

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET  
POPULAIRE

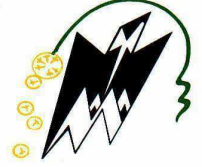
*Ministère de l'Enseignement Supérieur Et de la  
Recherche scientifique*



*Université MOULOUD MAMMERY, TIZI-OUZOU*

*Faculté des Sciences*

*Département de Mathématiques*



## Mémoire de fin d'étude

*Présenté pour l'obtention du diplôme de Master en mathématiques*

*Option : Mathématiques appliquées à la gestion*

### Optimisation globale avec applications

*Présenté par : - Louakhche Houssem  
- Hamlat Islem*

**Devant le Jury composé de:**

*Président : Mr. kourat*

*Rapporteur : M. Leslous*

*Examineur : Mr. Chebbah*

*Année universitaire : 2018 - 2019*

# Table Des Matière

<b>Introduction generale</b> .....	3
<b>1 Définitions et Notations</b>	
1.1 Introduction .....	5
1.2 Description d'un problème d'optimisation .....	5
1.2.1 Formulation mathématique .....	5
1.2.2 Minimum local. ....	6
1.2.3 Minimum global. ....	6
1.3 Outils fondamentaux. ....	8
1.3.1 Différentiabilité. ....	8
1.3.1.1 Dérivée partielle. ....	8
1.3.1.2 Gradient. ....	9
1.3.1.3 Matrice Jacobienne. ....	10
1.3.1.4 Matrice Hessienne. ....	11
1.3.1.5 Dérivée directionnelle. ....	13
1.3.1.6 Fonction différentiable. ....	14
1.3.1.7 Pente. ....	14
1.3.1.8 courbure. ....	15
1.3.1.9 Direction de descente. ....	15
1.3.1.10 Théorème de la Plus forte descente. ....	16
1.3.1.11 Théorème de la plus forte pente. ....	17
1.3.1.12 Développement de Taylor. ....	18
1.3.2 Convexité. ....	19
1.3.2.1 Ensemble convexe. ....	19
1.3.2.2 Propriétés des ensembles convexes. ....	21
1.3.2.3 Fonction convexe. ....	21
1.3.2.4 Problème convexe. ....	22
1.3.3 Caractérisation différentielle de la convexité. ....	23
1.3.3.1 caractérisation de la convexité. ....	23
1.3.3.2 caractérisation de la convexité stricte. ....	25
1.3.3.3 Matrice définie positive. ....	25
1.4 Conditions d'optimalité. ....	26
1.4.1 Conditions nécessaires d'optimalité. ....	26
1.4.1.1 Conditions nécessaires d'optimalité du premier ordre. ....	26
1.4.1.2 Conditions nécessaires d'optimalité du second ordre. ....	27
1.4.2 Conditions suffisantes d'optimalité. ....	27

<b>2</b>	<b>Méthode de séparation et évaluation (B&amp;B)</b>	
2.1	Introduction.	29
2.2	Présentation de la méthode Séparation et évaluation (B&B)	29
2.2.1	Historique.	29
2.2.2	Définition de la méthode.	30
2.3.	La procédure de séparation	31
2.3.1	Démarche de la séparation.	31
2.3.2	Règles de branchement.	32
2.4	Procédure d'évaluation	33
2.4.1	Borne supérieure	33
2.4.2	Borne inférieure	33
2.4.3	Calcul de la borne inférieure.	34
2.4.4	Calcul de la borne supérieure	35
2.4.5	Conditions d'élimination.	35
2.4.5.1	Éliminé par optimalité	35
2.4.5.2	Éliminé par borne	36
2.4.5.3	Éliminé par infaisabilité	37
2.5	La stratégie de parcours.	37
2.5.1	Profondeur D'abord	37
2.5.2	Largeur D'abord	38
2.5.3	Meilleur D'abord	38
2.6	Efficacité de l'algorithme B&B.	38
2.6.1	Choix d'une stratégie de recherche.	38
2.6.2	Choix de la borne inférieure.	39
2.6.3	Choix de la borne supérieur.	39
2.6.4	Règles de dominance	39
2.6.5	Choix du critère d'évaluation.	39
2.7	Méthode séparation et évaluation sur des exemples.	39
<b>3</b>	<b>La méthode d'Approximation Extérieure</b>	47
3.1	Introduction.	47
3.2	Le principe de la méthode d'Approximation Extérieure	47
3.3	L'algorithme de base de la méthode d'Approximation Extérieure	48
3.4	La convergence de la méthode d'Approximation Extérieure	49
3.4.1	Approximation polyèdre.	49
3.5	Construction de $l_k(\cdot)$	51
3.6	Exemple d'application d'AE	53
<b>4</b>	<b>Programmation dans LINGO.</b>	55
4.1	Introduction.	55
4.2	Présentation du logiciel.	55
4.3	Résolution des exemples sur LINGO.	58
	<b>Conclusion et perspective</b>	63
	<b>Bibliographie</b>	64

## Introduction général

Nous faisons tous de l'optimisation. Dans notre vie quotidienne, nous cherchons à optimiser notre temps de travail, nos espaces de rangement, ou encore le trajet que nous aurons à parcourir pour nous rendre quelque part, etc. Nous recherchons tous une meilleure solution aux problèmes qui jalonnent notre existence. De manière générale, l'optimisation va donc consister à trouver cette meilleure solution. [8].

Comme nous le rappelle l'adage populaire selon lequel: "les mathématiques permettent de mettre le monde en équation", il peut être tracé en parallèle entre l'optimisation quotidienne et celle, plus technique que l'on retrouve en science. En mathématique, la meilleure solution se recherche au sein d'un domaine initial. Cette solution est souvent soumise à des contraintes qui correspondent à des obligations ou des souhaits à respecter. [1][6][7].

L'optimisation mathématique va consister à rechercher dans le domaine initial une solution qui maximise ou minimise une fonction objectif tout en respectant des contraintes. Pour un domaine continu, on distingue classiquement deux types d'optimisation :

**L'optimisation locale** recherche une solution qui est la meilleure localement, c'est-à-dire que dans son voisinage aucune solution n'est meilleure qu'elle. Cette solution est appelée un optimum local. [6]

**L'optimisation globale** recherche quant à elle la meilleure solution du domaine en entier, c'est-à-dire que dans tout le domaine il n'existe aucune solution qui lui soit meilleure tout en respectant les contraintes. Cette solution est appelée l'optimum global. [6].

L'intérêt de l'optimisation globale par rapport à l'optimisation locale est épatant. Elle garantit en effet que personne ne peut avoir une solution meilleure que celle trouvée. Or, pour une entreprise, cette information a son importance, car la différence entre la solution globale et une solution locale est bien souvent significative. Mais l'intérêt n'est pas que compétitif. Dans de nombreux problèmes, l'optimum global est la seule solution mathématique correspondant à une réalité physique. [2][10].

De nos jours, afin de résoudre des problèmes d'optimisation globale avec contraintes, de nombreuses stratégies algorithmiques s'avèrent disponibles. Pour guider le choix de la meilleure stratégie à utiliser, il est nécessaire de regarder:

- 1-la taille du problème.
- 2-les propriétés de la fonction objectif et les contraintes (continuité, différentiabilité, linéarité, convexité...).
- 3-le temps disponible pour résoudre le problème.

Ces dernières années de nombreux travaux ont été effectués sur les différentes approches (tel que les méthodes multi-start, les algorithmes révolutionnaires, Les méthodes métaheuristiques et les méthodes déterministes globales ). [6].

Nous allons nous intéresser aux méthodes déterministes ,parmi elles la méthode Branch and Bound qui nous offre la certitude d'obtenir l'optimum global (si celui ci existe ).Ce type d'algorithme dit de séparation et évaluation qui est plus connue sous le nom Branch and Bound qui peut théoriquement résoudre n'importe quel problème. [5][6].

Dans ce mémoire notre attention est portée sur l'étude de certaines méthodes d'optimisation global plus exactement la méthode branch and bound et la méthode d'approximation extérieur.

Le premier chapitre traite des notions mathématiques fondamentales à maîtriser avant de s'intéresser à la résolution de tout problème d'optimisation :

- la description mathématique d'un problème d'optimisation
- la notion de solution locale et quelques éléments d'analyse convexe.

Ainsi, nous étudierons quelques résultats théoriques (existence, unicité de solutions) sur l'optimisation sans contrainte(s) [1][3][4].

Le deuxième chapitre présentera la méthode branch and bound (B&B) qui est très utilisée pour la résolution d'un grand nombre de problèmes mathématiques, avec leurs différentes structures. Cette méthode engendre deux suites convergentes des bornes inférieure et supérieure de la valeur minimale de la fonction objectif du problème donné. [5][6][2].

Le troisième chapitre est consacré quant a lui a l'étude de la méthode d'approximation extérieur. son principe est de faire des coupes linéaires sur l'ensemble réalisable, en se rapprochant de plus en plus de la solution optimale du problème donné. [11].

Le dernier chapitre sera consacré pour résoudre un exemple en utilisant le logiciel LINGO. Quelques notions de base seront données sur ce logiciel au début de ce chapitre. [12].

Nous terminerons notre travail par une conclusion générale.

# Chapitre 1

## Définitions et notation

### 1.1 Introduction

L'optimisation globale est une branche des mathématiques, le domaine de sa recherche est considéré beaucoup plus riche ces dernières années grâce aux ordinateurs développés par la nouvelle technologie qui peuvent résoudre des problèmes d'optimisation de taille plus grande.

Du point de vue complexité, le problème d'optimisation globale est NP difficile. Les méthodes d'optimisation globale s'intéressent à la recherche d'un optimum (maximum ou minimum) global d'une fonction définie sur par exemple un compact  $K$ , et non à la recherche d'un optimum local que les méthodes classiques permettent de le trouver dans un voisinage réalisable, et c'est ça qui fait la différence entre les deux méthodes.

L'intérêt de l'étude de cette classe de problèmes d'optimisation est qu'elle englobe un domaine très vaste d'applications réelles. [1]

### 1.2 Description d'un problème d'optimisation

#### 1.2.1 Formulation mathématique

On considère le problème d'optimisation suivant:

$$\begin{cases} \min f(x) \\ c_E(x) = 0 \quad E = 1, 2, \dots, p \\ c_I(x) \leq 0 \quad I = 1, 2, \dots, q \\ x \in \mathbb{R}^n \end{cases} \quad (1.1)$$

- La fonction  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  porte divers noms: fonction-coût ou simplement cout, fonction-objectif ou simplement objectif, critère,... [3].
- L'ensemble des contraintes  $X$  est défini par des égalités et des inégalités:

$$X = \{x \in \mathbb{R}^n \mid c_E(x) = 0, c_I(x) \leq 0\},$$

Appelé l'ensemble des points admissible ou réalisables.

-La fonction  $c_E \mathbb{R}^n \rightarrow \mathbb{R}^p$  représente p contraintes d'égalité.

-La fonction  $c_I \mathbb{R}^n \rightarrow \mathbb{R}^q$  représente q contraintes d'inégalité.

On va considérer deux cas pour  $X$

- 1)  $X \subset \mathbb{R}^n$  le problème dit problème d'optimisation avec contraintes.
- 2)  $X = \mathbb{R}^n$  le problème dit problème d'optimisation sans contraintes.

Dans ce dernier cas le problème d'optimisation sans contraintes s'écrit:

$$\begin{cases} \min f(x) \\ \text{sous la contrainte} \\ x \in \mathbb{R}^n \end{cases} \quad (1.2)$$

Si  $x^*$  est une solution du problème (1.2), on dit que :

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

Résoudre le problème (1.2) revient à chercher des points de minimum local ou global.

### 1.2.2 Minimum local

$x^*$  est un minimum local (ou relatif) de  $f$  sur  $X$  si

$$\exists \varepsilon > 0 \text{ tel que } f(x^*) \leq f(x), \forall x \in X \text{ avec } \|x - x^*\| < \varepsilon,$$

Où  $\|v\|$  désigne la norme du vecteur  $v$ .

$x^*$  minimum local strict de  $f$  sur  $X$  si

$$\exists \varepsilon > 0 \text{ tel que } f(x^*) < f(x), \forall x \in X \text{ avec } \|x - x^*\| < \varepsilon.$$

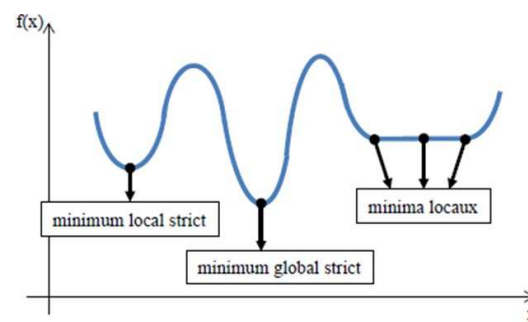
### 1.2.3 Minimum globale

$x^*$  est un minimum globale (ou absolu) de  $f$  sur  $X$  si

$$f(x^*) \leq f(x), \forall x \in X.$$

$x^*$  minimum global strict de  $f$  sur  $X$  si

$$f(x^*) < f(x), \forall x \in X.$$



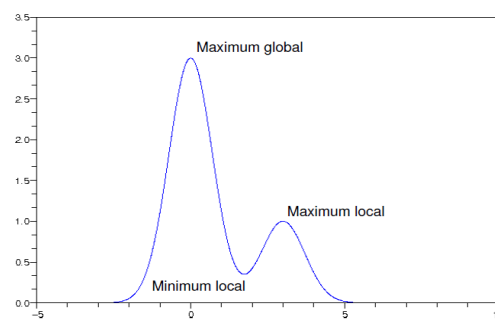
**FIGURE1.1:Exemple d'une fonction monodimensionnelle**

Les maxima locaux et globaux sont définis de manière similaire. Notons que  $x^*$  est un maximum local (respectivement globale) de la fonction  $f$  sur l'ensemble  $X$  si  $x^*$  est un minimum local (respectivement global) de  $-f$  sur  $X$ .

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  $-1$ . [4][5]

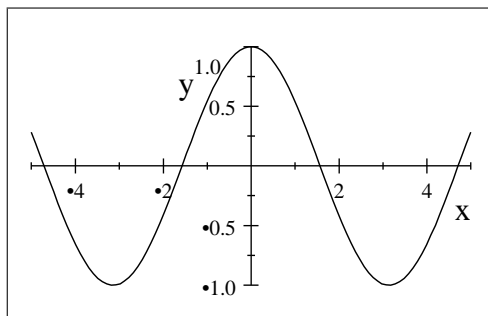
### Exemples :

1) Soit  $f_1(x) = 3 \exp(-x^2) + \exp(-(x-3)^2)$



**FIGURE1.2:Exemple de minima et maxima locaux  
Et globaux pour la fonction  $f_1$**

2) Soit  $f_2(x) = \cos(x)$  ou  $x \in X$



**FIGURE 1.3: Exemple de minima et maxima globaux pour la fonction  $f_2$**

Il existe une infinité de minima et maxima globaux.

### 1.3 Outils fondamentaux

Il y a deux outils mathématiques fondamentaux pour l'analyse des problèmes d'optimisation : la différentiabilité et la convexité [6].

#### 1.3.1 Différentiabilité

Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$ . On se place dans  $\mathbb{R}^n$ ,  $n < \infty$ , considère comme un espace vectoriel normé muni de la norme euclidienne notée  $\|\cdot\|$ . [7][8][9]

##### 1.3.1.1 Dérivée partielle [7]

Soit  $f$  fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}$ . La fonction notée  $\forall i f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ ,  
Egalement notée est appelée  $i^{\text{ième}}$  dérivée partielle de  $f$  est définie par :

$$\lim_{t \rightarrow 0} \left( \frac{f(x_1, \dots, x_i + t, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{t} \right)$$

Cette limite peut ne pas exister.

### Cas particulier n=1

On note :

$$f'(x) = \frac{\partial f(x)}{\partial x} = \lim_{t \rightarrow 0} \left( \frac{f(x+t) - f(x)}{t} \right) = \lim_{y \rightarrow x} \left( \frac{1}{y-x} \right) [f(y) - f(x)].$$

**Remarque 1.1 :** Si une fonction réelle d'une seule variable est dérivable en un point alors elle est continue en ce point, dans le cas d'une fonction de plusieurs variables l'existence des dérivées partielles en un point n'implique pas la continuité en ce point.[7]

Contre-exemple: On prend l'exemple de la fonction

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2}, & \text{si } (x, y) \neq (0, 0) \\ 0 & , \text{ si } (x, y) = (0, 0) \end{cases}$$

Qui n'est pas continue à l'origine en effet  $\lim_{(x,y) \rightarrow f(0+0)} f(x+y) = \lim_{r \rightarrow 0} \cos \theta \sin \theta$  bien que les dérivées partielles  $\frac{\partial f}{\partial x}(0,0) = 0$  et  $\frac{\partial f}{\partial y}(0,0) = 0$  existent.

### 1.3.1.2 Gradient

Le gradient joue un rôle essentiel dans le développement et l'analyse des algorithmes d'optimisation.

**Définition 1.1 [10] :**

La fonction notée  $\nabla f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$  est appelée Gradient de  $f$  au point

$x = (x_1, \dots, x_n)^t$  si toutes les dérivées partielles existent, on définit le gradient par la formule suivant:

$$\nabla f(x) = \left( \frac{\partial f(x)}{\partial x_i} \right)_{i=1, \dots, n} = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

On note dans  $\mathbb{R}^n$ ,  
 $\nabla f(x) = f'(x)$

### Exemples :

1) Soit  $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$  (fonction de Rosen-brock).

