

République Algérienne Démocratique et Populaire.  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique.



جامعة غليزان  
RELIZANE UNIVERSITY

# Optimisation

Cours et exercices

*Présenté par :*

***Dr. DJILALI Medjahed***

2021-2022

# Table des matières

Introduction Générale	1
<b>I Optimisation sans contraintes</b>	<b>4</b>
<b>1 Concepts et Considérations Théoriques</b>	<b>5</b>
1.1 Introduction . . . . .	5
1.2 Vecteurs et matrices . . . . .	5
1.2.1 Matrices et vecteurs partitionnés . . . . .	6
1.3 Discussion générale sur les solutions d'un système linéaire . . .	7
1.4 Propriétés des formes quadratiques . . . . .	8
1.4.1 Gradient d'une forme quadratique . . . . .	9
1.4.2 Forme quadratique définie et semi-définie positive . . .	10
1.4.3 Critère de Sylvester pour les formes quadratiques définies et semi-définies . . . . .	11
1.4.4 Propriétés des matrices définies et semi-définies positives	12
1.5 Éléments d'analyse convexe . . . . .	13
1.5.1 Ensembles convexes . . . . .	13
1.5.2 Fonctions convexes . . . . .	15
1.6 Semi-continuité inférieure . . . . .	17
1.7 Sur les polyèdres et les polytopes . . . . .	18
1.8 Exercices . . . . .	20
<b>2 Optimisation non linéaire sans contraintes</b>	<b>22</b>
2.1 Formulation mathématique d'un problème d'optimisation . . .	22
2.2 Minima locaux et globaux . . . . .	24
2.3 Théorèmes généraux d'existence . . . . .	26
2.4 Caractérisation des solutions optimales . . . . .	28

---

2.4.1	Conditions nécessaires d'optimalité . . . . .	28
2.4.2	Conditions suffisantes d'optimalité . . . . .	30
2.5	Étude de quelques exemples . . . . .	31
2.6	Exercices . . . . .	33
<b>3</b>	<b>Les méthodes numériques pour l'optimisation sans contraintes</b>	<b>37</b>
3.1	Introduction . . . . .	37
3.2	La méthode du gradient . . . . .	38
3.2.1	Interprétation physique du gradient . . . . .	38
3.2.2	Sélection des directions de descente . . . . .	39
3.2.3	Sélection du pas . . . . .	40
3.3	La méthode de Newton . . . . .	44
3.4	La méthode du gradient conjugué . . . . .	48
3.4.1	Cas linéaire . . . . .	48
3.4.2	Cas non linéaire . . . . .	51
3.5	La méthode de relaxation . . . . .	52
3.6	Exercices . . . . .	53

# Introduction Générale

*Euler : "Rien ne se passe dans le monde qui ne soit la signification d'un certain maximum ou d'un certain minimum."*

L'optimisation (ou programmation mathématique) apparaît comme l'une des branches des mathématiques les plus adaptées au développement d'outils pour l'ingénieur. L'optimisation c'est au moins trois choses :

- l'art de formuler des problèmes de décision grâce à des concepts et outils mathématiques précis ;
- une théorie mathématique ;
- une "cuisine" algorithmique. Une fois que le mathématicien a su caractériser une solution, et d'abord se prononcer sur son existence voire son unicité, l'ingénieur voudrait pouvoir calculer cette solution.

En 1949, Dantzig a proposé le terme "programmation linéaire" pour l'étude des problèmes théoriques et algorithmiques liés à l'optimisation de fonctions linéaires sous des contraintes linéaires.

Kuhn et Tucker proposent, en 1951, le terme de "programmation non linéaire" pour l'étude des problèmes d'optimisation non linéaire avec ou sans contraintes. La programmation dynamique est employée par R. Bellman en 1957 pour une méthode générale d'optimisation des systèmes dynamiques, c-à-d, évoluant au cours du temps.

La programmation en nombres entiers est suggérée par Gomory en 1958 pour les problèmes d'optimisation où les variables sont astreintes à ne prendre que des valeurs entières.

Cependant, malgré l'apparente diversité des thèmes abordés en les années 1945 et 1960, la prise de conscience progressive d'une affinité profonde, tant du point des structures que des méthodes, entre les différentes classes de problèmes amène rapidement à les intégrer au sein d'une nouvelle discipline plus vaste, la programmation mathématique, dont le terme apparaît officiellement en 1959 dans un symposium.

La programmation mathématique est aujourd'hui une branche particulièrement active des mathématiques appliquées et il y-a, à cela, de nombreuses raisons. La première est peut-être le nombre, la variété et l'importance de ses applications que ce soit dans les sciences de l'ingénieur ou dans d'autres domaines des mathématiques appliquées. Sans prétendre être exhaustif, on peut citer :

- En recherche opérationnelle : optimisation des systèmes technico-économiques, planifications, problèmes de transport, d'ordonnancement, de gestion des stocks, etc ...
- En analyse numérique : approximation, régression, résolution des systèmes linéaires et non linéaires, méthodes numériques liées à la mise en oeuvre des méthodes d'éléments finis, etc ...
- En automatique : identification des systèmes, commande optimale des systèmes, filtrage, ordonnancement d'atelier, commande de robots, etc ...
- En ingénierie : dimensionnement et optimisation des structures, conception optimale des systèmes techniques complexes tels que les systèmes informatiques, réseaux d'ordinateurs, réseaux de transport, de télécommunication, etc...
- En économie mathématique : résolution de grandes modèles macro-économiques, modèles d'entreprise, théorie de la décision et théorie des jeux.

Mais l'importance de la programmation mathématique vient aussi du fait qu'elle fournit un cadre conceptuel adéquat pour l'analyse et la résolution de nombreux problèmes des mathématiques appliquées.

Ce polycopié est destiné particulièrement aux étudiants de licence (L3) et de master en mathématiques.

Il regroupe les deux volets de l'optimisation, à savoir optimisation sans contraintes et optimisation avec contraintes. C'est un support de cours riche d'exercices et d'exemples numériques. Il est composé de 5 chapitres répartis en deux parties, l'une concerne l'optimisation sans contraintes et la seconde concerne l'optimisation avec contraintes. Dans le premier chapitre, on rappelle brièvement quelques notions mathématiques utiles pour la suite du cours. Les chapitres 2 et 3 sont consacrés à l'optimisation sans contraintes ; dans le premier, on présentera les conditions d'optimalité, les conditions d'existence et d'unicité, dans le cas d'un problème non linéaire sans contraintes. Puis, on développera les algorithmes les plus utilisés pour résoudre ce type de problèmes. Quand aux quatrième et cinquième chapitres, ils sont consacrés à la théorie et aux algorithmes de résolution d'un problème non linéaire soumis

à des contraintes.

Chaque chapitre est clôturé par un ensemble d'exercices. Les divers exercices accompagnent le document afin d'assimiler les notions plus théoriques vues en cours. Les solutions de certains de ces exercices feront l'objet d'un prochain polycopié.

# Première partie

## Optimisation sans contraintes

# Chapitre 1

## Concepts et Considérations Théoriques

### 1.1 Introduction

Dans ce chapitre, on donne quelques rappels sur l'algèbre linéaire et la programmation mathématique. On rappelle tout d'abord les propriétés essentielles des formes quadratiques, ainsi que la notion des ensembles et fonctions convexes.

### 1.2 Vecteurs et matrices

**Définition 1.2.1.** Soit  $n, m \in \mathbb{N}^*$ . Une matrice d'ordre  $m \times n$  à coefficients dans  $\mathbb{R}$  est un tableau à deux dimensions, ayant  $m$  lignes et  $n$  colonnes, représenté sous la forme suivante :

$$A = A(I, J) = (a_{ij}, i \in I, j \in J) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

où  $I = \{1, 2, \dots, m\}$  et  $J = \{1, 2, \dots, n\}$  représentent respectivement l'ensemble des indices des lignes et des colonnes de  $A$ . Pour des calculs pratiques,

la matrice  $A$  se note aussi

$$A = (a_1, a_2, \dots, a_j, \dots, a_n) = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_m \end{pmatrix},$$

où  $a_j = A(I, j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$  est un vecteur colonne de dimension  $m$  ;

$A_i = A(i, J) = (a_{i1}, a_{i2}, \dots, a_{in})$  est un vecteur ligne de dimension  $n$ .

Chaque vecteur, noté  $x = x(J) = (x_j, j \in J)$ , sera ainsi considéré comme un vecteur-colonne tandis que le vecteur-ligne sera noté  $x^T$ . La matrice transposée de  $A$  sera notée

$$A^T = A^T(J, I) = (a_{ji}, j \in J, i \in I).$$

Notons qu'un vecteur-colonne de dimension  $n$  peut être considéré comme une matrice d'ordre  $(n \times 1)$ , tandis qu'un vecteur ligne de dimension  $n$  peut être considéré comme une matrice d'ordre  $(1 \times n)$ .

La matrice  $A$  est dite carrée si on a  $m = n$  ; de plus, si  $A = A^T$ , la matrice est dite symétrique.

La matrice identité d'ordre  $n$  sera notée  $I_n$ .

### 1.2.1 Matrices et vecteurs partitionnés

On peut effectuer le produit d'une matrice  $A$  et d'un vecteur  $x$ , après les avoir partitionnés judicieusement. On dit alors qu'on a effectué le produit par blocs. En effet, si l'on a

$$A = [A_1|A_2], x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

alors on peut écrire :

$$Ax = [A_1|A_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A_1x_1 + A_2x_2.$$

### 1.3 Discussion générale sur les solutions d'un système linéaire 7

De même pour

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

l'équation  $Ax = b$  peut alors s'écrire :

$$\begin{cases} A_{11}x_1 + A_{12}x_2 = b_1, \\ A_{21}x_1 + A_{22}x_2 = b_2. \end{cases}$$

On peut partitionner une matrice d'une manière arbitraire. Par exemple, si  $A = A(I, J)$  est une matrice d'ordre  $(m \times n)$  et que  $J_B$  et  $J_N$  sont deux sous-ensembles quelconques de  $J$ , tels que

$$|J_B| = m, \quad J_B \cup J_N = J, \quad J_B \cap J_N = \emptyset,$$

alors on peut partitionner  $A$  de la façon suivante :

$$A = (a_1, a_2, \dots, a_j, \dots, a_n) = [A_B | A_N],$$

avec  $A_B = A(I, J_B)$ ,  $A_N = A(I, J_N)$ .

Si  $x = x(J) = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$ ,  $x_B = x(J_B)$ ,  $x_N = x(J_N)$ , alors on peut écrire

$$\begin{aligned} Ax &= \sum_{j=1}^n a_j x_j = \sum_{j \in J_B} a_j x_j + \sum_{j \in J_N} a_j x_j \\ &= A(I, J_B)x(J_B) + A(I, J_N)x(J_N) \\ &= A_B x_B + A_N x_N. \end{aligned}$$

### 1.3 Discussion générale sur les solutions d'un système linéaire

Soient  $m$  et  $n$  deux nombres entiers. Un système de  $m$  équations linéaires à  $n$  inconnues  $x_1, x_2, \dots, x_n$  s'écrit comme suit :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \vdots + \vdots + \vdots + \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \quad (1.1)$$

où les coefficients  $a_{ij}$  sont des réels. Les nombres  $b_1, b_2, \dots, b_m$  sont appelés les membres libres du système (1.1) ou les seconds membres. En posant

$$A = (a_{ij}, 1 \leq i \leq m, 1 \leq j \leq n), \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

Le système (1.1) peut s'écrire sous la forme matricielle suivante :

$$Ax = b. \tag{1.2}$$

Tout vecteur  $x$  vérifiant les équations (1.1) s'appelle solution du système. Le système (1.1) est dit compatible s'il possède une ou plusieurs solutions. Dans le cas contraire, il est dit incompatible ou impossible. Lorsque le vecteur  $b$  est nul, le système (1.2) est dit homogène. Tout système homogène possède la solution triviale  $x = 0$ .

**Définition 1.3.1.** Le système linéaire (1.2) est dit de rang complet en lignes si  $\text{rang}(A) = m, m \leq n$ , et de rang complet en colonnes si  $\text{rang}(A) = n, m \geq n$ .

**Lemme 1.3.1.** Soit  $m \leq n$  et  $\text{rang}(A) = m$ . Alors le système  $Ax = b$  admet toujours des solutions, quelque soit le second membre  $b$  :

- a) une et une seule solution si  $m = n$ ,
- b) une infinité de solutions si  $m < n$ .

## 1.4 Propriétés des formes quadratiques

**Définition 1.4.1.** Une fonction  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ , est dite forme quadratique de  $n$  variables  $x_1, x_2, \dots, x_n$  si elle s'écrit sous la forme suivante :

$$F(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = x^T A x \tag{1.3}$$

où  $x^T = (x_1, x_2, \dots, x_n)$  est un  $n$ -vecteur ligne et  $A = (a_{ij}, 1 \leq i, j \leq n)$  est une matrice carrée d'ordre  $n$ .

Pour  $i \neq j$ , le coefficient du terme  $x_i x_j$  s'écrit  $a_{ij} + a_{ji}$ . En vertu de cela, la matrice  $A$  peut être supposée symétrique. En effet, en définissant de nouveaux coefficients

$$d_{ij} = d_{ji}, \quad 1 \leq i, j \leq n.$$

On obtient une nouvelle matrice  $D$  symétrique telle que

$$D = (d_{ij}, 1 \leq i, j \leq n), \text{ avec } d_{ij} = d_{ji} = \frac{a_{ij} + a_{ji}}{2}.$$

Il est clair qu'après une redéfinition des coefficients, la valeur de la forme quadratique  $F(x)$  reste inchangée pour tout point  $x \in \mathbb{R}^n$  :

$$F(x) = x^T A x = x^T D x.$$

Pour cela, il est naturel de considérer que la matrice d'une forme quadratique est toujours symétrique.

