DÉPARTEMENT STPI

2ÈME ANNÉE MIC

Optimisation Continue

SÉBASTIEN TORDEUX

2006/2007

Chapitre 1

Notion élémentaire de topologie dans $\mathbb{R}^n$

1.1 Ensembles ouverts et ensembles fermés

1.1.1 Les ensembles ouverts

Définition 1.1 *Soit* $\mathbf{x}$ *un élément de* $\mathbb{R}^n$ *et soit* r *un réel strictement positif.*

La boule ouverte $B(\mathbf{x}, r)$ *de centre* $\mathbf{x}$ *et de rayon* r *est définie par*

$$B(\mathbf{x}, r) \ = \ \left\{ \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{x}\| < r \right\}. \tag{1.1}$$

Définition 1.2 *Soit* $O \subset \mathbb{R}^n$. *Un ensemble* O *est ouvert ssi*

$$\forall \mathbf{y} \in O \ \exists r > 0 \mid B(\mathbf{y}, r) \subset O. \tag{1.2}$$

Exemples. Les ensembles $\mathbb{R}^n$ et $\emptyset$ sont ouverts.

Proposition 1.1 *Les boules ouvertes sont des ouverts.*

Preuve. Soient $B(\mathbf{x}, R)$ cette boule et $\mathbf{y} \in B(\mathbf{x}, R)$. Soit r le réel strictement positif donné par $r = R - \|\mathbf{x} - \mathbf{y}\|$. Pour tout $\mathbf{z} \in B(\mathbf{y}, r)$,

$$\begin{aligned} \|\mathbf{x} - \mathbf{z}\| \ &\leq \ \|\mathbf{x} - \mathbf{y}\| \ + \ \|\mathbf{y} - \mathbf{z}\| \\ &< \ \|\mathbf{x} - \mathbf{y}\| \ + \ R \ - \ \|\mathbf{x} - \mathbf{y}\| \\ &< \ R. \end{aligned} \tag{1.3}$$

Ainsi, $\mathbf{z} \in B(\mathbf{x}, R)$.

Proposition 1.2 *Si* O_1 *et* O_2 *sont deux ouverts, alors* $O_1 \cap O_2$ *est ouvert.*

Preuve. Soit $\mathbf{x} \in O_1 \cap O_2$. Comme O_1 et O_2 sont ouverts, il existe $r_1 > 0$ et $r_2 > 0$ tel que

$$B(\mathbf{x}, r_1) \subset O_1 \quad \text{et} \quad B(\mathbf{x}, r_2) \subset O_2. \tag{1.4}$$

Désignons par r le minimum de r_1 et r_2. Ainsi, nous avons

$$B(\mathbf{x}, r) \subset O_1 \quad \text{et} \quad B(\mathbf{x}, r) \subset O_2 \tag{1.5}$$

et par conséquent

$$B(\mathbf{x}, r) \subset O_1 \cap O_2. \tag{1.6}$$

Ceci nous permet d'affirmer que $O_1 \cap O_2$ est ouvert.

Proposition 1.3 *Si $\{O_i\}$ désigne une famille d'ouverts, alors $\underset{i}{\cup} O_i$ est ouvert.*

Preuve. Soit $\mathbf{x} \in \underset{i}{\cup} O_i$. Il existe i tel que $\mathbf{x} \in O_i$. Comme O_i est ouvert, il existe $r > 0$ tel que

$$B(\mathbf{x}, r) \subset O_i \subset \underset{i}{\cup} O_i. \tag{1.7}$$

Ainsi, $\underset{i}{\cup} O_i$ est ouvert.

1.1.2 Les ensembles fermés

Définition 1.3 *Soit $\mathbf{x} \in \mathbb{R}^n$ et soit $r > 0$.*

La boule fermée de centre $\mathbf{x}$ et de rayon r est définie par

$$\overline{B}(\mathbf{x}, r) = \Big\{ \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{x}\| \leq r \Big\}. \tag{1.8}$$

Définition 1.4 *Soit $F \subset \mathbb{R}^n$. L'ensemble F est un fermé si son complémentaire est un ouvert.*

Exemples. L'ensemble $\mathbb{R}^n$, respectivement $\emptyset$, est fermé. En effet, son complémentaire est $\emptyset$, respectivement $\mathbb{R}^n$.

Proposition 1.4 *Soit $F \subset \mathbb{R}^n$ un fermé. Toute suite convergente d'éléments de F admet une limite dans F.*

Preuve. Considérons une suite $\mathbf{x}_n \in F$ qui converge vers $\mathbf{x} \in \mathbb{R}^n$. Soit O le complémentaire de F. Soit $\mathbf{y}$ un élément quelconque de O.

Comme O est ouvert, il existe $r > 0$ tel que $B(\mathbf{y}, r) \subset O$. Ainsi, pour tout $\mathbf{z} \in F$, $\mathbf{z} \notin B(\mathbf{y}, r)$ et par conséquent il suit

$$\|\mathbf{z} - \mathbf{y}\| \geq r. \tag{1.9}$$

Par inégalité triangulaire, nous avons

$$\|\mathbf{y}-\mathbf{x}\| \geq \|\mathbf{y}-\mathbf{x}_n\| - \|\mathbf{x}_n-\mathbf{x}\|. \tag{1.10}$$

Comme $\mathbf{x}_n$ est un élément de F, nous avons

$$\|\mathbf{y}-\mathbf{x}\| \geq r - \|\mathbf{x}_n-\mathbf{x}\|. \tag{1.11}$$

Ceci nous fournit pour n tendant vers 0

$$\|\mathbf{y}-\mathbf{x}\| \geq r > 0 \tag{1.12}$$

et par conséquent $\mathbf{y} \neq \mathbf{x}$. Comme ceci est vrai pour tout $\mathbf{y}$ dans O, $\mathbf{x} \notin O$, c'est-à-dire $\mathbf{x} \in F$.

Proposition 1.5 *Soit $F \subset \mathbb{R}^p$. Si toute suite convergente d'éléments de F admet une limite dans F alors F est fermé.*

Preuve. Supposons que toute suite convergente de F admet une limite dans F. Montrons que O, le complémentaire de F, est ouvert.

Raisonnons par l'absurde. Supposons que O ne soit pas ouvert

$$\forall r > 0 \quad B(\mathbf{y}, r) \not\subset O, \tag{1.13}$$

dit autrement

$$\forall r > 0 \quad B(\mathbf{y}, r) \cap F \neq \emptyset. \tag{1.14}$$

Pour $r = \frac{1}{n+1}$, nous construisons une suite $(\mathbf{x}_n)$ d'éléments de F telle que

$$\mathbf{x}_n \in B(\mathbf{y}, \frac{1}{n+1}). \tag{1.15}$$

Ceci peut être traduit en

$$\|\mathbf{x}_n - \mathbf{y}\| \leq \frac{1}{n+1}. \tag{1.16}$$

Il suit que $\mathbf{x}_n$ converge vers $\mathbf{y}$. Comme F est fermé, $\mathbf{y} \in F$. Ceci est impossible car $\mathbf{y} \in O$.

Proposition 1.6 *Si F_1 et F_2 sont deux fermés, alors $F_1 \cup F_2$ est fermé.*

Preuve. Soient O_1 et O_2 les complémentaires de F_1 et F_2. Comme O_1 et O_2 sont ouverts, $O_1 \cap O_2$ est ouvert. Le complémentaire $F_1 \cup F_2$ de $O_1 \cap O_2$ est donc fermé.

Proposition 1.7 *Si $\{F_i\}$ désigne une famille de fermés, alors $\underset{i}{\cap} F_i$ est fermé.*

Preuve. Soient O_i les complémentaires des F_i. Comme les O_i sont ouverts, $\underset{i}{\cup} O_i$ est ouvert. Le complémentaire $\underset{i}{\cap} F_i$ est fermé.

1.2 Propriétés des fonctions continues

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application.

Définition 1.5 *f est continue ssi*

$$\lim_{\mathbf{x}_n \longrightarrow \mathbf{x}} f(\mathbf{x}_n) \;=\; f(\mathbf{x}). \tag{1.17}$$

Définition 1.6 *Soit $E \subset \mathbb{R}^p$. L'image réciproque de E par f est définie par*

$$\overset{-1}{f}(E) \;=\; \Big\{ \mathbf{x} \in \mathbb{R}^n \;\Big|\; f(\mathbf{x}) \in E \Big\}. \tag{1.18}$$

Remarque 1.1 *Remarquons que $\overset{-1}{f}(\mathbb{R}^p) \;=\; \mathbb{R}^n$ et que pour tout A et tout B ensembles inclus dans $\mathbb{R}^p$*

$$\overset{-1}{f}(A \cup B) \;=\; \overset{-1}{f}(A) \cup \overset{-1}{f}(B) \quad \textit{et} \quad \overset{-1}{f}(A \cap B) \;=\; \overset{-1}{f}(A) \cap \overset{-1}{f}(B). \tag{1.19}$$

Dans la suite de ce paragraphe, (H) est une abréviation pour : Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application continue.

Proposition 1.8 *(H). Soit $O \subset \mathbb{R}^p$ un ouvert. L'image réciproque de O, $\overset{-1}{f}(O)$, est un ouvert.*

Preuve. Soit $\mathbf{x} \in \overset{-1}{f}(O)$ et $\mathbf{y} = f(\mathbf{x})$. Par définition de l'image réciproque, $\mathbf{y} \in O$. Comme O est ouvert, il existe $\varepsilon > 0$ tel que $B(\mathbf{y}, \varepsilon) \subset O$.

D'autre part, comme f est continue, on peut écrire sa continuité au point $\mathbf{x}$: il existe $\eta > 0$ tel que

$$\forall \mathbf{z} \in \mathbb{R}^n \quad \|\mathbf{z} - \mathbf{x}\| \;<\; \eta \quad \Longrightarrow \quad \|f(\mathbf{z}) - \mathbf{y}\| \;<\; \varepsilon. \tag{1.20}$$

Ainsi pour $\mathbf{z} \in B(\mathbf{x}, \eta)$, nous avons $f(\mathbf{z}) \in B(f(\mathbf{x}), \varepsilon) \subset O$; c'est-à-dire

$$B(\mathbf{x}, \eta) \subset \overset{-1}{f}(O). \tag{1.21}$$

$\overset{-1}{f}(O)$ est donc ouvert.

Proposition 1.9 *(H). Soit $F \subset \mathbb{R}^p$ un fermé. L'image réciproque de F, $\overset{-1}{f}(F)$, est un fermé.*

Preuve. Notons O le complémentaire de F. Comme F est fermé, O est ouvert. Notons que

$$\left\{\begin{array}{l} \overset{-1}{f}(F) \cup \overset{-1}{f}(O) \;=\; \overset{-1}{f}(F \cup O) \;=\; \overset{-1}{f}(\mathbb{R}^p) \;=\; \mathbb{R}^n, \\ \overset{-1}{f}(F) \cap \overset{-1}{f}(O) \;=\; \overset{-1}{f}(F \cap O) \;=\; \overset{-1}{f}(\emptyset) \;=\; \emptyset. \end{array}\right. \tag{1.22}$$

Ainsi, $\overset{-1}{f}(F)$ est le complémentaire de $\overset{-1}{f}(O)$ dans $\mathbb{R}^n$. D'après la proposition 1.8, $\overset{-1}{f}(O)$ est ouvert comme O est ouvert. Par conséquent, $\overset{-1}{f}(F)$ est fermé.

