

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mouloud Mammeri de Tizi ousou
Faculté des sciences
Département des mathématiques



Mémoire de fin d'étude



En vue de l'obtention du diplôme Master

Thème

OPTIMISATION SEMI-INFINIE

Les membres de jury :

OUANES Mohand

Mme RABIA

Mr KASDI.K

Réalisé par :

SINI Nabil

LOUNIS Ferhat

Promotion 2013-2014

Remerciement

Au terme de ce modeste travail, nous tenon à exprimer nos remerciements les plus sincères, tout d'abord au « BON DIEU » pour la patience et la santé qui nous a été utile tout au long de notre parcours.

Nous ne saurons trouver les termes qu'il faut pour exprimer notre profonde gratitude et la reconnaissance que nous devons à notre promoteur M^r OUANES Mohand auprès de la qu'il nous avons trouvé toujours des conseils bien veillant ainsi que l'aide, l'appui et la compréhension.

Nous tenons à remercier très respectueusement les membres du jury.

Nous exprimons nos remerciement à tous nos enseignants, et tous ceux qui ont aidé, de près ou de loin pour l'élaboration de ce travail.

Table des matières

0.1	Notations et abréviations	6
0.2	Introduction générale	7
1	Convexité et optimisation	10
1.1	La convexité	10
1.1.1	Définitions et propriétés	10
1.2	Optimisation non linéaire	13
1.2.1	Définitions et propriétés	13
1.2.1.1	Rappels sur les matrices	17
1.2.2	Optimisation non linéaire sans contraintes . .	19
1.2.2.1	Condition nécessaire d'optimalité du premier ordre	20
1.2.2.2	Condition nécessaire d'optimalité de second ordre	20
1.2.2.3	Conditions suffisantes d'optimalité de second ordre	21
1.2.3	Optimisation avec contraintes :	22
1.2.3.1	Condition de Khun-Tucker	24
1.2.3.2	La Dualité	25
1.2.3.3	Méthode du lagrangien	27
1.2.3.3.1	Introduction	27
1.2.3.3.2	Principe de la méthode	27
1.2.3.4	La méthode de direction admissible	29
1.2.3.4.1	Principe de la méthode	29

1.2.3.4.2	Algorithme de la méthode	30
1.3	Optimisation convexe	31
1.3.1	Optimisation convexe sans contrainte	31
1.3.1.1	Introduction	31
1.3.1.2	Condition nécessaire et suffisante	31
1.3.1.3	Algorithme du gradient à pas optimal	32
1.3.1.4	cas quadratique	32
1.3.1.5	Algorithme du gradient à pas optimal pour le cas quadratique	33
1.3.2	Optimisation convexe avec contraintes d'égalités	33
1.3.2.1	Position du problème	33
1.3.2.2	Condition nécessaire et suffisante	34
1.3.2.3	Cas quadratique	34
1.3.3	Optimisation convexe avec contraintes d'inégalité	35
1.3.3.1	Position du problème	35
1.3.3.2	Condition nécessaire et suffisante	35
1.3.3.3	Méthode d'Uzawa	36
1.3.3.3.1	Introduction	36
1.3.3.3.2	Algorithme d'Uzawa	36
1.3.3.4	Cas quadratique	37
1.3.3.5	La méthode de la stratégie des contraintes actives	37

2	Les principales méthodes en optimisation semi-infinie	39
2.1	Introduction	39
2.2	Formulation d'un problème semi-infini	40
2.2.1	Définitions et propriétés	41
2.3	Méthode du Lagrangien	43
2.3.1	Condition nécessaire du premier ordre	43
2.3.2	Condition du second ordre	43
2.4	Résolution par la discrétisation	44
2.4.1	Problème semi-infini discrétisé	44
2.4.2	Méthode discrétisation	45
2.4.2.1	Principe de la méthode	45
2.4.2.2	Test d'optimalité	45
2.4.3	Algorithme de résolution de (P.S.I) par la discrétisation	46
2.5	Méthode d'échange	47
2.5.1	Introduction	47
2.5.2	Le principe de la méthode d'échange	47
2.5.3	Test d'optimalité	47
2.5.4	Algorithme de résolution de (P.S.I) par la méthode d'échange	49
3	Paramétrisation dual en optimisation semi-infinie quadratique	51
3.1	Introduction	51
3.2	Formulation d'un problème semi-infini quadratique	52
3.3	Problème dual d'un problème semi-infini quadratique	53

3.4	Le dual paramétré d'un problème de programmation semi-infinie quadratique	56
3.5	Principe de la méthode	58
3.6	Algorithme de la méthode	59
3.7	La convergence de l'algorithme	60
4	Conclusion	62

0.1 Notations et abréviations

0_n désigne le vecteur nul dans l'espace euclidien \mathbb{R}^n .

$\|\cdot\|$ représente la norme euclidienne.

$\langle \cdot \rangle$: Produit scalaire.

vap : valeur propre.

v : valeur optimal.

SIP : Problème semi-infini.

LSIP : Problème semi-infini linéaire.

min : minimum.

max : maximum.

ssi : si et seulement si.

resp : respectivement.

CN : Condition nécessaire.

CS : Condition suffisante.

CNS : Condition nécessaire et suffisante.

0.2 Introduction générale

L'optimisation est un outil important en sciences et pour l'analyse des systèmes physiques. On doit d'abord identifier quelques objectifs (le profit, le temps, ou n'importe quelle quantité ou combinaison qui peut être représentée avec une valeur algébrique).

l'objectif est dépend de variables ou inconnues, notre but est de déterminer les valeurs des variables qui optimisent l'objectif.

Le processus « modélisation » permet de l'identification des objectifs, des variables. La connaissance et l'intuition du modélisateur sont importantes aussi bien dans la formulation d'un modèle avec un degré approprié de complexité que dans la l'interprétation des résultats du processus d'optimisation. Dès que le modèle a été formulé, un algorithme d'optimisation peut être utilisé pour la résolution du problème.

Bien que les origines de *SIP* sont liés à l'approximation de Chebyshev, le travail classique de Haar [19] sur des systèmes semi-infinis linéaires, et au Fritz Jean [21].

Le terme de *SIP* a été inventé en 1962 par Charnes, Hettich et Kortanek [14,16,17,18] dans certains articles consacrés à *LSIP*.

Gustavson et Kortanek [13,15] proposé, au cours de début des années 1970, les premières méthodes numériques pour les modèles de *SIP* dans les applications découlant.

La publication vers 1980 des six livres suivants converti *SIP* dans une sous-zone mature et indépendante en optimisation. Deux volumes de notes de cours sur Mathématiques, édité par Hettich [24] et par Fiacco et Kortanek [26], et quatre monographies par Tichatschke [25], Glashoff et Gustafson [23], Hettich et Zencke [20], et Brosowski [27] (consacré à la stabilité dans *SIP*). Plus récemment, Goberna et L'Opez [12] présentés dans une approche extensive à *LSIP*, y compris les aspects théoriques et numériques. Livres d'optimisation réputés consacrés certains chapitres sur *SIP* .

Nous mentionnons également l'article de Polak [22] (Les méthodes d'indications possibles sur les fondations des mathématiques), et Hettich et Kortanek [14] interrogées, dans une superbe manière, des résultats théoriques, méthodes et applications de *SIP*.

Récemment Goberna et L'Opez [12] ont examinés le modèle de *LSIP*.

Au cours des dernières années, les technologies cherchent de plus en plus à traiter des systèmes complexes, constitués par un grand nombre de paramètres liés les uns aux autres par une structure bien déterminée.

L'optimisation est un ensemble des méthodes qui permet d'obtenir le meilleur résultat. Optimiser revient donc à estimer des minimums et des maximums d'une fonction ou d'un système de fonction.

Les modèles de programmation linéaire et non linéaire permettent d'aborder un grand nombre de problèmes d'optimisation en apparence très différent, dans des contextes divers.

Il est important de bien connaître les caractéristiques du problème posé, afin d'identifier la technique appropriée pour sa résolution.

Les problèmes d'optimisation sont classés en fonction des caractéristiques mathématiques de la fonction objectif, des contraintes et des variables d'optimisation, les plus importantes classes d'optimisation sont :

*) Optimisation linéaire avec f est une fonction linéaire : $f(x) = \langle c, x \rangle$ et S est défini par des fonctions affines : $ax + b \geq 0$

*) Optimisation linéaire quadratique f est une fonction convexe quadratique : $f(x) = \frac{1}{2}x^T Ax + bx + c$, A est une matrice symétrique semi-définie positive, S est défini par des fonctions affines : $ax + b \geq 0$.

*) Optimisation convexe, f est une fonction convexe et S un domaine convexe.

*) Optimisation différentiable, f est une fonction différentiable et S est défini par des fonctions (= contraintes) différentiables.

*) Optimisation non différentiable, exemple : $f(x) = \max \{f_1(x), \dots, f_m(x)\}$

*) Optimisation en dimension infinie, exemple : problèmes variationnels $J(x) = \int_0^T L(t, x(t), \dot{x}(t)) dt$ avec x dans un ensemble de fonctions X , $x(0) = x_0$ et $x(T) = x_T$.

Les problèmes semi-infinis sont des problèmes qui ont une infinité de contraintes, et présentant des structures particulières de linéarité et convexité.

Notre travail s'inscrit dans cette problématique, nous nous intéressons aux problèmes semi-infinis non linéaire et le cas quadratique qui possèdent certaines particularités de structure.

Notre travail est organisé comme suit :

Dans le premier chapitre, nous présentons les notions de convexité et les notions de base sur l'étude de l'optimisation non linéaire, comme nous avons décrit quelques méthodes de résolution d'un problème non linéaire (la méthode de Lagrangien, la méthode de direction admissible), afin d'établir clairement l'optimisation convexe.

Le deuxième chapitre consacré à les principales méthodes en optimisation semi-infinie, nous présentons la méthode de discrétisation pour la résolution d'un problème semi-infini, c'est une méthode consiste à résoudre dans chaque une de ces itérations un sous problème fini, et aux méthodes d'échanges qui sont en général plus efficient que les algorithmes de discrétisation pures pour la résolution d'un problème non linéaire.

Entamons dans le troisième chapitre la méthode de paramétrisation dual pour les problèmes de programmation semi-infinie quadratique qui est développée récemment pour des problèmes de *SIP* avec une seule contrainte infinie. En outre, un algorithme de paramétrisation dual est également proposé pour la solution numérique de ces problèmes avec un résultat de convergence. L'efficacité de l'algorithme est illustré par la résolution d'un exemple numérique.

1 Convexité et optimisation

1.1 La convexité

1.1.1 Définitions et propriétés

Définition 1.1 :

Un ensemble de points est dit convexe si tout segment joignant deux points quelconques de l'ensemble appartient à l'ensemble.

Définition 1.2 :

Soit $S \subset \mathbb{R}^n$, on dit que :

S est convexe $\Leftrightarrow \lambda x + (1 - \lambda)y \in S, \forall x, y \in S \forall \lambda \in [0, 1]$

$\Leftrightarrow \sum_{i=1}^k \lambda_i x_i \in S \forall (x_i)_{i=1, \dots, k} \in S$ et $\forall \lambda_i \in \mathbb{R}_{\geq 0}$ t.q. $\sum_{i=1}^k \lambda_i = 1$

Définition 1.3 :

Soit $S \subset \mathbb{R}^n$ convexe, soit $f : S \rightarrow \mathbb{R}$, on dit que :

1) f est convexe sur $S \iff \forall x, y \in S$ et $\forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$

2) f est strictement convexe sur $S \iff \forall x, y \in S$ et $\forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$

Exemple 1.1 :

$f(x) = x^4$ et $f(x) = x^2 + y^2$ sont strictement convexes sur \mathbb{R} et \mathbb{R}^2 .