### 1.4.1 Gradient d'une forme quadratique

**Définition 1.4.2.** Soit  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction réelle continûment différentiable. Son gradient au point  $x$  est défini par :

$$\nabla F(x) = \begin{pmatrix} \frac{\partial F}{\partial x_1}(x) \\ \frac{\partial F}{\partial x_2}(x) \\ \vdots \\ \frac{\partial F}{\partial x_n}(x) \end{pmatrix} \quad (1.4)$$

Soit  $F$  une forme quadratique et  $D$  sa matrice symétrique associée :

$$F(x) = x^T D x. \quad (1.5)$$

En écrivant la matrice  $D$  sous forme de vecteurs colonnes

$$D = (d_1, d_2, \dots, d_n),$$

l'expression (1.5) peut se mettre sous la forme suivante :

$$F(x) = (x_1, x_2, \dots, x_j, \dots, x_n) \begin{pmatrix} d_1^T x \\ d_2^T x \\ \vdots \\ d_j^T x \\ \vdots \\ d_n^T x \end{pmatrix} = \sum_{j=1}^n x_j d_j^T x.$$

La dérivée partielle de  $F$  par rapport à chaque variable  $x_j$  est donnée par :

$$\begin{aligned}\frac{\partial F}{\partial x_j}(x) &= x_1 d_{1j} + \cdots + x_{j-1} d_{(j-1)(j)} + d_j^T x + x_j d_{jj} + \cdots + x_n d_{nj} \\ &= x_1 d_{1j} + \cdots + x_{j-1} d_{(j-1)(j)} + x_j d_{jj} + \cdots + x_n d_{nj} + d_j^T x \\ &= 2d_j^T x\end{aligned}$$

Par conséquent, le gradient de  $F(x)$  est :

$$\nabla F(x) = 2Dx. \quad (1.6)$$

**Définition 1.4.3.** Soit une fonction réelle de classe  $\mathcal{C}^2$ ,  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ . Le Hessian de la fonction  $F$  est défini par :

$$\begin{aligned}\nabla^2 F(x) &= \left( \nabla \frac{\partial F}{\partial x_1}(x), \nabla \frac{\partial F}{\partial x_2}(x), \dots, \nabla \frac{\partial F}{\partial x_j}(x), \dots, \nabla \frac{\partial F}{\partial x_n}(x) \right) \\ &= \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(x) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 F}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1}(x) & \frac{\partial^2 F}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 F}{\partial x_n^2}(x) \end{pmatrix}. \quad (1.7)\end{aligned}$$

**Définition 1.4.4.** Soit  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$ . La dérivée directionnelle de  $F$  dans la direction  $d$  au point  $x$  est :

$$\begin{aligned}F'(x; d) &= \lim_{t \rightarrow 0^+} \frac{F(x + td) - F(x)}{t} \\ &= \frac{\partial F}{\partial x_1}(x + td)|_{t=0} d_1 + \cdots + \frac{\partial F}{\partial x_n}(x + td)|_{t=0} d_n \\ &= (\nabla F(x))^T d.\end{aligned}$$

### 1.4.2 Forme quadratique définie et semi-définie positive

Soit  $F(x) = x^T D x$  une forme quadratique avec  $D$  symétrique.

**Définition 1.4.5.**  $F$  est dite définie positive si  $x^T D x > 0, \forall x \in \mathbb{R}^n$  et  $x \neq 0$ . elle est dite semi-définie positive ou définie non négative si  $x^T D x \geq 0, \forall x \in \mathbb{R}^n$ .

**Définition 1.4.6.**  $F$  est dite définie négative si  $x^T D x < 0, \forall x \in \mathbb{R}^n$  et  $x \neq 0$ . Elle est dite semi-définie négative ou définie non positive si  $x^T D x \leq 0, \forall x \in \mathbb{R}^n$ .

**Définition 1.4.7.** Une matrice symétrique  $D$  est dite matrice définie positive (respectivement non négative) et on note  $D \succ 0$  (respectivement  $D \succeq 0$ ) si elle est associée à une forme quadratique définie positive (respectivement non négative).

### 1.4.3 Critère de Sylvester pour les formes quadratiques définies et semi-définies

L'intérêt du critère du Sylvester est de caractériser une forme quadratique définie ou semi-définie. Pour cela, considérons la matrice symétrique suivante :

$$D = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{pmatrix}$$

Le mineur de la matrice  $D$ , formé des lignes  $i_1, i_2, \dots, i_p$  et les colonnes  $j_1, j_2, \dots, j_p$ , sera noté comme suit :

$$D \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} = \begin{vmatrix} d_{i_1 j_1} & d_{i_1 j_2} & \cdots & d_{i_1 j_p} \\ d_{i_2 j_1} & d_{i_2 j_2} & \cdots & d_{i_2 j_p} \\ \vdots & \vdots & & \vdots \\ d_{i_p j_1} & d_{i_p j_2} & \cdots & d_{i_p j_p} \end{vmatrix}.$$

Ce mineur est dit principal si :  $i_1 = j_1, i_2 = j_2, \dots, i_p = j_p$ , c'est-à-dire s'il est formé des lignes et des colonnes portant les mêmes numéros.

Les mineurs suivants :

$$D_1 = d_{11}, D_2 = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix}, \dots, D_n = \begin{vmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix},$$

sont appelés mineurs principaux successifs. Alors, le critère de Sylvester se formule comme suit :

**Théorème 1.4.1** (Critère de Sylvester). (i) Pour que la matrice  $D$  soit définie positive ( $D \succ 0$ ), il est nécessaire et suffisant que tous ses mineurs principaux successifs soient positifs :

$$D_1 > 0, D_2 > 0, \dots, D_n > 0; \quad (1.8)$$

(ii) Pour que la matrice  $D$  soit semi-définie positive ( $D \succeq 0$ ), il est nécessaire et suffisant que tous ses mineurs principaux soient non négatifs :

$$D \begin{pmatrix} i_1, i_2, \dots, i_p \\ i_1, i_2, \dots, i_p \end{pmatrix} \geq 0, \quad 1 \leq i_1 < i_2 < \dots < i_p \leq n, \quad p = 1, 2, \dots, n. \quad (1.9)$$

*Remarque 1.4.1.* Pour qu'une matrice  $D$  soit définie négative ( $D \prec 0$ ) ou non positive ( $D \preceq 0$ ), les conditions (1.8) et (1.9) se reformulent ainsi :

$$(i) \quad D \prec 0 \Leftrightarrow (-1)^p D_p > 0, \quad p = 1, 2, \dots, n.$$

$$(ii) \quad D \preceq 0 \Leftrightarrow (-1)^p D \begin{pmatrix} i_1, i_2, \dots, i_p \\ i_1, i_2, \dots, i_p \end{pmatrix} \geq 0, \quad 1 \leq i_1 < i_2 < \dots < i_p \leq n, \quad p = 1, 2, \dots, n.$$

*Remarque 1.4.2.* Le critère de Sylvester n'est valable que pour les matrices symétriques.

#### 1.4.4 Propriétés des matrices définies et semi-définies positives

Les matrices symétriques définies ont des propriétés très intéressantes. En voici quelques unes :

**Propriété 1.4.1.** Soit la matrice  $D$  partitionnée de la manière suivante :

$$D = \begin{pmatrix} m & k \\ D_{11} & D_{21} \\ D_{21} & D_{22} \end{pmatrix} \begin{matrix} m \\ k \end{matrix}, \quad m + k = n.$$

Si  $D > 0$  ( $D \geq 0$ ), alors les sous-matrices principales  $D_{11}$  et  $D_{22}$  sont définies positives (non négatives). D'une manière générale, toute sous-matrice principale d'une matrice définie positive (non négative) est aussi définie positive (non négative).

**Propriété 1.4.2.** Un élément de la diagonale d'une matrice symétrique  $D$  définie non négative ne peut s'annuler que si les autres éléments de la même ligne et colonne s'annulent aussi.

**Propriété 1.4.3.** Soit  $D$  une matrice symétrique définie non-négative. Si  $x \in \mathbb{R}^n$  est un point quelconque fixe tel que  $x^T D x = 0$ , alors on aura  $D x = 0$ .

## 1.5 Éléments d'analyse convexe

Dans les problèmes d'optimisation, avec ou sans contraintes, une notion va jouer un rôle très important : celle de convexité. En effet, pour la plupart des algorithmes, la convergence vers un optimum global ne pourra être démontrée qu'avec des hypothèses de convexité.

### 1.5.1 Ensembles convexes

**Définition 1.5.1.** Un ensemble  $S \subset \mathbb{R}^n$  est dit convexe si et seulement si :

$$\forall x, y \in S, \forall \lambda \in [0, 1], \text{ on a } \lambda x + (1 - \lambda)y \in S. \quad (1.10)$$

Cette définition peut s'interpréter en disant que  $S$  est convexe si et seulement si pour deux points quelconques  $x$  et  $y$  pris dans  $S$ , le segment  $[x, y]$  tout entier est contenu dans  $S$ .

**Définition 1.5.2.** Soit  $S = \{p_1, p_2, \dots, p_k\}$  un ensemble de  $k$  points de  $\mathbb{R}^n$ .

- Un point  $p \in \mathbb{R}^n$  peut être obtenu par :
  - Combinaison linéaire des points de  $S$  si

$$p = \sum_{i=1}^k \lambda_i p_i, \quad \lambda_i \in \mathbb{R}, \quad i = \overline{1, k}.$$

- Combinaison affine des points de  $S$  si

$$p = \sum_{i=1}^k \lambda_i p_i, \quad \sum_{i=1}^k \lambda_i = 0.$$

- Combinaison conique des points de  $S$  si

$$p = \sum_{i=1}^k \lambda_i p_i, \quad \lambda_i \geq 0 \text{ pour } i = \overline{1, k}.$$

– Combinaison convexe des points de  $S$  si

$$p = \sum_{i=1}^k \lambda_i p_i, \quad \sum_{i=1}^k \lambda_i = 1, \quad \lambda_i \geq 0, \quad i = \overline{1, k}.$$

- On définit l'enveloppe conique (respectivement convexe) des points de  $S$  comme l'ensemble des points de  $\mathbb{R}^n$  pouvant être obtenus par combinaison conique (respectivement convexe) des points de  $S$ .
- L'espace affine (respectivement espace vectoriel) engendré par les points de  $S$  est l'ensemble des points de  $\mathbb{R}^n$  pouvant être obtenus par combinaison affine (respectivement linéaire) des points de  $S$ .
- Les points de  $S$  sont dits affinement indépendants (respectivement linéairement indépendants) si aucun des points de  $S$  ne peut être obtenu par une combinaison affine (respectivement linéaire) des autres points de  $S$ .

**Propriété 1.5.1.** Un ensemble  $S$  de points de  $\mathbb{R}^n$  est convexe si et seulement si toute combinaison convexe de points de  $S$  est encore un point de  $S$ .

*Exemple 1.5.1.* L'ensemble  $S = \{x \in \mathbb{R}^n, x = x_0 + \alpha d, \alpha \geq 0\}$  est convexe, où  $d \in \mathbb{R}^n$  est un vecteur non nul, et  $x_0 \in \mathbb{R}^n$  est un point fixe.

En effet, pour tout  $x_1, x_2 \in S$  et pour tout  $\lambda \in [0, 1]$ , on a

$$x_1 = x_0 + \alpha_1 d, \quad x_2 = x_0 + \alpha_2 d,$$

où  $\alpha_1, \alpha_2 \in \mathbb{R}^+$ . Donc

$$\begin{aligned} \lambda x_1 + (1 - \lambda)x_2 &= \lambda(x_0 + \alpha_1 d) + (1 - \lambda)(x_0 + \alpha_2 d) \\ &= x_0 + [\lambda\alpha_1 + (1 - \lambda)\alpha_2]d. \end{aligned}$$

Comme  $\lambda\alpha_1 + (1 - \lambda)\alpha_2 \geq 0$ , alors  $\lambda x_1 + (1 - \lambda)x_2 \in S$ .

**Théorème 1.5.1.** Soient  $S_1$  et  $S_2$  deux ensembles convexes de  $\mathbb{R}^n$ . Alors

1.  $S_1 \cap S_2$  est convexe ;
2.  $S_1 \pm S_2 = \{x_1 \pm x_2 \mid x_1 \in S_1, x_2 \in S_2\}$  est convexe.

**Théorème 1.5.2.** Soit  $S \subset \mathbb{R}^n$  un ensemble convexe. Alors

1. l'intérieur de  $S$ , noté  $\text{Int}S$ , est un ensemble convexe ;
2. la clôture  $\overline{S}$  de  $S$  est un ensemble convexe.

### 1.5.2 Fonctions convexes

**Définition 1.5.3.** On dit qu'une fonction  $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , définie sur un ensemble convexe  $S$ , est convexe, si elle vérifie :

$$\forall x \in S, \forall y \in S, \forall \lambda \in [0, 1], \text{ on a : } f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Une fonction  $f$  est dite strictement convexe si l'inégalité ci-dessus est stricte pour  $x \neq y$  et  $\lambda \in ]0, 1[$ .

Une fonction  $f$  est dite concave si  $(-f)$  est convexe.

L'interprétation géométrique de cette définition est que le graphe d'une fonction convexe est toujours en dessous du segment reliant les points  $(x, f(x))$  et  $(y, f(y))$ .

*Exemple 1.5.2.* Un des exemples basiques de fonctions convexes est la fonction indicatrice. Soit  $S \subset \mathbb{R}^n$  un sous-ensemble non vide; la fonction indicatrice  $I_S : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  est définie par

$$I_S(x) = \begin{cases} 0, & \text{si } x \in S, \\ +\infty, & \text{ailleurs.} \end{cases}$$

La fonction indicatrice  $I_S$  est convexe si et seulement si  $S$  est convexe.

*Exemple 1.5.3.* On définit la fonction distance :

$$d_S(x) = \inf\{\|y - x\| \mid y \in S\},$$

où  $S \subset \mathbb{R}^n$  est un ensemble convexe non vide et  $\|\cdot\|$  est une norme quelconque dans  $\mathbb{R}^n$ . Alors  $d_S$  est une fonction convexe.

Une fonction convexe peut aussi être décrite par son épigraphe.

**Définition 1.5.4.** Soit  $S \subset \mathbb{R}^n$  un ensemble non vide et  $f : S \rightarrow \mathbb{R}$ . L'ensemble

$$\{(x, f(x)) : x \in S\} \subset \mathbb{R}^{n+1}$$

décrivant la fonction  $f$  est appelé le graphe de la fonction  $f$ .

