



HAL
open science

Contribution à l'étude des algorithmes de l'optimisation non convexe et non différentiable

Rachid Benacer

► **To cite this version:**

Rachid Benacer. Contribution à l'étude des algorithmes de l'optimisation non convexe et non différentiable. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1986. Français. NNT : . tel-00320986

HAL Id: tel-00320986

<https://theses.hal.science/tel-00320986v1>

Submitted on 12 Sep 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

Présentée à

**L'UNIVERSITE SCIENTIFIQUE, TECHNOLOGIQUE ET
MEDICALE DE GRENOBLE**

pour obtenir le grade de
docteur de l'Université Scientifique,
Technologique et Médicale de Grenoble
"mathématiques appliquées"

par

BENACER Rachid

**CONTRIBUTION A L'ETUDE DES
ALGORITHMES DE L'OPTIMISATION
NON CONVEXE ET NON DIFFERENTIABLE**

Thèse soutenue le 2 juillet 1986 devant la commission d'examen.

P.J. LAURENT

Président

J.P. CROUZEIX

J.B. HIRIART-URRUTY

D.T. PHAM

F. ROBERT

Examineurs



UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Année universitaire 1982-1983

Président de l'Université : M. TANCHE

MEMBRES DU CORPS ENSEIGNANT DE L'U.S.M.G.

(RANG A)

SAUF ENSEIGNANTS EN MEDECINE ET PHARMACIE

PROFESSEURS DE 1ère CLASSE

ARNAUD Paul	Chimie organique
ARVIEU Robert	Physique nucléaire I.S.N.
AUBERT Guy	Physique C.N.R.S.
AYANT Yves	Physique approfondie
BARBIER Marie-Jeanne	Electrochimie
BARBIER Jean-Claude	Physique expérimentale C.N.R.S. (labo de magnétisme)
BARJON Robert	Physique nucléaire I.S.N.
BARNOUD Fernand	Biosynthèse de la cellulose-Biologie
BARRA Jean-René	Statistiques - Mathématiques appliquées
BELORISKY Elie	Physique
BENZAKEN Claude (M.)	Mathématiques pures
BERNARD Alain	Mathématiques pures
BERTRANDIAS Françoise	Mathématiques pures
BERTRANDIAS Jean-Paul	Mathématiques pures
BILLET Jean	Géographie
BONNIER Jean-Marie	Chimie générale
BOUCHEZ Robert	Physique nucléaire I.S.N.
BRAVARD Yves	Géographie
CARLIER Georges	Biologie végétale
CAUQUIS Georges	Chimie organique
CHIBON Pierre	Biologie animale
COLIN DE VERDIERE Yves	Mathématiques pures
CRABBE Pierre (détaché)	C.E.R.M.O.
CYROT Michel	Physique du solide
DAUMAS Max	Géographie
DEBELMAS Jacques	Géologie générale
DEGRANGE Charles	Zoologie
DELOBEL Claude (M.)	M.I.A.G. Mathématiques appliquées
DEPORTES Charles	Chimie minérale
DESRE Pierre	Electrochimie
DOLIQUE Jean-Michel	Physique des plasmas
DUCROS Pierre	Cristallographie
FONTAINE Jean-Marc	Mathématiques pures
GAGNAIRE Didier	Chimie physique

.../...

GASTINEL Noël	Analyse numérique - Mathématiques appliquées
GERBER Robert	Mathématiques pures
GERMAIN Jean-Pierre	Mécanique
GIRAUD Pierre	Géologie
IDELMAN Simon	Physiologie animale
JANIN Bernard	Géographie
JOLY Jean-René	Mathématiques pures
JULLIEN Pierre	Mathématiques appliquées
KAHANE André (détaché DAFCO)	Physique
KAHANE Josette	Physique
KOSZUL Jean-Louis	Mathématiques pures
KRAKOWIAK Sacha	Mathématiques appliquées
KUPTA Yvon	Mathématiques pures
LACAZE Albert	Thermodynamique
LAJZEROWICZ Jeannine	Physique
LAJZEROWICZ Joseph	Physique
LAURENT Pierre	Mathématiques appliquées
DE LEIRIS Joël	Biologie
LLIBOUTRY Louis	Géophysique
LOISEAUX Jean-Marie	Sciences nucléaires I.S.N.
LOUP Jean	Géographie
MACHE Régis	Physiologie végétale
MAYNARD Roger	Physique du solide
MICHEL Robert	Minéralogie et pétrographie (géologie)
MOZIERES Philippe	Spectrométrie - Physique
OMONT Alain	Astrophysique
OZENDA Paul	Botanique (biologie végétale)
PAYAN Jean-Jacques (détaché)	Mathématiques pures
PEBAY PEYROULA Jean-Claude	Physique
PERRIAUX Jacques	Géologie
PERRIER Guy	Géophysique
PIERRARD Jean-Marie	Mécanique
RASSAT André	Chimie systématique
RENARD Michel	Thermodynamique
RICHARD Lucien	Biologie végétale
RINAUDO Marguerite	Chimie CERMAV
SENGEL Philippe	Biologie animale
SERGERAERT Francis	Mathématiques pures
SOUTIF Michel	Physique
VAILLANT François	Zoologie
VALENTIN Jacques	Physique nucléaire I.S.N.
VAN CUTSEN Bernard	Mathématiques appliquées
VAUQUOIS Bernard	Mathématiques appliquées
VIALON Pierre	Géologie

PROFESSEURS DE 2^{ème} CLASSE

ADIBA Michel	Mathématiques pures
ARMAND Gilbert	Géographie

AURIAULT Jean-Louis	Mécanique
BEGUIN Claude (M.)	Chimie organique
BOEHLER Jean-Paul	Mécanique
BOITET Christian	Mathématiques appliquées
BORNAREL Jean	Physique
BRUN Gilbert	Biologie
CASTAING Bernard	Physique
CHARDON Michel	Géographie
COHENADDAD Jean-Pierre	Physique
DENEUVILLE Alain	Physique
DEPASSEL Roger	Mécanique des fluides
DOUCE Roland	Physiologie végétale
DUFRESNOY Alain	Mathématiques pures
GASPARD François	Physique
GAUTRON René	Chimie
GIDON Maurice	Géologie
GIGNOUX Claude (M.)	Sciences nucléaires I.S.N.
GUITTON Jacques	Chimie
HACQUES Gérard	Mathématiques appliquées
HERBIN Jacky	Géographie
HICTER Pierre	Chimie
JOSELEAU Jean-Paul	Biochimie
KERCKOVE Claude (M.)	Géologie
LE BRETON Alain	Mathématiques appliquées
LONGEQUEUE Nicole	Sciences nucléaires I.S.N.
LUCAS Robert	Physiques
LUNA Domingo	Mathématiques pures
MASCLE Georges	Géologie
NEMOZ Alain	Thermodynamique (CNRS - CRTBT)
OUDET Bruno	Mathématiques appliquées
PELMONT Jean	Biochimie
PERRIN Claude (M.)	Sciences nucléaires I.S.N.
PFISTER Jean-Claude (détaché)	Physique du solide
PIBOULE Michel	Géologie
PIERRE Jean-Louis	Chimie organique
RAYNAUD Hervé	Mathématiques appliquées
ROBERT Gilles	Mathématiques pures
ROBERT Jean-Bernard	Chimie physique
ROSSI André	Physiologie végétale
SAKAROVITCH Michel	Mathématiques appliquées
SARROT REYNAUD Jean	Géologie
SAXOD Raymond	Biologie animale
SOUTIF Jeanne	Physique
SCHOOL Pierre-Claude	Mathématiques appliquées
STUTZ Pierre	Mécanique
SUBRA Robert	Chimie
VIDAL Michel	Chimie organique
VIVIAN Robert	Géographie



Je dédie ce travail

**à la mémoire de mon père,
à ma mère,
à mes frères et ma soeur.**



Cette thèse a été préparée au laboratoire TIM3 de l'Institut d'Informatique et de Mathématiques Appliquées de Grenoble.

Je tiens à exprimer ma profonde reconnaissance à Monsieur PHAM Dinh-Tao pour m'avoir proposé ce sujet . Ses encouragements constants et ses conseils m'ont beaucoup facilité la réalisation de ce travail.

Je suis très sensible à l'honneur que me fait Monsieur P.J.LAURENT en acceptant de présider ce jury. Qu'il veuille bien agréer l'expression de mes sincères et respectueuses considérations.

Messieurs J.P.CROUZEIX, professeur à l'Université de Clermont-Ferrand et J.-B.HIRIART-URRUTY, professeur à l'Université de Toulouse ont bien voulu juger ce travail et venir à Grenoble faire partie du jury. Qu'ils veuillent bien accepter mes sincères et respectueux remerciements.

Que Monsieur F.ROBERT reçoive mes sincères remerciements pour ses enseignements et pour son amabilité d'avoir bien voulu participer au jury.

Je remercie également tous les membres de l'Equipe Algorithmique Mathématique pour leur contribution que chacun d'eux a pu apporter et l'équipe du service de reprographie pour l'excellente qualité de son travail.

Enfin je voudrais exprimer ma gratitude à tous mes collègues et mes amis qui ont rendu mon séjour en France très agréable.



TABLES DES MATIERES

INTRODUCTION GENERALE	1
CHAPITRE I Méthodes de sous-gradients	7
I.1 Introduction.....	9
I.2 Maximisation d'une fonction convexe sur une partie compacte convexe.....	10
I.3 Minimisation de la différence de deux fonctions convexes sur une partie compacte convexe	22
I.4 Minimisation d'une fonction convexe sur une partie compacte convexe et une contrainte anticonvexe.....	27
CHAPITRE II Minimisation globale d'une fonction concave sur une partie compacte convexe	33
II.1 Introduction	35
II.2 Quelques méthodes de résolution connues dans la littérature.....	36
II.3 Méthodes de coupe pour le cas linéaire.....	38
II.4 Méthodes de résolution par partition de l'ensemble admissible	45
II.5 Méthodes de résolution par approximation extérieure de l'ensemble admissible.....	56
II.6 Essais numériques.....	59

CHAPITRE III	Minimisation globale d'une fonction linéaire sur une partie compacte convexe et une contrainte anticonvexe	65
III.1	Introduction.....	67
III.2	Réduction de quelques programmes mathématiques non convexes et applications.....	69
III.3	Minimisation d'une fonction linéaire sur un polyèdre convexe borné et une contrainte anticonvexe.....	76
III.4	Algorithmes de résolution dans le cas général.....	89
III.5	Essais numériques.....	92
CHAPITRE IV	Applications	97
IV.1	Introduction.....	99
IV.2	Maximisation globale d'une fonction quadratique non convexe	100
IV.3	Maximisation globale de la somme en valeur absolue des fonctions affines sur un polyèdre borné	104
IV.4	Programmation fractionnaire	106
IV.5	Essais numériques	110
CONCLUSION		113
REFERENCES		117

INTRODUCTION GENERALE



Depuis le début des années soixante , les problèmes d'optimisation non convexe ont pris une attention croissante qui reflète leur intérêt dans plusieurs domaines d'application. Un effort considérable se développe actuellement pour la classification des problèmes d'optimisation non convexe et les méthodes numériques de résolution, car on ne peut évidemment pas traiter tous les problèmes d'optimisation non convexe.

Il en résulte qu'une classe très large et très importante par ses applications , est composée de ces trois types de problèmes:

1-min $\{f(x) / x \in C\}$ (P1)
où f est concave sur \mathbb{R}^n et C une partie convexe compacte de \mathbb{R}^n .

2-min $\{f(x) - g(x) / x \in C\}$ (P2)
où f, g sont convexes sur \mathbb{R}^n et C une partie compacte convexe de \mathbb{R}^n .

3-min $\{f(x) / x \in C, g(x) \geq 0\}$ (P3)
où f, g et C vérifient les mêmes hypothèses que dans 2- .

Ces trois problèmes sont équivalents en ce sens que la résolution de l'un entraîne celle des deux autres [53,54]. Mais les propriétés particulières de chaque problème justifient l'étude séparément que nous allons faire sur ces trois problèmes.

Les sujets traités dans cette thèse font partie de deux thèmes de la recherche actuelle qui sont différents mais complémentaires. Le premier s'articule autour des conditions nécessaires et suffisantes d'optimalité (du 1^{er} et 2^{ème} ordre), la dualité en optimisation non convexe , les méthodes de sous-gradients basées sur cette étude qualitative. L'étude qualitative de cette classe de problèmes d'optimisation non convexe fait surtout l'objet des travaux de **B.N.Pschenichny [26]**, **J.F.Toland [48,49]**, **I.Singer [43-45]** et **I.Ekeland [11]**, **J.-B.Hiriart Urruty [17]**, **Pham Dinh Tao [31]**. Le deuxième thème correspond aux algorithmes qui permettent d'obtenir les solutions globales de cette classe de problèmes. Ces algorithmes, qui sont introduits par **H.Tuy [47,51-55]** et dont diverses variantes sont étudiées par **R.Horst [18]**, **K.K.Hoffman [13]**, **J.B.Rosen[36]**, ne font pas intervenir (en général) les caractérisations des solutions optimales dans leur description. Ils se situent dans le cadre de l'optimisation combinatoire. Pour la bibliographie

sur ces deux thèmes, le lecteur pourra consulter [1,12,17,47,55]. Nous allons passer en revue les principaux travaux effectués dans chacun des chapitres sur ces deux thèmes de recherche.

Le chapitre I est consacré à l'étude des conditions nécessaires d'optimalité du premier ordre et à la dualité de ces trois problèmes. Nous appliquons directement la dualité de l'optimisation convexe pour définir les problèmes duaux des problèmes (P1) et (P2). Nous étudions ensuite les conditions nécessaires d'optimalité à l'aide de la théorie des sous-gradients en optimisation. Ces caractérisations donnent la classe des méthodes de sous-gradients pour la résolution des trois problèmes ci-dessus. Ces méthodes de sous-gradients, introduites par Pham Dinh Tao pour la résolution du problème (P1), ont besoin de nouvelles caractérisations plus précises des solutions optimales pour une résolution complète de cette classe de problèmes.

Le chapitre II présente une classe de procédés qui constitue une synthèse des méthodes de résolution du problème (P1). Celle-ci conduit à des variantes intéressantes en pratique. Nous avons programmé les algorithmes de synthèse en langage Pascal et les avons mis en oeuvre sur un micro-ordinateur (Micral 90.50). A partir d'une série d'exemples qui ont été testés, on donne les résultats expérimentaux de quelques exemples qui permettent de faire la comparaison entre ces méthodes.

Le chapitre III traite la classe des problèmes d'optimisation non convexe qui peuvent être mis sous la forme du problème (P3) et les méthodes de sa résolution. On montre que le problème d'optimisation où la fonction objectif et les contraintes s'écrivent comme différence de fonctions convexes peut être transformé en un problème de type (P3). On explicite un problème d'application (loi du mélange, ELF) qui permet de justifier l'intérêt pratique de cette étude.

Pour sa résolution, nous nous sommes basés sur les méthodes d'optimisation convexe et les méthodes du chapitre II.

Notre première démarche concerne essentiellement le cas particulier suivant :

$$\min\{ \langle c, x \rangle / x \in D, g(x) \geq 0 \} \quad (P'3)$$

où D est un polytope non vide.

On dispose des caractérisations de ses solutions globales, qui donnent des algorithmes simples et finis (convergence en un nombre fini de pas). Nous adaptons ensuite ces algorithmes pour la résolution du problème général (P3) en se basant sur quelques méthodes d'accumulation des contraintes de l'optimisation convexe.

Le chapitre IV est consacré aux problèmes de nature particulière (programmation quadratique non convexe, programmation fractionnaire) et qui rentrent dans le cadre des problèmes étudiés dans les chapitres précédents. On montre que le problème de maximisation d'une forme quadratique non convexe sur un polyèdre borné non vide se ramène à un problème de type (P3). En se basant sur les procédés du chapitre III, on donne un algorithme qui trouve en un nombre fini de pas un point de Kuhn & Tucker associé à la solution globale. A la fin, on donne une correspondance entre les trois principaux problèmes traités dans cette thèse et une classe très large des problèmes de programmation fractionnaire. Ce chapitre certifie que les procédés décrits dans les trois chapitres précédents peuvent apporter des améliorations à des problèmes particuliers.



CHAPITRE I

Méthodes de sous-gradients



1.1. Introduction

Dans ce chapitre, on s'intéresse aux conditions nécessaires d'optimalité du premier ordre et aux algorithmes de sous-gradients pour ces trois types de problèmes d'optimisation non convexe:

$$1) \max \{f(x) / x \in C\}$$

$$2) \min \{g(x) - h(x) / x \in C\}$$

$$3) \min \{h(x) / x \in C, g(x) \geq 0\}$$

où f , h et g sont convexes sur \mathbb{R}^n et C une partie non vide de \mathbb{R}^n , compacte et convexe.

Le paragraphe 1.2, est consacré à l'étude du problème 1). Nous appliquons directement la dualité de l'optimisation convexe pour définir le problème dual. Nous étudions ensuite les liens entre les solutions des problèmes primal-dual et les conditions nécessaires d'optimalité en terme de sous gradients de f et de la fonction indicatrice de C . Ces caractérisations sont la base des méthodes de sous-gradients pour la résolution du problème 1). Ces méthodes de sous-gradients sont introduites par Pham Dinh Tao [28,30,31]. Elles peuvent être classées dans la catégorie des méthodes d'optimisation non différentiable et qui sont les analogues des méthodes de Frank-Wolfe et du gradient projeté en optimisation convexe.

Dans le paragraphe 1.3, on développe les résultats du paragraphe précédent pour le problème 2). Notre étude est complètement différente de celles faites dans [12,48,49]. Elle permet d'obtenir des caractérisations intéressantes des éléments optimaux. De plus ces caractérisations fournissent certaines idées importantes sur l'application des méthodes de sous gradients pour sa résolution.

Au dernier paragraphe, nous étudions le problème 3). On donne les caractérisations de ces solutions optimales. Nous présenterons un algorithme de type sous gradient qui généralise celui de H.Tuy [53,54] proposé pour le cas où la fonction g est différentiable.

L'inconvénient des méthodes de sous-gradients est que, d'une façon générale, on ne sait pas affirmer si les solutions locales fournies sont des solutions globales. Par contre, elles sont très faciles à programmer et sont les seules actuellement à pouvoir traiter des problèmes (de cette classe) de grande taille. Elles servent aussi à obtenir un bon choix de vecteurs initiaux pour une classe d'algorithmes dans lesquels n'interviennent pas les caractérisations des solutions optimales (chapitres II et III).

Enfin, pour une résolution complète de cette classe de problèmes, ces méthodes de sous-gradients ont besoin des nouvelles améliorations sur les conditions nécessaires d'optimalité.

1.2. Maximisation d'une fonction convexe sur une partie compacte convexe.

On considère le problème (P) d'optimisation non convexe suivant:

$$\lambda = \max\{f(x) \mid x \in C\} \quad (P)$$

avec les hypothèses suivantes

- 1- f est une fonction convexe sur $X = \mathbb{R}^n$ et à valeurs réelles
- 2- C est une partie non vide de X , compacte et convexe.

On notera l'ensemble des solutions de (P) par \mathcal{P} .

$$\mathcal{P} = \{x^* \in C \mid f(x^*) = \lambda\}$$

1.2.1. Problème dual associé à (P).

On sait que [22,34] par l'intermédiaire de la polaire f^* de f ;

$$f(x) = \max\{\langle x, y \rangle - f^*(y) \mid y \in Y\} = \max\{\langle x, y \rangle - f^*(y) \mid y \in \partial f(x)\} \quad (Y = \mathbb{R}^n)$$

Par suite on peut écrire :

$$\lambda = \max\{f(x) \mid x \in C\} = \max\{\max\{\langle x, y \rangle - f^*(y) : y \in Y\} \mid x \in C\} =$$

$$\max\{\max\{\langle x, y \rangle - f^*(y) : x \in C\} \mid y \in \partial f(C)\}$$

Le problème dual (Q) est ainsi défini par :

$$\lambda = \max \{ \chi_C^*(y) - f^*(y) / y \in Y \} = \max \{ \chi_C^*(y) - f^*(y) / y \in \partial f(C) \} \quad (Q)$$

On notera (Q) l'ensemble des solutions du problème dual (Q).

$$\mathcal{Q} = \{ y^* \in \partial f(C) / \chi_C^*(y^*) - f^*(y^*) = \lambda \}.$$

Pour chaque $y \in \partial f(C)$, considérons le problème (P_y) :

$$\lambda_y = \max \{ \langle x, y \rangle - f^*(y) / x \in C \} = \chi_C^*(y) - f^*(y) \quad (P_y)$$

dont l'ensemble des solutions (P_y) est exactement $\partial \chi_C^*(y)$.

On verra dans la suite lors de la description des méthodes de sous gradients les rôles des problèmes (P_y), en même temps que la justification possible de la convergence de ces méthodes vers une solution optimale de (P) lorsque le vecteur initial x^0 se trouve assez proche d'un élément de (P)

Proposition 2.1

Les hypothèses 1- et 2- entraînent que:

$$1- \mathcal{P} = \bigcup_{y^* \in \mathcal{Q}} \mathcal{P}_{y^*} = \bigcup_{y^* \in \mathcal{Q}} \partial \chi_C^*(y^*)$$

$$2- \mathcal{Q} = \bigcup_{x^* \in \mathcal{P}} \partial f(x^*)$$

$$3- \partial \chi_C^*(y^*) \subset \partial f^*(y^*), \forall y^* \in \mathcal{Q}.$$

Démonstration :

Si $y^* \in \mathcal{Q}$, alors pour tout $x^* \in \partial \chi_C^*(y^*)$ on a :

$$\langle x^*, y^* \rangle - f^*(y^*) = \chi_C^*(y^*) - f^*(y^*) = \lambda$$

et donc :

$$f(x^*) = \lambda \text{ et } x^* \in \partial f^*(y^*).$$

et par conséquent :

$$\partial\chi^*_C(y^*) \subset \partial f^*(y^*), \forall y^* \in \mathcal{Q} \text{ et } \bigcup_{y^* \in \mathcal{Q}} \partial\chi^*_C(y^*) \subset \mathcal{P}.$$

Inversement ; si $x^* \in \mathcal{P}$, alors pour tout $y^* \in \partial f(x^*)$ (qui est non vide car f est continue) on a :

$$f(x^*) = \langle x^*, y^* \rangle - f^*(y^*) = \lambda$$

et donc :

$$x^* \in \partial\chi^*_C(y^*) \text{ et } \chi^*_C(y^*) - f^*(y^*) = \lambda, \text{ i.e. } y^* \in \mathcal{Q}.$$

Nous avons ainsi montré 1-, 3- et l'inclusion $\bigcup_{x^* \in \mathcal{P}} \partial f(x^*) \subset \mathcal{Q}$.

Maintenant, si $y^* \in \mathcal{Q}$, alors on a :

$$\chi^*_C(y^*) - f^*(y^*) = \lambda \text{ et } \langle x^*, y^* \rangle - f^*(y^*) = \lambda \quad \forall x^* \in \partial\chi^*_C(y^*).$$

d'où , pour tout $x^* \in \partial\chi^*_C(y^*)$ (non vide car l'ensemble C est compact) , on a :

$$f(x^*) = \langle x^*, y^* \rangle - f^*(y^*) = \lambda$$

ceci est équivalent à $x^* \in \mathcal{P}$ et $y^* \in \partial f(x^*)$. \square

Corollaire 2.1

Soit f une fonction convexe sur X et soit C un ensemble convexe compact de X . Si f est strictement convexe sur chaque sous ensemble convexe de $\{x \in X / \partial f(x) \neq \emptyset\} = \text{dom}(\partial f)$ ou si f^* est différentiable sur $\partial f(C)$, alors

$$\partial\chi^*_C(y^*) = \partial f^*(y^*) = \{\nabla f^*(y^*)\}, \forall y^* \in \mathcal{Q}$$

Dans ce cas , on aura $\mathcal{P} = \bigcup_{y^* \in \mathcal{Q}} \{\nabla f^*(y^*)\}$.

Démonstration :

Conséquence immédiate de la proposition précédente. \square

Corollaire 2.2

\mathcal{P} est non vide si et seulement si \mathcal{Q} est non vide

\mathcal{P} est non vide si C est un ensemble non vide , compact et convexe .

Démonstration

Conséquence immédiate de la proposition précédente. \square

1.2.2. Caractérisation analytique des solutions optimales du problème(P).

Soit f une fonction convexe sur X et C une partie convexe compacte non vide de X . On suppose dans tout la suite que f n'est pas constante sur C .

On va décrire ci-dessous d'autres caractérisations des solutions du problème (P) analogues à celles de la proposition 2.1.

Proposition 2.2

- 1) $\mathcal{P} \subset \text{Frr } C$ (frontière relative de C)
- 2) $\partial f(x^*) \subset \partial \chi_C(x^*)$, $\forall x^* \in \mathcal{P}$.
- 3) $(x^*) = P_C(x^* + \rho \partial f(x^*))$, $\forall x^* \in \mathcal{P}$, $\forall \rho > 0$,
où P_C est la projection sur C .

Démonstration:

- 1) Est immédiate d'après la définition de la frontière relative de C et la convexité de la fonction f [34].
- 2) Il est clair que si $x^* \in \mathcal{P}$, alors $C \subset \{x \in X / f(x) \leq f(x^*)\}$.
C'est-à-dire ,analytiquement [34] $\partial f(x^*) \subset \partial \chi_C(x^*)$.
- 3) Est une simple traduction de la propriété 2). \square

Proposition 2.3

Les hypothèse 1-et 2- entraînent que :

$$1) \max \{ f'(x^*; x - x^*) / x \in C \} = \max \{ \chi_C^*(y) - \langle x^*, y \rangle / y \in \partial f(x^*) \}$$

$$2) \partial f(x^*) \subset \partial \chi_C(x^*) \text{ si et seulement si } \max \{ f'(x^*; x - x^*) / x \in C \} = 0.$$

3) Si $f'(x^*; d) < 0$ pour tout vecteur $d \neq 0$ de cône tangent $K_C(x^*)$, alors x^* est un maximum local strict .

Démonstration

Rappelons d'abord que $f'(x^*; x)$ est appelée dérivée directionnelle de f au point x^* par rapport à x et

$$f'(x^*; x) = \lim_{\substack{t \rightarrow 0 \\ t > 0}} (f(x^* + tx) - f(x^*)) / t.$$

D' après les hypothèses sur f , on a [2,35] :

$$1) f'(x^*; x - x^*) = \max \{ \langle x - x^*, y \rangle / y \in \partial f(x^*) \}$$

ainsi

$$\begin{aligned} & \max \{ f'(x^*; x - x^*) / x \in C \} \\ &= \max \{ \max \{ \langle x, y \rangle - \langle x^*, y \rangle / x \in C \} / y \in \partial f(x^*) \} \\ &= \max \{ \chi_C^*(y) - \langle x^*, y \rangle / y \in \partial f(x^*) \}. \end{aligned}$$

2) Il est clair que $\partial f(x^*) \subset \partial \chi_C(x^*)$ si et seulement si :

$$\langle x-x^*, y \rangle \leq 0, \forall x \in C \text{ et } \forall y \in \partial f(x^*).$$

Il est équivalent à :

$$\max(\max\{ \langle x-x^*, y \rangle / y \in \partial f(x^*) \mid x \in C \}) = \max\{ f'(x^*, x-x^*) / x \in C \} = 0.$$

3) Supposons que le point x^* ne soit pas un maximum local strict de f sur C , alors on peut trouver une suite $(x_n) \subset C$ telle que $\lim_{n \rightarrow \infty} (x_n) = x^*$ et $f(x_n) \geq f(x^*)$. (*)

posons $d_n = (x_n - x^*) / \|x_n - x^*\|$, la suite (d_n) converge vers un élément d^* vérifiant $\|d^*\| = 1$, c'est-à-dire $d^* \in K_C(x^*)$.

D'autre part, si $\lambda_n = \|x_n - x^*\|$ on a :

$$f'(x^*, d^*) = \lim_{n \rightarrow \infty} (f(x^* + \lambda_n d^*) - f(x^*)) / \lambda_n = \lim_{n \rightarrow \infty} (f(x^* + \lambda_n d_n) - f(x^*)) / \lambda_n = \lim_{n \rightarrow \infty} (f(x_n) - f(x^*)) / \lambda_n$$

Or, d'après (*) on a :

$$f'(x^*, d^*) \geq 0, \text{ ce qui contredit l'hypothèse. } \square$$

Pour chaque $u \in C$, considérons le problème suivant :

$$\alpha(u) = \text{Max} \{ f'(u; x-u) / x \in C \} \quad R(u)$$

dont l'ensemble des solutions est noté par $\mathfrak{R}(u)$; et $S(u)$ son problème dual (dans le sens de la dualité introduite au paragraphe 1.2.1).

$$\alpha(u) = \text{Max} \{ \chi_C^*(y) - \langle u, y \rangle / y \in \partial f(u) \} \quad S(u)$$

dont l'ensemble des solutions est noté par $\mathfrak{P}(u)$.