Le gradient de  $f$  est donné par:

$$\nabla f(x_1, x_2) = \begin{pmatrix} -400(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}$$

2) Soit  $f(x) = 4x^2 + \exp(2x) - \ln(x)$

Le gradient de  $f$  est donné par :

$$\nabla f(x) = f'(x) = 8x + 2 \exp(2x) - \frac{1}{x}$$

### 1.3.1.3 Matrice Jacobienne [8]

La matrice Jacobienne d'une fonction  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}$  en  $x$  notée  $J_f(x)$  qui est un élément de  $M_{1,n}(\mathbb{R})$  définie par:

$$(J_f(x))_i = \frac{\partial f}{\partial x_i}(x), \forall i = 1, \dots, n$$

### Remarque 1.2 :

Le gradient de  $f$  en  $x$  est défini comme la transposée de la matrice Jacobienne de  $f$  en  $x$ :

$$\nabla f(x) = (J_f(x))^T \in \mathbb{R}^n$$

Rappelons la formule:

$$\frac{\partial f}{\partial x}(h) = \langle \nabla f(x), h \rangle \forall x \in \Omega, \forall h \in \mathbb{R}^n$$

### 1.3.1.4 Matrice Hessienne

On appelle Hessienne de  $f$  la matrice carrée de  $M_n(\mathbb{R})$

$$H(x) = (\nabla^t f)(x) = \nabla^2 f(x) = \left( \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n}$$

Alors

$$H(x) = \begin{pmatrix} \left( \frac{\partial^2 f(x)}{\partial^2 x_1} \right) & \cdots & \left( \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \right) \\ \vdots & \ddots & \vdots \\ \left( \frac{\partial^2 f(x)}{\partial x_n \partial x_1} \right) & \cdots & \left( \frac{\partial^2 f(x)}{\partial^2 x_n} \right) \end{pmatrix}$$

On note dans  $\mathbb{R}^n \times \mathbb{R}^n$ :

$$H(x) = f''(x) \quad [8]$$

#### Remarque 1.3 :

Si  $f$  de classe  $C^2$  alors  $\nabla^2 f(x)$  est une matrice symétrique  $\forall x \in \Omega$

(Théorème de Schwarz).

#### Exemples :

1) Pour la fonction de Rosenbrock.

La hessienne de  $f$  est donnée par:

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} -400(x_2 - 3x_1) + 2 - 400x_1 & \\ -400x_1 & 200 \end{pmatrix}$$

2) Soit  $f(x) = 4x^2 + \exp(2x) - \ln(x)$

La hessienne de  $f$  est donnée par :  $\nabla^2 f(x) = f''(x) = 8 + 4 \exp(2x) - \frac{1}{x^2}$

#### Proposition 1.1 [3] :

a. Si  $f$  de classe  $C^2$  alors :

$$\nabla^2 f(x) = J_{\nabla f}(x) = \nabla J_f(x) \quad \forall x \in \Omega$$

## Chapitre 1: Définitions et Notations

la matrice Hessienne de  $f$  est le Jacobien du gradient de  $f$  ou le gradient de la Jacobienne de  $f$ .

b. Nous avons :

$$\nabla^2 f(x) = \nabla \langle \nabla f(x), h \rangle = \forall x \in \Omega, \forall h \in \mathbb{R}^n$$

Où le premier gradient dans le membre de droite de l'égalité est considéré par rapport à la variable  $x$ .

### Exemples :

1) Si  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  est une fonction constante alors  $\nabla f = \nabla^2 f = 0$ .

2) Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  définie par :  $f(x) = \langle a, x \rangle \forall x \in \mathbb{R}^n$

Où  $a \in \mathbb{R}^n$  est un vecteur donné (c'est à dire,  $f$  est une fonction linéaire), alors on calcule facilement  $\frac{\partial f}{\partial x_k} = a_k$ , donc

$$\nabla f = a$$

Le gradient est constant, Ceci nous donne

$$\nabla^2 f = 0$$

### Corollaire 1.1 [3] :

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  définie par  $f(x) = \langle Ax, x \rangle \forall x \in \mathbb{R}^n$ ,

Où  $A \in M_n$  est une matrice carrée réelle, de taille  $n$  (c'est à dire,  $f$  est la fonction quadratique associée à la matrice  $A$ ). Alors pour un  $p \in \{1, 2, \dots, n\}$  fixé, on peut écrire

$$f(x) = \sum_{l,j=1}^n A_{lj} x_l x_j = A_{pp} x_p^2 + \sum_{j=1, j \neq p}^n A_{pj} x_p x_j + \sum_{i=1, i \neq p}^n A_{ip} x_i x_p + \sum_{j,i=1, i \neq p, j \neq p}^n A_{ij} x_i x_j$$

Ce qui nous donne :

$$\frac{\partial f}{\partial x_p} = 2A_{pp} x_p + \sum_{j=1, j \neq p}^n A_{pj} x_j + \sum_{j=1, j \neq p}^n A_{ip} x_i + \sum_{j=1}^n A_{pj} x_j + \sum_{i=1}^n A_{ip} x_i = A(x)_p (A^T x)_p$$

Nous avons donc obtenu :

$$\nabla f(x) = (A + A^T) x, \quad \forall x \in \mathbb{R}^n$$

## Chapitre 1: Définitions et Notations

---

on peut aussi écrire :

$$\frac{\partial f}{\partial x_i}(x) = \sum_{k=1}^n (A + A^T)_{ik} x_k, \forall i = 1, \dots, n$$

on a alors :

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = (A + A^T)_{ij}, \forall i, j = 1, \dots, n$$

c'est à dire

$$\nabla^2 f(x) = A + A^T, \forall x \in \mathbb{R}^n$$

Donc la hessienne de  $f$  est constante. [7]

### 1.3.1.5 Dérivée directionnelle

On appelle dérivée directionnelle de  $f$  dans la direction  $d \in \mathbb{R}^n$  au point  $x$ , notée  $\delta f(x; d)$ , la limite ( éventuellement  $\pm \infty$  ) du rapport:

$$\frac{f(x + hd) - f(x)}{h} \text{ lorsque } h \text{ tend vers } 0$$

Autrement dit:

$$\delta f(x, d) = \lim_{h \rightarrow 0} \frac{f(x + hd) - f(x)}{h}$$

De plus, lorsque le gradient existe, la dérivée directionnelle est le produit scalaire entre le gradient de  $f$  et la direction  $d$ , i. e:

$$\nabla f(x, d) = \nabla^T f(x) d \quad [10]$$

#### Remarque 1.4 :

Si  $\|d\| = 1$  la dérivée directionnelle est le taux d'accroissement de  $f$  dans la direction  $d$  au point  $x$ .

-Le taux d'accroissement est maximal dans la direction du gradient.

-Le gradient indique la direction de la plus grande pente.

## Chapitre 1: Définitions et Notations

---

### Exemple :

Soit  $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

et soit  $d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$

La dérivée directionnelle de  $f$  dans la direction  $d$  est :

$$(d_1 d_2) \nabla f(x_1, x_2) = d_1 (-400(x_2 - x_1^2)x_1 - 2(1 - x_1)) + 200d_2(x_2 - x_1^2),$$

Où  $\nabla f(x_1, x_2)$  est donné par :

$$\nabla f(x_1, x_2) = \begin{pmatrix} -400(x_2 - x_1^2)x_1 - 2(1 - x_1) \\ -200(x_2 - x_1^2)x_1 \end{pmatrix}$$

### Remarque 1.5 :

Pour tout  $x \in \Omega$  et  $h \in \mathbb{R}^n$  on note

$$\frac{\partial f}{\partial d}(x) = \lim_{h \rightarrow 0} \frac{1}{h} [f(x + hd) - f(x)] = g'(0)$$

(C'est la dérivée directionnelle de  $f$  en  $x$  de direction  $d$ ) où on a noté  $g(h) = f(x + hd)$ .

### 1.3.1.6 Fonction différentiable

Soit  $f$  fonction continue de  $\mathbb{R}^n$  dans  $\mathbb{R}$ . Si pour tout  $d \in \mathbb{R}^n$  la dérivée directionnelle de  $f$  dans la direction  $d$  existe, alors la fonction  $f$  est dite différentiable. [7]

### 1.3.1.7 pente

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable. Soient  $x, d \in \mathbb{R}^n$ . La quantité:

$$\frac{d^T \nabla f(x)}{\|d\|}$$

Représente la pente de la fonction  $f$  en  $x$  dans la direction  $d$ . [7]

### Remarque 1.5

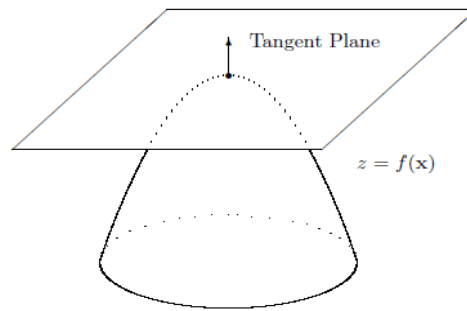
Dans le cas d'une fonction réelle d'une variable  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ , sa dérivée au point  $A \in$

$Df$ , notée  $f'(A)$ , représente la pente de la droite tangente à la courbe au point  $(A, f(A))$ .

Dans le cas d'une fonction réelle de deux variables, la dérivée n'a pas de sens et la représentation graphique de la fonction est une surface de l'espace : il existe une infinité de droite tangente la surface en un point ; en fait il y a un plan tangent.

## Chapitre 1: Définitions et Notations

---



### 1.3.1.8 courbure [7]

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction deux fois différentiable. Soient  $x, d \in \mathbb{R}^n$  La quantité :

$$\frac{d^T \nabla^2 f(x)}{\|d\|}$$

représente la courbure de la fonction  $f$  en  $x$  dans la direction  $d$ .

La dérivée directionnelle donne des informations sur la pente de la fonction dans la direction  $d$ , tout comme la dérivée donne des informations sur la pente des fonctions à une variable. En particulier,

- Si  $\delta f(x; d) > 0$  alors  $f$  est croissante dans la direction  $d$ .
- Si  $\delta f(x; d) < 0$  alors  $f$  est décroissante dans la direction  $d$ .

Dans ce dernier cas, on dira que  $d$  est une direction de descente de  $f$ .

### 1.3.1.9 Direction de descente [7]

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

$$f(x + td) < f(x), \forall 0 < t < \eta$$

### Proposition 1.2:

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable. Soient  $x \in \mathbb{R}^n$  tel que  $\nabla f(x) \neq 0$  et  $d \in \mathbb{R}^n$  un vecteur

## Chapitre 1: Définitions et Notations

---

- Si  $d^T \nabla f(x) < 0$  alors  $d$  est une direction de descente.
- Si  $d$  est une direction de descente alors  $d^T \nabla f(x) \leq 0$

### Preuve :

Soit  $d$  tel que  $d^T \nabla f(x) < 0$ . On écrit le développement de Taylor-Young de  $t \mapsto f(x + td)$  pour  $t$  assez petit,

$$f(x + td) = f(x) + t \nabla f(x)^T d + t \varepsilon(t), \text{ avec } \varepsilon(t) \xrightarrow{t \rightarrow 0} 0$$

Comme  $-\nabla f(x)^T d > 0$ , il existe  $\eta > 0$  tel que si  $|t| \leq \eta$  alors  $|\varepsilon(t)| < -\nabla f(x)^T d$

D'où  $\forall t \in ]0, \eta], f(x + td) - f(x) = t[\nabla f(x)^T d + \varepsilon(t)] < 0$ .

Ce qui montre la première partie de la proposition.

Soit maintenant  $d$  une direction de descente, supposons que  $\nabla f(x)^T d > 0$

en appliquant le même raisonnement que ci-dessus, on montre qu'il existe un  $\eta$  tel que pour tout  $t$  tel que  $t \leq \eta$  on a  $\varepsilon(t) > -\nabla f(x)^T d$  et ainsi  $f(x + td) - f(x) > 0$ , ce qui contredit le fait que  $d$  soit une direction de descente.

### 1.3.1.10 Théorème de la Plus forte descente

#### Proposition 1.3 [7] :

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable. Soient  $x \in \mathbb{R}^n$  et  $d^* = -\nabla f(x)$ :

Alors pour toute direction  $d$  de norme constante égale à  $\|d\| = \|\nabla f(x)\|$  on a:

$$(-\nabla f(x))^T \nabla f(x) \leq d^T \nabla f(x) \quad (1.3)$$

Ainsi, la direction  $d^* = -\nabla f(x)$  est appelée "direction de plus forte descente".

### Preuve :

Soit  $d \in \mathbb{R}^n$  une direction quelconque de norme  $\|d\| = \|\nabla f(x)\|$ .

On a alors :

$$(-d)^T \nabla f(x) \leq \| -d \| \|\nabla f(x)\|. \text{ (d'après l'inégalité de Cauchy-Schwarz)}$$

$$\leq \|\nabla f(x)\|^2 = \nabla f(x)^T \nabla f(x). \text{ (puisque } \|d\| = \|\nabla f(x)\| \text{)}$$

On en déduit donc (1.3)

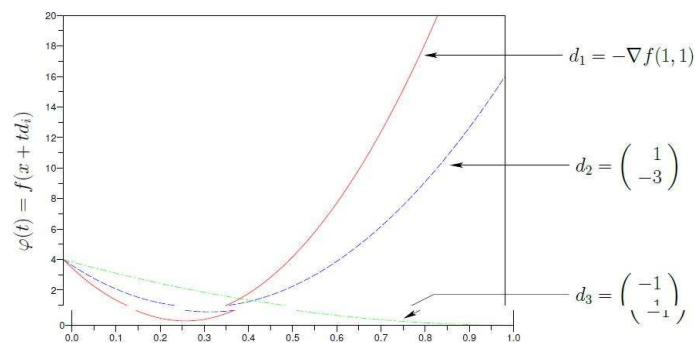
## Chapitre 1: Définitions et Notations

**Exemple :**

Soit  $f(x) = \frac{1}{2}x_1^2 + 2x_2^2$  et soit  $x = (1,1)^T$ .

Nous considérons trois directions

$$d_1 = -\nabla f(x) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, d_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, d_3 = \begin{pmatrix} -1 \\ -4 \end{pmatrix}$$



**FIGURE 1.4: Allure de la fonction  $f$**

La dérivée directionnelle de  $f$  dans chacune de ces directions sont:

$$d_1^T \nabla f(x) = -17, d_2^T \nabla f(x) = -5, d_3^T \nabla f(x) = -11$$

Sans surprise, si l'opposé du gradient correspond à la plus forte descente, le gradient correspond lui à la plus forte pente de montée :

### 1.3.1.11 plus forte pente

**Théorème 1.1** [7]

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable. Soient  $x \in \mathbb{R}^n$  et  $d^* = \nabla f(x)$ :

Alors pour tout  $d \in \mathbb{R}^n$  tel que  $\|d\| = \|\nabla f(x)\|$ , on a:

$$d^T \nabla f(x) \leq d^{T*} \nabla f(x) = \nabla^T f(x) \nabla f(x)$$

Ainsi, la direction  $d^* = \nabla f(x)$  est appelée "direction de plus forte pente".

## Chapitre 1: Définitions et Notations

---

**Exemple :**

Soit  $f(x) = \frac{1}{2}x_1^2 + 2x_2^2$  et soit  $x = (1,1)^T$ .

Nous considérons trois directions

$$d_1 = \nabla f(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, d_2 = \begin{pmatrix} -1 \\ -3 \end{pmatrix}, d_3 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

La dérivée directionnelle de  $f$  dans chacun de ces directions vaut :

$$d_1^T \nabla f(x) = 17, d_2^T \nabla f(x) = 5, d_3^T \nabla f(x) = 11$$

### 1.3.1.12 Développement de Taylor

La formule de Taylor est un outil important. Nous la rappelons dans le cas général.[8]

Soit  $\Omega \in \mathbb{R}$  ouvert,  $f: \Omega \rightarrow \mathbb{R}, a \in \Omega, h \in \mathbb{R}^n$ . Alors:

1) Si  $f \in C^1(\Omega)$ , alors :

#### 1-Formule de Taylor à l'ordre 1 avec reste intégral

$$f(a+h) = f(a) + \int_0^1 \langle \nabla f(a+th), h \rangle dt.$$

#### 2-Formule de Taylor-Maclaurin à l'ordre 1

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle.$$

#### 3-Formule de Taylor-Young à l'ordre 1

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + o(\|h\|)$$

2) Si  $f \in C^2(\Omega)$ , alors

#### 1-Formule de Taylor à l'ordre 2 avec reste intégral

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \int_0^1 (1-t) \langle \nabla^2 f(a+th), h \rangle dt.$$

## Chapitre 1: Définitions et Notations

---

### 2-Formule de Taylor-Maclaurin à l'ordre 2

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a + \theta h) h, h \rangle, \text{ avec } 0 < \theta < 1.$$

### 3-Formule de Taylor-Young à l'ordre 2

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a + \theta h) h, h \rangle + o(\|h\|^2)$$

### Théorème 1.2 [7] :

Soit  $f: \Omega \rightarrow \mathbb{R}$ , de classe  $C^{n+1}(\Omega)$ . Si le segment  $[a, a + h]$  est contenu dans  $\Omega$ , on a :

$$f(a + h) = f(a) + f'(a)h + \frac{1}{n!} f^{(n)}(a)(h)^n + \int_0^1 \frac{(1-t)^n}{n!} f^{(n+1)}(a + th)(h)^{n+1} dt.$$

### 1.3.2 Convexité

La convexité est à la base une propriété géométrique.

On voit assez bien ce qu'est un objet convexe dans un espace à deux ou trois dimensions.