**Définition 1.5.5.** Soit  $S \subset \mathbb{R}^n$  un ensemble non vide. Soit  $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . L'épigraphe de  $f$ , noté  $epif$ , est un sous-ensemble de  $\mathbb{R}^{n+1}$  défini par

$$epif = \{(x, \alpha) \mid f(x) \leq \alpha, x \in S, \alpha \in \mathbb{R}\}.$$

L'hypographe de  $f$ , noté  $hypf$ , est un sous-ensemble de  $\mathbb{R}^{n+1}$  défini par

$$hypf = \{(x, \alpha) \mid f(x) \geq \alpha, x \in S, \alpha \in \mathbb{R}\}.$$

Les théorèmes suivants indiquent la relation entre une fonction convexe et la convexité de  $epif$ .

**Théorème 1.5.3.** *Soient  $S \subset \mathbb{R}^n$  un ensemble convexe non vide et  $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Alors,  $f$  est convexe si et seulement si  $epif$  est un ensemble convexe.*

*Démonstration.* Supposons que  $f$  est convexe. Soient  $x_1, x_2 \in S$  et  $(x_1, \alpha_1), (x_2, \alpha_2) \in epif$ . D'après les définitions de la fonction convexe et de l'épigraph, il s'ensuit

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda \alpha_1 + (1 - \lambda)\alpha_2$$

pour tout  $\lambda \in [0, 1]$ . Comme  $S$  est un ensemble convexe,  $\lambda x_1 + (1 - \lambda)x_2 \in S$ . D'où

$$(\lambda x_1 + (1 - \lambda)x_2, \lambda \alpha_1 + (1 - \lambda)\alpha_2) \in epif,$$

ce qui implique que  $epif$  est convexe.

Supposons, maintenant, que  $epif$  est convexe, et soient  $x_1, x_2 \in S$ . On a  $(x_1, f(x_1)), (x_2, f(x_2)) \in epif$ . D'après la convexité de  $epif$ , on a

$$(\lambda x_1 + (1 - \lambda)x_2, \lambda f(x_1) + (1 - \lambda)f(x_2)) \in epif, \text{ pour } \lambda \in [0, 1].$$

Donc

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

pour tout  $\lambda \in [0, 1]$ . D'où  $f$  est convexe.  $\square$

**Théorème 1.5.4.** 1. *Soient  $f$  une fonction convexe sur un ensemble convexe  $S \subset \mathbb{R}^n$  et un nombre réel  $\alpha \geq 0$ . Alors  $\alpha f$  est aussi une fonction convexe sur  $S$ .*

2. *Soient  $f_1, f_2$  deux fonctions convexes sur un ensemble convexe  $S$ . Alors  $f_1 + f_2$  est aussi une fonction convexe sur  $S$ .*

3. *Soient  $f_1, f_2, \dots, f_m$  des fonctions convexes sur un ensemble convexe  $S$  et  $\alpha_1, \alpha_2, \dots, \alpha_m \geq 0$ , des nombres réels. Alors  $\sum_{i=1}^m \alpha_i f_i$  est aussi une fonction convexe sur  $S$ .*

Dans le cas où  $f : S \subset \mathbb{R} \rightarrow \mathbb{R}$ , on a le résultat suivant :

**Propriété 1.5.2.** Si  $f : S \rightarrow \mathbb{R}$  est deux fois continûment dérivable sur  $S$  convexe, alors  $f$  est convexe si et seulement si  $f''(x) \geq 0, \forall x \in S$ . Elle est strictement convexe si  $f''(x) > 0, \forall x \in S$ .

**Théorème 1.5.5.** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Si  $f$  est continûment différentiable, les conditions a) et b) ci-dessous sont équivalentes ;

Si  $f$  est deux fois continûment différentiable, les conditions a), b) et c) ci-dessous sont équivalentes :

a)  $f$  est convexe ;

b)  $\forall x, \forall y : f(y) \geq f(x) + \nabla f^T(x)(y - x)$  ;

c)  $\forall x$ , le hessien  $\nabla^2 f(x)$  est une matrice semi-définie positive ( $y^T \nabla^2 f(x) y \geq 0, \forall y \in \mathbb{R}^n$ ).

**Corollaire 1.5.1.** Soit  $f$  une forme quadratique définie par

$$f(x) = \frac{1}{2}x^T A x - b^T x.$$

Alors  $f$  est convexe si et seulement si  $A \succeq 0$ , et strictement convexe si et seulement si  $A \succ 0$ .

Cela provient du fait que  $\nabla^2 f(x) = A$ .

**Définition 1.5.6.** On dit qu'un problème de programmation mathématique est convexe s'il consiste à minimiser une fonction convexe (respectivement maximiser une fonction concave) sur un domaine convexe.

## 1.6 Semi-continuité inférieure

De façon générale, une fonction convexe peut être très compliquée sur le bord de son domaine. Par exemple, la fonction

$$f(x, y) = \begin{cases} 0 & \text{si } x^2 + y^2 < 1, \\ \phi(x, y) & \text{si } x^2 + y^2 = 1, \end{cases} \quad \text{avec } \phi(x, y) \geq 0 \quad (1.11)$$

est convexe. Cependant, son comportement sur le bord peut être arbitrairement complexe. Minimiser de telles fonctions est en général impossible. Cette remarque nous mène à nous restreindre à la classe de fonctions semi-continues inférieurement.

**Définition 1.6.1.** Un ensemble  $S \subset \mathbb{R}^n$  est compact si et seulement s'il est fermé et borné.

**Définition 1.6.2** (Fonction semi-continue inférieurement). Une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  est dite fermée ou semi-continue inférieurement si :

$$\forall x \in \mathbb{R}^n, \forall x_n \xrightarrow[n \rightarrow +\infty]{} x, \liminf f(x_n) \geq f(x).$$

**Proposition 1.6.1.** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ . Les propositions suivantes sont équivalentes :

- i)  $f$  est semi-continue inférieurement.
- ii) Quel que soit  $\alpha \in \mathbb{R}$ , l'ensemble de niveau  $\{x \in \mathbb{R}^n | f(x) \leq \alpha\}$  est fermé.
- iii)  $\text{epi}(f)$  est un ensemble fermé.

**Définition 1.6.3.** La fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est coercive si et seulement si

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

## 1.7 Sur les polyèdres et les polytopes

On note par  $S$  l'ensemble suivant :

$$S = \{x \in \mathbb{R}^n | Ax \leq b\},$$

où  $b$  est un vecteur de  $\mathbb{R}^m$  et  $A$  une matrice à  $m$  lignes et  $n$  colonnes,  $m$  et  $n$  étant des nombres finis.

On a

$$Ax \leq b \Leftrightarrow \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq b_2, \\ \vdots + \vdots + \vdots + \vdots \leq \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m. \end{cases}$$

**Définition 1.7.1.** L'ensemble  $S$  obtenu par l'intersection d'un nombre fini de demi-espaces de  $\mathbb{R}^n$  (chacun de ces demi-espaces correspondant à une des lignes du système matriciel  $Ax \leq b$ ), constitue ce que l'on nomme un polyèdre.

**Définition 1.7.2.** Un polyèdre borné, i.e. un polyèdre pour lequel il existe un nombre  $B$  tel que chaque point du polyèdre a des coordonnées comprises entre  $-B$  et  $B$ , est un polytope.

**Théorème 1.7.1.** *Un polyèdre est un ensemble convexe.*

**Définition 1.7.3.** Le polyèdre  $S$  est de dimension  $k$ , noté  $\dim(S) = k$ , si le nombre de points de  $S$  affinement indépendants est  $(k + 1)$ . Le polyèdre  $S$  a pleine dimension si  $\dim(S) = n$ .

**Définition 1.7.4.** Un point  $x$  d'un polyèdre  $S$  est un point extrême de  $S$  si, dans toute expression de  $x$  comme combinaison convexe de deux points  $x_1$  et  $x_2$  de  $S$ , on a forcément  $x_1 = x_2 = x$ .

**Définition 1.7.5.** Soit  $S = \{x \in \mathbb{R}^n \mid Ax \leq b\}$  un polyèdre dans  $\mathbb{R}^n$ .

- Une inéquation  $a^T x \leq a_0$  est une inéquation valide pour  $S$  si elle est vérifiée par chacun des points de  $S$ .  
Si  $a^T x \leq a_0$  est une inéquation valide pour  $S$ , elle définit une face  $F$  de  $S$ , avec  $F := \{x \in \mathbb{R}^n \mid x \in S, a^T x = a_0\}$ .
- Une face  $F$  de  $S$  est une face propre de  $S$  si  $F \neq \emptyset$  et  $F \neq S$ .
- Une face  $F$  d'un polyèdre  $S$  est une facette de  $S$  si  $\dim(F) = \dim(S) - 1$ .

**Théorème 1.7.2.** *Un point  $x$  est un point extrême d'un polyèdre  $S$  si et seulement si  $x$  est une face de dimension 0 de  $S$ .*

Une autre manière d'énoncer le théorème 1.7.2 est :

Un point  $x^0$  est un point extrême d'un polyèdre  $S$  si et seulement s'il existe un vecteur  $c$  tel que  $x^0$  est l'unique solution optimale du problème suivant :

$$c^T x \longrightarrow \max, x \in S.$$

**Théorème 1.7.3.** *Un polyèdre a un nombre fini de faces.*

**Définition 1.7.6.** Un ensemble convexe  $C$  de points de  $\mathbb{R}^n$  est un cône si  $x \in C$  implique  $(\lambda x) \in C$  pour tout  $\lambda \geq 0$ .

Un cône est pointé s'il possède un point extrême.

**Définition 1.7.7.** Le cône de récession du polyèdre  $S$  est le polyèdre  $S_0 = \{x \in \mathbb{R}^n \mid Ax \leq 0\}$ .

**Propriété 1.7.1.** Si  $S$  est un polytope, alors  $S_0$  ne contient que le point 0, i.e.  $\text{rang}(A) = n$ .

**Définition 1.7.8.** Si  $S$  n'est pas vide, alors  $r \in S_0 \setminus \{0\}$  est un rayon de  $S$ . Un rayon  $r$  de  $S$  est un rayon extrême de  $S$  s'il n'existe pas deux rayons de  $S$ ,  $r_1$  et  $r_2$ , tels que  $r_1 \neq \lambda r_2$  pour tout  $\lambda \geq 0$ , avec  $r = \frac{1}{2}r_1 + \frac{1}{2}r_2$ .

Une autre manière de caractériser les rayons extrêmes d'un polyèdre est donnée par le théorème suivant :

**Théorème 1.7.4.** *Soit  $S = \{x \in \mathbb{R}^n \mid Ax \leq b\}$  un polyèdre dans  $\mathbb{R}^n$ . Si  $S$  n'est pas vide, le vecteur  $r$  est un rayon extrême de  $S$  si et seulement si  $\{\lambda r \mid \lambda \in \mathbb{R}, \lambda \geq 0\}$  est une face de dimension 1 de  $S_0$ .*

**Corollaire 1.7.1.** *L'ensemble (polytope ou polyèdre) convexe :*

$$S = \{x \in \mathbb{R}^n \mid Ax \leq b\}$$

*a un nombre fini  $\tau(S)$  de points extrêmes et  $\tau(S) \leq C_n^m$ .*

**Corollaire 1.7.2.** *Tout point d'un polytope  $S \subset \mathbb{R}^n$  est combinaison convexe des points extrêmes de  $S$ .*

**Corollaire 1.7.3.** *Tout point d'un polyèdre  $S \subset \mathbb{R}^n$  est combinaison convexe des points extrêmes de  $S$ , à laquelle s'ajoute éventuellement une combinaison linéaire, à coefficients positifs, de rayons extrémaux.*

## 1.8 Exercices

*Exercice 1.8.1.* Soit  $A \in \mathbb{M}_n(\mathbb{R})$  telle que  $A^T A = A A^T$ . On suppose qu'il existe  $p \in \mathbb{N}$  tel que  $A^p = 0$ .

- a) Montrer que  $A^T A = 0$ .
- b) En déduire que  $A = 0$ .

*Exercice 1.8.2.* Soit  $A \in \mathbb{M}_n(\mathbb{R})$ . Montrer que la matrice  $A^T A$  est diagonalisable à valeurs propres positives.

*Exercice 1.8.3.* – Soit  $A$  une matrice carrée réelle d'ordre  $n$  et  $S = A^T A$ .

Montrer que  $S$  est symétrique semi-définie positive.

- Réciproquement, montrer que pour toute matrice  $S$  symétrique positive, il existe une matrice  $A$  carrée réelle de format  $n$  telle que  $S = A^T A$ . A t-on l'unicité de  $A$  ?
- Montrer que  $S$  est définie positive si et seulement si  $A$  est inversible.
- Montrer que  $\text{rang}(A) = \text{rang}(S)$ .

*Exercice 1.8.4.* 1. Soit  $A \in \mathbb{R}^{m \times n}$  et  $b \in \mathbb{R}^m$ . Utiliser la définition de la convexité pour montrer que l'ensemble  $S = \{x \in \mathbb{R}^n : Ax \geq b\}$  est un ensemble convexe.

2. Pour chacune des propositions suivantes, dire si elle est juste ou fausse :
- L'ensemble  $S \subset \mathbb{R}^n$  est dit fermé quand tout point de  $S$  est un point frontière.
  - Si  $S \subset \mathbb{R}^n$  est ouvert alors  $S$  est convexe.

*Exercice 1.8.5.* Etudier la convexité des fonctions suivantes :

1.  $f_1(X_1, X_2) = 10X_1 + 10X_2 + 100 - X_1^2 - 2X_1X_2 - X_2^2$
2.  $f_2(X_1, X_2) = X_1^2 + 2X_1X_2 + 100X_1 + 100X_2$
3.  $f_3(X_1, X_2) = X_1e^{X_1+X_2}, X_1, X_2 \geq 0$ .
4.  $f(X_1, X_2, X_3) = 3X_1^2 + 2X_2^2 + X_3^2 - 2X_1X_2 - 2X_1X_3 + 2X_2X_3 - 6X_1 - 4X_2 - 2X_3$

*Exercice 1.8.6.* Etudier la nature des formes quadratiques suivantes :

1.  $z = X_1^2 + X_2^2$  ;
2.  $z = (X_1 + X_2)^2$  ;
3.  $z = X_1^2 - X_2^2$ .

