

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



N°Réf :.....

Centre Universitaire
Abd Elhafid Boussouf Mila

Institut des sciences et de la technologie

Département de Mathématiques et Informatiques

Mémoire préparé en vue de l'obtention du diplôme de Master

en: Mathématiques

Spécialité : Mathématiques fondamentales et appliquées

Une recherche linéaire avec des fonctions majorantes dans certains problèmes d'optimisation

Préparé par : BOUDRAA Abderrahim
CHOUGUI Aziza

Soutenue devant le jury

Encadré par : BAICHE Kanzia.....M.A.B
Président : FADEL Wahida.....M.A.B
Examineur : BEN AOUICHA Lobna.....M.A.B

Année Universitaire : 2015/2016

Remerciement

*Nous remercierons dieu tout puissant pour nous avoir offert la force
et la patience durant toutes ces années.*

*Nous tenons à exprimer un remerciement particulier à notre
encadreur **Baiche Kenzia** pour avoir dirigé ce travail.*

Pour sa présence, son aide et surtout pour ses précieux conseils.

Mille merci nos trop chères familles, que dieu vous garde pour nous.

*Nous tenons à exprimer, nos sincères remerciements à tout le
personnel de l'institut des sciences et de la technologie surtout les
enseignants qui nous ont formé durant les années d'étude, et tous
ceux qui nous ont apporté une aide pour la réalisation de ce projet.*

*Sans oublier bien-sûre tous les amis et collègue d'études pour leur
enjouement et soutien moral.*

Abderrahim, Aziza

Dédicace

La femme La plus chère au monde, ma très chère mère La source de tendresse qui a tout donné sans recevoir (je remercie du fond de mon cœur, merci mamy Halima (fahima) .

Le plus cher homme du monde, Mon père la source de patience, je remercie du fond de mon cœur, merci papy Saleh Rahimaho Allah .
Mes chères sœurs : Ahlem, Amel, Meriem, Chems, Soumia, son mari Fateh et ma fille Djouri, pour ma soeur Fouzia, son mari Mohamed, son fils Anes et sa fille Oumaima, Pour mon mari Hicham, sa famille, mon frère Abdelmoumen Pour son soutien morale et ses sacrifices tout au long de ma formation, Mes grandes familles Ma grande mère Massouda Ma cousine Abla

Mes tantes, leurs Familles, mes oncles et leurs familles
Mes chers amies chacune à son nom Sabah, Sana, Chafiaa, Nafissa et Zineb.

Tous mes collègues de la promotion Mathématique 2014- 2016

*A mon binôme **Abderrahim***
A tous ceux que j'aime ; je dédie ce travail

AZIZA

Dédicace

*Je voudrai dédier cet humble travail à toute ma famille, à mes parents
qui ont veillé à ce que je sois ce que je suis devenu maintenant .*

Ma chère maman qui nous a quittés à jamais et Mon chère Papa que

Dieu le garde en bonne santé.

Mes frère : Yaakoub, Adennour, Roukia, Chaima, Amal, Moussa.

Mes tantes et leurs Familles et mes oncles et leurs familles.

à la famille : Boudraa et Benabiad.

Dédier ce travail à tous mes amis.

Tous mes collègues de la promotion Mathématique 2011- 2016.

*à mon binôme **AZIZA***

Abderrahim

TABLE DES MATIÈRES

Introduction Générale	iii
1 Amélioration des algorithmes projectifs pour la programmation linéaire	1
1.1 Préliminaires	2
1.1.1 Analyse convexe	2
1.1.2 Rappels d'optimisation	3
1.2 Programmation linéaire	8
1.2.1 Formulation d'un programmation linéaire	8
1.3 Dualité en Programmation Linéaire	10
1.3.1 Définition du problème dual	10
1.3.2 Exemple	12
1.4 Résolution d'un programme linéaire	13
1.4.1 L'aidée générale sur La méthode du simplexe	13
1.4.2 Méthodes de points intérieurs	14
2 Méthode projective de point intérieur pour la programmation SDP	20
2.1 Préliminaires matriciels	20
2.1.1 Produit scalaire, norme	21
2.1.2 Matrices (Semi définies positives)	21

2.1.3	Les cône	23
2.1.4	Formulation du problème	26
2.2	Dualité en programmation semi-définie	27
2.3	Exemple de problème SDP	32
2.3.1	Problème de programmation non linéaire	32
2.3.2	Problème de programmation min-max des valeurs propres	32
2.4	Résolution de SDP	34
2.4.1	Fonction barrières	35
3	Résolution d'un programme semi-défini par une approche barrière logarithmique	39
3.1	Étude du problème perturbé (D_r) associé au problème semi-défini (D) . . .	41
3.1.1	Calcul du gradient et du Hessien de la fonction barrière logarithmique f_r	41
3.2	Le problème D_r possède une solution unique	44
3.3	Comportement de la solution optimale $y(r)$ du problème (D_r) lorsque r tend vers 0	46
3.4	Direction de Newton et recherche linéaire	49
3.5	Quelques inégalités utiles	51
3.5.1	Première majorante θ_1	54
3.5.2	Une deuxième majorante θ_2	55
3.5.3	La troisième fonction majorante θ_3	57
3.5.4	Algorithme :	59
3.6	Tests Numériques	59
	Conclusion générale	63
	ANNEXE	64
	Bibliographie	68

INTRODUCTION GÉNÉRALE

Le sujet traité dans ce mémoire concerne le problème de programmation dit d'optimisation sous contraintes semi-définie, encore appelé problème SDP. Cette appellation est une conséquence de la terminologie anglaise SemiDefinite Programming. Ce problème est une extension naturelle de la programmation linéaire dans le sens où les variables sont des matrices symétriques semi-définies positives au lieu d'être des vecteurs. Ce qui explique d'ailleurs le transport du savoir faire de la programmation linéaire à la programmation semi-définie. L'étude de ce problème a commencé dès le début des années 60 en tant que problème mathématique non linéaire, il existe alors très peu de résultats publiés et ils sont principalement d'ordre théorique.

Depuis les années 90, grâce notamment aux travaux fondateurs de ALIZADEH, NEMIROVSKI, NESTEROV [4] entre autres, les méthodes de points intérieurs ont pu être étendues à la résolution de problèmes SDP tout en gardant la plupart des bonnes propriétés qui avaient été observées pour les programmes linéaires. En fait, de nombreux résultats sur les programmes linéaires, notamment en termes de dualité et d'optimalité, ont été étendus aux problèmes SDP.

Il a résulté de tout cela un grand nombre de domaines dans lesquels les problèmes SDP ont trouvé des applications. On cite : Optimisation combinatoire, Optimisation non linéaire, Théorie des graphes, commande robuste, etc,...

Actuellement, le problème (SDP) constitue l'un des sujets de recherche le plus convoité dans le domaine de l'optimisation numérique. Le but étant de développer une méthodologie adéquate. A ce propos, il est impératif de traiter des questions ouvertes, déjà étudiées

au niveau de la programmation linéaire : Initialisation, pas de déplacement, coût excessif de l'itération, qui sont ici plus complexes en raison de la structure complexe du problème.

proposons dans ce mémoire une approche barrière logarithmique dans laquelle, on introduit une procédure originale pour le calcul du pas de déplacement basée sur les fonctions majorants : On obtient une approximation explicite entraînant une décroissance significative de l'objectif, de plus elle est économique et robuste, contrairement aux méthodes classiques de recherche linéaire.

Présentation du mémoire :

Ce mémoire est réparti en trois chapitres, le premier chapitre présente un rappel des notions fondamentales : L'analyse convexe, quelques rappels d'optimisation, et la programmation linéaire. Le deuxième chapitre, nous présentons des notions de la programmation semi définie positive. Enfin, nous terminons le troisième chapitre par l'introduction des fonctions majorantes qui constituent l'originalité de notre étude, ce chapitre est étayé par des expérimentations numériques. Notre travail ouvre des perspectives pour d'autres classes de problèmes avec autres types de méthodes.

CHAPITRE 1

AMÉLIORATION DES ALGORITHMES PROJECTIFS POUR LA PROGRAMMATION LINÉAIRE

Introduction

Dans ce chapitre, nous rappelons brièvement quelques notions fondamentales sur l'analyse convexe, rappel d'optimisation et quelques notions en programmation linéaire.

1.1 Préliminaires

1.1.1 Analyse convexe

Définition 1.1.

Un ensemble $K \subset \mathbb{R}^n$ est dit convexe si :

$$\forall x, y \in K, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in K$$

Autrement dit, un ensemble convexe contient toujours le segment de droite joignant deux points quelconques $x, y \in K$.

Sachant que :

$$[x, y] = \{\lambda x + (1 - \lambda)y \in K / 0 \leq \lambda \leq 1\}$$

Définition 1.2.

Un ensemble $K \subset \mathbb{R}^n$ est dit affine si :

$$\forall x, y \in K, \forall \lambda \in \mathbb{R}, \lambda x + (1 - \lambda)y \in K$$

Remarque 1.1.

Tout ensemble affine est convexe. La réciproque est fausse.

Définition 1.3.

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

$$\forall x, y \in K, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

On dit qu'une fonction f est strictement convexe sur un ensemble convexe K :

$$\forall x, y \in K, x \neq y, \forall \lambda \in]0, 1[, f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Définition 1.4.

La fonction f est dit affine si :

$$\forall \lambda \in \mathbb{R}, f(\lambda x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda)f(y)$$

Définition 1.5.

Un ensemble K est un polyèdre si :

$$K = \{x \in \mathbb{R}^n : p_i^t x \leq \alpha_i, i = 1, \dots, m\}$$

Où p_i est un vecteur non nul de \mathbb{R}^n et α_i est un scalaire pour $i = 1, \dots, m$.

1.1.2 Rappels d'optimisation

Un problème d'optimisation est présenté par la forme suivante :

$$(PC) \begin{cases} \min f(x) \\ x \in C \end{cases} \quad (1.1)$$

1.1.2.1 Optimisation sans contraintes

Un problème d'optimisation sans contrainte est présenté par la forme suivante :

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

► Résultat d'existence-unicité

Théorème 1.1. *Existence*

Si f une fonction propre, continue et coercive sur \mathbb{R}^n Alors (1.2) admet au moins une solution.

Théorème 1.2. *Unicité*

Soit f strictement convexe. Alors la solution optimale de (1.2) si elle existe, est unique.

► Condition d'optimalité

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction deux fois différentiable.

Théorème 1.3. Conditions nécessaires d'optimalité

Soit $x^* \in \mathbb{R}^n$ un minimum local du problème (1.2).

1. **Condition nécessaire de première ordre** si : f est différentiable en x^* alors :

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

2. **Condition nécessaire de second ordre** si : f est deux fois différentiable au point x^* , alors :

$$\nabla^2 f(x^*) = \frac{\partial^2 f(x^*)}{\partial x_i \partial y_i}$$

est une matrice semi-définie positive.

$$\nabla^2 f(x^*) \geq 0 \iff x^t \nabla^2 f(x^*) x \geq 0, \forall x \in \mathbb{R}^n$$

Théorème 1.4. Conditions suffisantes d'optimalité

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^2$. On suppose que :

1. $\nabla f(x^*) = 0$.
2. $\nabla^2 f(x^*)$ est une matrice semi-définie positive, i.e $\forall x \in \mathbb{R}^n, x^t \nabla^2 f(x^*) x > 0$.

Alors x^* est un point minimum local.

Proposition 1.1. Condition nécessaire et suffisantes cas convexe

Soit f une fonction convexe de classe C^2 , définie sur \mathbb{R}^n et x^* un point de \mathbb{R}^n . Alors, x^* est un minimum local (global) de f si et seulement si $\nabla f(x^*) = 0$ et $\nabla^2 f(x^*) > 0$.

1.1.2.2 Optimisation avec contrainte

Un problème d'optimisation avec contrainte est présenté par la forme suivante :

$$(PC)' \begin{cases} \min f(x) \\ x \in C \end{cases} \quad (1.3)$$

Où C est l'ensemble des contraintes.

► Résultat d'existence-unicité

Théorème 1.5. *Existence*

Soit $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue sur un ensemble C , et si C fermé non vide.

On suppose que :

1. C est borné.
2. C et non borné et $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ (on dit alors que f est coercive).

alors f possède au moins un point minimum sur C .

Théorème 1.6. *Unicité*

Soit f une fonction strictement convexe sur C , et C convexe. Le minimum de f sur C , s'il existe, est unique.

► Condition d'optimalité

▷ Condition d'optimalité du première ordre

Considérons le problème (1.3) :

$$C = \{x \in \mathbb{R}^n / g_i(x) = 0, i = 1, \dots, p, \quad h_j(x) \leq 0, j = 1, \dots, q\}$$

- Tel que les contraintes $g_i(x) = 0$ sont appelées "contraintes d'égalité" et les contraintes $h_j(x) \leq 0$ sont appelées "contraintes d'inégalité".
- Un point x de C est une solution réalisable (ou admissible) de (1.3).