Nous allons maintenant montrer comment cette propriété peut aussi s'appliquer aux fonctions de  $\mathbb{R}^n$  dans  $\mathbb{R}$ . [4]

#### 1.4.2.1 Ensemble convexe

Soit l'ensemble  $C \subseteq \mathbb{R}^n$ ,  $C$  est convexe si:

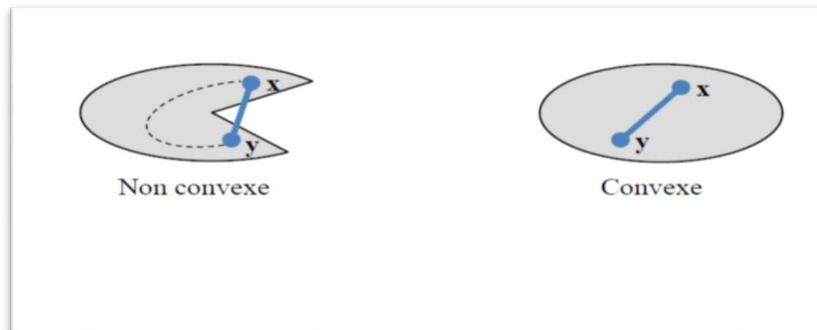
$$(1-t)x + ty \in C; \forall x, y \in C, \forall t \in [0,1]$$

Où

$$x, y \in C \Rightarrow [x, y] \subset C$$

## Chapitre 1: Définitions et Notations

---



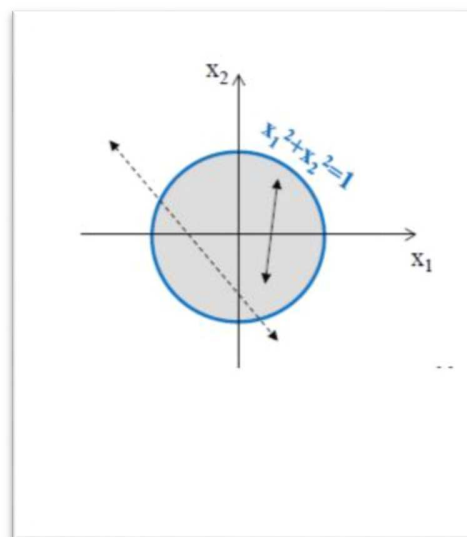
**FIGURE 1.5: Exemple d'un ensemble convexe et non convexe**

### Exemples :

1)  $\mathbb{R}^n$  est convexe.

2) Ensemble:  $X = \{(x, y) / x_1^2 + x_2^2 \leq 1\}$  est convexe.

Ensemble:  $Y = \{(x_1, x_2) / x_1^2 + x_2^2 \geq 1\}$  est non convexe



**FIGURE 1.6 : L'ensemble X et Y**

## Chapitre 1: Définitions et Notations

### 1.3.2.2 Propriétés des ensembles convexes

Soit  $x_1, x_2, \dots, x_k \in \mathbb{R}^n$  et  $t_j$  telle que  $t_j \geq 0$  et  $\sum_{j=1}^k t_j = 1$ .

Toute expression de la forme :

$$\sum_{j=1}^k t_j x_j$$

S'appelle combinaison convexe des points  $x_j$  ou barycentre.

-Tout entier est un ensemble convexe, de même qu'un singleton  $\{a\}$ .

-Soit la famille  $\{C_i\}_{1,\dots,p}$  d'ensemble convexes et  $S = \bigcap_{i=1}^p C_i$  Alors  $S$  c'est convexe.

### 1.3.2.3 Fonction convexe

Soit  $C \subseteq \mathbb{R}^n$  un ensemble convexe et  $f$  une fonction de  $C$  dans  $\mathbb{R}$ ,  
 $f$  est dit convexe sur  $C$  si,

$$\forall x, y \in C, \forall t \in [0,1] \\ f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

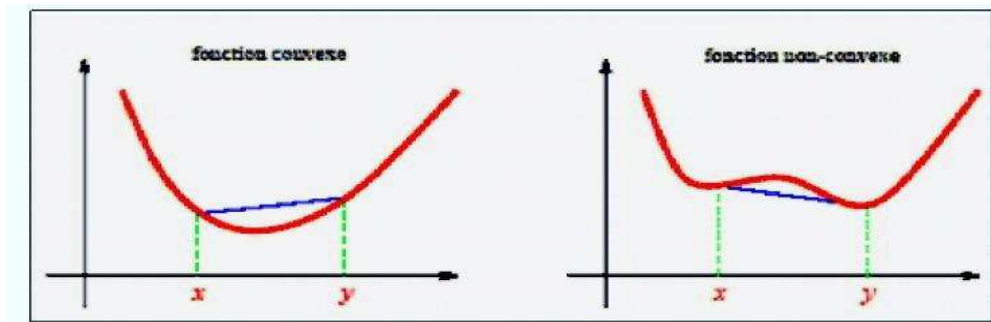


FIGURE 1.7: Exemple de fonction convexe et non convexe

## Chapitre 1: Définitions et Notations

---

$f$  est dit strictement convexe sur  $C$  si :

$$\begin{cases} x \neq y \\ \forall t \in ]0,1[ \\ f((1-t)x + ty) < (1-t)f(x) + tf(y) \end{cases}$$

### Propositions 1.4 :

Soit  $U \subset \mathbb{R}^n$  un ensemble convexe,  $p \in \mathbb{N}^*$ ,  $f_1, f_2, \dots, f_p: U \rightarrow \mathbb{R}$  des fonctions convexes et  $\gamma_1, \gamma_2, \dots, \gamma_n$  des constantes strictement positives.

On pose  $f(x) = \gamma_1 f_1 + \gamma_2 f_2 + \dots + \gamma_p f_p$ , on a :

- 1- La fonction  $f$  est convexe (donc toute combinaison linéaire avec des coefficients Strictement positifs de fonctions convexes est convexe).
- 2-Si au moins l'une des fonctions  $f_1, f_2, \dots, f_p$  est strictement convexe alors  $f$  est strictement convexe.
- 3-Si au moins l'une des fonctions  $f_1, f_2, \dots, f_p$  est fortement convexe alors  $f$  est fortement convexe.

La convexité est une notion globale, qui va donner des informations sur le caractère global d'un point de minimum.

### 1.3.2.4 Problème convexe

Soit  $X \subset \mathbb{R}^n$  convexe et  $f : X \rightarrow \mathbb{R}$  une fonction quelconque.

- Le problème  $\min_{x \in X} f(x)$  est dit convexe si et seulement si la fonction objective  $f$  est convexe et

l'ensemble des contraintes est convexe.

-Si  $f$  est strictement convexe et l'ensemble  $X$  est convexe, alors le problème  $\min_{x \in X} f(x)$  Est dit

strictement convexe. [4]

## Chapitre 1: Définitions et Notations

---

### 1.3.3 Caractérisation différentielle de la convexité

Dans le cas où la fonction  $f$  est différentiable (à l'ordre 1 ou à l'ordre 2) en tout point, on sait caractériser la convexité d'une fonction via son gradient ou sa hessienne [9].

#### 1.3.3.1 caractérisation de la convexité

##### Proposition 1.7 [3] :

Soit  $\Omega \subset \mathbb{R}^n$  ouvert,  $U \subset \Omega$  avec  $U$  convexe et  $f : \Omega \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ . Alors :

1) Les 3 propositions suivantes sont équivalentes :

1-  $f$  est convexe sur  $U$  :

2-  $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle, \forall x, y \in U$

3-  $\nabla f$  est monotone sur  $U$ , c'est-à-dire :  $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0, \forall x, y \in U$

2) Si de plus  $f$  est de classe  $C^2$  sur  $\Omega$  alors  $f$  est convexe sur  $U$  si et seulement si :  $\langle \nabla^2 f(x), (y - x), y - x \rangle \geq 0, \forall x, y \in U$

**Preuve :**

1) On montre ici l'équivalence entre 1, 2, 3

1)  $\Rightarrow$  2) Supposons  $f$  convexe. la définition de la convexité peut s'écrire

$$f(x + t(y - x)) - f(x) \leq t[f(y) - f(x)]$$

En fixant  $x, y$  en divisant par  $t$  et en faisant  $t$  tendre vers 0 (ce qui est possible car :  $t \in [0, 1]$ ) On obtient 2.

2)  $\Rightarrow$  3) de 2) on déduit

$$f(y) \geq f(x) + \langle \nabla f(y), x - y \rangle, \forall x, y \in U$$

Et aussi (en versant  $x, y$ ) :

$$f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle, \forall x, y \in U$$

En faisant la somme de ces 2 inégalités on obtient 3).

3)  $\Rightarrow$  1) Soient  $x, y \in U$  on introduit la fonction  $g : I \rightarrow \mathbb{R}$  définie par :

$$t \in I \rightarrow g(t) = f(ty + (1 - t)x) \in \mathbb{R}$$

Où  $I$  est un intervalle ouvert qui contient  $[0, 1]$ . Il est facile de voir que  $g$  est de classe  $C^1$  et on a :

$$g'(t) = \langle \nabla f(ty + (1 - t)x), y - x \rangle \forall t \in I$$

## Chapitre 1: Définitions et Notations

---

Soient  $t_1, t_2 \in [0,1]$  avec  $t_1 < t_2$  Alors :

$$\begin{aligned} g'(t_1) - g'(t_2) &= \langle \nabla f(x + t_2(y - x)) - \nabla f(x + t_1(y - x)), y - x \rangle \\ &= \langle \nabla f(x + t_2(y - x)) - \nabla f(x + t_1(y - x)), y - x, (t_2 - t_1)y - x \rangle \frac{1}{t_2 - t_1} \end{aligned}$$

Par hypothèse (3) le dernier terme de l'égalité précédente est  $\geq 0$ , ce qui montre que la fonction  $g'$  est une fonction croissante. On déduit alors que  $g$  est une fonction convexe sur  $[0,1]$ , ce qui nous donne pour tout  $t \in [0,1]$ :

$$f(t \cdot 1 + (1 - t) \cdot 0) \leq tg(1) + (1 - t)g(0),$$

C'est-à-dire

$$f(ty + (1 - t)x) \leq tf(y) + (1 - t)f(x),$$

Dont  $f$  est convexe.

2) on suppose ici  $f \in C^2(\Omega)$ .

$\Rightarrow$  Supposons que  $f$  est convexe

Soit  $h \in \mathbb{R}^n$  fixé et notons  $g: \Omega \rightarrow \mathbb{R}$  la fonction donnée par  $g(x) = \langle \nabla f(x), h \rangle, \forall x \in \Omega$ .

Nous avons :

$$\langle \nabla g(x), h \rangle = \frac{\partial g}{\partial h}(x) = \lim_{t \rightarrow 0} \frac{\langle \nabla f(x + th), h \rangle - \langle \nabla f(x), h \rangle}{t},$$

Ce qui nous donne :

$$\langle \nabla^2 f(x)h, h \rangle = \lim_{t \rightarrow 0} \frac{\langle \nabla f(x + th) - \nabla f(x), th \rangle}{t^2}$$

Considérons maintenant :

$x, y \in U$  arbitraires et  $h = y - x$ .

Comme  $x + t(y - x) \in U, \forall t \in [0,1]$ , de l'égalité précédente on déduit à l'aide de la monotonie de  $\nabla f$  que  $\langle \nabla^2 f(x)h, h \rangle \geq 0$

Soient  $x, y \in U$  fixées arbitraires, et considérons la fonction  $g_1: \Omega \rightarrow \mathbb{R}$  donnée par

$$g_1(z) = \langle \nabla f(z), x - y \rangle, \forall z \in \Omega. \text{ Alors:}$$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle = g_1(x) - g_1(y) = \langle \nabla g_1(y + 0(x - y)), x - y \rangle$$

## Chapitre 1: Définitions et Notations

Avec  $\theta \in ]0,1[$  (on a utilisé l'une des formules de Taylor).

D'autre part, nous avons :  $\forall g_1(z) = \nabla^2 f(z)(x - y)$ ,

Et ceci nous permet de déduire :

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle = \langle \nabla^2 f(y + \theta(x, y)), x - y \rangle \geq 0$$

Ceci nous donne la monotonie de  $\nabla f$  donc la convexité de  $f$ .

### 1.3.3.2 caractérisation de la convexité stricte

Soit  $\Omega \subset \mathbb{R}^n$  ouvert,  $U \subset \Omega$  avec  $U$  convexe et  $f: \Omega \rightarrow \mathbb{R}$  une fonction de classe  $C^1$

Alors :

1) Les 3 propositions suivantes sont équivalentes :

1.  $f$  est strictement convexe sur  $U$ .
2.  $f(y) > f(x) + \langle \nabla f(x), y - x \rangle, \forall x, y \in U$  avec  $x \neq y$ .
3.  $\nabla f$  est strictement monotone sur  $U$ , c'est-à-dire :

$$\langle \nabla f(y) - \nabla f(x) \rangle > 0 \quad \forall x, y \in U \text{ avec } x \neq y.$$

2) Si de plus  $f$  est de classe  $C^2$  sur  $\Omega$  alors si :

$$\langle \nabla^2 f(x)(y - x), y - x \rangle > 0, \forall x, y \in U \text{ avec } x, y$$

Donc  $f$  est strictement convexe.[3]

### 1.3.3.3 Matrice définie positive

Soit  $A \in \mathbb{R}^{n \times n}$

- $A$  définie positive  $\Leftrightarrow \forall d \in \mathbb{R}^n, d^T A d > 0$ .
- $A$  semi définie positive  $\Leftrightarrow \forall d \in \mathbb{R}^n, d^T A d \geq 0$ .

**Théorème 1.3 :** [3]

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^2$  On note  $H(x)$  sa hessienne en  $x$ .

- Si  $H(x)$  est semi définie positive pour tout  $x \in \mathbb{R}^n$ , alors  $f$  est convexe.
- Si  $H(x)$  est définie positive pour tout  $x \in \mathbb{R}^n$ , alors  $f$  est strictement convexe.

## Chapitre 1: Définitions et Notations

---

### 1.4 Conditions d'optimalité

Nous supposons dans tout ce paragraphe que  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  est une fonction une ou deux fois différentiable. On notera  $x^*$  un minimum (local) de  $f$ .

#### 1.4.1 Conditions nécessaires d'optimalité

Etant donné un point  $x^*$ , la propriété de différentiabilité continue de la fonction  $f$  fournit une première manière de caractériser une solution optimale.

##### 1.4.1.1 Conditions nécessaires d'optimalité du premier ordre

**Théorème 1.3 :** [3]

Soit :

$f: \mathbb{R}^n \rightarrow \mathbb{R}$  différentiable au point  $x^* \in \mathbb{R}^n$  est un minimum local de  $f$

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

**Preuve:**

On démontre par l'absurde,  
on suppose que  $\nabla f(x^*) \neq 0$

Si on suppose  $d = \nabla f(x^*)$

on obtient:  $\nabla f(x^*)^T d = -\|\nabla f(x^*)\|^2 < 0$ ,

il existe  $\eta = 0$

tel que  $f(x^* + td) < f(x^*), \forall 0 < t \leq \eta$

Ce qui donne une contradiction avec le fait que  $x^*$  est un minimum local,

d'où  $\nabla f(x^*) = 0$ .

**Example**

- 1) La fonction  $x \mapsto x^2$  admet un point critique en  $x = 0$  qui est aussi minimum local.
- 2) La fonction  $x \mapsto -x^2$  admet un point critique en  $x = 0$  qui est aussi maximum local.
- 3) La fonction  $(x, y) \mapsto x^2 - y^2$  admet un point critique en  $(x, y) = (0, 0)$  qui n'est ni minimum local ni maximum local, c'est ici un point-selle.

## Chapitre 1: Définitions et Notations

---

La fonction  $(x, y) \mapsto x^2 - y^2$  admet un point critique en  $(x, y) = (0, 0)$  qui n'est ni minimum local ni maximum local, c'est ici un point-selle.

### 1.4.1.2 Conditions nécessaires d'optimalité du second ordre

**Théorème 1.4:** [3]

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  deux fois différentiable au point  $x^* \in \mathbb{R}^n$ . Si  $x^*$  est un minimum local de  $f$  alors:  $\nabla f(x^*) = 0$  et la matrice hessienne de  $f$  au point  $x$  qu'on note  $H(x)$  est semi définie positive.

### 1.4.2 Conditions suffisantes d'optimalité

**Théorème 1.5** [3] :

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  deux fois différentiable au point  $x^* \in \mathbb{R}^n$ . Si

$\nabla f(x^*) = 0$  et  $H(x^*)$  est définie positive alors  $x^*$  est un minimum local strict de  $f$ .

**Remarque 1.5 :**

Cas particulier  $n = 1$ .

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$  une fonction de classe  $C^2$  si  $x^* \in \mathbb{R}$  vérifie les conditions:

$$\begin{cases} f'(x^*) = 0 \\ f''(x^*) > 0 \end{cases}$$

Alors,  $x^*$  est un minimum local strict de  $f$ .

**Proposition 1.5** [4] :

Soit  $U$  un ouvert de  $\mathbb{R}^n$ ,  $f: U \rightarrow \mathbb{R}$  de classe  $C^2(U)$  et soit  $x^* \in U$ .

- Si toutes les valeurs propres de la matrice hessienne  $H(x^*)$  sont strictement positives, alors  $f$  admet un minimum local.

## Chapitre 1: Définitions et Notations

---

- Si toutes les valeurs propres de la matrice hessienne  $H(x^*)$  sont strictement négatives, alors  $f$  admet un maximum local.
- S'il y a des valeurs strictement positives et d'autres strictement négatives, alors  $(x^*)$  est un point selle.

### **Théorème 1.4** [3] :

Soit  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  telle que  $f$  est convexe et différentiable. Alors  $x$  est un minimum globale de  $f$  si et seulement si  $\nabla f(x^*) = 0$ .

Si  $f$  est strictement convexe, alors tout minimum local devient non seulement global mais aussi unique.

# Chapitre 2

## Méthode de séparation et évaluation (B&B)

### 2.1 Introduction

Pour trouver la solution optimale, et être certain qu'il n'y a pas mieux, il faut utiliser une méthode exacte, qui demande un temps de calcul beaucoup plus long. Il n'existe pas une méthode exacte universellement plus rapide que toutes les autres. Chaque problème possède des méthodes mieux adaptées que d'autres. Dans ce chapitre nous allons présenter un exemple d'algorithme de ce type, nommé procédure par séparation et évaluation (PSE), ou en anglais branch and bound.