*Exercice 1.8.7.* Calculer la jacobienne de la fonction  $f$  définie de  $\mathbb{R}^2$  dans  $\mathbb{R}^3$  et pour tout  $(x, y)$  de  $\mathbb{R}^2$ , on a :

$$f(x, y) = (e^x + y^3, y, 3 \cosh x + \sinh y)^T.$$

*Exercice 1.8.8.* Montrer que la fonction suivante est convexe :

$$f(x) = 3x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 + 2x_2x_3 - 6x_1 - 4x_2 - 2x_3.$$

*Exercice 1.8.9.* Calculer le gradient et le hessien de la fonction de Rosenbrock :

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

# Chapitre 2

## Optimisation non linéaire sans contraintes

### 2.1 Formulation mathématique d'un problème d'optimisation

Il existe différents problèmes d'optimisation. Certaines caractéristiques permettent de les distinguer : comportent-ils des contraintes ? les fonctions objectifs sont-elles linéaires ? sont-elles quadratiques ? sont-elles convexes ? les domaines de définition des fonctions sont-ils continus ou discrets ? le problème contient-il une seule ou plusieurs fonctions objectifs ? Tous ces problèmes possèdent des structures différentes et ne peuvent être traités de la même façon.

Un problème d'optimisation continue se présente habituellement sous la forme suivante :

$$\begin{cases} f(x) \longrightarrow \min, \\ x \in S. \end{cases} \quad (2.1)$$

**Définition 2.1.1.** Le vecteur  $x = (x_1, \dots, x_n)$  est appelé vecteur des variables de décision du problème ; la fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est la fonction objectif ;  $S$  est généralement défini par une collection de contraintes exprimées sous forme d'égalités et d'inégalités, du genre  $S = \{x \in \mathbb{R}^n, g(x) \leq 0\}$ , où  $g = (g_1, \dots, g_m)^T$  est une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$ .

**Définition 2.1.2.** Un vecteur  $x^0$  est dit solution optimale du problème (2.1) si pour tous les vecteurs  $x \in S$ , on a  $f(x) \geq f(x^0)$ .

## 2.1 Formulation mathématique d'un problème d'optimisation 23

**Définition 2.1.3.** Le problème d'optimisation (2.1) est dit programme linéaire si les fonctions  $f, g_1, \dots, g_m$  sont linéaires, i.e., elles vérifient

$$\begin{aligned}f(\alpha x + \beta y) &= \alpha f(x) + \beta f(y), \\g_i(\alpha x + \beta y) &= \alpha g_i(x) + \beta g_i(y), \quad i = \overline{1, m},\end{aligned}$$

pour tous  $x, y \in \mathbb{R}^n$  et tous  $\alpha, \beta \in \mathbb{R}$ .

Ici, on a

$$S = \{x \in \mathbb{R}^n, g_i(x) \leq 0, \quad i = 1, \dots, m\}.$$

Si l'une au moins de ces fonctions n'est pas linéaire, le problème est dit programme non-linéaire.

**Définition 2.1.4.** Le problème d'optimisation (2.1) est dit convexe si la fonction objectif  $f$  et les contraintes  $g_1, \dots, g_m$  sont convexes, i.e., elles vérifient

$$\begin{aligned}f(\alpha x + \beta y) &\leq \alpha f(x) + \beta f(y), \\g_i(\alpha x + \beta y) &\leq \alpha g_i(x) + \beta g_i(y), \quad i = \overline{1, m},\end{aligned}$$

pour tous  $x, y \in \mathbb{R}^n$  et tous  $\alpha, \beta \in \mathbb{R}$ , avec  $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$ .

**Conséquence 2.1.1.** *Un programme linéaire est alors un programme d'optimisation convexe.*

*On peut considérer l'optimisation convexe comme une généralisation de la programmation linéaire.*

Dans ce chapitre, nous allons étudier le problème de minimisation d'une fonction non linéaire  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , qu'on notera

$$\min_{x \in \mathbb{R}^n} f(x). \tag{2.2}$$

*Exemple 2.1.1.* On voudrait installer une antenne pour connecter 4 clients importants. Cette antenne doit se trouver au plus proche de chaque client, en donnant priorité aux meilleurs clients.

Pour chaque client, on connaît :

- sa localisation : les coordonnées  $(x, y)$  ;
- le nombre d'heures de communication par mois.

Le problème consiste à trouver  $(x, y)$  tel que

$$\begin{aligned}\min & 200\sqrt{(x-5)^2 + (y-10)^2} + 150\sqrt{(x-10)^2 + (y-5)^2} \\ & + 200\sqrt{x^2 + (y-10)^2} + 300\sqrt{(x-12)^2 + y^2}\end{aligned}$$

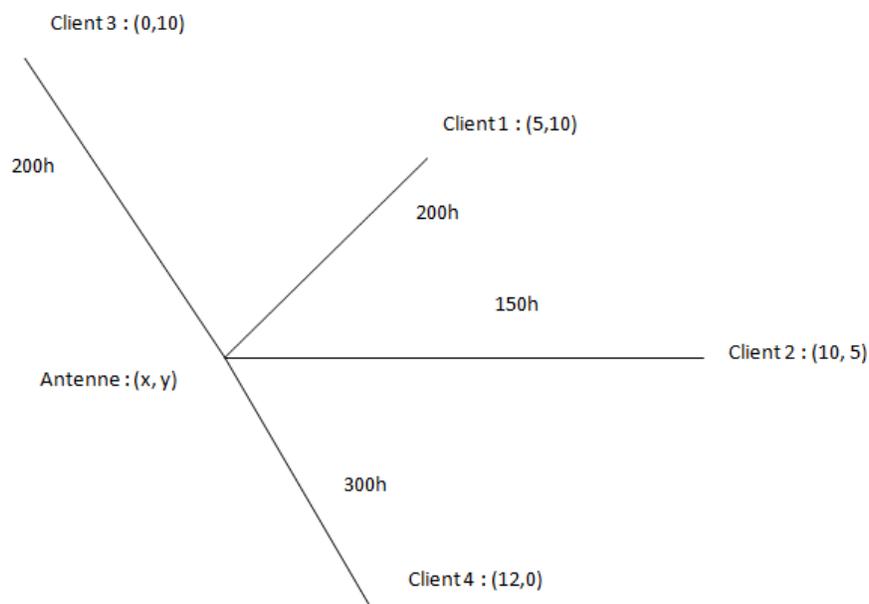


FIGURE 2.1 – Exemple

## 2.2 Minima locaux et globaux

Les minima locaux et globaux de  $f$  sur  $\mathbb{R}^n$  sont définis de la manière suivante :

**Définition 2.2.1.** Un vecteur  $x^0 \in \mathbb{R}^n$  est un *minimum local* de  $f$  sur  $\mathbb{R}^n$  si

$$\exists \epsilon > 0 \quad \text{tel que} \quad f(x^0) \leq f(x), \quad \forall x \in \mathbb{R}^n, \quad \text{avec} \quad \|x - x^0\| < \epsilon.$$

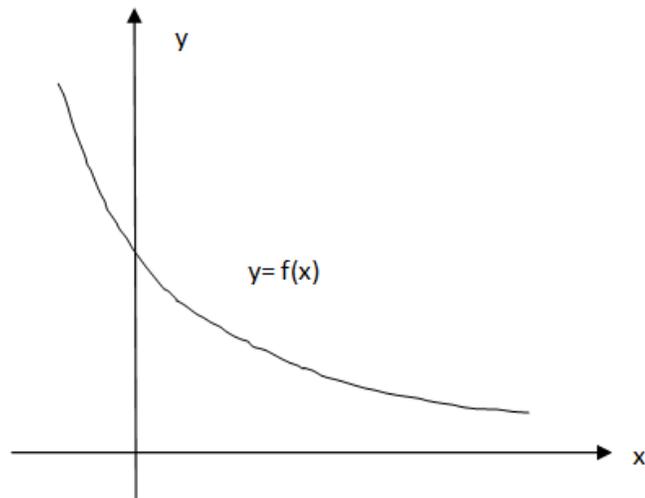
**Définition 2.2.2.** Un vecteur  $x^0 \in \mathbb{R}^n$  est un *minimum global* de  $f$  sur  $\mathbb{R}^n$  si

$$f(x^0) \leq f(x), \quad \forall x \in \mathbb{R}^n.$$

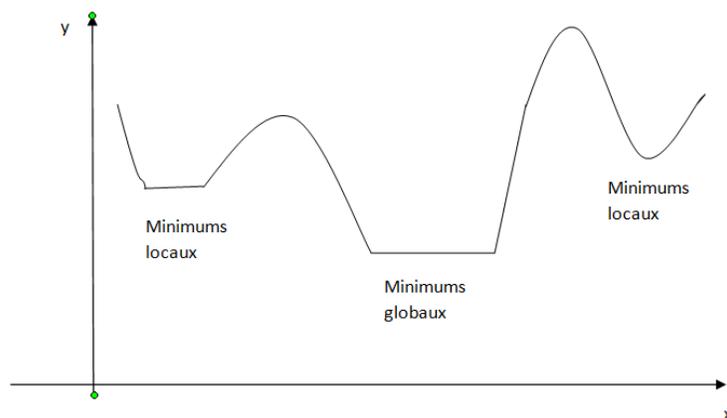
*Remarque 2.2.1.* Les maxima locaux et globaux sont définis de manière similaire. Notons que  $x^0$  est un maximum local (respectivement global) de la fonction  $f$  sur l'ensemble  $\mathbb{R}^n$  si  $x^0$  est un minimum local (respectivement global) de la fonction  $(-f)$  sur  $\mathbb{R}^n$ .

Il découle de cette observation que tout problème de maximisation peut être réduit immédiatement à un problème de minimisation (et inversement) en multipliant la fonction objectif par  $-1$ .

*Remarque 2.2.2.* Le problème (2.2) peut ne pas admettre de solutions optimales, comme dans la figure suivante :



*Remarque 2.2.3.* Le problème (2.2) peut aussi admettre une infinité de solutions optimales, comme dans la figure suivante :



*Remarque 2.2.4.* Dans le cas d'une fonction objectif convexe, il n'y a pas de distinction entre minimum local et global : tout minimum local est également global, comme l'établit le théorème suivant :

**Théorème 2.2.1.** *Soit  $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction convexe définie sur un ensemble convexe  $S$ . Alors, tout minimum local est également un minimum global. Si  $f$  est strictement convexe, alors il existe au plus un minimum global de  $f$ .*

## 2.3 Théorèmes généraux d'existence

**Théorème 2.3.1** (Weierstrass). *Soit  $S \subset \mathbb{R}^n$  un ensemble non vide. Si  $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction semi-continue inférieurement sur  $S$  compact (fermé et borné), alors il existe  $x^0 \in S$  tel que*

$$\min_{x \in S} f(x) = f(x^0).$$

**Théorème 2.3.2.** *Si  $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue sur  $\mathbb{R}^n$ . Si  $f$  est coercive, i.e.,*

$$\lim_{\|x\| \rightarrow \infty} f(x) = +\infty,$$

*alors le problème (2.2) admet une solution optimale  $x^0 \in \mathbb{R}^n$ .*

L'unicité résulte en général des propriétés de convexité de  $f$ . Avant de prouver les théorèmes 2.3.1 et 2.3.2, considérons les résultats suivants sur les fonctions coercives.

**Proposition 2.3.1.** *Soit  $S$  un sous ensemble fermé de  $\mathbb{R}^n$ . Si  $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est semi-continue inférieurement sur  $S$  alors  $\forall c \in \mathbb{R}$ , l'ensemble*

$$S(c) = \{x \in S / f(x) \leq c\}$$

*est fermé.*

*Démonstration.* Soit  $c \in \mathbb{R}$  et supposons que  $S(c) \neq \emptyset$ . Soit  $\{x^k\} \subset S(c)$  une suite de  $S(c)$  convergeant vers  $x$ .

Comme  $S$  est fermé, alors  $x \in S$ .

D'autre part, comme  $f$  est semi-continue inférieurement, alors de  $f(x_k) \leq c$ , on déduit

$$f(x) \leq \liminf_{k \rightarrow +\infty} f(x_k) \leq c.$$

D'où  $x \in S(c)$  et  $S(c)$  est fermé. □

**Proposition 2.3.2.** *Soit  $S$  un sous ensemble fermé de  $\mathbb{R}^n$ . Si  $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est semi-continue inférieurement et coercive, alors  $\forall c \in \mathbb{R}$ , l'ensemble*

$$S(c) = \{x \in S / f(x) \leq c\}$$

*est fermé et borné (compact).*

*Démonstration.* L'ensemble  $S(c)$  est fermé d'après la proposition 2.3.1. Démontrons que  $S(c)$  est borné. Supposons le contraire : il existe alors une suite  $\{x_k\} \subset S(c)$  de  $S(c)$  telle que

$$\lim_{k \rightarrow \infty} \|x^k\| = +\infty.$$

Comme  $f$  est coercive, alors

$$\lim_{k \rightarrow \infty} f(x^k) = +\infty.$$

Or

$$x^k \in S(c) \Rightarrow f(x^k) \leq c, \quad \forall k.$$

D'où

$$+\infty = \lim_{k \rightarrow \infty} f(x^k) \leq c < +\infty.$$

Ce qui nous amène à une contradiction. □

*Démonstration du théorème 2.3.1.* On suppose que  $S \neq \emptyset$ . Soit  $\{x_k\} \subset S$  une suite minimisante telle que

$$\lim_{k \rightarrow \infty} f(x_k) = \inf_{x \in S} f(x).$$

Comme  $S$  est borné, on peut extraire une sous-suite  $\{x_{k'}\}$  convergeant vers  $x^*$ .

Comme  $S$  est fermé, alors  $x^* \in S$ . Par ailleurs,  $f$  est semi-continue inférieurement, alors

$$f(x^*) \leq \lim_{k' \rightarrow \infty} f(x_{k'}) = \lim_{k \rightarrow \infty} f(x_k) = \inf_{x \in S} f(x).$$

D'où

$$f(x^*) = \inf_{x \in S} f(x).$$

□

*Démonstration du théorème 2.3.2.* Soit  $c \in \mathbb{R}$  et considérons l'ensemble

$$S(c) = \{x \in S / f(x) \leq c\}.$$

$f$  étant coercive et semi-continue inférieurement, alors d'après la proposition 2.3.2, l'ensemble  $S(c)$  est compact.