Remarque 1.2 *(H). Soit $F \subset \mathbb{R}^n$ un fermé. L'image directe de F*

$$f(F) \;=\; \Big\{\mathbf{y} = f(\mathbf{x}) \mid \mathbf{x} \in F\Big\} \tag{1.23}$$

n'est pas toujours fermée. Par exemple, pour $f : \mathbb{R} \longrightarrow \mathbb{R}$ tel que $f(x) = e^x$, nous avons

$$f(\mathbb{R}) \;=\; \mathbb{R}_+^* \;=\; \Big\{x > 0\Big\} \tag{1.24}$$

qui n'est pas un ensemble fermé.

Remarque 1.3 *(H). Soit O un ouvert, l'image directe de O, $f(O)$, n'est pas forcément ouvert. Par exemple, pour $f : \mathbb{R} \longrightarrow \mathbb{R}$ tel que $f(x) = 0$, nous avons*

$$f(\mathbb{R}) = \{0\} \tag{1.25}$$

qui n'est pas un ensemble ouvert.

1.3 Les ensembles compacts

Définition 1.7 *Soit $K \subset \mathbb{R}^p$. K est compact ssi toute suite $(\mathbf{x}_n)$ d'éléments de K admet une suite extraite convergente dont la limite est un élément de K.*

Proposition 1.10 *Soient a et b deux éléments de $\mathbb{R}$. L'intervalle fermé $[a, b]$ est compact.*

Preuve. Soit (x_n) une suite d'éléments de $[a, b]$. La suite (x_n) est bornée. D'après le théorème de Bolzano-Weierstrass, la suite (x_n) admet une sous-suite extraite convergente. Notons x sa limite. x vérifie

$$x_n \to x \text{ et } a \leqslant x_n \leq b \quad \Longrightarrow \quad a \leq x \leq b \tag{1.26}$$

Proposition 1.11 *Soient a et b deux éléments de $\mathbb{R}$. Le pavé $[a,b]^p \subset \mathbb{R}^p$ est compact.*

Preuve. Soit $(\mathbf{x}_n)$ une suite d'éléments de $[a,b]^p \subset \mathbb{R}^p$

$$\mathbf{x}_n = (\mathbf{x}_{n,1}, \mathbf{x}_{n,2}, \cdots, \mathbf{x}_{n,p}). \tag{1.27}$$

La suite $\mathbf{x}_{n,1}$ est une suite d'éléments de $[a,b]$. On peut donc en extraire une sous-suite extraite convergente

$$\mathbf{x}_{\sigma_1(n),1} \longrightarrow \mathbf{l}_1 \in [a,b]. \tag{1.28}$$

Considérons maintenant $\mathbf{x}_{\sigma_1(n)} = (\mathbf{x}_{\sigma_1(n),1}, \mathbf{x}_{\sigma_1(n),2}, \cdots, \mathbf{x}_{\sigma_1(n),p})$. La suite $(\mathbf{x}_{\sigma_1(n),2})$ est une suite d'éléments de $[a,b]$. Il existe une sous-suite extraite $(\mathbf{x}_{\sigma_2(\sigma_1(n)),2})$ convergente

$$\mathbf{x}_{\sigma_2(\sigma_1(n)),2} \longrightarrow \mathbf{l}_2 \in [a,b], \tag{1.29}$$

d'autre part

$$\mathbf{x}_{\sigma_2(\sigma_1(n)),1} \longrightarrow \mathbf{l}_1 \in [a,b]. \tag{1.30}$$

Par récurrence, nous construisons une suite $\mathbf{x}_{\sigma_p(\cdots(\sigma_2(\sigma_1(n))))}$ telle que

$$\left\{ \begin{array}{c} \mathbf{x}_{\sigma_p(\cdots(\sigma_2(\sigma_1(n)))),1} \longrightarrow \mathbf{l}_1 \in [a,b], \\ \mathbf{x}_{\sigma_p(\cdots(\sigma_2(\sigma_1(n)))),2} \longrightarrow \mathbf{l}_2 \in [a,b] \\ \cdots \\ \mathbf{x}_{\sigma_p(\cdots(\sigma_2(\sigma_1(n)))),p} \longrightarrow \mathbf{l}_p \in [a,b]. \end{array} \right. \tag{1.31}$$

Cette sous-suite extraite converge vers $\mathbf{l} = (\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_p) \in [a,b]^p$.

Définition 1.8 *(Ensemble borné) Soit $\mathcal{B} \subset \mathbb{R}^n$. L'ensemble $\mathcal{B}$ est borné ssi*

$$\exists A > 0 \mid \forall \boldsymbol{x} \in \mathcal{B} \quad \|\boldsymbol{x}\| < A. \tag{1.32}$$

Dit autrement, il existe une boule $B(0,A)$ qui contient $\mathcal{B}$.

Théorème 1.1 *Soit $K \subset \mathbb{R}^p$. L'ensemble K est compact ssi K est fermé et borné.*

Preuve. ($\Longleftarrow$) Soit K un fermé borné. Soit $(\mathbf{x}_n)$ une suite d'éléments de K.

Il existe un pavé fermé $\overline{P} = [a,b]^p$ qui contient K. Les pavés fermés étant compacts, la suite $\mathbf{x}_n$ admet une sous-suite extraite convergente dont la limite est dans $\overline{P}$. Comme K est fermé et $\mathbf{x}_n \in K$, cette limite est un élément de K.

($\Longrightarrow$) Soit K un compact.

Montrons que K est fermé. Soit $(\mathbf{x}_n)$ une suite convergente

$$\mathbf{x}_n \longrightarrow \mathbf{l}. \tag{1.33}$$

Nous allons montrer que $\mathbf{l} \in K$. Comme K est compact, elle admet une sous-suite extraite $\mathbf{x}_{\sigma(n)}$ qui converge dont la limite est dans K

$$\mathbf{x}_{\sigma(n)} \longrightarrow \mathbf{l}_\sigma \in K; \tag{1.34}$$

d'autre part

$$\mathbf{x}_{\sigma(n)} \longrightarrow \mathbf{l}. \tag{1.35}$$

Par unicité de la limite, $\mathbf{l} = \mathbf{l}_\sigma \in K$. K est donc fermé.

Montrons que K est borné. Raisonnons par l'absurde. Supposons que K est non borné. Nous pouvons construire une suite $(\mathbf{x}_n)$ d'éléments de K telle que $\|\mathbf{x}_n\|$ tend vers $+\infty$. Toute suite extraite $\mathbf{x}_{\sigma(n)}$ vérifie aussi $\|\mathbf{x}_{\sigma(n)}\| \to +\infty$ et est donc divergente. Ceci est impossible car K est compact.

1.4 Exercices

1.4.1 Ensembles ouverts et ensembles fermés

Exercice. Montrer que le demi-espace ouvert est ouvert

$$\mathbb{R}^p_+ \; = \; \Big\{ (\mathbf{x}_1, \cdots, \mathbf{x}_p) \in \mathbb{R}^p \;\Big|\; \mathbf{x}_1 > 0 \Big\}. \tag{1.36}$$

Exercice. Soient a et b deux réels avec $a < b$. Les intervalles sont-ils ouverts, fermés ?

$$[a,b], \quad [a,b[, \quad]a,b], \quad]a,b[, \quad]-\infty,b], \quad]-\infty,b[, \quad]a,+\infty[, \quad [a,+\infty[. \tag{1.37}$$

Exercice. Montrer que le demi-espace fermé est fermé

$$\overline{\mathbb{R}^p_+} \; = \; \Big\{ (\mathbf{x}_1, \cdots, \mathbf{x}_p) \in \mathbb{R}^p \;\Big|\; \mathbf{x}_1 \geq 0 \Big\}. \tag{1.38}$$

Exercice. Soient n et p deux entiers avec $0 < p < n$. Montrer que $\mathbb{R}^p$ n'est pas ouvert en tant qu'ensemble de $\mathbb{R}^n$. Montrer que $\mathbb{R}^p$ est fermé en tant qu'ensemble de $\mathbb{R}^n$. On entend par sous-ensemble de $\mathbb{R}^n$

$$\mathbb{R}^p \; = \; \Big\{ \mathbf{x} \in \mathbb{R}^n \quad \Big| \quad \mathbf{x}_{p+1} = \cdots = \mathbf{x}_n = 0 \Big\}. \tag{1.39}$$

1.4.2 Fonctions continues

Exercice. Montrer que l'application norme est continue.

Exercice. Montrer que toute boule fermée est fermée.

Exercice. Montrer que l'ensemble suivant est fermé

$$\left\{\mathbf{x} \in \mathbb{R}^3 \;\middle|\; \mathbf{x}_1 + \mathbf{x}_2^2 + \mathbf{x}_3^2 = 1\right\}. \tag{1.40}$$

Exercice. Montrer que l'ensemble suivant est ouvert

$$\left\{\mathbf{x} \in \mathbb{R}^2 \;\middle|\; -3 < 4\mathbf{x}_1^2 + 3\mathbf{x}_2^2 + 5\mathbf{x}_3^2 < 2\right\}. \tag{1.41}$$

1.4.3 Ensembles compacts

Exercice. $\mathbb{R}^n$ est-il compact ?

Exercice. Montrer que les boules fermées sont compactes.

Exercice. Soient a et b deux réels avec $a < b$. Les intervalles sont-ils compacts ?

$$[a,b], \quad [a,b[, \quad]a,b], \quad]a,b[, \quad]-\infty,b], \quad]-\infty,b[, \quad]a,+\infty[, \quad [a,+\infty[. \tag{1.42}$$

Exercice. Montrer que l'ensemble suivant est compact

$$\left\{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \in \mathbb{R}^3 \;\middle|\; 2\mathbf{x}_1^2 + \mathbf{x}_2^2 + 3\mathbf{x}_3^2 \leq 5\right\}. \tag{1.43}$$

Exercice. Soit $(\mathbf{x}_n)$ une suite d'éléments de $\mathbb{R}^p$ qui converge vers $\mathbf{x} \in \mathbb{R}^p$. Montrer que l'ensemble suivant est compact

$$\{\mathbf{x}_n \mid n \in \mathbb{N}\} \cup \{\mathbf{x}\}. \tag{1.44}$$

Exercice. Soit K un compact. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application continue. Montrer que $f(K)$ est fermé.

Exercice. Soit K un compact. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application continue. Montrer que $f(K)$ est compact.

Chapitre 2

Deux théorèmes d'existence de minimum et maximum d'une fonction

Nous nous intéressons ici à la recherche de $\mathbf{x}$ qui réalise le minimum d'une application continue $f : E \subset \mathbb{R}^p \longrightarrow \mathbb{R}$, c'est-à-dire nous cherchons

$$\mathbf{x} \in K \text{ tel que } f(\mathbf{x}) \leq f(\mathbf{y}) \quad \forall \mathbf{y} \in K. \tag{2.1}$$

Bien entendu la recherche d'un maximum peut se ramener à la recherche d'un minimum en considérant $-f$ à la place de f. C'est pourquoi nous porterons une attention plus particulière à la recherche du minimum.

2.1 Minimum, maximum, supremum, infimum

Définition 2.1 *(minorant, majorant) Soit* $E \subset \mathbb{R}$ *un ensemble.*

$m \in \mathbb{R} \cup \{-\infty, +\infty\}$ *est un minorant de* E *ssi*

$$\forall x \in E \quad m \leq x. \tag{2.2}$$

$M \in \mathbb{R} \cup \{-\infty, +\infty\}$ *est un majorant de* E *ssi*

$$\forall x \in E \quad M \geq x. \tag{2.3}$$

Définition 2.2 *(infimum, supremum) Soit* $E \subset \mathbb{R}$.