Théorème 1.7. *Condition nécessaire du première ordre*

Si f est une fonction différentiable et si C est un convexe fermé, alors toute solution x^* de (1.3) vérifiant les conditions nécessaire d'optimalité de premier ordre :

$$f(x) - f(x^*) \geq \nabla f(x^*)^t(x - x^*), \forall x \in C \quad (1.4)$$

Théorème 1.8. *Condition NS du première ordre dans le cas convexe*

Soit f une fonction convexe différentiable, et C convexe fermé de \mathbb{R}^n , soit x^* un point de C . La condition (1.4) est nécessaire et suffisante pour x^* est une solution de (1.3). x^* un point minimum de f sur C .

► **Condition de qualification**

Nous présentons trois conditions de **qualification** ou de **régularité**.

★ (CQ1) *C* est un polyèdre convexe :

C'est le cas lorsque les fonctions g_i et h_j sont affines.

★ (CQ2) **condition de Slater :**

La condition de qualification des contraintes de Slater est comme suit :

- 1) Les fonctions g_i sont convexes et les fonctions h_j sont affines.
- 2) Il existe \hat{x} tel que $g_i(\hat{x}) = 0$ et $h_j(\hat{x}) < 0; \forall i, j$.

★ (CQ3) **condition de Mangasarian-Fromowitz :**

Définition 1.6. (Régularité)

On dit que le point x^* est régulier pour les contraintes g et h si :

- x^* est réalisable $g(x^*) = 0, h(x^*) \leq 0$.
- Et si les vecteurs $\nabla g_i(x^*), \nabla h_j(x^*), 1 \leq i \leq p, j \in I(x^*)$, sont linéairement indépendants.
- Il existe $d \neq 0 \in \mathbb{R}^n$ tel que $\langle \nabla g_i(x^*), d \rangle = 0$ pour $i = 1, \dots, p$ et $\langle \nabla h_j(x^*), d \rangle < 0$ pour $\forall j \in I(x^*)$.

On assoie le problème suivant :

$$\left\{ \begin{array}{l} \min f(x) \\ g_i(x) = 0, i = 1, \dots, p \\ h_j(x) \leq 0, j = 1, \dots, q \\ \mu_j \geq 0, j = 1, \dots, q \\ \mu_j h_j(x) = 0, j = 1, \dots, q \end{array} \right.$$

La fonction $L : \mathbb{R}^n \times [0, \infty[^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ définie par :

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \mu_j h_j(x)$$

est appelée le " LAGRANGIEN ".

► Théorème de KKT

Le théorème suivant de **Karush-Kuhn-Tucker-Lagrange** donne une condition nécessaire d'optimalité :

Théorème 1.9.

Supposons que les fonctions f, g_i, h_j sont C^1 dans un voisinage de $x^* \in C$ et que les contraintes vérifient une condition de qualification. Si f est un minimum local en x^* sur C alors $\forall \lambda^* \in \mathbb{R}^p$ et $\exists \mu^* \in \mathbb{R}^q$ tel que :

$$\left\{ \begin{array}{l} \nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^q \mu_j^* \nabla h_j(x^*) = 0 \\ \nabla_\lambda L(x^*, \lambda^*, \mu^*) = g_i(x^*) = 0 \\ \nabla_\mu L(x^*, \lambda^*, \mu^*) = h_j(x^*) \leq 0 \\ \mu_j^* \geq 0, \quad \mu_j^* \nabla_\mu L(x^*, \lambda^*, \mu^*) = \mu_j^* h_j(x^*) = 0 \end{array} \right.$$

Les quantités λ_i et μ_j sont appelées **multiplicateur de Karush-Kuhn-Tucker**.

Lorsque le problème est convexe (f et g_i convexes et h_j affines) on a la condition suffisante d'optimalité.

Théorème 1.10.

Supposons que les fonctions f, g_i sont C^1 dans un voisinage de $x^* \in C$, convexes et que les fonctions h_j sont affines. Si $\exists \lambda^* \in \mathbb{R}^p$ et $\mu^* \in \mathbb{R}^q$ tel que :

$$\left\{ \begin{array}{l} \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^q \mu_j^* \nabla h_j(x^*) = 0 \\ g_i(x^*) = 0 \\ \mu_j^* \geq 0, \quad h_j(x^*) \leq 0, \quad \mu_j^* h_j(x^*) = 0 \end{array} \right.$$

alors f est un minimum local en x^* sur C .

1.2 Programmation linéaire

La programmation linéaire constitue la pierre angulaire de toute la recherche opérationnelle. Il faut bien sûr éviter de forcer tout modèle à être linéaire. Par contre, un très grand nombre de modèles constituent des extensions de programmes linéaires. Elle peut se définir comme une technique mathématique permettant de résoudre des problèmes de gestion et particulièrement ceux où le gestionnaire doit déterminer, face à différentes possibilités, l'utilisation optimale des ressources de l'entreprise pour atteindre un objectif spécifique comme la maximisation des bénéfices ou la minimisation des coûts.

1.2.1 Formulation d'un programmation linéaire

Un programme linéaire (PL) peut être écrit sous l'une des deux formulations suivantes :

a) **Forme canonique :**

$$\left\{ \begin{array}{l} \min c^t x = z \\ Ax \geq b \\ x \geq 0 \end{array} \right. \quad (1.5)$$

où A est une matrice ($m \times n$), $c \in \mathbb{R}^n$ et x un vecteur inconnu dans \mathbb{R}^n et $b \in \mathbb{R}^m$.

De nombreux modèles d'application se ramènent de manière naturelle à cette forme qui se prête mieux à l'analyse théorique.

b) **Forme standard :**

$$\left\{ \begin{array}{l} \min c^t x = z \\ Ax = b \\ x \geq 0 \end{array} \right. \quad (1.6)$$

C'est la forme à laquelle on se ramène le plus souvent à cause de sa convenance pour l'analyse théorique et le traitement numérique.

Tell que pour les deux formes de PL est :

- $b \in \mathbb{R}^m$, est le second membre, avec $b_i \geq 0, \forall i = 1, \dots, m$.
- $c^t x = \sum_{j=1}^n c_j x_j$, est la fonction objectif .

1.2.1.1 Exemple

Un menuisier désire fabriquer des tables et des chaises et des tabourets. Il dispose de 50 unités de bois et 90 heures de travail. La fabrication d'une table nécessite 10 unités de bois et 15 heures de travail. Celle d'une chaise, 5 unités de bois et 10 heures de travail. Cependant la fabrication d'un tabouret 1 unité de bois et 2 heures de travail. Une table vendue procure au menuisier un profit de 1000 DA ; une chaise, un profit 400 DA. La vente d'un tabouret procure au menuisier un bénéfice de 250 DA. Combien de tables, de chaises et de tabourets doit-il alors fabriquer pour maximiser son profit ?

Formulation mathématique de l'exemple

Pour set exemple, on identifie trois variables X_1, X_2 et X_3 .

X_1 : nombre de tables

X_2 : nombre de chaises

X_3 : nombre de tabourets

Il y a une contrainte rattachée au bois et une autre au temps de travail, d'où les inégalités :

$$\begin{cases} 10X_1 + 5X_2 + X_3 \leq 50 \\ 15X_1 + 10X_2 + 2X_3 \leq 90 \end{cases}$$

Les variables sont non négatives :

$$X_1 \geq 0, X_2 \geq 0 \quad \text{et} \quad X_3 \geq 0$$

La fonction objective F est à maximiser et s'écrit :

$$\max F = 1000X_1 + 400X_2 + 250X_3$$

Comme la fonction objective s'exprime par une équation linéaire et les contraintes s'expriment par des inéquations linéaires par rapport aux variables de décision, ce problème est aussi un modèle de PL.

$$\max F = 1000X_1 + 400X_2 + 250X_3$$

Sous les contraintes :

$$\left\{ \begin{array}{l} 10X_1 + 5X_2 + X_3 \leq 50 \\ 15X_1 + 10X_2 + 2X_3 \leq 90 \\ X_1 \geq 0, X_2 \geq 0 \text{ et } X_3 \geq 0. \end{array} \right.$$

Dans cet exemple, on a trois variables de décision ($n = 3$) et deux contraintes ($m = 2$).

1.3 Dualité en Programmation Linéaire

La dualité est une notion essentielle en programmation mathématique en particulier en programmation linéaire. A un problème de PL appelons le primal. l'opération de dualité associe un autre problème de PL son dual qui ne sera défini qu'à l'aide des seules données du problème primal. La dualité présente un double intérêt :

- d'autre part, le problème dual a une interprétation économique importante : il constitue une autre vision du problème initial, le primal.
- d'autre part, des propriétés mathématiques simples est fortes , lient les deux problèmes primal et dual.

1.3.1 Définition du problème dual

La dualité associé à tout problème linéaire un autre problème linéaire qui est appelé problème dual du problème initial : par opposition le problème initial est appelé problème primal.

Soit le programme linéaire écrit sous forme standard :

$$(P) \begin{cases} \min c^t x \\ Ax = b \\ x \geq 0 \end{cases}$$

On appelle dual (P), le programme linéaire :

$$(D) \begin{cases} \max b^t y \\ A^t y \leq c \\ y \in \mathbb{R}^m \end{cases}$$

qui est équivalent au problème suivant :

$$(D) \begin{cases} \max b^t y \\ A^t y + s = c \\ s \geq 0 \\ y \in \mathbb{R}^m \end{cases}$$

où s désigne une variable d'écart. Les ensembles des solutions réalisables de (PL) et (D) seront notés respectivement F_P et F_D . Ainsi, on a :

$$F_P = \{x \in \mathbb{R}^n, Ax = b, x \geq 0\}$$

$$F_D = \{(y, s) \in \mathbb{R}^m \times \mathbb{R}^m, A^t y + s = c, s \geq 0\}$$

- Si l'un des problèmes (PL) ou (D) possède une solution optimale fini, il en est de même pour l'autre et leurs valeurs optimales sont égales.
- Si l'un des problèmes (PL) ou (D) est non borné, alors le domaine réalisable de l'autre est vide.

1.3.2 Exemple

Soit le programme linéaire primal :

$$(P) \begin{cases} \min(2x_1 - 3x_2) \\ x_1 - x_2 \leq 1 \\ 2x_1 + 3x_2 \geq 4 \\ x_1 + x_2 = 3 \\ x_1 \geq 0, x_2 \in \mathbb{R} \end{cases}$$

Le dual de ce programme s'écrit :

$$(D) \begin{cases} \max(y_1 + 4y_2 + 3y_3) \\ y_1 + 2y_2 + y_3 \leq 2 \\ -y_1 + 3y_2 + y_3 = -3 \\ y_1 \leq 0, y_2 \geq 0, y_3 \in \mathbb{R} \end{cases}$$

Remarque 1.2.

1. Le dual du dual est le primal.
2. Il est utile d'écrire directement le dual d'un programme linéaire mais sous forme standard (le transformer en forme canonique). Cette définition de dualité s'étend donc à travers ces transformations à tout programme linéaire primal écrit ou non sous forme canonique.

Théorème 1.11. (Dualité faible)

Si x et y sont respectivement des solutions réalisables pour (P) et (D) alors

$$c^t x \geq b^t y$$

Théorème 1.12. (Dualité forte)

Si x^* et y^* sont respectivement des solutions réalisables pour (P) et (D) telles que

$$c^t x^* = b^t y^*$$

alors x^* et y^* sont des solutions optimales pour leurs problèmes respectifs (P) et (D) .

Théorème 1.13. (Théorème de dualité)

- Si l'un des problèmes (P) et (D) admet une solution optimale, il en est de même pour l'autre et leur valeurs optimales correspondantes sont égales.
- Si l'un des deux problèmes à une valeur optimale non bornée, l'autre n'a pas de solution optimale.

1.4 Résolution d'un programme linéaire

La solution d'un programme linéaire est toujours atteinte en un sommet (point extrémal) de la région admissible, c-à-d : un point réalisable satisfaisant exactement (m) contraintes linéairement indépendantes.

1.4.1 L'aidée générale sur La méthode du simplexe

Cette méthode est une procédure qui examine les points extrémaux (réalisables) d'une façon convenable, le passage d'un sommet à un autre entraîne souvent une diminution de l'objectif si le problème est un problème de minimisation.