La méthode de branch and bound (procédure par évaluation et séparation progressive) consiste à énumérer ces solutions d'une manière intelligente en ce sens que, en utilisant certaines propriétés du problème en question, cette technique arrive à éliminer des solutions partielles qui ne mènent pas à la solution que l'on recherche. De ce fait, on arrive souvent à obtenir la solution recherchée en des temps raisonnables. Bien entendu, dans le pire cas, on retombe toujours sur l'élimination explicite de toutes les solutions du problème. [2]

### 2.2 Présentation de la méthode Séparation et évaluation (B&B)

#### 2.2.1 Historique :

- Le premier exemple développé de procédure par séparation est l'algorithme proposé en 1960 par LAND et DOIG pour les programmes mixtes.
- En 1963, Little et al utilisent une PS pour résoudre le problème du voyageur de commerce et eut un grand retentissement.
- de façon à isoler une solution optimale dans l'un de ses sous-ensembles.

En 2002 Gutin et A.P. Punnen publiaient un livre "The Traveling Salesman Problem and Its Variations" Ce livre couvre tous les domaines importants de l'étude sur TSP, y compris la théorie des polyèdres symétriques et asymétriques pour le TSP, branch and bound et d'autres méthodes. [2]

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

### 2.2.2 Définition de la méthode

Les méthodes par séparation et évaluation sont des méthodes de résolution exacte de problèmes d'optimisation combinatoire introduite par Land et Doig. Une méthode de séparation et évaluation consiste à énumérer implicitement toutes les solutions dans  $S$  (l'espace de solutions) en examinant les sous-ensembles de  $S$ .

Il s'agit essentiellement de diviser (divide and conquer) l'ensemble de toutes les solutions réalisables (problème initial) en sous-ensembles plus petits (sous problèmes) et mutuellement exclusifs. C'est la phase « séparation » (Branch). Puis divers critères sont utilisés pour identifier les sous-ensembles qui peuvent contenir la solution optimale et les sous-ensembles qui ne doivent pas être explorés plus à fond car ils ne peuvent pas contenir la solution optimale. C'est la phase « évaluation » (Bound).

L'énumération des solutions du problème consiste à construire un arbre Branch and Bound dont les nœuds sont des sous-ensembles de solutions du problème, et les branches sont les nouvelles contraintes à respecter. La taille de l'arbre dépend de la stratégie utilisée pour la construire.

Pour ce faire, cette méthode se dote d'une fonction qui permet de mettre une borne inférieure (en cas de min) ou borne supérieure (en cas de max) sur certaines solutions pour soit les exclure soit les maintenir comme des solutions potentielles. Bien entendu, la performance d'une méthode de branch and bound dépend, entre autres, de la qualité de cette fonction (de sa capacité d'exclure des solutions partielles tôt). [13]

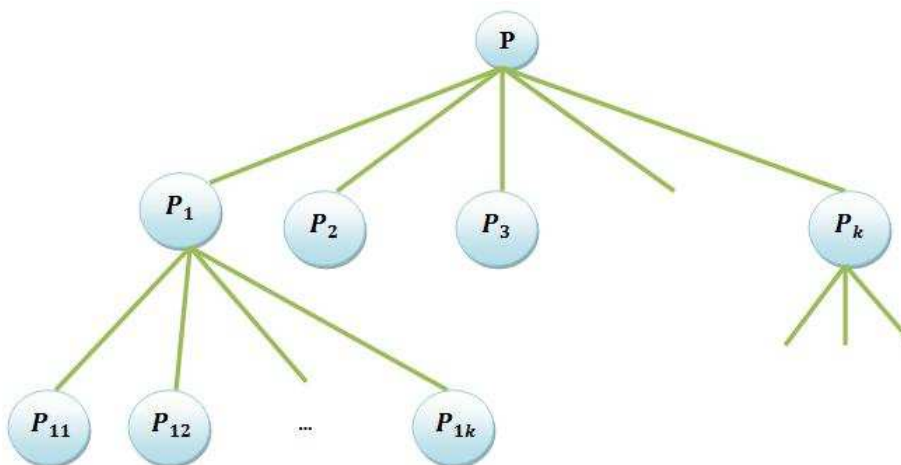


FIGURE 2.6 – Présentation de la méthode branch and bound

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

---

Une méthode de branch and bound est basé sur de trois axes principaux :

- La procédure de séparation ou bien Le branchement, qui consiste de partitionner un ensemble de solutions en sous ensembles.
- La procédure d'évaluation permettant le calcul d'une borne pour un ensemble de solutions .
- La stratégie de parcours ou procédure de cheminement d'exploration de l'arborescence de recherche".

### 2.3.1 La procédure de séparation

La séparation consiste à diviser le problème en un certain nombre de sous-problèmes qui ont chacun leur ensemble de solutions réalisables. En résolvant tous les sous-problèmes et en prenant la meilleure solution trouvée, on est assuré d'avoir résolu le problème initial. Ce principe de séparation est appliqué de manière récursive à chacun des sous-ensembles tant que celui-ci contient plusieurs solutions.[2]

#### 2.3.1 Démarche de la séparation :

##### **Définition [5] :**

un domaine est un ensemble fini ou infiniment dénombrable d'élément  $D = \{d_1, d_2, \dots, d_m\}$ . On fait l'hypothèse d'une minimisation de la fonction objective  $F$ . Posons  $W = \{\min F(X) | X \in D\}$ . Nous allons supposer que l'arbre construit est binaire (le degré extérieur de chacun de ses sommets est soit 0 soit 1 "voir l'exemple de séparation") pour simplifier la présentation de la méthode.

La généralisation à un arbre quelconque se ferait sans difficulté. A chaque sommet  $x_i$  de l'arbre correspond un domaine de solutions  $d_i \subset D$  de façon qu'à tout moment de la construction du graphe, l'arbre possède la propriété suivante :

Les  $d_i$  associés aux feuilles du graphe forment une partition de  $D$ . Pour ce faire, l'arbre est construit au départ d'un sommet (racine) dont le domaine associé est  $D$  tout entier. Ensuite, à chaque étape un sommet  $x_i$  de niveau  $N \geq 0$  et de domaine  $d_i$  est analysé. Cet algorithme lui associe 2 descendants  $y$  et  $y'$  de niveau  $N + 1$  dont les domaines respectifs forment une partition de  $d_i$ .

Les descendants du sommet  $x_i$  s'obtiennent en fixant une caractéristique binaire de la solution. Pour ce faire, on choisit une variable de séparation et on coupe en deux son domaine de valeurs possibles. Une fois le sommet  $x_i$  analysé et les sommets  $y$  et  $y'$  définis,  $x_i$  n'est plus candidat à une étude ultérieure.

### Exemple

Si  $S \subseteq \{0,1\}^3$  on peut construire l'énumération suivant :

D'abord on divise  $S$  en deux ensembles :

$S_0 = \{x \in S; x_1 = 0\}$  et  $S_1 = \{x \in S; x_1 = 1\}$  en suite

$S_{00} = \{x \in S; x_2 = 0\} = \{x \in S_0; x_1 = 0 = x_2 = 0\}$ ,  $S_{01} = \{x \in S_0; x_2 = 1\}$

et ainsi de suite...[24]

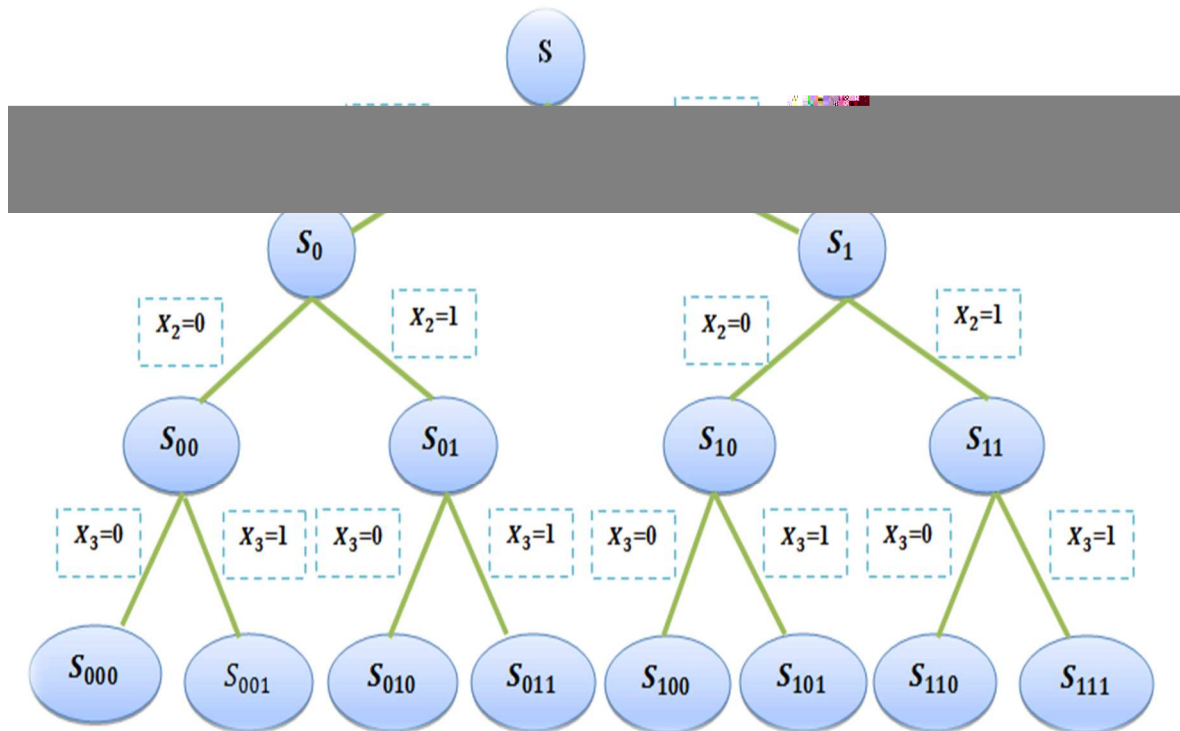


FIGURE 2.7 – Décomposition d'un problème binaire

### 2.3.2 Règles de branchement

Les règles de branchement dans un algorithme (B&B), sont utilisées pour subdiviser l'ensemble des solutions possibles.

Deux règles sont utilisées dans le branchement, qui consistent à affecter les tâches une à une à partir de la fin (branchement en arrière exemple Problème des 8 reines), ou au contraire, affecter les tâches à partir du début (branchement en avant).

Toutefois, il existe des heuristiques qui, une fois appliquées, déterminent à chaque étape, les tâches qui doivent être affectées au début, à la fin, directement après ou directement avant une tâche.[5]

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

---

### Remarque 2.1 :

Le processus de séparation d'un sous-ensemble  $p_i$  s'arrête dans l'un des cas suivants (cas Min) :

- Lorsque la borne inférieure de  $p_i$  est à la meilleure solution trouvée jusqu'à maintenant pour le problème initial.
- Lorsque  $p_i$  n'admet pas de solution réalisable.
- Lorsque  $p_i$  admet une solution complète du problème initial.[5]

### 2.4 Procédure d'évaluation

L'évaluation d'un nœud de l'arbre de recherche a pour but de déterminer l'optimum de l'ensemble des solutions réalisables associées au nœud en question ou, au contraire, de prouver mathématiquement que cet ensemble ne contient pas de solution intéressante pour la résolution du problème. Lorsqu'un tel nœud est identifié dans l'arbre de recherche, il est inutile d'effectuer la séparation de son espace de solutions. On peut distinguer pendant le déroulement de l'algorithme trois types de nœuds dans l'arbre de recherche : le nœud courant qui est le nœud en cours d'évaluation, des nœuds actifs qui sont dans la liste des nœuds qui doivent être traités, et des nœuds inactifs qui ont été élagués au cours du calcul.[ 6]

Donc la procédure d'évaluation consiste à analyser un nœud ( sous-problème), cette analyse vise à évaluer la valeur optimale de la fonction objectif de sous-problème, plus précisément à déterminer une borne inférieure ou la borne supérieur de cette valeur.

L'algorithme de Branch and Bound maintient une borne supérieure (UB) et inférieure (LB) au nœud racine.

#### 2.4.1 Borne supérieure

Est la meilleure solution réalisable connue. Une solution réalisable est trouvée par l'algorithme aux feuilles de l'arbre d'énumération, et parfois avec un peu de chance comme solution de la relaxation linéaire à un nœud. On peut bien sûr utiliser en plus des heuristiques pour en générer au cours de l'exploration et espérer accélérer la recherche.

#### 2.4.2 Borne inférieure

A chaque nœud  $x$ , la borne inférieure  $LB(x)$  pour le sous-problème correspondant est calculée, par relaxation linéaire. Ceci permet éventuellement d'éliminer le nœud de la recherche si  $(x) > UB$  . Ceci permet aussi (et surtout) de mettre à jour la borne inférieure de son père. En effet, la borne inférieure d'un nœud doit être supérieure à la plus petite des bornes inférieures de ces fils. Les bornes inférieures sont ainsi mis à jour de proche en proche, éventuellement jusqu'au nœud racine. Mais la valeur  $LB$  de la racine augmente en générale très lentement.

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

---

La connaissance d'une borne inférieure du problème et d'une borne supérieure de la fonction d'utilité de chaque sous-problème permet de stopper l'exploration d'un sous-ensemble de solutions qui ne sont pas candidates à l'optimalité.

Si pour un sous-problème la borne supérieure est plus petite que la borne inférieure du problème, l'exploration du sous-ensemble correspondant est inutile. D'autre part, lorsque le sous-ensemble est suffisamment « petit », on procède à une énumération explicite : on résout alors le sous-problème correspondant.[6]

\

### 2.4.3 Calcul de la borne inférieure

La procédure qui calcule la borne inférieure est l'élément clé de n'importe quel algorithme BB, car une qualité médiocre de la borne inférieure ne peut pas être compensée par de bons choix de la stratégie de recherche ou de la règle de branchement.[5]

Les techniques de relaxation sont souvent utilisées dans le calcul d'une borne inférieure. Appliquées à un problème de minimisation, elles fournissent une évaluation par défaut de l'optimum, en relâchant les contraintes les plus difficiles à satisfaire, c'est-à-dire en les supprimant, ou en les prenant partiellement en compte. [6].

Voici dans ce qui suit, les principales techniques de relaxation :

- **Relaxation des contraintes** : Une technique simple de relaxation consiste à ignorer certaines contraintes du problème. On obtient alors un problème dont la solution optimale est plus facile à calculer. [6]
- **Relaxation lagrangienne** : c'est une technique de relaxation qui consiste à supprimer des contraintes en les intégrant dans la fonction objectif en la pénalisant si ces contraintes ne sont pas respectées.
- **La relaxation continue** : c'est une relaxation naturelle qui est, pour la PLNE (par l'algorithme du simplexe, des points intérieurs ou du volume) et la programmation quadratique (avec matrice définie positive) très efficace. Néanmoins, nous verrons que la relaxation continue n'est pas forcément la meilleure.

Il est également possible d'améliorer la borne obtenue par relaxation continue en utilisant des idées spécifiques au problème. En ce qui concerne plus spécifiquement la PLNE, on peut améliorer la valeur de relaxation linéaire en ajoutant des contraintes, le renforcement ou des variables, la reformulation.[24]

- **Relaxation linéaire continue** : le problème (LP) obtenu par relaxation des contraintes d'intégrité d'un (ILP "integer linear programming") définit une relaxation de ce (ILP). Pour un problème de minimisation, la solution optimale du (LP) donne une borne inférieure de la valeur de la solution optimale du (ILP),

Il s'agit d'un relâchement de la contrainte  $x \in P \cap N^n$  en  $x \in P$  telle que :

$$P = \{x \in R^{n+} : Ax \leq b\}$$

Comme  $P \cap N^n$  et la fonction objective ne change pas alors il est clair qu'il est une relaxation. La relaxation linéaire ne donne pas seulement une borne inférieure mais des fois preuve l'optimalité. [24]

### 2.4.4 Calcul de la borne supérieure

La borne supérieure est un élément essentiel de l'algorithme B&B. Calculée initialement par une heuristique appropriée, la borne supérieure est toujours comparée avec les bornes inférieures trouvées au niveau des nœuds.

Dans le cas où une borne inférieure d'un nœud est supérieure à la borne supérieure, ce nœud est élagué. Une fois arrivée à un nœud feuille, la borne supérieure sera remplacée par la valeur de la solution calculée au niveau de la feuille. [6]

### 2.4.5 Conditions d'élimination

Le nœud de l'arborescence sondé si une des trois conditions suivantes sont :

- Éliminé par optimalité.
- Éliminé par borne.
- Éliminé par infaisabilité :  $P_i = \emptyset$ . [6]

#### 2.4.5.1 Éliminé par optimalité

-  
Dans la figure suivante on a une décomposition d'un problème  $P$  a deux sous-problèmes  $P_1$  et  $P_2$  avec ses bornes supérieures et inférieures.

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

On observe que :

la borne inférieure  $P_1$  égale,  $LB(P_1) = 17$

la borne supérieure  $P_2$  égale,  $UB(P_2) = 22$

Donc ne peut pas dériver  $P_1$  Alors la branche de sous-problème  $P_1$  est éliminée par optimalité.

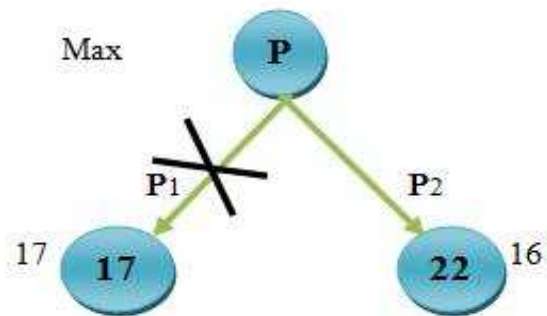


FIGURE 2.8 – Éliminer par optimalité

### 2.4.5.2 Éliminé par borne :

Dans la figure suivante on a une autre décomposition d'un problème  $P$  à deux sous-problèmes  $P_1$  et  $P_2$  avec ses bornes supérieures et inférieures.

On observe que :

la valeur optimale est minorée par 16, mais la borne supérieure  $P_1$  égale,  $UB(P_1) = 22$ . Donc le sous-problème  $P_1$  ne contient pas la solution optimale. Alors la branche de sous-problème  $P_1$  est éliminée à partir de borne.