L'infimum $\inf(E) \in \mathbb{R} \cup \{-\infty, +\infty\}$ *de* E *est le plus grand des minorants.*

Le supremum $\sup(E) \in \mathbb{R} \cup \{-\infty, +\infty\}$ *de* E *est le plus petit des majorants. On note aussi*

$$\inf_{x \in E}(x) \quad \textit{et} \quad \sup_{x \in E}(x). \tag{2.4}$$

Définition-Proposition 2.1 *(suite minimisante, suite maximisante) Soit* $E \subset \mathbb{R}$, $E \neq \emptyset$.

Il existe (x_n) *une suite d'éléments de* E *telle que*

$$x_n \longrightarrow \inf(E) \tag{2.5}$$

et une suite (y_n) *d'éléments de* E *telle que*

$$y_n \longrightarrow \sup(E). \tag{2.6}$$

La suite (x_n) *est une suite minimisante et la suite* (y_n) *est une suite maximisante.*

Preuve. Pour $E \neq \emptyset$, $\inf(E) \in \mathbb{R} \cup \{-\infty\}$ et $\sup(E) \in \mathbb{R} \cup \{+\infty\}$.

(i) Pour $\inf(E) \in \mathbb{R}$. Comme $\inf(E)$ est le plus grand des minorants de E, $\inf(E) + 1/n$ pour tout $n \in \mathbb{N}^*$ n'est pas un minorant de E. Il existe un élément $x_n \in E$ tel que

$$x_n \leq \inf(E) + 1/n. \tag{2.7}$$

Cette suite (x_n) d'éléments de E vérifie $\inf(E) \leq x_n \leq \inf(E) + 1/n$ et admet donc $\inf(E)$ comme limite.

(ii) Pour $\inf(E) = -\infty$, E admet seulement $-\infty$ comme minorant. Par conséquent pour tout $n \in \mathbb{N}$, il existe $x_n \in E$ tel que

$$x_n \leq -n. \tag{2.8}$$

La suite x_n ainsi construite converge vers $-\infty$.

La construction de la suite (y_n) est laissée en exercice.

Définition 2.3 *(minimum, maximum) Soit* $E \subset \mathbb{R}$.

Un infimum est un minimum si $\inf(E) \in E$.
Un supremum est un maximum si $\sup(E) \in E$.
Dans ce cas, on note

$$\min(E) = \inf(E) \quad \textit{et} \quad \max(E) = \sup(E). \tag{2.9}$$

Proposition 2.1 *Soit* $F \subset \mathbb{R}$ *(*$F \neq \emptyset$*) un fermé.*

Si F *est borné inférieurement alors* F *admet un minimum.*

Si F *est borné supérieurement alors* F *admet un maximum.*

Preuve. Soit (x_n) une suite minimisante

$$\lim_{n \to +\infty} x_n = \inf(F). \tag{2.10}$$

Comme F est fermé, $\inf(F) \in F$ et, par conséquent, $\inf(F)$ est le minimum de F. Le résultat sur le maximum est laissé en exercice.

2.2 Existence de minimum sur les ensembles bornés

Définition 2.4 *Soit $f : E \subset \mathbb{R}^p \longrightarrow \mathbb{R}$ une application et $K \subset E$. $\mathbf{x}$ est un argument-minimum de f sur K si*

$$\mathbf{x} \in K \quad \textit{et} \quad f(\mathbf{x}) \leq f(\mathbf{y}) \quad \forall \mathbf{y} \in K. \tag{2.11}$$

On dit aussi que $\mathbf{x}$ réalise le minimum de f sur K.

S'il n'existe qu'un argument-minimum, nous adoptons la notation suivante

$$\mathbf{x} \;=\; \underset{\mathbf{y}\in K}{\textit{argmin}}(f(\mathbf{y})). \tag{2.12}$$

Proposition 2.2 *Pour tout $\mathbf{x}$ argument-minimum de f sur K. $\mathbf{x}$ réalise le minimum de f*

$$f(\mathbf{x}) = \inf_{\mathbf{y}\in K}(f(\mathbf{y})). \tag{2.13}$$

Preuve. Comme $\mathbf{x} \in K$, il suit

$$f(\mathbf{x}) \geq \inf_{\mathbf{y}\in K}(f(\mathbf{y})). \tag{2.14}$$

D'autre part comme $f(\mathbf{x})$ est un minorant de $f(\mathbf{y})$ pour tout $\mathbf{y} \in K$

$$f(\mathbf{x}) \leq \inf_{\mathbf{y}\in K}(f(\mathbf{y})). \tag{2.15}$$

Ceci prouve l'égalité souhaitée.

Remarque 2.1 *Attention, il n'existe pas toujours d'argument-minimum. En effet pour $f : \mathbb{R} \longrightarrow \mathbb{R}$ avec $f(x) = \frac{1}{x^2+1}$. $\inf(f(\mathbf{y}))$ est zéro qui n'est atteint en aucun x.*

Attention, il n'y a pas toujours unicité de l'argument-minimum de f sur K. Par exemple, pour $f(\mathbf{x}) = 0$ tout $\mathbf{x} \in E$ est argument-minimum de f.

Tout $\mathbf{x}$ argument-minimum de f est caractérisé par

$$\mathbf{x} \;\in\; \overset{-1}{f}(\inf_{\mathbf{y}\in E}(f(\mathbf{y}))). \tag{2.16}$$

Théorème 2.1 *Soit $K \subset R^p$ un compact (un fermé borné), $K \neq \emptyset$. Soit $f : K \subset \mathbb{R}^p \longrightarrow \mathbb{R}$ une application continue.*

Il existe $\mathbf{x} \in K$ argument-minimum de f sur K. Dit autrement, il existe $\mathbf{x} \in K$ tel que

$$f(\mathbf{x}) \leq f(\mathbf{y}), \quad \forall \mathbf{y} \in K. \tag{2.17}$$

Preuve. Introduisons l'image directe de K par f

$$f(K) \;=\; \Big\{ f(\mathbf{y}) \;\Big|\; \mathbf{y} \in K \Big\}. \tag{2.18}$$

Considérons une suite minimisante dans $f(K)$. C'est-à-dire une suite de $(\mathbf{x}_n)$ d'éléments de K telle que

$$f(\mathbf{x}_n) \longrightarrow \inf_{\mathbf{y}\in K} f(\mathbf{y}). \tag{2.19}$$

Comme K est compact, il existe une sous-suite extraite $(\mathbf{x}_{\sigma(n)})$ qui converge vers $\mathbf{x} \in K$. Cette suite suite extraite vérifie

$$\mathbf{x}_{\sigma(n)} \longrightarrow \mathbf{x} \quad \text{et} \quad f(\mathbf{x}_{\sigma(n)}) \longrightarrow \inf_{\mathbf{y}\in K} f(\mathbf{y}). \tag{2.20}$$

Comme f est continue et par unicité de la limite, il suit

$$f(\mathbf{x}) \;=\; \inf_{\mathbf{y}\in K} f(\mathbf{y}) \text{ avec } \mathbf{x} \in K. \tag{2.21}$$

Ainsi, f réalise son minimum sur K.

2.3 Existence de minimum sur les ensembles non bornés

Définition 2.5 *Une application $f : F \subset \mathbb{R}^p \longrightarrow \mathbb{R}$ est infinie à l'infini ssi*

$$\forall A \in \mathbb{R}, \quad \exists R > 0 \mid \forall \mathbf{x} \in F \quad \|\mathbf{x}\| \geq R \Longrightarrow f(\mathbf{x}) \geq A \tag{2.22}$$

On note

$$\lim_{\|\mathbf{x}\| \longrightarrow +\infty} f(\mathbf{x}) = +\infty. \tag{2.23}$$

Pour montrer que f est infinie à l'infini on utilise souvent la

Proposition 2.3 *Soit $f : F \subset \mathbb{R}^p \longrightarrow \mathbb{R}$ une application et $g : \mathbb{R} \longrightarrow \mathbb{R}$ vérifiant*

$$f(\mathbf{x}) \geq g(\|\mathbf{x}\|) \quad \textit{avec} \quad \lim_{t\to+\infty} g(t) = +\infty. \tag{2.24}$$

Alors, f est infinie à l'infini.

Preuve. Comme g tend vers $+\infty$ en $+\infty$

$$\forall A \in \mathbb{R}, \quad \exists R > 0 \mid \forall t \in \mathbb{R} \quad t \geq R \Longrightarrow g(t) \geq A. \tag{2.25}$$

Avec $t = \|\mathbf{x}\|$ et comme $g(\mathbf{x}) \geq f(\|\mathbf{x}\|)$, nous obtenons (2.22).

Théorème 2.2 *Soit $F \subset \mathbb{R}^p$ un fermé, $F \neq \emptyset$. Soit $f : F \subset \mathbb{R}^p \longrightarrow \mathbb{R}$ une application continue infinie à l'infini.*

Il existe $\mathbf{x} \in F$ qui réalise le minimum de f sur F. Dit autrement, il existe $\mathbf{x} \in F$ tel que

$$f(\mathbf{x}) \leq f(\mathbf{y}), \quad \forall \mathbf{y} \in F. \tag{2.26}$$

Preuve. (i) Définissons K compact. Comme $F \neq \emptyset$, nous avons $\inf(F) \in \mathbb{R} \cup \{-\infty\}$. Il existe $A \in \mathbb{R}$, tel que $A > \inf_{\mathbf{y} \in E}(f(\mathbf{y}))$. Comme f est infinie à l'infini, il existe $R_1 > 0$ tel que pour $\mathbf{y} \in \mathbb{R}^p$

$$\|\mathbf{y}\| > R \quad \Longrightarrow \quad f(\mathbf{y}) > A. \tag{2.27}$$

Comme F est non vide, il existe $R_2 > 0$ tel que

$$\overline{B}(0, R_2) \cap F \neq \emptyset. \tag{2.28}$$

Choisissons $R = \max(R_1, R_2)$

$$\|\mathbf{y}\| > R \quad \Longrightarrow \quad f(\mathbf{y}) > A \quad \text{et} \quad \overline{B}(0, R) \cap F \neq \emptyset. \tag{2.29}$$

Soit $K = \overline{B}(0, R) \cap F$. L'ensemble K est un compact non vide car il est borné ($\|\mathbf{y}\| \leq R$) et fermé (intersection de deux fermés).