Définition 1.4 :

Soit $S \subset \mathbb{R}^n$ convexe, soit $f : S \rightarrow \mathbb{R}$, on dit que :

1) f est concave sur S

$\iff \forall x, y \in S$ et $\forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$

2) f est strictement concave sur S

$$\iff \forall x, y \in S \text{ et } \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$$

Remarque 1.1 :

Si f est concave alors $-f$ est convexe.

Définition 1.5 :

L'épigraphe d'une fonction f noté $\text{epi}(f)$ est l'ensemble défini par :

$$\text{epi}(f) = \{(\mu, x) \in \mathbb{R}^{n+1} \mid f(x) \leq \mu, x \in \mathbb{R}^n, \mu \in \mathbb{R}\}$$

Propriété 1.1 : [1]

Une fonction f est convexe si et seulement si $\text{epi}(f)$ est un ensemble convexe.

Définition 1.6 :

Un ensemble $C \subset \mathbb{R}^n$ est un cône s'il est stable pour la multiplication par des scalaires positifs et l'addition. ($x, y \in C \Rightarrow x + y \in C$ et $\lambda x \in C \forall \lambda \geq 0$)

Remarquons, que cette définition implique qu'un cône est un ensemble convexe.

Définition 1.7 :

Soit A un ensemble de \mathbb{R}^n on appelle enveloppe convexe de A , le plus petit ensemble convexe contenant A , noté $\text{Conv}(A)$, avec : $A \subset \text{Conv}(A)$.

Définition 1.8 :

Soient $a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ et $\beta \in \mathbb{R}$

alors l'ensemble $H = \{(x \in \mathbb{R}^n, ax = \beta)\}$ est appelé un hyperplan de \mathbb{R}^n .

Exemple 1.2 :

$H = \{(x_1, x_2) \in \mathbb{R}^2 / a_1 x_1 + a_2 x_2 = \beta \text{ avec } a_1, a_2, \beta \in \mathbb{R}\}$ est un Hyperplan de \mathbb{R}^2 qui est une droite.

Propriété 1.2 : [4]

Un Hyperplan est un ensemble convexe.

Définition 1.9 :

Soient : $a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ et $\beta \in \mathbb{R}$ on appelle :

Demi espace positif fermé, l'ensemble $H_+ = \{(x \in \mathbb{R}^n, ax \geq \beta)\}$,

Demi espace négatif fermé, l'ensemble $H_- = \{(x \in \mathbb{R}^n, ax \leq \beta)\}$,

Demi espace positif ouvert, l'ensemble $H_+^\circ = \{(x \in \mathbb{R}^n, ax > \beta)\}$,

Demi espace négatif ouvert, l'ensemble $H_-^\circ = \{(x \in \mathbb{R}^n, ax < \beta)\}$.

Définition 1.10 :

L'intersection d'un nombre fini de demi-espace fermé de \mathbb{R}^n est appelée polytope convexe.

Définition 1.11 :

On appelle polyèdre, un polytope convexe borné (c'est un compact convexe de \mathbb{R}^n).

Propriété 1.3 : [4]

Un ensemble de \mathbb{R}^n est dite compact, s'il est fermé et borné.

Exemple 1.3 :

1. Les compacts convexes dans \mathbb{R}^2 sont : les triangles, les cercles, \dots
2. Les compact convexes dans \mathbb{R}^3 sont : les sphères, \dots

1.2 Optimisation non linéaire

1.2.1 Définitions et propriétés

Définition 1.13 : (Fonction linéaire)

Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est dite linéaire si elle s'écrit comme suit :

$$f(x) = C^\top x = \sum_{i=1}^{i=n} c_i x_i$$

-) C : un vecteur de \mathbb{R}^n constant, indépendant de x .
-) $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est dite linéaire si chaque'une de ses composantes
-) $f_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, m$, est linéaire : $f(x) = Ax$ ou $A \in \mathbb{R}^{m \times n}$ est une matrice constante.

Définition 1.14 : (Fonction affine)

1) Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est dite affine si elle s'écrit :

$$f(x) = C^\top x + d$$

-) C : un vecteur de constante de \mathbb{R}^n .
-) d : une constante de \mathbb{R} .

2) Une fonction $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est affine si chaque'une de ses composantes : est affine.

$g_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, m$ est affine, tel que :

$$g_i(x) = Ax + b$$

-) A : une matrice de $\mathbb{R}^{m \times n}$.
-) b : un vecteur de constante de \mathbb{R}^m .

Définition 1.15 : (Fonction non linéaire)

toute fonction qui n'est pas affine est dite non linéaire.

Définition 1.16 : (Dérivées partielles)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue, la fonction notée $:\nabla_i f(x) : R \rightarrow \mathbb{R}$, également notée par :

$$\frac{\partial f(x)}{\partial(x_i)}$$

est appelée *i*^{ème} dérivée partielle de f , elle est définie par :

$$\lim_{\alpha \rightarrow 0^>} \frac{f(x_1, \dots, x_i + \alpha, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{\alpha}$$

Définition 1.17 : (Gradient)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiable, la fonction notée $\nabla f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$

est appelée le gradient de f définie par :

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial(x_1)} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial f(x)}{\partial(x_n)} \end{pmatrix}$$

Définition 1.18 : (Dérivée Directionnelle)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue et soient $x \in \mathbb{R}^n$, $d \in \mathbb{R}^n$, La dérivée directionnelle de f en x dans la direction d est donnée par :

$$D(x) = \lim_{\alpha \rightarrow 0^>} \frac{f(x + \alpha d) - f(x)}{\alpha}$$

Si la limite existe, de plus, le gradient de f existe, alors la dérivée directionnelle est le produit scalaire entre le gradient de f et la direction d :

$$D(x) = (\nabla f(x))^T d$$

Définition 1.19 : (Fonction différentiable)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue. Si pour tous $d \in \mathbb{R}^n$, La dérivée directionnelle de f dans la direction d existe, alors la fonction f est dite différentiable.

Exemple 1.4 :

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

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

le gradient de f est donnée par :

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

la dérivée directionnelle de f en $x = (x_1, x_2)$ dans la direction d est :

$$D(x) = (d_1, d_2) \nabla f(x_1, x_2) = d_1 + 2x_2 d_2$$

Remarque 1.1 :

La dérivée partielle égale a la dérivée directionnelle dans la direction des axes de coordonnées :

$$\frac{\partial f(x)}{\partial x_i} = \nabla f(x)^\top e_i$$

avec :

e_i : le $i^{\text{ème}}$ vecteur de la base canonique de \mathbb{R}^n .

Définition 1.20 : (Fonction quadratique)

Soit A une matrice symétrique d'ordre n , b est un vecteur d'ordre n et c un nombre réel. L'application q de \mathbb{R}^n dans \mathbb{R} définie par :

$$q(x) = c + b^\top x + \frac{1}{2} x^\top A x$$

s'appelle fonction quadratique

Définition 1.21 : (Direction de descente)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $x \in \mathbb{R}^n$ un vecteur $d \in \mathbb{R}^n$ est une direction de descente au point x si et seulement si :

$$\exists \theta > 0 : f(x + \lambda d) < f(x), \forall \lambda \in]0, \theta[$$

Théorème 1.1 :[17]

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable au point $x \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$ un vecteur qui vérifie :

$$d^\top \nabla f(x) < 0$$

alors : d est une direction de descente au point x .

Théorème 1.2 : (Convexité par le gradient)

1. $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiable sur un ensemble convexe X . f est convexe sur un ensemble convexe X si et seulement si :

$$f(y) - f(x) \geq (y - x)^\top \nabla f(x) , \forall x, y \in X$$

2. $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ est strictement convexe sur un ensemble convexe X si et seulement si :

$$f(y) - f(x) > (y - x)^\top \nabla f(x) , \forall x, y \in X$$

Définition 1.22 : (Matrice jacobienne)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

La fonction notée $J(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$ définie par :

$$J(x) = (\nabla f(x))^\top = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \dots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

est appelée matrice Jacobienne.

Définition 1.23 : (Matrice Hessienne)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ deux fois différentiable. La fonction notée $\nabla^2 f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ définie par :

$$H = \nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \dots & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \dots & \dots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

est appelée matrice Hessienne.

proposition 1.1 :[5]

1) $H_f(x)$ est une matrice symétrique car : $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ dès que f est deux fois dérivable (Schwarz).

2) On a le développement de Taylor suivant :

$$f(x+v) = f(x) + (\nabla f(x))^\top v + \frac{1}{2} v^\top H_f(x) v + \varepsilon(v)$$

avec $\frac{\|\varepsilon(v)\|}{\|v\|} \rightarrow 0$ quand $\|v\|^2 \rightarrow 0$.

Exemple 1.5 :

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

le gradient de f est donné par :

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

le Hessienne de f est donné par :

$$H = \nabla^2 f(x_1, x_2) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

1.2.1.1 Rappels sur les matrices .

Définition 1.24 :

Soit $M \in \mathbb{R}^{n \times n}$ symétrique.

1. On dit que M est semi définie positive ssi $x^t M x \geq 0 \forall x \in \mathbb{R}^n$.
2. On dit que M est définie positive ssi $x^t M x > 0 \forall x \in \mathbb{R}^n / \{0\}$.

Remarque 1.2 :

On définit de la même façon les matrices semi définie négative et définie négative en renversant les inégalités.

Théorème 1.3 : [18] (Convexité par le Hessien)

Soit $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction deux fois différentiable sur un ensemble convexe X , si $\nabla^2 f(x)$ est semi définie positive (resp. défini positive) pour tous x dans X alors f est convexe (resp. strictement convexe)

Remarque 1.3 :

1. Toute matrice symétrique semi définie positive n'admet pas des valeurs propres négatives.
2. Toute matrice symétrique définie positive, admet des valeurs propres strictement positives.

Définition 1.25 :

1. On appelle le rang colonne d'une matrice, le nombre maximum de colonnes linéairement indépendantes.
2. On appelle le rang ligne d'une matrice, le nombre maximum de lignes linéairement indépendantes.

Théorème 1.4 : [17]

Quelque soit la matrice A , le rang ligne égale au rang colonne et on le note par $\text{rang}(A)$

Proposition 1.2 : [19]

Soit $A \in \mathbb{R}^{r \times n}$ si $\text{rang}(A) = r \leq n$ alors :

$\forall b \in \mathbb{R}^r$, le système : $Ax = b$ admet une solution $x \in \mathbb{R}^n$.

Définition 1.26 :

On dit que $A \in \mathbb{R}^{n \times n}$ est non singulière si :

$$\text{rang}(A) = n$$

Définition 1.27 : (Point stationnaire)

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, un point x de X est dite stationnaire (ou critique) de f si : $\nabla f(x) = 0$. C'est donc un point où le gradient s'annule.

On note $S(f)$: l'ensemble des stationnaire de f .

Proposition 1.3 : [19]

Si f admet un extrémum local en x , alors x est un point stationnaire.

Remarque 1.4 :

-) Pour déterminer les extrêmes locaux, on cherche donc tout d'abord à déterminé les points stationnaires.
-) La matrice hessienne permet alors de conclure, s'il s'agit bien (ou non) d'un maximum ou d'un minimum.

1.2.2 Optimisation non linéaire sans contraintes**Définition 1.28 :**

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, on appelle problème de minimisation sans contraintes le problème (P_1) suivant :

$$(P_1) \Leftrightarrow \begin{cases} \min f(x) \\ x \in \mathbb{R}^n \end{cases}$$

-) f est appelée critère ou fonction de coût.