La résolution d'un programme linéaire par cette méthode est traditionnellement divisée en deux phases :

- **Etape1** : Est l'étape d'initialisation, elle consiste à chercher un premier sommet réalisable, et peut être posée (de plusieurs manières) sous forme d'un programme linéaire possédant une solution initiale triviale.
- **Etape2** : Construit une suite de sommets adjacents réalisables diminuant la valeur de l'objectif d'une itération à l'autre. Le nombre de sommets étant fini, la convergence de la méthode du simplexe vers l'optimum est garantie en un nombre fini d'opérations à condition que le problème ne soit pas dégénéré (i.e : aucune composante de la solution réalisable de base n'est pas nulle).

1.4.2 Méthodes de points intérieurs

Les méthodes de points intérieurs comme leur nom l'indique éviteront soigneusement la frontière de l'ensemble admissible et ne souffriront donc pas de l'aspect combinatoire inhérent à la méthode du simplexe. Khachiyan (1979) fut le premier à proposer un algorithme dit polynomial, c'est-à-dire dont le nombre d'itérations grandit d'une manière polynomiale avec la taille du problème. Malheureusement, cet algorithme s'est avéré inefficace en pratique, il a fallu donc attendre les travaux de Karmarkar (1984) pour susciter un engouement pour les méthodes de points intérieurs. Actuellement, l'importance de ces méthodes a dépassé le cadre de l'optimisation linéaire et elles sont beaucoup plus utilisées dans le cadre de l'optimisation convexe, grâce notamment aux travaux de Nesterov et Nemirovsky (1994). Actuellement, les méthodes de points intérieurs sont devenues compétitives à la méthode du simplexe surtout pour les problèmes de grande taille (> 1000 variables ou contraintes).

Il y a, principalement, quatre grandes catégories des méthodes de points intérieurs :

- Les méthodes affines.
- Les méthodes projectives.
- Les méthodes de réduction du potentiel.
- Les méthodes de trajectoire centrale.

1.4.2.1 Méthode projective version Ye-Lustig

Malgré sa complexité exponentielle, l'algorithme du Simplexe est resté longtemps l'algorithme de référence pour la programmation linéaire. La méthode est demeurée sans compétiteurs sérieux jusqu'en 1984 où un nouvel algorithme polynomial, l'algorithme des méthodes projectives a été proposé par Narendra Karmarkar aux laboratoires de AT et T Bell aux Etats Unis. Mais la mise en place et les transformations nécessaires pour résoudre un problème standard de programmation linéaire par la méthode de Karmarkar sont demeurées secrètes jusqu'en 1985. Cet algorithme a été essentiellement différent de la méthode du simplexe par le fait que, dans cette dernière, on se déplace en suivant la frontière du domaine réalisable, alors que pour la méthode de Karmarkar, on progresse

tout en restant strictement à l'intérieur du domaine réalisable, d'où la connotation "intérieur", et c'était grâce à lui qu'une recherche intense dans les méthodes du points intérieurs s'est déclenchée et a donné comme résultat une grande variété d'algorithmes de ce type. L'algorithme de Karmarkar avait un grand intérêt théorique puisqu'il est polynomial, en effet, il exige un $o(n^3L)$ opérations arithmétiques, d'où une complexité totale de $o(n^4L)$ opérations. Karmarkar a même été capable de la réduire à $o(n^{3,5}L)$ en utilisant un processus de mise à jour partiel.

Ces méthodes projectives n'ont pas été présentées seulement par Karmarkar mais aussi par Todd et Burrell, Anstreicher, Gay et Ye et Kojima. Ces chercheurs ont suggéré des modifications de l'algorithme original de Karmarkar afin d'aboutir à un algorithme très efficace et plus utile dans la pratique. Ces méthodes ont été mise en œuvre hors des laboratoires Bell par Adler et al, Gill et al, Monma et Morton et Mcshane et al, ces différentes mise en œuvre ont donné de bons résultats prouvant leur supériorité par rapport aux autres méthodes (simplexe, ellipsoïde ...).

Dans cette partie on étudie l'approche de Ye-Lustig.

L'approche de Ye-Lustig.

Soit le programme linéaire sous forme standard :

$$PL \left\{ \begin{array}{l} \min c^t x = z^* \\ Ax = b \\ x \geq 0 \end{array} \right.$$

où : c, x de \mathbb{R}^n , $A \in \mathbb{R}^{m \times n}$.

avec l'hypothèse suivante :

H1) La matrice des contraintes A est de plein rang ($\text{rg}(A) = m < n$).

Dans cette méthode, à chaque itération k , on utilise une transformation projective qui ramène la région admissible polyédrique $\{Ax = b, x \geq 0\}$ à un simplexe

$$S_{n+1} = \left\{ x \in \mathbb{R}_+^{n+1}, \sum_{i=1}^{n+1} x_i = 1 \right\}.$$

❖ CHAPITRE 1. AMÉLIORATION DES ALGORITHMES PROJECTIFS POUR LA PROGRAMMATION LINÉAIRE

Cette transformation est définie par :

$$T_a : \mathbb{R}_+^n \longrightarrow S_{n+1}$$

$$T_a(x^k) = y = \begin{cases} y_i = \frac{\frac{x_i^k}{a_i}}{1 + \sum_{i=1}^n \frac{x_i^k}{a_i}}, & i = 1, \dots, n \\ y_{n+1} = 1 - \sum_{i=1}^n y_i, & a \in \mathbb{R}_+^n \end{cases}$$

Le problème (PL) devient alors :

$$P_t \begin{cases} \min(D_k c, -z^*)^t y \\ A_k y = 0 \\ y \in S_{n+1} = \left\{ y \in \mathbb{R}^{n+1}, y \geq 0, \sum_{i=1}^{n+1} y_i = 1 \right\}. \end{cases}$$

Où $A_k = [AD_k - b] \in \mathbb{R}^{(n+1) \times m}$, $D_k = \text{diag}(x^k)$, matrice diagonale.

Comme le calcul d'une solution optimale d'un programme linéaire, sur une sphère est évident, on introduit à chaque itération la plus grande sphère inscrite dans le simplexe S_{n+1} .

On obtient le sous problème de P_t suivant :

$$P_s \begin{cases} \min(D_k c, -z^*)^t y \\ A_k y = 0 \\ \|y - e_{n+1}\|^2 \leq r^2 < 1 \end{cases}$$

Où $r = \frac{1}{\sqrt{n(n+1)}}$, est le rayon de la sphère.

$\frac{e_{n+1}}{n+1} = (\frac{1}{n+1}, \dots, \frac{1}{n+1})^t \in \mathbb{R}^{n+1}$, est le centre de la sphère.

Comme la valeur de z^* est généralement inconnue, Ye-Lustig approxime z^* chaque itération par des bornes supérieures $z^k = c^t x^k > z^*$.

Le problème P_s est remplacé par :

$$P_k \begin{cases} \min(D_k c, -c^t x^k)^t y \\ A_k y = 0 \\ \|y - e_{n+1}\|^2 \leq r^2 < 1 \end{cases}$$

En utilisant les conditions d'optimalités de KKT relatives au problème (P_k) , la solution optimale est donnée par :

$$y_k^* = \frac{e_{n+1}}{n+1} - \alpha^k r d^k$$

où α^k est le pas de déplacement, $0 < \alpha^k < 1$.

$d^k = \frac{p^k}{\|p^k\|}$, p^k est la projection du vecteur coût $(D_k c, -c^t x^k)^t$ sur le noyau de la matrice A_k .

$$P^k = (I - A_k^t (A_k A_k^t)^{-1} A_k) (D_k c, -c^t x^k)^t$$

La vitesse de convergence de cette méthode dépend du calcul de la projection p^k , et aussi du pas de déplacement.

Calcul de la projection

Nous avons

$$P^k = (I - A_k^t (A_k A_k^t)^{-1} A_k) (D_k c, -c^t x^k)^t$$

Posons

$$u_k = (A_k A_k^t)^{-1} A_k (D_k c, -c^t x^k)^t$$

donc

$$P^k = (D_k c, -c^t x^k) - A_k^t u_k$$

le calcul de p^k dépend ainsi de la manière par laquelle, on résout le système linéaire :

$$A_k A_k^t u_k = A_k (D_k c, -c^t x^k)^t$$

dont la matrice $A_k A_k^t$ est symétrique définie positive.

H2)

Initialisation

$x^0 > 0, k = 0.$

Tan que $\frac{\|d^k\|}{|c^t x^0|} > \varepsilon$ **faire**

1. Construire $D_k = \text{diag}(x^k), A_k = [AD_k, -b], B_k = [A_k, e_n]^T.$
2. Calcul de la projection p_k

$$P_k = [I - B_k^t (B_k B_k^t)^{-1} B_k] (D_k c, -c^t x^k)^t$$

3. Normaliser la projection p_k

$$d^k = \frac{p^k}{\|p^k\|}, r = \frac{1}{\sqrt{n(n+1)}}$$

4. Calculer l'itéré suivant

$$y^k = \frac{e_{n+1}}{n+1} - \alpha^k r d^k, \alpha^k \text{ le pas de déplacement}$$

5. Revenir au problème original par T_k^{-1}

$$x^{k+1} = T_k^{-1}(y^k) = \frac{D_k y^k [n]}{y_{n+1}^k}$$

$$k = k + 1$$

Fin tant que

Fin algorithme.

Calcul d'une solution initiale strictement réalisable

Un point strictement réalisable de (PL) est une solution du problème :

$$(PF) \begin{cases} Ax = b \\ x > 0 \end{cases}$$

En utilisant la technique de la variable artificielle, le problème (PF) est équivalent au problème d'optimisation suivant :

$$(P\lambda) \begin{cases} \min \lambda \\ Ax + \lambda q = b \\ x > 0, \lambda \geq 0 \end{cases}$$

tel que : $q = b - Aa$ et $a \in \mathbb{R}_+^n$.

(P_λ) est équivalent au programme linéaire :

$$(\tilde{P}\tilde{L}) \begin{cases} \min \tilde{c}^t \tilde{x} = w^* = 0 \\ \tilde{A}\tilde{x} = b \\ \tilde{x} > 0 \end{cases}$$

où $\tilde{x} = (x, \lambda)^t$, $\tilde{c} = (0, 0, 0, \dots, 1)^t \in \mathbb{R}^{n+1}$, $\tilde{A} = [A, b - Aa] \in \mathbb{R}^{m \times (n+1)}$.

Étant donné que $(a, 1)$ est une solution strictement réalisable de $(\tilde{P}\tilde{L})$ pour tout $a > 0$. Alors le calcul de la solution optimale de $(\tilde{P}\tilde{L})$ est donné par la phase 2 de l'algorithme de Ye-Lustig. L'algorithme correspondant se présente comme suit :

Algorithme d'initialisation

Initialisation

$x^0 = a, \lambda^0 = 1, \tilde{x} = (x^0, \lambda^0)$ et $k = 0$

1) Si $\|Ax^0 - b\| \leq \varepsilon$ Stop ; x^k est une solution approximative de (PF).

Sinon aller à l'étape (2)

2) Si $\lambda^k \leq \varepsilon$ Stop ; x^k est une solution approximative de (PF).

Sinon aller à l'étape (3)

3) Poser $D_k = \text{diag}(\tilde{x}^k)$, $B = [A, b - Aa]$ et $r = \frac{1}{\sqrt{(n+1)(n+2)}}$

Calculer $p_k = [I - B_k^t (B_k B_k^t)^{-1} B_k] (D_k \tilde{c}, \tilde{c}^t \tilde{x}^k)^t$

tel que $B_k = [BD_k, -b]$ et $\tilde{c} = (0, 0, 0, \dots, 1)^t \in \mathbb{R}^{n+1}$

calculer $y^{k+1} = \frac{e_{n+2}}{n+2} - \alpha^k r \frac{p_k}{\|p_k\|}$, $\tilde{x}^{k+1} = (y_{n+2}^{k+1})^{-1} D_k y^{k+1} [n+1]$

4) Faire $k = k + 1$ et retourner à l'étape (2).

Fin

CHAPITRE 2

MÉTHODE PROJECTIVE DE POINT INTÉRIEUR POUR LA PROGRAMMATION SDP

Introduction

Dans ce chapitre, nous rappelons quelques définitions de produit scalaire, les normes, et effectuons quelques propriétés des matrices symétriques. En fin nous étudions la programmation semi-définie positive.

2.1 Préliminaires matriciels

Notation : Soit

- $M_n(\mathbb{R}) = M_n$, l'ensemble des matrices carrées de taille n à coefficients dans \mathbb{R} .
- $S_n(\mathbb{R}) = S_n$, l'ensemble des matrices symétriques réelles de taille n c-à-d : les matrices $A \in M_n$ vérifiant $A = A^t$.

2.1.1 Produit scalaire, norme

Si $A, B \in M_n$ le produit scalaire de A et B , noté $A \bullet B$ est défini par :

$$A \bullet B = \langle A, B \rangle = \mathbf{Tr}(A^t B) = \sum_{i=1}^m \sum_{j=1}^m a_{ij} b_{ij} = B \bullet A$$

Où la trace $\mathbf{Tr}(A) = \sum_{i,j=1}^m a_{ij}$ est la somme des éléments diagonaux de A .