(ii) Minimisons f sur K. Comme f est continue et K est compact, f atteint son minimum sur K, c'est-à-dire

$$\exists \mathbf{x} \in K \quad | \quad f(\mathbf{x}) = \inf_{\mathbf{y} \in K}(f(\mathbf{y})), \tag{2.30}$$

(iii) Montrons que minimiser sur K revient à minimiser sur F. D'une part, nous avons

$$\inf_{\mathbf{y} \in F} f(\mathbf{y}) \ = \ \inf\left(\inf_{\mathbf{y} \in K} f(\mathbf{y}); \inf_{\mathbf{y} \in F \setminus K} f(\mathbf{y})\right). \tag{2.31}$$

D'autre part, comme pour $\mathbf{z} \notin K$ $\|\mathbf{z}\| \geq R$, nous avons

$$f(\mathbf{z}) > A > \inf_{\mathbf{y} \in F} f(\mathbf{y}) \tag{2.32}$$

et par conséquent

$$\inf_{\mathbf{y} \in F} f(\mathbf{y}) \ < \ \inf_{\mathbf{y} \in F \setminus K} f(\mathbf{y}). \tag{2.33}$$

D'où, il suit

$$\inf_{\mathbf{y} \in F} f(\mathbf{y}) \ = \ \inf_{\mathbf{y} \in K} f(\mathbf{y}) \tag{2.34}$$

et d'après (2.30) il vient

$$\exists \mathbf{x} \in K \subset F \quad | \quad f(\mathbf{x}) = \inf_{\mathbf{y} \in F}(f(\mathbf{y})). \tag{2.35}$$

2.4 Exercices

Exercice. Soit $R > 0$. Soit $\mathbf{b} \in \mathbb{R}^n$. Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique définie positive. Considérons le problème de minimisation de la fonctionnelle

$$J : \mathbb{R}^n \longrightarrow \mathbb{R} \quad \text{avec} \quad J(\mathbf{x}) \ = \ \mathbf{b}^T \, \mathbf{x} \tag{2.36}$$

sur l'ensemble $K \subset \mathbb{R}^n$

$$K \ = \ \Big\{ \mathbf{x} \in \mathbb{R}^n \ \Big| \ \frac{1}{2}\mathbf{x}^T A\mathbf{x} \leq R \Big\}. \tag{2.37}$$

1) Montrer qu'il existe $\lambda_{min} > 0$ et $\lambda_{max} > 0$ tels que

$$\lambda_{min} \, \|\mathbf{x}\|^2 \ \leq \ \mathbf{x}^T A\mathbf{x} \ \leq \ \lambda_{max} \, \|\mathbf{x}\|^2. \tag{2.38}$$

2) Montrer que K est fermé.

3) Montrer que K est borné.

4) Montrer que l'application J est continue.

5) Montrer que l'application J atteint son minimum sur K.

Exercice. Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique définie positive. Soit $\mathbf{b} \in \mathbb{R}^n$ un vecteur. Soit $c \in \mathbb{R}$.
Considérons l'application $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par

$$J(\mathbf{x}) \ = \ \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x} + c \quad \forall \mathbf{x} \in \mathbb{R}^n. \tag{2.39}$$

1) Montrer qu'il existe $\lambda_{min} > 0$ et $\lambda_{max} > 0$ tels que

$$\lambda_{min} \, \|\mathbf{x}\|^2 \ \leq \ \mathbf{x}^T A\mathbf{x} \ \leq \ \lambda_{max} \, \|\mathbf{x}\|^2. \tag{2.40}$$

2) Montrer que

$$|J(\mathbf{x}+\mathbf{h}) - J(\mathbf{x})| \ \leq \ \frac{1}{2}\lambda_{max}\|\mathbf{h}\|^2 + \|A\mathbf{x}\|\|\mathbf{h}\| + \|\mathbf{b}\|\|\mathbf{h}\|. \tag{2.41}$$

3) Montrer que J est continue.

4) Montrer la minoration suivante

$$J(\mathbf{x}) \ \geq \ \frac{1}{2}\lambda_{min}\|\mathbf{x}\|^2 - \|\mathbf{b}\|\|\mathbf{x}\| + c \quad \forall \mathbf{x} \in \mathbb{R}^n. \tag{2.42}$$

5) Montrer que J est infinie à l'infini.

6) Montrer que le minimum de J est atteint sur $\mathbb{R}^n$.

Exercice. Considérons le problème de maximisation de J

$$J : \mathbb{R}^2 \longrightarrow \mathbb{R} \quad \text{et} \quad J(x_1, x_2) = ch(x_1^2 + x_2) \tag{2.43}$$

sur

$$E \;=\; \Big\{ (x_1, x_2) \in \mathbb{R}^2 \;\Big|\; x_1^4 + x_2^2 = 1 \Big\}. \tag{2.44}$$

Montrer que J admet un argument-minimum sur E.

Exercice. Considérons le problème de minimisation dans $\mathbb{R}^3$ de

$$J : \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad J(\mathbf{x}) \;=\; \exp\left(\mathbf{x}_1^2 + \mathbf{x}_2^2 + \mathbf{x}_3^2\right) \;+\; \mathbf{x}_1. \tag{2.45}$$

Montrer que J atteint son minimum sur $\mathbb{R}^3$.

2.5 Problème : le théorème de d'Alembert

Soit P un polynôme de la variable complexe $z = x + iy$ non constant

$$P(z) = \sum_{n=0}^{N} a_n \, z^n, \quad \text{avec } a_n \in \mathbb{C} \text{ et } a_N \neq 0. \tag{2.46}$$

Le but de ce problème est de montrer que P est scindé dans $\mathbb{C}$

$$P(z) = a_N \prod_{n=1}^{N} (z - z_n) \quad \text{avec } z_n \in \mathbb{C}. \tag{2.47}$$

Les z_n sont les racines complexes du polynôme.

1. Expliquer pourquoi on peut identifier **C** à $\mathbb{R}^2$.

2. On note $\mathbf{z} = (x, y)$. Montrer que $\|\mathbf{z}\| = |z|$.

3. Montrer l'inégalité suivante

$$|P(z)| \geq |a_N||z^N|\Big(1 - \sum_{n=0}^{N-1} \frac{|a_n|}{|a_N|}|z|^{n-N}\Big). \tag{2.48}$$

4. Montrer que l'application suivante est infinie à l'infini

$$f : \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R}, \\ \mathbf{z} \longmapsto |P(z)|. \end{cases} \tag{2.49}$$

5. Montrer qu'il existe $\mathbf{z}_N$ qui réalise le minimum de f

$$\exists z_N \in \mathbb{C} \quad |P(z_N)| \leq |P(z)| \quad \forall z \in \mathbb{C}. \tag{2.50}$$

6. Montrer que $P(z_N + z)$ peut s'écrire sous la forme

$$P(z_N + z) \;=\; b_0 + b_1 z + \cdots + b_{N-1} z^{N-1} + a_N z^n \tag{2.51}$$

avec b_0, b_1, $\cdots$, b_{N-1} des complexes. Remarquer que $P(z_N) = b_0$.

7. Soit k le plus petit $n \in [\![1; N]\!]$ tel que $b_n \neq 0$. Soit $\xi \in \mathbb{C}$ une solution de l'équation

$$b_0 + b_k \xi^k = 0. \tag{2.52}$$

Pour r un réel positif, montrer

$$P(z_N + r^{\frac{1}{k}} \xi) = b_0(1 - r) + r\,\varepsilon(r) \quad \text{avec } \lim_{r \to 0} \varepsilon(r) = 0. \tag{2.53}$$

8. En utilisant le fait que $\mathbf{z}_N$ minimise $|P(z)|$, montrer que $b_0 = 0$ et que $P(z_N) = 0$. Montrer que P peut s'écrire

$$P(z) = (z - z_N) Q(z) \quad \text{avec } Q \text{ un polynôme de degré } N - 1. \tag{2.54}$$

9. Conclure.

Chapitre 3

Convexité et optimisation

3.1 Les fonctionnelles convexes

Définition 3.1 *Soit $C \subset \mathbb{R}^n$. L'ensemble C est convexe ssi*

$$\forall \mathbf{x}, \mathbf{y} \in C \quad \forall \lambda \in]0,1[\quad \lambda \mathbf{x} + (1-\lambda)\mathbf{y} \in C. \tag{3.1}$$

Dit autrement, si $\mathbf{x}$ et $\mathbf{y}$ sont deux éléments de C alors le segment qui relie $\mathbf{x}$ à $\mathbf{y}$ est inclus dans C.

Exemple. $\mathbb{R}^n$ est convexe.

Définition 3.2 *Soit $C \subset \mathbb{R}^n$ un convexe. Une fonctionnelle $J : C \longrightarrow \mathbb{R}$ est convexe ssi*

$$\forall \mathbf{x}, \mathbf{y} \in C \quad \forall \lambda \in]0,1[\quad J(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \ \leq \ \lambda J(\mathbf{x}) + (1-\lambda) J(\mathbf{y}). \tag{3.2}$$

Définition 3.3 *Soit $J : E \longrightarrow \mathbb{R}$ une fonctionnelle. $\mathbf{x} \in E$ est un argument-minimum local de J ssi*

$$\exists r > 0 \mid J(\mathbf{x}) \ \leq \ J(\mathbf{y}) \quad \forall \mathbf{y} \in E \cap B(\mathbf{x}, r). \tag{3.3}$$

$J(\mathbf{x})$ est un minimum local de E.

$\mathbf{x} \in E$ est un argument-minimum global de J ssi

$$J(\mathbf{x}) \ \leq \ J(\mathbf{y}) \quad \forall \mathbf{y} \in E. \tag{3.4}$$

$J(\mathbf{x})$ est un minimum global de E.

Proposition 3.1 *Soit $C \subset \mathbb{R}^n$ un convexe. Soit $J : C \longrightarrow \mathbb{R}$ une fonctionnelle convexe.*

Tout argument-minimum local de J sur C est argument-minimum global.

Preuve. Raisonnons par l'absurde. Soit $\mathbf{x} \in C$ un point qui réalise un minimum local de J sur C

$$\exists r > 0 \mid \forall \mathbf{y} \in C \text{ avec } \|\mathbf{y} - \mathbf{x}\| < r \quad J(\mathbf{x}) \leq J(\mathbf{y}). \tag{3.5}$$

Supposons qu'il ne réalise pas un minimum global

$$\exists \mathbf{z} \in C \mid J(\mathbf{z}) < J(\mathbf{x}). \tag{3.6}$$

Soit

$$\mathbf{y} = \lambda \mathbf{z} + (1 - \lambda)\mathbf{x} \quad \text{avec} \quad \lambda = \inf(\frac{1}{2}, \frac{r}{2\|\mathbf{x} - \mathbf{z}\|}). \tag{3.7}$$

D'une part, comme $0 < \lambda < 1$, $\mathbf{y} \in C$ et par conséquent comme J est convexe, nous avons

$$\left\{ \begin{array}{lcl} J(\mathbf{y}) & \leq & \lambda\, J(\mathbf{z}) + (1 - \lambda) J(\mathbf{x}) \\ & < & \lambda\, J(\mathbf{x}) + (1 - \lambda) J(\mathbf{x}), \\ J(\mathbf{y}) & < & J(\mathbf{x}). \end{array} \right. \tag{3.8}$$

D'autre part, comme

$$\mathbf{y} = \mathbf{x} + \lambda\,(\mathbf{z} - \mathbf{x}) \quad \text{et} \quad \|\lambda\,(\mathbf{z} - \mathbf{x})\| < r, \tag{3.9}$$

nous avons

$$\|\mathbf{x} - \mathbf{y}\| < r \tag{3.10}$$

et par suite il vient

$$J(\mathbf{y}) \geq J(\mathbf{x}). \tag{3.11}$$

Ceci est impossible, c.f. (3.8).

Définition 3.4 *Soit $C \subset \mathbb{R}^n$ un convexe. Une fonctionnelle $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ est strictement convexe ssi*

$$\forall \mathbf{x}, \mathbf{y} \in C, \quad \textit{avec } \mathbf{x} \neq \mathbf{y}, \quad \forall \lambda \in]0, 1[$$
$$J(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \; < \; \lambda J(\mathbf{x}) + (1 - \lambda) J(\mathbf{y}). \tag{3.12}$$

Théorème 3.1 *Soit $C \subset \mathbb{R}^n$ un convexe. Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle strictement convexe.*

L'argument-minimum de J sur C, s'il existe, est unique.

Preuve. Raisonnons par l'absurde. Soient $\mathbf{x}_1$ et $\mathbf{x}_2$ deux éléments de C réalisant le minimum de J.