Exemple 1.6 : (Un problème de production)

Une usine a besoin de n produits a_1, \dots, a_n en quantité x_1, \dots, x_n pour fabriquer un produit fini b en quantité $h(x_1, \dots, x_n)$.

Soit p_i le prix unitaire du produit i et p le prix de vente du produit fini. On désire calculer la production optimale. Il s'agit donc de maximiser la fonction $ph(x_1, \dots, x_n) - \sum_{i=1}^n x_i p_i$

Définition 1.29 :

- 1) On dit que x_0 est un minimum local si

$$\forall x \in V(x_0), f(x) \geq f(x_0)$$

2) On dit que x_0 est un minimum local strict si

$$\forall x \in V(x_0), f(x) > f(x_0)$$

3) On dit que x_0 est un minimum global si

$$\forall x \in \mathbb{R}^n, f(x) \geq f(x_0)$$

4) On dit que x_0 est un minimum global si

$$\forall x \in \mathbb{R}^n, f(x) > f(x_0)$$

On définit de la même façon un maximum local, maximum local strict, un maximum global, maximum global strict, en renversant les inégalités.

Propriété 1.4 : [9]

Si x_0 est un minimum global, alors c'est un minimum local.

Propriété 1.5 : [9]

Si on connaît tous les minima locaux alors le plus petit est le minimum global.

1.2.2.1 Condition nécessaire d'optimalité du premier ordre

Théorème 1.5 : [9]

Soit $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable au point $x \in D$, (où D est un ouvert) si x est un minimum local de (P_1) alors $\nabla f(x) = 0$.

1.2.2.2 Condition nécessaire d'optimalité de second ordre

Théorème 1.6 : [9]

Soit $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ deux fois différentiable au point $x \in D$, si x est un minimum local de (P_1) alors $\nabla f(x) = 0$ et la matrice Hessienne de f au point x , noté par $H(x)$, est semi définie positive.

1.2.2.3 Conditions suffisantes d'optimalité de second ordre

Théorème 1.7 : [5]

Soit $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ deux fois différentiable au point $x \in D$,

1.) Si $\nabla f(x) = 0$ (x est un point stationnaire de f) et $H(x)$ est définie positive, alors x est un minimum local de (P_1)
2.) Si $\nabla f(x) = 0$ et $H(x)$ est définie négative, alors x est un maximum local de (P_1)

Exemple 1.7 :

Soit $f(x, y) = mx^2 + y^2$ définie sur \mathbb{R}^2 dans $\mathbb{R}, m \neq 0$ pour déterminer les extrêmes de f , il faut d'abord déterminer les points stationnaires de f et la matrice Hessienne va déterminer la nature du point stationnaire.

*) Calcul de $\nabla f(x, y) = \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y} \right) = (2mx, 2y)$

$$(2mx, 2y) = (0, 0) \Rightarrow \begin{cases} 2mx = 0 \\ 2y = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0, \\ y = 0 \end{cases}$$

$S(f) = \{(0, 0)\}$. Ce point stationnaire est le seul extrémum éventuel.

a) Nature du point stationnaire $(0, 0)$

$$D^2 f(x, y) = \begin{pmatrix} 2m & 0 \\ 0 & 2 \end{pmatrix}$$

$$H(0, 0) = D^2 f(0, 0) = \begin{pmatrix} 2m & 0 \\ 0 & 2 \end{pmatrix}$$

-) Si $m > 0$, alors $Hf(x_0)$ est défini positive ($\text{vap} > 0$) donc par la condition suffisante $x_0 = (0, 0)$ est minimum local strict.
-) Si $m < 0$, alors les valeurs propre n'ont pas le même signe donc par la condition nécessaire, $x_0 = (0, 0)$ ne peut pas être un minimum local.

1.2.3 Optimisation avec contraintes :

Nous avons traité dans la partie précédente, la résolution d'un problème de minimisation sans contraintes.

Nous considérons l'ensemble $D = \{x \in \mathbb{R}^n \text{ t.q. } g_j(x) \leq 0 \forall j = 1, \dots, m\}$ une partie de \mathbb{R}^n décrite par des contraintes fonctionnelles d'inégalité.

Soit le problème suivant :

$$(P_2) \iff \begin{cases} f(x) \rightarrow \min \\ g_j(x) \leq 0, & J = \{1, \dots, m\} \end{cases}$$

-) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, de classe C^1 .
-) $g_j : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \forall j = 1, \dots, m$.

Exemple 1.8 :

Une firme fabrique des objets A et des objets B en utilisant des matières premières m_1 et m_2 .

pour fabriquer une unité de A , il faut 2 unités de m_1 et 1 unité de m_2 ,

et pour fabriquer une unité de B , il faut 1 unité m_1 et 2 unités de m_2 .

On dispose de 8 unités de m_1 et 7 unités de m_2 .

Enfin le bénéfice est de 4 par unité de A et de 5 par unité de B .

Le patron voudrait optimiser son bénéfice.

Il s'agit donc de maximiser sur S la fonction $4A + 5B$ où a (resp b) représente la quantité de produit A (resp B) fabriquée avec S défini par :

$$S = \{(a, b) \in \mathbb{N}^2, \text{ t.q. } 2a + b \leq 8 \text{ et } a + 2b \leq 7\}$$

Remarque 1.5 :

1. Si $\bar{x} \in \mathbb{R}^n$ est un minimum de f et si $\bar{x} \in D$ alors, \bar{x} est un minimum de (P_2) .

2. Si \bar{x} se trouve sur la frontière de D alors interviendront nécessairement les fonction g_j .

Définition 1.30 :

Soit la fonction

$$L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$$
$$(x, \lambda) \rightarrow L(x, \lambda) = f(x) + \lambda^\top g(x)$$

L est appelé le lagrangien associé au problème (P_2) , λ est appelé multiplicateur de (KKT) associé à x .

Définition 1.31 :

On appelle les $j \in J_0 = \{j/g_j(\bar{x}) = 0\}$ les indices des contraintes actives ou saturées en \bar{x} .

Si $j \notin J_0$ c'est-à-dire $g_j(\bar{x}) < 0$ alors au voisinage de \bar{x} par continuité, on a $g_j(x) < 0$ donc la contrainte $g_j(x) \leq 0$ ne joue pas localement aucun rôle.

Si $j \notin J_0$ c'est-à-dire $g_j(\bar{x}) = 0$ alors il est probable que \bar{x} est à la frontière de D et que l'on sort nécessairement de D en bougeant un petit peu autour de \bar{x} .

Donc seules les fonctions g_j correspondant aux indices actives interviendront dans les conditions nécessaires d'optimalité.

Définition 1.32 : (Contraintes qualifiées)

Nous dirons que, les Contraintes sont qualifiées au point x^* si :

- Les fonctions g_j sont affines.
- Ou bien les vecteurs $\nabla g_j(x^*)$, $j \in J_0$, sont linéairement indépendants où :

$$J_0 = \{j \in J, g_j(x^*) = 0\}$$

Exemple 1.9 :

On cherche à calculer le max de $-(x-2)^2 - (y-1)^2$ sous $D = \{-1 \leq x \leq 1 \text{ et } 0 \leq y \leq \frac{1}{2}\}$ ce qui revient à minimiser le problème (P) :

$$\begin{cases} ((x-2)^2 + (y-1)^2) \rightarrow \min \\ g_1(x, y) = x - 1 \leq 0 \\ g_2(x, y) = -x - 1 \leq 0 \\ g_3(x, y) = y - \frac{1}{2} \leq 0 \\ g_4(x, y) = -y \leq 0 \end{cases}$$

En $A = (1, 0)$, g_1 et g_4 sont saturés, $J_0 = \{j = 1, 4, g_j(A) = 0\}$.

Calculons le gradient du contraintes saturés au point A :

$\nabla g_1(A) = (1, 0)$ et $\nabla g_4(A) = (0, -1)$ forment un système libre donc la qualification de contrainte au point A est vérifiée.

Proposition 1.4 : [9]

Si les contraintes sont qualifiées au point x^* , alors : $y \in \mathbb{R}^n$ est une direction admissible, si et seulement si

$$\nabla g_j(x^*) \cdot y \leq 0, j \in J_0$$

1.2.3.1 Condition de Khun-Tucker .

On a le théorème fondamental de Khun-Tucker (1951) suivant, qui constitue le résultat fondamental de la programmation non linéaire :

Théorème 1.8 : [5] (Khun-Tucker)

Soit le problème (P_2)

-) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, de classe C^1 .
-) $g_j : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \forall j = 1, \dots, m$ de classe C^1 .

On suppose que les contraintes sont qualifiées au point \bar{x} .

Alors il existe des nombres positifs $\lambda_1, \dots, \lambda_m$, (appelés multiplicateurs de Kuhn-Tucker ou de Lagrange généralisés) tel que :

$$\begin{cases} \nabla f(\bar{x}) + \sum_{j=1}^{j=m} \lambda_j g_j(\bar{x}) = 0, \\ \lambda_j g_j(\bar{x}) = 0, \end{cases} \quad \forall j \in J$$

Remarque 1.6 :

Ce résultat n'est pas une équivalence au sens où ce n'est pas une condition nécessaire et suffisante de caractérisation du minimum.

Elle le devient par contre lorsque nous considérons un problème convexe, en effet :

Théorème 1.9 : [5] (Kuhn-Tucker)

soit le problème (P_2) :

-) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, de classe C^1 et convexe.
-) $g_j : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \forall j = 1, \dots, m$ de classe C^1 et convexe.

On suppose que les contraintes sont qualifiées au point \bar{x} , alors \bar{x} est une solution de problème (P_2) si et seulement s'il existe des nombres positifs $\lambda_1, \dots, \lambda_m$ tels que.

$$\begin{cases} \nabla f(\bar{x}) + \sum_{j=1}^{j=m} \lambda_j \nabla g_j(\bar{x}) = 0, \\ \lambda_j g_j(\bar{x}) = 0, \end{cases} \quad \forall j \in J$$

1.2.3.2 La Dualité

$$(P) \Leftrightarrow \begin{cases} f(X) \rightarrow \min, \\ g_j(X) \leq 0, \quad \forall j \in J \\ X \in \mathbb{R}^n \end{cases}$$

Le lagrangien de problème (P) est la fonction $L(X, \lambda) = f(X) + \sum_{j \in J} \lambda_j g_j(X)$

Définition 1.33 :

On définit la fonction duale de Lagrange par $Q(\lambda) = \inf_X L(X, \lambda)$

Remarque 1.7 :

Q est toujours concave.

Définition 1.34 :

Le problème dual (D) associe au problème primal (P)

$$(D) \Leftrightarrow \begin{cases} Q(\lambda) \rightarrow \max \\ \lambda_j \geq 0, j \in J \end{cases}$$

Le saut de dualité est : $v(P) - v(D)$

Définition 1.35 : (Point selle)

1) Point selle : minimal par rapport à une variable, maximal par rapport à l'autre.

2) Le problème primal (P) est équivalent à $\inf_X \sup_{\lambda \geq 0} L(X, \lambda)$

3) Le problème dual (D) est équivalent à $\sup_{\lambda \geq 0} \inf_X L(X, \lambda)$

4) (X^*, λ^*) est un point selle du Lagrangien si pour tout (X, λ) tels que $(\lambda^* \geq 0, \lambda \geq 0)$

$$L(X^*, \lambda) \leq L(X^*, \lambda^*) \leq L(X, \lambda^*)$$

Théorème 1.10 : [5]

(X^*, λ^*) est un point selle avec $\lambda^* \geq 0$ ssi X^* est une solution de (P) , λ^* est une solution de (D) et le saut de dualité est nul.

Remarque 1.8 :

1. Un point selle du lagrangien vérifie les conditions de KKT (sans hypothèse autre que f, g est de classe C^1).
2. Si le problème est convexe : point selle \Leftrightarrow conditions de KKT.