D'après le théorème 2.3.1, il existe  $x^* \in S(c)$  tel que

$$f(x^*) = \inf_{x \in S(c)} f(x) = \inf_{x \in S} f(x).$$

□

**Corollaire 2.3.1.** *Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction semi-continue inférieurement sur  $\mathbb{R}^n$  et coercive. Alors le problème*

$$\min_{x \in \mathbb{R}^n} f(x)$$

*a une solution optimale  $x^*$ .*

*Démonstration.* Soit  $x^0 \in \mathbb{R}^n$  quelconque. En supposant que la fonction  $f$  est coercive, il existe une constante  $M > 0$  telle que

$$\forall x \in \mathbb{R}^n : \|x\| \geq M \Rightarrow f(x) \geq f(x^0).$$

Le problème se ramène donc à un problème d'optimisation sur la boule compacte  $S = \{x \in \mathbb{R}^n : \|x\| \leq M\}$  qui est non vide et le théorème 2.3.1 s'applique. □

Pour beaucoup de problèmes d'optimisation sans contraintes, les principales méthodes de résolution connues ne permettent pas la détermination d'un minimum global, on se contente d'optimums locaux.

## 2.4 Caractérisation des solutions optimales

### 2.4.1 Conditions nécessaires d'optimalité

Étant donné un vecteur  $x^0$ , nous souhaiterions être capables de déterminer si ce vecteur est un minimum local ou global de la fonction  $f$ . La propriété de différentiabilité de  $f$  fournit une première manière de caractériser une solution optimale.

**Théorème 2.4.1** (Condition nécessaire du premier ordre). *Si  $x^0$  est un minimum local du problème (2.1) et supposons que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est continûment différentiable ( $f$  est continue et ses dérivées partielles  $\frac{\partial f}{\partial x_i}$  sont continues) sur un ensemble ouvert  $S \subset \mathbb{R}^n$  contenant  $x^0$ .*

Alors

$$\nabla f(x^0) = 0.$$

*Démonstration.* Soit  $d \in \mathbb{R}^n$  un vecteur quelconque. On a

$$g(\alpha) = f(x^0 + \alpha d) \geq f(x^0), \quad \forall \alpha \in \mathbb{R} / (x^0 + \alpha d) \in S.$$

En utilisant l'hypothèse que  $f$  est différentiable, on obtient

$$0 \leq \lim_{\alpha \rightarrow 0^+} \frac{f(x^0 + \alpha d) - f(x^0)}{\alpha} = \frac{dg(0)}{d\alpha} = \lim_{\alpha \rightarrow 0^-} \frac{f(x^0 + \alpha d) - f(x^0)}{\alpha} \leq 0.$$

D'où

$$\frac{dg(0)}{d\alpha} = d^T \nabla f(x^0) = 0.$$

Comme  $d$  est choisi arbitrairement, alors

$$\nabla f(x^0) = 0.$$

□

**Définition 2.4.1.** Les points  $x^0$  tels que  $\nabla f(x^0) = 0$  sont appelés points critiques (ou stationnaires) de  $f$ .

*Remarque 2.4.1.* Si  $f$  est convexe, la condition nécessaire du premier ordre est également suffisante pour que  $x^0$  soit un minimum global.

**Théorème 2.4.2** (Condition nécessaire du second ordre). *Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , deux fois continûment différentiable et  $x^0 \in \mathbb{R}^n$ . Si  $x^0$  est un minimum local (ou global) de  $f$ , alors  $\nabla f(x^0) = 0$  et  $\nabla^2 f(x^0)$  est semi-définie positive.*

*Démonstration.* Supposons que  $f$  est deux fois continûment différentiable et  $d \in \mathbb{R}^n$  un vecteur quelconque.

Pour tout  $\forall \alpha \in \mathbb{R}$ , le développement de Taylor d'ordre 2 s'écrit :

$$f(x^0 + \alpha d) - f(x^0) = \alpha [\nabla f(x^0)]^T d + \frac{\alpha^2}{2} d^T \nabla^2 f(x^0) d + o(\alpha^2).$$

En utilisant la condition  $\nabla f(x^0) = 0$  et le fait que  $x^0$  est un minimum local, on déduit l'existence d'un  $\epsilon > 0$  (suffisamment petit) tel que

$$\forall \alpha \in (0, \epsilon) : 0 \leq \frac{f(x^0 + \alpha d) - f(x^0)}{\alpha^2} = \frac{1}{2} d^T \nabla^2 f(x^0) d + \frac{o(\alpha^2)}{\alpha^2}.$$

Le passage à la limite quand  $\alpha \rightarrow 0$  et l'utilisation de la relation

$$\lim_{\alpha \rightarrow 0} \frac{o(\alpha^2)}{\alpha^2} = 0$$

entraînent

$$d^T \nabla^2 f(x^0) d \geq 0$$

ce qui démontre que  $\nabla^2 f(x^0)$  est une matrice semi-définie positive.  $\square$

### 2.4.2 Conditions suffisantes d'optimalité

Les conditions données précédemment sont nécessaires, c'est-à-dire qu'elles doivent être satisfaites pour tout minimum local (ou global) ; cependant, tout vecteur vérifiant ces conditions n'est pas nécessairement un minimum.

*Exemple 2.4.1.* Soit la fonction

$$f(x) = -x^4, \quad x \in \mathbb{R}.$$

Le point  $x^0 = 0$  satisfait les conditions nécessaires du premier et du second ordre :

$$f'(x^0) = -4x^3|_{x=0} = 0, \quad f''(x^0) = -12x^2|_{x=0} = 0 \geq 0.$$

Néanmoins, le point  $x^0 = 0$  ne constitue pas un minimum de la fonction  $f$ . Il est au contraire un maximum global :

$$\forall x \in \mathbb{R}, \quad f(x) = -x^4 \leq 0 = f(0).$$

Le théorème qui suit établit une condition suffisante pour qu'un vecteur soit un minimum local, si  $f$  est deux fois continûment différentiable.

**Théorème 2.4.3** (Condition suffisante du second ordre). *Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , deux fois continûment différentiable et  $x^0 \in \mathbb{R}^n$ . Si  $\nabla f(x^0) = 0$  et  $\nabla^2 f(x^0)$  est définie positive, alors  $x^0$  est un minimum local de  $f$ .*

*Démonstration.* Le développement de Taylor de la fonction  $f$  à l'ordre 2, au voisinage de  $x^0$  s'écrit :

$$f(x^0 + l) - f(x^0) = \frac{1}{2}l^T \nabla^2 f(x^0)l + o(\|l\|^2).$$

Toute direction  $l \in \mathbb{R}^n$ , tendant vers 0, peut être écrite sous la forme

$$l = \theta d, \text{ où } \|d\| = 1 \text{ et } \theta > 0 \text{ (assez petit).}$$

On aura alors

$$f(x^0 + l) - f(x^0) = f(x^0 + \theta d) - f(x^0) = \theta^2 \left[ \frac{1}{2}d^T \nabla^2 f(x^0)d + \frac{o(\theta^2)}{\theta^2} \right].$$

Comme

$$d^T \nabla^2 f(x^0)d > 0$$

et

$$\lim_{\theta \rightarrow 0} \frac{o(\theta^2)}{\theta^2} = 0,$$

alors pour tout nombre positif  $\theta > 0$  assez petit, on aura

$$f(x) = f(x^0 + \theta d) > f(x^0),$$

ce qui prouve que  $x^0$  est un minimum local de  $f$ . □

## 2.5 Étude de quelques exemples

*Exemple 2.5.1.* Soit  $f(x) = 1 - e^{-x^2}$ .

On a à résoudre le problème suivant :

$$\min_{x \in \mathbb{R}} f(x).$$

On a

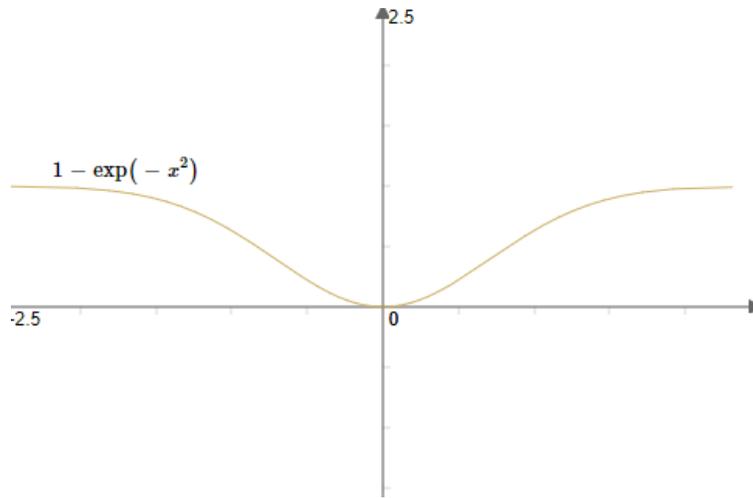
$$f'(x) = 0 \Leftrightarrow 2xe^{-x^2} = 0 \Leftrightarrow x = 0,$$

et

$$f''(x)|_{x=0} = 2e^{-x^2} - 4x^2e^{-x^2}|_{x=0} = 2 > 0.$$

Le point  $x^0 = 0$  est un minimum local de  $f$ .

On peut remarquer que le point  $x^0 = 0$  est un minimum global de  $f$ .



*Exemple 2.5.2.* Soit

$$f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2 + x_1 \rightarrow \min$$

sur  $\mathbb{R}^2$ .

On a

$$\nabla f(x) = 0 \Leftrightarrow \begin{cases} \frac{\partial f}{\partial x_1} = 2x_1 - 2x_2 + 1 = 0 \\ \frac{\partial f}{\partial x_2} = 2x_2 - 2x_1 = 0 \end{cases}$$

Les équations n'étant pas compatibles, il s'ensuit alors que le problème posé n'a pas de solution. Le maximum de  $f$  n'existe pas non plus, puisque la condition nécessaire du premier ordre est la même que pour le minimum.

*Exemple 2.5.3.* Soit

$$f(x_1, x_2) = x_1^2 - x_2^2 - 2x_1x_2 + x_1 \rightarrow \min$$

sur  $\mathbb{R}^2$ .

Cherchons les points stationnaires

$$\nabla f(x) = 0 \Leftrightarrow \begin{cases} \frac{\partial f}{\partial x_1} = 2x_1 - 2x_2 + 1 = 0 \\ \frac{\partial f}{\partial x_2} = -2x_2 - 2x_1 = 0 \end{cases} \Rightarrow x_1 = \frac{-1}{4}, \quad x_2 = \frac{1}{4}.$$

Calculons le hessien de  $f$  au point  $x^0 = \left( -\frac{1}{4} \quad \frac{1}{4} \right)$

$$\frac{\partial^2 f}{\partial x_1^2} = 2, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -2, \quad \frac{\partial^2 f}{\partial x_2^2} = -2.$$

Donc

$$\nabla^2 f(x^0) = \begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix} = M,$$

où

$$M_1 = 2 > 0; \quad M_2 = \det M = -8 < 0$$

Le hessien n'est pas semi-définie positif. Il s'ensuit alors que le point  $x^0 = \left( -\frac{1}{4} \quad \frac{1}{4} \right)$  ne constitue pas un minimum de  $f$ .

$$f\left(\frac{-1}{4}, \frac{1}{4}\right) = -\frac{1}{8} > -1 = f(0, 1).$$

*Exemple 2.5.4.* Soit  $f(x) = f(x_1, x_2) = e^{-x_1^2 - x_2^2}$  à minimiser sur  $\mathbb{R}^2$ .

On a

$$\begin{cases} \frac{\partial f}{\partial x_1} = -2x_1 e^{-x_1^2 - x_2^2} = 0 \\ \frac{\partial f}{\partial x_2} = -2x_2 e^{-x_1^2 - x_2^2} = 0 \end{cases} \Rightarrow x_1 = x_2 = 0.$$

D'autre part, au point  $x^0 = (0, 0)$ , on a

$$\nabla^2 f(x^0) < 0.$$

Le point  $x^0 = (0, 0)$  ne constitue pas donc un minimum pour la fonction  $f$ . La fonction  $f$  admet le point  $x^0 = (0, 0)$  comme maximum.

## 2.6 Exercices

*Exercice 2.6.1.* Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

- Démontrer que si  $f$  est convexe, alors tout minimum local est global.
- Supposons que  $f$  est une forme quadratique. Démontrer une condition nécessaire et suffisante pour que  $f$  soit strictement convexe.
- Déterminer les valeurs de  $a, b$  et  $c$  telles que la fonction

$$f(x, y, z) = x^2 + 2axy + by^2 + cz^2,$$

soit convexe sur  $\mathbb{R}^3$ .

d) Déterminer, si possible, les valeurs de  $a$  et  $b$  telles que

$$f(x) = x^3 + ax^2 + bx$$

admette :

1. un point maximum local au point  $x = -1$  et un point minimum local au point  $x = +1$ .
2. un point maximum local au point  $x = +1$  et un point minimum local au point  $x = 0$ .

*Exercice 2.6.2.* Soit  $\varphi$  la fonction définie sur  $\mathbb{R}$  par

$$\varphi(x) = x \exp(-x^2/2), \text{ pour } x \in \mathbb{R}$$

et  $f$  la fonction définie sur  $\mathbb{R}^2$  par

$$f(x, y) = xy \exp((-x^2 - y^2)/2), \text{ pour } (x, y) \in \mathbb{R}^2.$$

1. Déterminer les extrema de  $\varphi$  sur  $\mathbb{R}$  et préciser leurs natures.
2. Déterminer les extrema de  $f$  sur  $\mathbb{R}^2$  et préciser leurs natures.

*Exercice 2.6.3.* Supposons que  $f : \mathbb{R} \rightarrow \mathbb{R}$  est convexe, et  $a$  et  $b$  du domaine de définition de  $f$  avec  $a < b$ .

(a) Montrer que

$$f(x) \leq \frac{b-x}{a-x} f(a) + \frac{x-a}{x-b} f(b), \forall x \in [a, b].$$

(b) Montrer que

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x}, \forall x \in ]a, b[.$$

(c) Supposons que  $f$  est différentiable. Utiliser le résultat de (b) (ou la définition de  $f$  convexe) pour démontrer que

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

(d) Supposons que  $f$  est deux fois différentiable. Utiliser le résultat de (c) pour montrer que  $f''(a) \geq 0$  et  $f''(b) \geq 0$

*Exercice 2.6.4.* – Montrer que la fonction  $f$  définie sur  $\mathbb{R}^2$  par

$$f(x, y) = 2x^2 + y^2 - \sqrt{3}xy + 2x + 1, \quad (x, y) \in \mathbb{R}^2,$$

admet un unique point de minimum global sur  $\mathbb{R}^2$  que l'on déterminera.

