, une fonction est convexe ssi son graph est au-dessus de tous ses plans tangents, ou encore ssi les arcs reliant deux points du graph sont au-dessus du graph.

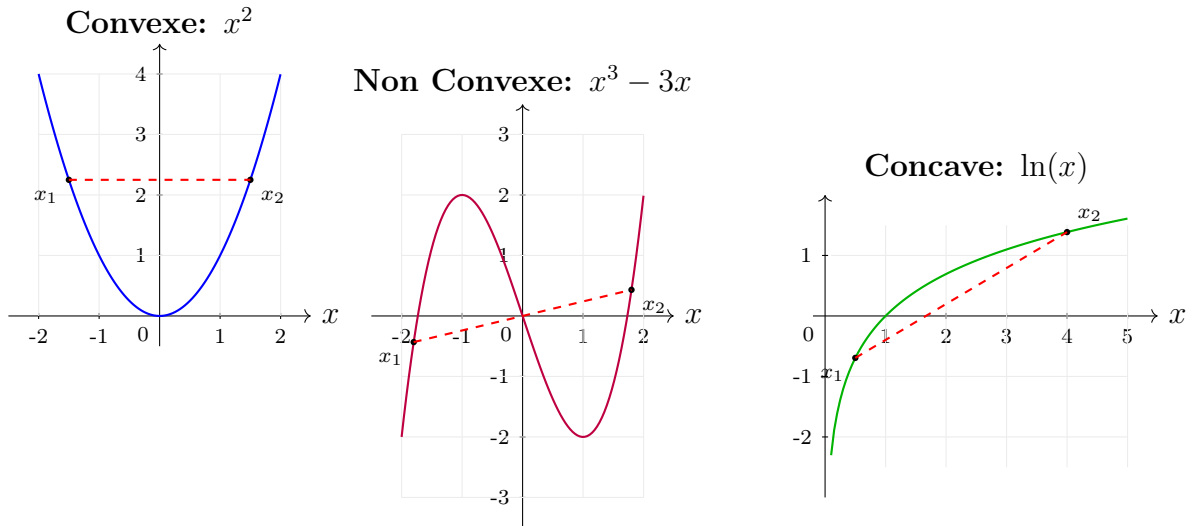


FIGURE 3. Exemples sur la convexité des fonctions en dimension un.

4.2.1. *Fonctions différentiables.* Lorsque  $f$  est  $\mathcal{C}^1$ , c'est-à-dire qu'elle est différentiable et que  $x \mapsto \nabla f(x)$  est continue, la convexité peut s'exprimer plus simplement en utilisant le gradient.

**Proposition 4.4** (Caractérisation de la convexité avec  $\nabla f$ ). *Soit  $K \subset \mathbb{R}^d$  un ensemble convexe et  $f : K \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$  dans un voisinage de  $K$ . Les assertions suivantes sont équivalentes :*

- (1)  $f$  est convexe sur  $K$ ,
- (2)  $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle, \forall x, y \in K$ ,
- (3)  $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0, \forall x, y \in K$ ,
- (4) (Si  $f$  est  $\mathcal{C}^2$ )  $\nabla^{\otimes 2} f(x) \geq 0, \forall x \in K$ .

La seconde assertion signifie que  $f$  est au-dessus de son développement d'ordre 1. On peut énoncer le même résultat en remplaçant les inégalités par des inégalités strictes dans les assertions précédentes.

*Proof.* Prouvons que (1) implique (2). On prend  $x, y \in K$ . On va se réduire au cas de dimension 1 en posant  $\xi(t) := f(x + t(y - x))$ . On a  $\xi(0) = f(x)$ ,  $\xi(1) = f(y)$ . Comme  $f$  est convexe,  $\xi(t) \leq t\xi(1) + (1-t)\xi(0)$ , donc la courbe de  $\xi$  est en-dessous de la droite passant par  $(0, \xi(0))$  et  $(1, \xi(1))$ . On a donc  $\frac{1}{t}(\xi(t) - \xi(0)) \leq \xi(1) - \xi(0)$  et en faisant  $t \rightarrow 0$  on obtient  $\xi'(0) \leq \xi(1) - \xi(0)$ . Or,  $\xi'(t) = \langle \nabla f(x + t(y - x)), y - x \rangle$  (voir Section (1.4.1)) donc  $\xi(0) = \langle \nabla f(x), y - x \rangle$  et on obtient le résultat.

Prouvons que (2) implique (3). On écrit (1) en échangeant  $x$  et  $y$  puis avec (1) on obtient

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x) \leq \langle \nabla f(y), y - x \rangle.$$

Prouvons que (2) implique (4). On prend  $y = x + tz$ , où  $t \in \mathbb{R}$ , on a

$$f(x + tz) = f(x) + t \langle \nabla f(x), y \rangle + \frac{t^2}{2} \langle z, (\nabla^{\otimes 2} f(x)) z \rangle + O(t^3)$$

donc

$$0 \leq \frac{1}{t^2} (f(x + tz) - f(x) - t \langle \nabla f(x), y \rangle) = \frac{1}{2} \langle z, (\nabla^{\otimes 2} f(x)) z \rangle + O(t)$$

on fait  $t \rightarrow 0$  et on obtient  $\langle z, (\nabla^{\otimes 2} f(x)) z \rangle \geq 0$ . C'est vrai pour tout  $z \in \mathbb{R}^d$  donc on peut conclure.  $\square$

La stricte convexité a la même caractérisation en remplaçant les inégalités par des inégalités strictes.