1.2.3.3 Méthode du lagrangien .

1.2.3.3.1 Introduction .

Pour résoudre un problème non linéaire plusieurs méthodes sont utilisées, On s'intéresse à l'une de ces méthodes : **la méthode des multiplicateurs de Lagrange** qui est une méthode d'optimisation, permettant de déterminer l'extrémum local d'une fonction f sous contraintes g .

On pose I :

$$(I) = \begin{cases} f(x_1, x_2, \dots, x_n) \rightarrow \min \\ g(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

-) $f : \mathbb{R}^n \rightarrow \mathbb{R}$.
-) $g : \mathbb{R}^n \rightarrow \mathbb{R}$.

On définit le lagrangien de problème I est la fonction L

$$L(x_1, x_2, \dots, x_n, \lambda) = f(x_1, x_2, \dots, x_n) - \lambda g(x_1, x_2, \dots, x_n)$$

La variable inconnue λ est appelée multiplicateur de Lagrange.

1.2.3.3.2 Principe de la méthode .

Les solutions du système S obtenu en annulant les dérivées partielles de L par rapport à toutes les variables $X = (x_1, x_2, \dots, x_n)$ et λ sont les points stationnaires avec :

$$(S) = \begin{cases} \frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda_i \left(\sum_{i=1}^m \frac{\partial g_i}{\partial x_1} \right) \\ \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda_i \left(\sum_{i=1}^m \frac{\partial g_i}{\partial x_2} \right) \\ \dots \\ \dots \\ \dots \\ \frac{\partial L}{\partial x_n} = \frac{\partial f}{\partial x_n} + \lambda_i \left(\sum_{i=1}^m \frac{\partial g_i}{\partial x_n} \right) \\ g_i(x_1, x_2, \dots, x_n) = 0, \quad i = 1, \dots, m. \end{cases}$$

Pour savoir la nature des points stationnaires en utilisant la matrice Hessienne.

$$H = \nabla_X L(x_1, x_2, \dots, x_n) = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \dots & \dots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \dots & \dots & \frac{\partial^2 L}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial^2 L}{\partial x_n \partial x_1} & \frac{\partial^2 L}{\partial x_n \partial x_2} & \dots & \dots & \frac{\partial^2 L}{\partial x_n^2} \end{pmatrix}$$

Si la matrice A définie positive alors les points stationnaires sont des minimums locaux.

Si la matrice A définie négative alors les points stationnaires sont des maximums locaux.

Si la matrice A est indéfinie alors les points stationnaires sont des points réguliers.

Exemple 1.10 :

Considérons le problème non linéaire suivant

$$\begin{cases} x_1^2 + x_2^2 \rightarrow \min \\ x_1 + x_2 = 1 \\ x_1, x_2 \in \mathbb{R} \end{cases}$$

Le Lagrangien est de forme :

$$L(x_1, x_2, \lambda) = x_1^2 + x_2^2 + \lambda(x_1 - x_2 - 1)$$

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

On résout le système de trois équations et trois inconnus on trouve :

$$x_1 = \frac{1}{2}, x_2 = \frac{1}{2} \text{ et } \lambda = -1$$

Calculons la Hessienne

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Les valeurs propres sont positive donc la hessienne est semi-définie positive.

alors le point $(\frac{1}{2}, \frac{1}{2})$ est un minimum de problème.

1.2.3.4 La méthode de direction admissible .

Nous nous intéressé au problème suivant

$$\begin{cases} f(x) \rightarrow \min \\ g_j(x) \geq 0, j \in J \\ x \in \mathbb{R}^n \end{cases}$$

1.2.3.4.1 Principe de la méthode .

Pour tenter de résoudre le problème (P) , on choisit de départ $x^0 \in X$ et on construit de façon itérative une suite x^k de X vérifiant $f(x^k) < f(x^{k+1})$, jusqu'à ce qu'on "estime" avoir obtenu une approximation satisfaisante.

à partir de x^k , on recherche une direction de descente d qui ne fasse pas sortir "immédiatement" de X . On cherche alors, en se déplaçant dans la direction d , un point x^{k+1} de X est meilleur que x^k (par exemple, $f(x^k + sd)$, pour $s > 0$) en minimisant avec la contrainte que $x^k + sd$ appartienne à X , si on sait résoudre ce nouveau problème, On recommence à partir de x^{k+1} tant qu'un certain critère d'arrêt n'est pas vérifié.

Pour choisir d on peut résoudre le problème linéaire suivant :

$$(F) \iff \begin{cases} \min_d d^\top \nabla f(x^k) \\ d^\top \nabla g_j(x^k) \geq 0, j \in J_0 = \{j \in J / g_j(x^k) = 0\} \\ d^\top d = 1. d \in \mathbb{R}^{|J_0|} \end{cases}$$

On obtient alors pour d la direction admissible de plus grande pente.

1.2.3.4.2 Algorithme de la méthode .

1) Initialisation : Choisir un point admissible x^0 de départ ,choisir $\varepsilon > 0$.

2) Itération : $k = 1, \dots, N$.

3) calculer la direction de descente admissible d^k de problème (F)

4) Déterminer une solution τ_k pour le problème

$$\min_{\tau} \{f(x^k + \tau d^k) / \tau > 0\}.$$

calculons :

$$x_{k+1} = x_k + \tau_k d^k$$

5) Test d'arrêt :

Arrêtez si $\|x_{k+1} - x_k\| < \varepsilon$ aller vers 6), sinon $k := k + 1$, aller au 2).

6) Écrire : " x^{k+1} est une solution optimale " et aller au 7).

7) FIN.

1.3 Optimisation convexe

Définition 1.37 :

Un problème d'optimisation est dit convexe si on veut $\min_{x \in S} f(x)$ avec f convexe et S convexe ou $\max_{x \in S} f(x)$ avec f concave et S convexe.

1.3.1 Optimisation convexe sans contrainte

1.3.1.1 Introduction .

Dans la section précédente, on a vu qu'on n'a pas toujours des CNS mais seulement des CN, CS pour les minima. On va voir que dans le cadre des fonctions convexes, les choses se passent beaucoup mieux. D'ailleurs le point fort des fonctions convexes : $\min local = \min global$.

Soit le problème suivant :

$$\begin{cases} \min f(x) \\ x \in \mathbb{R}^n \end{cases}$$

f convexe et de classe C^1 .

Théorème 1.11 : [6]

Soit S un convexe ouvert et f strictement convexe sur S alors il existe au plus un $\bar{x} \in S$ minimisant f sur S .

1.3.1.2 Condition nécessaire et suffisante .

Théorème 1.12 : [6]

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable sur \mathbb{R}^n et convexe,

$\nabla f(x^*) = 0$ si et seulement si x^* est un minimum global.

On va présenter un algorithme du gradient à pas optimal.

1.3.1.3 Algorithme du gradient à pas optimal .

- 1) Initialisation : fixer $\varepsilon > 0$, choisir $x^0 \in \mathbb{R}^n$
- 2) Itération $k = 0, 1, \dots, n$

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

avec : α_k solution du problème : $\min_{\alpha > 0} f(x^k - \alpha \nabla f(x^k))$

- 3) Si $\|x^{k+1} - x^k\| < \varepsilon$, x^{k+1} est optimal, aller en 4) sinon aller en 5).
- 4) Écrire : « x^k est la solution optimale de (P) », allez vers 6).
- 5) Posons : $k := k + 1$ et allez vers 2).
- 6) FIN.

1.3.1.4 cas quadratique .

Soit le problème quadratique suivant :

$$(\psi) \Leftrightarrow \begin{cases} \frac{1}{2} X^T A X + b^T X \\ X \in \mathbb{R}^n \end{cases}$$

avec A symétrique définie positive.

La solution du problème (ψ) vérifie $\nabla f(x^*) = 0 \Leftrightarrow Ax^* + b = 0$
 $\Rightarrow x^* = -A^{-1}b$

Remarque 1.9 :

En pratique, pour A de dimension très grande on ne calcule pas A^{-1} , mais on utilise une méthode itérative.

L'une des méthodes utilisée est la méthode du gradient à pas optimal.

1.3.1.5 Algorithme du gradient à pas optimal pour le cas quadratique .

- 1) Initialisation : fixer $\varepsilon > 0$, choisir $x^0 \in \mathbb{R}^n$
- 2) Itération $k = 0, 1, \dots, n$

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

avec : $\alpha_k = \frac{\|\nabla f(x^k)\|^2}{(\nabla f(x^k))^T A \nabla f(x^k)}$.

- 3) Si $\|x^{k+1} - x^k\| < \varepsilon$, x^{k+1} est optimal, aller en 4) sinon aller en 5).
- 4) Écrire : « x^k est la solution optimale de (P) », allez vers 6).
- 5) Posons : $k := k + 1$ et allez vers 2).
- 6) FIN.

1.3.2 Optimisation convexe avec contraintes d'égalités

1.3.2.1 Position du problème .

Soit le problème

$$(P) \Leftrightarrow \begin{cases} \min f(x) \\ BX = C, B \in \mathbb{R}^{n \times m} \\ X \in \mathbb{R}^n \end{cases}$$

avec f convexe et les contraintes linéaires.

Comme, le domaine $D = \{X \in \mathbb{R}^n / BX = C\}$ est convexe donc le problème (P) est convexe.

Remarque 1.10 :

1. Pour que le problème (P) soit convexe, il faut que les contraintes d'égalités doivent être linéaires.
2. Si $m < n$ et $\text{rang } B = m$ alors tout les points sont réguliers.
3. Lagrangien du problème (P) est convexe par rapport à X .

1.3.2.2 Condition nécessaire et suffisante .

Théorème 1.13 : [2]

Soient f convexe et $(g_i)_{i=1,\dots,m}$ affines,

$\nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) = 0$, avec $\lambda_i \in \mathbb{R}$ si et seulement si \bar{x} est un minimum.

Proposition 1.5 : [2]

$$x^* \text{ est optimal pour } (P) \Leftrightarrow \nabla f(x^*)(x - x^*) \geq 0 \forall x \in D$$

1.3.2.3 Cas quadratique .

soit le problème quadratique suivant

$$\begin{cases} \frac{1}{2}X^TAX + b^T X \\ BX = C \\ X \in \mathbb{R}^n \end{cases}$$

A est symétrique et semi définie positive, $b \in \mathbb{R}^n$, $B_{m \times n}$, $C \in \mathbb{R}^m$, $\text{rang } B = m$.

En appliquant le théorème de Lagrange $L(x, \lambda) = \frac{1}{2}X^TAX + b^T X + \lambda^T (BX - C)$

$$\nabla L(X, \lambda) = 0 \Leftrightarrow \begin{cases} AX + b + B^T \lambda = 0 \\ BX - C = 0 \end{cases}$$

$$\text{d'où } \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} X \\ \lambda \end{pmatrix} = \begin{pmatrix} -b \\ c \end{pmatrix}$$

Définition 1.38 :

la matrice $\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$ noté K est appelée matrice K.K.T

proposition 1.6 :[2]

Si A est définie positive et B de rang plein c'est-à-dire $\text{rang } B = m$

Alors la matrice $K = \begin{pmatrix} A & B^\top \\ B & 0 \end{pmatrix}$ est inversible.