Proposition 2.4

Soit f une fonction convexe dans X et C un ensemble compact convexe de X . Alors :

$$1) \mathfrak{R}(u) = \bigcup_{y^* \in \mathfrak{P}(u)} \partial \chi_C^*(y^*)$$

$$2) \mathfrak{P}(u) = \bigcup_{x^* \in \mathfrak{R}(u)} \partial \chi_{\partial f(u)}^*(x^* - u)$$

Démonstration :

Analogue à celle de la proposition 2.1. \square

1.2.3 Méthodes de sous-gradients

Méthodes du type -1-

Ces méthodes sont basées sur les caractérisations des éléments de l'ensemble (\mathcal{P}) données en proposition 2.1 et proposition 2.2.

a) Première forme de méthodes du type -1-

Partant de $x^0 \in \text{Frr } C$, on construit les deux suites (x^k) et (y^k) de la manière suivante :

$$x^0 \in \text{Frr } C \quad \mapsto \quad y^0 \in \partial f(x^0)$$

$$x^1 \in \partial \chi_C^*(y^0) \quad \mapsto \quad y^1 \in \partial f(x^1)$$

$$x^k \in \partial \chi_C^*(y^{k-1}) \quad \mapsto \quad y^k \in \partial f(x^k)$$

Proposition 2.5 (convergence de la méthode)

1) $f(x^k) \leq f(x^{k+1})$

Egalité si et seulement si $x^k \in \partial \chi_C^*(y^k)$ et $y^k \in \partial f(x^{k+1})$. Dans ce cas, on aura $\partial f(x^k) \cap \partial \chi_C(x^k) \neq \emptyset$ (i.e. x^k est un point critique au sens de Toland).

2) La suite $f(x^k)$ converge de façon croissante vers $\mu \leq \lambda$. Cette valeur μ dépend de x^0 et des sous gradients choisis pour la construction des suites (x^k) et (y^k) .

3) $\lim_{k \rightarrow \infty} [\chi_C^*(y^k) - \langle x^k, y^k \rangle] = 0$

4) Les ensembles des valeurs d'adhérence de (x^k) et (y^k) sont non vides. De plus pour toute valeur d'adhérence x^* de (x^k) (respectivement y^* de (y^k)) il existe une valeur d'adhérence y^* de (y^k) (respectivement x^* de (x^k)) telle que :

$$4.1 - x^* \in \partial \chi_C^*(y^*) \text{ et } y^* \in \partial f(x^*)$$

$$4.2 - \lim_{k \rightarrow \infty} f(x^k) = f(x^*) = \mu$$

$$4.3 - \chi_C^*(y^*) - f^*(y^*) = f(x^*) = \mu$$

Démonstration

La continuité de f (par hypothèse) et de celle de χ_C^* (car l'ensemble C est compact convexe et non vide) entraînent que les suites (x^k) et (y^k) sont bien définies.

De la convexité de f et χ_C on a :

$$f(x^{k+1}) \geq f(x^k) + \langle x^{k+1} - x^k, y^k \rangle$$

et

$$x^{k+1} \in \partial \chi_C^*(y^k) \Rightarrow \langle x^{k+1}, y^k \rangle \geq \langle x^k, y^k \rangle$$

d'où

$$f(x^{k+1}) \geq f(x^k)$$

Si $f(x^{k+1}) = f(x^k)$, alors

$$\chi_C^*(y^k) = \langle x^k, y^k \rangle \Leftrightarrow x^k \in \partial \chi_C^*(y^k)$$

et de plus

$$f(x^{k+1}) + f^*(y^k) = \langle x^k, y^k \rangle \quad (f(x^{k+1}) = f(x^k))$$

d'où

$$y^k \in \partial f(x^{k+1}) \cap \partial \chi_C(x^k)$$

Inversement, $x^k \in \partial \chi_C^*(y^k)$ et $y^k \in \partial f(x^{k+1})$ entraînent que :

$$\langle x^{k+1}, y^k \rangle = \langle x^k, y^k \rangle \text{ et } f(x^{k+1}) + f^*(y^k) = \langle x^k, y^k \rangle$$

ce qui donne

$$f(x^{k+1}) = f(x^k)$$

La propriété 2) est évidente car λ est fini

La propriété 3) est aussi immédiate

Puisque la suite $(x^k) \subset C$ et C est un ensemble compact non vide, alors l'ensemble des valeurs d'adhérence de la suite (x^k) est non vide. La même remarque pour la suite (y^k) , car la continuité de f et la compacité de C impliquent la compacité de $\partial f(C)$ [34].

Soit x^* la limite d'une sous suite (x^{k_i}) de la suite (x^k) et soit y^* la limite d'une sous suite (y^{k_i}) de la suite (y^k) .

De la propriété 3), on a $x^* \in \partial \chi_C^*(y^*)$.

Puisque la multi-application $\partial f: x \mapsto \partial f(x)$ est semi continue supérieurement sur X [22,34], alors elle est fermée et on a:

$x^{k_i} \mapsto x^*$, $y^{k_i} \mapsto y^*$ et $y^{k_i} \in \partial f(x^{k_i})$ entraînent que $y^* \in \partial f(x^*)$.

La propriété 4.2 est une conséquence de la continuité de f .

De la propriété 4.1, on a

$$\chi_C^*(y^*) = \langle x^*, y^* \rangle = f(x^*) + f^*(y^*)$$

ce qui donne la propriété 4.3. \square

Remarques

1) Choix de $y^k \in \partial f(x^k)$

On arrête l'algorithme si $\partial f(x^k) \cap \partial \chi_C(x^k) \neq \emptyset$. Dans le cas contraire, on prend un autre $y^k \in \partial f(x^k) - \partial \chi_C(x^k)$ et on relance cet algorithme.

2) Choix de x^k

Il est clair que le bon choix de l'élément x^k est le suivant:

$$f(x^k) = \max \{ f(x) / x \in \partial \chi_C^*(y^k) \} = \max \{ f(x) / x \in C, \langle x, y^k \rangle - \chi_C^*(y^k) = 0 \}.$$

b) *Deuxième forme des méthodes du type -1-*

D'après les indications sur le choix de sous gradients et le vecteur x^k et les caractérisations des maximums locaux en proposition 2.3, nous présentons la variante suivante.

Partant de $x^0 \in \text{Frr } C$, on construit les deux suites (x^k) et (y^k) de la façon suivante:

$$\begin{aligned} x^0 \in \text{Frr } C & \mapsto y^0 \in \mathcal{P}(x^0) \\ x^1 \in \mathcal{C}(y^0) & \mapsto y^1 \in \mathcal{P}(x^1) \\ \\ x^k \in \mathcal{C}(y^{k-1}) & \mapsto y^k \in \mathcal{P}(x^k) \end{aligned}$$

où $\mathcal{C}(y)$ est l'ensemble des solutions du problème $T(y)$ défini par :

$$\text{Max} \{ f(x) \mid x \in \partial \chi_C^*(y) \} \quad T(y)$$

Proposition 2.6 (convergence de la méthode)

1) $f(x^k) \leq f(x^{k+1})$

Egalité si et seulement si $\partial f(x^k) \subset \partial \chi_C(x^k)$ et $y^k \in \partial f(x^{k+1})$.

2) La suite $f(x^k)$ converge de façon croissante vers $\mu \in \lambda$. Cette valeur μ dépend uniquement du vecteur initial x^0 .

3) $\lim_{k \rightarrow \infty} (\chi_C^*(y^k) - \langle x^k, y^k \rangle) = 0$

4) Les ensembles de valeurs d'adhérence de (x^k) et (y^k) sont non vides. De plus pour toute valeur d'adhérence x^* de (x^k) (respectivement y^* de (y^k)) il existe une valeur d'adhérence y^* de (y^k) (respectivement x^* de (x^k)) telle que:

4.1- $x^* \in \partial \chi_C^*(y^*)$, $y^* \in \partial f(x^*)$

4.2- Si la multi-application $\mathcal{P}(x)$ est fermée sur l'ensemble de valeurs d'adhérence de (x^k) , alors $\partial f(x^*) \subset \partial \chi_C(x^*)$

4.3- Si la multi-application $\mathcal{C}(y)$ est fermée sur l'ensemble de valeurs d'adhérence de (y^k) , alors $\partial \chi_C^*(y^*) \subset \partial f^*(y^*)$

$$4.4- \lim_{k \rightarrow \infty} f(x^k) = f(x^*) = \mu$$

$$4.5- \chi_C^*(y^*) - f^*(y^*) = f(x^*) = \mu$$

Démonstration

La continuité de f (par hypothèse) et celle de χ_C^* (car l'ensemble C compact convexe et non vide) entraînent que les suites (x^k) et (y^k) sont bien définies

1) Soit $y^k \in \mathcal{P}(x^k)$ et $x^{k+1} \in \mathcal{C}(y^k)$.

De la convexité de f et χ_C on a

$$f(x^{k+1}) - f(x^k) \geq \langle x^{k+1} - x^k, y^k \rangle \quad \forall y^k \in \partial f(x^k).$$

et

$$x^{k+1} \in \partial \chi_C^*(y^k) \Leftrightarrow \langle x^{k+1} - x^k, y^k \rangle \geq 0$$

d'où $f(x^{k+1}) \geq f(x^k)$

$f(x^{k+1}) = f(x^k)$ si et seulement si

$$i) \langle x^{k+1}, y^k \rangle = \langle x^k, y^k \rangle \Leftrightarrow x^k \in \partial \chi_C^*(y^k) \Leftrightarrow y^k \in \partial f(x^k).$$

$$ii) f(x^k) + f^*(y^k) = \langle x^k, y^k \rangle = \langle x^{k+1}, y^k \rangle = f(x^{k+1}) + f^*(y^k) \Leftrightarrow y^k \in \partial f(x^{k+1})$$

D'autre part, de la définition du problème $\mathcal{P}(x^k)$ on a

$$y^k \in \mathcal{P}(x^k) \Leftrightarrow \langle x - x^k, y^k \rangle \leq 0 \quad \forall x \in C \text{ et } \forall y^k \in \partial f(x^k).$$

ceci est équivalent à

$$\partial f(x^k) \subset \partial \chi_C(x^k)$$

2) Est simple car λ est fini

3) Est une conséquence de la propriété 2)

4) Soit x^* et y^* des valeurs d'adhérence des suites (x^k) et (y^k) (existent toujours car C et $\partial f(C)$ sont compacts).

Par passage à des sous suites (x^{k_i}) et (y^{k_i}) qui convergent vers x^* et y^* respectivement on a :

$$x^* \in \partial \chi_C^*(y^*) \text{ et } y^* \in \partial f(x^*)$$

4.2- $\mathcal{P}(x^*)$ fermée sur $x^* \Rightarrow y^* \in \mathcal{P}(x^*) \Leftrightarrow$

$$\max_{x \in C} \langle y^*, x - x^* \rangle \geq \max_{x \in C} \langle y, x - x^* \rangle \quad \forall y \in \partial f(x^*)$$

puisque $x^* \in \partial \chi_C^*(y^*)$ d'après 4.1, alors

$$\langle y^*, x - x^* \rangle \leq 0 \quad \forall x \in C$$

d'où

$$\mathcal{P}(x^*) \text{ fermée sur } x^* \Rightarrow \langle y, x - x^* \rangle \leq 0 \quad \forall y \in \partial f(x^*), \forall x \in C \Leftrightarrow \partial f(x^*) \subset \partial \chi_C(x^*)$$

4.3- $\mathcal{E}(y^*)$ fermée sur $\partial f(C) \Rightarrow x^* \in \mathcal{E}(y^*) \subset \partial \chi_C^*(y^*)$ ([2,8]) \Leftrightarrow

$$f(x^*) + f^*(y^*) \geq f(x) + f^*(y^*) \quad \forall x \in C, \langle x^*, y^* \rangle = \langle x, y^* \rangle,$$

puisque $y^* \in \partial f(x^*)$ d'après 4.1, alors $f(x^*) + f^*(y^*) = \langle x^*, y^* \rangle$

donc

$\mathcal{E}(y^*)$ fermée sur $\partial f(C) \Rightarrow f(x) + f^*(y^*) = \langle x, y^* \rangle \quad \forall x \in \partial \chi_C^*(y^*) \Leftrightarrow$

$$\partial \chi_C^*(y^*) \subset \partial f^*(y^*)$$

La propriété 4.4 est une conséquence de la continuité de f .

De 4.1 on déduit que

$$\chi_C^*(x^*) = \langle x^*, y^* \rangle = f(x^*) + f^*(y^*)$$

ceci donne la propriété 4.5. \square

1.2.4 Méthodes du type-2-

(Analogue du gradient projeté en optimisation convexe).

Ces méthodes sont basées sur la caractérisation suivante

$$[x^*] = P_C(x^* + \rho \partial f(x^*)), \quad \forall x^* \in \mathcal{P}, \quad \forall \rho > 0.$$

On se donne $\rho > 0$, partant de $x^0 \in \text{Frr}(C)$, on construit (x^k) et (y^k) de la manière suivante:

$$y^k \in \partial f(x^k) \mapsto x^{k+1} \in P_C(x^k + \rho^k y^k).$$

Les paramètres ρ^k ($\rho^k > 0$), doivent se trouver dans un intervalle compact. Chaque $\rho(x^k)$ peut être choisie de façon que $x^k + \rho^k y^k \notin C$. Ce choix à pour effet d'accélérer la convergence de l'algorithme.

Proposition 2.7 (convergence de la méthode)

$$1) f(x^{k+1}) \geq f(x^k) + \langle x^{k+1} - x^k, y^k \rangle \geq f(x^k) + 1/\rho^k \|x^{k+1} - x^k\|^2$$

Egalité si et seulement si $x^k = x^{k+1} = P_C(x^k + \rho^k y^k)$.

2) La suite croissante $f(x^k)$ est convergente vers $\mu \leq \lambda$. Cette valeur μ dépend du x^0 et le choix des sous gradients y^k dans la construction de ces suites (x^k) et (y^k) . De plus, on a $\lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0$.

3) Les ensembles des points d'adhérence de (x^k) et (y^k) sont non vides. De plus, pour toute valeur d'adhérence x^* de x^k (respectivement

y^* de (y^k)) il existe une valeur d'adhérence y^* de (y^k) (respectivement x^* de (x^k)) telle que :

$$3.1- x^* = P_C(x^* + \rho y^*) , \text{ où } y^* \in \partial f(x^*) \text{ et } \rho > 0 .$$

$$3.2- \lim_{k \rightarrow \infty} f(x^k) = f(x^*) = \mu .$$

Démonstration

Par définition de la projection , on a :

$$\langle x^{k+1} - \rho^k y^k - x^{k+1} , x^k - x^{k+1} \rangle \leq 0$$

ainsi :

$$\langle x^{k+1} - x^k , y^k \rangle \geq 1 / \rho^k \| x^{k+1} - x^k \|^2 ,$$

d'où les propriétés 1) et 2) sont immédiates .

Montrons la propriété 3), puisque C et $\partial f(C)$ sont compacts non vides alors les ensembles des valeurs d'adhérence de (x^k) et (y^k) sont non vides . Soit x^* la limite d'une sous suite (x^{k_i}) de la suite (x^k) et y^* la limite d'une sous suite (y^{k_i}) de la suite (y^k) telle que :

$$x^* = \lim_{k_i \rightarrow \infty} x^{k_i} , y^* = \lim_{k_i \rightarrow \infty} y^{k_i} \text{ et } y^* \in \partial f(x^*)$$

puisque f et P_C sont continues en x^k et $\lim_{k_i \rightarrow \infty} \|x^{k+1} - x^k\| = 0$, alors

$$x^* = P_C(x^* + \rho^* y^*) , \text{ où } y^* \in \partial f(x^*) \text{ et } \rho^* = \lim_{k_i \rightarrow \infty} \rho^{k_i} \geq \rho > 0 . \quad \square$$

Remarques générales

1- Dans les méthodes du type-2- , on a seulement le choix du sous gradient $y^k \in \partial f(x^k)$ quand $x^{k+1} = x^k$.

2- $\lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0$ pour les méthodes du type -1- si $f \in C^2(X)$ et fortement convexe , par contre cette propriété est toujours vérifiée pour les méthodes du type -2-. Dans ce cas, si l'ensemble des valeurs d'adhérences de (x^k) est fini alors la suite (x^k) est convergente .

3-D'après les caractérisations des solutions optimales (proposition 2.1-2.3), si x^0 est assez voisin d'une solution optimale alors l'algorithme converge vers une solution optimale. Un tel vecteur x^0 peut être obtenu à l'aide des méthodes de chapitres II et III .

4- Si C est un polyèdre borné, alors l'algorithme de la première forme

est fini (car chaque élément de la suite (x^k) est un sommet du polyèdre C).

5- Nous n'avons pas encore pu démontrer que dans la deuxième forme des méthodes du type -1- on a $\partial f(x^*) \subset \partial \chi_C(x^*)$ et $\partial \chi_C^*(y^*) \subset \partial f^*(y^*)$ pour tout couple (x^*, y^*) de l'ensemble des valeurs d'adhérence de (x^k, y^k) . Mais en pratique on peut utiliser la remarque précédente.

1.3 Minimisation de la différence de deux fonctions convexes sur une partie compacte convexe.

Nous allons étudier dans ce paragraphe un problème qui généralise celui que nous avons étudié au paragraphe 1.2. On utilisera les mêmes notations. Considérons le problème d'optimisation non convexe suivant:

$$\lambda = \min \{ f(x) = g(x) - h(x) \mid x \in C \} \quad (P)$$

avec les hypothèses suivantes

- 1- h et g sont convexes sur $X = \mathbb{R}^n$ et à valeurs réelles.
- 2- C est une partie compacte, convexe et non vide de X .

1.3.1 Problème dual associé à (P).

Tenant compte de la section 1.2.1, on peut écrire (P) de la manière suivante.

$$\begin{aligned} \lambda &= \min \{ g(x) - h(x) \mid x \in C \} = \min \{ g(x) - \sup \{ \langle x, y \rangle - h^*(y) \mid y \in \partial h(x) \} \mid x \in C \} \\ &= \min \{ \min \{ g(x) + h^*(y) - \langle x, y \rangle \mid y \in \partial h(x) \} \mid x \in C \} \\ &= \min \{ \min \{ g(x) + h^*(y) - \langle x, y \rangle \mid x \in C \} \mid y \in \partial h(C) \} \end{aligned}$$

or

$$(g + \chi_C)^*(y) = \sup \{ \langle x, y \rangle - (g + \chi_C)(x) \mid x \in X \} = \sup \{ \langle x, y \rangle - g(x) \mid x \in C \}$$

d'où

le problème dual (Q) du problème (P) est le suivant:

$$\lambda = \min \{ h^*(y) - (g + \chi_C)^*(y) \mid y \in \partial h(C) \} \quad (Q)$$

On notera par \mathcal{P} (respectivement \mathcal{Q}) l'ensemble des solutions optimales du problème (P) (respectivement du problème dual (Q)).

$$\mathcal{P} = \{x^* \in C / f(x^*) = g(x^*) - h(x^*) = \lambda\},$$

et

$$\mathcal{Q} = \{y^* \in \partial h(C) / h^*(y^*) - (g + \chi_C)^*(y^*) = \lambda\}.$$

1.3.2 Caractérisation des solutions optimales.

Proposition 3.1

Si les hypothèses ci-dessus sur le problème (P) sont vérifiées, alors on aura:

$$1) \min\{g(x) - \langle y^*, x \rangle / x \in C\} = g(x^*) - \langle y^*, x^* \rangle \quad \forall x^* \in \mathcal{P} \text{ et } y^* \in \partial h(x^*)$$

$$2) \partial h(x^*) \subset \partial(g + \chi_C)(x^*) = \partial g(x^*) + \partial \chi_C(x^*) \quad \forall x^* \in \mathcal{P}$$

$$3) \partial(g + \chi_C)^*(y^*) \subset \partial h^*(y^*) \quad \forall y^* \in \mathcal{Q}$$

$$4) \bigcup_{x^* \in \mathcal{P}} \partial h(x^*) = \mathcal{Q}$$

$$5) \bigcup_{y^* \in \mathcal{Q}} \partial(g + \chi_C)^*(y^*) = \mathcal{P}$$

Démonstration:

(1)- Soient $x^* \in \mathcal{P}$ et $y^* \in \partial h(x^*)$ (qui est non vide car h est continue).

De la convexité de h on déduit :

$$g(x) - h(x) + h(x^*) - g(x^*) \leq g(x) - g(x^*) - \langle y^*, x - x^* \rangle \quad \forall x \in C$$

d'où la propriété 1).

Tenant compte de 1) et la définition de la polaire de $g + \chi_C$ (qui est fini d'après les hypothèses 1- et 2-) on a:

$$\lambda = g(x^*) - h(x^*) = g(x^*) - \langle x^*, y^* \rangle + h^*(y^*) = -(g + \chi_C)^*(y^*) + h^*(y^*)$$

ce qui donne

$$y^* \in \mathcal{Q} \text{ et } y^* \in \partial(g + \chi_C)(x^*) = \partial g(x^*) + \partial \chi_C(x^*).$$

d'où

$$\partial h(x^*) \subset \partial \chi_C(x^*) + \partial g(x^*) \quad \forall x^* \in \mathcal{P}$$

Nous avons aussi montré les inclusions

$$\bigcup_{x^* \in \mathcal{P}} \partial h(x^*) \subset \mathcal{Q} \text{ et } \mathcal{P} \subset \bigcup_{y^* \in \mathcal{Q}} \partial(\chi_C + g)^*(y^*).$$

Maintenant, soient $y^* \in \mathcal{Q}$ et $x^* \in \partial(\chi_C + g)^*(y^*)$ (qui est non vide car les hypothèses 1 et 2 entraînent la continuité de $(\chi_C + g)^*$)

$$\langle x^*, y^* \rangle - h^*(y^*) = (\chi_C + g)^*(y^*) + g(x^*) - h^*(y^*) = -\lambda + g(x^*) = h(x^*)$$

d'où

$$x^* \in \partial h^*(y^*) \Leftrightarrow y^* \in \partial h(x^*) \text{ et } \partial(\chi_C + g)^*(y^*) \subset \partial h^*(y^*)$$

par conséquent on a aussi

$$\mathcal{Q} \subset U_{x^* \in \mathcal{P}} \partial h(x^*) \text{ et } U_{y^* \in \mathcal{Q}} \partial(g + \chi_C)^*(y^*) \subset \mathcal{P}$$

D'où la vérification des cinq propriétés. \square

Proposition 3.2

Les hypothèses ci-dessus entraînent que

1- $x^* \in \partial(g + \chi_C)^*(y^*)$ si et seulement si x^* est solution optimale du problème d'optimisation convexe suivant
 $\text{Min}\{g(x) - \langle x, y^* \rangle / x \in C\}$

2- $\partial h(x^*) \subset \partial g(x^*) + \partial \chi_C(x^*)$ si et seulement si

$$\text{Min}\{g'(x^*; x-x^*) - h'(x^*; x-x^*) / x \in C\} = 0$$

3- Si $g'(x^*, d) - h'(x^*, d) > 0$ pour tout vecteur $d \neq 0$ de cône tangent $K_C(x^*)$, alors x^* est un minimum local strict du problème (P).

Démonstration:

1) Est une simple traduction de la condition nécessaire et suffisante d'optimalité du problème d'optimisation convexe $\text{min}\{g(x) - \langle x, y^* \rangle / x \in C\}$.

2) De 1), on a

$$\partial h(x^*) \subset \partial g(x^*) + \partial \chi_C(x^*) \Leftrightarrow \text{min}\{g(x) - \langle x, y \rangle / x \in C\} = g(x^*) - \langle x^*, y \rangle \quad \forall y \in \partial h(x^*)$$

$$\Leftrightarrow g'(x^*, x-x^*) - \langle x-x^*, y \rangle \geq 0 \quad \forall x \in C, \forall y \in \partial h(x^*) \text{ (car la fonction } g(x) - \langle x, y \rangle \text{ est convexe)} \Leftrightarrow \text{min}\{g'(x^*, x-x^*) - h'(x^*, x-x^*) / x \in C\} = 0.$$

3) Analogue à celle de la proposition 2.3 du paragraphe 1.2

1.3.3 Méthodes de sous-gradients

Remarquons que sans perte de généralité, on peut supposer que les fonctions g et h sont fortement convexes (car $\text{min}\{g(x) - h(x)\} = \text{min}\{(g(x) + \|x\|^2) - (h(x) + \|x\|^2)\}$). On trouve l'intérêt de cette hypothèse sur la convergence de la méthode de sous-gradient décrite ci-dessous (voir les propriétés 1-, 4.3- et 4.6- de la proposition 3.3)

Pour chaque $y \in \partial h(C)$, on notera par (P_y) le problème d'optimisation convexe suivant:

$$\text{Min}\{g(x) - \langle x, y \rangle / x \in C\}$$

(P_y)

dont l'ensemble des solutions (\mathcal{P}_y) est singleton.

A partir de $x^0 \in C$, on construit les deux suites (x^k) et (y^k) de la façon suivante :

$$\begin{aligned} x^0 \in C & \quad \mapsto \quad y^0 \in \mathcal{P}(x^0) \\ x^1 \in \mathcal{P}_{y^0} & \quad \mapsto \quad y^1 \in \mathcal{P}(x^1) \\ \\ x^k \in \mathcal{P}_{y^{k-1}} & \quad \mapsto \quad y^k \in \mathcal{P}(x^k) \end{aligned}$$

où $\mathcal{P}(x)$ est l'ensemble des solutions du problème d'optimisation non convexe $S(x)$ défini par

$$\max\{(g + \chi_C)^*(y) - \langle x, y \rangle \mid y \in \partial h(x)\}.$$

Proposition 3.3 (convergence de la méthode)

- 1- $f(x^{k+1}) = g(x^{k+1}) - h(x^{k+1}) \leq f(x^k) = g(x^k) - h(x^k)$
 Egalité si et seulement si $x^k = x^{k+1}$. Dans ce cas on aura $\partial h(x^k) \subset \partial(\chi_C + g)(x^k)$ et $\nabla(g + \chi_C)^*(y^k) = \nabla h^*(y^k)$.
- 2- La suite $f(x^k)$ converge vers $\mu \geq \lambda$. Cette valeur μ dépend du vecteur initial x^0 .
- 3- $\lim_{k \rightarrow \infty} \{ (g + \chi_C)^*(y^k) + g(x^k) - \langle x^k, y^k \rangle \} = 0$
- 4- Les ensembles des valeurs d'adhérence de (x^k) et de (y^k) sont non vides. Pour toute valeur d'adhérence x^* de (x^k) (respectivement y^* de (y^k)) il existe une valeur d'adhérence y^* de (y^k) (respectivement x^* de (x^k)) telle que:
 - 4.1- $y^* \in \partial h(x^*) \cap \partial(g + \chi_C)(x^*)$
 - 4.2- Si la multi-application $\mathcal{P}(x)$ est fermée sur l'ensemble de valeurs d'adhérence de (x^k) , alors $\partial h(x^*) \subset \partial(g + \chi_C)(x^*)$
 - 4.3- $x^* \in \partial(g + \chi_C)^*(y^*) = \nabla(g + \chi_C)^*(y^*) \in \partial h^*(y^*) = \nabla h^*(y^*)$
 - 4.4- $\lim_{k \rightarrow \infty} f(x^k) = f(x^*) = \mu$

$$4.5 - h^*(y^*) - (g + \chi_C)^*(y^*) = f(x^*) = \mu$$

$$4.6 - \lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0.$$

Démonstration

1- La forte convexité de h entraîne que :

$$f(x^{k+1}) - f(x^k) \leq g(x^{k+1}) - g(x^k) - \langle y^k, x^{k+1} - x^k \rangle - \rho \|x^{k+1} - x^k\|^2 \quad (\rho > 0)$$

et par définition du point x^{k+1} on a :

$$f(x^{k+1}) - f(x^k) \leq -\rho \|x^{k+1} - x^k\|^2 < 0.$$

d'où l'algorithme s'arrête quand $x^{k+1} = x^k$ et dans ce cas par construction, x^k est solution du problème (P, y^k)

c'est-à-dire

$$g(x) - g(x^k) \geq \langle y^k, x - x^k \rangle \quad \forall x \in C$$

d'après le choix du y^k on a

$$g(x) - g(x^k) \geq \langle y, x - x^k \rangle \quad \forall y \in \partial h(x^k) \quad \forall x \in C$$

ce qui revient à dire

$$\partial h(x^k) \subset \partial (g + \chi_C)(x^k)$$

Comme la fonction $g + \chi_C$ est strictement convexe sur C et la fonction h est strictement convexe sur \mathbb{R}^n (d'après les hypothèses), alors $\partial (g + \chi_C)^*(y) = \nabla (g + \chi_C)^*(y)$ pour tout $y \in \text{int}(\text{dom}(g + \chi_C)^*)$ et $\partial h^*(y) = \nabla h^*(y)$ pour tout $y \in \text{int}(\text{dom } h^*)$ (cf. [34], section 26). Ceci donne $\nabla (g + \chi_C)^*(y^k) = \nabla h^*(y^k)$

2- Car λ est fini

3- Par définition du point x^{k+1} , on a

$$f(x^{k+1}) \leq g(x^{k+1}) - g(x^k) - \langle y^k, x^{k+1} - x^k \rangle = -(g + \chi_C)^*(x^k) - g(x^k) + \langle x^k, y^k \rangle + f(x^k)$$

d'après 2-, ceci donne la propriété 3-

4- Soit x^* et y^* des valeurs d'adhérence des suites (x^k) et (y^k) (existent toujours car C et $\partial h(C)$ sont compacts).