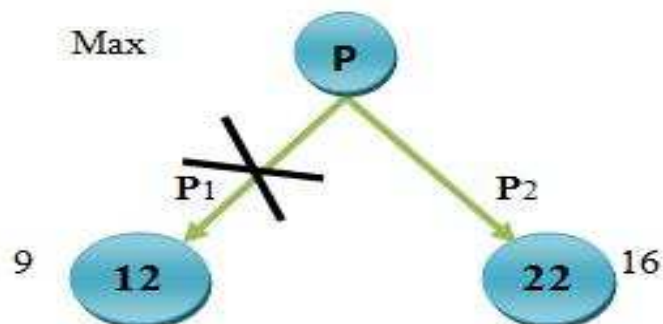


FIGURE 2.9 – Éliminer par borne

### 2.4.5.3 Éliminé par infaisabilité

Dans la figure suivante on a une autre décomposition d'un problème  $P$  a deux sous-problèmes  $P_1$  et  $P_2$  avec ses bornes supérieur et inférieur. On observe que : la borne inférieure  $P_1$  égale,  $LB(P_1) = \emptyset$ , la borne supérieure  $P_2$  égale,  $UB(P_2) = 22$

Donc le sous problème  $P_1$  infaisabilité. Alors la branche de sous-problème  $P_1$  est éliminée par infaisabilité

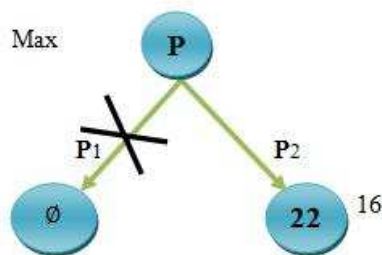


FIGURE 2.10 – Éliminer par infaisabilité.

## 2.5 La stratégie de parcours

La plupart des méthodes de parcours d'un arbre de recherche sont connues depuis déjà plusieurs années. La différence entre deux stratégies de recherche réside principalement dans la taille de l'arbre de recherche généré et donc dans le temps de résolution du problème traité. Les stratégies appliquées sont la plupart du temps différentes selon la nature du problème.

On distingue les problèmes d'optimisation combinatoire avec une fonction objectif et un nombre généralement important de solutions et les problèmes de satisfaction de contraintes pour lesquels on ne cherche la plupart du temps qu'une seule solution [6]. On peut citer différentes stratégies de recherche. Pour les méthodes de parcours de l'arbre de recherche, on peut citer, entre autres, les politiques suivantes :

### 2.5.1 Profondeur D'abord

La méthode depth-first search est une méthode couramment utilisée. Elle consiste en une exploration en profondeur de l'arbre de recherche. Les variables sont fixées une à une itérativement, jusqu'à obtenir une solution ou jusqu'à ce qu'une variable ait un domaine vide. Lorsque c'est le cas, on revient sur la dernière décision qui a été prise et on prend une autre décision, si cela est possible. Dans le cas contraire, on remonte plus haut dans les décisions prises. Cette technique est appelée Backtrack "marché en arrière".

Lorsque le nombre de solutions est important, cette méthode permet de trouver très rapidement une solution. Cette méthode est très classique dans l'implémentation d'un Branch and Bound. En terme d'occupation mémoire, cette méthode est la plus intéressante.[6]

### 2.5.2 Largeur D'abord

Cette méthode est la méthode opposée à la méthode depth-first search. Elle consiste à explorer l'arbre en largeur. Le parcours en largeur correspond à un parcours par niveau de nœuds de l'arbre. Un niveau est un ensemble de nœuds ou de feuilles situés à la même distance du nœud racine. Ce type de parcours est rarement utilisé, du fait de l'occupation mémoire trop importante.

Une fois l'arbre créé, elles permettent d'ordonner une liste de nœuds à traiter. Elles ne modifient en aucun cas la structure propre de l'arbre. Néanmoins, deux politiques de parcours d'arbre de recherche peuvent donner des résultats différents en termes de rapidité d'exécution et en termes d'occupation mémoire. En effet, on peut associer à ces méthodes des calculs de bornes qui pourront amener à la troncature de sous-espaces de l'arbre de recherche (par exemple, lorsque la borne inférieure calculée à un nœud est supérieure à la borne supérieure). Ainsi, si la structure de l'arbre n'est pas modifiée, les espaces explorés de l'arbre peuvent être modifiés par les schémas d'exploration.[6]

### 2.5.3 Meilleur D'abord

On choisit parmi les nœuds restant à traiter celui qui a la plus petite borne inférieure. Cette borne peut être calculée par exemple, par relaxation de contraintes, par une méthode heuristique, . . . Ce choix est efficace lorsque les bornes calculées sont de bonne qualité. Elle présente par contre l'inconvénient de traiter des nœuds successifs très différents (valeurs et variables différentes) et donc de ne présenter que peu d'incrémentalité dans les calculs, notamment les calculs de bornes inférieures.

De plus, cette méthode n'est applicable que lorsqu'une borne est calculable, ce qui peut ne pas être le cas pour certains problèmes de satisfaction de contraintes, dans lesquels il n'y a pas de fonction objectif.[6]

## 2.6 Efficacité de l'algorithme B&B

Différentes techniques sont utilisées pour rendre l'algorithme BB plus efficace. On peut en citer quelques-unes :

### 2.6.1 Choix d'une stratégie de recherche

La stratégie « profondeur d'abord » mariée avec le choix du nœud de plus petite borne inférieure, est utilisée en pratique car elle nécessite peu d'espace mémoire. Par contre, la stratégie « largeur d'abord » trouve une solution optimale plus rapidement que la stratégie « profondeur d'abord ».

Il est donc primordial de faire un bon choix de la stratégie de recherche, qui dépend généralement de la taille du problème et des ressources informatiques disponibles. [6]

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

---

### 2.6.2 Choix de la borne inférieure

Une borne inférieure doit être proche de l'optimum. Le choix de la borne inférieure dépend pleinement du problème à résoudre et de la technique utilisée (relaxation, heuristiques. . . ). Les bonnes sont celles qui éliminent un grand nombre de nœuds le plus rapidement possible, ce qui réduit considérablement l'arbre de recherche.

### 2.6.3 Choix de la borne supérieure

la borne supérieure peut être une heuristique qui permet d'élaguer certaines branches de l'arbre de recherche.

### 2.6.4 Règles de dominance

Les règles de dominance sont des conditions posées sur un nœud, pour pouvoir l'éliminer si elles sont vérifiées (on dit que ce nœud est dominé), et ce avant de calculer sa borne inférieure. Les règles de faisabilité sont appliquées en premier aux nœuds, pour décider de l'admissibilité de la séquence traversant ce nœud, évidemment, le nœud est éliminé si la séquence n'est pas admissible. Les règles de dominance sont utiles pour réduire au mieux l'arbre de recherche, en éliminant les nœuds dominés. [5]

### 2.6.5 Choix du critère d'évaluation

Lors d'une application de branch and bound, on a besoin d'avoir une fonction qui permet d'évaluer chaque nœud traité, et éventuellement élaguer ceux qui sont inutiles. De même, un nœud peut être élagué dans trois cas possibles : dans le premier cas, on arrive à un stade où la valeur de la borne inférieure d'un nœud courant est plus grande ou égale à la valeur de la borne supérieure qu'on avait établie auparavant.

Dans le deuxième cas, la solution n'est pas réalisable, l'un des critères d'évaluation n'a pas été respecté. Le troisième cas est le cas où on a obtenu une solution réalisable, tous les critères sont valides, mais la solution obtenue est supérieure à la borne inférieure. [5]

## 2.7 Méthode séparation et évaluation sur des exemples

La caractéristique principale de cette technique consiste de trouver dans quelle branche les solutions ont le plus de chance d'être trouvées pour ne pas continuer inutilement et voilà deux exemples pour comprendre le parcours des solutions.

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

---

### Exemple 1 [14]

Soit le PLNE (noté  $P$ ) :  $\min x_1 - 2x_2$

$$S.C \begin{cases} -4x_1 + 6x_2 \leq 9 \\ x_1 + x_2 \leq 4 \\ x_1, x_2 \geq 0, \text{ est entiers} \end{cases}$$

$Z^*$  sera le coût optimal de la relaxation linéaire (borne inférieure)

- Si la solution de la relaxation est entière, pas besoin de partitionner le sous problème.
- Sinon, on choisit un non entier, et on crée deux sous problèmes en ajoutant les contraintes :

$$x_* = [x_i^*] \text{ et } x_i \geq [x_i^*]$$

La solution optimale réelle obtenue par le simplexe (de  $P$  relaxé) est

$$x_* = (1.5, 2.5)$$

et

$$Z^* = -3.5$$

En cas de minimisation cette solution représente donc une borne inférieure (LB) :

$$Z_{LB} = Z^* = -3.5.$$

l'arrondir de cette solution donne une solution réalisable (borne supérieure (UB))  
on a :

$$x = (1, 2) \text{ et } Z_{UB} = -3.$$

Ainsi on peut borner la solution optimale du PLNE comme suit :

$$-3,5 \leq Z_{UB} \leq -3$$

### Création des sous problèmes $P_1$ et $P_2$

En rajoutant les contraintes  $x_2 \leq 2$  et  $x_2 \geq 3$

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

$P_1$	$P_2$
$\min x_1 - 2x_2$ S.C $4x_1 + 6x_2 \leq 9$ $x_1 + x_2 \leq 4$ $x_2 \geq 3$ $x_1, x_2$ , est entiers	$\min x_1 - 2x_2$ S.C $-4x_1 + 6x_2 \leq 9$ $x_1 + x_2 \leq 4$ $x_2 \leq 2$ $x_1, x_2$ , est entiers

TABLEAU 2.2 – Sous problèmes  $P_1$  et  $P_2$ .

Liste des sous problèmes actifs:  $\{P_1 \text{ et } P_2\}$

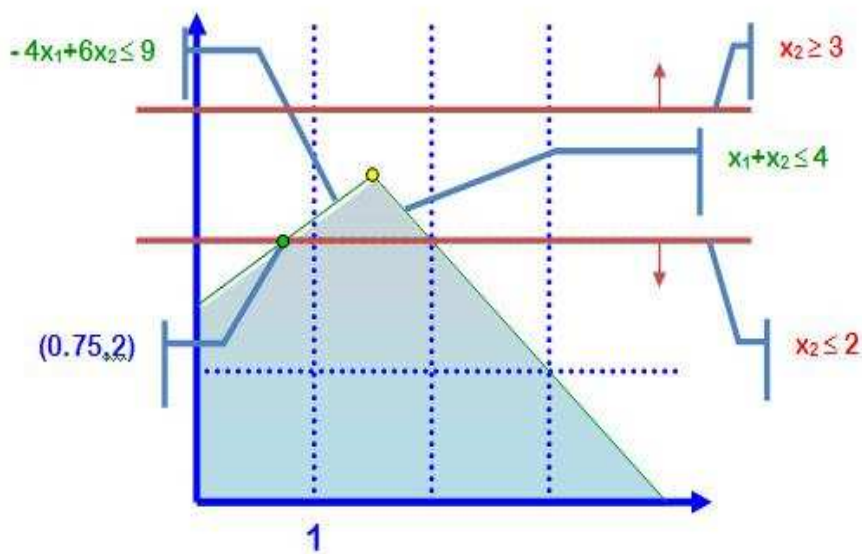


FIGURE 2.11 – Sous problèmes actifs :  $\{P_1 \text{ et } P_2\}$

On observe que :

Le problème  $P_1$  n'est pas réalisable

Liste des problèmes actifs  $P_2$

La solution optimale réelle de  $P_2$  relaxé est :  $x^* = (0.75, 2)$  et  $Z^* = -3.25$

La borne inférieure trouvée  $LB = (-3.25)$  étant inférieure à la meilleure solution trouvée (-3).

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

Création des sous problèmes  $P_3$  et  $P_4$ :

En rajoutant les contraintes  $x_1 \leq 0$  et  $x_1 \leq 1$

$P_3$	$P_4$
$\min x_1 - 2x_2$ S.C $4x_1 + 6x_2 \leq 9$ $x_1 + x_2 \leq 4$ $x_2 \leq 2$ $x_1 \geq 1$ $x_1, x_2$ , est entiers	$\min x_1 - 2x_2$ S.C $-4x_1 + 6x_2 \leq 9$ $x_1 + x_2 \leq 4$ $x_2 \leq 2$ $x_1 \leq 0$ $x_1, x_2$ , est entiers

TABLEAU 3.3 – Sous problèmes  $P_3$  et  $P_4$

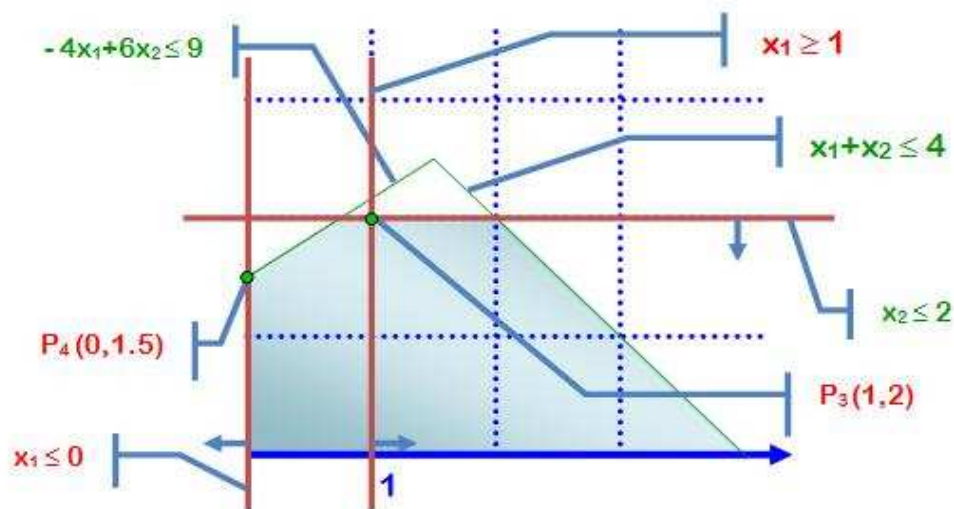


FIGURE 2.12 – Sous problèmes actifs :  $\{P_3$  et  $P_4\}$

Liste des problèmes actifs :  $\{P_3$  et  $P_4\}$

La solution optimale de  $P_3$  relaxé est entière donc :

$$x^* = (1, 2) \text{ et } Z = -3.$$

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

La solution optimale réelle de  $P_4$  relaxé est :

$$x^* = (0, 1.5) \text{ et } Z^* = -3.$$

Cette branche est arrêtée car  $Z^*(LB) \geq -3$ .

La solution optimale est donc :

$$x_{PLNE} = (1,2) \text{ avec } Z_{opt} = -3$$

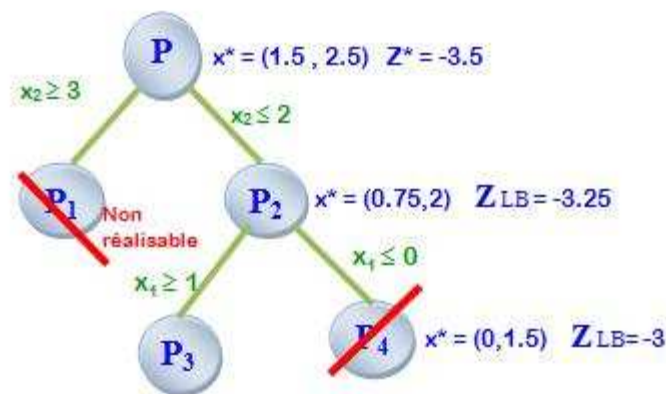


FIGURE 2.13 – Présentation du parcours de solution.

### Exemple 2 :

Problème du sac à dos (problème. de maximisation) – Deux simplifications

1. Les variables sont binaires.
2. La relaxation linéaire peut être résolue efficacement par un algorithme simple : prendre d'abord les articles à meilleur rendement, jusqu'à atteindre la capacité.

- Une société dispose de 1 400 000 DA à investir.
- Les experts proposent 4 investissements possibles :

	Coût	Bénéfice	Rendement
Inv. 1	500 000	1 600 000	3.20
Inv. 2	700 000	2 200 000	3.14
Inv. 3	400 000	1 200 000	3.00
Inv. 4	300 000	800 000	2.67

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

Résolution :

$a_i$	$c_i$	$Rdt$	$p$	$p_1$	$p_3$	$p_4$	$p_5$	$p_6$	$p_2$	$p_7$	$p_8$	$p_9$	$p_{10}$	$p_{11}$	$p_{12}$
500 000	1 600 000	3.20	1	1	1	$\frac{3}{5}$	0	1	1	1	1	1	$\frac{5}{4}$	0	1
700 000	2 200 000	3.14	1	$\frac{5}{7}$	0	1	1	1	1	1	$\frac{6}{7}$	0	1	1	1
400 000	1 200 000	3.00	$\frac{1}{2}$	1	1	1	1	1	0	0	0	0	0	0	0
300 000	800 000	2.67	0	0	1	0	1	1	$\frac{2}{3}$	0	1	1	1	1	1

Commentaires :

▪ **Relaxation de  $p$  :**

- Solution réelle  $x^* = (1, 1, 0.5, 0), Z^* = 4\,400\,000$
- 
- Solution arrondie  $\bar{x} = (1, 1, 0, 0), \bar{Z} = 3\,800\,000$

On enregistre cette solution comme meilleure solution rencontrée :  $x_{best} = \bar{x}$  et  $Z_{best} = 3\,800\,000$

$$\text{Nous avons } 3\,800\,000 \leq Z_{opt} \leq 4\,400\,000$$

Création de deux sous problèmes  $P_1$

**Relaxation de  $p_1$**

- Solution réelle  $x^* = (1, 5/7, 1, 0), Z^* = 4\,371\,429 > 3\,800\,000$
- Création de deux sous problèmes  $p_3 = (x_3=0)$  et  $p_4(x_2 = 1)$
- **Note :** On va appliquer comme stratégie de parcours le DFS (Parcours en profondeur)

▪ **Relaxation de  $p_3$**

- Solution entière  $\bar{x} = (1, 0, 1, 1), \bar{Z} = 3\,600\,000$

Arrêt de cette branche, puisque nous avons obtenu une solution entière (qui n'est pas optimale  $\bar{Z} < Z_{best}$ )

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

---

- **Relaxation de  $p_4$**

- Solution réelle  $x^* = \left(\frac{3}{5}, 1, 1, 0\right) Z^* = 4\,360\,000 > 3\,800\,000$
- Création de deux sous problèmes  $p_5(x_1 = 0)$  et  $p_6(x_1 = 1)$

- **Relaxation de  $p_5$**

- Solution entière  $\bar{x} = (1,0,1,1), \bar{Z} = 4\,200\,000$

Arrêt de cette branche, puisque nous avons obtenu une solution entière. Mise à jour de  $x_{best}$  et  $Z_{best} = 4\,200\,000$

- **Relaxation de  $p_6$**

- Problème non réalisable Arrêt de la branche.