Comme C est convexe, $(\mathbf{x}_1 + \mathbf{x}_2)/2 \in C$. Comme J est convexe, il suit

$$\left\{ \begin{array}{lcl} J(\dfrac{\mathbf{x}_1 + \mathbf{x}_2}{2}) & < & \dfrac{1}{2} J(\mathbf{x}_1) \ + \ \dfrac{1}{2} J(\mathbf{x}_2) \\ & < & \dfrac{1}{2} \inf\limits_{\mathbf{y} \in C} J(\mathbf{y}) \ + \ \dfrac{1}{2} \inf\limits_{\mathbf{y} \in C} J(\mathbf{y}) \\ & < & \inf\limits_{\mathbf{y} \in C} J(\mathbf{y}). \end{array} \right. \tag{3.13}$$

Ceci est impossible.

Corollaire 3.1 *Soit $K \subset \mathbb{R}^n$, $K \neq \emptyset$, un ensemble compact et convexe. Soit $J : K \longrightarrow \mathbb{R}$ une fonctionnelle continue strictement convexe. Alors le minimum de J est atteint en un unique $\mathbf{x} \in K$.*

Corollaire 3.2 *Soit $F \subset \mathbb{R}^n$, $F \neq \emptyset$, un ensemble fermé et convexe. Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle continue strictement convexe infinie à l'infinie. Alors le minimum de J est atteint en un unique $\mathbf{x} \in F$.*

3.2 Gradient et extremum

Rappelons tout d'abord la définition du gradient.

Définition 3.5 *Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle. J est différentiable au point $\mathbf{x}$ ssi il existe $\nabla J(\mathbf{x}) \in \mathbb{R}^n$ et $\mathbf{h} \mapsto \varepsilon(\mathbf{h}, \mathbf{x})$ vérifiant*

$$J(\mathbf{x} + \mathbf{h}) = J(\mathbf{x}) + [\nabla J(\mathbf{x})]^T \mathbf{h} \ + \ \|\mathbf{h}\| \ \varepsilon(\mathbf{h}, \mathbf{x}), \quad \forall \mathbf{h} \in \mathbb{R}^n, \tag{3.14}$$

avec

$$\varepsilon(\mathbf{h}, \mathbf{x}) \underset{\mathbf{h} \to 0}{\longrightarrow} 0. \tag{3.15}$$

Lorsqu'il existe, le gradient peut être exprimé à l'aide des dérivées partielles

$$\nabla J(\mathbf{x}) \ = \ \begin{bmatrix} \partial_{\mathbf{x}_1} J(\mathbf{x}) \\ \partial_{\mathbf{x}_2} J(\mathbf{x}) \\ \cdots \\ \partial_{\mathbf{x}_n} J(\mathbf{x}) \end{bmatrix}. \tag{3.16}$$

Définition 3.6 *Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle. J est de classe C^1.*
- *J est différentiable au point $\mathbf{x}$,*
- *chacune des composantes de ∇J est continue.*

Théorème 3.2 *Soit $O \subset \mathbb{R}^n$ un ouvert. Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle différentiable.*

Si $\mathbf{x} \in O$ réalise le minimum (respectivement maximum) de J alors

$$\nabla J(\mathbf{x}) = 0. \tag{3.17}$$

Preuve. Traitons le cas où $\mathbf{x}$ réalise le minimum de J.

Comme O est ouvert et $\mathbf{x} \in O$, il existe $r > 0$ tel que $B(\mathbf{x}, r) \subset O$. Ainsi, pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $\|\mathbf{h}\| < r$, $\mathbf{x} + \mathbf{h} \in O$. Comme $\mathbf{x}$ réalise le minimum de J sur O

$$J(\mathbf{x} + \mathbf{h}) - J(\mathbf{x}) \geq 0. \tag{3.18}$$

Traduisons cette expression à l'aide du gradient

$$[\nabla J(\mathbf{x})]^T \mathbf{h} \ + \ \|\mathbf{h}\| \, \varepsilon(\mathbf{h}, \mathbf{x}) \geq 0. \tag{3.19}$$

Choisissons $\mathbf{h} = -\frac{\nabla J(\mathbf{x})}{\|\nabla J(\mathbf{x})\|}\lambda$ avec $0 < \lambda < \frac{r}{2}$, cette expression s'écrit

$$-\lambda \|\nabla J(\mathbf{x})\| \ + \ \lambda \, \varepsilon(-\frac{\nabla J(\mathbf{x})}{\|\nabla J(\mathbf{x})\|}\lambda, \mathbf{x}) \geq 0. \tag{3.20}$$

C'est-à-dire

$$-\|\nabla J(\mathbf{x})\| \ + \ \varepsilon(-\frac{\nabla J(\mathbf{x})}{\|\nabla J(\mathbf{x})\|}\lambda, \mathbf{x}) \geq 0. \tag{3.21}$$

En faisant tendre λ vers 0, nous tirons

$$-\|\nabla J(\mathbf{x})\| \geq 0. \tag{3.22}$$

Or comme $\|\nabla J(\mathbf{x})\| \geq 0$, ceci implique $\nabla J(\mathbf{x}) = 0$.

Ceci conclut la preuve.

Dans le cas d'un maximum, on effectuera la même preuve avec $\mathbf{h} = \frac{\nabla J(\mathbf{x})}{\|\nabla J(\mathbf{x})\|}\lambda$.

Remarque 3.1 *Tout point où le gradient est nul n'est pas nécessairement un extremum. En effet pour $f : \mathbb{R} \longrightarrow \mathbb{R}$ avec $f(x) = x^3$ admet en $x = 0$ un point de dérivée nulle qui n'est pas un extremum (même local).*

3.3 Hessienne et convexité

Définition 3.7 *On dit que $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ est deux fois différentiable au point $\mathbf{x} \in \mathbb{R}^n$ ssi*

– *J est différentiable ;*

– *chacune des composantes de ∇J est différentiable.*

Définition 3.8 *Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ deux fois différentiable. La Hessienne de J définie à l'aide de ses dérivées partielles d'ordre 2*

$$H[J](\mathbf{x}) \;=\; [\partial_i \partial_j J(\mathbf{x})]_{1\leq i,j\leq n}. \tag{3.23}$$

Définition 3.9 *Une fonctionnelle $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ est de classe $\mathcal{C}^2$ ssi*
– *J est deux fois différentiable ;*
– *chacune des composantes de $H[J]$ est continue.*

Théorème 3.3 *(de Schwarz) Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^2$.*

La matrice $H[J](\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^n$, est symétrique.

Théorème 3.4 *(Développement de Taylor d'ordre 2) Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^2$. Si*

$$J(\mathbf{x}+\mathbf{h}) \;=\; J(\mathbf{x}) \;+\; [\nabla J(\mathbf{x})]^T\mathbf{h} \;+\; \frac{1}{2}\mathbf{h}^T H[J](\mathbf{x})\mathbf{h} \;+\; \|\mathbf{h}\|^2\,\varepsilon(\mathbf{x},\mathbf{h}) \tag{3.24}$$

avec

$$\varepsilon(\mathbf{x},\mathbf{h}) \underset{\mathbf{h}\to 0}{\longrightarrow} 0 \tag{3.25}$$

La Hessienne a notamment une importance vis-à-vis du résultat suivant.

Théorème 3.5 *Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^2$.*

Si la hessienne $H[J](\mathbf{x})$ de J est une matrice symétrique définie positive pour tout $\mathbf{x} \in \mathbb{R}^n$, alors J est strictement convexe.

Si $H[J](x)$ est une matrice symétrique positive pour tout $\mathbf{x} \in \mathbb{R}^n$, alors J est convexe.

Pour montrer ce résultat nous avons besoin du lemme technique suivant.

Lemme 3.1 *Soit $f \in \mathcal{C}^2([0,1])$.*
Si $f^{(2)}(x) > 0$, alors $f(\lambda) < (1-\lambda)f(0) + \lambda f(1)$ pour $\lambda \in]0,1[$.
Si $f^{(2)}(x) \geq 0$, alors $f(\lambda) \leq (1-\lambda)f(0) + \lambda f(1)$.

Preuve. Comme $f^{(2)}(x) > 0$ (resp. $\geq$) pour $0 \leq x \leq 1$, f' est strictement croissante (resp. croissante) et par conséquent

$$\begin{aligned} f(\lambda) - f(0) &= \int_0^\lambda f'(t)dt < \lambda f'(\lambda) \quad \text{(resp. } \leq\text{)}, \\ f(1) - f(\lambda) &= \int_\lambda^1 f'(t)dt > (1-\lambda) f'(\lambda) \quad \text{(resp. } \geq\text{)}. \end{aligned} \tag{3.26}$$

En groupant ces deux inégalités, nous tirons

$$\frac{f(\lambda)-f(0)}{\lambda} < \Big(f'(\lambda) < \Big)\frac{f(1)-f(\lambda)}{1-\lambda} \quad \text{(resp. } \leq\text{);} \tag{3.27}$$

c'est-à-dire

$$f(\lambda) < (1-\lambda)f(0) + \lambda f(1) \quad \text{(resp. } \leq\text{).} \tag{3.28}$$

Preuve du théorème 3.5. Supposons $H[J](\mathbf{z})$ symétrique définie positive. Soient $\mathbf{x}$ et $\mathbf{y}$ deux éléments distincts de $\mathbb{R}^n$. Introduisons la fonction $f : [0,1] \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ définie par

$$f(\lambda) = J(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}). \tag{3.29}$$

Utilisons le développement à l'ordre deux de J pour calculer la dérivée seconde de f

$$f(\lambda+h) = J((\lambda+h)\mathbf{x} + (1-\lambda-h)\mathbf{y}) = J(\lambda\mathbf{x} + (1-\lambda)\mathbf{y} + h(\mathbf{x}-\mathbf{y})). \tag{3.30}$$

Avec $\mathbf{t} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$, nous avons

$$\begin{aligned} f(\lambda+h) = J(\mathbf{t}) + [\nabla J(\mathbf{t})]^T h(\mathbf{x}-\mathbf{y}) + \frac{1}{2}h(\mathbf{x}-\mathbf{y})^T H[J](\mathbf{t}) h(\mathbf{x}-\mathbf{y}) \\ + h^2\|\mathbf{x}-\mathbf{y}\|^2\varepsilon(\mathbf{t}, h(\mathbf{x}-\mathbf{y})), \end{aligned} \tag{3.31}$$

et, par conséquent

$$\begin{aligned} f(\lambda+h) = f(\lambda) + h\Big([\nabla J(\mathbf{t})]^T(\mathbf{x}-\mathbf{y})\Big) \\ + \frac{h^2}{2}\Big((\mathbf{x}-\mathbf{y})^T H[J](\mathbf{t})(\mathbf{x}-\mathbf{y})\Big) + h^2\varepsilon(h) \end{aligned} \tag{3.32}$$

Comme f est de classe $\mathcal{C}^2$, la dérivée première et la dérivée seconde de f sont donc données par

$$\begin{aligned} f'(\lambda) &= [\nabla J(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})]^T(\mathbf{x}-\mathbf{y}), \\ f^{(2)}(\lambda) &= (\mathbf{x}-\mathbf{y})^T \underbrace{H[J]((\lambda\mathbf{x} + (1-\lambda)\mathbf{y})}_{\in\mathbb{R}^{n\times n}}(\mathbf{x}-\mathbf{y}). \end{aligned} \tag{3.33}$$

Donc, la fonction $f^{(2)}(\lambda)$ est strictement positive car $H[J](\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$ est symétrique définie positive et $\mathbf{x} \neq \mathbf{y}$.