Par passage à des sous suites (x^{k_i}) et (y^{k_i}) qui convergent vers x^* et y^* respectivement on a :

$$x^* \in \partial(g + \chi_C)^*(y^*) \text{ et } y^* \in \partial h(x^*)$$

Il est facile de vérifier la propriété 4.2 si la multi-application $\mathcal{P}(x^*)$ est fermée sur x^* (c'est-à-dire $y^* \in \mathcal{P}(x^*)$ [2,8]).

Le reste de la proposition est facile.

Remarques

1- Notre algorithme est basé sur la linéarisation de la fonction h , on doit ainsi résoudre à chaque itération un problème d'optimisation convexe et non convexe.

2- Remarquons que sans perte des propriétés de convergence de l'algorithme décrit ci-dessus, il n'est pas nécessaire de résoudre complètement à chaque itération le problème (P_y^k) . En fait il suffit de prendre x^{k+1} comme un point qui minimise strictement la fonction $g(x) - \langle y^k, x \rangle$ suivant une direction de descente ([2,27]).

3- Le meilleur choix de l'arrêt de cet algorithme est $\partial h(x^k) \subset \partial(g + \chi_C)(x^k)$.

4- Il est évident que cette classe d'algorithmes dépend de la décomposition de la fonction f en une différence de deux fonctions convexes g et h . Malheureusement, la détermination d'une meilleure décomposition reste encore ouverte.

1.4. Minimisation d'une fonction convexe sur une partie compacte convexe et une contrainte anticonvexe

Dans ce paragraphe, on s'intéresse à l'étude des programmes mathématiques qui peuvent s'écrire sous la forme suivante:

$$\lambda = \min\{f(x) / x \in C, g(x) \geq 0\} \quad (P)$$

avec les hypothèses suivantes

- 1- f et g sont convexes sur $X = \mathbb{R}^n$ et à valeurs réelles.
- 2- C est un ensemble d'intérieur non vide de X , compact et convexe.
- 3- Il existe un point $w \in C$ tel que $g(w) < 0$ et $f(w) < \lambda$.
(l'hypothèse 3- implique que (P) est un problème d'optimisation non convexe)

On notera par $G = \{x \in \mathbb{R}^n / g(x) \geq 0\}$ et $\partial G = \{x \in \mathbb{R}^n / g(x) = 0\}$.

1.4.1 Caractérisation des solutions optimales

Proposition 4.1

Si les hypothèses 1-, 2-, 3- sont vérifiées, alors on aura :

$$1- \min \{ f(x) / x \in C, g(x) \geq 0 \} = \min \{ f(x) / x \in C, g(x) = 0 \}.$$

2- Chaque solution locale (ou globale) du problème (P) est un élément de $C \cap \partial G$.

3- Une condition nécessaire pour que x^* soit une solution optimale du problème (P) est que $\partial g(x^*) \subset \text{cône } \partial f(x^*) + \partial \chi_C(x^*)$,

$$\text{où cône } \partial f(x^*) = \bigcup_{\lambda > 0} \lambda \partial f(x^*).$$

Démonstration:

1) Soit x^* un minimum global du problème (P) tel que $g(x^*) > 0$. Alors, l'hypothèse $g(w) < 0$ et la convexité de C entraînent que il existe un élément $y^* = \lambda x^* + (1-\lambda)w$ tel que $g(y^*) = 0$, $0 < \lambda < 1$ et $y^* \in C$.

Or de la convexité de f et l'hypothèse $f(w) < \lambda$, on déduit que $f(y^*) < f(x^*)$.

D'où la contradiction.

2) Si x^* est un minimum local alors il existe un réel $\eta > 0$ tel que

$$f(x) \geq f(x^*) \quad \forall x \in C \cap B(x^*, \eta),$$

où $B(x^*, \eta)$ est une boule de centre x^* et de rayon η .

D'autre part si $g(x^*) > 0$, alors la continuité de g entraîne qu'il existe un réel $\rho > 0$ tel que

$$g(x) \geq g(x^*) > 0 \quad \forall x \in B(x^*, \rho).$$

Donc, si x^* est un minimum local tel que $g(x^*) > 0$, il existe un réel $\varepsilon > 0$ ($\varepsilon = \min(\eta, \rho)$) tel que

$$\forall x \in C \cap B(x^*, \varepsilon) \Rightarrow f(x) \geq f(x^*) \text{ et } g(x) \geq g(x^*).$$

Or, de la convexité de C on a :

tout $x \in C \cap B(x^*, \varepsilon)$ peut s'écrire sous la forme

$$x = \lambda x^* + (1-\lambda)w, \quad \text{où } 0 < \lambda < 1.$$

De l'hypothèse 3- et la convexité de f on déduit que :

$$f(x) < f(x^*) \quad \forall x \in C \cap B(x^*, \varepsilon).$$

D'où la contradiction avec

$$f(x) \geq f(x^*), \forall x \in C \cap \bar{B}(x^*, \epsilon).$$

3) De la convexité de g on a :

$$D(y) = \{x \in \mathbb{R}^n / \langle y, x - x^* \rangle \geq 0\} \subset \{x \in \mathbb{R}^n / g(x) \geq 0\} \quad \forall y \in \partial g(x^*) \text{ et } g(x^*) = 0$$

Soit x^* une solution optimale du problème (P).

Pour tout $y^* \in \partial g(x^*)$ (qui est non vide car g est continue) on a :

$$\min \{f(x) / x \in C / \langle x - x^*, y^* \rangle \geq 0\} = f(x^*)$$

Ce qui revient à dire [2,35] que pour tout $y \in \partial g(x^*)$ il existe un réel $\alpha \geq 0$ tel que $0 \in \partial f(x^*) - \alpha y + \partial \chi_C(x^*)$.

De l'hypothèse 3-, on aura $\alpha > 0$

(car $\alpha = 0$ implique que x^* est aussi solution du problème $\min\{f(x) / x \in C\}$) ceci est équivalent à

$$\partial g(x^*) \subset \text{cône } \partial f(x^*) + \partial \chi_C(x^*).$$

1.4.2 Méthodes de sous-gradients

La méthode que nous présenterons est basée sur la caractérisation des solutions optimales en proposition 4.2. Elle consiste à construire à partir de $x^0 \in C \cap \partial G$ la suite $(x^k) \subset C \cap \partial G$ de la façon suivante :

Supposons calculé x^k , alors :

1) $x^k \mapsto z^{k+1}$ par la relation

$$x^k \mapsto y^{k+1} \in \mathcal{P}(x^k) \mapsto z^{k+1} \in \mathcal{P}(y^{k+1})$$

2) Si $g(z^{k+1}) > 0$, alors on définit x^{k+1} par la relation

$$x^{k+1} = \lambda z^{k+1} + (1-\lambda)w, \text{ tel que } 0 < \lambda < 1 \text{ et } g(x^{k+1}) = 0$$

3) Si $g(z^{k+1}) = 0$ et $f(z^{k+1}) = f(x^k)$, alors l'algorithme s'arrête.

Sinon on pose $x^{k+1} = z^{k+1}$

où $\mathcal{P}(x^k)$ est l'ensemble des solutions du problème d'optimisation non convexe défini par :

$$\max \{ \chi_C^*(y) - \langle x^k, y \rangle \mid y \in \partial g(x^k), f(x) \leq f(x^k) \}$$

$\mathcal{P}(y^{k+1})$ est l'ensemble des solutions du problème d'optimisation convexe défini par:

$$\min \{ f(x) \mid x \in C, \langle x - x^k, y^{k+1} \rangle \geq 0 \}.$$

Proposition 4.2 (convergence de la méthode)

1) $f(x^{k+1}) \leq f(x^k)$

Egalité si et seulement si $f(x^k) = f(z^{k+1})$ et $g(z^{k+1}) = 0$. Dans ce cas on aura $\partial g(x^k) \subset \text{cône } \partial f(x^k) + \partial \chi_C(x^k)$, $z^{k+1} \in C \cap \partial G$

2) $f(x^k)$ converge vers $\mu \geq \lambda$. Cette valeur dépend du vecteur initial x^0 .

3) Les ensembles des valeurs d'adhérence de (x^k) , (y^k) , (z^k) sont non vides. Pour toute valeur d'adhérence x^* de (x^k) (resp y^* de (y^k)) il existe une valeur d'adhérence y^* de (y^k) (resp x^* de (x^k)) telle que:

3.1- $y^* \in \text{cône } \partial f(x^*) + \partial \chi_C(x^*)$

3.2- Si la multi-application $\mathcal{P}(x)$ est fermée sur C , alors $\partial g(x^*) \subset \text{cône } \partial f(x^*) + \partial \chi_C(x^*)$, $y^* \in \mathcal{P}(x^*)$, $x^* \in \mathcal{P}(y^*)$.

Démonstration

1) L'hypothèse 3- et la convexité de f entraînent que

$$f(x^{k+1}) \leq \lambda f(z^{k+1}) + (1-\lambda) f(w) < f(z^{k+1}) \leq f(x^k).$$

et donc on a l'égalité si et seulement si

$$f(x^k) = \min \{ f(x) \mid x \in C, \langle y^{k+1}, x - x^k \rangle \geq 0 \}$$

c'est -à-dire ; pour tout $x \in C$ on a

$$f(x) < f(x^k) \Rightarrow \langle y^{k+1}, x - x^k \rangle < 0$$

d'après le choix du $y^{k+1} \in \mathcal{P}(x^k)$ on a

$$x \in C, f(x) < f(x^k) \Rightarrow \langle y, x - x^k \rangle < 0 \quad \forall y \in \partial g(x^k)$$

ceci est équivalent à :

$$\partial g(x^k) \subset \text{cône } \partial f(x^k) + \partial \chi_C(x^k).$$

2) Est évidente car λ est fini.

3) Les hypothèses 1- et 2- entraînent que les ensembles des valeurs d'adhérence des suites (x^k) , (y^k) et (z^k) sont non vides et que les multi-applications \mathcal{P} est fermée [2,8]. Soit x^* la limite d'une sous suite (x^{k_i}) de la suite (x^k) , y^* la limite d'une sous suite (y^{k_i}) de la suite (y^k) et z^* la limite d'une sous suite (z^{k_i}) de la suite (z^k) où $z^* = \lambda w + (1-\lambda)x^*$.

Par construction de la suite (z^{k_i}) on a:

$$f(z^{k_i}) \leq f(x) \quad \forall x \in C \text{ et } \langle y^{k_i}; x - x^{k_i} \rangle \geq 0$$

Par passage à la limite on aura:

$$f(z^*) \leq f(x) \quad \forall x \in C \text{ et } \langle y^*; x - x^* \rangle \geq 0$$

ceci donne

$$\min\{f(x) / \forall x \in C \text{ et } \langle y^*; x - x^* \rangle \geq 0\} = f(x^*) \Leftrightarrow y^* \in \text{cône } \partial f(x^*) + \partial \chi_C(x^*)$$

Si la multi-application \mathcal{P} est fermée (i.e. $y^* \in \mathcal{P}(x^*)$) alors

$$f(x^*) = \min\{f(x) / x \in C, \langle y; x - x^* \rangle \geq 0\} \quad \forall y \in \partial g(x^*)$$

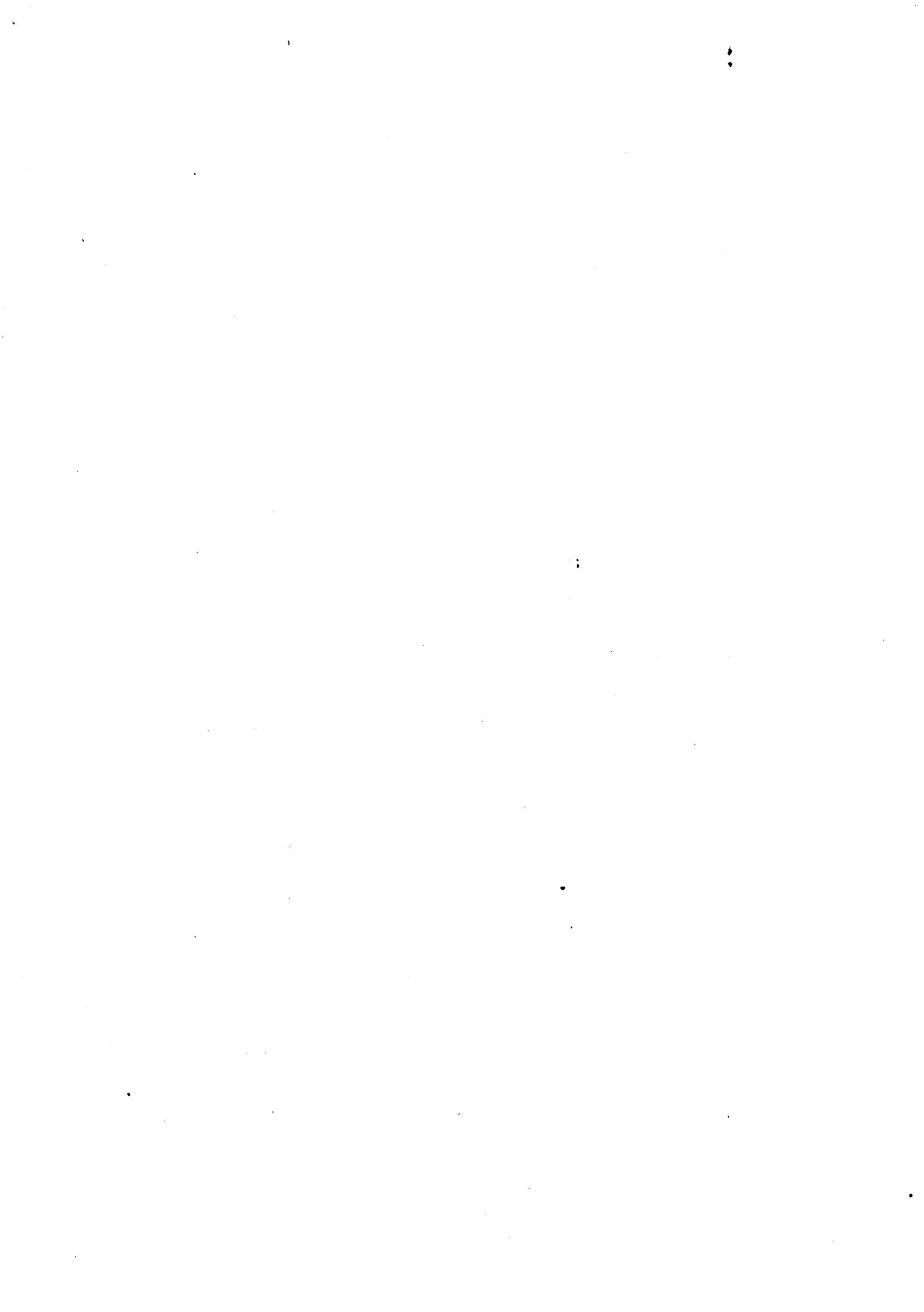
$$\Leftrightarrow \partial g(x^*) \subset \text{cône } \partial f(x^*) + \partial \chi_C(x^*). \quad \square$$

Remarques

1- Notre algorithme est basé sur la linéarisation de la contrainte anticonvexe " $g(x) \geq 0$ ".

2- Sans perdre les propriétés de la convergence de l'algorithme décrit ci-dessus, il n'est pas nécessaire de résoudre complètement à chaque itération le problème $\min\{f(x) / x \in C, \langle y^{k+1}, x - x^k \rangle \geq 0\}$. En fait il suffit de prendre z^{k+1} comme un point admissible qui minimise la fonctionnelle f suivant une direction de descente (il existe si x^k n'est pas une solution locale).

3) On donne au chapitre III, les caractérisations complètes sur les solutions globales du problème (P), lorsque C est un polytope dans \mathbb{R}^n .



CHAPITRE II

**Minimisation globale d'une fonction
concave sur une partie compacte convexe**



11.1 Introduction

Nous nous sommes attachés dans ce chapitre à étudier les méthodes qui permettent d'obtenir les solutions globales du problème (P) suivant :

$$\text{Min}\{f(x) / x \in C\} \quad (P)$$

où f est concave sur $X = \mathbb{R}^n$ et C est une partie convexe compacte de X .

On trouvera au paragraphe 11.2 une bibliographie des méthodes de résolution qui sont bien connues dans la littérature. Ceci nous conduit à classer ces méthodes en trois catégories.

Au paragraphe 11.3, nous présentons un algorithme général qui constitue une synthèse des algorithmes qui utilisent les coupes planes introduites par H. Tuy [51] pour le cas où C est polyédral. On met l'accent en particulier sur la détermination des coupes pour le cas de la dégénérescence des sommets du polyèdre C .

Le paragraphe 11.4 est consacré à des méthodes de résolution par partition de l'ensemble admissible C . Elles sont basées sur l'introduction de procédés de subdivision de C et le calcul d'un certain minorant de $f(x)$ dans chaque sous-domaine, et ceci jusqu'à ce que la solution globale soit mise en évidence. On donne quelques variantes intéressantes en pratique, en particulier pour le cas où C est polyédral.

Dans le paragraphe 11.5, on étudie les méthodes de type approximation extérieure de l'ensemble admissible C . L'idée générale de cette classe de méthodes est de remplacer (P) par une suite de sous problèmes (P_k) de la forme $\text{min}\{f(x) / x \in S_k\}$ (où $C \subset S_k$) plus simples à mettre en oeuvre. Dans le but de donner un procédé plus efficace, nous présentons un algorithme qui généralise celui de J.E. Falk et K.R. Hofman [13] proposé pour le cas où C est polyédral.

Ces algorithmes ont été programmés sur un micro-ordinateur (Micral 90.50) et la discussion des expériences numériques sur quelques exemples fait l'objet du dernier paragraphe.

11.2 Quelques méthodes de résolution connues dans la littérature

Le pionnier dans ce domaine est sans doute H.Tuy [51] avec sa première méthode de résolution du problème (P) sous l'hypothèse que C est un polyèdre convexe, non vide et non dégénéré. Pour son importance historique, cette méthode est rapidement rappelée.

-On commence avec l'identification d'un optimum local x^* du problème (P) puis on détermine une coupe de Tuy qui est l'hyperplan passant par les points x^1, \dots, x^n qui se trouvent sur les arêtes issues de x^* vers les points extrêmes adjacents et tels que $f(x^1) = f(x^2) = \dots = f(x^n) = f(x^*)$.

-On calcule le point y^* le plus éloigné du polyèdre C par rapport à cette coupe. Si y^* appartient au simplexe des sommets x^*, x^1, \dots, x^n , alors x^* est une solution globale du problème (P). Sinon, on subdivise le domaine C en un nombre fini de sous-domaines. On recommence la procédure sur chacun de ces domaines et ceci jusqu'à ce que la solution globale soit obtenue.

Hoang Tuy ne fournit pas la preuve de finitude de sa méthode et d'ailleurs Zwart [57] donne un contre-exemple sur lequel elle boucle. Cependant l'idée s'est largement répandue et a inspiré de nombreux algorithmes sous différentes formes des problèmes d'optimisation non convexe. Des fines méthodes apparentées à celle de Tuy sont d'ailleurs proposées dans [9,46,58] pour le cas où C est polyédral borné non dégénéré et dans [25] si cette dernière hypothèse est omise.

Dans le domaine de la programmation en nombre entiers, on montre dans [32] que le problème de programmation en nombres bivalents:

$$\text{Min}\{ \langle c, x \rangle / Ax \leq b, x \in (0, 1)^n \}$$

peut se ramener à un problème de type (P):

$$\text{Min}\{ cx + M \langle x, e - x \rangle / Ax \leq b, 0 \leq x \leq e \}$$

où M est un réel positif arbitrairement grand et e est un vecteur dont toutes les coordonnées sont égales à 1.

E.Balas trouve [3] l'application de la méthode de coupe de la convexité développée dans [14] et fondée sur les idées de Tuy pour la résolution de cette classe de problèmes.

Une autre adaptation des coupes de Tuy pour résoudre les problèmes de programmation bilinéaire :

$$\text{Min } \{f(x,y) / x \in C_1, y \in C_2\}$$

où C_1 et C_2 sont des polyèdre convexe non vides et pour x (resp y) fixé $f(.,y)$ (resp $f(x,.)$) est linéaire est donnée dans [21,39].

Une autre approche différente de celle de Tuy et qui est de type approximation extérieure se trouve dans [13] sous l'hypothèse de non dégénérescence du polyèdre C . l'idée de base est de le remplacer par une suite finie de programmes linéaires où la fonction objectif est l'enveloppe convexe de la fonction à minimiser sur un ensemble qui contient la région réalisable. A chaque itération, l'ensemble des variables de l'enveloppe convexe augmente et l'ensemble approché devient plus réduit. Dans [36] l'auteur développe une méthode de résolution en nombre fini de pas sous l'hypothèse que C est polyédral. En partant, par le calcul d'une minorante et une majorante estimation, de la solution globale par la résolution de $2n$ sous problèmes de programmation linéaire :

$$\text{Max } \{ \langle +u_i, x \rangle, \langle -u_i, x \rangle / x \in C, i=1, \dots, n \}$$

où u_i ($i=1, \dots, n$) sont les vecteurs propres de la matrice des dérivées secondes. Une partition du domaine réalisable en nombre fini de sous-domaines et une adaptation d'algorithme proposé en [13] dans chaque sous domaine sont décrites.

On connaît aussi une classe de méthodes qui traite le cas où l'ensemble C est une partie non vide de \mathbb{R}^n , convexe et compacte. Le principe de ces méthodes est la subdivision du domaine réalisable en sous-domaines, et que par la construction des simplexes qui contiennent ces sous-domaines, on calcule la minorante estimation de la solution optimale. Le processus de subdivision est arrêté dès que l'on trouve un élément de l'ensemble admissible dont la valeur de la fonction objectif approche la minorante estimation à ϵ - près. Les critères de la subdivision de la région réalisable pour assurer la convergence de cette classe de méthodes sont données dans [18]. On remarque également, dans les récents articles de H.Tuy [47,55] pour ce sujet, des améliorations intéressantes de son premier travail, où il s'est basé sur ces critères de subdivision.

11.3 Méthodes de coupe pour le cas linéaire

Dans ce paragraphe nous ferons l'hypothèse que l'ensemble C est un polyèdre convexe borné. Sans perte de généralité, on supposera que C est écrit sous la forme standard.

Traditionnellement C est défini de la manière suivante:

$$C = \{x \in \mathbb{R}^n / Ax = b, x \geq 0\} \quad (3.1)$$

où A est une matrice de type (m,n) , $(m \leq n)$, et b est un vecteur de type $(m,1)$ et le rang de la matrice A est (m) .

L'idée générale des algorithmes qui convergent en un nombre fini d'itérations est basée essentiellement sur la caractérisation suivante des solutions globales.

$$\text{Min}\{f(x) / x \in C\} = \text{Min}\{f(x) / x \in E(C)\} \quad (3.2)$$

où $E(C)$ est l'ensemble des sommets du polyèdre C.

C'est la raison pour laquelle la programmation linéaire est un outil pour la mise en oeuvre de cette classe de méthodes. Nous présentons une méthode qui constitue une synthèse de certains algorithmes qui convergent en un nombre fini de pas.

11.3.1 Rappels et notations[38]

Définition 3.1.

On dit que le polyèdre C est écrit sous la forme canonique par rapport à une base réalisable $J \subset \{1, \dots, n\}$, si C peut s'exprimer sous la forme suivante :

$$C = \{x \in \mathbb{R}^n / A^* x_J + U x_{J'} = b^*, x_J \geq 0, x_{J'} \geq 0\} \quad (3.3)$$

où

- A^* est une matrice de type $(m, n-m)$, U est la matrice unité de type (m, m) et b^* est un vecteur de type $(m, 1)$ non négatif.

- J désigne l'ensemble des indices de base, J' désigne l'ensemble des indices hors base, $\text{card}(J) = m$ et $\text{card}(J') = n-m$.

Définition 3.2

Etant donnée une base réalisable J du polyèdre C , la solution $x_J = b^*$ et $x_{J^c} = 0$ est appelée sommet du polyèdre C ou solution de la base réalisable J .

Définition 3.3

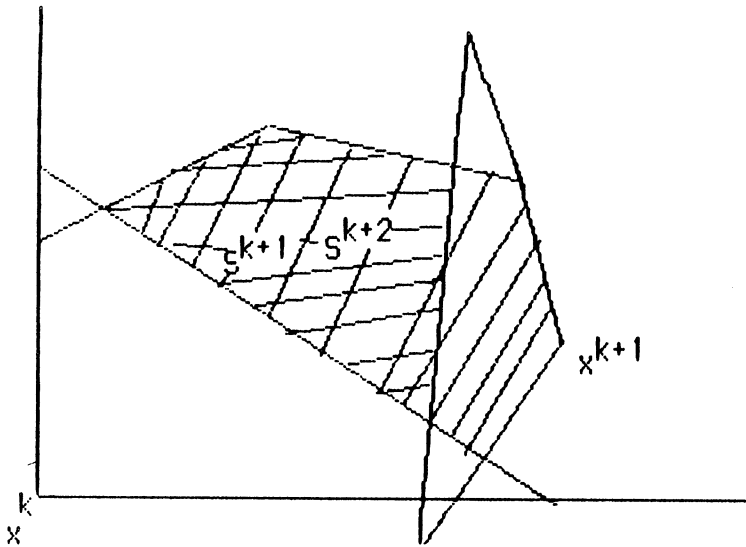
On dit que $x^* = (x_J, x_{J^c}) = (b^*, 0)$ est un sommet non dégénéré s'il y a exactement m -composantes de x^* qui sont strictement positives. Le sommet x^* est dégénéré s'il a moins de m -composantes strictement positives.

11.3.2 Description de l'algorithme général.

On initialise avec un optimum local x^0 (cf.1.2) et on pose $x^* = x^0, S^0 = C$ et $k=0$.

Itération k : ($k=0,1,\dots$)

- Etape 1: On détermine une coupe $\langle a_k, x \rangle \geq \alpha_k$ qui sépare le sommet x^k du polyèdre S^k des autres sommets y du polyèdre C tels que $f(y) < f(x^k)$ (cf.11.3.3). On ajoute cette contrainte aux contraintes du polyèdre S^k et ce nouveau polyèdre est noté par S^{k+1} .
- Etape 2: Si aucun sommet y de S^{k+1} tel que $\langle a_k, y \rangle > \alpha_k$ n'est sommet du polyèdre original C , alors x^* est un minimum global. Sinon, on détermine un nouveau sommet x^{k+1} du polyèdre C et qui appartient à S^{k+1} (cf.11.3.3).
- Etape 3: On calcule l'ensemble $E(x^{k+1})$ des sommets adjacents au sommet x^{k+1} . On remplace x^* par y si $f(x^*) > f(y) = \min\{ f(x) / x \in E(x^{k+1}) \}$, $k \rightarrow k+1$ et on retourne à l'itération k .



11.3.3 Mise en oeuvre de l'algorithme.

a) Construction d'une coupe.

Pour formaliser cette construction, on suppose que S^k est écrit sous la forme canonique par rapport à une base réalisable $I=J \cup H$ dont la solution de base correspondant est le sommet (x^k, y^k) , où x^k est le sommet du polyèdre C et y^k désigne les variables d'écart de nouvelles contraintes. De façon plus précise, (x^k, y^k) est la solution triviale du système suivant:

$$\begin{aligned} A^*x_J + Ux_H &= b^* \\ D^*x_J + U'y_H &= d^* \end{aligned} \quad x=(x_J, x_H) \geq 0, y_H \geq 0 \quad (3.4)$$

où D^* est une matrice de type $(k, n-m)$, U' est la matrice unité de type (k, k) , $D^*x_J + U'y_H = d^*$, $y_H \geq 0$ et $H = \{n+1, \dots, n+k\}$ définissent l'ensemble des nouvelles contraintes à la $k^{\text{ème}}$ itération.

Nous allons étudier le cas de non dégénérescence et de dégénérescence du sommet x^k .