**Proposition 4.5** (Caractérisation de la forte convexité avec  $\nabla f$ ). *Soit  $K \subset \mathbb{R}^d$  un ensemble convexe et  $f : K \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$  dans un voisinage de  $K$ . Les assertions suivantes sont équivalentes :*

- $f$  est  $\alpha$ -fortement convexe sur  $K$ , avec  $\alpha > 0$
- $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2} \|y - x\|^2, \forall x, y \in K$
- $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \alpha \|y - x\|^2, \forall x, y \in K$ .
- (Si  $f$  est  $\mathcal{C}^2$ )  $\nabla^{\otimes 2} f(x) \geq \alpha, \forall x \in K$

La condition  $\nabla^{\otimes 2} f(x) \geq \alpha$  est équivalente à ce que les valeurs propres de  $A^T + A$  soient supérieures à  $\alpha$ , ou encore à ce que tous les éléments diagonaux de  $D(x)$  (définie en (6)) soient supérieurs à  $\alpha$ .

**Définition 4.6** (Fonction elliptique). *On appelle fonction elliptique une fonction  $f : K \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$  et fortement convexe.*

4.2.2. *Dimension 1.* Les propriétés se simplifient en dimension 1. Soit  $I$  un intervalle et  $f : I \rightarrow \mathbb{R}$  une fonction deux fois différentiable. Alors on a

$$\begin{aligned} f \text{ convexe} &\iff f' \text{ croissante} \iff f'' \geq 0, \\ f \text{ strictement convexe} &\iff f' \text{ strictement croissante} \iff f'' > 0, \\ f \text{ fortement convexe} &\iff f'' \geq \alpha > 0. \end{aligned}$$

**Proposition 4.7** (Propriétés des fonctions convexes en dimension 1). *Soit  $I$  un intervalle et  $f : I \rightarrow \mathbb{R}$  convexe. Alors  $f$  est continue et localement lipschitzienne sur  $\overset{\circ}{I}$ . En tout point  $a \in \overset{\circ}{I}$  (qui signifie l'intérieur de  $I$ ),  $f$  admet une dérivée à gauche  $f'_g(a)$  et une dérivée à droite  $f'_d(a)$ , pas forcément égales. De plus, pour tout  $a, b \in \overset{\circ}{I}$ , on a*

$$f'_g(a) \leq f'_d(a) \leq \frac{f(b) - f(a)}{b - a} \leq f'_g(b) \leq f'_d(b).$$

En particulier, les dérivées gauches et droites  $f'_g$  et  $f'_d$  sont croissantes. Nous ne donnons pas la preuve. La fonction  $f$  n'est pas forcément continue sur les bords de  $I$ .

### 4.3. Exemples.

**Proposition 4.8** (Combinaisons linéaires et compositions).

- (1) *Soient  $f_j : \mathbb{R}^d \rightarrow \mathbb{R}$  des fonctions convexes, et  $\alpha_j \in \mathbb{R}_+$ . Alors  $\sum_{j=1}^n \alpha_j f_j$  est convexe.*
- (2) *Soit  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  convexe et  $g : \mathbb{R} \rightarrow \mathbb{R}$  convexe croissante. Alors  $g \circ f$  est convexe.*

*Preuve de (2).* Comme  $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$  pour tout  $x, y \in \mathbb{R}$ , alors,

$$g \circ f(tx + (1-t)y) \underset{g \text{ croissante}}{\leq} g(tf(x) + (1-t)f(y)) \underset{g \text{ convexe}}{\leq} tg \circ f(x) + (1-t)g \circ f(y).$$

□

La composée de deux fonctions convexes n'est pas forcément convexe, par exemple  $f(x) = -x$  et  $g(x) = e^x$  sont convexes mais  $g \circ f(x) = e^{-x}$  est concave.

- (1) Les normes sont convexes

*Proof.* Soit  $N : \mathbb{R}^d \rightarrow \mathbb{R}$  une norme. Pour  $t \in [0, 1]$  on a

$$N(tx + (1-t)y) \underset{\substack{\text{ineg.} \\ \text{triang.}}}{\leq} N(tx) + N((1-t)y) = tN(x) + (1-t)N(y).$$

□

- (2) La fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) := x^2$  est fortement convexe avec  $\alpha = 2$ .



*Proof.* On a pour tout  $x \in \mathbb{R}$ ,  $\nabla f(x) = f'(x) = 2x$  et  $f''(x) = 2 > 0$ .  $\square$

- (3) La fonction  $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$ ,  $f(x) := x^p$  est convexe ssi  $p \geq 1$  ou  $p \leq 0$ .