– Minimiser la fonction suivante sur son domaine de définition :

$$g(x_1, x_2, \dots, x_n) = \prod_{i=1}^n x_i^{x_i}.$$

*Exercice 2.6.5.* Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$f(x_1, x_2) = x_1^3 - 3x_1(1 + x_2^2).$$

a)  $f$  admet-elle des points critiques ? des extrêma locaux ? justifier.

b) Soit  $g : [0, 2\pi] \rightarrow \mathbb{R}$  la fonction définie par

$$g(t) = f(\cos t, \sin t)$$

1. Montrer que  $g'(t) = -6 \sin t \cos(2t)$
2. Trouver les points critiques de  $g$ .
3. Etudier la variation de  $g$ .
4. Déduire les extrêma de  $g$ .

*Exercice 2.6.6.* Considérons la fonction suivante :

$$h(\theta) = \frac{4\theta - 7}{\theta^2 + \theta - 2}$$

- Identifier tous les points stationnaires de  $h$  ;
- Parmi ceux-ci, identifier tous les minima locaux de  $h$  ;
- Identifier des points où la fonction n'est pas bornée.

*Exercice 2.6.7.* Considérons la fonction suivante :

$$h(\theta) = 1 - 12\theta + 7.5\theta^2 - \theta^3.$$

- Identifier tous les points stationnaires de  $h$  ;
- Parmi ceux-ci, identifier tous les minima locaux de  $h$ .

*Exercice 2.6.8.* a) Montrer que si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est strictement convexe et possède un minimum, alors celui-ci est unique.

b) L'affirmation suivante, spécifier si elle est vraie ou fausse. Justifier la réponse.

Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  et  $x^*$  tel que  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est semi définie positive. Alors  $x^*$  est un minimum local de  $f$ .

*Exercice 2.6.9.* – Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  définie par

$$\forall x \in \mathbb{R}, f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}.$$

On fixe  $x_1, \dots, x_n \in \mathbb{R}$ . Résoudre le problème suivant :

$$\max_{\mu \in \mathbb{R}} \left[ \prod_{i=1}^n f(x_i) \right].$$

– Déterminer les extrema, s'ils existent, de la fonction suivante :

$$f(x, y) = x^2 y^2 + x^2 + y^2 + 2axy \quad (a \geq 0).$$

*Exercice 2.6.10.* Soient  $n \geq 2$  et  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  la fonction définie par

$$f(x) = (1 + x_n)^3 \sum_{i=1}^{n-1} x_i^2 + x_n^2.$$

Montrer que 0 est le seul point critique de  $f$ , que  $f$  y atteint un minimum local strict, mais pas global.

# Chapitre 3

## Les méthodes numériques pour l'optimisation sans contraintes

### 3.1 Introduction

Comme la stationnarité de  $f$  est une condition nécessaire d'optimalité, pratiquement toutes les méthodes de minimisation sans contraintes dans  $\mathbb{R}^n$  consistent à rechercher un point  $x^*$  stationnaire ( $\nabla f(x^*) = 0$ ). Ce problème est équivalent à la résolution d'un système d'équations non linéaires :

$$g_i(x) = \frac{\partial f}{\partial x_i}(x) = 0, \quad i = 1, \dots, n.$$

On peut chercher à résoudre directement ce système, ce qui conduit à la méthode de Newton. Cependant, il faut se rappeler que cette méthode présente quelques inconvénients :

- i) la divergence de la méthode, si le point de départ est trop éloigné de  $x^*$ ,
- ii) elle exige que la fonction  $f$  soit deux fois continûment différentiable et exige le calcul des dérivées secondes.

C'est pourquoi, les méthodes les plus couramment utilisées sont basées sur un principe général assez simple, mais dont l'analyse des conditions de convergence sont complexes : on les appelle méthodes de descente. Il s'agit de procédures itératives de construction d'une suite de points  $\{x_k\}$  telle que la fonction  $f$  décroît d'une itération à une autre, i.e.,

$$f(x^{k+1}) < f(x^k), \quad k = 0, 1, \dots.$$

Le processus itératif général de construction de la suite  $\{x_k\}$  a la forme suivante :

$$x^{k+1} = x^k + \theta_k d_k, \quad k = 0, 1, \dots \quad (3.1)$$

où  $d_k$  est une direction d'amélioration ou de descente et  $\theta_k$  le pas le long de cette direction.

Le problème restant posé et qui différencie les différentes méthodes numériques est le choix du pas  $\theta_k$  et de la direction  $d_k$  à chaque itération.

## 3.2 La méthode du gradient

### 3.2.1 Interprétation physique du gradient

Pour une fonction différentiable  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , le gradient  $\nabla f(x^0)$  de  $f$  au point  $x^0$  se définit comme suit :

$$f(x^0 + \Delta x) - f(x^0) = (\nabla f(x^0))^T \Delta x + o(\|\Delta x\|), \quad (3.2)$$

où

$$\lim_{\|\Delta x\| \rightarrow 0} \frac{o(\|\Delta x\|)}{\|\Delta x\|} = 0.$$

Une autre interprétation du gradient peut se faire de la manière suivante : soit  $x^0 \in \mathbb{R}^n$ . De  $x^0$ , on choisit une direction de déplacement  $d \in \mathbb{R}^n$ ,  $\|d\| = 1$ , et on pose :

$$x(\theta) = x^0 + \theta d, \quad \theta \geq 0.$$

Le point  $x(\theta)$  est le point atteint après un déplacement le long de la direction  $d$  avec un pas  $\theta$ . Estimons la variation de la valeur de la fonction  $f$  en passant du point  $x^0$  au point  $x^0 + \theta d$ . La relation (3.2) prend la forme suivante :

$$f(x^0 + \theta d) - f(x^0) = \theta \left[ (\nabla f(x^0))^T d + \frac{o(\theta)}{\theta} \right].$$

Pour des valeurs de  $\theta \geq 0$  voisines de zéro, le signe de l'expression  $f(x^0 + \theta d) - f(x^0)$  dépendra du signe de  $(\nabla f(x^0))^T d$ .

Pour avoir la relation  $f(x^0 + \theta d) < f(x^0)$ , ie., une diminution de la valeur de  $f(x)$  lors du passage du point  $x^0$  au point  $x^0 + \theta d$ , on devrait choisir  $\theta \geq 0$  assez petit et la direction  $d \in \mathbb{R}^n$  de façon à avoir  $(\nabla f(x^0))^T d < 0$ .

**Définition 3.2.1.** La direction  $d \in \mathbb{R}^n$  est appelée direction de descente, si

$$(\nabla f(x^0))^T d < 0 \quad (3.3)$$

Le processus itératif (3.1) de construction de la suite  $\{x_k\}$  prend la forme

$$x^{k+1} = x^k + \theta_k d_k, \quad k = 0, 1, \dots \quad (3.4)$$

avec, si  $\nabla f(x^k) \neq 0$ , la direction  $d_k$  est choisie de façon que

$$[\nabla f(x^k)]^T d_k < 0, \quad (3.5)$$

et le pas  $\theta_k$  est positif. Si  $\nabla f(x^k) = 0$ , alors le processus itératif est arrêté et  $x^{k+1} = x^k$  (ce qui est équivalent à choisir  $d_k = 0$ ).

En vertu de la relation (3.5) de  $d_k$  et du gradient  $\nabla f(x^k)$ , on appelle les algorithmes de ce type "méthodes du gradient". Certains auteurs réservent le terme "méthode du gradient" pour le cas particulier où  $d_k = -\nabla f(x^k)$ .

Il existe une large variété de possibilités pour choisir la direction  $d_k$  et le pas  $\theta_k$  dans les méthodes du gradient.

### 3.2.2 Sélection des directions de descente

Plusieurs méthodes du gradient s'écrivent sous la forme :

$$x^{k+1} = x^k - \theta_k D_k \nabla f(x^k), \quad (3.6)$$

où  $D_k$  est une matrice symétrique définie positive. Dans ce cas, la direction  $d_k$  s'écrit

$$d_k = -D_k \nabla f(x^k)$$

et la condition de descente

$$[\nabla f(x^k)]^T d_k < 0$$

peut s'écrire

$$- [\nabla f(x^k)]^T D_k \nabla f(x^k) < 0$$

et qui est vérifiée puisque  $D_k > 0$ .

On donne quelques exemples de choix de la matrice  $D_k$ .

### Méthode de la descente la plus rapide (méthode de la plus forte pente, steepest descent)

On pose  $D_k = I$ ,  $k = 0, 1, 2, \dots$  où  $I$  est la matrice identité  $n \times n$ . Dans ce cas, le processus itératif (3.6) prend la forme

$$x^{k+1} = x^k - \theta_k \nabla f(x^k), \quad k = 0, 1, 2, \dots \quad (3.7)$$

Nous allons justifier ci-après l'appellation de méthode de la plus forte pente. En partant du point  $x^k$  et en se mouvant le long de la direction  $d$ , la fonction  $f(x)$  va varier avec une vitesse donnée par la dérivée directionnelle de  $f$  au point  $x^k$  dans la direction  $d$ , soit :

$$f'(x^k; d) = \lim_{\theta \searrow 0} \frac{f(x^k + \theta d) - f(x^k)}{\theta} = [\nabla f(x^k)]^T d.$$

Pour des directions  $d$  telles que  $\|d\| = 1$ , la vitesse de variation de la fonction  $f$  sera maximale pour

$$d_k = \frac{\nabla f(x^k)}{\|\nabla f(x^k)\|} \quad (3.8)$$

En effet, on a d'après l'inégalité de Cauchy-Schwartz

$$[\nabla f(x^k)]^T d \leq \|\nabla f(x^k)\| \|d\| = \|\nabla f(x^k)\|$$

et pour  $d_k$  vérifiant (3.8), on a

$$[\nabla f(x^k)]^T d_k = \frac{[\nabla f(x^k)]^T [\nabla f(x^k)]}{\|\nabla f(x^k)\|}.$$

Ainsi, le gradient  $\nabla f(x^k)$  constitue la meilleure direction de croissance à partir du point  $x^k$ . La direction opposée à celle du gradient ( $-\nabla f(x^k)$ ) est la meilleure direction de descente à partir de  $x^k$ .

L'inconvénient de la méthode de descente rapide est sa lente convergence.

#### 3.2.3 Sélection du pas

Il existe un certain nombre de règles de choix du pas  $\theta_k$  dans la méthode du gradient. Nous allons présenter les plus couramment utilisées.

**a) Règle du minimum**

Elle consiste à choisir, à l'itération  $k$ , le pas  $\theta_k$  minimisant la fonction  $f$  dans la direction  $d_k$ , c-à-d

$$f(x^k + \theta_k d_k) = \min_{\theta \geq 0} f(x^k + \theta d_k). \quad (3.9)$$

**b) Règle du minimum réduit**

C'est une version de la règle du minimum qui est la plus utilisée. On fixe un scalaire  $s > 0$  et  $\theta_k$  est défini par

$$f(x^k + \theta_k d_k) = \min_{\theta \in [0, s]} f(x^k + \theta d_k). \quad (3.10)$$

*Remarque 3.2.1.* La résolution des problèmes d'optimisation (3.9) et (3.10) se fait par des méthodes d'optimisation unidimensionnelle.

**c) Règle du pas constant**

Ici, on fixe pour toutes les itérations

$$\theta_k = s, \quad k = 0, 1, \dots \quad (3.11)$$

La règle du pas constant est la plus simple. Toutefois, si le pas est grand, il y-a risque de divergence et quand le pas est trop petit, la vitesse de convergence serait très lente. Ainsi la règle du pas constant n'est utile que pour les problèmes où la valeur appropriée du pas est connue ou peut être facilement calculée.

*Exemple 3.2.1.* On considère la fonctionnelle quadratique suivante :

$$f(x_1, x_2) = 8x_1^2 + 4x_1x_2 + 5x_2^2.$$

On applique la méthode du gradient à pas optimal, en partant du point initial

$$x^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \end{pmatrix},$$

pour trouver le minimum de  $f$ .

- On calcule le gradient de  $f$  :

$$\nabla f(x) = \begin{pmatrix} 16x_1 + 4x_2 \\ 4x_1 + 10x_2 \end{pmatrix}.$$

- On calcule la valeur du gradient au point  $x^{(0)}$  :

$$\nabla f(x^{(0)}) = \begin{pmatrix} 200 \\ 140 \end{pmatrix}.$$

- On calcule le point  $x^{(1)}$  :

$$x^{(1)} = x^{(0)} - \alpha^{(0)} \nabla f(x^{(0)}) = \begin{pmatrix} 10 \\ 10 \end{pmatrix} - \alpha^{(0)} \begin{pmatrix} 200 \\ 140 \end{pmatrix} = \begin{pmatrix} 10 - 200\alpha^{(0)} \\ 10 - 140\alpha^{(0)} \end{pmatrix}.$$

- On cherche la valeur optimale de  $\alpha^{(0)}$ . Pour ce, on calcule  $f(x^{(1)})$  puis on détermine  $\alpha^{(0)}$  permettant de minimiser  $f(x^{(1)})$ . On a

$$\begin{aligned} f(x^{(1)}) &= 8(10 - 200\alpha^{(0)})^2 + 4(10 - 200\alpha^{(0)})(10 - 140\alpha^{(0)}) + 5(10 - 140\alpha^{(0)})^2 \\ &= -59600 + 1060000\alpha^{(0)} \\ &= \varphi(\alpha^{(0)}). \end{aligned}$$

On calcule  $\alpha^{(0)}$  tel que  $\varphi'(\alpha^{(0)}) = 0$ , d'où  $\alpha^{(0)} = 0.056$ .

- On trouve ainsi

$$x^{(1)} = \begin{pmatrix} -1.20 \\ 2.16 \end{pmatrix}.$$

- On cherche de la même façon tous les points  $x^{(k)}$  jusqu'à ce que le critère d'arrêt soit satisfait.