Pour $j=1, \dots, n-m$
on calcule

$$\omega_j = \min (b_i^* / a_{i,j}^* , i= 1, \dots, m \text{ et } a_{i,j}^* > 0) = b_{s,j}^* / a_{s,j}^*$$

et on définit la direction a''_j de la façon suivante:

$$a''_j = \begin{cases} -a_{i,j}^* & \text{pour } i= 1, \dots, m \\ +1 & \text{pour } i=J_j \\ 0 & \text{pour } i \in J'-(J_j) \end{cases}$$

où a''_j est un vecteur de type $(n, 1)$ et J_j est la $j^{\text{ème}}$ élément de J' .

On détermine

$$\rho_j = \max (\rho / f(x^k + \rho a''_j) \geq f(x^*) , \omega_j \leq \rho \leq M). \quad (3.5)$$

avec M une constante assez grande.

Premier cas

Si $\rho_j > 0$ pour tout j alors la construction de la coupe est basée sur le lemme de Glover suivant.

$$\text{Soit } S = \{ x \in \mathbb{R}^n / x = \sum_{j=1}^{n-m} (\lambda_j \rho_j) a''_j + x^k , \lambda_j \geq 0 , \sum_{j=1}^{n-m} \lambda_j \leq 1 \}$$

Lemme de Glover [14]

Pour tout $y \in S \cap C$, on a:

$$1 - y = x^k + \sum_{j=1}^{n-m} x_{j,j} a''_j \quad \text{où} \quad \sum_{j=1}^{n-m} x_{j,j} / \rho_j \leq 1.$$

2-La coupe $\sum_{j=1}^{n-m} x_{j,j} / \rho_j \geq 1$ sépare le sommet x^k des autres sommets y

du polyèdre C telle que $f(y) < f(x^k)$.

Définition 3.4

Le demi-espace fermé $\sum_{j=1}^{n-m} x_{j,j} / \rho_j \geq 1$ est appelé coupe de Tuy.

Proposition 3.1

Si x^k est un sommet non dégénéré ou si x^k est un sommet dégénéré mais qui est un minimum local, alors $\rho_j > 0$ pour tout $j=1, \dots, n-m$.

Démonstration

1) la non dégénérescence entraîne que :

$$\omega^s_j > 0 \quad \forall j=1, \dots, n-m$$

et donc de (3.5) on a $\rho_j > 0 \quad \forall j=1, \dots, n-m$.

2) si x^k est un minimum local, alors il existe une boule $B(x^k, \eta)$ de centre x^k et de rayon $\eta > 0$ tel que

$$f(x) \geq f(x^k) \geq f(x^*) \quad \forall x \in B(x^k, \eta)$$

et ceci implique que $\rho_j > 0 (j=1, \dots, n-m)$. \square

Deuxième cas

Dans le cas où il existe un indice j tel que $\rho_j = 0$, la construction d'une coupe qui sépare le sommet x^k des autres sommets est basée sur le théorème suivant.

Théorème 3.1 [39]

Etant donné un polyèdre convexe borné:

$$S = \{ x \in \mathbb{R}^n / E x_1 + U x_2 = f, x \geq 0 \},$$

où E est une matrice de type $(m, n-m)$ et U est une matrice unité de type (m, m) . On définit pour chaque $f_j \neq 0 (j=1, \dots, m)$ l'indice $r(j)$ et la constante $\epsilon_j(r)$ par les relations suivantes

$$e_{j,r(j)} = \max \{ e_{j,i} : i=1, \dots, n-m \}$$

(3.6)

et

$$\epsilon_j(r(j)) = \min \{ f_i / e_{i,r(j)} / e_{i,r(j)} > 0 \}$$

La coupe $\sum x_i \geq \epsilon = \min \{ \epsilon_j / j=1, \dots, m, f_j \neq 0 \}$ sépare le sommet $l \in l$

x^k défini par $x_1 = f, x_2 = 0$ des autres sommets du polyèdre S .

Démonstration

Pour chaque $f_j \neq 0 (j=1, \dots, m)$, considérons le problème (P_j) de programmation linéaire suivant

$$\alpha(j) = \min \{x_{1j} / x \in S, \sum_{i \in I} x_i \leq \xi < \epsilon\} \quad (P_j)$$

$$\text{où } x_{1j} = f_j - \sum_{i=1}^{n-m} e_{j,i} x_i$$

D'après le choix de l'indice $r(j)$ et la constante ξ (3.6), on peut affirmer par les critères d'optimalité de la méthode du simplexe [38] que pour tout j on a

$$\alpha(j) = f_j - \xi e_{j,r(j)} > f_j - \epsilon_j(r(j)) e_{j,r(j)} \geq 0. \quad (1)$$

D'autre part, on sait que [38] x'' est un sommet différent du sommet x' si et seulement s'il existe au moins un indice $j \in I$ tel que $x''_j = 0$ et $x'_j \neq 0$.

Compte tenu de (1), le théorème est démontré. \square

Remarques 3.1

1- Par construction, cette coupe est complètement différente des coupes précédentes (du type H.Tuy). Elle peut être utilisée aussi dans le cas non dégénéré.

2- L'inconvénient de cette coupe est qu'elle ne sépare aucun sommet adjacent y tel que $f(y) \geq f(x^*)$.

3- D'après la proposition 3.1, on peut éliminer ce cas par les méthodes de sous-gradients (chapitre 1) ou par passage à l'un du sommet adjacent.

b) Détermination d'un sommet du polyèdre C

L'inconvénient de cet algorithme réside dans le fait que pour assurer la convergence en un nombre fini d'itérations il faut déterminer à chaque $k^{\text{ème}}$ itération un sommet du polyèdre original C qui appartient à S^{k+1} . Le lecteur pourra consulter l'article de A.Mathay et Winston [25] où ils proposent une procédure finie basée sur la programmation linéaire. Il consiste à trouver une base réalisable dont les variables hors bases sont les variables du polyèdre C par la résolution d'un nombre fini de problèmes du type programmation linéaire. On montre dans [9] que ce sous problème peut se ramener à un problème de programmation en nombres entiers. Cependant la recherche d'un algorithme pratiquement efficace reste encore ouverte.

11.3.4 Variante de l'algorithme général

Il existe plusieurs problèmes d'optimisation non convexe importants en applications où il est facile de définir une fonction affine h qui approche bien la fonction objectif f tel que $f(x) \leq h(x)$ pour tout $x \in C$. On pourra consulter, en particulier, sur ce sujet [46]. Il est préférable dans ce cas de remplacer l'algorithme général par:

On commence avec le point minimum x^0 de h dans C . Si $h(x^0) \leq f(x^0)$, alors x^0 est un minimum global. Sinon, on pose $x^* = x^0$, $S^0 = C$ et $k=0$.

Itération k :

Pas 1: Comme l'étape 1

Pas 2: On applique la méthode duale du simplexe pour déterminer le point x' qui réalise le minimum de h sur S^{k+1} . Si $h(x') \leq f(x^*)$ alors on s'arrête et x^* est une solution optimale. Sinon, on détermine un nouveau sommet x^{k+1} du polyèdre C adjacent à l'un des sommets testés et qui appartient à S^{k+1} , on passe au pas 3

Pas 3: Comme l'étape 3.

Théorème 3.2

L'algorithme général et sa variante trouvent un minimum global en un nombre fini de pas.

Démonstration

Par description et la finitude de nombre des sommets du polyèdre C . \square

II.4 Algorithmes de résolution par partition de l'ensemble admissible.

Nous faisons les hypothèses suivantes

H1- l'ensemble C est une partie convexe, compacte et définie par:

$$C = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i=1, \dots, m\} \quad (4.1)$$

où $g_i (i=1, \dots, m): \mathbb{R}^n \rightarrow \mathbb{R}$ sont convexes continues dans \mathbb{R}^n .

H2- L'origine 0 de \mathbb{R}^n appartient à l'intérieur de l'ensemble C (équivalent à la condition de Slater).

II.4.1 Description de l'algorithme général

On initialise avec :

- Un n-simplexe S [s^1, s^2, \dots, s^{n+1}] (de dimension n et de sommets $s^i (i=1, \dots, n+1)$) qui contient l'origine 0.
- $\mathcal{K}_0 = (K_{0,1}, \dots, K_{0,n+1})$ l'ensemble fini des cônes polyédriques de sommet l'origine et tel que chaque cône $K_{0,j}(s^1, \dots, s^{j-1}, s^{j+1}, \dots, s^{n+1})$ soit défini par :

$$\{x \in \mathbb{R}^n / x = \lambda_1 s^1 + \dots + \lambda_{j-1} s^{j-1} + \lambda_{j+1} s^{j+1} + \dots + \lambda_{n+1} s^{n+1}, \lambda_i \geq 0\}$$
- Pour tout $K \in \mathcal{K}_0$, on définit une minorante estimation $\mu(K)$ de $f(x)$ dans $C \cap K$, i.e. $\mu(K) \leq \inf \{f(x) / x \in C \cap K\}$ (cf. II.4.3).
- On pose $\beta_1 = f(x^1) = \min \{f(x^0), f(0)\}$, où x^0 est un minimum local du problème (P) (cf. I.2).

Itération h : (h=1, ...)

Etape 1: On pose $\mathcal{K}_h = \{K \in \mathcal{K}_{h-1} / \mu(K) < \beta_h\}$. Si $\mathcal{K}_h = \emptyset$, alors on arrête l'algorithme et x^h est un minimum global.

Sinon on prend $K_h(v^1, \dots, v^n) = \arg \min \{\mu(K) / K \in \mathcal{K}_h\}$ (le cône associé à la plus petite minorante estimation).

Etape 2: On subdivise le cône $K_h(v^1, \dots, v^n)$ en deux sous cônes $K_{h,1}$ et $K_{h,2}$ de la manière suivante :

On calcule $w = (v^r + v^s) / 2$ où $\|v^r - v^s\| = \max \{\|v^i - v^j\|, i \neq j\}$

et on pose :

$$K_h(v^1, \dots, v^n) = K_{h,1}(v^1, \dots, v^{r-1}, v^{r+1}, \dots, v^n) \cup K_{h,2}(v^1, \dots, v^{s-1}, v^{s+1}, \dots, v^n).$$

Etape 3: Pour chaque $i=1,2$ on calcule $\mu(K_{h,i})$ (cf. II.4.3) et on détermine

$$\pi(w) = \rho^* w, \text{ où } \rho^* = \max\{\rho \geq 0 / \rho w \in C\} \quad (\partial C \text{ est la frontière de } C).$$

Etape 4: On prend $x^{h+1} = \arg \min \{f(x^h), f(\pi(w))\}$ et on pose $\beta_{h+1} = f(x^{h+1})$.

On remplace K_h dans \mathcal{K}_h par $(K_{h,1}, K_{h,2})$, h par $h+1$ et on retourne à l'itération h .

II.4.2 Condition de la convergence

Définition 4.1

Etant donné un cône polyédrique K de sommet l'origine, on dit qu'une partition du cône K est **exhaustive** si toute sous suite infinie décroissante (K_i^q) de la suite de cônes $K_i \subset K$ tend vers une demi droite d'origine O , i.e. $\cap K_i^q = \{x = \rho x^* / x^* \in K, \rho \geq 0\}$.

Définition 4.2

Etant donné un cône polyédrique K de sommet l'origine, on dit que le procédé de minorante estimation de $f(x)$ dans $C \cap K$ est **compatible**, si pour toute sous suite (K_i^q) infinie décroissante de la suite (K_i) qui tend vers une demi droite on a $\mu(K_i^q) - \beta_i^q \rightarrow 0$ quand $q \rightarrow \infty$.

Définition 4.3

On dit que le procédé de la subdivision sur l'ensemble admissible est **complet**, si le processus de partition par les cônes est **exhaustif** et le calcul de minorante estimation est **compatible**.

Théorème 4.1

Si le procédé de subdivision de l'ensemble admissible C décrit dans l'algorithme ci-dessus est complet, alors on aura les propriétés suivantes.

1- Si l'algorithme s'arrête à l'itération h , alors x^h est un minimum global du problème (P).

2- Tout point d'adhérence de la suite infinie (x^h) est un minimum global.

3-Si x^* est minimum global du problème (P), alors
 $f(x^h) - f(x^*) \leq \beta_h - \mu(K_h)$.

Démonstration

Soient $\mu^* = \min(f(x) / x \in C)$ et $\mu(K_h) = \min(\mu(K) / K \in \mathcal{K}_h)$.

Par définition du $\mu(K_h)$, on a

$$\mu(K_h) \leq \mu^*.$$

1) On s'arrête à la $h+1$ ème itération si $\mathcal{K}_{h+1} = \emptyset$ et dans ce cas on aura:

$$\mu^* \geq \mu(K_h) \geq \beta_h = f(x^h).$$

D'où par définition du μ^*

$$\mu^* = \beta_h = f(x^h).$$

2) Puisque le processus de subdivision est complet, alors il existe une suite infinie (K_h^q) de cônes qui tend vers une demi-droite et $\mu(K_h^q) - \beta_h^q \rightarrow 0$ quand $q \rightarrow \infty$.

Par construction, on a:

$$f(x) \geq \mu(K_h^q), \forall x \in C.$$

Puisque f est continue et C compact, alors on déduit que:

la limite x^* de la sous suite (x_h^q) de la suite (x^h) appartient à C ,
 $f(x_h^q) = \beta_h^q \rightarrow f(x^*) = \beta^*$ et $f(x) \geq f(x^*) = \beta^* = \mu^*$.

3) Par construction.

11.4.3 Mise en oeuvre de l'algorithme.

Nous donnons maintenant quelques procédés de la subdivision de C qui sont complets.

a) Partition exhaustive

Théoriquement, il existe plusieurs possibilités pour définir des subdivisions exhaustives [18,47,55]. Nous rappelons ici, une procédure qui nous parait compatible avec la pratique dans le cas où C est convexe quelconque.

Soit $K_h(v^1, \dots, v^n)$ le cône de la subdivision à la $h^{\text{ème}}$ itération.

On définit le vecteur w par la relation

$$w = (v^r + v^s) / 2 \text{ avec } \|v^r - v^s\| = \max (\|v^i - v^j\|, i \neq j)$$

(i.e. le point milieu de la grande arête du simplexe $S[v^1, \dots, v^n]$)

On attache à ce vecteur les deux sous cônes suivants:

$$K_{h,1}(v^1, \dots, v^{s-1}, w, v^{s+1}, \dots, v^n) \text{ et } K_{h,2}(v^1, \dots, v^{r-1}, w, v^{r+1}, \dots, v^n).$$

Lemme 4.1 [18,47].

$K_h = K_{h,1} \cup K_{h,2}$ et le processus de subdivision est exhaustif.

b) Compatibilité de la minorante estimation.

On remarque que l'efficacité de cette classe d'algorithmes dépend de la procédure du calcul de la minorante estimation. Nous allons présenter deux procédures compatibles.

Considérons le cône $K_h(v^1, \dots, v^n)$ de sommet l'origine et tel que $K_h \cap C \neq \emptyset$. La question est de calculer $\mu(K_h) \leq \min \{f(x) / x \in K_h \cap C\}$ telle que cette minoration réalise une bonne estimation de la solution effective du problème $\min \{f(x) / x \in K_h \cap C\}$ et que le processus de calcul soit compatible.

Première procédure

Pour chaque $j=1, \dots, n$, on détermine l'élément

$$\pi(v^j) = \theta^j v^j \text{ avec } \theta^j = \max \{ \theta / \theta v^j \in \partial C / \theta \geq 0 \}. \quad (4.2)$$

On pose

$$\beta_{h+1} = \min \{ \beta_h, f(\pi(v^j)) \}, j=1, \dots, n \quad (\beta_h \text{ est donné}) \quad (4.3)$$

et pour tout $j=1, \dots, n$ on définit

$$y^j = \lambda^j v^j \text{ avec } \lambda^j = \max \{ \lambda / M \geq \lambda \geq 1, f(\lambda v^j) \geq \beta_{h+1} \}.$$

où M est une constante assez grande.

(i.e. On détermine un simplexe $S[0, y^1, \dots, y^n]$ où la valeur minimale de $f(x)$ est égale à β_{h+1})

On note par $H(y^1, \dots, y^n) = \{x \in \mathbb{R}^n / \langle a, x \rangle = 1\}$ l'hyperplan qui passe par les points $y^i (i=1, \dots, n)$. On détermine le point $z \in \text{CN}K_h$ qui est le plus éloigné de cet hyperplan $H(y^1, \dots, y^n)$.

Autrement dit [47], on résout le sous problème suivant:

$$\alpha_h = \max\{ \langle B^{-1}x, e_h \rangle / x \in \text{CN}K_h \} \quad (4.4)$$

où $e_h = (1, \dots, 1)$ et B est la matrice de vecteurs colonnes (y^1, \dots, y^n) .

On prend

$$\mu(K_h) = \begin{cases} \beta_{h+1} & \text{si } \alpha_h \leq 1 \\ \min \left(\frac{\sum_{i=1}^n \alpha_i f(z^i)}{\sum_{i=1}^n \alpha_i} - 1, \langle x - \sum_{i=1}^n \alpha_i z^i, e \rangle \right) & \text{si } \alpha_h > 1 \end{cases} \quad (4.5)$$

où $z^i = \alpha_h y^i (i=1, \dots, n)$, V est la matrice carrée de colonnes $z^i (i=1, \dots, n)$ et $\alpha = (\alpha_1, \dots, \alpha_n)$.

Lemme 4.2

La valeur de $\mu(K_h)$ décrite ci-dessus réalise une minorante estimation et le procédé de calcul est compatible.

Démonstration

Puisque l'hyperplan $H(y^1, \dots, y^n) = \{x \in \mathbb{R}^n / \langle a, x \rangle = 1\}$ passe par les points $y^i (i=1, \dots, n)$, alors :

$-H(y^1, \dots, y^n)$ peut s'écrire sous la forme $\{x \in \mathbb{R}^n / \langle B^{-1}x, e \rangle = 1\}$

où $e = (1, \dots, 1)$,

- $\langle B^{-1}x', e \rangle \leq 1$ si $x' \in S[0, y^1, \dots, y^n]$,

- $\langle B^{-1}x'', e \rangle > 1$ si $x'' \in \text{CN}K_h$ et $x'' \in S[0, y^1, \dots, y^n]$.

Par définition du sous problème (4.4), on a:

*Si $\text{CN}K_h \subset S[0, y^1, \dots, y^n]$, alors $\alpha_h \leq 1$.

La construction des points $y^i (i=1, \dots, n)$ et la concavité de f entraînent que $\mu(K_h) = \beta_{h+1}$ est une borne inférieure.

**Si $S[0, y^1, \dots, y^n] \not\subset \text{CN}K_h$, alors on a $\alpha_h > 1$. Par construction des points $z^i (i=1, \dots, n)$ le simplexe $S[0, z^1, \dots, z^n]$ contient $\text{CN}K_h$.

De la concavité de f on déduit que $\mu(K_h)$ défini en (4.5) est une minorante estimation.

Soit (K_h^q) une suite infinie décroissante de cônes telle que

$$\cap K_h^q = \{x = \theta x^* : x^* \in \partial C\}$$

Nous allons montrer que $f(x^*) = \beta^*$ et $\mu(K_h^q) - \beta_h^q \rightarrow 0$ quand $q \rightarrow \infty$

On note par $\pi(v_h^{q,i}), y_h^{q,i}, z_h^{q,i}$ et z_h^q les vecteurs $\mu(v^i), y^i, z^i$ et z définis comme ci-dessus pour le cône $K_h = K_h^q$.

Par construction et grâce aux hypothèses faites sur la suite (K_h^q) et C on a:

$$\mu(v_h^{q,i}) \rightarrow x^* \text{ pour } i=1, \dots, n \text{ et } z_h^q \rightarrow x^* \text{ quand } q \rightarrow \infty. \quad (1)$$

Si on note par y_h^q l'élément de l'intersection du segment $[0, z_h^q]$ avec le simplexe $S(y_h^{q,1}, \dots, y_h^{q,n})$, alors on a d'après (1)

$$y_h^q \rightarrow x^* \text{ quand } q \rightarrow \infty \quad (2)$$

La relation (2) donne

$$\alpha_h^q \rightarrow 1, y_h^{q,i} \text{ et } z_h^{q,i} \rightarrow x^* \text{ (} i=1, \dots, n \text{) quand } q \rightarrow \infty$$

Puisque

$$\min\{f(z_h^{q,i}), i=1, \dots, n\} \leq \mu(K_h^q) < \beta_h^q \leq f(y_h^q)$$

alors

$$\beta^* = f(x^*) \text{ et } \mu(K_h^q) - \beta_h^q \rightarrow 0 \text{ quand } q \rightarrow \infty.$$

Deuxième procédure [55]

Solent $\pi(v^j)$ ($j=1, \dots, n$) et β_{h+1} définis comme en (4.2) et (4.3) de la première procédure.

On prend

$$z = \pi(v^S) \text{ où } f(\pi(v^S)) = \min\{f(\pi(v^j)), j=1, \dots, n\}.$$

On définit l'ensemble \mathcal{J} d'indices par

$$\mathcal{J} = \{i / g_i(z) = 0, i=1, \dots, m\} (\neq \emptyset, \text{ car } z \in \partial C)$$

A tout $i \in \mathcal{J}$, on associe l'hyperplan

$$H_i(z) = \{x \in \mathbb{R}^n / \langle p^i, x-z \rangle \leq 0\}$$

où $p^i \in \partial g_i(z)$ (sous différentiel de g_i en z).

La convexité de g_i et l'hypothèse $g_i(0) < 0$ entraînent que :

$C \subset H_j(z)$ et $\langle p^j, z \rangle > 0$ pour tout $i \in J$.

Ceci nous conduit à penser à la minorante estimation suivante:

On note par

$$J = \{ i \in J / \langle p^j, v^j \rangle > 0, \forall j=1, \dots, n \}$$

-Si $J \neq \emptyset$, alors pour tout $i \in J$

on calcule

$$\alpha_i(K_h) = \min (\beta_{h+1}, f(z^1), \dots, f(z^n))$$

où $z^j = (\langle p^j, z \rangle / \langle p^j, v^j \rangle) v^j$ pour $j=1, \dots, n$

on prend

$$\mu(K_h) = \max (\alpha_i(K_h), i \in J).$$

-Dans le cas contraire, on prend

$$\mu(K_h) = \mu(K_{\text{préc}})$$

($\mu(K_{\text{préc}})$ la minorante estimation du cône précédent qui contient K_h)

Lemme 4.3

Le procédé de minorante estimation décrit en deuxième procédure est compatible.

Démonstration.

Analogue à celle du lemme 5 (page 506) de l'article [55].

Remarque sur le cas où C n'est pas borné [55]

On remarque que l'hypothèse de l'ensemble admissible C borné dans le deuxième procédure est nécessaire seulement pour la détermination des points $\pi(v^j)$ ($j=1, \dots, n$) (4.2). On peut supprimer cette hypothèse de la manière suivante:

Cas 1) Si pour tout j ($j=1, \dots, n$) on a $\theta^j \rightarrow \infty$, alors l'hypothèse de concavité de la fonction f entraîne que:

-Soit il existe un indice j tel que $f(\theta v^j) < f(0)$ pour tout $\theta > 0$ et dans ce cas f n'est pas bornée sur C et $\inf (f(x) / x \in C) \rightarrow -\infty$.

-Si pour tout j ($j=1, \dots, n$) on a $f(\theta v^j) > f(0)$ pour tout $\theta \geq 0$, alors on

peut prendre $\mu(k_h) = f(0)$.

Cas 2) S'il existe un indice j tel que θ^j est fini, alors on définit $z = \pi(v^S)$ où $f(\pi(v^S)) = \min \{ f(\pi(v^j)) / v^j < \infty, j = 1, \dots, n \}$ et on applique le reste de la deuxième procédure.

11.4.4 Variante de l'algorithme général dans le cas où C est polyédral

Nous supposons que C est un polyèdre convexe borné et défini par:

$$C = \{x \in \mathbb{R}^n / Ax \leq b, x \geq 0\}. \quad (4.6)$$

où A est une matrice de type (m,n).

Sans perte de généralité, on suppose que l'origine est un sommet non dégénéré. Nous utiliserons la caractérisation (3.2) pour diminuer le nombre d'itérations et en même temps on démontre que la convergence est en un nombre fini de pas.

Description de la variante

Initialisation On détermine les sommets (x^1, \dots, x^n) adjacents au sommet $x^0 = 0$. On pose $\beta_0 = f(x^*) = \min \{ f(0), f(x^i), i = 1, \dots, n \}$

$$\mathcal{K}_0 = \{K_0(x^1, \dots, x^n)\}.$$

Itération h: ($h = 0, 1, \dots$).

Etape -1-: On prend le cône polyédral $K_h(v^1, \dots, v^n) \in \mathcal{K}_h$, on calcule les points y^i ($i = 1, \dots, n$) qui se trouvent sur les n-arêtes du cône K_h tels que $f(y^i) \geq \beta_h$ ($i = 1, \dots, n$).

Etape -2-: On résout le sous problème suivant:

$$\alpha_h = \max \{ \langle e, Y^{-1}x \rangle / x \in C \cap K_h \} \quad (SP_h)$$

où Y est la matrice de vecteurs colonnes y^i ($i = 1, \dots, n$) et $e = (1, \dots, 1)$.

Etape -3-: Soit $x^{h+1} = \sum \lambda_j y^j$ la solution du problème (SP_h) et

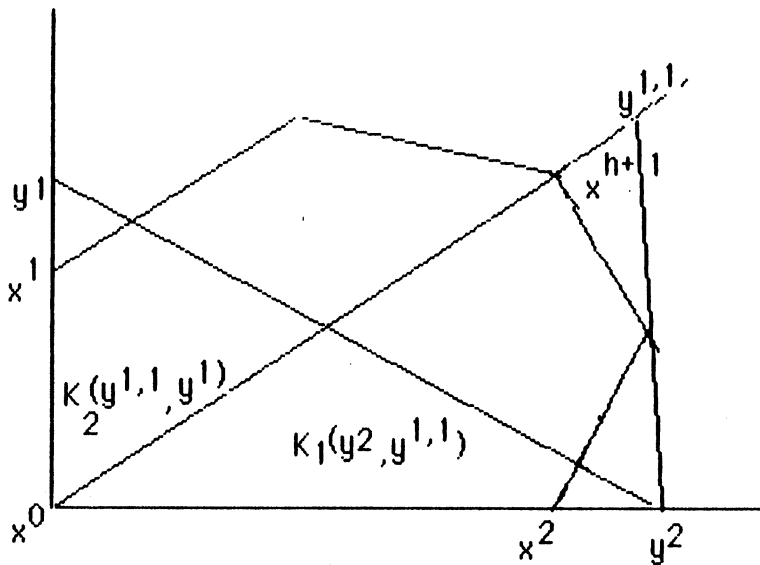
$$\beta_{h+1} = f(x^*) = \min \{ f(x^{h+1}), \beta_h \}.$$

Si $\alpha_h \leq 1$, alors on pose $\mathcal{X}_{h+1} = \mathcal{X}_h - K_h$. Sinon, on associe pour chaque $\lambda_j > 0$ le cône K_j défini par:

$K_j(v^1, \dots, v^{j-1}, x^{h+1}, v^{j+1}, \dots, v^n)$ et on remplace \mathcal{X}_h par $\mathcal{X}_{h+1} = \mathcal{X}_h \cup (K_j / \lambda_j > 0) - K_h$.

Etape -4-: Si $\mathcal{X}_{h+1} = \emptyset$ alors on change h par $h+1$ et on passe à l'itération h .

Dans le cas contraire on arrête l'algorithme et x^* est un minimum global.



Théorème 4.2

Cette variante converge en un nombre fini d'itérations vers une solution globale du problème (P).

Lemme 4.4

Soit $K(v^1, \dots, v^n)$ un cône polyédral du sommet l'origine et engendré par n -vecteurs linéairement indépendants v^1, \dots, v^n . Soit $w = \sum_{i \in I} \lambda_i v^i \in K$ où $I \neq \emptyset \subset \{1, \dots, n\}$ et $\lambda_i > 0, \forall i \in I$.

Alors $K(v^1, \dots, v^n) = \bigcup_{h \in I} K_h(v^1, \dots, v^{h-1}, w, v^{h+1}, \dots, v^n)$.

Démonstration:

Il est clair que $K_h \subset K$ pour tout $h \in I$ et qu'il suffit de montrer que pour tout $x \in K$ il existe un indice $h \in I$ tel que $x \in K_h$.

Soit $x \in K$, i.e. $x = \sum \alpha_i v^i = \sum_{i \in I'} \alpha_i v^i + \sum_{i \in I} \alpha_i v^i$ (I' est le complémentaire de I)

Soit $h \in I$ tel que

$$\alpha_h / \lambda_h = \min \{ \alpha_i / \lambda_i : i \in I \} \quad (1)$$

Il est facile de vérifier que x peut s'exprimer sous la forme suivante:

$$x = \sum_{i \in I'} \alpha_i v^i + \sum_{i \in I - \{h\}} (\alpha_i - (\alpha_h / \lambda_h) \lambda_i) v^i + \alpha_h / \lambda_h w.$$

D'après le choix de l'indice h (1), on a $x \in K_h$. \square

Lemme 4.5

Soit x^{h+1} solution du problème (SP_h) avec $\alpha_h > 1$. Pour tout sous cône $K_r \subset K_h$ ($r > h$) obtenu par la subdivision décrite dans la variante ci-dessus (II.4.4) on a la propriété suivante:
Si x^{h+1} est solution du problème (SP_r) alors $\alpha_r \leq 1$.