*Proof.* On a  $f''(x) = p(p-1)x^{p-2}$ . Or,  $p(p-1) \geq 0$  ssi  $p \geq 1$  ou  $p \leq 0$ .  $\square$

- (4) La fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) := e^{\lambda x}$  avec  $\lambda > 0$  est strictement convexe mais pas fortement convexe.

*Proof.* On a  $f''(x) = \lambda^2 e^{\lambda x} > 0$  mais  $f''(x) \rightarrow 0$  quand  $x \rightarrow -\infty$ .  $\square$

- (5) La fonction  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $f(x) := \|x\|^2$  est fortement convexe avec  $\alpha = 2$ .

*Proof.* On a  $\nabla f(x) = 2x$ , on déduit  $\langle \nabla f(y) - \nabla f(x), y - x \rangle = 2\|y - x\|^2$ .  $\square$

- (6) Plus généralement, soit  $f$  une forme quadratique avec les mêmes notations que la proposition 1.1. On a

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle = \langle (A + A^T)(y - x), y - x \rangle = 2 \langle A(y - x), (y - x) \rangle.$$

Par conséquent,

- (a)  $A$  est semi-définie positive  $\Leftrightarrow f$  convexe.
- (b)  $A$  est définie positive de plus petite valeur propre  $\lambda_{\min} > 0 \Leftrightarrow f$  fortement convexe avec  $\alpha = \lambda_{\min}$ .

#### 4.4. Fonctions coercives.

**Definition 4.9** (Coercivité). Soit  $\Omega \subset \mathbb{R}^d$  un ensemble non borné (par exemple,  $\mathbb{R}^d$ ). Une fonction  $f : \Omega \rightarrow \mathbb{R}$  est dite coercive sur  $\Omega$  si

$$\lim_{\substack{x \in \Omega \\ \|x\| \rightarrow +\infty}} f(x) = +\infty.$$

Ceci peut s'écrire de façon équivalente :

- $\forall (x_k)_{k \in \mathbb{N}} \in \Omega^{\mathbb{N}}, \quad \|x_k\| \rightarrow +\infty \implies f(x_k) \rightarrow +\infty,$
- $\forall M > 0, \exists R > 0, \forall x \in \Omega \quad \|x\| \geq R \implies f(x) \geq M.$

La coercivité ne dépend pas de la norme choisie. Comme toutes les normes sont équivalentes sur  $\mathbb{R}^d$ , en pratique, on choisit la norme la plus adaptée à la fonction  $f$  étudiée.

On verra que la coercivité est importante dans le théorème 5.4.

**Proposition 4.10.** Soit  $K \subset \mathbb{R}^d$  un ensemble convexe non borné. Si  $f : K \rightarrow \mathbb{R}$  est  $\mathcal{C}^1$  et fortement convexe sur  $K$ , alors  $f$  est coercive sur  $K$ .

On peut résumer en disant “fortement convexe  $\implies$  coercive”.

*Proof.* La seconde assertion de la Proposition 4.5 pour les fonctions fortement convexes nous permet d'écrire pour tout  $x, y \in K$ ,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2} \|y - x\|^2. \quad (14)$$

De plus, par l'inégalité de Cauchy-Schwarz,

$$\langle \nabla f(x), y - x \rangle \leq \|\nabla f(x)\| \|y - x\|, \quad \text{donc} \quad \langle \nabla f(x), y - x \rangle \geq -\|\nabla f(x)\| \|y - x\|.$$

En injectant la dernière inégalité dans (14),

$$f(y) \geq f(x) + \|y - x\| \left( \frac{\alpha}{2} \|y - x\| - \|\nabla f(x)\| \right).$$

En fixant  $x \in K$  et en faisant  $\|y\| \rightarrow +\infty$  (ce qui implique  $\|y - x\| \rightarrow +\infty$ ), on obtient le résultat.  $\square$

Exemples de fonctions coercives:

(1) Par la Proposition 4.10, les formes quadratiques, i.e. les fonctions de la forme

$$f(x) = \langle Ax, x \rangle + \langle b, x \rangle + c, \quad (15)$$

sont fortement convexes, donc coercives, si  $A$  est définie positive.

(2) Toute fonction minorée par une fonction coercive est coercive.

Un autre exemple simple nécessite une preuve.

**Lemma 4.11.** *Soient  $f_i : \mathbb{R} \rightarrow \mathbb{R}$  des fonctions minorées et coercives. La fonction  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  donnée par  $f(x) = \sum_{i=1}^n f_i(x_i)$  est coercive.*

*Proof.* Pour tout  $i \in \{1, \dots, n\}$ ,  $f_i$  est minorée par une constante  $m_i$ . Posons  $m = \max_{1 \leq i \leq n} |m_i|$  et soit  $M > 0$  fixé. Comme chaque  $f_i$  est coercif, il existe une constante  $R_i > 0$  telle que

$$\forall x_i \in \mathbb{R}, \quad |x_i| \geq R_i \Rightarrow f_i(x_i) \geq M + nm.$$