*Exemple 3.2.2.* Utiliser la méthode de la plus forte pente puis de la plus forte pente accélérée (p=2) pour trouver le minimum de la fonction  $f$  définie de  $\mathbb{R}^2$  dans  $\mathbb{R}$  comme suit :

$$f(x_1, x_2) = x_1^4 - 2x_1^3 + \frac{1}{6}x_2^2 - 3x_1 - x_1x_2.$$

On a

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1^3 - 6x_1^2 - 3 - x_2 \\ \frac{1}{3}x_2 - x_1 \end{pmatrix}.$$

Soit  $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . On trouve la direction de la plus forte pente

$$-\nabla f(x^0) = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

On minimise donc

$$g_0(\lambda) = f(x^0 - \lambda \nabla f(x^0)) = 81\lambda^4 - 54\lambda^3 - 6\lambda.$$

On trouve

$$\lambda_0 \approx 0.582.$$

Notre premier pas nous conduit donc au point

$$x^1 = x^0 - \lambda_0 \nabla f(x^0) = \begin{pmatrix} 1.746 \\ 0 \end{pmatrix}.$$

On réitère pour obtenir les valeurs suivantes :

$i$	$x^i$	$-\nabla f(x^i)$	$g_i(\lambda)$	$\lambda_i$	$x^i$
0	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 0 \end{pmatrix}$	$81\lambda^4 - 54\lambda^3 - 6\lambda$	0.582	$\begin{pmatrix} 1.746 \\ 0 \end{pmatrix}$
1	$\begin{pmatrix} 1.746 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1.746 \end{pmatrix}$	$0.508\lambda^2 - 3.049\lambda - 6.590$	3	$\begin{pmatrix} 1.746 \\ 5.238 \end{pmatrix}$
2	$\begin{pmatrix} 1.746 \\ 5.238 \end{pmatrix}$	$\begin{pmatrix} 5.238 \\ 0 \end{pmatrix}$	$752\lambda^4 + 716\lambda^3 + 214\lambda^2 - 27\lambda - 11$	0.05	$\begin{pmatrix} 2.01 \\ 5.238 \end{pmatrix}$
3	$\begin{pmatrix} 2.01 \\ 5.238 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.264 \end{pmatrix}$	$0.012\lambda^2 - 0.07\lambda - 11.904$	3	$\begin{pmatrix} 2.01 \\ 6.03 \end{pmatrix}$

Pour la plus forte pente accélérée, après avoir fait 2 pas de la plus forte pente standard, nous devons faire un pas accéléré, en utilisant la direction  $x^2 - x^0$ . On minimise donc la fonction

$$g_2(\lambda) = f(x^0 + \lambda(x^2 - x^0)) = 9.29\lambda^4 - 10.65\lambda^3 - 4.57\lambda^2 - 5.238\lambda.$$

On trouve

$$\lambda_2 = 1.172$$

et on atteint le point

$$x^3 = x^0 + \lambda_2(x^2 - x^0) = \begin{pmatrix} 2.05 \\ 6.137 \end{pmatrix},$$

qui est près du minimum global.

### 3.3 La méthode de Newton

La méthode de Newton est obtenue en posant dans (3.6)

$$D_k = (\nabla^2 f(x^k))^{-1}, \quad k = 0, 1, \dots \quad (3.12)$$

ce qui donne

$$x^{k+1} = x^k - \theta_k (\nabla^2 f(x^k))^{-1} \nabla f(x^k), \quad k = 0, 1, \dots$$

en supposant que  $\nabla^2 f(x^k)$  est une matrice définie positive. Dans le cas où  $\nabla^2 f(x^k)$  n'est pas définie positive, une certaine modification est nécessaire.

L'idée de la méthode de Newton peut s'expliquer de deux manières :

- Lors de la minimisation d'une fonction sur  $\mathbb{R}^n$ , le problème consistait à chercher un point  $x^* \in \mathbb{R}^n$  tel que

$$\nabla f(x^*) = 0.$$

La méthode numérique de Newton appliquée à l'équation

$$g(x^*) = \nabla f(x^*) = 0$$

construit les approximations successives de la manière suivante :

$$x^{k+1} = x^k - \left[ \frac{\partial g(x^k)}{\partial x} \right]^{-1} g(x^k), \quad k = 0, 1, \dots \quad (3.13)$$

où

$$g = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}, \quad \frac{\partial g(x^k)}{\partial x} = \left( \frac{\partial g_i(x^k)}{\partial x_j} \right)_{1 \leq i, j \leq n}.$$

En tenant compte du fait que

$$\frac{\partial g(x^k)}{\partial x} = \nabla^2 f(x^k) = H(x^k),$$

la formule itérative (3.13) s'écrit alors

$$x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1} \nabla f(x^k), \quad k = 0, 1, \dots \quad (3.14)$$

Le point  $x^{k+1}$  peut être interprété comme la racine de l'approximation linéaire de la fonction

$$g(x) = \nabla f(x)$$

au voisinage de  $x^k$ . On a

$$g(x) \simeq g(x^k) + \frac{\partial g(x^k)}{\partial x}(x - x^k) = 0,$$

d'où

$$x - x^k = - \left[ \frac{\partial g(x^k)}{\partial x} \right]^{-1} g(x^k)$$

et

$$x = x^{k+1} = x^k - \left[ \frac{\partial g(x^k)}{\partial x} \right]^{-1} g(x^k)$$

ou

$$x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1} \nabla f(x^k).$$

- Une autre interprétation de la méthode de Newton est la suivante : elle consiste à minimiser l'approximation quadratique  $q(x)$  de  $f(x)$  au voisinage de  $x^k$ .

$$f(x) \simeq q(x) = f(x^k) + [\nabla f(x^k)]^T (x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f(x^k)(x - x^k).$$

Si le hessien  $H(x^k) = \nabla^2 f(x^k)$  est définie positif, le point  $x^{k+1}$  de la méthode de Newton est alors pris comme étant le minimum de la fonction  $q(x)$  :

$$0 = \nabla q(x^{k+1}) = \nabla f(x^k) + \nabla^2 f(x^k)(x^{k+1} - x^k) = 0.$$

D'où la formule itérative

$$x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1} \nabla f(x^k).$$

Cette formule itérative est une itération de la méthode de Newton. Elle correspond à l'itération plus générale

$$x^{k+1} = x^k - \theta_k D_k \nabla f(x^k), \quad k = 0, 1, \dots \quad (3.15)$$

où  $\theta_k = 1$ ,  $D_k = [\nabla^2 f(x^k)]^{-1}$ , ou encore

$$x^{k+1} = x^k + \theta_k d_k$$

avec  $\theta_k = 1$ ,  $d_k = -[\nabla^2 f(x^k)]^{-1} \nabla f(x^k)$ .

*Remarque 3.3.1.* Lorsque le hessien  $\nabla^2 f(x^k)$  n'est pas défini positif, la direction de déplacement  $d_k$  dans la méthode de Newton peut ne pas être une direction de descente. En effet, l'expression

$$[\nabla f(x^k)]^T d_k = - [\nabla f(x^k)]^T [\nabla^2 f(x^k)]^{-1} \nabla f(x^k) = -d_k^T \nabla^2 f(x^k) d_k$$

peut être éventuellement positive et dans ce cas, la direction  $d_k$  ne serait plus une direction de descente.

*Exemple 3.3.1.* On cherche à obtenir le zéro de la fonction

$$F : (x, y, z) \mapsto \begin{pmatrix} x^2 + y^2 + z^2 - 3 \\ x^2 + y^2 - z - 1 \\ x + y + z - 3 \end{pmatrix}.$$

La méthode de Newton s'écrit comme suit :

$$\begin{cases} x^{(0)} \text{ est donné} \\ x^{(k+1)} = x^{(k)} - (\nabla F(x^{(k)}))^{-1} F(x^{(k)}). \end{cases}$$

Dans notre cas, on a

$$\nabla F(x) = \begin{pmatrix} 2x & 2y & 2z \\ 2x & 2y & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

**Premier cas :** On choisit  $x^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ .

*Itération 1 :* On a

$$F(x^{(0)}) = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

et

$$\nabla F(x^{(0)}) = \begin{pmatrix} 2 & 0 & 2 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

d'où

$$x^{(1)} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}.$$

*Itération 2* : On a

$$F(x^{(1)}) = \begin{pmatrix} \frac{-1}{2} \\ \frac{3}{2} \\ -1 \end{pmatrix}$$

et

$$\nabla F(x^{(1)}) = \begin{pmatrix} 3 & 1 & 2 \\ 3 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

d'où

$$x^{(2)} = \begin{pmatrix} \frac{5}{4} \\ \frac{3}{4} \\ 1 \end{pmatrix}.$$

*Itération 3* : On a

$$F(x^{(2)}) = \begin{pmatrix} \frac{1}{8} \\ \frac{1}{8} \\ 0 \end{pmatrix}$$

et

$$\nabla F(x^{(2)}) = \begin{pmatrix} 5/2 & 3/2 & 2 \\ 5/2 & 3/2 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

d'où

$$x^{(3)} = \begin{pmatrix} \frac{9}{8} \\ \frac{1}{8} \\ 1 \end{pmatrix}.$$

On a ainsi

$$F(x^{(3)}) = \begin{pmatrix} \frac{1}{32} \\ \frac{1}{32} \\ 0 \end{pmatrix}.$$

**Deuxième cas** : On choisit  $x^{(0)} = (0 \ 0 \ 0)$ .

*Itération 1* : On a

$$F(x^{(0)}) = \begin{pmatrix} 3 \\ -1 \\ -3 \end{pmatrix}$$

et

$$\nabla F(x^{(0)}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

On tombe ainsi sur l'un des cas pathologiques de la méthode de Newton, lorsque la matrice  $\nabla F$  est singulière.

## 3.4 La méthode du gradient conjugué

Dans cette méthode, on utilise un modèle quadratique de  $f$ , nous allons donc d'abord présenter l'algorithme appliqué à une fonction fortement convexe quadratique.

### 3.4.1 Cas linéaire

Soit  $f(x) = \frac{1}{2}x^T Ax - bx$ , avec  $A$  symétrique définie positive. On sait alors qu'il existe un minimum unique sur  $\mathbb{R}^n$  donné par  $x^* = A^{-1}b$ . Dans la suite, on note  $r(x) = Ax - b = \nabla f(x)$  le résidu.

**Définition 3.4.1.** Les vecteurs non nuls  $\{d_1, \dots, d_p\}$  sont dits conjugués par rapport à la matrice  $A$  si

$$\text{pour tout } i \neq j \in \{1, \dots, p\} : d_i^T A d_j = 0.$$

**Lemme 3.4.1.** *Un ensemble de vecteurs conjugués par rapport à  $A$  est un ensemble de vecteurs linéairement indépendants.*

Posons  $\phi(\rho) = f(x + \rho d)$ , alors

$$\phi'(\rho) = \rho d^T A d + A x d - b d$$

et

$$f(x + \rho d) = \frac{1}{2}(x + \rho d)^T A(x + \rho d) - b(x + \rho d).$$

D'où

$$\phi'(\rho) = 0 \Rightarrow \rho = -\frac{r(x)d}{d^T A d}, \quad (d \neq 0). \quad (3.16)$$

**Théorème 3.4.1.** *Soit  $x^{(0)} \in \mathbb{R}^m$  et  $\{d_0, \dots, d_{m-1}\}$  un ensemble de vecteurs conjugués par rapport à  $A$ . On considère la suite définie par*

$$x^{(k+1)} = x^{(k)} + \rho_k d_k, \quad \text{où} \quad \rho_k = -\frac{r(x^{(k)})d_k}{d_k^T A d_k}.$$

*Alors  $r(x^{(k)})d_i = 0$  pour  $i = 0, \dots, k-1$  et  $x^{(k)}$  minimise  $f$  et la suite  $(x^{(k)})$  converge en au plus  $n$  itérations.*

Pour pouvoir utiliser la méthode itérative du théorème précédent, il faut construire un ensemble de vecteurs conjugués par rapport à  $A$ . Or, afin de réduire le nombre d'opérations, on se propose de déduire  $d^{(k)}$  à partir de  $d^{(k-1)}$  uniquement.

Dans l'algorithme du gradient conjugué linéaire, on prend

$$d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$$

avec  $\beta_k$  tel que  $d^{(k)T} Ad^{(k-1)} = 0$ , on en déduit

$$\begin{aligned} d^{(k)T} Ad^{(k-1)} &= (-\nabla f(x^{(k)}) + \beta_k d^{(k-1)})^T Ad^{(k-1)} \\ &= -r(x^{(k)T})Ad^{(k-1)} + \beta_k d^{(k-1)T} Ad^{(k-1)} \end{aligned}$$

d'où

$$\beta_k = \frac{r(x^{(k)T})Ad^{(k-1)}}{d^{(k-1)T} Ad^{(k-1)}}. \quad (3.17)$$

Comme  $d^{(k)T} \nabla f(x^{(k)}) = -\|r(x^{(k)})\|_2^2$ , on déduit bien une direction de descente.

*Remarque 3.4.1.* On rappelle que s'il existe un vecteur  $d$  tel que  $d^T \nabla f(x^*) < 0$ , alors  $d$  est une direction de descente de  $f$  en  $x^*$ .

### Algorithme : Gradient conjugué linéaire

1. Initialisation Choisir  $x^{(0)}$  et poser  $k = 0$ ,  $r^{(0)} = Ax^{(0)} - b$ ,  $d^{(0)} = -r^{(0)}$ .
2. Itération  $k$  Tant que  $\|r(x^{(k)})\| > \epsilon$ , faire

$$\begin{aligned} \rho_k &= \frac{r^{(k)T} r^{(k)}}{d^{(k)T} Ad^{(k)}} \\ x^{(k+1)} &= x^{(k)} + \rho_k d^{(k)} \\ r^{(k+1)} &= r^{(k)} + \rho_k Ad^{(k)} \\ \beta_{k+1} &= \frac{r^{(k+1)T} r^{(k+1)}}{r^{(k)T} r^{(k)}} \\ d^{(k+1)} &= -r^{(k+1)} + \beta_{k+1} d^{(k)} \\ k &= k + 1 \end{aligned}$$

Fin tant que

Notons que les formules pour  $r^{(k)}$  et  $\beta_{k+1}$  ont changé, ceci est possible grâce au théorème suivant qui affirme en particulier que les  $d^{(k)}$  sont conjugués.