Démonstration

Par définition du problème (SP_h) on a

$\alpha_h > 1$ entraîne que x^{h+1} ne peut être un élément d'une arête du cône K_h .

D'après le procédé de subdivision on a:

$x^{h+1} \in K_r(v^1, \dots, v^n)$ si et seulement s'il existe un indice $j \in \{1, \dots, n\}$ tel que $v^j = x^{h+1}$.

D'où la vérification.

Démonstration du Théorème 4.2

Par définition de chaque sous problème (SP_h) (cf. II.4.4), on a:

Si $\alpha_h \leq 1$, alors $C \cap K_h \subset S[0, y^1, \dots, y^n]$.

D'autre part, La construction des points y^i ($i=1, \dots, n$) (étape 1) et la

concavité de f entraînent que:

$$f(x) \geq f(x^*) \quad \forall x \in K_h \cap C \quad (1)$$

D'après le choix du K_0 et le lemme 4.4 on a:

$$C \subset \bigcup_{h=0}^{\infty} U(K_j \subset K_0 / \alpha_j \leq 1, 0 < j < h) \quad \forall h \geq 0 \quad (2)$$

(1) et (2) montre que:

$$K_h = \emptyset \Rightarrow f(x) \geq f(x^*) \quad \forall x \in C.$$

La convergence est en un nombre fini d'itérations, car on a:

D'après le lemme 4.5, le nombre d'éléments de toute suite décroissante de cônes $(K_j) \subset K_0$ est inférieur strictement au nombre de sommets du polyèdre C (i.e. (K_j) fini).

D'autre part, comme la subdivision de cône à chaque itération est aussi finie par construction alors le nombre de suites décroissantes (K_j) est aussi fini.

II.5 Méthodes de résolution par approximation extérieure de l'ensemble admissible.

Nous supposons que l'ensemble admissible C est défini comme en (4.1). On notera par $E(D)$ l'ensemble des points extrémaux de tout polyèdre convexe non vide $D \subset \mathbb{R}^n$.

II.5.1 Description de l'algorithme

On initialise avec un polyèdre borné S^0 contenant C et un point w de l'intérieur de C .

Itération k : ($k=0, 1, \dots$)

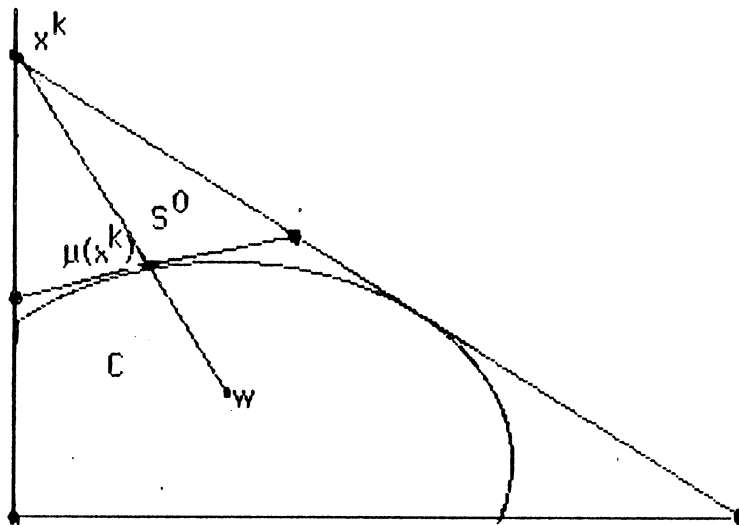
Etape 1: On calcule $f(x^k) = \max\{f(x) / x \in E(S^k)\}$. Si $x^k \in C$, alors on arrête et x^k est une solution globale du problème (P).

Sinon, on détermine le point $\mu(x^k) = \theta w + (1-\theta)x^k$ tels que $0 < \theta < 1$ et $h(\mu(x^k)) = \max\{g_i(\mu(x^k)) : i=1, \dots, m\} = 0$. On passe à l'étape 2.

Etape 2: On prend le sous différentiel e^k de h en $\mu(x^k)$ et on calcule l'ensemble $E(S^{k+1})$ (cf. II.5.2) du polyèdre

$$S^{k+1} = S^k \cap \{x \in \mathbb{R}^n / \langle e^k, x - \mu(x^k) \rangle \leq 0\}.$$

On pose $k=k+1$ et on retourne à l'étape 1.



II.5.2 Mise en oeuvre de l'algorithme

Le sous problème principal dans cet algorithme consiste à déterminer à chaque itération l'ensemble des sommets du polyèdre S^{k+1} . Par construction de cette suite de polyèdres, la résolution de chaque sous problème est basée sur le lemme suivant.

Lemme 5.1[4]

Soient $S \subset \mathbb{R}^n$ un polyèdre borné non vide et $H^- \subset \mathbb{R}^n$ un demi espace fermé tel que $S \cap H^+ \neq \emptyset$, où $H^+ = \mathbb{R}^n - H^-$.

Le point $v \in \mathbb{R}^n$ est sommet du polyèdre $S \cap H^-$ et non du polyèdre S si et seulement s'il est de la forme $v =]v^1, v^2[\cap \partial H^-$, où v^1 et v^2 sont deux sommets adjacents du polyèdre S tels que $v^1 \in H^+$, $v^2 \in H^- - \partial H^-$ et $]v^1, v^2[$ est l'arête joignant ces deux sommets.

En pratique, on détermine les éléments de l'ensemble $E(S^{k+1})$ par une succession d'opérations de pivotages[38] de la manière suivante:

On pose

$$S = S^k, H^- = \{x \in \mathbb{R}^n / \langle p^k, x - \mu(x^k) \rangle \leq 0\},$$

$$H^+ = \{x \in \mathbb{R}^n / \langle p^k, x - \mu(x^k) \rangle > 0\} \text{ et } \partial H^- = \{x \in \mathbb{R}^n / \langle p^k, x - \mu(x^k) \rangle = 0\}.$$

- On détermine l'ensemble $V = \{v \in E(S^k) / v \notin S^{k+1}\}$.

- Pour chaque $v \in V$, on calcule les éléments de l'ensemble $A(v)$ suivant:

$$A(v) = \{u / u =]v, y[\cap \partial H^-, y \in E(v)\}$$

où $E(v)$ est l'ensemble des sommets adjacents au sommet v séparés de celui-ci par l'hyperplan ∂H^- .

- On prend

$$E(S^{k+1}) = E(S^k) \cup (\cup_{v \in V} A(v)) - V.$$

II. 5.3 Convergence de l'algorithme

Théorème 5.1

Pour toute suite (x^k) générée par l'algorithme ci-dessus on a les propriétés suivantes :

- 1- Si l'algorithme s'arrête à l'itération k , alors x^k est un optimum global du problème (P).
- 2- Si les contraintes $g_i(x) \leq 0$ ($i=1, \dots, m$) sont linéaires alors l'algorithme converge en un nombre fini d'itérations.
- 3- Tout point d'adhérence de la suite (x^k) est solution du problème (P)

Démonstration

1) Par construction, on a:

$$C \subset \dots \subset S^{k+1} \subset S^k \subset \dots \subset S^0 \text{ pour tout } k > 0$$

Comme f est concave, alors pour tout $k \geq 0$, x^k réalise le minimum global de $f(x)$ dans S^k .

D'où, x^k est une solution globale si on s'arrête à l'itération k .

2) Si les fonctions g_i ($i=1, \dots, m$) sont linéaires alors chaque $k^{\text{ème}}$ itération consiste à ajouter une contrainte g_j violée en x^k .

Comme le nombre de contraintes est fini alors le nombre d'itérations est inférieur ou égal à m .

3) D'après (1), il suffit de montrer que tout point d'adhérence de la suite (x^k) appartient à C .

Comme les suites (x^k) , $(\mu(x^k))$ et (ρ^k) sont bornées alors elles possèdent toujours des points d'adhérences.

Soit $a^k(x) = \langle \rho^k, x - \mu(x^k) \rangle$.

Par construction, on a les inégalités suivantes:

$$a^k(x^k) > 0 \text{ pour tout } k \geq 0 \text{ et } a^k(x^j) \leq 0 \text{ pour tout } j > k \quad (1)$$

Supposons qu'il existe un point limite x^* de la suite (x^k) tel que $x^* \in C$.

Autrement, il existe une sous suite (k_l) d'indices telle que:

$$x^{k_l} \rightarrow x^*, a^{k_l}(x^{k_l}) \rightarrow a(x^*) > 0 \text{ et } a^{k_l}(x^{k_j}) \rightarrow a(x^*) \text{ quand } l, j \rightarrow \infty$$

Or, par passage à la limite dans (1) on a:

$$a(x^*) < 0 \text{ ce qui contredit } a(x^*) > 0. \quad \square$$

II.6 Essais numériques

Ces méthodes ont été mises au point sur Micral 90.50 en langage Pascal pour la résolution du problème (P) dans le cas où C est un polyèdre convexe borné. A partir d'une série d'exemples testés, nous présentons les résultats expérimentaux de trois exemples qui englobent les principales remarques sur la mise en oeuvre de ces algorithmes.

Nous avons pris, le problème de maximisation d'une forme quadratique symétrique définie positive sur un polytope.

$$\text{Max } (f(x) = \langle x, C x \rangle + \langle d, x \rangle + \alpha / Ax \leq b, x \geq 0) \quad (P)$$

Exemple 1

$$C = \begin{pmatrix} 0.75 & -2.25 \\ -2.25 & 8.75 \end{pmatrix} \quad A = \begin{pmatrix} -1 & 3 \\ 1.0 & -1.0 \\ -1.0 & 7.0 \\ 1.0 & -5.0 \end{pmatrix}$$

$$d = [9.0, -27.0], b = [6.0, 14.0, 22.0, 2], \alpha = 27.0$$

Le maximum global est $x^* = (17.0, 3.0)$ et $f(x^*) = 165.0$

Exemple 2 [57]

Nous prenons le contre-exemple de Zwart [57] pour la première méthode de H. Tuy [49].

$$C = \begin{pmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{pmatrix} \quad A = \begin{pmatrix} 1.0 & 1.0 & -1.0 \\ -1.0 & 1.0 & -1.0 \\ 12.0 & 5.0 & 12.0 \\ 12.0 & 12.0 & 7.0 \\ -6.0 & 1.0 & 1.0 \end{pmatrix}$$

$$d = [-2.0, 0.0, -2.0], b = [1.0, -1.0, 34.8, 29.1, -4.1], \alpha = 2.$$

Le maximum global est $x^* = [1.0, 0.0, 0.0]$ et $f(x^*) = 1.0$

Exemple 3 [28,29]

Calcul du maximum d'une forme quadratique définie positive sur la boule unité de la norme φ_∞ .

$$\text{Max } \langle y, C y \rangle / -1 \leq y \leq 1.$$

On fait le changement de variable $x=y+e$ où $e=(1, \dots, 1)$, on obtient

$$\text{Max } \langle x, Cx \rangle - 2\langle Ce, x \rangle + \langle e, Ce \rangle / 0 \leq x \leq 2$$

$$C = \begin{bmatrix} 22.0 & 1.0 & -7.0 & 6.0 \\ 1.0 & 33.0 & 4.0 & -7.0 \\ -7.0 & 4.0 & 27.0 & -1.0 \\ 6.0 & -7.0 & -1.0 & 38.0 \end{bmatrix} \quad A = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$d = [-44.0, -62.0, -46.0, -72.0], b = [2.0, 2.0, 2.0, 2.0], \alpha = 112.0$$

Les solutions globales sont:

$$x^*_1 = [2.0, 0.0, 0.0, 2.0], x^*_2 = [0.0, 2.0, 2.0, 0.0] \text{ et } f(x^*_i) = 168.0 \text{ (} i=1,2 \text{)}$$

Et l'on a les tableaux de résultats :

Algorithme de coupe (II.3.2)

exemple	(x^0)	$f(x^0)$	le nombre d'itérations	le temps(sec)
1	[0.0, 0.0]	27	2	5.52
	[2.0, 0.0]	48	3	9.5
	[17.0, 3.0]	165	1	5.16
2	[1.0, 0.0, 0.0]	1.0	1	4.02
	[1.9, 0.0, 0.9]	0.82	3	3.84
	[1.9, 0.9, 0.9]	0.82	2	2.52
3	[0.0, 0.0, 0.0, 0.0]	112.0	7	9.74
	[2.0, 0.0, 0.0, 2.0]	168.0	6	9.64
	[2.0, 2.0, 0.0, 2.0]	128.0	5	7.66

-Tableau I -

Algorithme de résolution par partition (II.4.2)

Première procédure

exemple	le cône initial	le nombre de cônes	le temps(sec)
1	[2.0 ,0.0] [0.0 ,2.0]	19	22.64
2	[0.9 ,0.0,-0.27] [0.0, 0.9, 00] [0.9,0.9,0.27]	9	7.38
3	[2.0 ,0.0 ,0.0 ,0.0] [0.0 ,2.0 ,0.0 ,0.0] [0.0 ,0.0 ,2.0 ,0.0] [0.0 ,0.0 ,0.0 ,2.0]	235	344.76

-Tableau 2 -

Deuième Procédure

exemple	le cône initial	le nombre de cônes	le temps(sec)
1	[2.0 ,0.0] [0.0 ,2.0]	35	1.48
2	[0.9 ,0.0,-0.27] [0.0, 0.9, 00] [0.9,0.9,0.27]	63	5.12
3	[2.0 ,0.0 ,0.0 ,0.0] [0.0 ,2.0 ,0.0 ,0.0] [0.0 ,0.0 ,2.0 ,0.0] [0.0 ,0.0 ,0.0 ,2.0]	1131	172.66

-Tableau 3 -

Variante de l'algorithme général de résolution par partition. (II.4.4)

exemple	le cône initial	le nombre de cônes	le temps(sec)
1	[2.0 ,0.0] [0.0 ,2.0]	5	0.7
2	[0.9 ,0.0,-0.27] [0.0, 0.9, 00] [0.9,0.9,0.27]	3	0.9
3	[2.0 ,0.0 ,0.0 ,0.0] [0.0 ,2.0 ,0.0 ,0.0] [0.0 ,0.0 ,2.0 ,0.0] [0.0 ,0.0 ,0.0 ,2.0]	41	14.68

-Tableau 4 -

Algorithme de résolution par approximation extérieure

exemple	(x ⁰)	f(x ⁰)	nobre d'itérations	E(S ^k)	temps(sec)
1	[0.0 ,0.0]	27	4	6	1.52
	[17.0 ,3.0]	165	2	4	1.18
	[20.0 ,6.0]	120	3	5	1.34
2	[0.72,0.0, 0.27]	0.60	3	8	1.24
	[1.9 ,0.0 , 0.9]	0.82	3	8	1.78
	[1.0 ,0.9 ,0.9]	0.82	5	7	2.42
3	[0.0 ,0.0 ,0.0 ,0.0]	112	5	16	4.12
	[2.0 ,0.0 ,0.0 ,2.0]	168	5	16	4.02
	[2.0 ,2.0 ,2.0 ,2.0]	112	5	16	4.12

-Tableau 5 -

Comparaison

Ces algorithmes ont été testés sur plusieurs exemples. La comparaison entre eux est basée sur deux critères principaux: l'occupation de mémoire nécessaire et le temps de calcul.

On constate que :

- La recherche à chaque itération d'un nouveau sommet du polyèdre de départ dans l'algorithme de coupe (pour assurer la convergence en un nombre fini de pas) augmente le temps de calcul par rapport à la variante II.4.4 et l'algorithme d'approximation extérieure II.5.1. Il dépend légèrement du vecteur de départ (voir les tableaux 1,4 et 5). Par contre la taille de la mémoire nécessaire est moindre que celle des méthodes de résolution par partition de l'ensemble admissible.
- L'inconvénient principal des algorithmes de résolution par les subdivisions est la taille de la place mémoire. Signalons que l'occupation de mémoire se fait à l'exécution, lors des subdivisions. Puisqu'on ne peut pas définir la taille nécessaire au départ, nous avons utilisé les pointeurs qui permettent de créer lors de la subdivision la mémoire nécessaire de chaque cône par l'instruction "new", et détruire les mémoires de cônes $K \in \mathcal{K}_h$ où $\mu(K) \geq \beta_h$ ($h > 0$) par l'instruction "dispose". Le temps de calcul de l'algorithme général avec les deux procédures est très grand; ceci peut être expliqué par le choix de la subdivision d'après la variante II.4.4 et le calcul de la minorante estimation utilisé (la question reste à l'étude). La durée de calcul pour la variante II.4.4 est satisfaisante (Tableau 4).
- L'algorithme d'approximation extérieure est plus simple à mettre en oeuvre que les autres algorithmes; on utilise seulement des opérations de pivotages. Le temps de calcul et la taille de mémoire augmentent très rapidement avec la taille des problèmes, les vecteurs de départs, le choix du simplexe (ou polytope) initial qui contient le domaine (Tableau 5).

Conclusion

- Nous avons étudié trois types différents de méthodes : méthodes de coupe du type Hoang Tuy , méthodes de résolution par partition de l'ensemble admissible et méthodes de type approximation extérieure .

- Nous avons mis en oeuvre sur un micro-ordinateur les algorithmes de synthèse décrits dans les paragraphes précédents .

- On peut dire que la variante de l'algorithme général de résolution par partition II.4.4 et l'algorithme de type approximation extérieure II.5.1 donnent des résultats satisfaisants pour la résolution des problèmes de petite taille de maximisation d'une fonction convexe (ou minimisation d'une fonction concave) sous des contraintes linéaires.

- Nous pensons que pour les problèmes de grande taille peuvent être traités de la manière suivante:

On subdivise le domaine admissible en plusieurs sous domaines , on applique simultanément dans chaque partie l'algorithme d'approximation extérieure ou de la variante II.4.4. Le point admissible qui réalise la meilleure borne par rapport aux bornes de ces sous domaines est évidemment une solution globale.

CHAPITRE III

**Minimisation globale d'une fonction
linéaire sur une partie compacte convexe
et une contrainte anticonvexe**



III.1 Introduction

Dans ce chapitre, on s'intéresse aux solutions globales des programmes mathématiques non convexes du type:

$$\min(f(x) = \langle c, x \rangle / x \in C) \quad (P)$$

où l'ensemble des contraintes est représenté sous la forme suivante:

$$C = \{ x \in \mathbb{R}^n / x \in D, g(x) \geq 0 \}$$

avec $g: \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe sur \mathbb{R}^n et D est une partie convexe fermée non vide.

Notre but dans ce travail est d'étudier le problème (P) dans les deux directions suivantes:

- Donner une classe générale de problèmes qui peuvent être mis sous la forme du problème (P) et ses applications.

- Adaptation des méthodes d'optimisation convexe et non convexe du chapitre II à la résolution de cette classe de problèmes.

Dans le paragraphe III.2, on montre que le problème de la minimisation de la différence de deux fonctions convexes sous des contraintes qui s'écrivent comme différence de fonctions convexes peut être ramené à la forme du problème (P). On explicite un problème concret d'application (loi de mélange) qui permet de justifier l'intérêt pratique de cette étude.

Le paragraphe III.3 est consacré aux méthodes de résolution dans le cas où D est polyédral. On dispose des caractérisations intéressantes de ses solutions globales, qui donnent des algorithmes simples de mise en oeuvre (en général) et finis (convergence en un nombre fini de pas).

Au paragraphe III.4, nous étudions le cas où D est convexe quelconque. Nous nous inspirons des méthodes d'accumulation des contraintes (J.Kelley [20], D.M.Topkis [50], ...) pour ramener la résolution du problème général (P) à celle d'une suite de sous-problèmes du paragraphe III.3. Nous adaptons l'algorithme d'approximation extérieure du chapitre II (cf. II.5.1) au problème (P).

On présente au dernier paragraphe quelques essais numériques.

Notations et définitions

Nous donnons quelques notations et définitions qui seront utilisées dans la suite.

Dans l'écriture de l'ensemble des contraintes on trouvera les notations suivantes :

- $S = \{ x \in \mathbb{R}^n / Ax = b, x \geq 0 \}$ est un polyèdre convexe non vide et A est une (m,n) -matrice et b un m -vecteur.

- $G = \{ x \in \mathbb{R}^n / g(x) \geq 0 \}$, $G^c = \{ x \in \mathbb{R}^n / g(x) < 0 \}$, $H = \{ x \in \mathbb{R}^n / h(x) \leq 0 \}$ et $C = S \cap G \cap H$.

- pour tout polyèdre convexe non vide S , on note l'ensemble des sommets et l'ensemble des arêtes de S par $E(S)$ et $A(S)$ respectivement.

- On note l'enveloppe convexe et la frontière de C par $co(C)$ et $\partial(C)$ respectivement pour tout ensemble non vide $C \subset \mathbb{R}^n$.

Définitions 2.1

Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est dite d.c s'il existe deux fonctions convexes f_1 et $f_2: \mathbb{R}^n \rightarrow \mathbb{R}$ telles que $f(x) = f_1(x) - f_2(x)$ pour tout $x \in \mathbb{R}^n$.

Définition 2.2

Une contrainte du type $g(x) \geq 0$ où $g: \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe, est dite contrainte anticonvexe (ou complémentirement convexe).

Définition 2.3

Soit le problème d'optimisation suivant:

Minimiser $f(x) / f_1(x) \leq 0 (i=1, \dots, r), g_1(x) = 0 (i=1, \dots, l)$

où les fonctions f, f_1 et g_1 sont continues sur \mathbb{R}^n et à valeurs réelles.

On dit que ce problème est du type programmation mathématique convexe lorsque les fonction f et $f_1 (i=1, \dots, r)$ sont convexes et $g_1 (i=1, \dots, l)$ sont affines. Dans le cas contraire on parle du type programmation mathématique non convexe.

Dans une intention d'éviter les complications de notations, on note par (Q) le problème d'optimisation non convexe suivant:

$$\text{Min}\{f_0(x) / x \in R \cap S\} \quad (Q)$$

où $R = \{x \in \mathbb{R}^n / f_i(x) \leq 0, i=1, \dots, k\}$ et les fonctions f_0 et f_i ($i=1, \dots, k$) sont continues et non linéaires.

Définition 2.4

Un programme mathématique d.c est un problème d'optimisation non convexe qui peut s'exprimer sous la forme du problème (Q) et dans lequel chaque fonction f_i ($i=0, \dots, k$) est décomposable en une différence de deux fonctions convexes.

III.2 Réduction de quelques programmes mathématiques non convexes et applications.

III.2.1 Programmation mathématique d.c.

Fonction d.c.

La classe des fonctions représentables comme différence de deux fonctions convexes définies sur un compact Ω de \mathbb{R}^n a attiré une attention croissante ces dernières années. Un effort considérable a été fait pour les propriétés de ces fonctions et les problèmes d'applications qui se ramènent à résoudre des programmes mathématiques non convexes d.c. Nous rappelons ici quelques exemples et propriétés des fonctions d.c. On pourra consulter sur ce sujet le récent travail dans [1, 12, 17, 52].

Soit Ω un ensemble compact convexe non vide de \mathbb{R}^n .

- Toute fonction convexe (resp concave) définie sur Ω est d.c.
- Toute fonction C^2 sur Ω est d.c. sur Ω . En particulier toute forme quadratique est d.c sur \mathbb{R}^n .
- Toute fonction sous- C^2 sur Ω est d.c sur Ω .
(i.e. toute fonction qui est localement l'enveloppe supérieure sur Ω d'une famille de fonctions de classe C^2) [12].

- Si l'espace $C(\Omega)$ des fonctions continues est muni de la topologie de la convergence uniforme, alors l'espace $DC(\Omega)$ des fonctions d.c sur Ω est dense dans $C(\Omega)$ (i.e. $\overline{DC(\Omega)} = C(\Omega)$) (cf. [12]).

Proposition 2.1

Soient f_1, \dots, f_p un nombre fini de fonctions d.c sur Ω , admettant chacune une décomposition de la forme $f_i(x) = g_i(x) - h_i(x)$ ($i = 1, \dots, p$).

Alors $\max_{1 \leq i \leq p} f_i(x)$ et $\min_{1 \leq i \leq p} f_i(x)$ sont d.c sur Ω .

Démonstration

On sait que

$$\min_{1 \leq i \leq p} a_i = \sum_{i=1}^p a_i - \max_{1 \leq i \leq p} \sum_{j \neq i}^p a_j$$

donc on peut écrire

$$\min_{1 \leq i \leq p} f_i = \sum_{i=1}^p g_i - \max_{1 \leq i \leq p} (h_i + \sum_{j \neq i}^p g_j)$$

et

$$\max_{1 \leq i \leq p} f_i = \max_{1 \leq i \leq p} (g_i + \sum_{j \neq i}^p h_j) - \sum_{i=1}^p h_i \quad \square$$

Proposition 2.2

Tout programme mathématique d.c peut se ramener à la forme du problème (P).

Démonstration

Par introduction d'une coordonnée supplémentaire x_{n+1} , le problème (Q) peut s'écrire sous la forme suivante:

$$\min x_{n+1}$$

sous les contraintes :

$$(x, x_{n+1}) \in [Sx[\alpha_1, \alpha_2] \quad g_i(x) - h_i(x) \leq 0 \quad (i=1, \dots, k), \quad g_0(x) - h_0(x) - x_{n+1} \leq 0]$$

où

$$\alpha_2 \geq \inf \{ f_0(x) / x \in \text{SNR} \} \geq \alpha_1$$

Posons

$$F(y) = F(x, x_{n+1}) = \max \{ g_i(x) - h_i(x) \quad (i=1, \dots, k), \quad g_0(x) - h_0(x) - x_{n+1} \}$$

De la proposition 2.1 on a :

$$F(y) = p(y) - q(y),$$

où p et q sont convexes continues et définies par:

$$p(y) = \max_{1 \leq i \leq k} (g_0(x) + \sum_{j=1}^k h_j - x_{n+1}, g_i(x) + \sum_{j=1}^k h_j) \text{ et } q(y) = \sum_{j=0}^k h_j$$

D'où, le problème (Q) peut s'écrire sous la forme du problème (P).

$$\min(x_{n+1} / (x, x_{n+1}, t) \in Sx[\alpha_1, \alpha_2] \times [\beta_1, \beta_2], p(y) - t \leq 0, q(y) - t \geq 0) \text{ (Q')}$$

où

$$\beta_1 \leq \inf(p(y) / y \in Sx[\alpha_1, \alpha_2]) \text{ et } \beta_2 \leq \sup(q(y) / y \in Sx[\alpha_1, \alpha_2]) \square$$

Les problèmes (Q) et (Q') sont liés de la manière suivante.

Proposition 2.3

Si x^* est solution du problème (Q) alors $(x^*, x_{n+1}^* = f_0(x^*), t^* = p(x^*, x_{n+1}^*) = q(x^*, x_{n+1}^*))$ est solution du (Q').
Inversement, si (x^*, x_{n+1}^*, t^*) est solution du problème (Q') alors x^* est solution du problème (Q) et on a $f_0(x^*) = x_{n+1}^*$.

Démonstration

Conséquence immédiate de la transformation. \square

Remarque 2.1

La transformation du problème (Q) défini dans l'espace \mathbb{R}^n à un problème (Q') dans l'espace \mathbb{R}^{n+2} donnée ci-dessus est très générale. Mais, il est évident que la meilleure transformation de chaque problème de programmation mathématique d.c est relative à la décomposition de fonctions d.c et aux hypothèses. Comme exemple, considérons le problème de *minimisation d'une fonction convexe sous les contraintes d'égalité* :

$$\min(f_0(x) / x \in S, f_i(x) = 0 (i=1, \dots, l)) \text{ (Q)}$$

avec les hypothèses suivantes :

les fonctions $f_i (i= 1, \dots, l)$ sont convexes sur \mathbb{R}^n et la fonction f_0 est linéaire.

C'est -à-dire que le problème (Q) est du type :

$$\min\{f_0(x) = \langle c, x \rangle / x \in S, f_i(x) \leq 0, f_i(x) \geq 0 (i=1, \dots, l)\} \quad (Q)$$

D'après la proposition 2.1, le problème (Q) peut se ramener à un problème équivalent (Q') dans l'espace \mathbb{R}^n défini par:

$$\min\{f_0(x) / x \in S, \sum_{i=1}^l f_i(x) \geq 0, \max_{1 \leq i \leq l} f_i(x) \leq 0\} \quad (Q')$$

III.2.2 Exemple d'application (ELF)

Rappelons tout d'abord brièvement la position du problème pour constater ensuite qu'il se ramène naturellement à un problème du type (P).

a) Position du problème (Loi de mélange)

Le problème consiste à déterminer à prix minimum d'un mélange de plusieurs bases conforme aux spécifications données à partir de caractéristiques connues de ces bases.