Soit  $R = \max_{1 \leq i \leq n} R_i$ . Alors pour tout  $x \in \mathbb{R}^d$  avec  $\|x\|_\infty \geq R$ , il existe un indice  $i \in \{1, \dots, n\}$  tel que  $|x_i| \geq R \geq R_i$  et donc  $f_i(x_i) \geq M + nm$ . Pour  $j \neq i$ ,

$$f_j(x_j) \geq m_j \geq -m.$$

On a donc  $f(x) \geq M$  et on a montré que

$$\forall M \geq 0, \exists R > 0 \text{ tel que } \forall x \in \mathbb{R}^d, \|x\|_\infty \geq R \Rightarrow f(x) \geq M.$$

Par conséquent,  $\lim_{\|x\|_\infty \rightarrow \infty} f(x) = +\infty$  ce qui prouve la coercivité de  $f$ .  $\square$

**4.5. Vocabulaire.** Cherchons les minimas locaux de  $f : K \rightarrow \mathbb{R}$  où  $K \subset \mathbb{R}^d$ .

**4.5.1. Sans ou sous contraintes.**

- Si  $K$  est un ouvert de  $\mathbb{R}^d$ , on parle de minimisation sans contrainte pour le problème  $\inf_{x \in K} f(x)$ . Les extremas sont alors dit libres car ils vérifient l'équation d'Euler (16).
- Sinon, on parle de minimisation sous contrainte. Les extremas sont alors dit liés.

Par exemple si  $K = \{x = (x_1, \dots, x_d) \in \mathbb{R}^d \mid x_1 = 0, x_2 \geq x_3\}$  on est dans le cadre d'une minimisation sous contraintes.

On parle d'optimisation convexe si  $f$  est une fonction convexe et si  $K$  est un ensemble convexe.

4.5.2. *Infimum.* Dans cette section on prendra toujours

$$f : K \rightarrow \mathbb{R}, \quad \inf_{x \in K} f(x).$$

On cherche à minimiser  $f$  sur  $K$ .

**Definition 4.12** (Fonction minorée). *On dit que  $f : K \rightarrow \mathbb{R}$  est minorée sur  $K \subset \mathbb{R}^d$  s'il existe  $m \in \mathbb{R}$  tel que  $\forall x \in K, m \leq f(x)$ . On dit que  $m$  est un minorant de  $f$  sur  $K$ .*

**Definition 4.13** (Infimum). *Supposons qu'il existe  $\xi \in ]-\infty, +\infty[ \cup \{-\infty\}$  tel que*

- $\forall x \in K, \xi \leq f(x)$
- *il existe  $(x_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$  telle que  $f(x_n) \xrightarrow{n \rightarrow +\infty} \xi$ .*

*Alors  $\xi$  est appelé infimum de  $f$  sur  $K$  et noté  $\xi = \inf_{x \in K} f(x) = \inf_K f$ .*

L'infimum sur  $K$  existe toujours, il est fini si et seulement si  $f$  est minorée sur  $K$ . Si  $f$  n'est pas minorée sur  $K$ , alors  $\inf_K f = -\infty$ .

**Definition 4.14** (Suite minimisante). *Soit  $(x_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$  telle que*

$$f(x_n) \xrightarrow{n \rightarrow +\infty} \inf_K f.$$

*Alors  $(x_n)_{n \in \mathbb{N}}$  est appelée suite minimisante du problème de minimisation.*

4.5.3. *Minimum.*

**Definition 4.15** (Minimum global). *S'il existe  $x_* \in K$  tel que*

$$\forall x \in K, \quad f(x_*) \leq f(x),$$

*alors on dit que  $f(x_*)$  est un minimum global de  $f$  sur  $K$ , on note  $f(x_*) = \min_{x \in K} f(x) = \min_K f$ . On dit que  $f$  atteint son minimum en  $x_*$  et que le problème admet  $x_*$  comme solution.*

Par abus de langage, on dit parfois que  $x_*$  est un minimum mais on devrait toujours dire que  $x_*$  est un minimiseur. Un minimum existe si et seulement si un minimiseur existe. Le minimum n'existe pas toujours, et on va chercher à trouver des conditions de son existence. S'il existe, le minimum est un infimum.

**Definition 4.16** (Minimum local). *Supposons qu'il existe  $x_* \in K$  et un voisinage  $V$  de  $x_*$  dans  $\Omega$  tels que*

$$\forall x \in V \cap K, \quad f(x_*) \leq f(x).$$

*Alors  $x_*$  est appelé minimum local de  $F$  sur  $K$  et  $f(x_*)$  est appelé minimum local.*

Un minimum global est un minimum local mais le contraire est faux en général. Quand on dit juste "minimum", c'est implicitement qu'on dit minimum global.

## 5. LES PRINCIPALES CONDITIONS D'OPTIMALITÉ

5.1.  $K$  ouvert.

Pour tout ensemble  $\Omega \subset \mathbb{R}^d$ , on appelle  $\partial\Omega := \overline{\Omega} \setminus \overset{\circ}{\Omega}$  son bord.

**Théorème 5.1: Condition suffisante,  $K$  ouvert**

On suppose que  $K$  est un ouvert borné, que  $f$  est continue sur  $\overline{K}$ , et qu'il existe un point  $x_0 \in K$  tel que  $\forall x \in \partial K, f(x_0) < f(x)$ . Alors  $f$  admet un minimum global sur  $K$ .