Remarque 1.11 :

Résoudre le problème (P) revient à résoudre un système linéaire de $(n + m)$ équations à $(n + m)$ inconnus, ça revient à résoudre un problème quadratique sans contraintes

$$\begin{cases} \min \frac{1}{2} Y^\top K Y - \xi^\top Y \\ Y \in \mathbb{R}^{n+m} \end{cases}$$

avec $Y = \begin{pmatrix} X \\ \lambda \end{pmatrix}$, $\xi = \begin{pmatrix} -b \\ c \end{pmatrix}$

1.3.3 Optimisation convexe avec contraintes d'inégalités

1.3.3.1 Position du problème .

Soit le problème

$$(P) \Leftrightarrow \begin{cases} \min f(x) \text{ (} f \text{ convexe)} \\ g_i(x) \leq 0, \ i = 1, \dots, m \text{ (} g \text{ convexe)} \\ X \in \mathbb{R}^n \end{cases}$$

Si $\exists \bar{x}$ tel que $g_i(\bar{x}) < 0$ alors la qualifications des contraintes est vérifier partout.

1.3.3.2 Condition nécessaire et suffisante .

Théorème 1.14 : [2]

Dans le problème (P) , on a les fonctions f et g convexes et supposons qu'ils sont de classe C^1 , La condition de KKT fournit une condition nécessaire et suffisante pour qu'un élément de \mathbb{R}^n soit minimum de (P) .

$$(KKT) \Leftrightarrow \begin{cases} L(X, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) \\ \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) = 0, \ \lambda \geq 0 \end{cases}$$

1.3.3.3 Méthode d'Uzawa .

1.3.3.3.1 Introduction .

La méthode d'Uzawa est basée sur la formulation duale et la recherche des points selles, et comme le problème (P) est convexe donc le point selle équivalent à la condition KKT qui est nécessaire et suffisante pour l'optimalité.

Nous faisons dans cette partie les hypothèses suivantes :

1) f est classe C^1 alors il existe $\alpha > 0$ tel que $\langle f(x) - f(y), x - y \rangle \geq \alpha \|x - y\|^2, \forall x, y \in \mathbb{R}^n$

2) Comme les $g_i, i = 1, \dots, m$ sont convexes et supposons que g_i est de classe C^1 , donc il existe $M > 0$ tel que

$$\|g_i(x) - g_i(y)\| \leq M \|x - y\|, \forall x, y \in \mathbb{R}^n$$

1.3.3.3.2 Algorithme d'Uzawa .

1) Poser $k = 0$, choisir $p^0 \in \mathbb{R}_+^m$ par exemple $(p^0 = 0)$ et on choisit $\varepsilon > 0$ assez petit.

2) Calculer $x^k \in \mathbb{R}^n$ qui minimise sur \mathbb{R}^n la fonction $x \in \mathbb{R}^n \rightarrow f(x) + \sum_{j=1}^m p_j^k g_j(x)$

Si $(k \geq 1)$ alors

Si $\|x^k - x^{k-1}\| \leq \varepsilon$ alors x^k est le point minimum recherché, aller au 5), sinon aller au 3)

3) $p_i^{k+1} = \max \{p_i^k + \rho_k g_i(x^k), 0\}$ tq : $i = 1, \dots, m$, aller au 4),

4) $k := k + 1$, aller au 2).

5) FIN.

Théorème 1.15 :[2] (Convergence de l'algorithme d'Uzawa)

Soient $a_1, a_2 \in \mathbb{R}$ tels que

$$0 < a_1 \leq a_2 < \frac{2\alpha}{M^2}$$

Supposons que $a_1 \leq \rho_k \leq a_2, \forall k \in \mathbb{N}$

Alors la suite x^k générée par l'algorithme d'Uzawa converge vers x^* l'unique solution du problème de minimisation (P) .

1.3.3.4 Cas quadratique .

Soit le problème quadratique suivant :

$$\begin{cases} \frac{1}{2}X^TAX + b^T X \\ BX \leq C \\ X \in \mathbb{R}^n \end{cases}$$

Remarque 1.12 :

- 1) On peut ajouter des variables d'écart et se ramener de cas avec contraintes d'égalité.
- 2) On peut utiliser la dualité lagrangienne pour résoudre ce problème.

1.3.3.5 La méthode de la stratégie des contraintes actives .

Le principe est le suivant :

- 1) On sélectionne un ensemble de contraintes appelée contraintes candidates comme si elles étaient actives et on résout le problème avec ces contraintes égalités

On peut utiliser la matrice de KKT.

Remarque 1.13 :

Si on connaissait les contraintes actives à l'optimum, il suffit de résoudre le problème avec ces contraintes.

- 2) On calcul les multiplicateurs λ_i^* associés aux contraintes si $\lambda_i^* < 0$.

La contrainte associé à λ_i^* sort de l'ensemble de travail.

- 3) On vérifie que les contraintes qui ne sont pas prise en compte pour l'instant sont satisfaites avec cette solution.

Celles qui ne sont pas satisfaites on les faites entrés dans l'ensemble de travail.

- 4) On recommence la procédure jusqu'à ce que tous $\lambda_i^* \geq 0$ et toutes les contraintes sont vérifiées.

Exemple 1.11 : (en utilisant la méthode de la stratégie des contraintes actives)

Soit le problème suivant :

$$\begin{cases} \min |x_1 - x_2|^2 \\ x_1 - 1 \leq 0, x_1, x_2 \in \mathbb{R} \quad (1) \\ 2 - x_2 \leq 0, x_1, x_2 \in \mathbb{R} \quad (2) \end{cases}$$

On commence par la contrainte (1) :

$$\begin{cases} \min (x_1 - x_2)^2 \\ x_1 = 1, x_1, x_2 \in \mathbb{R} \end{cases}$$

$$L(X, \lambda) = (x_1 - x_2)^2 + \lambda(x_1 - 1)$$

En appliquant la CNS

$$\nabla L(x, \lambda) = 0 \Leftrightarrow \begin{cases} 2x_1 - 2x_2 + \lambda = 0 \\ 2x_2 - 2x_1 = 0 \\ x_1 - 1 = 0 \end{cases}$$

on aura $X^* = (1, 1)$, $\lambda^* = 0$

On a $\lambda^* = 0 \geq 0$ donc la contrainte ($x_1 = 1$) reste dans l'ensemble de travail.

La contrainte (2) n'est pas vérifiée.

donc, on va l'inclure dans l'ensemble de travail

On résout le problème suivant :

$$\begin{cases} \min (x_1 - x_2)^2 \\ x_1 = 1 \\ x_2 = 2, x_1, x_2 \in \mathbb{R} \end{cases}$$

les solutions sont $X^* = (1, 2)$, $\lambda_1^* = 2$ $\lambda_2^* = 2$

tous les multiplicateurs sont positifs

toutes les contraintes sont vérifiées

Donc $X = (1, 2)$ est optimal.

2 Les principales méthodes en optimisation semi-infinie

2.1 Introduction

Un grand intérêt attire les chercheurs pour la classe des problèmes d'optimisation **Semi-infinie**, qui sont caractérisés par un nombre fini de variables et un nombre infini de contraintes. Des tels problèmes apparaissent par exemple dans la réduction de pollution atmosphérique, dans la solution des équations intégrale faiblement singulières, dans la distribution de probabilités.etc...

Il est clair que la théorie existe pour des problèmes de programmation semi-infinie, et des algorithmes existent aussi.

On va présenter deux méthodes : la méthode de discrétisation et la méthode d'échange .

2.2 Formulation d'un problème semi-infini

Considérons le problème suivant :

$$\begin{cases} f(x) \rightarrow \min & , \\ g(x, s) \geq 0, & \forall s \in S \\ X \in \mathbb{R}^n \end{cases} \quad (1)$$

-) S : est un compact de \mathbb{R}^p .
-) $f : \mathbb{R}^n \rightarrow \mathbb{R}$.
-) $g : \mathbb{R}^n \times S \rightarrow \mathbb{R}$.

*) Si f et g sont linéaires par rapport à X , on dit qu'on a un problème semi-infini linéaire (*P.S.I.L*).

*) Si f est convexe par rapport à X et/ou g concave par rapport à X , on dit qu'on a un problème semi-infini convexe (*P.S.I.C*).

*) Dans tous les autres cas, on a problème semi-infini non linéaire (*P.S.I.N.L*).

Remarque 2.1 :

Dans ce qui suit on va étudier les problème S.I.N.L qui ont la forme générale suivante :

$$(P) \Leftrightarrow \begin{cases} f(x) \rightarrow \min & , \\ g(x, s) \geq 0, & \forall s \in S. \\ X \in \mathbb{R}^n \end{cases} \quad (2)$$

-) S : Un produit cartésien de N intervalle fermés de \mathbb{R} ($S \subset \mathbb{R}^N$)
-) $g : S \times \mathbb{R}^N \rightarrow \mathbb{R}$, deux fois différentiable par rapport à ces variables.
-) $f : \mathbb{R}^N \rightarrow \mathbb{R}$, deux fois différentiable.

2.2.1 Définitions et propriétés

Définition 2.1 :

On pose :

-) $f_j(x)$: les dérivées partielles de f par rapport à x_j , $j = 1, \dots, n$

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^\top$$

-) $g_j(x)$: les dérivées partielles de g par rapport à x_j , $j = 1, \dots, n$

Définition 2.2 :

Soit $x^* \in \mathbb{R}^n$ une solution réalisable de problème (P)

x^* est un point stationnaire de (P) si :

Il existe des points $s_i^* \in S$ avec $g(x^*, s_i^*) = 0$, $i = 1, \dots, t$ et il existe des scalaires positifs λ_i^* , $i = 1, \dots, n$ tels que :

$$f_j(x) - \sum_{i=1}^m \lambda_i^* g(x^*, s_i^*) = 0, \quad j = 1, \dots, n,$$

Définition 2.3 :

Soit $x^* \in \mathbb{R}^n$ une solution réalisable de problème (P)

Les variables s_i tels que $g(x^*, s_i^*) = 0$, $i = 1, \dots, t$

représentent, les minimums locaux de la fonction $g(x, s)$ au voisinage de x^* .

Définition 2.4 :

soit x dans \mathbb{R}^n , on définit l'ensemble :

$E(x)$: l'ensemble des points minimums de g dans S qui satisfont la condition $g(x, s) \geq -\varepsilon$ avec $\varepsilon > 0$.

$$E(x) = \{s_i \in S \text{ tq : } g(x, s_i) = 0, \text{ et } \exists \varepsilon > 0, \text{ tq : } g(x, s_i) \geq -\varepsilon\} \quad (3)$$

Définition 2.5 :

Pour $x = (x_1, \dots, x_n)$ dans \mathbb{R}^n et $s = (s_1, \dots, s_N) \in E(x) \subset \mathbb{R}^N$, on définit :

1. $\sigma = (\sigma_1, \dots, \sigma_l)$: les composantes sur la frontière de S , $l \leq N$.
2. $\nabla_1 g(x, s) = \left(\frac{\partial g(s, x)}{\partial s_1}, \dots, \frac{\partial g(s, x)}{\partial s_l} \right)$.
3. $\nabla_2 g(x, s) = \left(\frac{\partial g(s, x)}{\partial s_{l+1}}, \dots, \frac{\partial g(s, x)}{\partial s_N} \right)$.
4. $\nabla_2^2 g(x, s)$: la $(N-l, N-l)$ matrice définit comme suit :

$$\nabla_2^2 g(x, s) = \begin{pmatrix} \frac{\partial g(s, x)}{\partial^2 s_{l+1}} & \frac{\partial g(s, x)}{\partial s_{l+1} \partial s_{l+2}} & \dots & \dots & \frac{\partial g(s, x)}{\partial s_{l+1} \partial s_N} \\ \frac{\partial g(s, x)}{\partial s_{l+2} \partial s_{l+1}} & \frac{\partial g(s, x)}{\partial^2 s_{l+2}} & \dots & \dots & \frac{\partial g(s, x)}{\partial s_{l+2} \partial s_N} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial g(s, x)}{\partial s_N \partial s_{l+1}} & \frac{\partial g(s, x)}{\partial s_N \partial s_{l+2}} & \dots & \dots & \frac{\partial g(s, x)}{\partial^2 s_N} \end{pmatrix}$$

Définition 2.6 :

Soit B un sous ensemble ouvert et non vide de \mathbb{R}^n tels que :

- 1) Il existe un nombre fini de points $s \in S$ qui vérifient :

$$\nabla_2 g(x, s) = 0$$

- 2) A chaque point s qui vérifie la condition 1) correspond à un minimal local de g , donc $\nabla_2^2 g$ est définie positive et chaque composantes est non nulle.