On la propriété importante suivante :

$$\forall A, B \in M_n(\mathbb{R}), \mathbf{Tr}(AB) = \mathbf{Tr}(BA)$$

A ce produit scalaire, on associe une norme dit de Frobenius :

$$\|A\|_F = \|A\| = \sqrt{A \bullet A}$$

On utilisera également la norme spectrale :

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^t A)}$$

Pour une matrice symétrique A , on obtient facilement les résultats suivants :

$$\begin{cases} \|A\|_F = \max |\lambda_i(A)| & (\text{rayon spectral de } A) \\ \|A\|_2 = \left(\sum_{i,j=1}^m a_{ij}^2 \right)^{\frac{1}{2}} & = \left(\sum_{i=1}^m \lambda_i(A)^2 \right)^{\frac{1}{2}} \end{cases}$$

2.1.2 Matrices (Semi définies positives)

Définition 2.1.

Soit $A \in S_n(\mathbb{R})$

1. A est dite semi-définie positive, et l'on notera $A \in S_n^+(\mathbb{R})$ ou $A \geq 0$ si

$$\forall x \in \mathbb{R}^n, x^t A x \geq 0.$$

2. A est dite définie positive, et l'on notera $A \in S_n^{++}(\mathbb{R})$ ou $A > 0$ si

$$\forall x \in \mathbb{R}^n \setminus \{0\}, x^t A x > 0.$$

Lemme 2.1.

On considère des matrices symétriques $A_i \in S_{n_i}(\mathbb{R})$. Alors la matrice diagonale par blocs

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_m \end{bmatrix}$$

est définie positive (resp semi-définie positive) si et seulement si chacune des A_i est positive.

Notation :

1. $S_n^+(\mathbb{R}) = S_n^+ = \{A \in S_n / A \succeq 0\}$.
2. $S_n^{++}(\mathbb{R}) = S_n^{++} = \{A \in S_n / A \succ 0\}$.

Proposition 2.1.

Soit $B \in M_n$ inversible, alors :

$$A \in S_n^+ \Leftrightarrow B^t A B \in S_n^+.$$

et

$$A \in S_n^{++} \Leftrightarrow B^t A B \in S_n^{++}.$$

Théorème 2.1.

Soit $A \in S_n(\mathbb{R})$. Les conditions suivantes sont équivalentes :

1. $A \in S_n^+(\mathbb{R})$ (resp. $A \in S_n^{++}(\mathbb{R})$).
2. $\lambda_{\min}(A) \geq 0$ (resp. $\lambda_{\min}(A) > 0$).
3. $\exists P \in M_n(\mathbb{R}), A = P^t P$ (resp. $\exists P \in M_n(\mathbb{R}), \text{rg}(P) = n$ tel que $A = P^t P$).

Théorème 2.2 (Complément de Schur).

Soient $A \in S_m^{++}(\mathbb{R}), C \in S_n(\mathbb{R})$ et $B \in M_{m,n}(\mathbb{R})$. On a alors les équivalences suivantes :

$$\begin{bmatrix} A & B \\ B^t & C \end{bmatrix} > 0 \Leftrightarrow C - B^t A^{-1} B > 0$$

et

$$\begin{bmatrix} A & B \\ B^t & C \end{bmatrix} \geq 0 \Leftrightarrow C - B^t A^{-1} B \geq 0$$

2.1.3 Les cônes

2.1.3.1 Le cône convexe

Définition 2.2.

Un ensemble $S_n = \mathbb{R}_+^n$ est un cône si :

$$x, y \in S_n \Rightarrow \forall \lambda \geq 0 : \lambda(x + y) \in S_n$$

Un cône est dit pointé si $S_n \cap (-S_n) = \{0_n\}$.

Proposition 2.2.

1. S_n^+ un cône fermé pointé dans S_n si $S_n^+ \cap (-S_n^+) = \{0_n\}$ (0_n une matrice nulle).
2. S_n^{++} n'est pas un cône car $0_n \notin S_n^{++}$.
3. S_n^{++} est l'intérieur de S_n^+ (c'est l'intérieur relatif).

On rappelle que $S_n \subset \mathbb{R}^n$, $S_n \neq \emptyset$, l'intérieur relatif de S_n est :

$$ri(S_n) = \{aff(S_n) / \exists \varepsilon > 0, (x + \varepsilon B) \cap aff(S_n) \subseteq S_n\}$$

et tel que :

$$aff(S_n) = \{x \in \mathbb{R}^n / x = \sum_{i=1}^n \lambda_i x_i, x_i \in S_n, \sum_{i=1}^n \lambda_i = 1\}$$

B est la boule unité euclidienne de \mathbb{R}^n :

$$B = \{x \in \mathbb{R}^n / \|x\| \leq 1\}$$

et si

$$aff(S_n) = \mathbb{R}^n \Rightarrow ri(S_n) = int(S_n) / int(S_n) = \{x \in S_n / \exists \varepsilon > 0, x + \varepsilon B \subseteq S_n\}$$

et S_n un convexe de \mathbb{R}^n .

4. La structure du cône de S_n^+ induit un ordre partiel sur S_n dit : “ Ordre de Löwner ” :

$$A \succeq B \Leftrightarrow A - B \in S_n^+(\mathbb{R})$$

$$A \succ B \Leftrightarrow A - B \in S_n^{++}(\mathbb{R})$$

(\succ, \succeq : relation d'ordre partiel).

Proposition 2.3.

Soient $A, B \in S_n^+$. Alors :

1. $A + B \geq B$.
2. $A^{\frac{1}{2}}BA^{\frac{1}{2}} \geq 0$.
(si $A \in S_n \Rightarrow \exists! M \in S_n^+ / A = M^2$ de plus M commute avec A et vérifie $rg(A) = rg(M) \Rightarrow M = A^{\frac{1}{2}}$).
3. $\mathbf{Tr}(AB) \leq \mathbf{Tr}(A)\mathbf{Tr}(B)$.
4. Les valeurs propres de la matrice AB sont positives.

Lemme 2.2.

1. Soient $A, B \in S_n^+$. Alors $A \bullet B \geq 0$.
De plus, on a l'équivalence suivante : $A \bullet B = 0 \Leftrightarrow AB = 0$.
2. Soient $A, B \in S_n^+$. Alors :

$$\lambda_{\min}(A)\lambda_{\max}(B) \leq \lambda_{\min}(A)\mathbf{Tr}(B) \leq A \bullet B \leq \lambda_{\max}(A)\mathbf{Tr}(B) \leq n\lambda_{\max}(A)\lambda_{\max}(B)$$

2.1.3.2 Cône des matrices symétriques positives

Définition 2.3.

Espace vectoriel des matrices symétriques :

$$S_n = \{A \in \mathbb{R}^{n \times n} / A = A^t\}$$

Cône des matrices symétriques positives de dimension n :

$$S_n^+ = \{A \in S_n / A \succeq 0\}$$

Ici \succeq ordre de Löwner, ou des formes quadratiques associées :

$$A \leq B \iff x^t Ax \leq x^t Bx, \forall x \in \mathbb{R}^n$$

Le cône S_n^+ est convexe et fermé et on a (λ_{\min} plus petite valeur propre) :

$$A \geq 0 \iff \lambda_{\min}(A) \geq 0$$

2.1.3.3 Cône de Lorentz

Définition 2.4.

Ensemble de \mathbb{R}^n défini par :

$$S_n = \{x \in \mathbb{R}^n, x_n \geq \sqrt{x_1^2 + \dots + x_{n-1}^2}\}$$

En dimension 3 : forme d'un "cornet de glace".

Si $n = 2$, on peut identifier S_2^+ au cône de Lorentz de dimension 3, utilisant les matrices de Pauli

$$A = x \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + y_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + y_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad x, y_1, y_2 \in \mathbb{R}$$

constatant que

$$\lambda_{\min}(A) \geq 0 \Leftrightarrow x \geq \sqrt{y_1^2 + y_2^2}.$$

Le cône des matrices SDP 2×2 , S_2^+ , coïncide avec le cône de Lorentz.

2.1.3.4 Cône de récession

Soit C un convexe non vide de \mathbb{R}^n et $a \in C$, on pose

$$C_\infty(a) = \{d \in \mathbb{R}^n / a + \lambda d \in C, \forall \lambda > 0\}$$

Alors $C_\infty(a)$ est un cône convexe non vide. En général, $C_\infty(a)$ dépend de a .

Définition 2.5. On appelle cône de récession (ou asymptote) de C l'ensemble

$$C_\infty = \bigcap_{a \in C} C_\infty(a)$$

Un élément $d \in C_\infty$ est appelée direction de récession.

2.1.3.5 Le cône polaire

Définition 2.6.

Le cône polaire S_n^* d'un cône S_n est l'ensemble :

$$S_n^* = \{y / \forall x \in S_n, x \bullet y \geq 0\}$$

Lemme 2.3.

$$(S_n^+(\mathbb{R}))^* = S_n^+(\mathbb{R}).$$

Ce lemme est équivalent au théorème de la trace de Fejer, que nous donnons en corollaire :

Corollaire :

$$A \in S_n^+(\mathbb{R}) \Leftrightarrow \forall B \in S_n^+(\mathbb{R}), A \bullet B \geq 0.$$

2.1.4 Formulation du problème

Les programmes semi définis généralisent la programmation linéaire au cône des matrices semi définies positives. Un programme semi défini (SDP) s'écrit :

$$(SDP) \begin{cases} \min C \cdot X = \mathbf{Tr}(C^t X) \\ A_i \cdot X = b_i, i = 1, \dots, m \\ X \in S_n^+(\mathbb{R}) \end{cases} \quad (2.1)$$

ou : $C, A_i \in M_n, b = (b_1, b_2, \dots, b_m)^t \in \mathbb{R}^m$.

La programmation semi-définie est un cas particulier de la programmation conique, et une généralité de la programmation linéaire, ici on manipule des matrices au lieu des vecteurs. Effectivement on peut ramener un problème de programmation linéaire en un problème de programmation semi-définie. Pour ce fait, soit le programme linéaire suivant :

$$(PL) \begin{cases} \min c^t x \\ Ax = b \\ x \geq 0, x \in \mathbb{R} \end{cases}$$

posons

$$C = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}, \quad X = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{pmatrix}$$

$$A_i = \begin{pmatrix} a_{i1} & 0 & \cdots & 0 \\ 0 & a_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{in} \end{pmatrix}, \quad b_i = b_i$$

on a donc :

$$\begin{cases} c^t x & = \langle C, X \rangle & = \mathbf{Tr}(C^t X) \\ Ax = b & \Leftrightarrow A_i \cdot X = b_i, & i = 1, \dots, m \\ x \geq 0 & \Leftrightarrow x_i \geq 0 & \forall i \Rightarrow X \succeq 0 \end{cases}$$

d'où :

$$(PL) \Leftrightarrow \begin{cases} \min c^t x \\ Ax = b \\ x \geq 0, x \in \mathbb{R} \end{cases} \Rightarrow \begin{cases} \min C \cdot X \\ A_i \cdot X = b_i, i = 1, \dots, m \\ X \succeq 0 \end{cases} \Leftrightarrow (SDP)$$

De plus l'ensemble des contraintes vérifiant $X \succeq 0$ est infini contrairement à la programmation linéaire ou il est fini ($x \succ 0$).

Exemple

La programmation linéaire est un cas particulier de la programmation semi définie :

$$(PL) \begin{cases} \min x_1 + 3x_2 \\ 2x_1 - x_2 = 1 \\ 5x_1 + 7x_2 = -6 \end{cases} \Leftrightarrow (SDP) \begin{cases} \min x_1 + 3x_2 \\ \begin{bmatrix} 2x_1 - x_2 - 1 & 0 \\ 0 & 5x_1 + 7x_2 + 6 \end{bmatrix} \succeq 0 \end{cases} = 0$$

2.2 Dualité en programmation semi-définie

La dualité en SDP est très similaire à la dualité classique en programmation linéaire à quelques différences près.