La condition  $\forall x \in \partial K, f(x_0) < f(x)$  signifie que si  $f$  a un minimum sur  $\overline{K}$ , il ne sera pas atteint sur le bord, mais dans l'intérieur.

*Proof.* Comme l'ensemble  $\overline{K}$  est compact, la fonction continue  $f$  admet un minimum  $x_*$  sur  $\overline{K}$  par le théorème de Weierstrass (qu'on reverra au théorème 5.4). Cet élément est donc tel que

$$\forall x \in \overline{K}, f(x_*) \leq f(x).$$

Montrons par l'absurde que  $x_* \notin \partial K$ . Supposons que  $x_* \in \partial K$ . On a  $f(x_*) \leq f(x)$  pour tout  $x \in \overline{K}$  donc  $f(x_*) \leq f(x_0)$ . Or,  $f(x) > f(x_0)$  pour tout  $x \in \partial K$ , donc pour  $x = x_*$ ,  $f(x_*) > f(x_0)$ . Cette contradiction nous permet de conclure que  $x_* \notin \partial K$  et donc  $x_*$  est bien dans l'ouvert  $K$ .  $\square$

**Exercice 5.1.** Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x) := x_1^2 + x_2^2 + e^{x_2}$ . Montrer que  $f$  a un minimum sur  $B(0, 1)$  la boule unité ouverte.

*Solution.* On peut paramétrer le bord  $\partial B(0, 1)$ , on a, pour tout  $\theta \in [0, 2\pi[$ ,

$$g(\theta) := f(\cos \theta, \sin \theta) = 1 + e^{\sin \theta} > 1 = f(0).$$

On applique le théorème 5.1 avec  $x_0 = 0$ .  $\square$

**Théorème 5.2: Condition nécessaire,  $K$  ouvert**

Soit  $K \subset \mathbb{R}^d$  un ouvert et  $f : K \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$ . Si  $x_*$  est un minimiseur local de  $f$  sur  $K$ , alors

$$\nabla f(x_*) = 0. \tag{16}$$

Si en plus  $f$  est  $\mathcal{C}^2$  au voisinage de  $x_*$ , alors

$$\nabla^{\otimes 2} f(x_*) \geq 0. \tag{17}$$

*Proof.* Soit  $y$  un vecteur quelconque de  $\mathbb{R}^d$ . Comme  $x_* \in K$  et que  $K$  est ouvert, il existe  $h_0 > 0$  assez petit tel que  $B(x_*, h_0) \subset K$ , donc le point  $x_* + hy$  appartient à  $K$ . Or  $x_*$  étant un minimum local du problème, on a  $0 \leq f(x_* + hy) - f(x_*)$ . Comme

$$f(x_* + hy) - f(x_*) = h \langle \nabla f(x_*), y \rangle + O(h^2),$$

en divisant l'inégalité ci-dessus par  $h > 0$ , on obtient  $0 \leq \langle \nabla f(x_*), y \rangle + O(h)$  donc en faisant  $h \rightarrow 0$  on a

$$0 \leq \langle \nabla f(x_*), y \rangle.$$

Comme cette inégalité est vraie pour tout  $y \in \mathbb{R}^d$ , elle est également vraie pour  $-y$ . Donc  $\langle \nabla f(x_*), -y \rangle \geq 0$ , d'où  $\langle \nabla f(x_*), y \rangle = 0$ . On a montré

$$\langle \nabla f(x_*), y \rangle = 0, \quad \forall y \in \mathbb{R}^d,$$

c'est-à-dire que  $\nabla f(x_*) = 0$ . Si  $f$  est  $\mathcal{C}^2$ , on a en plus

$$\begin{aligned} f(x_* + hy) - f(x_*) &= \langle \nabla f(x_*), hy \rangle + \frac{1}{2} \langle hy, \nabla^{\otimes 2} f(x_*) hy \rangle + O(h^3) \\ &= \frac{1}{2} \langle hy, \nabla^{\otimes 2} f(x_*) hy \rangle + O(h^3) \end{aligned}$$

On divise par  $h^2$  et on obtient  $h^{-2}(f(x_* + hy) - f(x_*)) = \frac{1}{2} \langle y, \nabla^{\otimes 2} f(x_*) y \rangle + O(h)$ , or  $x_*$  est un minimiseur local donc  $f(x_* + hy) - f(x_*) \geq 0$  et en faisant tendre  $h \rightarrow 0$  on obtient la conclusion.  $\square$

(16) est appelée équation d'Euler. Ce théorème est très important, il montre que quand on est **sur un ouvert, il faut chercher les minimiseurs parmi les points annulant le gradient**. Cette procédure donne les minimums locaux, il faudra les comparer entre eux pour trouver le/les minimums globaux.