Remarque 2.2 :

Soit x dans B , pour chaque minimum local s_i de $g(x, s)$, $g(x, s_i)$ est une fonction différentiable de x définie par :

$$\nabla_2 g(x, s_i) = 0, \quad i = 1, \dots, p \quad (4)$$

p : le nombre de minimum local de $g(x, s)$ avec $x \in B$, donc $p \leq t$.

2.3 Méthode du Lagrangien

2.3.1 Condition nécessaire du premier ordre

On définit le lagrangien :

$$L(x, \lambda) = f(x) + \lambda_i h_i(x) \quad (5)$$

avec :

$$h_i(x) = g(s_i(x), x), \quad i = 1, \dots, p \quad (6)$$

et les fonctions $s_i(x)$, $i = 1, \dots, p$ sont définies par (4)

Si $x^* \in B$, de (4) et (6) on aura :

$$\nabla_x L(x^*, \lambda^*) = 0 \quad (7)$$

2.3.2 Condition du second ordre

Théorème 2.1 : [22]

Soit $x^* \in B$, la solution du problème (2) tel que (7) est vérifié alors :

$$s^\top [\nabla_x^2 L(x^*, \lambda^*)] s \geq 0, \quad \forall s : s^\top \nabla h_i(x^*) = 0, \quad i = 1, \dots, t$$

Théorème 2.2 : [22]

Soit $x \in B$, une solution réalisable qui satisfait (7), si :

$$s^\top [\nabla_x^2 L(x^*, \lambda^*)] s \geq 0, \quad \forall s : s^\top \nabla h_i(x^*) = 0, \quad i = 1, \dots, t$$

alors : x^* est une solution locale de (2)

2.4 Résolution par la discrétisation

La discrétisation est une règle permettant d'utiliser un sous ensemble fini de points à partir d'un ensemble infini. Et résoudre avec les méthodes déjà existantes.

On aura alors à la fin une suite de solutions qui va converger vers la solution optimale.

2.4.1 Problème semi-infini discrétisé

Soit le problème suivant :

$$(P) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s) \geq 0, \quad \forall s \in S \\ X \in \mathbb{R}^n \end{cases}$$

Et soit (P_m) le sous problème suivant :

$$(P_m) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s_i) \geq 0 \quad s_i \in S_m = \{s_1, s_2, \dots, s_m\} \subset S \\ X \in \mathbb{R}^n \end{cases}$$

(P_m) est appelé problème discrétisé de (P)

Posons :

$$D = \{x \in \mathbb{R}^n, g(x, s) \geq 0 \quad s \in S\}$$
$$D_m = \{x \in \mathbb{R}^n, g(x, s_i) \geq 0 \quad s_i \in S_m\}$$

et de fait que $S_m \subset S$, on a bien $D \subset D_m$ et $v(P_m) \leq v(P)$

Définition 2.7 :

(P) est appelé fini réductible s'il existe un sous ensemble (S_m) tel que :

$$v(P_m) = v(P)$$

(P) est dit faiblement discrétisable s'il existe une suite de grilles S_k telle que

$$v(P_k) \rightarrow v(P)$$

2.4.2 Méthode discrétisation

2.4.2.1 Principe de la méthode .

Considérons le problème suivant :

$$(P_m) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s_i) \geq 0 \quad s_i \in S_m = \{s_1, s_2, \dots, s_m\} \subset S \\ X \in \mathbb{R}^n \end{cases}$$

Il s'agit à chaque itération m de résoudre le problème discrétisé (P_m)

Supposons qu'on est à l'itération m , le problème (P_m) peut être résolu par des méthodes classiques d'optimisation.

Soit $X^m = (x_1^m, x_2^m, \dots, x_n^m)$ la solution de (P_m) et une petite valeur $\varepsilon > 0$,

2.4.2.2 Test d'optimalité .

-) Si $g(X^m, s) \geq -\varepsilon, \forall s \in S$

Donc $X^m \in D$

Notons par $X^* = (x_1^*, x_2^*, \dots, x_n^*)$ la solution optimale de (P) on a alors

$$f(X^*) \leq f(X^m) \quad (8)$$

Et d'autre part $D \subset D_m$ et $v(P_m) \leq v(P)$

$$f(X^m) \leq f(X^*) \quad (9)$$

De (8) et (9), (X^m) est une solution optimale de (P)

-) Si $g(X^m, s^m) < -\varepsilon$, alors X^m n'est pas dans D .

On doit chercher une autre solution pour (P) . qui sera réalisable et optimale.

-) Sélectionnez une discrétisation plus fine tel que :

$$\max_{s \in S} \min_{s^{m+1} \in S^{m+1}} \|s - s^{m+1}\| < \max_{s \in S} \min_{s^m \in S^m} \|s - s^m\|$$

Considérons le nouveau problème discrétisé

$$(P_{m+1}) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s_i) \geq 0 \quad s_i \in S_{m+1} \\ X \in \mathbb{R}^n \end{cases},$$

Soit $X^{m+1} = (x_1^{m+1}, x_2^{m+1}, \dots, x_n^{m+1})$ la solution du problème (P_{m+1}) .

On utilise le même test précédent pour vérifier l'optimalité de X^{m+1} pour (P) , et ainsi de suite jusqu'à l'obtention de la solution X^* du problème.

2.4.3 Algorithme de résolution de (P.S.I) par la discrétisation

- 1) initialisation : choisir un ensemble $s_0 \subset S$ et une petite valeur $\varepsilon > 0$.
- 2) itération : $k = 1, \dots, N$.
- 3) calculer la solution x^k de problème (P_k) :

$$(P_k) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s_i) \geq 0 \quad s_i \in S_k = \{s_1, s_2, \dots, s_k\} \subset S \\ X \in \mathbb{R}^n \end{cases},$$

Soit x^k sa solution.

- 4) Si $g(X_k, s) \geq -\varepsilon, s \in S$, aller en 5)
- sinon aller en 6)
- 5) Écrire : X_k est la solution optimale de (P) ; aller vers 7)
- 6) Posons : $k := k + 1$, sélectionnez une discrétisation plus fine tel que :

$$\max_{s \in S} \min_{s^{k+1} \in S^{k+1}} \|s - s^{k+1}\| < \max_{s \in S} \min_{s^k \in S^k} \|s - s^k\|$$

Aller vers 2).

- 7) FIN.

2.5 Méthode d'échange

2.5.1 Introduction

Nous présentons également la méthode d'échange qui est souvent plus efficace que la méthode de discrétisation pure. Ce procédé peut être considéré comme un compromis entre la méthode de discrétisation et l'approche de la réduction continue.

2.5.2 Le principe de la méthode d'échange

Considérons le problème suivant :

$$(P_m) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s_i) \geq 0 \quad s_i \in S_m = \{s_1, s_2, \dots, s_m\} \subset S \\ X \in \mathbb{R}^n \end{cases},$$

Il s'agit à chaque itération m de résoudre le problème discrétisé (P_m)

Supposons qu'on est à l'itération m , le problème (P_m) peut être résolu par des méthodes classiques.

Soit $X_m = (x_1^m, x_2^m, \dots, x_n^m)$ la solution de (P_m) et une petite valeur $\varepsilon > 0$.

2.5.3 Test d'optimalité

calculons :

$$Q(X^m) : \min_{s \in S} g(X^m, S)$$

Soit l'ensemble des solutions de problème $Q(X^m)$ est :

$$E(X^m) = \{s_j, (j \geq 1)\}$$

Supposons que s_1 est un minimum global.

On distingue deux cas :

Cas 1 :

Soit :

$$g(X^m, s_1) \geq -\varepsilon$$

Donc $X^m \in D$

Notons par $X^* = (x_1^*, x_2^*, \dots, x_n^*)$ la solution optimale de (P) on a alors

$$f(X^*) \leq f(X^m) \quad (10)$$

Et d'autre part $D \subset D_m$ et $v(P_m) \leq v(P)$

$$f(X^m) \leq f(X^*) \quad (11)$$

De (10) et (11), (X^m) est une solution optimale de (P) .

Cas 2 :

Si :

$$g(X^m, s_1) < -\varepsilon$$

alors X^m n'est pas dans D .

On doit chercher une autre solution pour (P) . qui sera réalisable et optimale.

Posons $S_{m+1} = S_m \cup E(X^m)$ et passons à l'itération suivant :

Considérons le nouveau problème discrétisé

$$(P_{m+1}) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s_i) \geq 0 \quad s_i \in S_{m+1} \\ X \in \mathbb{R}^n \end{cases},$$

Soit $X^{m+1} = (x_1^{m+1}, x_2^{m+1}, \dots, x_n^{m+1})$ la solution du problème (P_{m+1}) .

On utilise le même test précédent pour vérifier l'optimalité de X^{m+1} pour (P) , et ainsi de suite jusqu'à l'obtention de la solution X^* du problème.

2.5.4 Algorithme de résolution de (P.S.I) par la méthode d'échange

1) initialisation : choisir un ensemble $S_0 \subset S$ et une petite valeur $\varepsilon > 0$.

2) itération : $k = 1, \dots, N$

3) calculer la solution x^k de problème (P_k) :

$$(P_k) \Leftrightarrow \begin{cases} f(x) \rightarrow \min \\ g(x, s_i) \geq 0 \quad s_i \in S_k \\ X \in \mathbb{R}^n \\ S_k = \{s_1, s_2, \dots, s_k\} \subset S \end{cases}$$

Soit X^k sa solution.

4) calculons

$$Q(X^k) : \min_{s \in S} g(X^k, s)$$

5) Supposons $E(X^k)$ sa solution

$$E(X^k) = \{s_j, j = 1, \dots, j(j \geq 1)\}$$

et soit s_1 est le minimum global.

6) si $g(X^k, s_1) \geq -\varepsilon$, aller en 7) sinon aller en 8).

7) Écrire : " X^k est la solution optimale de (P) " ; allez vers 9).

8) Posons : $k := k + 1$, et posons $S_{k+1} = S_k \cup E(X^k)$, allez vers 2).

9) FIN.

Thérème 2.3 : [22]

soit $(X^m)_m$ la suite des solutions des problèmes discrétisés :

$$(p_1), (p_2), \dots, (p_m), (p_{m+1}) \dots$$

avec : $S_1 \subset S_2 \subset \dots \subset S_m \subset S_{m+1} \subset \dots \subset S$ de problème (P) .

la suite $(X^m)_m$ converge vers la solution optimale X^* du problème (P) .

$$\lim_{m \rightarrow \infty} (X^m) = X^*$$

Exemple 2.1 :

Soit le problème semi-infini suivant :

$$(P) \Leftrightarrow \begin{cases} x^2 + y^2 \rightarrow \min \\ -x - y \geq -s, s \in [-1, 0] \\ x, y \in \mathbb{R} \end{cases}$$

Posons que $S_1 = 0$

donc le problème discrétisé est :

$$(P_1) \Leftrightarrow \begin{cases} x^2 + y^2 \rightarrow \min \\ -x - y \geq 0 \end{cases}$$

qui a pour solution $X^1 = (0, 0)$

On a le $\min_{s \in [-1, 0]} g(X^1, s) = \min_{s \in [-1, 0]} (s) = -1 < 0$, donc X^1 n'est pas la solution de (P)

$\min_{s \in [-1, 0]} (s) = -1 < 0$, alors $s_1 = -1$.