C'est-à-dire à trouver une solution $x^* = (x_1^*, \dots, x_n^*)$ du problème suivant:

$$\min \sum_{j=1}^n c_j x_j$$

Sous les contraintes:

implicites : $\sum x_j = 1, x_j \geq 0, j=1, \dots, n$

dues aux spécifications: $f_j(x) \leq \alpha_j, j=1, \dots, l$

$f_l(x) = \alpha_l, l=1, \dots, r$

de quantités : $\langle Q_k, x \rangle \leq q_k$

où n =nombre de bases,

x_j =fraction volumique de la base j ($x_j = \text{vol } j / \text{vol } t$) et

c_j =coût de la base j .

La première démarche de résolution est par la programmation linéaire. Elle consiste à supposer que chaque fonction f_j est linéaire et définie par:

$$f_j(x) = \sum_{i=1}^n a_{j,i} x_i \quad (j=1, \dots, l+s)$$

où $a_{j,i}$ désigne la $i^{\text{ème}}$ caractéristique de la base i et α_j désigne la $j^{\text{ème}}$ spécification.

Mais, en réalité, l'approche linéaire de certaines contraintes de spécifications est fautive. Par exemple la loi de masse volumique de chaque spécification dans un mélange "du pool carburant" est donnée par la formule suivante:

$$mv_{\text{spé}} = \sum mv_i + \sum \beta_{i,j} x_i x_j$$

où mv_i est la masse volumique de la base i et $\beta_{i,j}$ est le coefficient d'interaction quadratique entre les bases i et j .

Par conséquent, l'introduction de contraintes quadratiques (ou non linéaires) représentent mieux la réalité de ce problème. Ce modèle est appelé modèle quadratique ou modèle de Morris.

b) Réduction du modèle quadratique (Modèle de Morris)

Le modèle quadratique se formule de manière générale sous la forme suivante:

$$\begin{aligned} \min \langle c, x \rangle \\ \text{sous les contraintes:} \end{aligned} \quad (\text{PA})$$

$$\begin{aligned} x \in \text{SCR}^n \\ m_j \leq f_j(x) = \langle x, A_j x \rangle + \langle x, b_j \rangle \leq M_j \quad (j=1, \dots, m) \end{aligned}$$

où les matrices A_j ($j=1, \dots, m$) sont symétriques et S est un polytope convexe non vide.

On remarque tout d'abord que (PA) peut s'écrire aussi sous la forme :

$$\begin{aligned} & \min \langle c, x \rangle \\ & x \in S \\ & f_j(x) = \langle x, A_j + \alpha U x \rangle + \langle x, b_j \rangle - \alpha \langle x, Ux \rangle \leq M_j \\ & f_j^*(x) = \langle x, -A_j x + \alpha Ux \rangle - \langle x, b_j \rangle - \alpha \langle x, Ux \rangle \leq -m_j \\ & j = 1, \dots, m, \alpha \in \mathbb{R} \text{ et } U \text{ est la matrice unité} \end{aligned}$$

Si on fait l'hypothèse suivant sur α :

$$\alpha \geq \max_{j=1, \dots, m} (\max_{i=1, \dots, n} |\lambda_i(A_j)|)$$

où $\lambda_i(A_j)$ ($i=1, \dots, n$) sont les valeurs propres de la matrice symétrique A_j .

Alors on aura:

$$1\text{- Les fonctions } f(x) = \sup_{i=1}^m \langle x, A_j + \alpha U x \rangle + \langle x, b_j \rangle - \alpha \langle x, Ux \rangle - M_j \quad \text{et}$$

$$g(x) = \alpha \langle x, Ux \rangle \text{ sont convexes continues sur } \mathbb{R}^n \text{ (car les matrices } A_j + \alpha U, -A_j + \alpha U \text{ et } \alpha U \text{ sont symétriques définies positives).}$$

2- Le problème (PA) peut s'exprimer analytiquement sous la forme:

$$\begin{aligned} & \min \langle c, x \rangle \\ & x \in S \\ & f(x) - g(x) \leq 0 \end{aligned}$$

Donc, par l'introduction d'une variable supplémentaire t , on transforme (PA) en un problème équivalent (PA)* du type (P).

$$\begin{aligned} & \min \langle c, x \rangle \\ & (x, t) \in S \times [\beta_1, \beta_2] \\ & f(x) - t \leq 0, g(x) - t \geq 0 \end{aligned} \quad \text{(PA)*}$$

où $\beta_1 = \min (m_j - M_j, j=1, \dots, m)$ et β_2 est une constante plus grande que la borne supérieure de $g(x)$ dans l'ensemble admissible du problème (PA).

Les solutions des problèmes (PA) et (PA)* sont liées de la manière suivante:

Proposition 2.4

Si (x^*, t^*) est une solution du problème $(PA)^*$, alors x^* est une solution du problème (PA) . Inversement, si x^* est une solution du problème (PA) alors pour tout réel t^* tel que $f(x^*) \leq t^* \leq g(x^*)$, (x^*, t^*) est solution du problème $(PA)^*$.

Démonstration

D'après la transformation. \square

Remarques

1-Ce modèle quadratique a été étudié par B.Chague (ELF CRES). Il a proposé des algorithmes de résolution par programmation linéaire pas à pas, ses méthodes sont de type approximation linéaire des contraintes quadratiques qui restreint le domaine d'intérêt (car les fonctions f_j sont non convexes) et d'autres algorithmes sont dus à la nature du problème dans la pratique.

2-La réalisation d'un logiciel pour la résolution de ce problème par les méthodes dont on dispose et les méthodes de B.Chague est en cours par deux étudiants de l'ENSIMAG (Projet de 3^{ème} année ENSIMAG).

III.3 Minimisation d'une fonction linéaire sur un polyèdre convexe borné et une contrainte anticonvexe.

Dans ce paragraphe nous étudions le problème suivant:

$$\text{Min}\{f(x)=\langle c,x \rangle / x \in S \cap G\} \quad (\text{SP})$$

avec les hypothèses que S est un polyèdre convexe borné et que les ensembles G et G^c sont non vides. Le lecteur pourra consulter aussi [5,6,15,16].

III.3.1 Caractérisation des solutions globales

Théorème 3.1 [15]

L'enveloppe convexe de $S \cap G$ est un polyèdre borné et on a

$$\text{co}(S \cap G) = \text{co}(A(S) \cap G) = \text{co}(\cup_{i \in I} (A_i(S) \cap G)),$$

où $(A_i(S), i \in I)$ est l'ensemble des arêtes du polyèdre S dont l'intersection avec G est non vide.

Théorème 3.2

S'il existe un point $x' \in S \cap G^c$ tel que $f(x') < \min\{f(x) / x \in S \cap G\}$, alors $\min\{f(x) / x \in S \cap G\} = \min\{f(x) / x \in S \cap \partial G\} = \min\{f(x) / x \in A(S) \cap \partial G\}$.

Démonstration

Soit x^* un minimum global du problème (SP). Supposons que $x^* \notin \partial G$, c'est-à-dire $g(x^*) > 0$.

Puisque $g(x') < 0$, alors d'après la convexité de la fonction g et l'ensemble S il existe un élément

$$w = \lambda x^* + (1-\lambda)x' \text{ avec } 0 < \lambda < 1 \text{ et } y \in S \cap \partial G.$$

Comme la fonction $f(x)=cx$ est linéaire alors on a:

$$f(w) = \lambda f(x^*) + (1-\lambda)f(x') < f(x^*) \quad (\text{d'après l'hypothèse sur } x')$$

ce qui contredit le fait que x^* est un minimum global. D'où la première égalité

$$\min\{f(x) / x \in S \cap G\} = \min\{f(x) / x \in S \cap \partial G\} \quad (1)$$

Du théorème 1, on déduit:

$$E(\text{co}(S \cap G)) = E(\text{co}(A(S) \cap G)) \subset A(S) \cap G \subset S \cap G.$$

Puisque la fonction objectif est linéaire alors on a:

$$\min\{f(x) / x \in S \cap G\} \leq \min\{f(x) / x \in \text{co}(S \cap G)\} = \min\{f(x) / x \in E(\text{co}(A(S) \cap G))\} \leq \min\{f(x) / x \in S \cap G\} \quad (2)$$

De (1) et (2) on déduit:

$$\min\{f(x) / x \in S \cap G\} = \min\{f(x) / x \in A(S) \cap \partial G\}. \quad \square$$

Théorème 3.3

Si x^* est une solution du problème $\min\{f(x) / x \in E(S) \cap G\}$, alors $\min\{f(x) / x \in S \cap G\} = \min\{f(x) / x \in \mathcal{P}(x^*)\}$, où

$$\mathcal{P}(x^*) = \bigcup_{y \in E(S'')} (A_y(S') \cap \partial G)$$

avec $S' = \{x \in S / f(x) \leq f(x^*)\}$, $S'' = \{x \in S' / f(x) = f(x^*)\}$ et $A_y(S')$ est l'ensemble des arêtes du polyèdre S' dont le sommet y est un élément d'extrémité.

Démonstration

Soit x^* un optimum global du problème (SP). En particulier, d'après le théorème 3.2 on suppose que x^* est un élément de l'arête $[x_1, x_2]$.

Par définition de x^* , on a:

$$f(x_1) < f(x^*) \leq f(x') \leq f(x_2), g(x_1) < 0, g(x_2) > 0 \text{ et } g(x^*) = 0 \quad (1)$$

Si $f(x') = f(x_2)$, alors $x_2 \in E(S'')$

et ceci entraîne que

$$[x_1, x_2] \in A_{x_2}(S') \text{ et } x^* \in \mathcal{P}(x^*).$$

Dans le cas contraire ($f(x') \neq f(x_2)$), il existe d'après (1) un élément

$y \in [x_1, x_2]$ tel que $f(x') = f(y)$ et dans ce cas on a deux cas possibles.

1- si $f(x') = f(y) = f(x^*)$, alors $x^* \in E(S) \cap \partial G \subset \mathcal{P}(x')$.

2- Sinon, $f(x_1) < f(x^*) < f(x') = f(y)$.

D'après la convexité de g et S , il existe $\lambda ; \lambda \in]0, 1[$ tel que:

$$x^* = \lambda y + (1-\lambda)x_1 \text{ et } g(y) > 0.$$

ce qui donne $x^* \in [x_1, x_2] \cap \partial G \subset \mathcal{P}(x')$. \square

III.3.2 Algorithme de coupe du type H.Tuy

On généralise l'algorithme général du paragraphe (II.3) pour la résolution du problème (SP).

III.3.2.1) Description de l'algorithme:

Initialisation

On détermine une solution x^1 du problème $\min \{f(x) / x \in S\}$.

Si $x^1 \in G$, alors x^1 est aussi un optimum global du (SP). Sinon, on calcule $y^1 \in E(S) \cap G$ (cf. III.3.2.2), on pose $S^1 = S$, $\alpha_1 = f(y^1)$ et $k=1$.

Itération k: (k=1, 2, ...)

Etape 1: - Si x^k est un sommet dégénéré dans S^k , alors on utilise l'idée de Balas (cf. [3]) qui consiste à construire un polyèdre S^k qui contient S^k tel que x^k devient un sommet non dégénéré dans S^k .

- Soit $\{x^{k,1}, \dots, x^{k,n-m}\}$ l'ensemble des sommets adjacents au sommet x^k dans S^k ou S^k .

Pour $j = 1, \dots, n-m$ on définit

$$\begin{aligned} z^{k,j} &= x^{k,j}, & \text{si } g(x^{k,j}) \leq 0 \\ z^{k,j} &= x^k + \rho_j (x^{k,j} - x^k), & \text{si } g(x^{k,j}) > 0, \text{ où } g(z^{k,j}) = 0. \end{aligned}$$

Etape 2: On détermine une coupe $\langle a_k, x \rangle \geq \beta_k$ qui sépare le sommet x^k de l'ensemble admissible, et

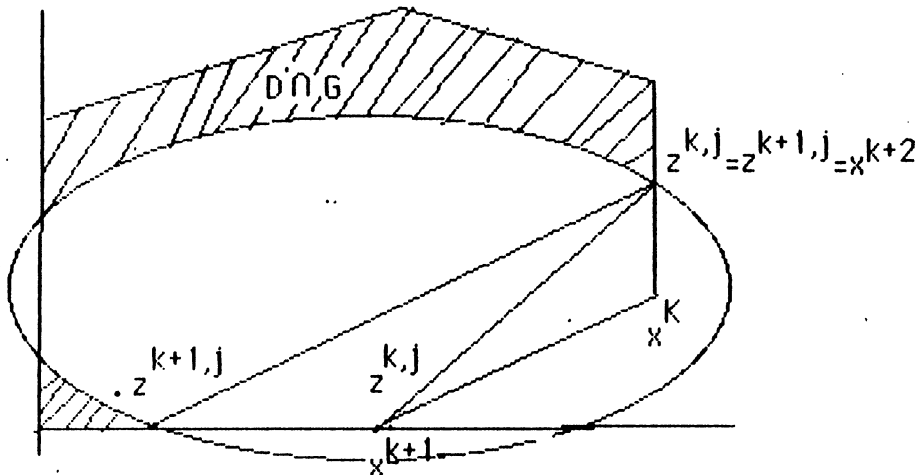
$$\alpha_{k+1} = f(y^{k+1}) = \min \{f(y^k), (f(z^{k,j}) / g(z^{k,j}) = 0)\}$$

où y^{k+1} réalise le minimum à la $(k+1)$ ème itération. On applique la méthode duale du simplexe pour trouver la solution x^{k+1} du problème $\min \{f(x) / x \in S^{k+1}\}$, où

$$S^{k+1} = S^k \cap \{x \in \mathbb{R}^n / \langle a_k, x \rangle \geq \beta_k\}.$$

Si $x^{k+1} \in S \cap G$, alors on arrête l'algorithme et x^{k+1} est un minimum global du problème (SP). Sinon, on passe à l'étape 3.

Etape 3 : On détermine le sommet x^{k+1} tel que $f(x^{k+1}) = \min\{f(x) / x \in E(x^1, \dots, x^k) \cap S^{k+1}\}$, où $E(x^1, \dots, x^k)$ est l'ensemble des sommets adjacents aux sommets de S testés. Si $f(x^{k+1}) \geq f(y^{k+1})$ alors y^{k+1} est une solution globale. Sinon on passe à l'itération $k=k+1$.



III.3.2.2 Mise en oeuvre de l'algorithme.

1) Initialisation:

On peut prendre y^1 comme solution du problème:

$$\min \{f(x) / x \in S, g(x^1) + \langle p(x^1), x - x^1 \rangle \geq 0\} \quad (p(x^1) \in \partial g(x^1)).$$

2) Construction d'une coupe

Pour formaliser cette construction de coupe, nous utiliserons les mêmes notations qu'au paragraphe II.3.3.

Le polyèdre S^k est écrit sous la forme canonique par rapport à une base réalisable:

$$\begin{aligned} A^* x_J + U x_{J'} &= b^* & (x = (x_J; x_{J'}) \geq 0), z \geq 0 \\ D^* x_I + U' z_h &= d^* \end{aligned}$$

dont la solution correspondant est le sommet (x^k, z^k) défini par:
 $(x^k_{J'} = b^*, x^k_{J''} = 0 \text{ et } z^k_h = d^*)$

Pour $j=1, \dots, n-m$,
 on définit le vecteur

$$a_j = \begin{cases} -a^*_{i,j} & \text{pour } i=1, \dots, m \\ +1 & \text{pour } i=J_j \text{ (} J_j \text{ désigne la } j^{\text{ème}} \text{élément de } J') \\ 0 & \text{Sinon.} \end{cases}$$

On calcule

$$\theta_j = \min \begin{cases} b^*_i / a^*_{i,j}, b^*_i > 0, a^*_{i,j} > 0, i=1, \dots, m \\ d^*_i / d^*_{i,j}, d^*_i > 0, d^*_{i,j} > 0, i=1, \dots, k \end{cases}$$

et

$$\rho_j = \max \{ \rho / g(x^k + \rho a_j) \leq 0, 0 < \rho \leq \theta_j \}$$

Proposition 3.1

$\rho_j > 0$ pour tout $j=1, \dots, n-m$ et le demi espace fermé $\sum_{j=1}^{n-m} x_{J_j} / \rho_j \geq 1$
 sépare le sommet x^k de l'ensemble admissible SNR

Démonstration

1) D'après le choix de x^k à l'étape 3 on a $g(x^k) < 0$. Puisque la fonction g est continue, alors il existe une boule $B(x^k, \delta)$ de centre x^k et de rayon $\delta > 0$ telle que pour tout $x \in B(x^k, \delta)$ on a $g(x) < 0$.

D'où $\rho_j \geq \delta > 0$ pour tout $j=1, \dots, n-m$.

2) Posons

$$R = \{ x \in \mathbb{R}^n / \sum x_{J_j} / \rho_j \leq 1 \}, D = SNR \text{ et } \text{int } R = \{ x \in \mathbb{R}^n / \sum x_{J_j} / \rho_j < 1 \}.$$

Par définition de l'ensemble D , on remarque que D est le simplexe des sommets $(x^k, z^{k,1}, \dots, z^{k,n-m})$ où $z^{k,l} = x^k + \rho_j a_j, (j=1, \dots, n-m)$.

C'est-à-dire, tout $y \in D$ peut s'écrire sous la forme suivante:

$$y = \lambda_0 x^k + \sum_{j=1}^{n-m} \lambda_j z^{k,j} \quad \text{avec } \lambda_j \geq 0 \quad \text{et } \sum_{j=0}^{n-m} \lambda_j = 1.$$

Par construction de points $z^{k,j}$ ($j=1, \dots, n-m$) et la convexité de g on a:

$$g(y) \leq \lambda_0 g(x^k) + \sum \lambda_j g(z^{k,j}) \leq 0.$$

D'où, pour tout $y \in S \cap \text{int } R$ (i.e. $\lambda_0 \neq 0$), on a $g(y) < 0$.

Ceci est équivalent à $D \cap \text{int } R \subset G^c$. \square

III.3.2.3 Convergence de l'algorithme

Théorème 3.4

L'algorithme décrit ci-dessus converge en un nombre fini de pas vers un optimum global du problème (SP).

Démonstration

Si l'algorithme s'arrête à l'étape 2, alors x^{k+1} est un minimum global, car par construction on a $S \cap G \subset S^k$ pour tout k .

Si l'algorithme s'arrête à l'étape 3, alors y^{k+1} est un optimum global ; en effet:

par construction du point y^{k+1} on a :

$$f(y^{k+1}) = \min\{f(x) / x \in \bigcup_{i \in I} A_i(S) \cap \partial G\},$$

où $\{A_i(S) / i \in I\}$ est l'ensemble des arêtes du polyèdre S dont chaque $A_i(S)$ a un élément de l'ensemble (x^1, \dots, x^k) comme point d'extrémité.

Si on s'arrête à l'étape 2, alors il n'existe aucun sommet s du polyèdre S tel que $f(s) < f(y^{k+1})$.

ceci est équivalent à dire que pour toute arête $[s_1, s_2]$ du polyèdre S dont les éléments appartiennent à S^{k+1} , on a pour tout $y \in [s_1, s_2] \cap G$:

$$f(y) \geq f(y^{k+1}).$$

Comme le nombre de sommets de S est fini, alors l'algorithme trouve un minimum global en un nombre fini d'itérations. \square

III.3.3 Algorithmes de type approximation extérieure

Cette classe d'algorithmes est composée de deux phases alternatives.

- Dans la phase 1, de $y^k \in E(S)$ tel que $g(x^k) > 0$, on calcule $x^{k+1} \in A(S) \cap \partial G$ tel que $f(x^{k+1}) \leq f(y^k)$.

- Dans la phase 2, on teste x^{k+1} pour l'optimalité globale et on détermine le point initial y^{k+1} de la phase 1 tel que $f(y^{k+1}) < f(x^{k+1})$ dans le cas où x^{k+1} n'est pas un minimum global.

III.3.3.1 Description de l'algorithme général.

Etape 0 : On commence par identifier un sommet $y^1 \in E(S) \cap G$ (cf. III.3.2.2) et on pose $k=1$.

Itération k: ($k=1, \dots$)

Phase 1: i) Du point y^k , on applique le processus de la méthode du simplexe pour le problème $\min\{f(x) \mid x \in S\}$ jusqu'à ce qu'on trouve deux sommets x_1 et $x_2 \in E(S)$ tels $g(x_1) \geq 0$, $g(x_2) < 0$ et $f(x_1) > f(x_2)$.

ii) On résout le problème suivant:

$$\lambda^* = \min\{\lambda \mid g(x_2 + \lambda(x_1 - x_2)) \geq 0, \lambda \geq 0\}$$

On pose $x^{k+1} = x_2 + \lambda^*(x_1 - x_2)$ et on passe à la phase 2.

Phase 2: a) On détermine un simplexe S_{k+1}^0 (cf. III.3.3.3) du sommet x^{k+1} et qui contient le polyèdre $S^i = \{x \in S \mid f(x) \leq f(x^{k+1})\}$ et on fait $i=0$.

b) Si pour chaque $s \in E(S_{k+1}^i) - \{x^{k+1}\}$ on a :

$g(s) < 0$ ou $g(s) = 0 \Rightarrow f(s) = f(x^{k+1})$, alors x^{k+1} est un minimum global. Sinon on passe à c).

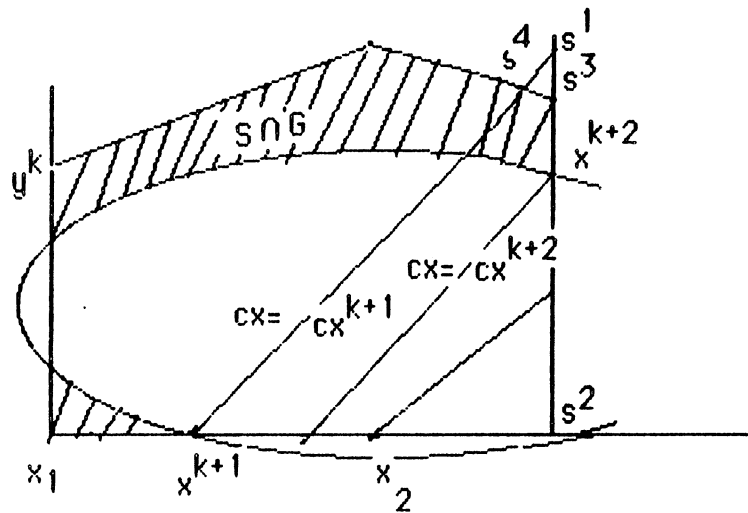
c) S'il existe un sommet s^* de S_{k+1}^i tel que $s^* \in S \cap G$ et $s^* \neq x^{k+1}$,

alors on pose $y^{k+1} = s^*$, $k:=k+1$ et on retourne à la phase 1.

Sinon, on prend $s^* = \arg \min\{f(s) \mid s \in E(S_{k+1}^i), g(s) \geq 0\}$ et on détermine l'indice "j" de la variable s^*_j qui est négative

et on passe à d).

d) On pose $S_{k+1}^{i+1} = S_{k+1}^i \cap \{x \in \mathbb{R}^n \mid x_j \geq 0\}$ et on calcule les éléments de l'ensemble $E(S_{k+1}^{i+1})$. On prend $i=i+1$ et on retourne à b).



III.3.3.2 Quelques variantes de l'algorithme général

D'après la caractérisation d'un ensemble qui contient un minimum global en terme d'un sommet qui réalise la solution du problème:

$$\min\{f(x) / x \in E(S) \cap G\}$$

(Théorème 3.3 et les conditions nécessaires d'optimalité du problème (SP) (cf. I.4)), nous proposons les deux variantes suivantes.

Première variante

Etape 1: On applique l'étape 0 et la phase I de l'algorithme général.

Etape 2: a) On détermine un simplexe S^0_1 (cf. III.3.3.3) du sommet x^1 et qui contient le polyèdre $S' = \{x \in S / f(x) \leq f(x^1)\}$ et on fait $i=0$.

b) Si pour chaque $s \in E(S^i_1) - \{x^1\}$ on a :

$g(s) < 0$ ou $g(s) = 0 \Rightarrow f(s) = f(x^1)$, alors x^1 est un minimum global. Sinon on passe à c).

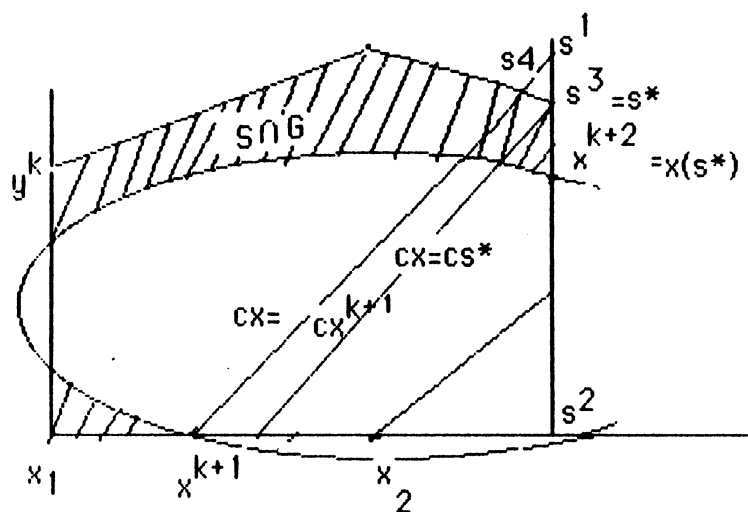
c) On calcule $s^* = \arg \min\{f(s) / s \in E(S^i_1), g(s) \geq 0\}$. Si $s^* \in E(S')$, alors on passe à l'étape 3-. Sinon on passe à d).

d) On pose $S^{i+1}_1 = S^i_1 \cap \{x \in \mathbb{R}^n / x_j \geq 0\}$, où "j" est l'indice d'une variable négative de s^* . On calcule les éléments de l'ensemble $E(S^{i+1}_1)$. On prend $i=i+1$ et on retourne à b).

Etape 3: On détermine une solution x^* du problème

$$\min\{f(x) / x \in \bigcup_{y \in E(S^*)} A_y(S^*) \cap \partial G\}, \text{ où } S^* = \{x \in S / f(x) = f(s^*)\}$$

(cf. III.3.3.3). Si $x^* \in S \cap \partial G$ alors x^* est un minimum global du problème (SP), sinon on retourne au pas d) de l'étape 2 avec $s^* = x^*$.



Deuxième variante

Etape 0: Comme dans l'algorithme général.

Itération k:

Phase 1: Du point y^k , on applique l'algorithme de sous-gradient du chapitre I pour obtenir un minimum local x^k et on passe à la phase 2.

Phase 2: On teste x^k pour l'optimalité globale comme dans l'algorithme général.

III.3.3.3 Mise en oeuvre de l'algorithme général et ces variantes

a) Construction du simplexe S_k^0 :

Pour formaliser la construction du simplexe S_k^0 qui contient le polyèdre $S^* = \{x \in S / f(x) \leq f(x^k)\}$, nous utiliserons les notations suivantes:

$$S^* = \{s \in \mathbb{R}^{n+1} / s = (x, x_{n+1}), s_j = b^* - A^* s_j, s_2 \geq 0\},$$

où A^* est une matrice du type $(m+1, n-m)$, s_J (resp. J) désigne l'ensemble des variables (resp. des indices) de base, $s_{J'}$ (resp. J') désigne l'ensemble des variables (resp. des indices) hors base et x_{n+1} désigne la variable d'écart de la contrainte $f(x) \leq f(x^k)$, $\text{card}(J') = n-m$ et $\text{card}(J) = m+1$. La solution $s^0_J = b^*$, $s^0_{J'} = 0$ est le sommet $s^0 = (x^k, x^k_{n+1})$.

Soit

$$K = \{s \in \mathbb{R}^{n+1} / s_J = b^* - A^*s_{J'}, s_{J'} \geq 0\},$$

le cône polyédral du sommet s^0 .

On définit S^0_k par :

$$S^0_k = \{s \in K / \sum s_{J'} \leq \alpha\}, \text{ où } \alpha = \max\{\sum s_{J'} / s \in S'\}.$$

Proposition 3.2

Le simplexe $S^0_k[s^0, \dots, s^{n-m}]$ contient le polyèdre S' et les sommets s^i ($i=0, 1, \dots, n-m$) sont définis par:

$$\begin{aligned} s^i_{J'1} &= \alpha \\ s^i_{J'} &= b^* - \alpha a^*_i \\ s^i_{J'-(J'1)} &= 0 \end{aligned} \quad (i=1, \dots, n-m)$$

avec a^*_i (resp. $J'1$) est la i^{em} colonne (resp. i^{em} élément) de la matrice A^* (resp. de l'ensemble J').

Démonstration

Puisque dans le cône K on élimine les contraintes $s_{J'} \geq 0$, on a donc $S' \subset K$.

Par définition du scalaire α , on a $S' \subset \{s \in \mathbb{R}^{n+1} / \sum s_{J'} \leq \alpha\}$.