On notera

$$Z := \{x \in K \mid \nabla f(x) = 0\} = K \cap ((\nabla f)^{-1}(\{0\})) \quad (18)$$

l'ensemble des zéros du gradient, aussi appelé ensemble des points critiques. En définissant

$$\mathcal{M} := \{x \in \mathbb{R}^2 \mid x \text{ est un minimiseur local}\},$$

le théorème dit que

$$\mathcal{M} \subset Z$$

mais ces ensembles ne sont pas égaux en général. En effet,  $Z$  contient

- les minimums locaux
- les maximums locaux
- les points selle

**Exercice 5.2.** Trouver les minimiseurs locaux et les minimiseurs globaux de  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$f(x) := x_1^2 + 2x_1x_2 - x_2^2 + x_2^4.$$

*Solution.* On cherche  $\mathcal{M}$ , par le théorème 5.2 on sait que  $\mathcal{M} \subset Z$  donc on commence par chercher  $Z$ . On calcule

$$\nabla f(x) = 2 \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 + 2x_2^3 \end{pmatrix}, \quad \nabla^{\otimes 2} f(x) = 2 \begin{pmatrix} 1 & 1 \\ 1 & 6x_2^2 - 1 \end{pmatrix}.$$

On commence par l'analyse. Si  $x \in Z$ , alors  $x_1 = -x_2$  et  $x_1 - x_2 + 2x_2^3 = 0$ , ce qui implique  $x_1 = -x_2$  et  $x_2(x_2^2 - 1) = 0$  et donc en définissant  $S := \{(0, 0), (1, -1), (-1, 1)\}$  on a  $Z \subset S$ . On vérifie facilement que  $S \subset Z$  et donc  $Z = S$ . Trouvons  $\mathcal{M}$  par deux méthodes différentes.

• Méthode utilisant la Hessienne. Comme  $\mathcal{M} \subset Z$ , il faut vérifier pour quels points de  $Z$  la Hessienne est définie positive. On a

$$\nabla^{\otimes 2} f(0, 0) = 2 \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \nabla^{\otimes 2} f(1, -1) = \nabla^{\otimes 2} f(-1, 1) = 2 \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}.$$

On voit que  $\nabla^{\otimes 2} f(0,0)$  a une valeur propre négative donc  $(0,0)$  n'est pas un minimiseur local, alors que les valeurs propres de  $\nabla^{\otimes 2} f(1,-1)$  sont toutes les deux positives, donc  $(1,-1)$  et  $(-1,1)$  sont des minimiseurs locaux.

• Méthode sans la Hessienne. On a  $f(0, x_2) - f(0,0) = x_2^2(x_2^2 - 1)$  qui est négatif pour  $x_2$  petit, donc  $(0,0)$  n'est pas un minimiseur local. On a

$$\begin{aligned} f(x_1, x_2) - f(1, -1) &= f(x_1, x_2) - f(-1, 1) = x_1^2 + 2x_1x_2 - x_2^2 + x_2^4 + 1 \\ &\geq -2x_2^2 + x_2^4 + 1 \geq 0, \end{aligned}$$

donc  $(1,-1)$  et  $(-1,1)$  sont des minimiseurs locaux.

Dans les deux cas on trouve donc  $\mathcal{M} = \{(1,-1), (-1,1)\}$  et on remarque que  $\mathcal{M} \neq Z$ ,  $(0,0)$  est en fait un point selle.

On définit

$$\mathcal{G} := \{x \in \mathbb{R}^2 \mid x \text{ est un minimiseur global}\}.$$

On sait que  $\mathcal{G} \subset \mathcal{M}$ , on a  $f(1,-1) = f(-1,1)$ . On ne peut pas conclure parce que  $f$  pourrait décroître en l'infini. Par l'inégalité de Young avec  $p = 3$  et  $q = 3/2$ , on a  $|2x_1x_2| \leq \frac{2}{3}(|x_1|^{3/2} + 2|x_2|^3)$  donc

$$f(x) \geq \left(x_1^2 - \frac{2}{3}|x_1|^{3/2}\right) + \left(x_2^4 - x_2^2 - \frac{4}{3}|x_2|^3\right).$$

On voit que  $f$  est coercive et donc  $\mathcal{G} = \mathcal{M}$ . □

#### 5.1.1. Condition d'ordre 2.

##### Théorème 5.3: Condition suffisante d'ordre 2

On suppose que  $K$  est un ouvert et que  $f$  est  $\mathcal{C}^2$  au voisinage de  $x_* \in K$ . On suppose aussi que  $\nabla f(x_*) = 0$ . Si  $\nabla^{\otimes 2} f(x_*)$  est définie positive, alors  $x_*$  est un minimum local.

*Proof.* Prenons  $y \in \mathbb{R}^d$ , on a  $\langle y, \nabla^{\otimes 2}(x_*)y \rangle > 0$ . Prenons  $t > 0$  assez petit pour que  $x_* + ty$  reste dans  $K$ . Le développement d'ordre 2 donne

$$f(x_* + ty) - f(x_*) = \frac{t^2}{2} \langle y, \nabla^{\otimes 2}(x_*)y \rangle + O(t^3)$$

On divise par  $t^2$  et on obtient  $t^{-2}(f(x_* + ty) - f(x_*)) = \frac{1}{2} \langle y, \nabla^{\otimes 2}(x_*)y \rangle + O(t)$ , donc pour  $t$  assez petit, on a  $t^{-2}(f(x_* + ty) - f(x_*)) \geq 0$ . □

Si on a une direction  $y \in \mathbb{R}^d$  dans laquelle la Hessienne est dégénérée, i.e. si  $\langle y, \nabla^{\otimes 2} f(x_*)y \rangle = 0$ , alors  $f(x_* + y) = f(x_*) + O(\|y\|^3)$ . L'ordre 2 ne suffit alors pas pour connaître le signe de  $f(x_* + y) - f(x_*)$ . Il faut donc regarder l'ordre 3.