Le lemme 3.1, s'écrit ($f(0) = J(\mathbf{y})$, $f(1) = J(\mathbf{x})$ et $f(\lambda) = J(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$)

$$J(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) < (1-\lambda)J(\mathbf{y}) + \lambda J(\mathbf{x}). \tag{3.34}$$

Ceci prouve que J est strictement convexe.

Si pour tout $\mathbf{x}$, $H[J](\mathbf{x})$ est positive une preuve similaire nous permet de montrer que J est convexe.

Proposition 3.2 *Soit $O \subset \mathbb{R}^n$ un ouvert. Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^2$. Soit $\mathbf{x} \in O$ qui réalise le minimum de J sur O. Alors, J vérifie*

$$\nabla J(\mathbf{x}) \ = \ 0 \quad \text{et} \quad H[J](\mathbf{x}) \text{ est positive.} \tag{3.35}$$

Remarque 3.2 *La matrice $H[J](\mathbf{x})$ peut ne pas être définie voire même être nulle ($f(x) = x^4$ par exemple).*

Preuve. Soit $\mathbf{x} \in O$ tel que $J(\mathbf{x}) \leq J(\mathbf{z})$ pour tout $\mathbf{z} \in O$.

$$J(\mathbf{x}+\mathbf{h}) \ = \ J(\mathbf{x}) \ + \ [\nabla J(\mathbf{x})]^T \mathbf{h} + \mathbf{h}^T H[J](x)\mathbf{h} + \|\mathbf{h}\|^2 \, \varepsilon(\mathbf{x},\mathbf{h}) \tag{3.36}$$

avec

$$\lim_{\mathbf{h}\to 0} \varepsilon(\mathbf{x},\mathbf{h}) = 0. \tag{3.37}$$

Par conséquent, comme $\mathbf{x}$ réalise le minimum de F sur O

$$J(\mathbf{x}+\mathbf{h}) \ - \ J(\mathbf{x}) \ \geq \ 0 \quad \text{et} \quad \nabla J(\mathbf{x}) = 0 \tag{3.38}$$

Pour $\mathbf{y} \in \mathbb{R}^n$ et $\lambda > 0$ suffisament petit, $\mathbf{x} + \lambda\mathbf{y} \in O$. Nous tirons pour $\mathbf{h} = \lambda\mathbf{y}$

$$\lambda\mathbf{y}^T H[J](x)\lambda\mathbf{y} + \|\lambda\mathbf{y}\|^2 \, \varepsilon(\mathbf{x},\lambda\mathbf{y}) \ \geq \ 0. \tag{3.39}$$

Comme $\lambda^2 > 0$, nous avons

$$\mathbf{y}^T H[J](\mathbf{x})\mathbf{y} + \varepsilon(\lambda) \ \geq \ 0 \quad \text{avec } \lim_{\lambda\to 0} \varepsilon(\lambda) = 0. \tag{3.40}$$

En passant à la limite, nous obtenons le résultat

$$\mathbf{y}^T H[J](\mathbf{x})\mathbf{y} \geq 0 \quad \forall \mathbf{y} \in \mathbb{R}^n. \tag{3.41}$$

Définition 3.10 *soit $E \subset \mathbb{R}^n$. Soit $J : E \longrightarrow \mathbb{R}$ une fonctionnelle. Le point $\mathbf{x} \in \mathbb{R}^n$ est un point de minimum local ssi*

$$\exists r > 0 \quad \forall y \in B(\mathbf{x}, r) \cap E \quad J(\mathbf{y}) \geq J(\mathbf{x}). \tag{3.42}$$

Théorème 3.6 *Soit O un ouvert Soit $J : O \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^2$.*

Si $\mathbf{x} \in O$ vérifie

$$\nabla J(\mathbf{x}) = 0 \quad \text{et} \quad H[J](\mathbf{x}) \text{ est symétrique définie positive} \tag{3.43}$$

alors $\mathbf{x}$ est un point de minimum local.

Remarque 3.3 *Comme J est de classe $\mathcal{C}^2$, il n'est pas nécessaire de vérifier que la matrice $H[J](\mathbf{x})$ est symétrique.*

3.4 Exercices

Exercice. Montrer que les boules ouvertes $B(\mathbf{x}, r)$ et les boules fermées $\overline{B}(\mathbf{x}, r)$ sont convexes.

Exercice. Montrer que l'intersection de deux convexes est convexe. Montrer que l'union de deux convexes n'est pas nécessairement convexe.

Exercice. Soit $A \in \mathbb{R}^{n\times n}$ et $b \in \mathbb{R}^n$. Calculer le gradient des fonctionnelles $\mathbb{R}^n \longrightarrow \mathbb{R}$

$$\mathbf{x} \longmapsto \mathbf{b}^T\mathbf{x} \quad \text{et} \quad \mathbf{x} \longmapsto \mathbf{x}^T A\mathbf{x} \tag{3.44}$$

Exercice. Montrer que les applications suivantes sont convexes

$$\mathbf{x} \longmapsto \|\mathbf{x}\| \quad \text{et} \quad \mathbf{x} \longmapsto \|\mathbf{x}\|^2. \tag{3.45}$$

Exercice. soit $f \in \mathcal{C}^2(\mathbb{R})$ tel que $f^{(2)}$ est strictement positive. Montrer que f est strictement convexe.

Exercice. Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique définie positive. Soient $\mathbf{b} \in \mathbb{R}^n$ et $c \in \mathbb{R}$. Considérons la fonctionnelle $J : \mathbb{R}^n \longrightarrow \mathbb{R}$

$$J(\mathbf{x}) \ = \ \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x} + c. \tag{3.46}$$

1) Pour $\alpha > 0$, $\beta \in \mathbb{R}$, $\gamma \in \mathbb{R}$ et $f(\lambda) = \alpha\lambda^2 + \beta\lambda + \gamma$, Montrer que

$$f(\lambda) < \lambda\, f(0) \ + \ (1-\lambda)f(1) \quad \forall\lambda \in]0,1[. \tag{3.47}$$

2) Montrer que J est strictement convexe.

3) Montrer que le minimum de J est atteint en un unique point.

4) Caractériser ce minimum en calculant le gradient de J.

Chapitre 4

Optimisation avec contrainte

4.1 Minimisation avec une contrainte d'égalité

Nous cherchons à résoudre le problème suivant

$$\left\{ \begin{array}{l} \text{Chercher } \mathbf{x} \text{ argument-minimum de } J : \mathbb{R}^n \longrightarrow \mathbb{R} \text{ sur} \\ \qquad C \;=\; \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; f(\mathbf{x}) = 0 \right\}. \end{array} \right. \tag{4.1}$$

On dit que C est l'ensemble des contraintes (c'est ici une contrainte d'égalité).

Définition 4.1 *Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle.*

On appelle lagrangien de J sous la contrainte f la fonctionnelle

$$\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \longrightarrow \mathbb{R}, \quad \mathcal{L}(\mathbf{x}, \lambda) \;=\; J(\mathbf{x}) + \lambda\, f(\mathbf{x}). \tag{4.2}$$

$\lambda \in \mathbb{R}$ est appelé multiplicateur de Lagrange.

Calcul du gradient de $\mathcal{L}$.

Supposons que J et f sont de classe $\mathcal{C}^1$. Calculons le gradient de

$$\mathcal{L} : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}, \quad (\mathbf{x}, \lambda) \in \mathbb{R}^{n+1} \tag{4.3}$$

$$\nabla_{x,\lambda}\mathcal{L}(\mathbf{x}, \lambda) \;=\; \begin{bmatrix} \nabla_x \mathcal{L}(\mathbf{x}, \lambda) \in \mathbb{R}^n \\ \partial_\lambda \mathcal{L}(\mathbf{x}, \lambda) \in \mathbb{R} \end{bmatrix} \;=\; \begin{bmatrix} \partial_{\mathbf{x}_1} \mathcal{L}(\mathbf{x}, \lambda) \in \mathbb{R} \\ \cdots \\ \partial_{\mathbf{x}_n} \mathcal{L}(\mathbf{x}, \lambda) \in \mathbb{R} \\ \partial_\lambda \mathcal{L}(\mathbf{x}, \lambda) \in \mathbb{R} \end{bmatrix} \tag{4.4}$$

où on entend par

– $\partial_{\mathbf{x}_i}$ la dérivée suivant $\mathbf{x}_i$ avec, pour $j \neq i$, $j \in [\![1; n]\!]$, $\mathbf{x}_j$ et λ fixés.
– ∂_λ la dérivée suivant λ avec, pour tout $j \in [\![1; n]\!]$, $\mathbf{x}_j$ fixés.

Le calcul direct nous fournit

$$\left\{\begin{array}{lcl} \partial_{\mathbf{x}_1}\mathcal{L}(\mathbf{x},\lambda) & = & \partial_{\mathbf{x}_1}J(\mathbf{x}) + \lambda\,\partial_{\mathbf{x}_1}f(\mathbf{x}), \\ \dots & & \\ \partial_{\mathbf{x}_n}\mathcal{L}(\mathbf{x},\lambda) & = & \partial_{\mathbf{x}_n}J(\mathbf{x}) + \lambda\,\partial_{\mathbf{x}_n}f(\mathbf{x}), \\ \partial_\lambda\mathcal{L}(\mathbf{x},\lambda) & = & f(\mathbf{x}). \end{array}\right. \tag{4.5}$$

Ainsi, il suit

$$\nabla_{\mathbf{x},\lambda}\mathcal{L}(\mathbf{x},\lambda) \;=\; \begin{bmatrix} \nabla_{\mathbf{x}}J(\mathbf{x}) \;+\; \lambda\,\nabla_{\mathbf{x}}f(\mathbf{x}) \\ f(\mathbf{x}) \end{bmatrix}. \tag{4.6}$$

Nous avons prouvé la proposition

Proposition 4.1 *Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^1$. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^1$.*

Si $\nabla_{\mathbf{x},\lambda}\mathcal{L}(\mathbf{x},\lambda) = 0$ alors $f(\mathbf{x}) = 0$.

On a alors le résultat suivant

Théorème 4.1 *Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^1$. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^1$. Soit $C = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = 0\}$.*

Si $\mathbf{x}$ est un argument-minimum de J sur C vérifiant $\nabla_{\mathbf{x}}f(\mathbf{x}) \neq 0$ alors

$$\exists\lambda \in \mathbb{R} \quad | \quad \nabla_{\mathbf{x},\lambda}\mathcal{L}(\mathbf{x},\lambda) \;=\; 0. \tag{4.7}$$

Remarque 4.1 *D'après le théorème précédent, les $\mathbf{x}$ minimisants J sur C vérifient soit*

$$\exists\lambda \in \mathbb{R} \quad | \quad \nabla_{\mathbf{x},\lambda}\mathcal{L}(\mathbf{x},\lambda) = 0 \tag{4.8}$$

soit

$$f(\mathbf{x}) = 0 \text{ et } \nabla f(\mathbf{x}) = 0. \tag{4.9}$$

Pour déterminer les $\mathbf{x}$ arguments-minima de J sur C, on utilise la technique suivante

1. *On montre l'existence d'un argument-minimum.*
2. *On trouve l'ensemble E des $\mathbf{x}$ vérifiant (4.9) ou (4.8).*
3. *Les arguments-minima de J sur C sont les arguments-minima de J sur E. On calcule donc $J(\mathbf{x})$ pour $\mathbf{x} \in E$.*

Exemple : Minimisation d'une forme linéaire sur une sphère

Pour $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{b} \neq 0$, minimisons

$$J : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad J(\mathbf{x}) \;=\; \mathbf{b} \cdot \mathbf{x} \tag{4.10}$$

sur $C = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|^2 = 1\}$.