Soit le problème SDP linéaire primal sous la forme standard :

$$(SDP) \quad m_p = \min_x \left[\langle C, X \rangle, \langle A_i, X \rangle = b_i, i = 1, \dots, m, X \in S_n^+ \right]$$

Pour obtenir le problème dual de SDP, on considère la fonction Lagrangienne :

$$q(y) = \min_{x \in S_n^+} \left[\langle C, X \rangle + \sum_{i=1}^m (b_i - \langle A_i, X \rangle) y_i, y \in \mathbb{R}^m \right]$$

d'où

$$\max_{y \in \mathbb{R}^m} q(y) = \max_{y \in \mathbb{R}^m} \min_{x \in S_n^+} \left[\langle (C - \sum_{i=1}^m y_i A_i), X \rangle + \sum_{i=1}^m b_i y_i \right]$$

donc

$$\max_{y \in \mathbb{R}^m} q(y) = \begin{cases} \max_{y \in \mathbb{R}^m} \sum_{i=1}^m b_i y_i & \text{Si } C - \sum_{i=1}^m y_i A_i \in S_n \\ -\infty & \text{ailleurs} \end{cases}$$

D'où par convention le dual du problème SDP est un problème SDP défini par :

$$m_d = \begin{cases} \max b^t y \\ C - \sum_{i=1}^m y_i A_i \in S_n^+ \\ y \in \mathbb{R}^m \end{cases} \iff \begin{cases} \max b^t y \\ C - \sum_{i=1}^m y_i A_i = S \\ y \in \mathbb{R}^m, S \in S_n^+ \end{cases} \quad (DSDP)$$

Définition 2.7.

Une solution réalisable de DSDP est le couple $(y, S) \in \mathbb{R}^m \times S_n^+$ tel que :

$$C - \sum_{i=1}^m y_i A_i = S$$

De même le couple (y, S) est dit strictement réalisable pour DSDP si

$$C - \sum_{i=1}^m y_i A_i \in S_n^{++}$$

On note $F(\hat{F})$ l'ensemble des solutions réalisables pour DSDP (resp. l'ensemble des solutions strictement réalisables pour DSDP).

Définition 2.8.

La valeur optimale de DSDP est définie par :

$$d^* = m_d = \sup \left\{ b^t y : C - \sum_{i=1}^m y_i A_i \in S_n^+, y \in \mathbb{R}^m \right\}$$

avec

$$(y^*, S^*) \in F \text{ et } b^* y^* = d^*$$

et

$$S^* = C - \sum_{i=1}^m y_i A_i$$

Proposition 2.4. "Dualité faible"

Soit $X \in Y$ et $(y, S) \in F$ alors :

$$C \bullet X - b^t y = X \bullet S \geq 0$$

Définition 2.9.

La différence $C \bullet X - b^t y \geq 0$ est appelée écart de dualité.

Nous allons étudier quelles conditions donnent lieu à une dualité forte, c'est-à-dire à un écart de dualité de solutions optimales nul ($p^* - d^* = 0$).

Théorème 2.3. "Dualité fore"

Supposons qu'il existe $y \in \mathbb{R}^m$ tel que $\sum_{i=1}^m y_i A_i \succ 0$. Alors $p^* = d^*$.

Remarque 2.1.

1. $X \bullet S$ est appelé l'écart ou Saut de dualité.
2. Certains résultats de la programmation linéaire ne sont pas valables en SDP.

Exemple 2.1.

$E = S_3$ l'espace des matrice symétriques 3×3 , $m = 2$

$$m_p = \min_x [C \bullet X, X \in S_3^+, \langle A_i, X \rangle = b_i, i = 1, \dots, m] \quad (SDP)$$

ou :

$$C = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$X = \begin{pmatrix} x_{11} & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & x_{33} \end{pmatrix} \text{ et } b = (0, -1)^t$$

Ce problème s'écrit aussi sous la forme suivante :

$$m_p = \min_x \left[x_{12}, X = \begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1 + x_{12} \end{pmatrix} \geq 0 \right] \quad (SDP)$$

L'ensemble des solutions réalisables correspond à l'ensemble

$$Y = \{(x_{12}, x_{22}) \in \mathbb{R}^2 : x_{12} = 0, x_{22} \geq 0\}$$

❖ CHAPITRE 2. MÉTHODE PROJECTIVE DE POINT INTÉRIEUR POUR LA PROGRAMMATION SDP

En outre l'ensemble des solutions optimales est

$$X^* = \left\{ \left(\begin{array}{ccc} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1 + x_{12} \end{array} \right), x_{22} \geq 0 \right\}$$

Il s'ensuit que

$$m_p = 0$$

Écrivons le problème dual

$$m_d = \max_{y \in \mathbb{R}^2} \left[b^t y : C - \sum_{i=1}^2 y_i A_i \in S_3^+ \right] \quad (DSDP)$$

qui correspond à la forme

$$m_d = \max_{y \in \mathbb{R}^2} \left[-y : \left(\begin{array}{ccc} -y_1 & \frac{1-y_2}{2} & 0 \\ \frac{1-y_2}{2} & 0 & 0 \\ 0 & 0 & y_2 \end{array} \right) \in S_3^+ \right] \quad (DSDP)$$

l'ensemble des solutions réalisables correspond à l'ensemble

$$\{(y_1, y_2) \in \mathbb{R}^2 : y_1 \leq 0, y_2 = 1\}$$

En outre l'ensemble des solutions optimales est

$$Y^* = \{(y_1, 1), y_1 \leq 0\}$$

Il s'ensuit que

$$m_d = -1 \neq 0 = m_p$$

Les problèmes (SDP) et (DSDP) sont tous deux réalisables, X^* et Y^* sont tous deux non vides mais

$$-\infty < m_d < m_p < +\infty$$

Ainsi la réalisabilité des deux problèmes ne suffit pas pour avoir les mêmes valeurs optimales.

Pour conserver le résultat de la dualité forte en SDP, une condition plus forte est indispensable, il s'agit de la stricte réalisabilité comme le montre le théorème suivant :

Théorème 2.4.

1. Si le problème (SDP) primal est strictement réalisable, c-à-d :

$$\exists X \in S_n^{++}(\mathbb{R}) : \langle A_i, X \rangle = b_i, i = 1, \dots, m$$

alors

$$m_d = m_p$$

2. Si en outre m_p est fini, alors l'ensemble des solutions optimales du problème dual (DSDP) est compact non vide. Si le problème (DSDP) est strictement réalisable, c-à-d :

$$\exists y \in \mathbb{R}^m : C - \sum_{i=1}^m y_i A_i \in S_n^{++}$$

alors

$$m_d = m_p$$

3. Si en outre m_d est fini, alors l'ensemble des solutions optimales du problème primal (SDP) est compact non vide.

Définition 2.10.

Une matrice $X \in S_n$ est dite réalisable (resp. strictement réalisable) pour SDP si :

$$\langle A_i, X \rangle = b_i, i = 1, \dots, m \quad \text{et} \quad X \in S_n^+$$

$$\langle A_i, X \rangle = b_i, i = 1, \dots, m \quad \text{et} \quad X \in S_n^{++}$$

On note Y (resp \hat{Y}) l'ensemble des solutions réalisables (resp l'ensemble des solutions strictement réalisables) pour SDP :

Définition 2.11.

La valeur optimale (primal) de SDP est définie par :

$$p^* = \inf \left\{ \langle C, X \rangle \mid \langle A_i, X \rangle = b_i, i = 1, \dots, m, X \in S_n^+ \right\}$$

$$X^* \in Y \quad \text{et} \quad \langle C, X^* \rangle = p^*$$

2.3 Exemple de problème SDP

2.3.1 Problème de programmation non linéaire

Soit le programme non linéaire suivante :

$$\begin{cases} \min(x_1^3 + x_2) \\ x_1^3 x_2 \geq 1 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

Ce problème s'écrit sous forme d'un problème semi-défini (SDP) comme suit :

$$\min[\langle C, X \rangle : X \in K, \langle A_1, X \rangle = b_1]$$

$$\text{avec } C = I, X = \begin{pmatrix} x_1^3 & x_3 \\ x_3 & x_2 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, b_1 = 1$$

2.3.2 Problème de programmation min-max des valeurs propres

Ce problème a été étudié vers les années 1950 en Algèbre linéaire comme suit :

Chercher une valeur optimale

$$\lambda^* = \min_{y \in \mathbb{R}^n} \lambda_{\max}(C + A(y)) \quad (2.2)$$

Ou $C \in M_n$

$$\begin{aligned} A_{\text{opérateur}} : \mathbb{R}^n &\longrightarrow M_n \\ y &\longmapsto A(y) = \sum_{i=1}^n A_i y_i, \quad A_i \in M_n \end{aligned}$$

En fait, trouver la plus grande valeur propre d'une matrice symétrique A correspond au problème suivant :

$$\max_x [x^t A x : \|x\| = 1]$$

Ou $x \in \mathbb{R}^n - \{0\}$ est un vecteur propre de la matrice A .

Soit encore en considérant sur les matrices symétriques le produit scalaire :

$$\langle A, B \rangle = \text{Tr}(AB)$$

$$\max_x [\langle A, x x^t \rangle : \langle I, x x^t \rangle = 1] \quad (2.3)$$

Ou encore, X étant symétrique :

$$\max_X [\langle A, X \rangle : \langle I, X \rangle = 1, X \succeq 1 \text{ de range } 1] \quad (2.4)$$

Si on considère le problème semi-défini linéaire suivant :

$$\max_X [\langle A, X \rangle : \langle I, X \rangle = 1, X \succeq 1]$$

Ce problème admet des solutions optimales car l'ensemble des solutions réalisables est compact. La fonction objective étant linéaire et l'ensemble des solutions réalisables étant convexe, l'optimum est atteint au moins en un point extrémal, donc en un X matrice de rang 1. X est aussi solution optimale du problème (2.3). Bien sur, toute solution optimale de (2.3) est solution optimale de (2.4).

Le problème (2.2) s'écrit sous forme d'un problème SDP comme suit :

$$\lambda^* = \min_{y \in \mathbb{R}^n} \lambda_{\max}(C + A(y)) \Leftrightarrow \begin{cases} \min \lambda \\ \lambda I \geq C + A(y) \\ y \in \mathbb{R}^n \end{cases} \Leftrightarrow \begin{cases} \min \lambda \\ \lambda I - C - A(y) \geq 0 \\ y \in \mathbb{R}^n \end{cases}$$

Le dual de ce problème est un problème SDP formulé par :

$$\begin{cases} \max \langle C, X \rangle \\ A^t(X) = 0 \\ x \succeq 0 \\ \mathbf{Tr}(X) = 1 \end{cases}$$

2.4 Résolution de SDP

Nous avons déjà dit que l'ensemble S_n^+ est un cône non-polyédrique, et la notion de sommet n'est plus valable pour les problèmes SDP ce qui favorise d'avantage l'extension des méthodes de points intérieurs en programmation linéaire pour SDP.

Plusieurs travaux ont été réalisés depuis les années 90 et le grand nombre d'articles parus dans des revues internationales en témoigne, tout particulièrement les travaux de Shapiro, Fletcher-Craven (1996) qui s'intéressent aux conditions d'optimalité du problème SDP, les travaux de Ramana et Wolkowic (1997) qui ont étudié la dualité forte pour ces problèmes. Du point de vue algorithmique, on rencontre les méthodes de points intérieurs qui sont relativement nouvelles et qui s'apparentent à la méthode projective de Karmarkar pour la programmation linéaire. Ces dernières années, plusieurs chercheurs ont proposé des méthodes pour résoudre les problèmes SDP qui sont généralement des extensions des méthodes de points intérieurs pour la programmation linéaire. En effet, en 1994 Farid ALIZADEH est le 1^{er} chercheur qui a fait une étude profonde pour SDP et a proposé un algorithme primal-dual de points intérieurs du type projectif appelé "Algorithme de réduction du potentiel". Depuis, plusieurs algorithmes sont proposés dans la littérature, on cite par exemple :

1. F. Alizadeh (1995) propose une méthode projective primale-duale.
2. Vanderbeghe et Boyd (1996) ont proposé un algorithme primal-dual.
3. Monteiro (1997) propose une méthode de type trajectoire centrale.
4. Todd et All (1998) ont proposé des variantes newtoniennes pour résoudre le problème de complémentarité linéaire.
5. J. Ji et All (1999) ont étudié la convergence de la méthode de prédicteur-correcteur.
6. M. et autres (2002) ont étudié la convergence de la méthode de trajectoire centrale.
7. Dj. Benterki, J.P. Crouzeix et B. Merikhi (2004) ont proposé une étude théorique et numérique de certains algorithmes projectifs de Alizadeh.

Toutes ces méthodes de différents types à savoir les méthodes de trajectoire centrale réalisable et non-réalisable, primale, duale et primale-duale ont un problème majeur commun

celui de l'initialisation. En 2007 Dj. Benterki, J.P. Crouzeix et B. Merikhi ont proposé une méthode de point intérieur réalisable pour résoudre SDP. L'avantage principal de cette méthode est le calcul d'une solution réalisable initiale pour SDP.