Pour un point critique  $x_* \in K$ ,

- Si  $\nabla^{\otimes 2} f(x_*)$  a toutes ses valeurs propres strictement positives alors  $x_*$  est un minimum local
- Si  $\nabla^{\otimes 2} f(x_*)$  a toutes ses valeurs propres strictement négatives alors  $x_*$  est un maximum local
- Si  $\nabla^{\otimes 2} f(x_*)$  a au moins une valeur propre nulle, on ne peut pas conclure pour l'existence d'un maximum ou minimum local

- Si  $\nabla^{\otimes 2} f(x_*)$  a au moins une valeur propre positive et au moins une négative, alors  $x_*$  est un point selle

## 5.2. $K$ fermé.

### **Théorème 5.4: Condition suffisante, $K$ fermé**

Soit  $K \subset \mathbb{R}^d$  un ensemble non-vidé et fermé dans  $\mathbb{R}^d$  et  $f : K \rightarrow \mathbb{R}$  une fonction continue. Si  $K$  est borné ou  $f$  est coercive sur  $K$ , alors  $f$  admet un minimum global sur  $K$ .

C'est pour ce théorème que la coercivité est une propriété importante des fonctions. Rappelons que dans le cas où  $K$  est borné,  $f$  a aussi un maximum sur  $K$ .

*Proof.* Dans le premier cas,  $K$  est fermé et borné, donc il est compact. Comme  $f$  est continue, le théorème de Weierstrass assure que  $f$  est bornée sur  $K$  et elle atteint ses bornes. Donc il existe au moins un minimiseur.

Prouvons le second cas. Soit  $x_0 \in K$  fixé. La coercivité de  $f$  sur  $K$  entraîne qu'il existe  $r > 0$  tel que

$$\forall x \in K, \quad \|x\| \geq r \Rightarrow f(x) > f(x_0). \quad (19)$$

On note  $\overline{B(0, r)}$  la boule fermée de  $\mathbb{R}^d$  de centre 0 et de rayon  $r$ . Sans perte de généralité, on prend  $r$  assez grand pour que  $x_0 \in \overline{B(0, r)}$ . Nous avons

$$\inf_{x \in K} f(x) = \inf_{x \in K \cap \overline{B(0, r)}} f(x).$$

Comme l'ensemble  $K \cap \overline{B(0, r)}$  est fermé et borné et que  $f$  est continue, le théorème de Weierstrass assure que  $f$  atteint ses bornes dans  $K \cap \overline{B(0, r)}$ . Ceci assure l'existence d'un minimum  $x_*$  dans  $K \cap \overline{B(0, r)}$ . Ce minimum est aussi le minimum sur  $K$ . En effet, pour tout  $x \in K$ :

- (1) soit  $x \in K \cap \overline{B(0, r)}$ , et alors  $f(x) \geq f(x_*)$  car  $f$  atteint son minimum sur  $K \cap \overline{B(0, r)}$  en  $x_*$ ,
- (2) soit  $x \notin K \cap \overline{B(0, r)}$ , auquel cas on a  $f(x) > f(x_0) \geq f(x_*)$  où la seconde inégalité vient du fait que  $x_0$  est dans la boule.

Ceci montre donc que  $x_*$  est un minimum de  $f$  sur  $K$ .  $\square$

**Exercice 5.3.** Soit  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $f(x) := -x_1 x_2^2 + |x_1|^3 + e^{x_2} + |x_3|$ . Montrer que  $f$  admet un minimum sur  $\mathbb{R}^3$ .

*Solution.* On a

$$f(x) \geq -2(x_1^2 + x_2^4) + |x_1|^3 + e^{x_2} + |x_3| = x_1^2(|x_1| - 2) + (e^{x_2} - 2x_2^4) + |x_3|,$$

donc  $f$  est coercive, et on peut conclure en utilisant le théorème 5.4.  $\square$

En pratique, sur un fermé borné  $K$ , pour trouver les minimums, il faudra chercher en deux étapes :

- dans l'intérieur  $\overset{\circ}{K}$  qui est un ouvert. On cherchera donc dans ce cas les points qui annulent le gradient, on note  $Z := \{x \in \overset{\circ}{K} \mid \nabla f(x) = 0\}$ .
- sur le bord  $\partial K$ .