D'où $S' \subset S^0_k$.

Par construction du S^0_k , ces sommets sont définis par l'intersection, des $(n-m)$ contraintes parmi les $(n+1-m)$ contraintes $s_{J'} = 0$ et $\sum s_{J'} = \alpha$ ([14]).

D'après la définition des points $s^i (i=0, \dots, n-m)$, on a:
 s^0 est l'intersection de ces $(n-m)$ contraintes $s_j=0$ et chaque $s^i (i \geq 1)$
réalise l'intersection de l'hyperplan $\sum s_j = \alpha$ avec ces $(n-m-1)$ contraintes
 $s_{j'-(j')}$. \square

Remarque

Si toutes les coefficients d'une ligne de la matrice A^* sont strictement positives alors la contrainte associée à cette ligne peut remplacer la contrainte qu'on ajoute (i.e. $\sum s_j \leq \alpha$).

b) Détermination de l'ensemble $E(S^{i+1}_k)$

Voir le procédé décrit au paragraphe (II.5) du chapitre II.

c) L'étape 3 de la première variante

- On détermine l'ensemble $E(S'')$ par les opérations de pivotage [38].
- Pour chaque $s \in E(S'')$, on détermine l'ensemble (s_1, \dots, s_h) des sommets du polyèdre S^i_1 qui lui sont adjacents et qui ne sont pas dans G .
- Dans chaque segment (arête) $[s, s_j] (j=1, \dots, h)$ on applique l'étape ii) donnée à la phase I pour déterminer $s^*_j = s + \lambda^*_j(s_j - s)$ et on pose:

$$x^*(s) = \arg \min \{ f(s^*_j), j=1, \dots, h \}.$$

Ensuite, on prend

$$x^* = \arg \min \{ f(x^*(s)) / s \in E(S'') \}$$

comme solution globale du problème $\min \{ f(x) / x \in \bigcup_{y \in E(S'')} A_y(S^i_1) \cap \partial G \}$.

III.3.3.4 Convergence

Théorème 3.4

L'algorithme général et ses variantes convergent en un nombre fini d'itérations vers un minimum global du problème (SP).

Démonstration

Par construction de l'ensemble des polyèdres S^i_k ($i \geq 0, k \geq 1$), on a :

$$S^{m+1}_k = S^i \text{ pour tout } k \geq 0.$$

D'où la phase 2 est finie .

Par construction de la suite (y^k) dans l'algorithme général et la deuxième variante , la suite $f(y^k)$ est strictement décroissante et chaque y^k est un sommet du polyèdre S^i tel que $f(x^{k+1}) < f(y^k)$.

L'étape 3 de la première variante est finie d'après le procédé décrit en (III.3.3.3 c).

D'où la convergence de ces algorithmes en un nombre fini de pas . Nous allons montrer que chaque point obtenu par l'un de ces algorithmes réalise un minimum global du (SP) .

Du théorème 3.2 (III.3.1) et du procédé décrit en III.3.3.3.c , il est clair que chaque x^* obtenu à l' étape 3 de la première variante est un minimum global du problème $\min\{f(x) / x \in \bigcup_{y \in E(S^i)} A_y(S^i_1) \cap \partial G\}$. Puisque par construction $S^i \subset S^i_1$, alors on arrête l'algorithme dans le cas où x^* est réalisable .

Pour l'algorithme général et la deuxième variante , si à la $k^{\text{ème}}$ itération l'algorithme s'arrête à la phase 2 alors x^{k+1} est un minimum global , car par l'hypothèse de convexité de g et les conditions d'arrêt (étape 2) il est facile de vérifier que pour tout $y \in S^i \subset S^i_k$ on a :

$$f(y) < f(x^{k+1}) \Rightarrow g(y) < 0 \text{ (i.e. } y \notin G)$$

et

$$g(y) = 0 \Rightarrow f(y) = f(x^{k+1}) \text{ (i.e. } y \text{ est une autre solution globale) . } \square$$

III.4 Algorithmes de résolution dans le cas général

Nous proposons deux algorithmes de type approximation extérieure pour la résolution des problèmes d'optimisation non convexe qui peuvent s'écrire sous la forme du problème (P).

-Le premier algorithme est basé sur les méthodes de coupe (ou accumulation de contraintes) qui sont bien connues dans la littérature pour la résolution des problèmes d'optimisation convexe sous plusieurs variantes intéressantes en pratique (I.L.Kelley [20], D.M.Topkis [50], B.Martinet [24], ...).

-Le deuxième algorithme est considéré comme une généralisation de l'algorithme (II.5.1) (chapitre II) pour la résolution du problème (P).

Nous faisons l'hypothèse qu'il existe $w \in \text{SNG}$ tel que $h(w) < 0$ et que le domaine admissible est compact.

III.4.1 Description de l'algorithme 1

Initialisation

On résout le sous problème (SP_1) par l'algorithme (III.3.2.1):

$$\min \{ f(x) / x \in \text{SNG} \} \quad (SP_1)$$

Soit x^1 la solution, si $x^1 \in C$ (i.e. $x^1 \in H$) alors x^1 est un minimum global du problème (P). Sinon, on pose $S^1 = S$, $k=1$ et on passe à l'itération k .

Itération k : ($k=1, \dots$)

Etape 1: On détermine l'élément y^k de l'intersection du segment $[w, x^k]$ avec la frontière de H (i.e. $h(y^k)=0$). On pose

$$S^{k+1} = S^k \cap \{ x \in \mathbb{R}^n / \langle p(y^k), x - y^k \rangle \leq 0 \} \quad (p(y^k) \in \partial h(y^k))$$

Etape 2: On applique l'algorithme (III.3.2.1) pour la résolution du problème (SP_k) suivant:

$$\min \{ f(x) / x \in S^{k+1} \cap \text{NG} \} \quad (SP_k)$$

Soit x^{k+1} la solution du (SP_k) , si $x^{k+1} \in H$ alors on arrête l'algorithme et x^{k+1} est un minimum global. Sinon, on retourne à l'étape 1 avec $k=k+1$.

III.4.2 Description de l'algorithme 2.

On initialise avec un polytope S^1 contenant C et le point w .

Itération k : ($k=1, \dots$)

Etape 1: On calcule $f(x^k) = \min \{f(x) / x \in E(S^k) \cap G\}$. Si $x^k \in H \cap S \cap \partial G$, alors x^k est un minimum global.

Si $x^k \notin H \cap S$ on passe à l'étape 2 et dans le cas contraire à l'étape 3:

Etape 2: i) Si $x^k \notin H$, alors on détermine y^k l'élément de l'intersection du segment $[w, x^k]$ avec ∂H . Si $y^k \in S$ alors on pose

$$S^{k+1} = S^k \cap \{x \in \mathbb{R}^n / \langle p(y^k), x - y^k \rangle \leq 0\}$$

et on passe à l'étape 4. Dans le cas contraire, on remplace x^k par y^k et on passe à ii).

ii) On détermine la contrainte violée en x^k du polyèdre S , par exemple $\langle a_j, x \rangle \leq b_j$. On pose

$$S^{k+1} = S^k \cap \{x \in \mathbb{R}^n / \langle a_j, x \rangle \leq b_j\}$$

et on passe à l'étape 4.

Etape 3: On applique l'étape 3 du paragraphe (III.3.3.3,c). Soit z^k la solution obtenue. Si $z^k \in H \cap S$ alors on arrête l'algorithme et z^k est un minimum global, sinon on retourne à l'étape 2 après le remplacement de x^k par z^k .

Etape 4: On détermine l'ensemble $E(S^{k+1})$ des sommets du polytope S^{k+1} (cf. II.5.2) et on passe à l'itération $k=k+1$.

III.4.3 Convergence

Théorème 4.1

Pour toute suite (x^k) engendrée par l'algorithme 1 ou l'algorithme 2 décrits ci-dessus on a les propriétés suivantes:

1- Si l'algorithme 1 (resp l'algorithme 2) s'arrête à l'itération k , alors x^k est une solution globale du problème (P).

2- Tout point d'adhérence de la suite (x^k) est un minimum global du problème (P).

Démonstration

Analogue à celle du théorème 5.1 du chapitre II. \square

III.4.4 *Remarques*

1- les variantes de ces algorithmes généraux dépendent des hypothèses du problème initial et de la façon de le ramener à la forme du problème (P)(cf.III.2).

2-On peut étendre sans difficulté les techniques d'accumulations des contraintes et l'élimination des contraintes non actives qui sont connues dans [24 ,50] et d'autres articles pour certaines classes des problèmes d'optimisation convexe .

3-Le choix du polytope initial S^0 est important en pratique et il y a plusieurs façons de le déterminer.Par conséquent, dans l'exemple d'application (III.2.2) on peut prendre S^0 comme le simplexe du sommet l'origine s^0 et des sommets $s^i(i=1, \dots, n)$ définis par:

$$s^i_j=0 \text{ pour } i \neq j, s^i_j=1 \text{ pour } i=j.$$

III.5 Essais numériques

Les algorithmes III.3.2.1 et III.3.3.1 de résolution des problèmes de programmation linéaire avec une contrainte anticonvexe ont été programmés sur Micral 90.50 en langage Pascal. Pour faciliter la compréhension des procédures et leurs interprétations géométriques, nous présentons les détails de résultats numériques de deux exemples suivants.

Exemple 1[15]

$$\max f(x) = 1.5x_1 - x_2$$

sous les contraintes :

$$x_1 \leq 8,$$

$$x_2 \leq 6$$

$$x_1 - x_2 \leq 6,$$

$$-2x_1 + x_2 \leq 5$$

$$g(x) = x_1^2 + 4x_2^2 - 8x_1 - 16x_2 + 7 \geq 0$$

$$x_1 \geq 0, x_2 \geq 0$$

Le maximum global est $x^* = (8, 3.5)$ avec $f(x^*) = 8.5$ et $g(x^*) = 0.0$

Exemple 2[15]

$$\max f(x) = x_1 + x_2$$

sous les contraintes :

$$x_1 - x_2 \leq 6.9$$

$$3x_1 + x_2 \leq 8.1$$

$$x_2 \leq 2.1$$

$$g(x) = 2(x_1 - 2)^2 + (x_2 - 0.1)^2 + (x_2 - 1.1)^2 - 2 \geq 0$$

$$x_1 \geq 0, x_2 \geq 0$$

La solution globale est $x^* = (1.38, 1.2)$ avec $f(x^*) = 2.6$ et $g(x^*) \approx 0.0$

Algorithme III.3.2.1

k	$(x^k, g(x^k))$	$(x^{k,j}, g(x^{k,j}))$	$(z^{k,j}, g(z^{k,j}))$
1	(8, 2, -9)	(8, 6, 55) (6, 0, -5)	(8, 3.5, 0.0) (6, 0, -5)
2	(6, 0, -5)	(8, 3.5, 0.0) (0, 0, 7.0)	(8, 3.5, 0.0) (1.0, 0, 0)
3	(8, 3.5, 0)	-	-

-Exemple 1 -

k	$(x^k, g(x^k))$	$(x^{k,j}, g(x^{k,j}))$	$(z^{k,j}, g(z^{k,j}))$
1	(2.3, 1.2, -0.6)	(2.5, 0.65, -0.12) (0.0, 1.2, 7.22)	(2.5, 0.65, -0.12) (1.36, 1.2, 0.0)
2	(2.5, 0.65, -0.12)	(2.3, 0.0, -0.6) (1.36, 1.2, 0.0)	(2.3, 0.0, -0.6) (1.36, 1.2, 0.0)
3	(1.36, 1.2, 0.0)	-	-

-Exemple 2 -

Algorithme III.3.3.1

k	i	$(y^k, g(y^k))$	$(x^k, g(x^k))$	$v(S_k^1)$	s^*	$g(s^*)$
0		(0, 0, 7)	(1, 0, 0)	(8, 10.5) (8, 0)	(8, 0)	278 7
1	1	"	"	(8, 10.5) (6, 0) (8, 2)	(8, 10.5)	278 -5 -9
2		"	"	(5, 6) (6, 0) (8, 2) (8, 6)	(8, 6)	40 -5 -9 55
0		(8, 6, 55)	(8, 3.5, 0.0)	(8, 0) (5.6, 0)	(8, 0)	7 -6
2	1	"	"	(8, 2) (5.6, 0) (6, 0, 0)		-9 -6 -5

-Exemple 1-

Remarque

Dans l'exemple 1, le sommet $s^*=(8,6)$ est la solution du problème $\max\{f(x) / x \in E(S) \cap G\}$ (cf III.3.3.3). En utilisant l'étape 3 de la première variante, on obtient:

$$E(S'') = \{(8,6), (4,0)\}$$

$$U_{v \in E(S'')} A(S') = \{[(8,6), (8,2)], [(4,0), (6,0)], [(8,6), (4,0)]\}$$

$$\mathcal{P}(s^*) = \{(8,3.5), (6.8, 4.2)\}$$

$$x^* = \arg \max\{f(x) / x \in \mathcal{P}(s^*)\} = (8,3.5)$$

k	i	$(y^k, g(y^k))$	$(x^k, g(x^k))$	$V(S_k^i)$	s^*	$g(s^*)$
0		(0, 0, 7.22)	(1.38, 0, 0)	(5.7, 0) (-2.93, 4.31)	(5.7, 0.0)	26.6 74.83
1	1	"	"	(2.26, 0.0) (-2.93, 4.31) (2.7, 1.4)	(2.7, 1.4)	-0.63 74.83 1.65
2		"	"	(2.26, 0.0) (-2.93, 4.31) (2.48, 0.65) (2.1, 1.8)	(2.1, 1.8)	-0.63 74.83 -1.02 1.4
3		"	"	(2.26, 0.0) (0.18, 1.2) (2.48, 0.65) (2.3, 1.2)	(0.18, 1.2)	-0.63 5.82 -1.02 -0.6
2	0	(0.18, 1.2)	(1.38, 1.2, 0.0)	(2.3, 0.24) (2.6, 1.2)	(2.6, 1.2)	-1.0 0.109
1		"	"	(2.3, 0.24) (2.4, 0.65) (2.3, 1.2)		-1.0 -1.02 -0.6

-Exemple 2-



CHAPITRE IV

Applications



IV.1 Introduction

Le but de ce chapitre est de prouver que les procédés décrits dans le chapitre III peuvent apporter des raffinements à des problèmes de nature particulière et de montrer qu'ils peuvent être utilisés pour la résolution d'une classe très large de problèmes de programmation fractionnaire .

Le problème de maximisation d'une fonction quadratique symétrique définie négative (respectivement définie positive) a été très largement étudié par plusieurs chercheurs comme un cas particulier de la programmation convexe (respectivement de maximisation d'une fonction convexe , chapitre II).Pour la bibliographe, le lecteur pourra consulter [7,21,52,56] .

Par contre le problème de maximisation (ou minimisation) d'une fonction quadratique indéfinie est très peu étudié ,récemment des algorithmes ont été proposés [40,52] pour sa résolution comme un problème de maximisation de la différence de deux fonctions concaves (ou la minimisation de la différence de deux fonctions convexes). Dans le paragraphe IV.2 ,on transforme le problème de maximisation d'une fonction quadratique non convexe sur un polyèdre borné en un problème de type programmation linéaire avec une contrainte anticonvexe. En se basant sur les procédés du chapitre III , on donne un algorithme qui trouve le point de Kuhn & Tucker associé à la solution globale en un nombre fini de pas .

Au paragraphe IV .3 ,nous signalons que le problème de maximisation de la somme en valeur absolue de fonctions affines sous des contraintes linéaires [42] peut se résoudre complètement par les procédés du chapitre précédent.

Dans le paragraphe IV.4, on s'intéresse aux problèmes de programmation fractionnaire qu'on rencontre dans plusieurs domaines d'applications (en particulier en économie). On montre que les problèmes de minimisation du rapport de deux fonctions d.c sous des contraintes d.c peuvent se ramener aux formes des trois problèmes traités dans cette thèse .

En fin de chapitre nous donnerons quelques exemples numériques traités avec leurs résultats .

IV.2 Maximisation globale d'une fonction quadratique non convexe

On applique le théorème de Kuhn & Tucker pour réduire la maximisation d'une fonction quadratique sous des contraintes linéaires à un problème de forme programmation linéaire avec une contrainte anticonvexe. Nous adaptons ensuite l'algorithme III.3.3.1 pour sa résolution

IV.2.1 Réduction du problème d'optimisation quadratique

Considérons le problème d'optimisation quadratique suivant:

$$\max f(x) = 1/2 \langle x, Cx \rangle + \langle d, x \rangle + \alpha$$

sous les contraintes: (P)

$$x \in D = \{x \in \mathbb{R}^n / Ax \leq b, x \geq 0\}$$

où C est une matrice carrée symétrique (n,n), A est une matrice (m,n) et α est une constante réelle.

On fait les hypothèses que la borne supérieure atteinte et que le polyèdre D est non vide.

On sait que le théorème de Kuhn & Tucker (voir par exemple [23]) donne les conditions nécessaires d'optimalité du problème (P) et que leur application peut se formuler de la manière suivante :

Pour que $x^* \in \mathbb{R}^n$ soit une solution optimale du problème (P), il faut qu'il existe un vecteur $(u^*, x^*, y^*, v^*) \in \mathbb{R}^{2(m+n)}$ qui vérifie les relations suivantes:

$$\begin{aligned} Ax + y &= b \\ -A^t u + Cx + v &= -d \\ \langle u, y \rangle + \langle v, x \rangle &= 0 \\ u \geq 0, y \geq 0, x \geq 0, v \geq 0 \end{aligned} \quad (2.1)$$

Supposons (2.1) vérifié en (u^0, x^0, y^0, v^0) , alors

$$\begin{aligned}
 f(x^0) &= 1/2 \langle x^0, Cx^0 \rangle + \langle d, x^0 \rangle + \alpha \\
 &= 1/2 (\langle -d, x^0 \rangle + \langle u^0, Ax^0 \rangle - \langle v^0, x^0 \rangle) + \langle d, x^0 \rangle + \alpha \\
 &= 1/2 (\langle d, x^0 \rangle + \langle u^0, b \rangle - \langle u^0, y^0 \rangle + \langle v^0, x^0 \rangle) + \alpha \\
 &= 1/2 (\langle d, x^0 \rangle + \langle b, u^0 \rangle) + \alpha \quad (2.2)
 \end{aligned}$$

Par conséquent, si x^0 n'est pas une solution globale, alors il existe un autre point de Kuhn & Tucker (u^1, x^1, y^1, v^1) tel que :

$$f(x^1) = 1/2 (\langle d, x^1 \rangle + \langle b, u^1 \rangle) + \alpha > f(x^0) = 1/2 (\langle d, x^0 \rangle + \langle b, u^0 \rangle) + \alpha \quad (2.3)$$

D'autre part, on sait que [54] la contrainte non linéaire $\langle u, y \rangle + \langle v, x \rangle = 0$ avec $u, x, y, v \geq 0$ peut s'exprimer sous la forme :

$$g(u, x, y, v) = \sum_{i=1}^m \min(u_i, y_i) + \sum_{i=1}^n \min(x_i, v_i) \leq 0 \quad (2.4)$$

où la fonction g est concave.

Donc, la contrainte non linéaire dans (2.1) peut être remplacée par une contrainte anticonvexe $-g(u, x, y, v) \geq 0$ dans $\mathbb{R}^{2(n+m)}$.

Compte tenu de (2.1)-(2.4), on en déduit que le problème (P) peut se transformer en un problème équivalent (P') (dans le sens que la résolution d'un problème entraîne celle de l'autre) défini par:

$$\max \langle p, z \rangle + \alpha$$

sous les contraintes : (P')

$$(z, w) \in S = \{ t = (z, w) \in \mathbb{R}^{2h} / Mz + w = q, t \geq 0 \}$$

et

$$(z, w) \in G = \{ t = (z, w) \in \mathbb{R}^{2h} / -g(z, w) \geq 0 \}$$

où

$$M = \begin{pmatrix} 0 & A \\ -A^t & C \end{pmatrix}, p = 1/2 \begin{pmatrix} b \\ d \end{pmatrix}, q = \begin{pmatrix} b \\ -d \end{pmatrix}, z = \begin{pmatrix} u \\ x \end{pmatrix}, w = \begin{pmatrix} y \\ v \end{pmatrix} \text{ et } h = 2(n+m)$$

Difficultés du problème équivalent (P')

On trouve dans la littérature un nombre considérable de travaux

consacrés à l'élaboration d'algorithmes finis pour la résolution des problèmes de type " *complémentaire linéaire* " .

En particulier pour les fonctions quadratiques , le problème de complémentaire linéaire s'exprime sous la forme:

$$\text{" trouver } z \text{ et } w \text{ tels que : } Mz + w = q, z \geq 0, w \geq 0 \text{ et } \langle z, w \rangle = 0 \text{" (2.5)}$$

Le lecteur pourra consulter sur ce sujet [56] . Malheureusement , ces méthodes qui sont basées sur les techniques de programmation linéaires (en particulier les opérations de pivotage) ne permettent pas d'affirmer (en général) qu' un tel point réalise la solution globale ou non .

De façon plus précise, on ne peut pas appliquer les critères d'optimalité globale de la méthode du simplexe [38] sur la fonction objectif $\langle p, z \rangle + \alpha$.

L'exemple suivant montre les principales difficultés:

Exemple

Soit le problème d'optimisation quadratique non convexe suivant :

$$\max\{ f(x) = 1/2 (x_1^2 + x_2^2) - 0.5 x_1 - x_2 \mid 0 \leq x_i \leq 2, (i=1,2) \}$$

dont la fonction $f(x)$ est convexe et $x^* = (2, 0)$ est un maximum global.

On trouve que :

1- Le point $(u^*, x^*, y^*, v^*) = \{ (3/2, 0), (2, 0), (0, 2), (0, 1) \}$ est le point de Kuhn & Tucker associé à la solution globale x^* .

2- Le critère d'optimalité de la méthode du simplexe n'est pas vérifié ,
 $\langle p, z \rangle = -0.75 y_1 + v_1 + v_2 - 1.5 x_2$.

3- Si la contrainte non linéaire (anticonvexe) est omise , la valeur de la fonction objectif donnée par l'algorithme du simplexe peut être infinie.

Relation entre les solutions de deux problèmes .

Théorème 2.1

x^* est un maximum global du problème (P) si et seulement s'il existe (z^*, w^*) tel que:

$$\langle p, z^* \rangle + \alpha = \max \{ \langle p, z \rangle + \alpha \mid (z, w) \in E(S), g(z, w) = 0 \},$$

où $E(S)$ est l'ensemble des sommets du polyèdre S .

Démonstration

Conséquence immédiate de la transformation du problème (P) et du fait que $S \cap \partial G = E(S) \cap \partial G$. \square

Remarque

C'est le théorème 2.1 qui met en évidence le rôle important de réduction du problème (P). On a montré en effet que:

-Si le problème (P) a une solution globale, alors il a un optimum global qui correspond à un sommet du polyèdre S.

-L'importance de l'introduction de la contrainte non linéaire sous la forme d'une contrainte anticonvexe est qu'elle permet de tester chaque point de Kuhn & Tucker à l'optimalité globale du problème (P') (ou (P)).

4.2. Algorithme

On suppose que $\max\{f(x) / x \in D\} \leq \beta < \infty$.

Initialisation:

On initialise avec un point de Kuhn & Tucker $t^0 = (z^0, w^0)$ par l'une des méthodes proposées dans [37,56]. On pose $S^1 = \{s = (t, r) \in \mathbb{R}^{2(h+1)} / t \in S, -\langle p, z \rangle + r_2 = -\langle p, z^0 \rangle, \langle p, z \rangle + r_1 = \beta - \alpha\}$, on détermine un simplexe S^1 du sommet $s^0 = (t^0, r^0)$ tel que $F' \subset S^1$ et on passe à l'itération k avec $k=1$.

Itération k:

Etape 1: Si pour tout $s \in E(S^k)$ on a:

$g(s) > 0$ ou $g(s) = 0 \Rightarrow \langle p, z \rangle + \alpha = \langle p, z^0 \rangle + \alpha$. Alors t^0 est une solution globale du problème (P'). Sinon, on passe à l'étape 2.

Etape 2: On calcule $s^* = \arg \max\{\langle p, z \rangle + \alpha / s = (t, r) \in E(S^k), g(t) \leq 0\}$. Si $t^* \in E(S) \cap \partial G$, alors on arrête l'algorithme et t^* est une solution globale du problème (P'). Sinon, on détermine le variable s_i qui est strictement négative (i.e. une contrainte violée) et on pose

$$S^{k+1} = S^k \cap \{s \in \mathbb{R}^{2(h+1)} / s_i \geq 0\}.$$

Etape 3: On détermine l'ensemble $E(S^{k+1})$ (cf II.5.2) et on pose $k = k+1$.

Quelques indications sur la mise en oeuvre de l'algorithme.

1-Pour la construction du simplexe S^1 du sommet $t^0=(z^0, w^0)$ qui contient S^0 , voir paragraphe III.3.3.3.

2-On remarque que le sommet s^0 est dégénéré dans S^1 , en pratique, il est intéressant de mettre la variable d'écart r_2 de la contrainte $\langle p, z \rangle \geq \langle p, z^0 \rangle$ dans l'ensemble des variables hors base pour assurer que $\langle p, z \rangle \geq \langle p, z^0 \rangle$ pour tout $s \in E(S^k)$.

3-Voir II.5.2 pour la détermination de l'ensemble des sommets du polyèdre S^k .

4-On peut adapter sans difficulté les variantes du paragraphe III.3.3.2

IV.2.3 Convergence

Théorème 2.2

L'algorithme décrit ci-dessus trouve le point de Kuhn & Tucker associé à une solution globale du problème (P) en un nombre fini d'itérations.

Démonstration

Analogue à celle du théorème 4.1 du chapitre III. □

IV.3 Maximisation globale de la somme en valeur absolue des fonctions affines sur un polyèdre

Considérons le problème d'optimisation suivant (cf. [33, 42]).

$$\max f(x) = \sum_{i=1}^m |f_i(x)| \tag{P}$$

sous les contraintes:

$$x \in D = \{x \in \mathbb{R}^n / Ax \leq b, x \geq 0\}$$

où les fonctions $f_i(x) = \langle c_i, x \rangle - d_i$ ($i=1, \dots, m$) sont linéaires et D est un polyèdre convexe borné.

L'objet de ce paragraphe est de signaler que cette classe de problèmes peut être traitée par les algorithmes du paragraphe III.3. En effet: on montre dans [42] que si l'on fait le changement de variable

$$u_i - v_i = f_i(x) , u_i, v_i \geq 0 , i=1 , \dots , m$$

le problème (P) peut se formuler de la manière suivante:

$$\max \sum_{i=1}^m u_i + v_i$$

sous les contraintes:

$$Ax + y = b$$

$$Cx - u + v = d$$

$$\langle u, v \rangle = 0$$

$$x, y, u, v \geq 0$$

(Q')

où C est une matrice (m,n) dont chaque i^{ème} ligne est le vecteur c_i et

$u = (u_1, \dots, u_m)$ (resp. $v = (v_1, \dots, v_m)$) est un m-vecteur colonne.

Comme il est mentionné au paragraphe précédent, la contrainte $\langle u, v \rangle = 0$ peut être remplacée par une contrainte anticonvexe et que les procédés du chapitre III peuvent être utilisés pour trouver les solutions globales de ce type de problèmes.

Remarque

Dans [33,42] les auteurs trouvent les mêmes difficultés que celles que nous avons signalées dans le paragraphe précédent pour la résolution de ce type de problèmes.

IV.4 Programmation fractionnaire

Les problèmes d'optimisation fractionnaire ont attirés dans ces dernières années une attention qui reflète l'intérêt de ces problèmes dans plusieurs domaines d'applications. Il existe à l'heure actuelle de très nombreuses publications sur les problèmes d'applications et les méthodes numériques de résolution, nous n'en citerons que [19,40]. Pour une bibliographie complète, le lecteur pourra consulter S.Schaible [41].

L'objet du paragraphe est de démontrer qu'une classe très large de ce type de problèmes se ramène à la forme de l'un de trois problèmes principaux étudiés dans cette thèse. L'étude des conditions nécessaires d'optimalité et la dualité de cet exemple dans le contexte de la transformation que nous utilisons est en cours.

IV.4.1 Notations-Définitions-Hypothèses

Définition 1

un programme fractionnaire est un problème d'optimisation du type :

$$\min F(x) = f(x) / g(x) \quad (P)$$

sous les contraintes : $x \in S$

où la fonction g est strictement positive sur S et
 $S = \{x \in \mathbb{R}^n / h_i(x) \leq 0, i=1, \dots, m\}$.

Définition 2

Un programme fractionnaire d.c est un problème d'optimisation du type (P) dans lequel les fonctions $f, g, h_i, i=1, \dots, m$ sont continues et admettent chacune une décomposition en une différence de deux fonctions convexes.

Nous faisons les hypothèses suivantes :

H1- S est un compact non vide.