- **Relaxation de  $p_2$**

- Solution réelle  $x^* = (1, 1, 0, 2/3) Z^* = 4\,333\,333 > 4\,200\,000$  Création de deux sous problèmes  $P_7(x_4 = 0)$  et  $p_8(x_4 = 1)$

- **Relaxation de  $p_7$**

- Solution entière  $\bar{x} = (1,1,0,1), \bar{Z} = 3\,800\,000$

Arrêt de cette branche.

- **Relaxation de  $p_8$**

- Solution réelle  $x^* = (1, 6/7, 0, 1) Z^* = 4\,285\,714 > 4\,200\,000$  Création de deux sous problèmes  $p_9(x_2 = 0)$  et  $p_{10}(x_2 = 1)$

- **Relaxation de  $p_9$**

- Solution  $\bar{x} = (1,0,0,1), \bar{Z} = 2\,400\,000$

Arrêt de cette branche, car solution entière.

- **Relaxation de  $p_{10}$**

- Solution réelle  $x^* = (4/5, 1, 0, 1) Z^* = 4\,280\,000 > 4\,200\,000$
- Création de deux sous problèmes  $p_{11}(x_1 = 0)$  et  $p_{12}(x_1 = 1)$

- **Relaxation de  $p_{11}$**

- Solution entière  $\bar{x} = (0,1,0,1), \bar{Z} = 3\,000\,000$

Arrêt de cette branche.

- **Relaxation de  $p_{12}$**

## Chapitre 2 : Méthode de séparation et évaluation (B&B)

Problème non réalisable  $\Rightarrow$  Arrêt de la branche.

La meilleure solution trouvée est donc la solution optimale (Nœud  $p_5$ ), ainsi Solution optimale :

$$x_{opt} = (0, 1, 1, 1), Z_{opt} = 4\,200\,000$$

### Note :

Seules 7 combinaisons ont été considérées ( $p_3, p_5, p_6, p_7, p_9, p_{11}, p_{12}$ ), une énumération complète aurait considéré  $2^4 = 16$  combinaisons.

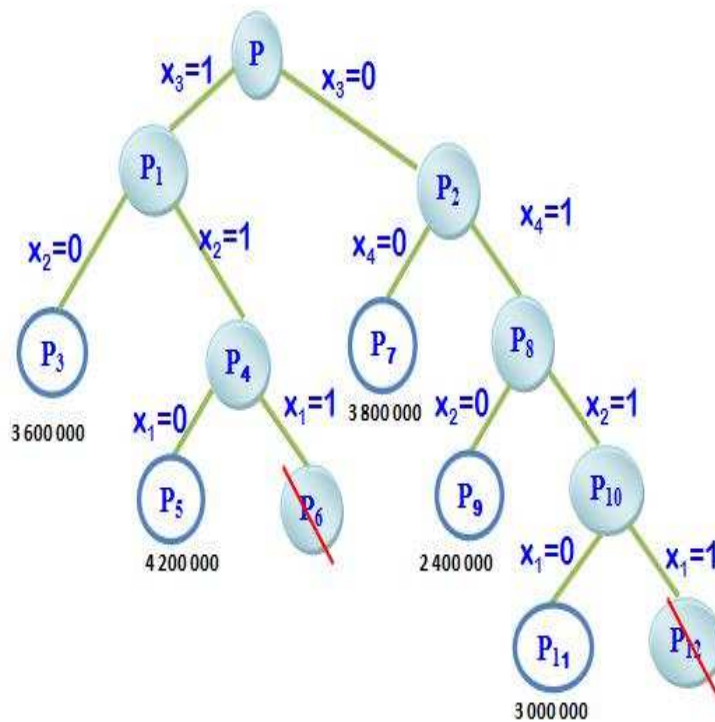


FIGURE 2.14 – Présentation de parcours de la solution.



# Chapitre 3

## La méthode d'Approximation Extérieure

### 3.1 Introduction

Les méthodes de plans de coupe constituent des outils de base dans différentes branches d'optimisation. Un plan est utilisé pour couper un ensemble de telle façon que cela n'exclut pas des points optimaux du problème d'optimisation. Nous considérons, dans cette section, les coupes qui n'enlèvent jamais aucun point réalisable. Il s'agit donc des méthodes d'Approximation Extérieure. En dehors des coupes linéaires bien connues dans la programmation convexe, des coupes continues i.e. des coupes déterminées par des fonctions continues non nécessairement linéaires seront admises.

Ces méthodes d'Approximation Extérieure, leur grand principe est de faire des coupes linéaires sur un ensemble de telle sorte qu'on ne coupe aucun point de l'ensemble réalisable, en se rapprochant de plus en plus de la solution optimale du problème donné.

On utilise des coupes linéaires, ainsi le problème donné sera remplacé par une suite de problèmes ayant des contraintes linéaires.[11]

### 3.2 Le principe de la méthode d'Approximation Extérieure

Soit (P) le problème d'optimisation à résoudre :

$$(P) \begin{cases} \min f(x) \\ x \in H \end{cases}$$

Où :  $f$  continue, linéaire ou convexe et  $H$  le domaine du problème (P).

$H = \{x : g_i(x) \leq 0\}$ , où  $g_i(x)$  sont des contraintes convexes, avec  $i = 1, 2, \dots$ .

Soit  $D_1$  un hyper rectangle tel que  $H \subset D_1$ .

On résout le problème

$$(P) \begin{cases} \min f(x) \\ x \in D_1 \end{cases}$$

$P_1$  est un problème linéaire ou convexe.

On trouve une solution  $x^1$ .

## Chapitre 3 : Méthode d'Approximation Extérieure

---

Si elle est admissible pour  $(P)$  alors elle est optimale.

Sinon, on ajoute une contrainte linéaire qui coupe la solution  $x^1$  dans  $D_1$  mais sans couper aucune solution du domaine  $H$ .

Donc on obtiendra une suite de problèmes convexes  $(P_k)$  avec des contraintes linéaires, avec  $k = 1, 2, \dots$

### Ajout des contraintes pour les problèmes $(P_k)$ :

On ajoute une contrainte linéaire au problème  $(P_k)$  de façon à ce qu'on coupe la solution  $x^k$ , sans couper aucune solution du domaine réalisable  $H$  du problème de départ  $(P)$ .

### 3.3 L'algorithme de base de la méthode d'Approximation Extérieure

#### 1. Initialisation :(1ère étape)

Choisir  $D_1 \supset H$ , poser  $k = 1$  et passer à la 2ème étape.

#### 2. Itérations $k = 1, 2, \dots$ :( 2ème étape)

- Résoudre le problème :

$$(P_k) \begin{cases} \min f(x) \\ x \in D_k \end{cases}$$

pour obtenir la solution  $x^k$

- Si  $x^k \in H$ , stop  $x^k$  est une solution optimale globale
- Sinon, construire une fonction  $I_k(x)$  telle que :
  - 1)  $I_k(x) \leq 0 \quad \forall x \in H$
  - 2)  $I_k(x^k) > 0$ .

et définir :

$$\begin{aligned} D_{k+1} &= D_k \cap \{x: I_k(x) \leq 0\} \\ &= D_k \cap H \end{aligned}$$

- Poser :  $k = k + 1$  et retourner à l'itération  $k + 1$

#### Commentaire :

L'idée de l'approximation extérieure  $AE$  est de relaxer et remplacer le problème  $(P)$  par une suite de problèmes  $(P_k)$  qui sont plus faciles à résoudre et dont les solutions convergent vers une solution optimale. Cette idée vient de la méthode de Cheney et Goldstein pour la résolution d'un programme convexe. Le premier algorithme  $AE$  pour la minimisation d'une fonction concave sur un convexe a été proposé par Falk et Hoffman. Depuis, beaucoup de travaux ont été effectués. Pour la réalisation d'un algorithme  $AE$ ,

## Chapitre 3 : Méthode d'Approximation Extérieure

deux questions importantes sont à résoudre :

- Construire des problèmes  $(P_k)$  qui sont faciles à résoudre.
- Construire  $I_k(\cdot)$  pour que la suite  $\{x^k\}$  converge.

### 3.4 La convergence de la méthode d'Approximation Extérieure

#### 3.4.1 Approximation polyédrale

Soit  $D_1$  un ensemble polyédral (convexe) et  $I_k(x) = \langle a^k, x \rangle + b_k$  une fonction affine. Alors  $S_k = \{x \in \mathbb{R}^n : I_k(x) = 0\}$  est un hyperplan séparant strictement  $x^k$  de  $H$ . Tous ensemble  $D_k$  construit ainsi, est un polyédral et on obtient une suite décroissante de polyèdres qui approximent  $H$ . Par construction  $a^k \neq 0$ , donc on peut supposer que  $\|a^k\| = 1$ . [11]

**Définition 3.1.** On dit qu'une suite de plans  $\{S_k\}$  converge vers un plan :

$S = \{x : I(x) = \langle a, x \rangle + b = 0\}$  si

$$a^k \rightarrow a, \quad b^k \rightarrow b \quad (k \rightarrow \infty).$$

#### Théorème 3.1 (H. Tuy, 1983)

Dans le processus d'approximation extérieure, soit  $D_1$  un polyèdre compact et supposons que pour chaque sous-suite convergente  $\{x^{kq}\}$  de  $\{x^k\}$  telle que  $x^{kq} \rightarrow \bar{x}$  ( $\bar{x} \notin H$ ) et  $S_{kq} \rightarrow S$  on a  $S$  sépare strictement  $\bar{x}$  de  $H$ .

Alors tout point d'accumulation de  $\{x^k\}$  appartient à  $H$  et résout ainsi  $(P)$ . [11]

**Preuve :**

Soit  $\bar{x}$  un point d'accumulation de  $\{x^k\}$  et  $x^{kq} \rightarrow \bar{x}$ .

On a  $\|a^k\| = 1 \forall k$  donc  $\{a^k\}$  est bornée. D'autre part :

$$-\langle a^k, x \rangle \geq b_k > -\langle a^k, x^k \rangle \quad \forall x \in H$$

Puisque  $D_1$  est un compact, il existe une constante  $M$  telle que  $\|x\| \leq M$  pour tout  $x \in D_1$ .

Par suite

$$|\langle a^k, x^k \rangle| \leq M, \quad |\langle a^k, x \rangle| \leq M \quad \forall x \in H$$

ce qui implique

$$|b^k| \leq \max\{|\langle a^k, x^k \rangle|, |\langle a^k, x \rangle|\} \leq M$$

i.e.  $\{b^k\}$  est bornée aussi.

En prenant une sous-suite on peut donc supposer  $a^{kq} \rightarrow a, b^{kq} \rightarrow b$ , autrement dit  $S^{kq} \rightarrow S$ . Supposons que  $\bar{x} \notin H$ . Alors  $S$  va séparer strictement  $\bar{x}$  de  $H$ , i.e.

$l(\bar{x}) = \langle a, \bar{x} \rangle + b > 0$ . D'autre part  $l_{kq}(x^{kq+r}) \leq 0 \forall r$  et en faisant  $r \rightarrow \infty$  on obtient  $l_{kq}(\bar{x}) \leq 0$ . Lorsque  $S_{kq} \rightarrow S$  on a  $l(\bar{x}) \leq 0$  ce qui est absurde. Donc  $\bar{x} \in H$ .

Puisque  $f(x^k) \leq f(x) \forall x \in D_k$  on a, par la continuité de  $f$ ,  $f(\bar{x}) \leq f(x) \forall x \in H$ , i.e.  $\bar{x}$  est une solution optimale de  $(P)$ . [11]

## Chapitre 3 : Méthode d'Approximation Extérieure

### Théorème 3.2. (R.Horst, Thoai N.V, and H. Tuy, 1989)

Dans le processus d'approximation extérieure, soit  $D_1$  un compact. Supposons que  $l_k$  est continue et pour chaque suite  $\{x^q\} \subset \{x^k\}$  telle que  $x^q \rightarrow \bar{x}$  il existe une sous-suite  $\{x^{q_1}\}$  de  $\{x^q\}$  telle que

1.  $l_{q_1}(x) \rightarrow l(x) \quad \forall x$ .
2.  $l_{q_1}(x^{q_1}) \rightarrow l(\bar{x})$  où  $l(\bar{x}) = 0$  implique  $\bar{x} \in H$ .

Alors tout point d'accumulation de  $\{x^k\}$  appartient à  $H$  et résout  $(P)$ .

#### Preuve :

Soit  $x^q$  une suite convergente vers  $\bar{x}$  ( $\bar{x}$  un point d'accumulation de  $\{x^k\}$ ) et  $\{x^{q_1}\}$  la sous-suite qui satisfait les hypothèses 1 et 2 ; Puisque  $l_{q_1}(x^{q_1+r}) \leq 0 \quad \forall r$ , en faisant  $r \rightarrow \infty$  on a  $l_{q_1}(\bar{x}) \leq 0$ . Par suite,  $l(\bar{x}) \leq 0$ .

D'autre part  $l_{q_1}(x^{q_1}) > 0 \quad \forall l$  et  $l_{q_1}(x^{q_1}) \rightarrow l(\bar{x})$  implique  $l(\bar{x}) \geq 0$ .

Donc  $l(\bar{x}) = 0$  et en vertu de 2)  $\bar{x} \in H$ .

D'une façon analogue comme ci-dessus, on peut conclure que  $\bar{x}$  est une solution optimale de  $(P)$

En particulier, si  $l_k$  sont des fonctions Lipschitz, on a le résultat suivant :

### Théorème 3.3 (H. Tuy, 1983)

Supposons que  $l_k(\cdot)$  est continue et

1.  $|l_k(z) - l_k(x)| \leq L \|z - x\|, \quad \forall z, x \in D_1$  ( $L$  étant une constante).
2. Pour chaque  $\{x^q\} \subset \{x^k\}$  telle que

$$x^q \rightarrow \bar{x}, \quad l_q(x^q) \rightarrow 0$$

On a  $\bar{x} \in H$

Alors tout point d'accumulation de  $\{x^k\}$  appartient à  $H$  et résout  $(P)$ . [11]

#### Preuve :

En vertu de 1) on a  $l_q(x^q) \leq l_q(\bar{x}) + L \|x^q - \bar{x}\|$ .

Par construction,  $l_q(x^j) \leq 0, \quad \forall j > q$ , ce qui entraîne  $l_q(\bar{x}) \leq 0$

Donc  $0 < L_q(x^q) \leq L \|x^q - \bar{x}\|$ .

Quand  $x^q \rightarrow \bar{x}$ , on a  $l_q(x^q) \rightarrow 0$  et  $\bar{x} \in H$ .

Donc, pour assurer la convergence d'un algorithme  $AE$ , les coupes  $l_q(x)$  (qui par définition vérifient 1) et 2) ) devraient satisfaire les hypothèses 1, 2 du théorème 3.

## Chapitre 3 : Méthode d'Approximation Extérieure

### 3.5 Construction de $l_k(\cdot)$

$H$  est un convexe

Dans ce cas, on peut utiliser l'approximation polyédrale. Etant donné un point  $x^k \notin H$ . Il nous faut donc construire un hyperplan (ou une fonction affine  $l_k(\cdot)$ ) séparant  $x^k$  de  $H$ .

Il ya deux méthodes :

1. méthode de sous-gradient.
2. méthode de projection.

En utilisant la première méthode :

#### Méthode de sous-gradient

Soit  $H$  un convexe fermé et  $H^0 = \{x : g(x) < 0\}$ . Cette hypothèse est vérifiée, par exemple, si  $H^0 \neq \emptyset$  où la condition de Slater est satisfaite.

En effet, supposons que  $x \in H^0$ ,  $y \notin H^0$ ,  $p \in \partial g(x)$  alors

$$g(y) \geq 0, \quad g(x) < 0 \quad \text{et} \quad (g(x) - g(y)) \geq \langle p, x - y \rangle$$

ce qui entraîne  $\langle p, x - y \rangle < 0$ . Cela signifie que  $p \neq 0$ , i.e.  $\partial g(x) \setminus \{0\} \neq \emptyset$ .

On considèrera des coupes (linéaires) de la forme

$$l_q(x) = \langle p^k, x - y^k \rangle + \beta_k$$

#### Théorème 3.4 (R.Horst, Thoai N.V, and H. Tuy, 1989)

Soit  $K$  un sous-ensemble compact quelconque de  $H^0$ . Choisissons

$$y^k \in Y^k = \text{conv}(K \cup \{x^k\}) \setminus H^0, \quad p^k \in \partial g(y^k) \setminus \{0\}, \quad \beta_k = g(y^k)$$

( $x^k \text{ argmin } f(D_k)$ ). Alors  $l_k(\cdot)$ , vérifie les conditions (i), (ii) de l'algorithme d'AE et toutes hypothèses du théorème 3.2 [33]

#### Preuve.

1)  $l_k(x) = \langle p^k, x - y^k \rangle + g(y^k) \leq g(x) \leq 0, \quad \forall x \in H.$

2) Si  $H^0 \neq \emptyset$  alors  $K \neq \emptyset$  donc  $y^k = x^k$  et  $l^k(x^k) = g(x^k) < 0$  car  $x^k \notin H$ .