1. On montre l'existence d'un argument-minimum.

Notons $f(\mathbf{x}) = \|\mathbf{x}\|^2 - 1$. L'ensemble C nous est donné par

$$C = \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; f(\mathbf{x}) = 0 \right\} = \overset{-1}{f}(\{0\}). \tag{4.11}$$

Remarquons que C est fermé car f est continue et $\{0\}$ est fermé. D'autre part, C est borné car $\|\mathbf{x}\| \leqslant 1$ pour tout $\mathbf{x} \in C$.

Comme C est compact (fermé et borné) et f est continue, il existe un argument-minimum de f sur C.

2. On trouve l'ensemble E des $\mathbf{x}$ vérifiant (4.9) ou (4.8).

Soit $\mathcal{L}(\mathbf{x}, \lambda) = J(\mathbf{x}) + \lambda f(\mathbf{x})$. Cherchons l'ensemble des $\mathbf{x}$ vérifiant

$$f(\mathbf{x}) = 0 \quad \text{et} \quad \nabla_{\mathbf{x}} f(\mathbf{x}) = 0. \tag{4.12}$$

Calculons le gradient de f

$$\left\{ \begin{array}{rcl} f(\mathbf{x} + \mathbf{h}) & = & (\mathbf{x} + \mathbf{h}) \cdot (\mathbf{x} + \mathbf{h}) - 1 \\ & = & \mathbf{x} \cdot \mathbf{x} - 1 + \mathbf{x} \cdot \mathbf{h} + \mathbf{h} \cdot \mathbf{x} + \mathbf{h} \cdot \mathbf{h} \\ & = & \|\mathbf{x}\|^2 - 1 + 2\, \mathbf{x} \cdot \mathbf{h} + \|\mathbf{h}\|^2. \end{array} \right. \tag{4.13}$$

En identifiant avec le développement de Taylor d'ordre 1

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \nabla_{\mathbf{x}} f(\mathbf{x}) \cdot \mathbf{h} + \|\mathbf{h}\|\, \varepsilon(\mathbf{h}), \tag{4.14}$$

on tire avec $\varepsilon(\mathbf{h}) = \|h\|$ (qui tend vers 0 avec $\mathbf{h}$)

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = 2\, \mathbf{x}. \tag{4.15}$$

L'ensemble des $\mathbf{x}$ tels que $f(\mathbf{x}) = 0$ et $\nabla_{\mathbf{x}} f(\mathbf{x}) = 0$ est vide. En effet, il n'existe pas de $\mathbf{x}$ vérifiant

$$\|\mathbf{x}\| = 1 \quad \text{et} \quad 2\, \mathbf{x} = 0. \tag{4.16}$$

Déterminons les $\mathbf{x}$ vérifiant $\nabla_{\mathbf{x},\lambda} \mathcal{L}(\mathbf{x}, \lambda) = 0$. Calculons le gradient suivant $\mathbf{x}$

$$\nabla_{\mathbf{x}} \mathcal{L}(x, \lambda) = \nabla_{\mathbf{x}} J(\mathbf{x}) + \lambda \nabla_{\mathbf{x}} f(\mathbf{x}). \tag{4.17}$$

Calculons le gradient de J

$$J(\mathbf{x}+\mathbf{h}) = \mathbf{b}\cdot(\mathbf{x}+\mathbf{h}) = \mathbf{b}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{h} \tag{4.18}$$

d'où (en vérifiant que $\varepsilon(\mathbf{h}) = 0$ tend bien vers 0 avec $\mathbf{h}$)

$$\nabla_{\mathbf{x}} J(\mathbf{x}) \;=\; \mathbf{b}. \tag{4.19}$$

Ainsi, nous avons

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x},\lambda) = \mathbf{b} \;+\; \lambda\, 2\,\mathbf{x} \tag{4.20}$$

D'autre part, on peut calculer la dérivée partielle suivant λ

$$\partial_\lambda \mathcal{L}(\mathbf{x},\lambda) = f(\mathbf{x}) = \|\mathbf{x}\|^2 - 1 \tag{4.21}$$

D'où $\nabla_{\mathbf{x},\lambda}\mathcal{L}(\mathbf{x},\lambda) = 0$ implique

$$2\lambda\,\mathbf{x} \;=\; -\mathbf{b} \quad \text{et} \quad \|\mathbf{x}\| = 1 \tag{4.22}$$

Par conséquent on obtient

$$\lambda = \pm\frac{\|\mathbf{b}\|}{2} \quad \text{et} \quad \mathbf{x} = \pm\frac{\mathbf{b}}{\|\mathbf{b}\|}. \tag{4.23}$$

Notons $E = \{\frac{\mathbf{b}}{\|\mathbf{b}\|}, -\frac{\mathbf{b}}{\|\mathbf{b}\|}\}$.

3. Les arguments-minima de J sur C sont les arguments-minima de J sur E. On calcule donc $J(\mathbf{x})$ pour $\mathbf{x} \in E$.

Calculons $J(\mathbf{x})$ pour $\mathbf{x} \in E$

$$J(\frac{\mathbf{b}}{\|\mathbf{b}\|}) = \frac{\mathbf{b}\cdot\mathbf{b}}{\|\mathbf{b}\|} = \|\mathbf{b}\| \quad \text{et} \quad J(-\frac{\mathbf{b}}{\|\mathbf{b}\|}) = -\frac{\mathbf{b}\cdot\mathbf{b}}{\|\mathbf{b}\|} = -\|\mathbf{b}\|. \tag{4.24}$$

Ainsi, l'argument-minimum de J existe et est unique et vaut $\mathbf{x} = -\frac{\mathbf{b}}{\|\mathbf{b}\|}$.

La preuve du théorème 4.1 est basée sur le théorème des fonctions implicites.

Lemme 4.1 *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonctionnelle de classe $\mathcal{C}^1$. Soit $\mathbf{x} \in \mathbb{R}^n$ tel que*

$$\nabla f(\mathbf{x}) \neq 0 \quad \textit{et} \quad f(\mathbf{x}) = 0. \tag{4.25}$$

Soit $i \in [\![1; N]\!]$ tel que $\partial_{\mathbf{x}_i} f(\mathbf{x}) \neq 0$.

Il existe un ouvert O de $\mathbb{R}^{n-1}$ avec

$$(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n) \in O \tag{4.26}$$

et $g : O \subset \mathbb{R}^{n-1} \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ tel que

$$\mathbf{x}_i \ = \ g(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n) \tag{4.27}$$

et pour tout $(\mathbf{y}_1, \cdots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \cdots, \mathbf{y}_n) \in O$

$$f\begin{pmatrix} \mathbf{y}_1 \\ \cdots \\ \mathbf{y}_{i-1} \\ g(\mathbf{y}_1, \cdots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \cdots, \mathbf{y}_n) \\ \mathbf{y}_{i+1} \\ \cdots \\ \mathbf{y}_n \end{pmatrix} = 0. \tag{4.28}$$

D'autre part, on a

$$\nabla_{\mathbf{x}} f(\mathbf{y}) \ = \ \partial_{\mathbf{x}_i} f(\mathbf{y})\Big(\mathbf{e}_i - \nabla_{\mathbf{x}} g(\mathbf{y}_1, \cdots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \cdots, \mathbf{y}_n)\Big) \tag{4.29}$$

où e_i désigne le ième vecteur de la base canonique et $\nabla_{\mathbf{x}} g(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n)$ désigne le gradient de g en tant que fonction de la variable $\mathbf{y} = (\mathbf{y}_1, \cdots, \mathbf{y}_n) \in \mathbb{R}^n$.

On se référera au cours de fonctions de plusieurs variables pour la preuve.

Preuve du théorème 4.1 Soit x un argument-minimum de J sur C tel que

$$\nabla f(\mathbf{x}) \neq 0. \tag{4.30}$$

Réduction du problème à $\mathbb{R}^{n-1}$. Soit $i \in [\![1, n]\!]$ tel que $\partial_{\mathbf{x}_i} f(\mathbf{x}) \neq 0$.

Il existe un ouvert O de $\mathbb{R}^{n-1}$ avec

$$(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n) \in O \tag{4.31}$$

et $g : O \subset \mathbb{R}^{n-1} \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ tel que

$$\mathbf{x}_i \ = \ g(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n); \tag{4.32}$$

et pour tout $(\mathbf{y}_1, \cdots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \cdots, \mathbf{y}_n) \in O$

$$\begin{pmatrix} \mathbf{y}_1 \\ \cdots \\ \mathbf{y}_{i-1} \\ g(\mathbf{y}_1, \cdots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \cdots, \mathbf{y}_n) \\ \mathbf{y}_{i+1} \\ \cdots \\ \mathbf{y}_n \end{pmatrix} \in C. \tag{4.33}$$

Introduisons la fonctionnelle $J_C : O \subset \mathbb{R}^{n-1} \longrightarrow \mathbb{R}$

$$J_C \begin{pmatrix} \mathbf{y}_1 \\ \cdots \\ \mathbf{y}_{i-1} \\ \mathbf{y}_{i+1} \\ \cdots \\ \mathbf{y}_n \end{pmatrix} = J \begin{pmatrix} \mathbf{y}_1 \\ \cdots \\ \mathbf{y}_{i-1} \\ g(\mathbf{y}_1, \cdots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \cdots, \mathbf{y}_n) \\ \mathbf{y}_{i+1} \\ \cdots \\ \mathbf{y}_n \end{pmatrix}. \tag{4.34}$$

Comme $\mathbf{x}$ est un argument-minimum de J sur C,

$$(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n) \tag{4.35}$$

est un argument-minimum de J_C sur O. Comme O est ouvert et J_C est de classe $\mathcal{C}_1$, nous avons

$$\nabla J_C(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n) = 0 \in \mathbb{R}^{n-1}. \tag{4.36}$$

Calcul du gradient de J_C. Le théorème de composition des dérivées nous fournit

$$\nabla J_C \begin{pmatrix} \mathbf{x}_1 \\ \cdots \\ \mathbf{x}_{i-1} \\ \mathbf{x}_{i+1} \\ \cdots \\ \mathbf{x}_n \end{pmatrix} = \begin{pmatrix} \partial_{\mathbf{x}_1} J(\mathbf{x}) + \partial_{\mathbf{x}_i} J(\mathbf{x}) \partial_{\mathbf{x}_1} g(\mathbf{z}) \\ \cdots \\ \partial_{\mathbf{x}_{i-1}} J(\mathbf{x}) + \partial_{\mathbf{x}_i} J(\mathbf{x}) \partial_{\mathbf{x}_{i-1}} g(\mathbf{z}) \\ \partial_{\mathbf{x}_{i+1}} J(\mathbf{x}) + \partial_{\mathbf{x}_i} J(\mathbf{x}) \partial_{\mathbf{x}_{i+1}} g(\mathbf{z}) \\ \cdots \\ \partial_{\mathbf{x}_n} J(\mathbf{x}) + \partial_{\mathbf{x}_i} J(\mathbf{x}) \partial_{\mathbf{x}_n} g(\mathbf{z}) \end{pmatrix} \tag{4.37}$$

avec $\mathbf{z} = (\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_n)$. D'autre part (4.29) implique que

$$\nabla J_C(\mathbf{z}) = \begin{pmatrix} \partial_{\mathbf{x}_1} J(\mathbf{x}) - \dfrac{\partial_{\mathbf{x}_i} J(\mathbf{x})}{\partial_{\mathbf{x}_i} f(\mathbf{x})} \partial_{\mathbf{x}_1} f(\mathbf{x}) \\ \cdots \\ \partial_{\mathbf{x}_{i-1}} J(\mathbf{x}) - \dfrac{\partial_{\mathbf{x}_i} J(\mathbf{x})}{\partial_{\mathbf{x}_i} f(\mathbf{x})} \partial_{\mathbf{x}_{i-1}} f(\mathbf{x}) \\ \partial_{\mathbf{x}_{i+1}} J(\mathbf{x}) - \dfrac{\partial_{\mathbf{x}_i} J(\mathbf{x})}{\partial_{\mathbf{x}_i} f(\mathbf{x})} \partial_{\mathbf{x}_{i+1}} f(\mathbf{x}) \\ \cdots \\ \partial_{\mathbf{x}_n} J(\mathbf{x}) - \dfrac{\partial_{\mathbf{x}_i} J(\mathbf{x})}{\partial_{\mathbf{x}_i} f(\mathbf{x})} \partial_{\mathbf{x}_n} f(\mathbf{x}) \end{pmatrix} \tag{4.38}$$