H2- $\min \{f(x) / x \in S\}$ est strictement positive .

On verra dans la suite le rôle de l'hypothèse H2 dans la classification des problèmes de programmation fractionnaire.

Relation entre programmation fractionnaire d.c et programmation mathématique d.c.

Soit le changement de variable introduit par Charnes et Cooper dans le cas où f et g sont linéaires. Il a été utilisé par Schaible dans le cas général [40]:

$$y = x / g(x) \quad , t = 1 / g(x) \quad (4.1)$$

On définit le problème (P') par :

$$\min F(y,t) = t f(y/t) \quad (P')$$

sous les contraintes:

$$(y,t) \in \{ (y,t) \in \mathbb{R}^n \times \mathbb{R} \mid t > 0, y/t \in S, tg(y/t) = 1 \} \quad (4.2)$$

Les problèmes (P) et (P') sont liés de la manière suivante:

Lemm 4.1

Si (P) est un problème de programmation fractionnaire d.c, alors les hypothèses H1 et H2 entraînent que:

1-Le problème (P') a une solution optimale

2-La contrainte d'inégalité $tg(y/t) \geq 1$ peut remplacer la contrainte d'égalité $tg(y/t) = 1$.

3-Les solutions des problèmes (P) et (P') sont liées par (4.1).

Démonstration

1- La compacité de S et la continuité de g entraînent que l'ensemble admissible du problème (P') est compact. Ceci affirme l'existence d'une solution optimale du problème (P').

2- Soit $(y', t') \in \mathbb{R}^n \times \mathbb{R}_+$ tel que $y'/t' \in S$ et $t'g(y'/t') > 1$.

Si on définit (y'', t'') par:

$$y''/t'' = y'/t' \quad \text{et} \quad t'' = 1 / g(y'/t')$$

alors

$$t'' < t' \quad \text{et} \quad y''/t'' \in S.$$

Compte tenu de l'hypothèse H2, on a:

$$t''f(y''/t'') \leq t'f(y''/t'') = t'f(y'/t')$$

ce qui donne:

$$\min\{tf(y/t) / t > 0, y/t \in S, tg(y/t) = 1\} = \min\{tf(y/t) / t > 0, y/t \in S, tg(y/t) \geq 1\}$$

3-Immédiate.

Lemme 4.2

Si (P) est du type programmation fractionnaire d.c, alors les hypothèses H1 et H2 entraînent qu'il se ramène à un problème de programmation mathématique d.c (Q) du type:

$$\min\{F(y,t) = tf(y/t) / 1/\alpha_1 \geq t \geq 1/\alpha_2, y/t \in S, tg(y/t) \geq 1\} \quad (Q)$$

$$\text{où } \alpha_1 \leq \min\{g(x) / x \in S\} \text{ et } \alpha_2 \geq \max\{g(x) / x \in S\}$$

Démonstration

Il est clair qu'il suffit de montrer que :

Si f est convexe pour tout $x \in \mathbb{R}^n$, alors $tf(y/t) = F(y,t)$ est convexe pour tout $(y,t) \in \mathbb{R}^n \times \mathbb{R}_+^*$. En effet;

soient $(y',t'), (y'',t'') \in \mathbb{R}^n \times \mathbb{R}_+^*$, et $0 \leq \mu \leq 1$

$$F(\mu y' + (1-\mu)y'', \mu t' + (1-\mu)t'') = \mu t' + (1-\mu)t'' f(\mu y' + (1-\mu)y'' / \mu t' + (1-\mu)t'')$$

$$= \mu t' + (1-\mu)t'' f((\mu t' / \mu t' + (1-\mu)t'') y' / t' + ((1-\mu)t'' / \mu t' + (1-\mu)t'') y'' / t'')$$

Comme la fonction f est convexe, alors

$$F(\mu y' + (1-\mu)y'', \mu t' + (1-\mu)t'') \leq \mu t' f(y'/t') + (1-\mu)t'' f(y''/t'') = \mu F(y',t') + (1-\mu)F(y'',t'').$$

Ceci montre que si f est du type d.c alors $tf(y/t)$ est aussi du type d.c

Proposition 4.1

Si l'hypothèse H2 n'est pas satisfaite, alors l'hypothèse H1 et la transformation (4.1) entraînent que tout problème de programmation fractionnaire d.c du type:

$$\min\{F(x) = f(x) / g(x) \quad / x \in S\}$$

peut se transformer en un problème de programmation mathématique d.c du type:

$$\min\{F(y,t) = tf(y/t) / y/t \in S, 1/\alpha_1 \geq t \geq 1/\alpha_2, tg(y/t) \leq 1\}$$

$$\text{où } \alpha_1 \leq \min\{g(x) / x \in S\} \text{ et } \alpha_2 \geq \max\{g(x) / x \in S\}$$

Démonstration

De la même manière que la propriété 2 du lemme 4.1.

Classification du problème de programmation fractionnaire d.c

Théorème général

L'hypothèse H1 et la transformation (4.1) entraînent que:

1-Si les fonctions $f, -g$ (resp. g), $h_i (i=1, \dots, m)$ sont convexes et l'hypothèse H2 est satisfaite (resp. est non satisfaite), alors le problème (P) se ramène à un problème de programmation convexe.

2-Si les fonctions $-g$ (resp. g), $h_i (i=1, \dots, m)$ sont convexes et f est concave alors (P) se transforme à un problème de minimisation d'une fonction concave sur une partie compacte convexe dans le cas où l'hypothèse H2 est satisfaite (resp. dans le cas où l'hypothèse H2 n'est pas satisfaite).

3-Si les fonctions $-g$ (resp. g), $h_i (i=1, \dots, m)$ sont convexes et la fonction f est non convexe mais du type d.c, alors (P) se ramène à un problème de minimisation de la différence de deux fonctions convexes sur une partie compacte convexe dans le cas où l'hypothèse H2 est satisfaite (resp. dans le cas où l'hypothèse H2 n'est pas satisfaite).

4-Si (P) est du type programmation fractionnaire d.c non convexe alors (P) se ramène à un problème de minimisation d'une fonction linéaire sous des contraintes convexes et une contrainte anticonvexe.

Démonstratin

1-, 2-, 3- sont immédiates en vertu des lemmes 4.1, 4.2 et de la proposition précédente.

4- On a déjà montré dans le paragraphe III.3 que tout problème de programmation mathématique d.c non convexe peut être ramené à un problème de la minimisation d'une fonction linéaire sous des contraintes convexes et une contrainte anticonvexe. D'après les lemmes 4.1, 4.2 et la proposition 4.1, on en déduit la propriété 4-.

IV.5. Essais numériques

L'algorithme de maximisation d'une fonction quadratique non convexe sur un polytope programmé en langage Pascal a été mis au point sur Micral 90.50. Nous présentons les résultats expérimentaux de deux exemples où la fonction objectif est une forme quadratique indéfinie.

Exemple 1[4]

$$\begin{aligned} \max f(x) &= -x_1 + 1/2 x_2 + 2/3 x_1^2 - 1/2 x_2^2 \\ \text{sous les contraintes:} \\ 2x_1 - x_2 &\leq 3 \\ x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

Le problème (P') s'exprime sous la forme suivante

$$\begin{aligned} \max 1.5u_1 - 1/2x_1 + 0.25x_2 \\ \text{sous les contraintes:} \\ 2x_1 - x_2 + y_1 &= 3 \\ -2u_1 + 4/3x_1 + v_1 &= 1 \\ u_1 - x_2 + v_2 &= -0.5 \\ u, x, y, v \geq 0 \end{aligned}$$

$$g(u, x, y, v) = \min(u_1, y_1) + \min(x_1, v_1) + \min(x_2, v_2) \leq 0$$

Le maximum global est $x^* = (2.25, 1.5)$ et le point de Kuhn & Tucker associé est $t^* = (u^*, x^*, y^*, v^*) = (1, (2.25, 1.5), 0, (0, 0))$.

Nous avons obtenu, à partir du point de Kuhn & Tucker $t^0 = (0, (0, 0.5), 3.5, (1, 0))$ (où $f(x^0) = 1/8$) et $\beta = 6 \geq \max \{f(x) / x \in D\}$, le tableau de résultats suivant:

k	$E(S^k)=(u,x,y,v,r_1,r_2)=(z,w,r_1,r_2)=(t,r)$	s^*	$g(t)$	$\langle p,z \rangle$
1	$s^1=(18.64,0.0,19.14,22.14,38.28,0.0,-26.75,32.62)$		18.64	
	$s^2=(9.32,32.62,9.82,-52.42,-23.85,0.0,5.87,0.0)$	(*)	-76.28	0.12
	$s^3=(-4.66,0.0,28.46,31.46,-8.32,32.65,5.87,0.0)$		15.48	
2	$s^1=(18.64,0.0,19.14,22.14,38.28,0.0,-26.75,32.62)$		18.64	
	$s^2=(15.87,9.68,16.37,0.0,18.83,0.0,-17.06,22.93)$		9.68	
	$s^3=(-4.66,0.0,28.46,31.46,-8.32,32.65,5.87,0.0)$		15.48	
	$s^4=(0.58,2.04,1.08,0.0,-0.55,0.0,5.87,0.0)$	(*)	-0.55	0.12
	$s^5=(0.58,12.23,21.47,0.0,-14.14,20.38,5.87,0.0)$		6.24	
3	$s^1=(18.64,0.0,19.14,22.14,38.28,0.0,-26.75,32.62,)$		18.64	
	$s^2=(15.87,9.68,16.37,0.0,18.83,0.0,-17.06,22.93)$		9.68	
	$s^3=(0.50,0.0,26.80,29.80,0.0,26.80,0.504,5.82)$		26.30	
	$s^4=(1.00,2.25,1.50,0.0,0.0,0.0,5.25,0.62)$	(*)	0.00	0.75
	$s^5=(6.95,11.17,19.35,0.0,0.0,11.90,-3.65,99.55)$		11.90	
	$s^6=(-0.50,0.0,3.50,6.50,0.0,3.50,5.87,0.0)$		3.00	
	$s^7=(0.37,1.31,0.87,1.25,0.0,0.0,5.87,0.0)$		0.37	

Exemple 2[57]

Nous avons pris le contre-exemple de Zwart [57] pour l'algorithme de Ritter.

$$\max f(x) = 2x_1^2 + x_1x_2 - 10x_1 + 2x_2 - 20$$

sous les contraintes:

$$x_1 + x_2 \leq 11$$

$$1.5x_1 + x_2 \leq 11.4$$

$$x_1, x_2 \geq 0$$

Le problème (P') se formule de la façon suivante:

$$\max 5.5u_1 + 5.7u_2 - 5x_1 + x_2$$

sous les contraintes:

$$\begin{aligned} x_1 + x_2 + y_1 &= 11 \\ 1.5x_1 + x_2 + y_2 &= 11.4 \\ -u_1 - 1.5u_2 + 4x_1 + x_2 + v_1 &= 10 \\ -u_1 - u_2 + x_1 + v_2 &= -2 \\ u, x, y, v &\geq 0 \end{aligned}$$

$$g(u, x, y, v) = \sum_{i=1}^2 \min(u_i, y_i) + \sum_{i=1}^2 \min(x_i, v_i) \leq 0$$

Le maximum global est $x^* = (7.6, 0)$ ($f(x^*) = 19.52$) et le point de Kuhn & Tucker associé est $t^* = (u^*, x^*, y^*, v^*) = \{(0.0, 13.6), (7.6, 0.0), (3.4, 0.0), (0.0, 4.0)\}$.

A partir du vecteur $t^0 = (u^0, x^0, y^0, v^0) = \{(2, 0), (0, 11), (0, 0.4), (1, 0)\}$ et $\beta = 40 \geq \max \{f(x) / x \in D\}$, on obtient.

k	$s^* = (u^*, x^*, y^*, v^*, r_1^*, r_2^*) = (z^*, w^*, r_1^*, r_2^*) = (t^*, r^*)$	$g(t^*)$	$\langle p, z^* \rangle - 20$
1	$(-67.63, 67.2, 0.0, 11, 0.0, 0.4, 32.2, -2.4, 38.0, 0.0)$	-69.6	2
2	$(0.0, 11.6, 0.0, -44.5, 55.6, 55.9, 72.0, 9.7, 38.0, 0.0)$	-32.8	2
3	$(24.2, 0.0, 11.0, 0.0, 0.0, -5.1, -9.8, 11.2, -18.2, 56.2)$	-14.8	58.2
4	$(10.9, 0.0, 7.6, 0.0, 3.4, 0.0, -9.4, 1.3, 38.0, 0.0)$	-6.09	2
5	$(0.0, 13.6, 7.6, 0.0, 3.4, 0.0, 0.0, 4.0, 20.4, 17.5)$	0.0	19.5

CONCLUSION



Le travail de cette thèse consiste, principalement, à étudier des méthodes existantes dans la littérature et à proposer de nouveaux algorithmes pour la résolution d'une classe très large de problèmes d'optimisation non convexe et non différentiable.

Nous avons étudié à l'aide de la théorie de l'analyse convexe (fonctionnelle indicatrice, polaire, sous-différentiel, etc...) les conditions nécessaires d'optimalité de trois types différents de problèmes. Cette étude qualitative permet de justifier et de développer la classe des méthodes de sous-gradients. Nous pensons qu'un effort considérable doit être porté, d'une part, sur les raffinements de ces conditions d'optimalité et, d'autre part, sur leurs utilisations dans les algorithmes de type combinatoire (chapitres II et III).

Après les résultats numériques d'un grand nombre d'algorithmes qui permettent d'obtenir les solutions globales, on constate que certains algorithmes donnent des résultats satisfaisants pour la résolution des problèmes de petite taille. Il reste encore beaucoup à faire dans ce domaine soit pour apporter des améliorations à les algorithmes (surtout dans le cas où l'ensemble admissible est non polyédral), soit pour traiter les problèmes de grande taille.

Dans le but de prouver que les algorithmes que nous avons proposés dans le troisième chapitre peuvent être adaptés à des problèmes particuliers, nous avons étudié la programmation quadratique non convexe. Enfin, Nous avons signalé que les procédés décrits dans cette thèse peuvent être utilisés pour la résolution d'une classe de problèmes de programmation fractionnaire (non convexe et non différentiable).



REFERENCES



- [1] **G.Auchmuty (1983)**
Duality for non-convex variational principles.
Journal of differential equations ,vol 50 ,p 80-145.

- [2] **A.Auslender (1976)**
Optimisation -méthodes numériques .Masson

- [3] **E.Balas (1971)**
Intersection Cuts,-A new type of cutting planes for integer programming -.Oper.Res.,Vol 19 ,p 19-39

- [4] **E.Balas &C.A.Burdet (1975)**
Nonconvex quadratic programming via generalized polars
SIAM J.Appl.Math.,Vol 28 , p 325-349

- [5] **R.Benacer & Pham Dinh Tao (1985)**
Linear programs with reverse convex constraints .Submitted to Math.Programming .

- [6] **R.Benacer & Pham Dinh Tao (1985)**
Two general algorithms for solving linear programming with an additional reverse convex constraint. Submitted to Mathematical Programming Study (devoted to the analysis and optimization of d.c.functions)edited by J.-B.Hiriart Urruty and H.Tuy.

- [7] **R.Benacer & Pham Dinh Tao (1985)**
Global maximization of a nondefinite quadratic function over a convex polyhedron. Submitted to JOURNEES FERMAT 85 . Mathematics for Optimization .Math .Studies .North-Holland.

- [8] **C.Berge (1959)**
Espaces Topologiques et fonctions multivoques .

- [9] **V.Cabot & L.Francis (1974)**
Variation on a cutting plane method for solving concave minimization problems with linear constraints. Naval Research Logistics Quaterly ,Vol 21,p 265-274

- [10] **M.E.Deyer & L.G.Proll.(1977)**
An algorithm for determining all extreme points of a convex polytope. Math.Programming.,Vol 12,p 81-96.

- [11] **I.Ekeland (1979)**
Non convex duality. Analyse non convexe (1977.Pau)
Bull. Soc. Math. France. Mémoire 60, p 45-55.
- [12] **R. Ellala (1984)**
Contribution à l'analyse et l'optimisation de différence de
fonctions convexes .Thèse 3^{eme} cycle . Toulouse .
- [13] **J.E.Falk & K.K.Hoffman (1976)**
A successive underestimation method for concave minimization.
Math. Oper. Res. Vol 1,p 251-259
- [14] **E.Glover (1973)**
Convexity cuts and cut serach. Oper. Res,Vol 21,p 123-134.
- [15] **R.J.Hillestad & S.E.Jacobsen (1980)**
Reverse convex programming . Appl. MATH. Optim. ,Vol 6 ,p 63-78.
- [16] **R.J.Hillestad & S.E.Jacobsen (1980)**
Linear Programs with an additional reverse convex constraint.
Appl.Math.Optim. Vol 6 ,p 257-269 .
- [17] **J.-B.Hiriart Urruty (1985)**
Generalized differentiability,duality and optimization for problems
dealing with differences of convex functions.
Lecture Note in Economics and Math. Systems 256, 37-70.
- [18] **R.Horst (1976)**
An algorithm for nonconvex programming problems .Math
Programming. Vol 10, p312-321 .
- [19] **T.Ibaraki & S.Schaible (1983)**
Invited review ."Fractional programming ".
Europ. Jour .Oper . Res ,Vol 12, p 325-338
- [20] **J.E.Kelley (1960)**
the cutting plane method for solving convex programs.
J.Soc. Ind.Appl. Math. Vol 8,p 703-712 .

- [21] **H.Konno (1976)**
Maximisation of a convex quadratic function under linear constraints .Math .Programming .,Vol 11,p 117-127
- [22] **P.-J.Laurent (1972)**
Approximation et optimisation .Hermann ,Paris .
- [23] **O.L.Mangasarian (1969)**
Nonlinear Programming .Mc.Graw -will ,New York .
- [24] **B.Martinet (1972)**
Algorithmes pour la resolution de problèmes d'optimisation et de minimax .Thèse d'Etat ,USMG (Grenoble) .
- [25] **A.Mat jhay &A.Whinston (1974)**
Quasiconcave minimization subject to linear constraints .Discrete .Math. Vol 9,p 35-59 .
- [26] **B.N.Pchéritchny (1971)**
In contrôle optimal et jeux différentiels, Chiers de l'I.R.I.A. N 4
- [27] **B.N.Pchéritchény & Daniline (1977)**
Méthodes numériques dans les problèmes d'extrémum .Edition Mir.
- [28] **Pham Dinh Tao (1981)**
Contribution à la théorie de normes et ses applications à l'analyse numérique .Thèse d'Etat .USMG (Grenoble).
- [29] **Pham Dinh Tao (1984)**
Algorithmes de calcul du maximum d'une forme quadratique sur la boule unité de la norme du maximum.Num.Math. Vol 45,pp 163-183.
- [30] **Pham Dinh Tao (1984)**
Convergence of a subgradient method for computing the bound norm of matrices .Linea .Alg .Appl ,Vol 62,p 163-182
- [31] **Pham Dinh Tao (1985)**
Algorithms for solving a class of nonconvex optimizations problems.Methods of sous gradients.Submitted to Fermat Days Mathematics for Optimization. Math Studies. North-Holland .

- [32] **M.Raghavachari (1969)**
On connection between zero-one integer programming and concave programming under linear constraints .Oper .Res .Vol 17,p 680-684
- [33] **M.R.Rao (1973)**
Somme comments on "linear" programming with absolute value functionals .Oper .Res .,Vol 21 ,p 373-374
- [34] **R.T.Rockafellar (1970)**
Convex analysis .Princeton U.P.
- [35] **R.T.Rockafellar (1979)**
La théorie des sous gradients et ses applications à l'optimisation.
Les presses de l'Université de Montréal.
- [36] **J.B.Rosen (1983)**
Global minimization of a linearly constrained concave function by partition of feasible domain.
Math. Oper. Res., Vol 8, p 215-230 .
- [37] **P.P.Rough (1979)**
The indefinite quadratic programming problem .
Oper .Res . ,Vol.27 , p 516-533 .
- [38] **M.Sakarovitch (1983)**
Programmation linéaire .E.N.S.I.M.A.G
- [39] **S.Z.Selim & C.M.Shetty (1984)**
Optimization of linear -convex programs .
(A paraitre dans Math.Programming)
- [40] **S.Schaible (1981)**
Fractional programming : Applications and algorithms
Europ .Jou .Oper .Res .,Vol 7,p 111-120
- [41] **S.Schaible (1982)**
Bibliography in fractional programming
Zeits .Oper .res . ,Vol 7 ,(1982) ,p 211-241

- [42] **D.F.Shanno & R .L.Weil (1971)**
Linear programming with absolute value functionals .
Oper .Res. ,Vol 19 ,p 120-124
- [43] **I.Singer (1979)**
A Fenchel-Rockafellar type duality theorem for maximization
Bull. of the Austral .Math Soc . ,V 20 ,p 193-198
- [44] **I.Singer (1980)**
Maximization of lower semi-continuous convex functionals on
bounded subsets of locally convex spaces. II: Quasi-lagrangian
duality theorems. Resultate der Mathematik (3) ,p 235-248 .
- [45] **I.Singer (1983)**
Optimization by level set methods : Duality formulae.
Optimization theory and algorithms .
Lecture notes in pure and applied mathematics ,vol 86.
Edited by J.-B.Hiriart Urruty ,W.Oettli,J.Stoer.
- [46] **H .A.Taha (1973)**
Concave minimisation over a convex polyedron.
Naval Research Logistics Quaterly ,Vol 20,p 533-548
- [47] **Ng.V.Thoai & H.Tuy (1980)**
Convergent algorithms for minimizing a concave function. Math.
Oper. Res. Vol 5,p 556-566
- [48] **J.F.Toland (1978)**
Duality in nonconvex optimisation .
J.Math. Anl .App .,Vol 66, p 399 - 415
- [49] **J.F.Toland (1979)**
A duality principle for non-convex optimization and the calculus of
variations. Arch. Rational. Mech. Analy ,Vol 71,p 41-61
- [50] **D.M.Topkis (1970)**
Cutting plane methods without nested constraint sets .Oper.Res.,
Vol 18, p 404-413 .

- [51] **H.Tuy (1964)**
Concave programming under linear constraints .Doklad.Nauk.Vol 159 ,pp 32-35 ,(1964) English translation in Soviet Mathematics , Vol 5,p 1437-1440 .

- [52] **H.Tuy (1984)**
Global minimization of a difference of two convex functions. Submitted to Math.Programming.Study.

- [53] **H.Tuy (1984)**
Convex programs with an additional reverse convex constraint . Prinprint. Institute of Mathematics ,Hanoi.

- [54] **H.Tuy (1985)**
A general deterministic approach to global optimization via d.c.Programming.Prinprint,Institute of Mathematics,Hanoi.

- [55] **H.Tuy & T.V.Thieu N.Q.Thai (1985)**
A conical algorithm for globally minimizing a concave function over a closed convex set ,Math .Oper .Res . ,p 498-514

- [56] **R.Werner & R.Wcetzal (1985)**
Complementary pivoting algorithms involving extreme rays .Math .Oper .Res . ,Vol 10 .,N 2 ,p 195-206 .

- [57] **P.B.Zwart (1973)**
Nonlinear programming :Counterexamples to two global optimization algorithms .Oper .Res .,N 21,Vol 6 ,p 1960-1966

- [58] **P.B.Zwart (1974)**
Global maximisation of a convex function with linear inequality constraints.Oper .Res .,Vol 22 ,p 602-608

AUTORISATION DE SOUTENANCE

DOCTORAT 3ème CYCLE, DOCTORAT-INGENIEUR, DOCTORAT USMG

Vu les dispositions de l'arrêté du 16 avril 1974,

Vu les dispositions de l'arrêté du 5 juillet 1984,

Vu les rapports de M. *J. P. CROUZEIX*.....

M. *J.-B. HIRIART-URRUTY*.....

M. *B. E. N. A. C. E. R.* *Rachid*..... est autorisé
à présenter une thèse en vue de l'obtention du *Doctorat... U.S.M.G.*.....
...en... Mathématiques... Appliquées.....

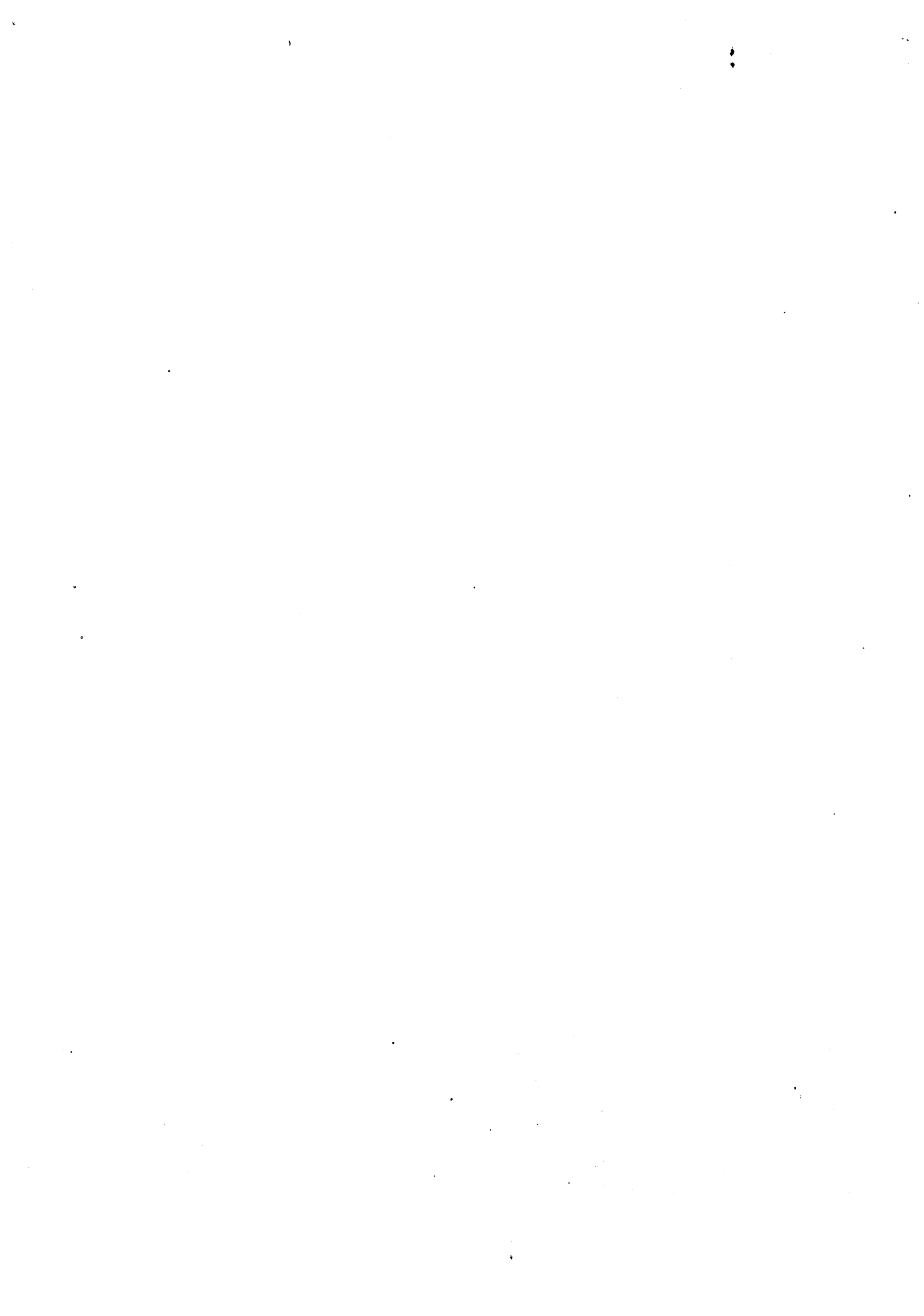
Grenoble, le **10 JUIN 1986**.....

Le Président de l'Université Scientifique
et Médicale



Tanche

M. TANCHE



RESUME

Ce travail concerne l'étude théorique et algorithmique des problèmes d'optimisation non convexes et non différentiables des types suivants :

$$1-\text{Max}\{f(x) / x \in C\}$$

$$2-\text{Min}\{f(x)-g(x) / x \in C\}$$

$$3-\text{Min}\{f(x) / x \in C, g(x) \geq 0\}$$

où f, g sont convexes définies sur \mathbb{R}^n et C est une partie compacte convexe non vide de \mathbb{R}^n

Nous étudions les conditions nécessaires d'optimalité du premier ordre, la dualité, les méthodes de sous-gradients qui convergent vers des solutions optimales locales, et les algorithmes qui permettent d'obtenir les solutions globales (méthodes de coupe, méthodes de partition, méthodes d'approximation extérieure).

Nous donnons quelques résultats numériques et applications des algorithmes présentés.

Mots-clés:

Optimisation non convexe,
Maximisation d'une fonction convexe,
Programmation mathématique d.c.,
Contraintes anti-convexes,
Conditions d'optimalité,
Méthodes numériques.