Supposons que  $H^0 \neq \emptyset$  et  $K \neq \emptyset$

Alors tout point  $y^k \in \text{conv}(K \cup \{x^k\}) \setminus H^0$  peut s'écrire sous la forme :

$$y^k = \lambda_k z^k + (1 - \lambda_k)x^k$$

### Chapitre 3 : Méthode d'Approximation Extérieure

Où  $z^k \in K \subset H^0$  et  $\lambda_k \in [0,1]$ .

Puisque  $K$  est compact, il existe  $\delta > 0$  tel que,  $g(z) \leq -\delta, \forall z \in K$ .

C'est facile de voir que :

$$x^k - y^k = \alpha_k(y^k - z^k).$$

Où  $\alpha_k = \lambda_k/(1 - \lambda_k) \geq 0$  et

$$\langle p^k, z^k - y^k \rangle \leq \langle p^k, z^k - y^k \rangle + g(y^k) \leq g(z^k) \leq -\delta < 0.$$

Donc si  $\alpha_k > 0$  alors

$$l_k(x^k) = \langle p^k, z^k - y^k \rangle + g(y^k) = \alpha_k \langle p^k, y^k - z^k \rangle \geq \alpha_k \delta + g(y^k) > 0.$$

Si  $\alpha_k = 0$  on a  $x^k = y^k$  et  $l_k(x^k) = g(x^k) > 0$ .

Soit  $x^q \rightarrow \bar{x}$ . Comme la suite  $\{x^q\}$  est bornée, il en est de même de la suite  $\{y^q\}$  et donc ainsi de la suite  $\{p^q \partial g(y^q)\}$ . On peut donc extraire une sous-suite  $\{x^{q_1}\}$  telle que

$$p^{q_1} = \bar{p}, \lambda_{q_1} \rightarrow \bar{\lambda}, z^{q_1} \rightarrow \bar{z} \in K$$

Lorsque  $l \rightarrow \infty$ , d'où

$$y^{q_1} = \lambda_{q_1} z^{q_1} + (1 - \lambda_{q_1}) x^{q_1} \rightarrow \bar{y} = \bar{\lambda} \bar{z} + (1 - \bar{\lambda}) \bar{x}, \beta_{q_1} = g(y^{q_1}) \rightarrow g(\bar{y}) = \bar{\beta}.$$

Noter que  $\bar{\lambda} < 1$ , car sinon  $\bar{y} = \bar{z} \in K \subset H^0$ . De plus, par la fermeture de l'opérateur multivoque  $\partial g(y)$  on a  $\bar{p} \in \partial g(\bar{y})$ .

Définissons  $l(x) = \langle \bar{p}, x - \bar{y} \rangle + \bar{\beta}$ . Alors

- 1-  $l_{q_1}(x) = \langle p^{q_1}, x - y^{q_1} \rangle + \beta_{q_1} \rightarrow l(x) \forall x$ .
- 2-  $l_{q_1}(x^{q_1}) = \langle p^{q_1}, x^{q_1} - y^{q_1} \rangle + g(y^{q_1}) \rightarrow \langle \bar{p}, \bar{x} - \bar{y} \rangle + \bar{\beta} = l(\bar{x})$ .

D'autre part

$$l_{q_1}(x^{q_1}) = \alpha_{q_1} \langle p^{q_1}, y^{q_1} - z^{q_1} \rangle + g(y^{q_1}) \rightarrow \frac{\bar{\lambda}}{1 - \bar{\lambda}} \langle \bar{p}, \bar{y} - \bar{z} \rangle + \bar{\beta}.$$

Donc si  $l(\bar{x}) = 0$  alors

$$\frac{\bar{\lambda}}{1 - \bar{\lambda}} \langle \bar{p}, \bar{y} - \bar{z} \rangle + \bar{\beta} = 0$$

Puisque

$$\langle p^q, x^q - y^q \rangle \geq \delta > 0 \implies \langle \bar{p}, \bar{x} - \bar{y} \rangle \geq \delta > 0.$$

Par conséquent, on doit avoir

$$\lambda = 0 \text{ et } \beta = g(\bar{y}) = 0 \text{ ou } \bar{x} = \bar{y} \in H.$$

## Chapitre 3 : Méthode d'Approximation Extérieure

### 3.6 Exemple d'application d'AE [11]

**Exemple 1.1.** Soit à résoudre le problème (P) suivant :

$$(P) \begin{cases} \min f(x, y) \\ (x, y) \in H \end{cases}$$

Où  $f(x, y) = x + y$ , et  $H = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \leq 1\}$

**Application de l'algorithme de la méthode d'AE :**

Choisir un hyper rectangle  $D_1$  telle que  $D_1 \supset H$ .

Soit  $D_1$  un carré,  $D_1 = \{(x, y) \in \mathbb{R}^2 / -1 \leq x \leq 1, -1 \leq y \leq 1\}$

Soit à résoudre le problème (P1) suivant :

$$(P_1) \begin{cases} \min f(x, y) \\ (x, y) \in D_1 \end{cases} \Leftrightarrow (P_1) \begin{cases} \min x + y \\ x \leq 1 \\ -x \leq 1 \\ y \leq 1 \\ -y \leq 1 \end{cases}$$

Après la résolution du problème (P<sub>1</sub>) avec la méthode des points intérieurs, la solution trouvée est  $X^1 = (-1.000, -1.0000)$ , et  $f(X^1) = -2.0000$ .

Comme  $X^1 \notin H$ , alors on construit la coupe linéaire  $L_1(X)$  de la manière suivante :

Soit  $g(x, y) = x^2 + y^2 - 1$ ,  $\nabla g(x, y) = (2x, 2y)$ .

$L_1(X) = g(X^1) + \nabla g(X^1)(X - X^1)$ , avec  $X = (x, y)$ .

$X^1 = (-1, -1) \Rightarrow L_1(X) = -2x - 2y - 3 \leq 0$

On ajoute la contrainte linéaire  $L_1(X)$  au problème (P<sub>1</sub>), et comme ça on aura le problème (P<sub>2</sub>)

$$(P_2) \begin{cases} \min x + y \\ x \leq 1 \\ -x \leq 1 \\ y \leq 1 \\ -y \leq 1 \\ -2x - 2y \leq 3 \end{cases}$$

A résoudre toujours avec la méthode des points intérieurs, et on obtiendra la solution  $X^2$ .

$X^2 = (-0.7500, -0.7500)$ ,  $f(X^2) = -1.5000$

Si  $X^2 \in H$ , alors  $X^2$  est la solution optimale globale.

Sinon, on construit une autre coupe linéaire  $L_2(X)$ , et ainsi de suite.

Comme  $X^2 \notin H$ , alors  $L(X) = \frac{3}{2}x - \frac{3}{2}y - \frac{17}{8} \leq 0$ .

On ajoute la contrainte linéaire  $L_2(X)$  au problème (P<sub>2</sub>), et on aura le problème (P<sub>3</sub>),

### Chapitre 3 : Méthode d'Approximation Extérieure

---

suisant :

$$(P_3) \left\{ \begin{array}{l} \min x + y \\ x \leq 1 \\ -x \leq 1 \\ y \leq 1 \\ -y \leq 1 \\ -2x - 2y \leq 3 \\ -\frac{3}{2}x - \frac{3}{2}y \leq \frac{17}{8} \end{array} \right.$$

La résolution du problème  $(P_3)$ , nous donnera la solution suivante :

$$X^3 = (-0.7083, -0.7083), \quad f(X^3) = -1.4167.$$

$$X^3 \notin H \Rightarrow L_3(X) = -1.4166x - 1.4166y - 2.0034 \leq 0.$$

On ajoute la contrainte linéaire  $L_3(X)$  au problème  $(P_3)$ , et on aura le problème  $(P_4)$

suisant :

$$(P_4) \left\{ \begin{array}{l} \min x + y \\ x \leq 1 \\ -x \leq 1 \\ y \leq 1 \\ -y \leq 1 \\ -2x - 2y \leq 3 \\ -\frac{3}{2}x - \frac{3}{2}y \leq \frac{17}{8} \\ -1.416x - 1.4166y \leq 2.0034 \end{array} \right.$$

La résolution du problème  $(P_4)$ , nous donnera la solution suivante :

$$X^4 = (-0.7071, -0.7071), \quad f(X^3) = -1.4142.$$

$X^4 \in H$ , car  $g(X^4) = 1$  d'où  $X^4$  est la solution optimale globale du problème de départ  $(P)$ .



# Chapitre 4

## Programmation dans LINGO

### 4.1 Introduction

La programmation est un ensemble d'outils et de technique permettant de résoudre des problèmes mathématiques par ordinateurs, elle sert à trouver une solution optimale de n'importe quel type de problèmes.

Le processus de résoudre un problème mathématiques exige un grand nombre de calculs donc il est mieux de l'exécuter par machine. Pour cela on a choisi le logiciel LINGO qui est spécifié pour résoudre des problèmes d'optimisations (linéaire, non linéaire, convexe, non convexe, etc).

Il comporte un éventail de commandes qui peuvent être appelés à tout moment.

### 4.2 Présentation du logiciel

LINGO est livré avec un jeu de solveurs pour l'optimisation. En entrée, ce solveur LINGO doit recevoir un modèle mathématique et il possède quatre solveurs qu'il utilise afin de résoudre les différents types de modèles :

- Solveur direct.
- Solveur linéaire.
- Solveur non linéaire.
- Méthode de type séparation et évaluation.

De plus, LINGO est:

- Un moyen pour confirmer que l'optimum trouvé est l'optimum global.
- Un moyen amélioré pour résoudre beaucoup de type de problèmes.
- Un moyen de décomposition si un modèle contient des sous-modèles.
- Possible de résoudre les problèmes plus rapidement.

**Les fonctions utilisées dans un modèle de LINGO sont :**

@FOR-utilisée pour produire des contraintes

@SUM-calcul de la somme.

@MAX-recherche le maximum.

@MIN-recherche le minimum.

## Chapitre 4 : Programmation dans LINGO

### Type de variables dans LINGO

Toutes les variables dans un modèle LINGO sont considérées non négatives et continues. Les fonctions variables d'un modèle de LINGO sont: @GIN-toute valeur positive de nombre entier.  
@BIN-une valeur binaire (0 ou 1).  
@FREE-toute valeur positive ou négative réelle.  
@BND-toute valeur bornée par des limites indiquées.

Sa forme générale pour la déclaration d'une variable  $x$  en utilisant la fonction @BND qui inclut les bornes inférieures et supérieures est:

@BND (borne inférieure,  $x$ , borne supérieure)

Pour savoir plus sur ce logiciel, regardez les références: [4],[6],[8]

On peut sélectionner les options à suivre pour résoudre un modèle en cliquant sur LINGO options. Dans la fenêtre suivante on sélectionne "global solver".

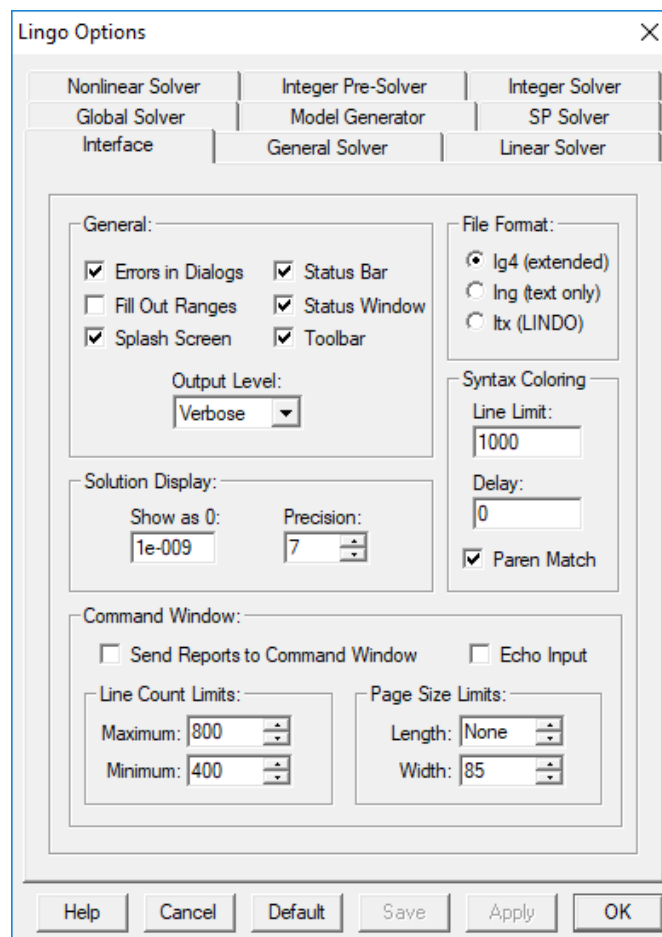


Fig. 4. 1

## Chapitre 4 : Programmation dans LINGO

Ensuite on fait rentrer le modèle de problèmes à résoudre.

Pour ouvrir un modèle enregistré en cliquant sur "open" dans la barre des menus, la fenêtre suivante apparaît :

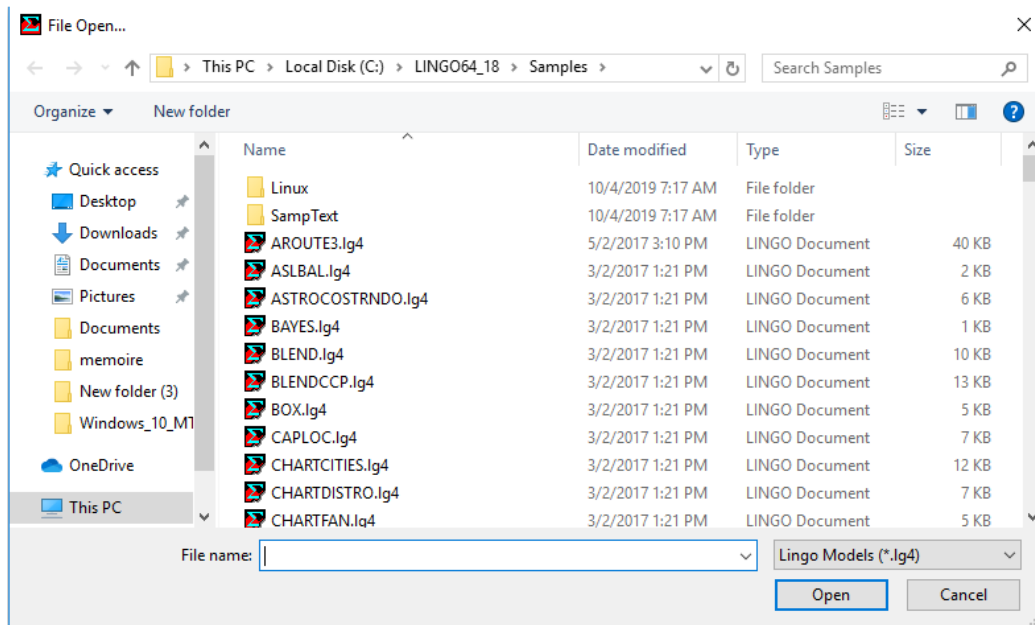


Fig. 4. 2

On sélectionne le modèle à ouvrir en cliquant sur "ouvrir" dans la fenêtre suivante le modèle sélectionné s'ouvre dans la fenêtre "LINGO model":

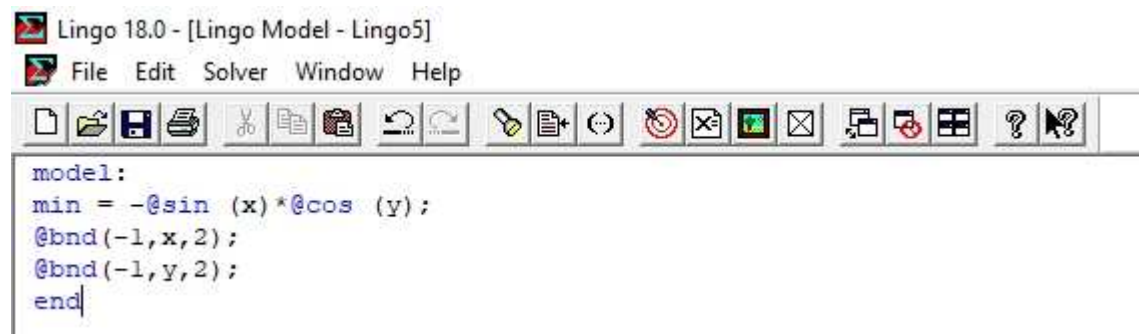


Fig. 4. 3

## Chapitre 4 : Programmation dans LINGO

Si aucune erreur n'est signalée alors la fenêtre suivante apparaît :

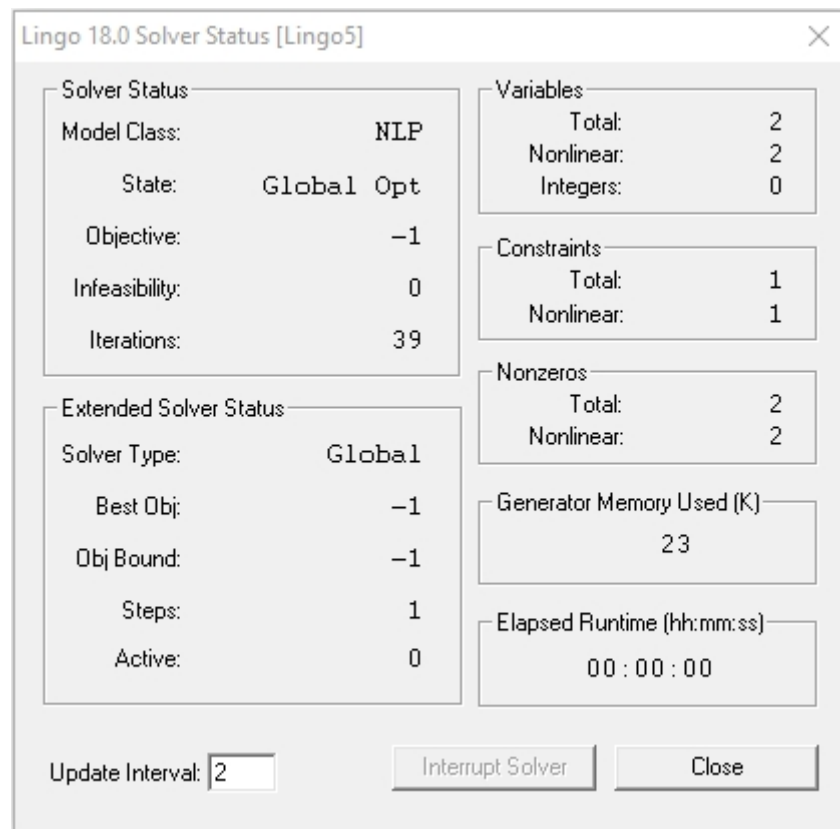


Fig. 4. 4

Cette fenêtre fournit des informations sur le nombre de variables non linéaires, le type de la solution obtenue (locale ou globale) la valeur de la fonction objectif, le nombre d'itérations requise pour résoudre le modèle.