**Théorème 3.4.2.** Soit  $x^{(k)} \neq x^*$ , obtenu après la  $k^{\text{ème}}$  itération de l'algorithme du gradient conjugué linéaire avec  $r^{(k)}$  et  $\beta_{k+1}$  définis par les équations (3.16) et (3.17) respectivement. Alors, on a

- (i)  $r^{(k)}r^{(i)} = 0$ , pour  $i = 0, \dots, k-1$  ;  
(ii)  $\text{Vect}\{r^{(0)}, \dots, r^{(k)}\} = \text{Vect}\{r^{(0)}, Ar^{(0)} \dots, A^k r^{(k)}\}$  ;  
(iii)  $\text{Vect}\{d^{(0)}, \dots, d^{(k)}\} = \text{Vect}\{d^{(0)}, Ad^{(0)} \dots, A^k d^{(k)}\}$  ;  
(iv)  $d^{(k)T} Ad^{(i)} = 0$  pour  $i = 0, \dots, k-1$  ;

Commentaires. • On a

$$\rho_k = -\frac{r^{(k)T} d^{(k)}}{d^{(k)T} Ad^{(k)}}, \quad (3.18)$$

et dans l'algorithme, on a

$$\rho_k = \frac{r^{(k)T} r^{(k)}}{d^{(k)T} Ad^{(k)}} \quad (3.19)$$

Montrons l'équivalence de (3.18) et (3.19). On a  $d^{(k)} = -r^{(k)} + \beta_k d^{(k-1)}$ , donc (3.18) s'écrit comme suit :

$$\begin{aligned} \rho_k &= -\frac{r^{(k)T}}{d^{(k)T} Ad^{(k)}} (-r^{(k)} + \beta_k d^{(k-1)}) \\ &= \frac{r^{(k)T} r^{(k)}}{d^{(k)T} Ad^{(k)}} - \beta_k \frac{r^{(k)T} d^{(k-1)}}{d^{(k)T} Ad^{(k)}} \\ &= \frac{r^{(k)T} r^{(k)}}{d^{(k)T} Ad^{(k)}} \end{aligned}$$

car  $r^{(k)T} d^{(k-1)} = 0$ .

- On a  $x^{(k+1)} = x^{(k)} + \rho_k d^{(k)}$  et  $r^{(k)} = Ax^{(k)} - b$  d'où

$$\begin{aligned} r^{(k+1)} - r^{(k)} &= (Ax^{(k+1)} - b) - (Ax^{(k)} - b) \\ &= (Ax^{(k+1)} - x^{(k)}) \\ &= \rho_k Ad^{(k)} \end{aligned}$$

d'où

$$r^{(k+1)} = r^{(k)} + \rho_k Ad^{(k)}.$$

- Montrons que

$$\beta_k = \frac{r^{(k)T} Ad^{(k-1)}}{d^{(k-1)T} Ad^{(k-1)}}$$

et

$$\beta_k = \frac{r^{(k)T} r^{(k)}}{r^{(k-1)T} r^{(k-1)}}$$

sont équivalentes.

On a

$$r^{(k)} - r^{(k-1)} = \rho_{k-1} Ad^{(k-1)}$$

donc

$$r^{(k)T} Ad^{(k-1)} = r^{(k)T} \frac{r^{(k)} - r^{(k-1)}}{\rho_{k-1}}$$

d'où

$$\beta_k = r^{(k)T} \frac{r^{(k)} - r^{(k-1)}}{\rho_{k-1} d^{(k-1)T} Ad^{(k-1)}}$$

or

$$\rho_k = \frac{r^{(k)T} r^{(k)}}{d^{(k)T} Ad^{(k)}} \Rightarrow \rho_{k-1} d^{(k-1)T} Ad^{(k-1)} = r^{(k-1)T} r^{(k-1)}$$

d'où

$$\beta_k = r^{(k)T} \frac{r^{(k)} - r^{(k-1)}}{r^{(k-1)T} r^{(k-1)}} = \frac{r^{(k)T} r^{(k)}}{r^{(k-1)T} r^{(k-1)}}$$

car

$$r^{(k)T} r^{(k-1)} = 0, \quad r^{(k)} = d^{(k)} - \beta_{k-1} d^{(k-1)}$$

et  $r^{(k)}$  est orthogonal à toutes les directions  $d^{(0)}, d^{(1)}, \dots, d^{(k)}$ .

□

- Remarque 3.4.2.*
1. Dans la démonstration du théorème, il est essentiel de choisir  $d^{(0)} = -\nabla f(x^{(0)})$ . C'est à partir de  $d^{(0)}$  que l'on construit des gradients orthogonaux et des directions conjuguées.
  2. L'algorithme du gradient conjugué linéaire est surtout utile pour résoudre des grands systèmes creux.
  3. La convergence peut être assez rapide, si  $A$  admet seulement  $p$  valeurs propres distinctes, la convergence a lieu en au plus  $p$  itérations.

### 3.4.2 Cas non linéaire

Pour pouvoir appliquer l'algorithme du gradient conjugué linéaire au cas d'une fonction  $f$  quelconque, il faut :

- remplacer  $r(x^{(k)})$  par  $\nabla f(x^{(k)})$ ;
- déterminer  $\rho_k$  par une minimisation unidimensionnelle.

#### Algorithme du Gradient conjugué non linéaire (Fletcher-Reeves)

1. Initialisation Choisir  $x^{(0)}$ , poser  $k = 0$  et calculer  $f^{(0)} = f(x^{(0)})$  et  $\nabla f^{(0)} = \nabla f(x^{(0)})$ ,  $d^{(0)} = -\nabla f^{(0)}$ .

2. Itération  $k$  Tant que  $\|\nabla f^{(k)}\| > \epsilon$ , faire
- Déterminer  $\rho_k$
  - Calculer  $x^{(k+1)} = x^{(k)} + \rho_k d^{(k)}$
  - Calculer  $\nabla f^{(k+1)} = \nabla f(x^{(k+1)})$
  - Calculer  $\beta_{k+1} = \frac{\nabla f^{(k+1)T} \nabla f^{(k+1)}}{\nabla f^{(k)T} \nabla f^{(k)}}$
  - Calculer  $d^{(k+1)} = -\nabla f^{(k+1)} + \beta_{k+1} d^{(k)}$
  - Poser  $k = k + 1$ .
- Fin tant que

### 3.5 La méthode de relaxation

La dernière méthode que nous présentons permet de ramener un problème de minimisation dans  $\mathbb{R}^n$  à la résolution successive de  $n$  problèmes de minimisation dans  $\mathbb{R}$ , à chaque itération.

On cherche à minimiser  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . On pose  $X = (x_1, \dots, x_n)$ .

Le principe de la méthode est le suivant : étant donné un itéré  $X^k$  de coordonnées  $(x_1^k, \dots, x_n^k)$ , on fixe toutes les composantes sauf la première et on minimise sur la première :

$$\min f(x, x_2^k, \dots, x_n^k), x \in \mathbb{R}.$$

On obtient ainsi la première coordonnée de l'itéré suivant  $X^{k+1}$  que l'on note  $x_1^{k+1}$  ; on peut, pour effectuer cette minimisation dans  $\mathbb{R}$ , utiliser par exemple la méthode de Newton dans  $\mathbb{R}$ . On recommence ensuite en fixant la première coordonnée à  $x_1^{k+1}$  et les  $n - 2$  dernières comme précédemment. On minimise sur la deuxième et ainsi de suite.

L'algorithme obtenu est :

#### Algorithme de relaxation

1. Initialisation Poser  $k = 0$  et choisir  $X^0 \in \mathbb{R}^n$ .
2. Itération  $k$  Pour  $i$  variant de 1 à  $n$ , on calcule la solution  $x_i^{k+1}$  de

$$\min f(x_1^{k+1}, x_2^{k+1}, \dots, x_i^{k+1}, x, x_{i+1}^{k+1}, \dots, x_n^{k+1}), x \in \mathbb{R}.$$

3. Critère d'arrêt – Si  $\|x_{k+1} - x_k\| < \epsilon$ , stop.
  - Sinon, poser  $k = k + 1$  et aller à 2.

## 3.6 Exercices

*Exercice 3.6.1.* Considérons un problème de minimisation d'une fonction donnée sous la forme suivante :

$$f(x) = \frac{1}{2} \sum_{i=1}^m f_i(x)^2;$$

où  $f_i$  est deux fois différentiable.

- Calculer le gradient et le Hessien de  $f$  en un point  $x$ .
- Donner l'algorithme de Newton ou autre appliqué pour cette fonction  $f$ .
- Donner l'algorithme en prenant

$$f_i(x) = \frac{1}{2} x^T A_i x + b_i^T x + 1.$$

*Exercice 3.6.2.* Soit le problème d'optimisation sans contraintes suivant où la fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  :

$$\min f(x, y) = 3x^2 + 3y^2.$$

Pour chacun de ces algorithmes, appliquer deux itérations de la méthode en prenant comme point de départ  $(x^0, y^0) = (1, 1)$ .

- La méthode du gradient à pas variable ;
- La méthode de Newton ;
- La méthode des gradients conjugués.

*Exercice 3.6.3.* Dans cet exercice, on étudie une méthode de minimisation sans contraintes d'une fonction quadratique de la forme :

$$f(x) = \frac{1}{2} x^T A x + x^T b + c.$$

*Définition 3.6.1.* On dit que les vecteurs  $d_i, i = 0, \dots, n - 1$  sont conjugués par rapport à une matrice symétrique définie positive  $A$  si :

$$\forall i, j \in \{i = 0, \dots, n - 1\}, i \neq j \Rightarrow d_i^T A d_j = 0.$$

• On note  $\{d_i, i = 0, \dots, n - 1\}$  une famille quelconque de vecteurs conjugués par rapport à une matrice symétrique définie positive  $A$  de  $\mathbb{R}^{n \times n}$ . On se donne par ailleurs un vecteur  $b \in \mathbb{R}^n$ .

Soit un vecteur initial  $x_0 \in \mathbb{R}^n$  arbitraire.

A l'étape  $(k+1)$ , supposant construits les vecteurs  $x_1, x_2, \dots, x_k$ , ce qui sous-entend que les vecteurs  $(Ax_l + b)$ ,  $l = 0, \dots, k-1$  sont tous différents de zéro, deux cas peuvent se présenter :

- ou bien  $Ax_k + b = 0$  et l'algorithme est terminé,
- ou bien  $Ax_k + b \neq 0$ , auquel cas, on définit le nombre

$$\rho_k = \frac{d_k^T \cdot (Ax_k + b)}{d_k^T A d_k},$$

puis le vecteur

$$x_{k+1} = x_k - \rho_k \cdot d_k.$$

1. Justifier le choix de  $\rho_k$ .
2. En supposant que  $Ax_k + b \neq 0, \forall k = 1, \dots, n-1$ , exprimer  $x_l$ , pour  $l = 1, \dots, n$  en fonction de  $x_0, (\rho_i, i = 0, \dots, n-1)$  et  $(d_i, i = 0, \dots, n-1)$ .
3. Montrer que  $d_k^T Ax_k = d_k^T Ax_0, \forall k = 0, \dots, n-1$ .
4. - Calculer  $d_k^T (Ax_n + b)$ .  
- Déduire que  $x_n$  est la solution du système  $Ax = -b$ .  
- Que peut-on dire sur la convergence de la méthode.
5. Appliquer l'algorithme du gradient conjugué pour résoudre le problème d'optimisation sans contraintes suivant :

$$\min_{(x_1, x_2) \in \mathbb{R}^2} f(x_1, x_2) = 10x_1 + 10x_2 + 100 - x_1^2 - x_2^2 - 2x_1x_2.$$

*Exercice 3.6.4.* Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$f(x_1, x_2) = x_1^3 - 3x_1(1 + x_2^2).$$

- a)  $f$  admet-elle des points critiques ? des extrêma locaux ? justifier.
- b) Soit  $g : [0, 2\pi] \rightarrow \mathbb{R}$  la fonction définie par

$$g(t) = f(\cos t, \sin t)$$

1. Montrer que  $g'(t) = -6 \sin t \cos(2t)$
2. Trouver les points critiques de  $g$ .
3. Etudier la variation de  $g$ .

4. Déduire les extrêma de  $g$ .

*Exercice 3.6.5.* Pour chacune des affirmations suivantes, spécifier si elles sont vraies ou fausses. Justifier la réponse.

1. Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction deux fois continûment différentiable et soit  $x_k$  tel que  $\nabla^2 f(x_k)$  est définie positive. Alors la direction de Newton  $d_N$  est une direction de descente pour  $f$  en  $x_k$ .
2. Soit  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  telle que

$$q(x) = \frac{1}{2}x^T Qx + g^T x + c,$$

avec  $Q \in \mathbb{R}^{n \times n}$  symétrique définie positive,  $g \in \mathbb{R}^n$  et  $c \in \mathbb{R}$ .

On suppose  $d_1, \dots, d_n$   $Q$ -conjuguées. Si  $Q = \alpha I$  avec  $\alpha > 0$ , alors  $d_1, \dots, d_n$  sont orthogonales.

3. Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  et  $x^*$  tel que  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est semi définie positive. Alors  $x^*$  est un minimum local de  $f$ .

*Exercice 3.6.6.* Soit le problème d'optimisation sans contraintes suivant :

$$\min_{(x,y) \in \mathbb{R}^2} f(x,y) = 3x^2 + 3y^2.$$

1. Appliquer deux itérations de la méthode du gradient à pas fixe en prenant comme point de départ  $(x_0, y_0) = (1, 1)$ .
2. Appliquer deux itérations de la méthode du gradient à pas optimal en prenant comme point de départ  $(x_0, y_0) = (1, 1)$ .

*Exercice 3.6.7.* Appliquer un algorithme pour résoudre le problème d'optimisation sans contraintes suivant :

$$\min_{(x_1, x_2) \in \mathbb{R}^2} f(x_1, x_2) = 10x_1 + 10x_2 + 100 - x_1^2 - x_2^2 - 2x_1x_2.$$

*Exercice 3.6.8.* Soit  $A \in \mathcal{M}_n(\mathbb{R})$  et  $J$  la fonction définie de  $\mathbb{R}^n$  dans  $\mathbb{R}$  par  $J(x) = e^{\|Ax\|^2}$ , où  $\|\cdot\|$  désigne la norme euclidienne sur  $\mathbb{R}^n$ .

1. Montrer que  $J$  admet un minimum.
2. On suppose que la matrice  $A$  est inversible, montrer que ce minimum est unique.
3. Ecrire l'algorithme du gradient à pas optimal pour la recherche de ce minimum.
4. A quelle condition suffisante cet algorithme converge-t-il ?