Autrement dit,

$$\min_{x \in K} f(x) = \min_{x \in Z \cap \partial K} f(x).$$

**5.3.  $K$  convexe.** Quand  $K$  est convexe, on a des conditions nécessaires et suffisantes utilisant le gradient.

**Théorème 5.5:  $K$  convexe**

Soit  $K$  un ensemble convexe et  $f : K \rightarrow \mathbb{R}$  une fonction  $\mathcal{C}^1$ . Si  $x_*$  est un minimum local de  $f$  sur  $K$ , alors

$$\forall y \in K, \quad \langle \nabla f(x_*), y - x_* \rangle \geq 0. \quad (20)$$

Si, en plus  $f$  est convexe, alors,

$$(20) \iff x_* \text{ est un minimum global de } f \text{ sur } K.$$

L'inéquation (20) est appelée inéquation d'Euler. Elle signifie que partant de  $x_*$ ,  $f$  croît dans toutes les directions qui restent dans  $K$ .

Avant de présenter ce résultat, il est important de voir que la condition (20) se réduit à l'équation d'Euler

$$\nabla f(x_*) = 0.$$

lorsque  $K$  est un ouvert (en particulier, lorsque  $K = \mathbb{R}^d$ ).

*Proof.*

- Prouvons la première partie. Soit  $y \in K$  et  $h \in ]0, 1]$ . Alors,  $x_* + h(y - x_*) \in K$  car  $K$  est convexe. De plus, comme  $x_*$  est un minimum local de  $K$ ,

$$\frac{f(x_* + h(y - x_*)) - f(x_*)}{h} \geq 0.$$

Or on a le développement de Taylor à l'ordre 1

$$\begin{aligned} f(x_* + h(y - x_*)) &= f(x_*) + (d_{x_*} f)(h(y - x_*)) + O(h^2) \\ &= f(x_*) + h \langle \nabla f(x_*), y - x_* \rangle + O(h^2), \end{aligned}$$

donc  $\langle \nabla f(x_*), y - x_* \rangle + O(h) \geq 0$ . Enfin, on fait  $h \rightarrow 0$ .

- Prouvons la seconde partie. Le sens  $\Leftarrow$  se fait avec la partie précédente. Montrons le sens  $\Rightarrow$ . Comme  $f$  est convexe sur  $K$ , pour tout  $y \in K$ ,

$$f(y) \geq f(x_*) + \langle \nabla f(x_*), y - x_* \rangle$$

grâce à la formule (ii) de la Proposition 4.4. Comme  $\langle \nabla f(x_*), y - x_* \rangle \geq 0$ , on a donc  $f(y) \geq f(x_*)$  pour tout  $y \in K$  ce qui prouve que  $x_*$  est un minimum global de  $f$  sur  $K$ .  $\square$

Dans le cas où  $K$  est convexe et  $f$  est non seulement convexe, mais strictement ou fortement convexe, nous avons les deux résultats suivants qui permettront de garantir existence, unicité et caractérisation du minimiseur.



**Théorème 5.6: Existence et unicité,  $K$  convexe**

Soit  $K \subset \mathbb{R}^d$  un ensemble convexe et  $f : K \rightarrow \mathbb{R}^d$  une fonction strictement convexe. S'il existe un minimiseur sur  $K$ , alors il est unique.  
Si en plus  $K$  est fermé et si  $f$  est  $\mathcal{C}^1$  et fortement convexe, alors il existe un unique minimiseur global.

*Proof.*

• Nous allons raisonner par l'absurde. Soient  $x_1$  et  $x_2 \in K$  avec  $x_1 \neq x_2$  deux points de minimum de  $f$  sur  $K$ . Nous avons donc  $f(x_1) = f(x_2) \leq f(x)$  pour tout  $x \in K$ . Comme  $f$  est strictement convexe et que  $(x_1 + x_2)/2 \in K$  car  $K$  est convexe, nous avons

$$f\left(\frac{x_1+x_2}{2}\right) < \frac{1}{2}f(x_1) + \frac{1}{2}f(x_2) = f(x_1)$$

Ceci contredit le fait que  $x_1$  soit un minimum.

• Comme  $f$  est fortement convexe et de classe  $\mathcal{C}^1$ , elle est continue et coercive. L'existence du minimum est garantie par le théorème 5.4.  $\square$

Par le théorème 5.5, ce minimiseur vérifie l'inéquation d'Euler (20).

**Exercice 5.4.** On définit  $f(x_1, x_2) := (x_1 + x_2)^2 + x_1x_2 - 4x_1 - 3x_2$ , montrer que  $f$  a un unique minimum sur  $K := \{x \in \mathbb{R}^2 \mid x_1 + x_2 \leq 3\}$ .

*Solution.*  $K$  et  $f$  sont convexes,  $K$  est fermé,  $f$  est une forme quadratique avec  $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$  qui a ses valeurs propres strictement positives donc  $f$  est fortement convexe donc on peut appliquer le théorème 5.6.  $\square$

## 6. THÉORÈME KKT

## REFERENCES

[1] J.-C. CULIOLI, *Introduction à l'optimisation*, Ellipses, 2012.

LABORATOIRE AGM - UMR 8088, CNRS - CY CERGY PARIS UNIVERSITÉ, 2 AVENUE ADOLPHE CHAUVIN, CERGY-PONTOISE, 95302 CEDEX, FRANCE

Email address: louis.garrigue@cyu.fr