Posons $S_2 = S_1 \cup \{-1\} = \{-1, 0\}$, et considérons le problème discrétisé

$$(P_2) \Leftrightarrow \begin{cases} x^2 + y^2 \rightarrow \min \\ -x - y \geq -s, s \in \{-1, 0\} \end{cases}$$

c'est-à-dire

$$(P_2) \Leftrightarrow \begin{cases} x^2 + y^2 \rightarrow \min \\ -x - y \geq 0, s \in \{-1, 0\} \\ -x - y \geq 1 \end{cases}$$

La solution du (P_2) est $X^2 = \left(\frac{-1}{2}, \frac{-1}{2}\right)$

Calculons $\min_{s \in [-1, 0]} g(X^2, s) = \min_{s \in [-1, 0]} [1 + s] = 0 \geq 0$

donc X^2 est la solution optimale de problème (P) .

3 Paramétrisation dual en optimisation semi-infinie quadratique

3.1 Introduction

Nous notons que le procédé de paramétrisation dual est différent des méthodes de discrétisation existantes pour les problèmes SIP.

Les méthodes de discrétisation remplacent généralement les contraintes infinies et avec un nombre fini de contraintes. En résolvant le problème fini d'optimisation, une solution approximative, qui est généralement non réalisable.

La méthode de paramétrisation dual fonctionne avec le problème dual et essaie de trouver une solution approchée. Une fois cela fait, une recherche locale mènera à la solution exacte (jusqu'à la précision de l'ordinateur) de problème de SIP.

Nous présenterons d'abord la méthode de paramétrisation dual pour le problème (P).

La théorie de la méthode de paramétrisation dual au cas de plusieurs contraintes infinies est similaire au cas d'une seule contrainte infinie. Par conséquent, la plupart des résultats seront indiqués sans preuve.

L'efficacité de l'algorithme adaptatif est représenté par la résolution d'un exemple numérique.

3.2 Formulation d'un problème semi-infini quadratique

Considérons le problème semi-infini suivant :

$$P \iff \begin{cases} f(X) = \frac{1}{2}X^\top QX + P^\top X \rightarrow \min \\ A(s)X - b(s) \leq 0 \quad s \in S \\ X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n \\ P = \{p_1, p_2, \dots, p_n\} \in \mathbb{R}^n \end{cases}$$

-) $Q \in \mathbb{R}^{n \times n}$ est matrice symétrique définie positive,
-) S : est compact de \mathbb{R}^m ,
-) $A : S \rightarrow \mathbb{R}^{n \times m}$, différentiable et continu,
-) $b : S \rightarrow \mathbb{R}^m$, différentiable et continu.

Dans ce qui suit on suppose que la qualification de contrainte est vérifiée c'est-à-dire

$$\exists x^0 \in \mathbb{R}^n / A(s)x^0 - b(s) < 0 \quad \forall s \in S.$$

On note $C(S)$ l'espace de Banach des fonctions réelles continues sur S équipé de la norme sup, et par $M(S)$ l'espace dual de $C(S)$.

Soit : $A : \mathbb{R} \rightarrow C(s)$ l'opérateur définie par :

$$(Ax) s = (As) x, \quad \forall s \in S \tag{12}$$

3.3 Problème dual d'un problème semi-infini quadratique

Désignons par A^* le dual de A . Notez que nous avons utilisé le même symbole à la fois pour la fonction de matrice et de l'opérateur correspondant. Toutefois, cela ne devrait pas causer d'une confusion dans le contexte.

D'après (12), le problème (P) peut être représenté comme suit :

$$(P) \iff \begin{cases} f(X) = \frac{1}{2}X^\top QX + P^\top X \rightarrow \min \\ AX - b \leq 0 \\ X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n \\ P = \{p_1, p_2, \dots, p_n\} \in \mathbb{R}^n \end{cases}$$

le problème dual (D) correspondant s'écrit comme suit :

$$(D) \iff \begin{cases} \min_{x, \lambda} \frac{1}{2}X^\top QX + \langle \lambda, b \rangle \\ QX + p + A^*\lambda = 0 \\ \lambda \geq 0 \end{cases}$$

où

$$\langle \lambda, b \rangle = \int_S b(s) d\lambda(s)$$

Lemme 3.1 : [6] (Condition KKT)

Supposons le minimum de problème (P) est obtenu en $x^* \in \mathbb{R}^n$ si et seulement si x^* est faisable et existe un $\lambda^* \in M(S)$ de telle sorte que :

$$Qx^* + p + A^*\lambda = 0$$

$$\langle \lambda^*, Ax^* - b \rangle$$

$$\lambda^* \geq 0$$

Lemme 3.2 : [11]

Soit X un sous-ensemble de \mathbb{R}^n . si $x \in \text{cône } X$, x peut être représenté comme une combinaison linéaire non négative des points de X , c'est-à-dire, il existe n nombres $\lambda_i \geq 0$ telle que $x = \sum_{i=1}^n \lambda_i x_i$

Lemme 3.3 : [6]

Supposons que le minimum de problème de (P) est obtenu en $x^* \in \mathbb{R}^n$. alors λ^* est un multiplicateur qui satisfait la condition KKT si et seulement si (x^*, λ^*) est une solution du problème (D) .

Théorème 3.1 :

supposons que le minimum de problème de (P) est obtenu en $x^* \in \mathbb{R}^n$. alors, la solution du problème dual (D) contient une paire de solution (x^*, λ^*) tel que le multiplicateur de KKT a un support fini de pas plus de n points.

Preuve :

Compte tenu du lemme 3.2, il suffit de prouver qu'il existe un support fini mesurer λ^* , avec pas plus de n points d'appui, tels que (x^*, λ^*) satisfait à la Condition KKT . Comme x^* est une solution primale du théorème KKT, il existe une mesure $\bar{\lambda}$ de telle sorte que $(x^*, \bar{\lambda})$ satisfait les conditions KKT. Notons l'ensemble actif de problème (P) en x^* par $S(x^*)$

$$S(x^*) = \{s \in S : A(s)x^* - b(s) = 0\}$$

De la continuité de $A(s)$ et $b(s)$ et la compacité de S , on voit que $S(x^*)$ soit compact ou vide. Il résulte de la Condition KKT que le soutien de $\bar{\lambda}$ satisfait $SUPP(\bar{\lambda}) \in S(x^*)$ Ainsi,

$$Qx^* + p + \int_{S(x^*)} (A(s))^\top d\bar{\lambda} = 0$$

D'où $-(Qx^* + p) = \int_{S(x^*)} (A(s))^\top d\bar{\lambda} \in \text{cône} \{a^i(s) : s \in S(x^*), i = 1, \dots, m\}$.

Où $a^i(s)$ est la $i^{\text{ème}}$ colonne de $(A(s))^\top$. Du lemme 3.2, il existe un entier $k \leq n$, k vecteurs $a^{ij}(s_j^*)$, $j = 1, \dots, k$ et k constantes

$$\alpha_j^* \geq 0, j = 1, \dots, k \text{ tel que : } -(Qx^* + p) = \sum_{j=1}^k \alpha_j^* a^{ij}(s_j^*)$$

où $1 \leq i \leq m$ et $s_j^* \in S(x^*)$, $j = 1, \dots, k$. Soit λ_j^* la dimension m vecteur dont les éléments sont toutes nulles à l'exception de son $i^{\text{ème}}$ élément, qui est un α_j^* . Définir un m -dimensionnelle mesure λ^* , avec un support fini $\{s_j^* : j = 1, \dots, k\}$ par $\lambda(s_j^*) = \lambda_j^*$. La mesure λ^* ainsi définie à un support fini de $k \leq n$ de points d'appui et satisfait λ^* , $j = 1, \dots, k$.

$$\begin{aligned}
& Qx^* + p + \int_{S(x^*)} (A(s))^\top d\lambda^*(s) \\
&= (Qx^* + p) + \sum_{j=1}^k (A(s_j^*))^\top \lambda_j^* \\
&= (Qx^* + p) + \sum_{j=1}^k \alpha_j^* a^{ij}(s_j^*) = 0
\end{aligned}$$

et, puisque $s_j^* \in S(x^*)$, $\langle \lambda^*, Ax^* - b \rangle = \sum_{j=1}^k A(s_j^*) x^* - b(s_j^*) \lambda_j^* = 0$.
Ainsi (x^*, λ^*) satisfait les conditions KKT .

3.4 Le dual paramétré d'un problème de programmation semi-infinie quadratique

Considérons le problème suivant :

$$P \iff \begin{cases} f(X) = \frac{1}{2}X^\top QX + P^\top X \rightarrow \min \\ A(s)X - B(s) \geq 0 \quad s \in S \\ X \in \mathbb{R}^n \\ P \in \mathbb{R}^n \end{cases}$$

Soit $\tau = \{s_1, s_2, \dots, s_k\}$ un sous ensemble de S , et $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ est un multiplicateur de KKT.

Donc, d'après le théorème (3,1), le dual paramétré correspondant peut s'écrire comme suit :

$$(D_k) \iff \begin{cases} \min_{x, \lambda, \tau} \frac{1}{2}X^\top QX + \sum_{j=1}^k b(s_j)^\top \lambda_j \\ QX + p + \sum_{j=1}^k A(s_j)^\top \lambda_j = 0 \quad s_j \in S \\ \lambda_j \geq 0 \quad j = 1, \dots, k \end{cases}$$

$$\text{avec } \sum_{j=1}^k A(s_j)^\top \lambda_j = \sum_{j=1}^k \lambda_j \alpha^{ij}(s_j)$$

sachant que i représente la $i^{\text{ème}}$ colonne de $A(s)^\top$, et j représente la $j^{\text{ème}}$ ligne de $A(s)^\top$.

donc l'écriture suivante est équivalente à (D_k) :

$$(D_k) \iff \begin{cases} \min_{x, \lambda, \tau} \frac{1}{2}X^\top QX + \sum_{j=1}^k b(s_j)^\top \lambda_j \\ QX + p + \sum_{j=1}^k \lambda_j \alpha^{ij}(s_j) = 0 \quad s_j \in S \\ \lambda_j \geq 0 \quad j = 1, \dots, k \end{cases}$$

Une solution du problème (D_k) sera notée par : $\{\tau, \lambda, x\}$

tels que :

-) $\tau = \{s_1, s_2, \dots, s_k\}$,
-) $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$,
-) $X = \{x_1, x_2, \dots, x_n\}$.

Lemme 3.4 : [11]

Pour le problème (D_k) , ce qui suit est vraie.

-) $v(D_{k^*}) = v(D_k)$ pour tout $k \geq k^*$

et si $k \geq 1$ $v(D_{k^*-1}) > v(D_{k^*})$.

-) k^* est l'entier minimum tel que pour $k \geq k^*$, la solution du problème (D_k) fournit la solution de problème (P) dans le sens où si (x^*, λ^*, τ^*) est une solution globale de problème (D_k) , alors x^* est une solution de problème (P) .
-) Le nombre k^* satisfait $0 \leq k^* \leq n$
-) Si $0 < k < k^*$,

alors : $v(D_k) > v(D_{k+1})$.

Remarque 3.1 :

l'identification de k n'est pas facile avant de connaître la solution du problème (D) .

3.5 Principe de la méthode

Dans la section précédente, nous avons réduit le problème (D) à un problème fini.

Par conséquent, afin de résoudre le problème (P) , il suffit de résoudre le problème (D_k) . Cependant, le problème (D_k) est non linéaire et le défi est de trouver la solution globale. un algorithme est proposé pour calculer une solution globale du problème (D_k) .