En fermant cette fenêtre on peut alors regarder la fenêtre de rapport de solution.

### 4.3 Résolution des exemples sur LINGO avec la méthode Branch-and-Bound

#### Exemple 1:

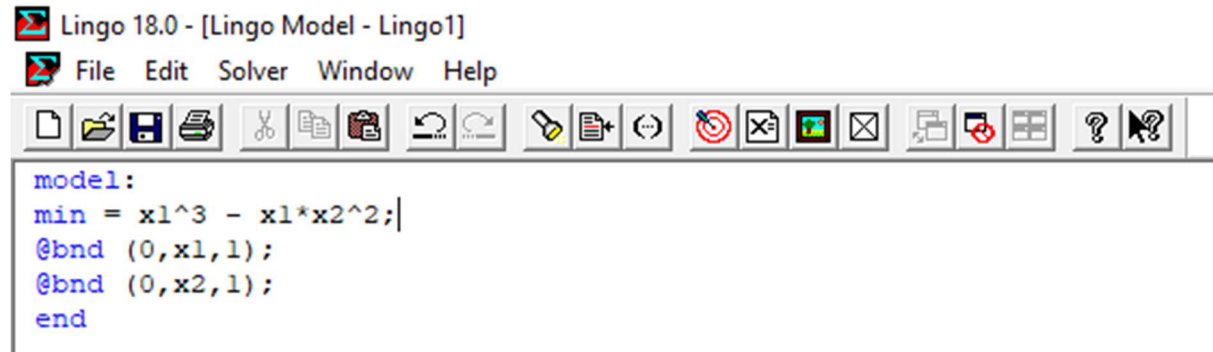
Soit à résoudre le problème suivant:

$$(P) \begin{cases} \min f(x) \\ X \in H \end{cases}$$

$$\text{Où } f(x_1, x_2) = x_1^3 - x_1 x_2^2$$

$$X = (x_1, x_2) \in H, H = [0,1] \times [0,1], \alpha = 1.25, \epsilon = 10^{-5}$$

Le modèle d'exécution sur LINGO est le suivant:



The screenshot shows the LINGO 18.0 application window titled "Lingo 18.0 - [Lingo Model - Lingo1]". The menu bar includes "File", "Edit", "Solver", "Window", and "Help". The toolbar contains various icons for file operations and solving. The main text area displays the following LINGO model code:

```
model:  
min = x1^3 - x1*x2^2;  
@bnd (0,x1,1);  
@bnd (0,x2,1);  
end
```

Fig. 4. 5-modèle d'exécution

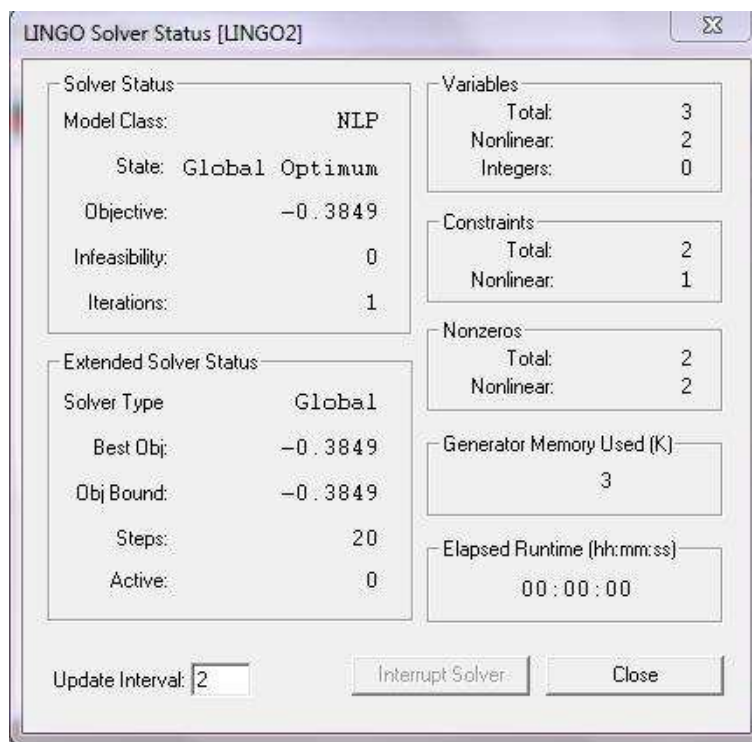


Fig. 4. 6 – Résultat de l'exécution du model

## Chapitre 4 : Programmation dans LINGO

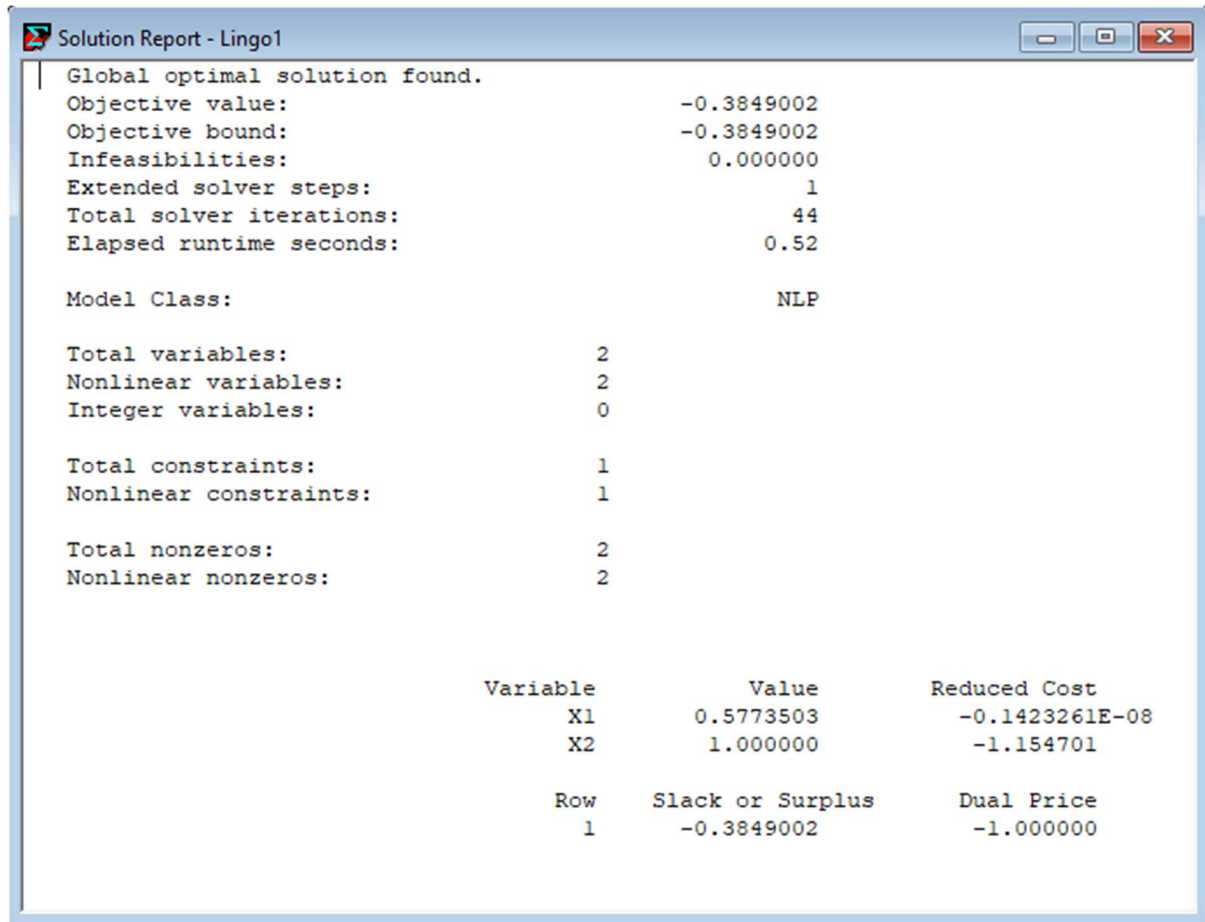


Fig. 4.7 – Résultat de l'exécution du model

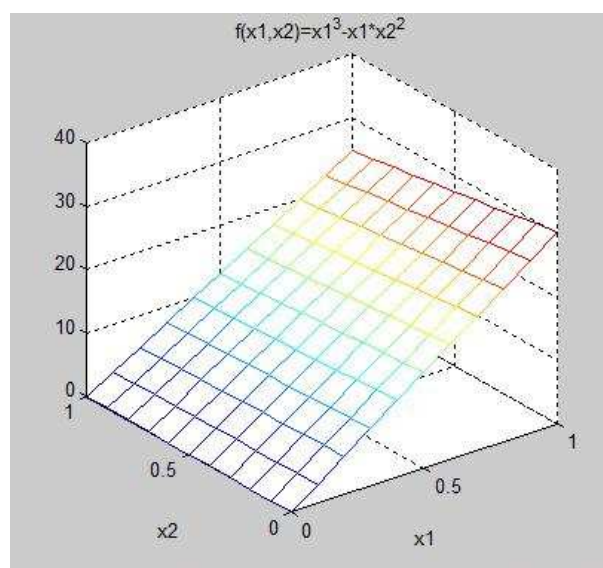


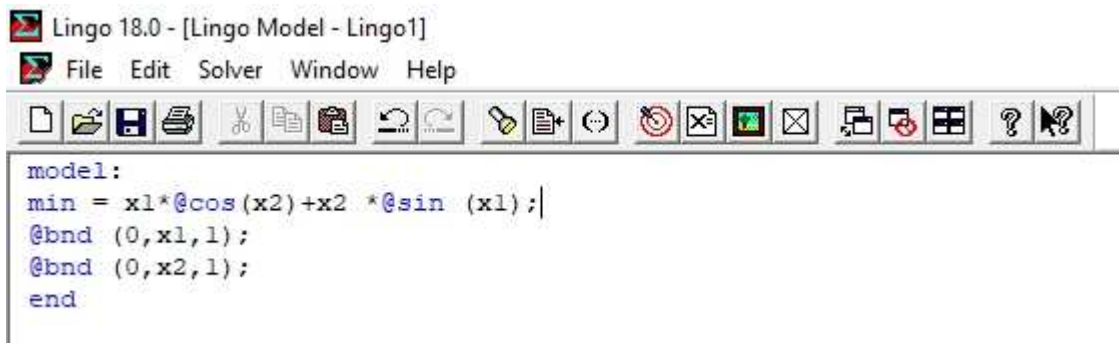
Fig. 4.8 – Le graphe de la fonction

## Chapitre 4 : Programmation dans LINGO

### Exemple 2 :

$f(x_1, x_2) = x_1 \cos x_2 + x_2 \sin x_1$  avec  $(x_1, x_2) \in H, H = [0,1] \times [0,1], \alpha = 0.635, \epsilon = 10^{-5}$

Le modèle d'exécution sur LINGO est le suivant:

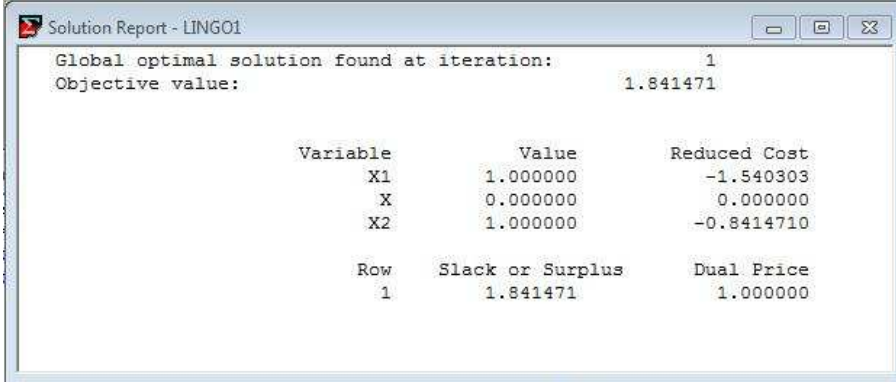


```
model:
min = x1*@cos(x2)+x2 *@sin(x1);
@bnd(0,x1,1);
@bnd(0,x2,1);
end
```

Fig. 4. 9 – Résultat de l'exécution du model



Fig. 4. 10 – Résultat de l'exécution du model



Solution Report - LINGO1

Global optimal solution found at iteration: 1  
Objective value: 1.841471

Variable	Value	Reduced Cost
X1	1.000000	-1.540303
X	0.000000	0.000000
X2	1.000000	-0.8414710

Row	Slack or Surplus	Dual Price
1	1.841471	1.000000

Fig. 4.11 – Résultat de l'exécution du model

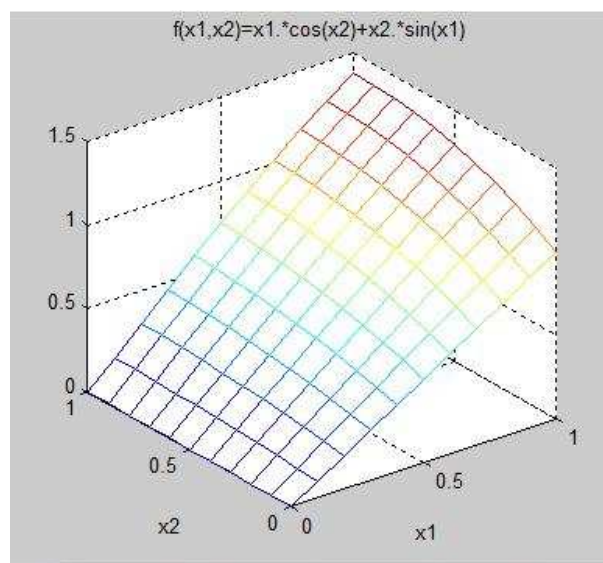


Fig. 4.12 – Le graphe de la fonction

## Conclusion général

Dans ce projet, on c'est intéresser a l'optimisation combinatoire plus précisément l'optimisation global.

On présenté la description d'un problème d'optimisation et on parler des outils fondamentaux les plus importantes lié a l'optimisation tel que la différentiabilité, la convexité

Au deuxième chapitre nous avons présenté, étudié, développé et évalué la résolution des problèmes d'optimisation globale en exploitons la méthode par séparation et évaluations (branch and bound) ainsi que ces procédures d'évaluation, la stratégie du parcours et l'efficacité de son algorithme tout en traitons quelques exemples numériques avec une constante  $\alpha$ , la vérification des résultats a été faite sur le logiciel LINGO en programmant un exemple au dernier chapitre

Au troisième chapitre on a présenté la méthode d'approximation extérieure en parlant de son principe et expliqué comment fonctionne son algorithme avec la présentations de quelque exemples.

L'optimisation est un domaine très vaste et l'usage de l'optimisation global est un peut délaissé ces dernière année et tout ce focalise sur l'optimisation local

Dans notre travail on a montrer l'efficacité de la méthode de séparation et évaluations (branch and bound) et la méthode d'approximation extérieure a trouver la solution optimum pour la résolution d'un problème d'optimisation.

# Bibliographie

- [1] Sid Sihem mémoire de master : étude et comparaison d'algorithmes d'optimisation (université l'arbi ben mehdi-oum el bouaghi 2016-2017)
- [2] METHODE DE BRANCH AND BOUND (CHAPITRE 06) .
- [3] M. Bergounioux. Optimisation et contrôle des systèmes linéaires, Dunod, Paris,(2001)
- [4] M. Bierlaire. Introduction à l'optimisation différentiable. Presses polytechniques et universitaires romandes, Lausanne, (2006).
- [5] Melle BOUZEBODJA Tassadit. Optimisation globale avec application pour fonction lipschiziennes. Ingénieur d'état en recherche opérationnelle.
- [6] Melle Larachiche imene et Melle remadelia amina. Algorithme Branch and bound appliquée au problème de sac à dos . université Djilali Bounaama (2018)
- [7] R. Danchin. Cours de calcul différentiel. Licence de Mathématiques, 3ème année, (2010).
- [8] J. Royer. Calcul Différentiel et intégral. Université Toulouse 3, (2013-2014).
- [9] Mr OUANES .Mohand ."Optimisation semi-infinie et Optimisation Globale Théorie et Algorithmes". Thèse d'état, UMMTO,(2006).
- [10] L. Pujo-Menjouet. Cours de calcul différentiel. Université Claude Bernard, Lyon, France.
- [11] Houria BEKRI. Optimisation globale avec application pour les fonctions de holder. Mémoire master UMMTO 2012.
- [12] GATER Naima et CHIKHAOUI Manel : Optimisation globale des fonctions deux fois différentiable. Mémoire master. UMMTO 2012.
- [13] METHODE DE BRANCH AND BOUND (CHAPITRE 03) .
- [14] OPTIMISATION EN NOMBRE ENTIER .