Conclusion. Soit $\lambda = -\frac{\partial_{\mathbf{x}_i} J(\mathbf{x})}{\partial_{\mathbf{x}_i} f(\mathbf{x})} \in \mathbb{R}$. Pour $j \neq i$, de (4.38) et (4.36), nous tirons

$$\partial_{\mathbf{x}_j} J(\mathbf{x}) \; + \; \lambda \, \partial_{\mathbf{x}_j} f(\mathbf{x}) = 0. \tag{4.39}$$

Pour $j = i$, par défintion de λ, nous avons aussi

$$\partial_{\mathbf{x}_i} J(\mathbf{x}) \ + \ \lambda \, \partial_{\mathbf{x}_i} f(\mathbf{x}) = 0. \tag{4.40}$$

Ceci s'écrit sous forme vectorielle

$$\nabla J(\mathbf{x}) + \lambda \nabla f(\mathbf{x}) = 0. \tag{4.41}$$

D'autre part, $\mathbf{x} \in C$ se traduit en

$$f(\mathbf{x}) = 0. \tag{4.42}$$

Enfin, (4.41) et (4.42) s'écrivent sous forme compacte

$$\nabla_{\mathbf{x},\lambda} \mathcal{L}(\mathbf{x}, \lambda) = 0. \tag{4.43}$$

Exemple. Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique définie positive. Soit $\mathbf{b} \in \mathbb{R}^n$. Minimisons dans $\mathbb{R}^n$

$$J(\mathbf{x}) = \frac{1}{2} \mathbf{x} \cdot A\mathbf{x} \tag{4.44}$$

sous la contrainte

$$f(\mathbf{x}) = 0 \quad \text{ avec } f(\mathbf{x}) = \mathbf{b} \cdot \mathbf{x} - 1. \tag{4.45}$$

1. On montre l'existence d'un argument-minimum.

Comme A est symétrique définie positive, A admet une base orthonormale de vecteurs propres ($\mathbf{v}_i \in \mathbb{R}^n$) associée à des valeurs propres $(\lambda_i > 0)$.

$$A\mathbf{v}_i \ = \ \lambda_i \, \mathbf{v_i}. \tag{4.46}$$

Tout $\mathbf{x} \in \mathbb{R}^n$ se représente sous la forme

$$\mathbf{x} \ = \ \sum_{i=1}^{n} X_i \, \mathbf{v}_i \quad \text{ avec } X_i \in \mathbb{R}. \tag{4.47}$$

On peut à l'aide de cette décomposition montrer que J est infinie à l'infini

$$J(\mathbf{x}) = \frac{1}{2} \Big(\sum_{i=1}^{n} X_i \, \mathbf{v}_i \Big) \cdot A \Big(\sum_{i=1}^{n} X_i \, \mathbf{v}_i \Big) \ = \ \frac{1}{2} \sum_{i=1}^{n} \lambda_i \, X_i^2 \geqslant \frac{1}{2} \, \lambda_{min} \, \|\mathbf{x}\|^2. \tag{4.48}$$

D'autre part, $C = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{b} \cdot \mathbf{x} - 1 = 0\} = \overset{-1}{f}(\{0\})$ est fermé car $\{0\}$ est fermé et f est continue.

Comme C est fermé non vide ($\frac{\mathbf{b}}{\|\mathbf{b}\|^2} \in C$) et J est continue infinie à l'infini, on obtient l'existence d'un argument-minimum.

2. On trouve l'ensemble E des x vérifiant (4.9) ou (4.8).

Définissons le lagrangien de J sous la contrainte f

$$\mathcal{L}(\mathbf{x}, \lambda) \;=\; J(\mathbf{x}) + \lambda\, f(\mathbf{x}). \tag{4.49}$$

Déterminons l'ensemble des $\mathbf{x} \in \mathbb{R}^n$ tels que

$$\exists \lambda \in \mathbb{R} \quad | \quad \nabla_{\mathbf{x},\lambda} \mathcal{L}(\mathbf{x}, \lambda) = 0. \tag{4.50}$$

Ceci s'écrit

$$\left\{ \begin{aligned} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda) \;&=\; \nabla J(\mathbf{x}) + \lambda \nabla f(\mathbf{x}) = A\mathbf{x} + \lambda \mathbf{b} = 0, \\ \partial_\lambda \mathcal{L}(\mathbf{x}, \lambda) \;&=\; f(\mathbf{x}) \;=\; \mathbf{b} \cdot \mathbf{x} - 1 \;=\; 0. \end{aligned} \right. \tag{4.51}$$

donc $\mathbf{x} = -\lambda\, A^{-1}\mathbf{b}$ (A est inversible car définie). Par conséquent, nous avons

$$1 = \mathbf{b} \cdot \mathbf{x} = -\lambda\, \mathbf{b} \cdot A^{-1}\mathbf{b} \quad \Longrightarrow \quad \lambda = \frac{-1}{\mathbf{b} \cdot A^{-1}\mathbf{b}}. \tag{4.52}$$

Nous trouvons donc $\mathbf{x} = \frac{A^{-1}\mathbf{b}}{\mathbf{b} \cdot A^{-1}\mathbf{b}}$. Déterminons l'ensemble des $\mathbf{x} \in \mathbb{R}^n$ tels que $f(\mathbf{x}) = 0$ et $\nabla f(\mathbf{x}) = 0$. La seconde condition s'écrit $\mathbf{b} = 0$. Cet ensemble est donc vide.

$$E \;=\; \left\{ \frac{A^{-1}\mathbf{b}}{\mathbf{b} \cdot A^{-1}\mathbf{b}} \right\}. \tag{4.53}$$

3. Les arguments-minima de J sur C sont les arguments-minima de J sur E. On calcule donc $J(\mathrm{x})$ pour $\mathrm{x} \in E$.

Comme E ne contient qu'un point, il s'agit de l'agument-minimum de J sur C. Pour $\mathbf{x} = \frac{A^{-1}\mathbf{b}}{\mathbf{b} \cdot A^{-1}\mathbf{b}}$, $J(\mathbf{x})$ prend la valeur

$$J(\mathbf{x}) \;=\; \frac{1}{2} \frac{A^{-1}\mathbf{b}}{\mathbf{b} \cdot A^{-1}\mathbf{b}} \cdot \frac{A\, A^{-1}\mathbf{b}}{\mathbf{b} \cdot A^{-1}\mathbf{b}} \;=\; \frac{1}{2} \frac{1}{\mathbf{b} \cdot A^{-1}\mathbf{b}}. \tag{4.54}$$

4.2 Minimisation avec une contrainte d'inégalité

Nous nous intéressons à la résolution du problème suivant

$$\left\{ \begin{aligned} &\text{Chercher } \mathbf{x} \text{ argument-minimum de } J : \mathbb{R}^n \longrightarrow \mathbb{R} \text{ sur} \\ &\qquad C \;=\; \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; f(\mathbf{x}) \leq 0 \right\}. \end{aligned} \right. \tag{4.55}$$

On dit que C décrit une contrainte d'inégalité.

Théorème 4.2 *Soit* $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ *une fonctionnelle de classe* $\mathcal{C}^1$. *Soit* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *une fonctionnelle de classe* $\mathcal{C}^1$. *Soit* $C = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leqslant 0\}$.

Si $\mathbf{x}$ *est un argument-minimum de* J *sur* C *alors* $\mathbf{x}$ *vérifie soit*

$$\exists \lambda \in \mathbb{R} \quad | \quad \nabla_{\mathbf{x},\lambda} \mathcal{L}(\mathbf{x}, \lambda) \ = \ 0 \quad \textit{avec} \quad \mathcal{L}(\mathbf{x}, \lambda) = J(\mathbf{x}) + \lambda f(\mathbf{x}) \tag{4.56}$$

soit

$$f(\mathbf{x}) = 0 \quad \textit{et} \quad \nabla_{\mathbf{x}} f(\mathbf{x}) = 0 \tag{4.57}$$

soit

$$\nabla J(\mathbf{x}) = 0 \textit{ et } f(\mathbf{x}) < 0. \tag{4.58}$$

Mode d'emplois.

1. On montre l'existence d'un argument-minimum de J sur C.

2. On recherche l'ensemble

$$E = \Big\{ \mathbf{x} \in \mathbb{R}^n \text{ vérifiant (4.56) ou (4.57) ou (4.58)} \Big\}. \tag{4.59}$$

3. On minimise J sur E. Les arguments-minima de J sur E sont les arguments-minima de J sur C.

Preuve. Remarquons que C admet la décomposition suivante

$$C = C_{=} \cup C_{<} \tag{4.60}$$

avec

$$C_{=} = \Big\{ \mathbf{x} \in \mathbb{R}^n \mid f(x) = 0 \Big\} \quad \text{et} \quad C_{<} = \Big\{ \mathbf{x} \in \mathbb{R}^n \mid f(x) < 0 \Big\}. \tag{4.61}$$

Ainsi, les arguments-minima de J sur C sont des arguments-minima de J sur $C_{=}$ ou des arguments-minima de J sur $C_{<}$.

D'après le théorème 4.1, les argument-minima de J sur $C_{=}$ vérifient (4.56) ou (4.57).

Pour conclure, il nous suffit de montrer que les arguments-minima de J sur $C_{<}$ vérifient (4.58).

(i) L'ensemble $C_{<}$ est ouvert. En effet, f est continue, $]-\infty, 0[$ est ouvert et

$$C_{<} \ = \ \overset{-1}{f}(]-\infty, 0[). \tag{4.62}$$

(ii) Comme $C_{<}$ est ouvert et J est de classe C^1 tout $\mathbf{x}$ argument-minimum de J sur $C_{<}$ vérifie d'après le théorème 3.2

$$\nabla J(\mathbf{x}) = 0. \tag{4.63}$$

(iii) Bien entendu, l'ensemble des $\mathbf{x}$ dans $C_<$ vérifient

$$f(\mathbf{x}) < 0. \tag{4.64}$$

Exercice. Soit $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{b} \neq 0$. Soit A une matrice symétrique définie positive. Caractériser les arguments-minima de

$$J : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad J(\mathbf{x}) \;=\; \mathbf{b} \cdot \mathbf{x} \tag{4.65}$$

sur $C = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot (A\mathbf{x}) \leq 1\}$.

Exercice. Soit $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{b} \neq 0$.

Caractériser les arguments-minima de

$$J : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad J(\mathbf{x}) \;=\; \|\mathbf{x}\|^2 \tag{4.66}$$

sur $C = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{b} \cdot \mathbf{x} \geq 1\}$.