2.4.1 Fonction barrières

Où $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et X est un ensemble fermé. L'ensemble des points admissibles est

$$\mathcal{F} = \{x \in \mathbb{R}^n \setminus x \in X, g(x) \leq 0\} \quad (2.5)$$

Dans ce cadre-ci, nous appellerons l'ensemble des points intérieurs l'ensemble S définie par :

$$S = \{x \in \mathbb{R}^n \setminus x \in X, g(x) < 0\} \quad (2.6)$$

Notre qu'il s'agit d'un abus de langage et que, si S est non vide, les points ne sont intérieurs que lorsque l'on se restreint à l'ensemble X . Nous supposerons que

- $S \neq \emptyset$.
- tout point admissible peut être approché arbitrairement par un point intérieur, c'est-à-dire que, pour tout $x \in \mathcal{F}$ et pour tout $\varepsilon > 0$, il existe $\tilde{x} \in S$ tel que

$$\|\tilde{x} - x\| \leq \varepsilon$$

si X est un ensemble convexe et g une fonction convexe, cette hypothèse est toujours vérifiée.

Lemme 2.4.

Soit $X \subset \mathbb{R}^n$ un ensemble convexe fermé .

Soit $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction convexe, soient \mathcal{F} défini par(2.5) et S défini par (2.6).

Pour tout $x \in \mathcal{F}$ et pour tout $\varepsilon > 0$ il existe $\tilde{x} \in S$ tel que :

$$\|\tilde{x} - x\| \leq \varepsilon$$

les méthodes de points intérieurs utilisant des fonctions dites fonctions barrières afin d'obliger les algorithmes à rester dans S .

Définition 2.12. " Fonction barrière "

Soit $X \subset \mathbb{R}^n$ un ensemble convexe fermé .

Soit $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction convexe, soit \mathcal{F} défini par(2.5), la fonction $B : S \rightarrow \mathbb{R}$ est une fonction barrière si elle est continue et si :

$$\lim_{X \in S, d(x) \rightarrow 0} B(x) = +\infty$$

Exemple 2.2. " Fonction barrière "

Les fonctions barrières les plus utilisées sont la fonction logarithmique.

$$B(x) = - \sum_{i=1}^m \ln(-g_i(x))$$

et la fonction inverse

$$B(x) = - \sum_{j=1}^m \frac{1}{g_j(x)}$$

Considérons l'ensemble des contraintes définies par $1 \leq x \leq 3$ sont définies par :
 $g : \mathbb{R} \rightarrow \mathbb{R}^2$, avec $g_1(x) = 1 - x$ et $g_2(x) = x - 3$, la fonction barrière logarithmique pour ces contraintes s'écrit :

$$- \ln(x - 1) - \ln(3 - x)$$

Définition 2.13. (Fonction barrière de type primal et de type dual pour PL)

Le programme linéaire (PL) sous forme standard suivant :

$$\left\{ \begin{array}{l} \min_x c^t x \\ Ax = b \\ x \geq 0 \end{array} \right. \quad (P)$$

et son dual

$$\left\{ \begin{array}{l} \max_y b^t y \\ A^t y \leq c \\ y \in \mathbb{R}^n \end{array} \right. \quad (D)$$

tels que $A \in \mathbb{R}^{m \times n}$ une matrice donnée et $b \in \mathbb{R}^m$ et $c \in \mathbb{R}^n$.

❖ CHAPITRE 2. MÉTHODE PROJECTIVE DE POINT INTÉRIEUR POUR LA PROGRAMMATION SDP

Une des idées derrière cette méthode est l'observation qu'une des difficultés de la programmation linéaire provient des inéquations $x \geq 0$. En effet les variables non astreintes, une fois dans la base, n'en sortent jamais et de plus dès qu'un coût réduit est non nul dans la fonction objectif, elle peut entrer dans la base. Pour cette raison on va convertir le PL en un problème avec seulement des équations, en utilisant une fonction barrière qui va empêcher une variable d'atteindre la frontière $x_i = 0$. On y parvient en ajoutant les termes $\ln x_i$ à la fonction objectif. Ces termes vont faire augmenter cette fonction vers $+\infty$ quand x_i va tendre vers 0. Nous introduisons la fonction barrière suivante :

$$B(x) = -\mu \sum_{i=1}^n \ln(-g_i(x))$$

est appelé fonction barrière logarithmique et μ désigne le paramètre de pénalité.

Donc le problème barrière ayant comme origine le primal est le suivant :

$$\left\{ \begin{array}{l} \min_x c^t x \\ Ax = b \\ x \geq 0 \end{array} \right\} \implies \left\{ \begin{array}{l} \min_x c^t x - \mu \sum_{i=1}^n \ln(-g_i(x)) \\ \mu > 0 \end{array} \right. \quad (\mathbf{P}_\mu)$$

Pour tout $\mu > 0$ le problème (\mathbf{P}_μ) admet une solution optimale $x(\mu)$.

Le problème barrière ayant comme origine le dual est le suivant :

$$\left\{ \begin{array}{l} \max_y b^t y \\ A^t y \leq c \\ y \in \mathbb{R}^n \end{array} \right\} \implies \left\{ \begin{array}{l} \max_y b^t y - \mu \sum_{i=1}^n \ln(-g_i(x)) \\ \mu > 0 \end{array} \right. \quad (\mathbf{D}_\mu)$$

Exemple 2.3.

Considérons le problème :

$$\left\{ \begin{array}{l} \min_x x \\ x \geq 0 \end{array} \right.$$

La fonction barrière est dans ce cas :

$$B(x, \mu) = x - \mu \ln x$$

Exemple 2.4.

Soit le problème suivante :

$$\left\{ \begin{array}{l} \min_x x_2 \\ x_1 + x_2 + x_3 = 1 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{array} \right.$$

La fonction barrière est dans ce cas :

$$B(x, \mu) = x_2 - \mu \ln x_1 - \mu \ln x_2 - \mu \ln x_3$$

CHAPITRE 3

RÉSOLUTION D'UN PROGRAMME SEMI-DÉFINI PAR UNE APPROCHE BARRIÈRE LOGARITHMIQUE

Introduction

On propose dans ce chapitre une approche barrière pour résoudre SDP. La méthode est basée sur une fonction de pénalisation logarithmique donnant lieu à un algorithme de type Newton dont l'efficacité est mise en valeur à travers des expérimentations numériques encourageantes. On s'intéresse dans ce chapitre au problème suivant :

$$m_d = \inf_y [b^t y : \sum_{i=1}^m y_i A_i - C \in K, y \in \mathbb{R}^m] \quad (D)$$

Où K désigne le cône des matrices $n \times n$ symétriques semi-définies positives et les matrices $C, A_i, i = 1, \dots, m$, sont des matrices symétriques données.

Le problème (D) est le dual du problème suivant :

$$m_p = \max_X [\langle C, X \rangle : X \in K, \langle A_i, X \rangle = b_i, \forall i = 1, \dots, m] \quad (P)$$

❖ CHAPITRE 3. RÉOLUTION D'UN PROGRAMME SEMI-DÉFINI PAR UNE APPROCHE BARRIÈRE LOGARITHMIQUE

où $\langle C, X \rangle = \text{Trace}(CX)$. On remarque que $\langle \cdot, \cdot \rangle$ définit un produit scalaire sur l'ensemble des matrices $n \times n$ symétriques et la norme associée

$$|||X||| = \langle X, X \rangle^{\frac{1}{2}} = \left(\sum_{i,j} X_{ij}^2 \right)^{\frac{1}{2}}$$

Un des avantages du problème (D) par rapport au problème (P) est que l'argument de la fonction que l'on minimise est un vecteur, alors qu'il est une matrice pour (P).

Dans toute la suite, on utilisera les notations suivantes :

$$Y = \left\{ y \in \mathbb{R}^m : \sum_{i=1}^m y_i A_i - C \in K \right\}, \quad \hat{Y} = \left\{ y \in \mathbb{R}^m : \sum_{i=1}^m y_i A_i - C \in \text{int}(K) \right\}$$

$$F = \{ X \in K : \langle A_i, X \rangle = b_i \forall i = 1, \dots, m \}, \quad \hat{F} = \{ X \in F : X \in \text{int}(K) \}$$

On rappelle que $\text{int}(K)$ est l'ensemble des matrices $n \times n$ symétriques définies positives.

Y et \hat{Y} sont les ensembles des solutions réalisables et strictement réalisables du problème (D) tandis que F et \hat{F} ceux du problème (P). Tous ces ensembles sont convexes.

La contrainte $\sum_{i=1}^m y_i A_i - C \in K$ sera traitée en considérant la famille de problèmes pénalisés :

$$m(r) = \inf [f_r(y) : y \in \mathbb{R}^m] \tag{D_r}$$

avec $r > 0$ paramètre de pénalisation et $f_r : \mathbb{R}^m \rightarrow]-\infty, +\infty]$ définie par :

$$f_r(y) = \begin{cases} b^t y + nr \ln r - r \ln [\det(\sum_{i=1}^m y_i A_i - C)] & \text{Si } y \in \hat{Y} \\ +\infty & \text{Sinon} \end{cases}$$

Pour étudier le problème D_r les hypothèses suivantes sont nécessaires :

- H1) Le système d'équations $\langle A_i, X \rangle = b_i; i = 1, \dots, m$ est de rang m .
- H2) les ensembles \hat{Y} et \hat{F} sont non vides.

On sait alors que :

- a) $-\infty < m_p = m_d < +\infty$.
- b) les ensembles de solutions optimales de (P) et (D) sont des convexes compacts non vides.

-c) si \bar{X} est solution optimale de (P) , alors \bar{y} est solution optimale de (D) si et seulement si

$$\bar{y} \in Y \text{ et } \left(\sum_{i=1}^m \bar{y}_i A_i - C \right) \bar{X} = 0$$

-d) si \bar{y} est solution optimale de (D) , alors \bar{X} est solution optimale de (P) si et seulement si

$$\bar{X} \in F \text{ et } \left(\sum_{i=1}^m \bar{y}_i A_i - C \right) \bar{X} = 0$$

Sous les hypothèses H1) et H2), la résolution du problème (D) permet d'obtenir la résolution du problème (P) .

3.1 Étude du problème perturbé (D_r) associé au problème semi-défini (D)

3.1.1 Calcul du gradient et du Hessien de la fonction barrière logarithmique f_r

La fonction f_r fait intervenir du déterminant. Pour calculer son gradient nous utiliserons le résultat suivant dont nous rappelons la démonstration.

Proposition 3.1.

Soit B une matrice symétrique définie positive et H une matrice symétrique alors

$$\det(B + H) = (1 + \langle B^{-1}, H \rangle) \det(B) + |||H||| \varepsilon(H)$$

où $\varepsilon(H) \rightarrow 0$ si $|||H||| \rightarrow 0$

Preuve :

Étant donné que B est symétrique définie positive, il existe une matrice triangulaire inférieure inversible L tel que $B = LL^t$, on a alors

$$\frac{\det(B + H) - \det(B)}{\det(B)} = \det(I + L^{-1}H(L^{-1})^t) \tag{3.1}$$

❖ CHAPITRE 3. RÉOLUTION D'UN PROGRAMME SEMI-DÉFINI PAR UNE APPROCHE BARRIÈRE LOGARITHMIQUE

On pose $K = L^{-1}H(L^{-1})^t$

$$\det[I + K] = \det \begin{bmatrix} 1 + k_{11} & k_{12} & k_{13} & \cdots & k_{1n} \\ k_{21} & 1 + k_{22} & k_{23} & \cdots & k_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ k_{n1} & k_{n2} & k_{n3} & \cdots & 1 + k_{nn} \end{bmatrix}$$

On a $\|K\| \rightarrow 0$ et $\|H\| \rightarrow 0$.

En utilisant l'expression du déterminant :

$$\det(B) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{i_1} \cdot a_{i_2} \cdot \dots \cdot a_{i_n}$$

où $\epsilon(\sigma)$ représente la signature de la permutation $\sigma = (i_1, i_2, \dots, i_n)$ et S_n l'ensemble des permutations de $(1, 2, \dots, n)$. En négligeant les termes de degré 2 en k_{ij} , on obtient

$$\det(I + K) \sim \prod_{i=1}^n (1 + k_{ii}) \sim 1 + \sum_{i=1}^n k_{ii} = 1 + \text{trace}(K)$$

Remplaçons dans (3.1) nous obtenons le résultat. □

Les notations suivantes sont utilisées dans l'expression du gradient et du Hessien : Pour $y \in \hat{Y}$, on introduit une matrice $B(y)$ symétrique définie positive de taille m , et une matrice $L(y)$ triangulaire inférieure telles que :

$$B(y) = \sum_{i=1}^m y_i A_i - C$$

cette matrice est symétrique définie positive pour tout $y \in \hat{Y}$.

$$B(y) = \sum_{i=1}^m y_i A_i - C = L(y)L^t(y)$$

Et pour $i, j = 1, 2, \dots, m$, on définit

$$\hat{A}_i(y) = [L(y)]^{-1} A_i [L^t]^{-1}$$

$$b_i(y) = \text{trace}(\hat{A}_i(y)) = \text{trace}(A_i B^{-1}(y))$$

$$\Delta_{ij}(y) = \text{trace}(B^{-1}(y) A_i B^{-1}(y) A_j) = \text{trace}(\hat{A}_i(y) \hat{A}_j(y))$$

Avec $b(y)$ est un vecteur de \mathbb{R}^m et $\Delta(y)$ est une matrice symétrique de taille m .