Considérons le problème suivant :

$$(D_k(\tau)) \iff \begin{cases} f(X) = \frac{1}{2}X^T QX + \sum_{j=1}^k b(s_j)^T \lambda_j \rightarrow \min_{x, \lambda} \\ QX + p + \sum_{j=1}^k A(s_j)^T \lambda_j = 0 \quad s_j \in S \\ \lambda_j \geq 0 \quad j = 1, \dots, k \end{cases}$$

À la solution optimale, augmenter k et affiner le sous ensemble $\tau = \{s_1, s_2, \dots, s_k\}$ de tel façon que $\max_{s \in S} \min_{1 \leq j \leq k} \|s_j - s\| \rightarrow 0$ quand $k \rightarrow \infty$.

Le nouveau problème a résoudre $(D_k(\tau))$. répétez cette procédure jusqu'à satisfaire le critère d'arrêt. A la fin, un vecteur de points d'indices approximatives τ^* avec la solution optimale (x^*, λ^*) correspondante au problème $(D_k(\tau))$ qui est obtenu.

Le triplet (x^*, λ^*, τ^*) donne une solution approchée du problème (D_k) . Si cette solution approchée est assez bon, il est prévu que (x^*, λ^*, τ^*) se trouve dans le bassin d'une solution globale de problème (D_k) .

Lemme 3.5 : [11]

Soit $\tau = \{s_1, s_2, \dots, s_k\}$ un sous ensemble S , alors le problème $(D_k(\tau))$ est le dual de $(P_k(\tau))$ tel que :

$$(P_k(\tau)) \iff \begin{cases} f(X) = \frac{1}{2}X^T QX + P^T X \rightarrow \min \\ A(s_i) X - B(s_i) \geq 0 \\ i = 1, 2, \dots, k \end{cases}$$

Soit x est la solution du problème $(P_k(\tau))$ si et seulement s'il existe un $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ telle que $\{\lambda, x\}$ est une solution du problème $(D_k(\tau))$.

Les valeurs optimales des problèmes $(P_k(\tau))$ et $(D_k(\tau))$ satisfaisante :

$(P_k(\tau)) = (D_k(\tau))$, c'est-à-dire le saut de dualité est nul.

3.6 Algorithme de la méthode

1) Choisir un x^0 , un petit nombre $\varepsilon > 0$, un entier N , et un ensemble de paramétrage

$S_i = \{s_j^i : j = 1, 2, \dots, k_i\}$ Vérifiant :

$$\max_{s \in S} \min_{1 \leq i \leq k_i} \|s - s_j^i\| \rightarrow 0.$$

2) Soit $E^0 = \emptyset$ et $i = 0$.

3) $i = i + 1$, $V^i = \{s \in S_i \text{ tq } A(s)x^{i-1} - b(s) \geq 0\} \cup E^{i-1}$.

Supposons V^i à l_i éléments $V^i = \{v_1, v_2, \dots, v_{l_i}\}$ et réglez $\tau^i = \{v_1, v_2, \dots, v_{l_i}\}$

4) Calculer la solution de problème $(D_{l_i}(\tau^i))$:

$$(D_{l_i}(\tau^i)) \iff \begin{cases} f(X) = \frac{1}{2}X^\top QX + \sum_{j=1}^{l_i} b(s_j)^\top \lambda_j \rightarrow \min \\ QX + p + \sum_{j=1}^{l_i} A(s_j)^\top \lambda_j = 0 \quad s_j \in \tau^i \\ \lambda_j \geq 0 \quad j = 1, \dots, l_i \end{cases}$$

soit x_i sa solution.

5) Si $i \leq N$ ou $f(x^i) - f(x^{i-1}) \geq \varepsilon$,

réglez $E^i = \{s \in V_i \text{ tq } A(s)x^i - b(s) = 0\}$, Aller vers 3).

6) Sinon la solution optimale de problème (P) est x^i et aller vers 7).

7) FIN.

3.7 La convergence de l'algorithme

Théorème 3.2 :[11]

Si $\max_{s \in S} \min_{1 \leq i \leq k} \|s - s_j^i\| \rightarrow 0$ quand $k \rightarrow \infty$, est satisfaite, alors la séquence $\{x^i\}$ obtenue à partir de l'algorithme précédent converge vers la solution du problème (P) . Par conséquent, si ε et N sont convenablement choisis, le x^* obtenu dans l'étape (6) de l'algorithme est la solution optimale du problème (P) .

Exemple 3.1 :

Soit le problème suivant :

$$\begin{cases} \min x^\top Qx \\ A(s)x - b(s) \leq 0 \forall s \in [0, 5] \\ Q \in R^{16 \times 16}, x \in R^{16} \end{cases}$$

où

$$Q = \begin{bmatrix} 4 & 1 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \cdots & 0 \\ & & & \vdots & \\ 0 & 0 & 0 & \cdots & 4 \end{bmatrix}$$

$$A(s) = \left(-e^{-(-s - (\frac{5}{16}))^2}, -e^{-(-s - 2(\frac{5}{16}))^2}, \dots, -e^{-(-s - 16(\frac{5}{16}))^2} \right) \in R^{16}$$

$$b(s) = -3 - 4.5 \sin(4.7\pi(s - 1.23)/8) \in R$$

Afin de résoudre ce problème en utilisant l'algorithme adaptatif développé dans la section précédente, nous avons choisi la séquence de paramétrage $\{S_i\}$ comme suit : $S_1 = \{0, 2.5, 5\}$ pour $i \geq 2$, S_i consiste des milieux de toutes les paires de points adjacents dans $\bigcup_{j=1}^{i-1} S_j$. Les résultats des calculs sont présentés dans le tableau 1.

Nous notons que l'actif points d'indice de contrainte à la solution est identifié comme $s_1^* = 2,06165$ et $s_2^* = 5.00000$. La solution optimale coïncide également avec les calculs précédents jusqu'à sept chiffres décimaux. Nous avons testé différentes valeurs pour x^0 . Tous les résultats sont similaires à celles présentées dans Tableau 1.

<i>L'approximation de X</i>	[0.084044	0.142695	0.258183	0.436827	
	0.707666	1.046537	1.345125	1.453519	
	1.303190	0.987357	0.708656	0.664137	
	0.914220	1.451353	1.763188	2.606933]	24.350055
<i>L'approximation de λ</i>	0.000000	1.975311	16.463732	0.000000	
<i>x optimal</i>	[0.076461	0.204917	0.486116	0.927634	
	1.437805	1.808159	1.845673	1.530850	
	1.039785	0.610216	0.404887	0.495781	
	0.855648	1.458993	1.798618	2.668698]	
Objectif optimal	154.116154				

Tableau 1. Résultats pour l'exemple 1.

4 Conclusion

Nous nous sommes intéressé dans notre modeste travail à une classe particulière des problèmes d'optimisation nommés problèmes d'optimisation semi-infinie, à savoir :

Les problèmes semi-infinie non linéaires et quelques techniques de résolution et pour cela on a présenté un problème semi-infini non linéaire et la méthode de discrétisation,

en suite, la méthode d'échange enfin on a présenté le problème semi-infini quadratique qui fait appel à l'algorithme de Paramétrisation du dual.

Et comme perspectives futures, Pour la résolution des problèmes semi-infinis non linéaires, on peut utiliser l'algorithme de **Branch and Bound (Séparation et évaluation)**, et l'algorithme de **gradient projeté**.

Références

- [1] J. Amaya, J. A. Gómez, Strong duality for inexact linear programming problems, *Optimisation* 49 (2001) 243 - 269
- [2] R. Horst, *Deterministic Optimization, with partitions sets whose feasibility is not known. Application to concave Minimization, Reverse convex constraints, D.C-programming and Lipschitzization Theory and Applications*, (1988)
- [3] A. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards, Section B, Mathematical Sciences*, 49 : 263-265 ; 1952.
- [4] G.B. Dantzig. *Linear Programming and extensions*. Princeton University Press, Princeton, 1963
- [5] E.J. Anderson, P. Nas, *Linear Programming in Infinite Dimensional Spaces*, Wiley, New York, 1987.
- [6] Stephen Boyd and Vandenberg. *Convex Optimization* (2006).
- [7] Irène Charon, Olivier Hudry, *Optimisation non linéaire*.
- [8] J.F. Bonnans, J. CH. Gilbert, C. Lemarichal, et C. Sagastizabal. *Numerical Optimisation : Theoretical and numerical aspects*. Série Universitext, Springer-Verlag, Berlin, 2003.
- [9] Marco LOPEZ, George STILL *Semi-Infinite Programming*, 2005.
- [10] P. Ciarlet, B. Miara, J.M. Thomas, *Exercices l'analyse numérique matricielle et d'optimisation*, Masson, Paris, 1987.
- [11] Y. LIU et K.L. TEO (*Journal of Global Optimization* global 24 : 205-217, 2002) *An Adaptive Dual Parametrization Algorithm for Quadratic Semi-infinite Programming Problems*.
- [12] Goberna and Lopez. *Conditions for closedness of the characteristic cone associated with an infinite linear system*. Depart. de maths Faculté des sciences Univ. Alicante Espagne. 1982.
- [13] Gustafson and Kortanek. *Semi-infinite programming and application*. Royal institut of technology Stockholm Sweden. 1984.
- [14] Hettich, Kortanek « *semi infinite programming* » *Theory, methods and applications*. SIAM rev.35.3,(1993) 380-429.
- [15] Gustafson and Kortanek. *numerical solution of a class of semi-infinite programming problems*. *Naval Res. logist* 20(1973) 477-504.
- [16] A. Charnes, W.W. Cooper, K.O.Kortanek, *Duality, Haar programs and finite sequence spaces*, *Proceedings of the National Academy of Science* 48 (1962) 783-786.

- [17] A. Charnes, W.W. Cooper, K.O.Kortanek, Duality in semi-infinite programs and some works of Haar and Carath'eodory, *Management Sciences* 9 (1963) 209-228.
- [18] A. Charnes, W.W. Cooper, K.O. Kortanek, On the Theory of Semi-Infinite Programming and a Generalization of the Kuhn-Tucker Saddle Point Theorem for Arbitrary Convex Functions, *Naval Research Logistics Quarterly* 16 (1969) 41-51.
- [19] A. Haar, lineare Ungleichungen, *Acta Math. Szeged* 2 (1924) 1-14.
- [20] R. Hettich, P. Zencke, *Numerische Methoden der Approximation und der semi-infiniten Optimierung*, Teubner, Stuttgart, 1982.
- [21] F. John, Extremum problems with inequalities as subsidiary conditions, *Studies and Essays : Courant Anniversary Volume*, Interscience Publisher (Wiley), New York, 1948, 187-204.
- [22] Polak. preserving discretization strategies for semi infinite programming and optimal control (1989).
- [23] Glashoff and Gustafson; *linear optimization and approximation*; *Amplitude Math.Sciences*, Springer-Verlag 1983.
- [24] R. Hettich (Ed.), *Semi-infinite programming*, *Lecture Notes in Control and Information Sciences*, Springer, Berlin, 1979.
- [25] R. Tichatschke, *Lineare Semi-Infinite Optimisation dans l'Approximations theorie*, *Karl-Marx-Stadt : Wissenschaftliche Schriftenreihe der Technischen Hochschule*, 1981.
- [26] A.V. Fiacco, K.O. Kortanek (Eds.), *Semi-Infinite Programming and Applications*, *Lecture Notes in Economics and Mathematical Systems* 215, 1983.
- [27] B. Brosowski, *Parametric semi-infinite optimization*, *Verlag Peter Lang*, Frankfurt, 1982.