Théorème 3.1.

La fonction f_r est deux fois continument différentiable sur \hat{Y} .

Plus précisément, pour tout $y \in \hat{Y}$ on a :

- a) $\nabla f_r(y) = b - rb(y)$.
- b) $\nabla^2 f_r(y) = r\Delta(y)$.
- c) La matrice $\nabla^2 f_r(y)$ est définie positive.

Preuve :

- a) Notons par (e_1, e_2, \dots, e_m) la base canonique de \mathbb{R}^m , $i \in \{1, \dots, m\}$ et $z_i \in \mathbb{R}$, $z_i \neq 0$ alors

$$\begin{aligned} \frac{f_r(y + z_i e_i) - f_r(y)}{z_i} &= b_i - \frac{r}{z_i} [\ln \det(B(y + z_i e_i)) - \ln \det(B(y))] \\ &= b_i - \frac{r}{z_i} [\ln \det(L(y)[I + z_i \hat{A}_i]L^t(y)) - \ln \det(B(y))] \\ &= b_i - \frac{r}{z_i} \ln \det(I + z_i \hat{A}_i(y)) \\ &= b_i - \frac{r}{z_i} \ln[1 + z_i \text{trace}(\hat{A}_i(y)) + z_i \varepsilon(z_i)] \end{aligned}$$

où la fonction ε est telle que $\varepsilon(z) \rightarrow 0$ lorsque $z \rightarrow 0$.

Passant à la limite lorsque $z_i \rightarrow 0$, et en négligeant les termes de degré 2 en z , on en déduit pour tout $i = 1, \dots, m$.

$$[\nabla f_r(y)]_i = b_i - r \langle B^{-1}(y), A_i \rangle$$

- b) De même manière qu'en a), on considère pour tout $i, j \in \{1, \dots, m\}$

$$\frac{b_i(y + z_j e_j) - b_i(y)}{z_j} = \frac{-1}{z_j} [\text{trace}(A_i[B^{-1}(y + z_j e_j) - B^{-1}(y)])]$$

Mais

$$\begin{aligned} B^{-1}(y + z_j e_j) - B^{-1}(y) &= [B(y) + z_j A_j]^{-1} - B^{-1}(y) \\ &= [B(y)(I + z_j B^{-1}(y)A_j)]^{-1} - B^{-1}(y) \\ &= [(I + z_j B^{-1}(y)A_j)^{-1} - I]B^{-1}(y) \end{aligned}$$

En négligeant les termes du second ordre en z_j , nous obtenons

$$\frac{b_i(y + z_j e_j) - b_i(y)}{z_j} \sim \text{trace}(A_i B^{-1}(y) A_j B^{-1}(y))$$

Passant à la limite lorsque $z_j \rightarrow 0$. En outre l'égalité

$$\text{trace}(A_i B^{-1}(y) A_j B^{-1}(y)) = \text{trace}(\hat{A}_i(y) \hat{A}_j(y))$$

est immédiate.

- c) Soit $d \neq 0$. puis on pose $M = \sum_{i=1}^m d_i \hat{A}_i(y)$. En raison de l'hypothèse (H1), on a $M \neq 0$, D'autre part

$$\langle \nabla^2 f_r(y) d, d \rangle = r \text{trace} \left(\sum_{ij} d_i d_j \hat{A}_i(y) \hat{A}_j(y) \right) = r \text{trace}(M^2) > 0$$

Ainsi la matrice $\nabla^2 f_r(y)$ est définie positive. Il s'ensuit que la fonction f_r est strictement convexe sur \hat{Y} . □

Puisque f_r est strictement convexe (D_r) possède au moins une solution optimale.

3.2 Le problème D_r possède une solution unique

Puisque la fonction f_r prend la valeur $+\infty$ sur la frontière du domaine réalisable et différentiable à l'intérieur, alors elle est semi-continue inférieurement.

Pour prouver que D_r possède une solution unique, il suffit de prouver que le cône de récession est réduit à l'origine, avant ça on a le résultat suivant :

Proposition 3.2.

$$[b^t d \leq 0 \text{ et } \sum_{i=1}^m d_i A_i \in K] \Rightarrow d = 0$$

Preuve :

Admettons que $d \neq 0$, $b^t d \leq 0$ et $C = \sum_{i=1}^m y_i A_i \in K$. Alors d'une part (H1) implique que $C \neq 0$. D'autre part d'après (H2) il existe \hat{X} , telle que $\hat{X} \in \hat{F}$, et

$$0 < \langle C, \hat{X} \rangle = \sum_{i=1}^m d_i \langle A_i, \hat{X} \rangle = b^t d$$

D'où la proposition est prouvée. □

Théorème 3.2. $d = 0$ et $(f_r)_\infty \leq 0$.

Preuve :

D'après (H2), il existe $\bar{y} \in \hat{Y}$. La fonction de récession $(f_r)_\infty$ de $f - r$ est définie par

$$(f_r)_\infty(d) = \lim_{t \rightarrow +\infty} [\xi(t) = \frac{f_r(y + td) - f_r(y)}{t}]$$

Soit $B = B(y) = \sum_{i=1}^m y_i A_i - C$, puisque B est une matrice symétrique définie positive, alors il existe une matrice non singulière triangulaire inférieure telle que $B = LL^t$, pour $d \in \mathbb{R}^m$ on pose $H(d) = \sum_{i=1}^m d_i A_i$. Alors pour tout t telle que la matrice $B + tH(d)$ est définie positive on a :

$$\begin{aligned} \xi(t) &= b^t d - rt^{-1} [\ln \det(B + tH(d))] - [\ln \det(B)] \\ &= b^t d - rt^{-1} [\ln \det(I + tE(d))] \end{aligned}$$

Avec $E(d) = L^{-1}H(d)(L^{-1})^t$, on en déduit que

$$\xi(t) = \begin{cases} b^t d - rt^{-1} \ln \det(I + tE(d)) & \text{Si } I + tE(d) \in \widehat{K} \\ +\infty & \text{Sinon} \end{cases}$$

La condition $[f_r]_\infty \leq 0$ est aussi équivalente à dite que $H(d)$ est semi-défini positive et par suite $E(d)$ est aussi définie positive et

$$b^t d \leq r \lim_{t \rightarrow \infty} \frac{1}{t} \ln \det(I + tE(d)) = r \lim_{t \rightarrow \infty} \sum_{i=1}^n \frac{1}{t} \ln(1 + t\lambda_i(d)) = 0$$

où les $\lambda_i(d)$ désignant les valeurs propres de $E(d)$. On a

$$\det(I + tE(d)) = \prod_{i=1}^n (1 + t\lambda_i(d))$$

Passant à la limite lorsque t tend vers $+\infty$, on a

$$\lim_{t \rightarrow \infty} rt^{-1} \ln \det(I + tE(d)) = \lim_{t \rightarrow \infty} r \sum_{i=1}^n \frac{\ln(1 + t\lambda_i(d))}{t} = 0$$

on obtient par la suite que $[f_r]_\infty(d) \leq 0$ implique $\langle b, d \rangle \leq 0$ et $\sum_{i=1}^n d_i A_i$ semi définie positive.

La proposition(1.2) nous dit alors que $d = 0$. Ainsi f_r est inf-compacte.

Théorème 3.3.

Pour tout $r > 0$, (D_r) admet une solution optimale unique appartenant à \hat{Y} . On la notera $y_r = y(r)$.

3.3 Comportement de la solution optimale $y(r)$ du problème (D_r) lorsque r tend vers 0

Dans ce qui suit, on s'intéresse au comportement de la valeur optimal $m(r)$ et la solution optimale $y(r)$ du problème (D_r) , pour cela on considère la fonction

$h : \mathbb{R}^m \times \mathbb{R} \rightarrow]-\infty, +\infty[$ défini par :

$$h(y, t) = \begin{cases} b^t y - \ln \det[\sum_{i=1}^m y_i A_i - tC] & \text{Si } \sum_{i=1}^m y_i A_i - tC \in (K) \\ +\infty & \text{Sinon} \end{cases}$$

Il est facile de vérifier que la fonction h est semi-continue inférieurement sur $\mathbb{R}^m \times \mathbb{R}$. On définit ensuite la fonction

$$\phi(y, t, r) = \begin{cases} rh(r^{-1}y, r^{-1}t) & \text{Si } r > 0 \\ h_\infty(y, t) & \text{Si } r = 0 \\ +\infty & \text{Si } r < 0 \end{cases}$$

On sait, Rockafellar([7]), que ϕ est semi-continue inférieurement convexe sur $\mathbb{R}^m \times \mathbb{R} \times \mathbb{R}$.

On prend ensuite $f : \mathbb{R}^m \times \mathbb{R} \rightarrow]-\infty, +\infty[$ définie par :

$$f(y, r) = \phi(y, 1, r)$$

f est alors convexe et semi-continue inférieurement par construction

$$f(y, r) = \begin{cases} f_r(y) & \text{Si } r > 0 \\ b^t y & \text{Si } r = 0, y \in Y \\ +\infty & \text{ailleurs} \end{cases} \quad (3.2)$$

On définit la fonction $m : [0, \infty] \rightarrow]-\infty, +\infty[$ par $m(r) = \inf_y [f(y, r) : y \in \mathbb{R}^m]$. Cette fonction est convexe. En outre on a

$$m(r) = \begin{cases} \inf_y [f_r(y) : y \in \hat{Y}] & \text{Si } r > 0 \\ \inf_y [b^t y : y \in Y] & \text{Si } r = 0 \\ +\infty & \text{Si } r < 0 \end{cases}$$

Pour $r > 0$, on retrouve le problème (D_r) et pour $r = 0$, le problème (D) . Il est claire que pour $r > 0$ on a

$$m(r) = f_r(y(r)) = f(y(r), r)$$

et

$$0 = \nabla f_r(y(r)) = \nabla_y f(y(r), r) = b - rb(y_r)$$

On s'intéresse maintenant à la différentiabilité des fonctions $m(r)$ et $y(r)$ sur $]0, +\infty[$.

Théorème 3.4.

Les fonctions m et y sont continument différentiables sur $]0, +\infty[$. On a pour $r > 0$

$$r\Delta(y_r)y'(r) - b(y_r) = 0$$

$$m'(r) = n + n \ln(r) - \ln(\det(B(y_r)))$$

En outre

$$m_d = m(0) \leq b^t y(r) \leq m(0) + nr \quad (3.3)$$

Preuve :

Soit $\bar{r} > 0$, puisque $y(\bar{r}) \in \hat{Y}$ est solution optimal de $(D_{\bar{r}})$ alors

$$\nabla_y f(y(\bar{r}), \bar{r}) = 0$$

La fonction f est deux fois continument différentiable sur $\hat{Y} \times]0, +\infty[$, et la matrice $\nabla_{yy}^2 f(y(r), r)$ est définie positive. En appliquant le théorème des fonctions implicites à l'équation $0 = T(y, r) = \nabla_y f(y, r)$ au point $(y(\bar{r}), \bar{r})$, on déduit qu'au voisinage de \bar{r} la fonction y est continument différentiable et on a

$$\nabla_{yy}^2 f(y_r, r)y'(y) - b(y_r) = 0$$

Donc

$$y'(r) = -[\nabla_{yy}^2 f(y(r), r)]^{-1}(b - r\langle B^{-1}(y(r)), A \rangle)$$

On a $m(r) = f(y(r), r)$ et y est différentiable d'où

$$\begin{aligned} m'(r) &= f'_y(y(r), r)y'(r) + f'_r(y(r), r) \\ &= f'_r(y(r), r) \\ &= n(\ln(r) + 1) - \ln[\det(\sum_{i=1}^m y_i(r)A_i - C)] \end{aligned}$$

Puisque la fonction m est convexe

$$m(0) \geq m(r) + (0 - r)m'(r)$$

Pour laquelle, on obtient

$$+\infty > m_d = m(0) \geq b^t y(r) - nr > -\infty$$

D'autre part $y_r \in \hat{Y}$. Donc $b^t y(r) \geq m(0) = m_d$. Finalement on obtient

$$m_d \leq b^t y(r) \leq m_d + nr \quad \square$$

Désignons par S_D l'ensemble des solutions optimales de (D) , on sait que cet ensemble est un convexe compact non vide. La distance du point y à S_D est défini par :

$$d(y, S_D) = \inf_z [\|y - z\| : z \in S_D]$$

Le résultat suivant concerne le comportement de y_r et $m(r)$ lorsque r tend vers 0.

Théorème 3.5.

Lorsque $r \rightarrow 0$, $d(y, S_D) \rightarrow 0$ et $m(r)$ tend vers m_d .

Preuve :

Considérons la multiplication S défini sur \mathbb{R} par

$$S(r) = \{y \in Y : b^t y \leq m_d + nr\}$$

Son graphe est fermé, $S(r) = \emptyset$ si $r < 0$, $\emptyset \neq S(0) = S_D \subset S(r)$, $y_r \in S(r)$ et $S(r)$ est un ensemble convexe fermé. Le cône de récession de $S(r)$, $r > 0$, coïncide avec le cône de récession de $S(0)$, qui est un compact non vide et $S(r)$ est donc aussi un compact non vide. On en déduit que la multiplication S est semi-continue supérieurement(USC) sur $[0, +\infty]$, et puisque $y_r \in S(r)$, alors $d(y_r, S(0)) \rightarrow 0$ lorsque $r \rightarrow 0$.

Montrons que $m(r) \rightarrow m_d = m(0)$, puisque m est une fonction convexe, il suffit de prouver qu'elle est semi-continue inférieurement en 0. On procède par contradiction, alors il existe $\lambda < m_d$ et une suite $\{r_k\}$ de nombres positifs convergeant vers 0 telle que $m(r_k) > \lambda$.

Soit $y_{r_k} = y(r_k)$. Alors il existe $\bar{y} \in S_d$ est une sous suite $\{y_{r_{k_l}}\}$ convergeant vers \bar{y} . Puisque la fonction f est semi continue inférieurement sur $\mathbb{R}^m \times \mathbb{R}$ et $f(\bar{y}, 0) = m_d > \lambda$, on a pour l suffisamment grand

$$\lambda > m(r_{k_l}) = f(y_{k_l}, r_{k_l}) > \lambda$$

D'où la contradiction. □

3.4 Direction de Newton et recherche linéaire

La présence de la fonction barrière logarithmique dans la fonction objectif, ramène le problème (D) à un problème sans contrainte (D_r) , qu'on peut le résoudre par une méthode de descente classique de Newton. Comme f_r prend la valeur $+\infty$ sur la frontière alors les itérés y sont à l'intérieur du domaine réalisable \hat{Y} . Ainsi la méthode proposée est une méthode de points intérieurs.

Admettons que notre itéré $y \in \hat{Y}$. Alors pour direction de descente d , on choisit la direction de Newton qui est solution du système

$$[\nabla^2 f_r(y)]d = -\nabla f_r(y) \tag{3.4}$$

En utilisant le théorème (3.1), le système linéaire est équivalent au système

$$\Delta(y)d = b(y) - \frac{1}{r}b \tag{3.5}$$

$B(y)$, $b(y)$ et $\Delta(y)$ sont définies en (3.1.1)

La matrice $\Delta(y)$ étant symétrique, définie positive, la méthode de Choleskin est appropriée pour la résolution du système linéaire (3.5).

Pour la suite, on admet que $\nabla f(y) \neq 0$ (si non l'optimum est atteint). Il s'ensuit que $d \neq 0$. La direction de descente d ayant été calculé, notre prochaine étape consiste à choisir \bar{t} donnant une décroissance significative à f_r tout en conservant la définie positivité de la matrice $B(y^k + \bar{t}d)$.

Pour faire, on considère la fonction

$$\begin{aligned}\theta(t) &= \frac{1}{r}[f_r(y)(y + td) - f_r(y)], \quad y + td \in \hat{Y} \\ \theta(t) &= \frac{1}{r}tb^t d - \ln \det(B(y + td)) + \ln \det(B(y))\end{aligned}$$

Puisque $\nabla^2[f_r(y)]d = -\nabla f_r(y)$ on a

$$d^t \nabla^2 f_r(y) d = -d^t \nabla f_r(y) = d^t b(y) - r d^t b$$

Pour simplifier les notations, on prend

$$B = B(y) = \sum_{i=1}^m y_i A_i - C \text{ et } H = \sum_{i=1}^m d_i A_i$$

B étant symétrique définie positive, alors il existe une matrice triangulaire inférieure L (décomposition de Choleski) telle que $B = LL^t$.

Par suite, on pose

$$E = L^{-1}H(L^{-1})^t$$

Puisque $d \neq 0$, l'hypothèse (H1) implique que $H \neq 0$ et par suite $E \neq 0$. En prenant compte de cette notation, pour tout $t > 0$, telle que $I + tE$ est définie positive, on a

$$\theta(t) = t[\mathbf{Tr}(E) - \mathbf{Tr}(E^2)] - \ln \det(I + tE) \quad (3.6)$$

Désignons par λ_i les valeurs propres de la matrice symétrique E , que nous ne cherchons pas explicitement, nous obtenons

$$\theta(t) = \sum_{i=1}^m [t(\lambda_i - \lambda_i^2) - \ln(1 + t\lambda_i)], \quad t \in [0, \hat{t}]$$

avec

$$\hat{t} = \sup\{t : 1 + t\lambda_i > 0 \forall i\} = \sup\{t : y + td \in \hat{Y}\} \quad (3.7)$$

observons que $\hat{t} = +\infty$ si E est semi positive et $0 < \hat{t} < \infty$ sinon. Il est clair que θ est convexe sur $[0, \hat{t}]$, $\theta(t) = 0$ et

$$0 < \sum_{i=1}^m \lambda_i^2 = \theta''(0) = -\theta'(0)$$

en plus $\theta(t) \rightarrow +\infty$ lorsque $t \rightarrow \hat{t}$. Il s'ensuit, qu'il existe un point unique t_{opt} telle que $\theta'(t_{opt}) = 0$, θ atteint son minimum en ce point.

Malheureusement, d'une part, il n'existe pas de formules explicites pour le calcul de t_{opt} , d'autre part la résolution de l'équation $\theta'(t_{opt}) = 0$ par des méthodes itératives nécessite à chaque itération le calcul de θ et θ' qui est coteaux du point de vu calcul, du fait que l'expression de θ en (3.2) contient le déterminant qui n'est pas facile à calculer et (3.3) nécessite la connaissance de la valeur des λ_i de façon exacte. Ces difficultés nous a poussé a chercher d'autres alternatives. A partir de la donnée de la matrice E , il est facile d'obtenir

$$\mathbf{Tr}(E) = \sum_{i=1}^m e_{ii} = \sum_{i=1}^m \lambda_i \quad \text{et} \quad \mathbf{Tr}(E^2) = \sum_{i,j=1}^m e_{ij}^2 = \sum_{i=1}^m \lambda_i^2$$

Dans la section 6, on tire profit de ces données pour proposer des bornes inférieures pour \hat{t} et des majorantes pour la fonction θ . Avant ça, on introduit quelques inégalités utiles donnant information sur un ensemble de nombres réels positifs qu'on connait sa moyenne et son écart type.

3.5 Quelques inégalités utiles

On suppose connus la moyenne \bar{x} et l'écart type σ_x d'une série statistique x_1, x_2, \dots, x_n de n nombres réels et, on cherche certaines informations sur les x_i à partir de ces données. On rappelle que

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{et} \quad \sigma_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Le résultat suivant est du à H.Wolkowicz et G.P.H.Styan dans [5], voir aussi J.-P.Crouzeix et Alberto Seeger dans [6] pour des résultats complémentaires.

Proposition 3.3.

$$\begin{aligned} \bar{x} - \sigma_x \sqrt{n-1} &< \min_i x_i \leq \bar{x} - \frac{\sigma_x}{\sqrt{n-1}} \\ \bar{x} + \frac{\sigma_x}{\sqrt{n-1}} &\leq \max_i x_i \leq \bar{x} + \sigma_x \sqrt{n-1} \end{aligned}$$

En particulier, dans le cas où tout les x_i sont positifs en on réduit que

$$n \ln(\bar{x} - \sigma_x \sqrt{n-1}) \leq \sum_{i=1}^m \ln(x_i) \leq n \ln(\bar{x} + \sigma_x \sqrt{n-1})$$

où par convention, $\ln(t) = \infty$ si $t < 0$.

Théorème 3.6.

On admet que $x_i > 0$ pour tout $i = 1, \dots, n$. Alors

$$n \ln(\bar{x} - \sigma_x \sqrt{n-1}) \leq A \leq \sum_{i=1}^m \ln(x_i) \leq B \leq n \ln(\bar{x})$$

avec

$$A = (n-1) \ln\left(\bar{x} - \frac{\sigma_x}{\sqrt{n-1}}\right) + \ln(\bar{x} - \sigma_x \sqrt{n-1})$$

et

$$B = \ln(\bar{x} + \sigma_x \sqrt{n-1}) + (n-1) \ln\left(\bar{x} - \frac{\sigma_x}{\sqrt{n-1}}\right)$$

on posera

$$\bar{\lambda} = \frac{1}{n} \sum_{i=1}^n \lambda_i \quad \text{et} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n \lambda_i^2 - \bar{\lambda}^2$$

Nous obtenons

$$\theta(t) = t(n\bar{\lambda} - \|\lambda\|^2) - \sum_{i=1}^n \ln(1 + t\lambda_i)$$

Pour l'étude de la fonction θ on distingue deux cas :

Premier cas :

$$\lambda_i \geq 0.$$

pour tout i , θ est définie sur l'intervalle $[0, \infty[$ et admet un minimum unique sur $[0, \infty[$.

Deuxième cas :

$$\min \lambda_i < 0$$

θ est alors définie sur l'intervalle $[0, \hat{t}]$

avec

$$\hat{t} = -[\min \lambda_i]^{-1}$$

puisque $\theta(t)$ tend vers $+\infty$ lorsque t tend vers \hat{t} , θ admet un minimum unique sur $[0, \hat{t}]$.

Rappelons que nous ne voulons pas calculer les λ_i de façon explicite, ce qui fait que nous ne déterminons pas ce \hat{t} .

Nous allons proposer deux stratégies :

- La première stratégie consiste à chercher une minoration \tilde{t} de \hat{t} et faire ensuite une recherche linéaire de type Goldstein-Price sur l'intervalle $[0, \tilde{t}]$.

a) $\hat{t}_1 = -[\bar{\lambda} - (n-1)^{\frac{1}{2}}\sigma]^{-1}$. Cette valeur est obtenue à partir de la proposition (3.6) donnée en annexe.

b) Dans le cas $\lambda_i < 0$, on a pour tout i , $-\lambda_i \leq -\lambda_{\min} \leq \|\lambda\|$ alors $\|\lambda\|^{-1} \leq (-\lambda_{\min})^{-1}$. On prend donc comme deuxième minoration $\hat{t}_2 = \|\lambda\|^{-1}$.

Et comme on a $\|\lambda\|^{-1} = [(\sum_{i=1}^n \lambda_i^2)^{\frac{1}{2}}]^{-1} = [(\mathbf{Tr}(E^2))^{\frac{1}{2}}]^{-1}$, le calcul de \hat{t}_2 est aussi facilement obtenu à partir la matrice E .

c) On sait qu'une matrice symétrique dont les éléments diagonaux sont positifs et qui est à diagonale strictement dominante est définie positive.

Une première minoration est obtenue en prenant

$$\hat{t}_3 = \max[t : 1 + tE_{ii} > 0 \text{ et } |1 + tE_{ii}| > \sum_{i \neq j} |E_{ij}|] = \min_i [(s_i)^{-1} : s_i > 0]$$

avec

$$s_i = \sum_{i \neq j} |E_{ij}| - E_{ii}$$

Le calcul de \hat{t} est facilement calculé à partir de la matrice E .

En fait on pourra choisir

$$\tilde{t} = \max[\hat{t}_1, \hat{t}_2, \hat{t}_3]$$

- L'autre consiste à trouver une majorante de θ sur l'intervalle $[0, \hat{t}]$ que l'on sait minimiser et prendre la valeur t pour laquelle le minimum est facilement obtenu, et prendre comme pas de déplacement cette valeur.

Étude théorique de la fonction θ

Puisque la fonction

$$\theta(t) = \frac{1}{r} [f_r(y + td) - f_r(y)]$$

et que la fonction f_r est inf-compacte et strictement convexe, θ est aussi infcompacte est strictement convexe sur son domaine de définition, θ atteint son minimum en un point unique, en outre on a

$$\theta'(t) = n\bar{\lambda} - \|\lambda\|^2 - \sum_{i=1}^n \frac{\lambda_i}{1 + t\lambda_i} \text{ alors } \theta'(0) = -\|\lambda\|^2$$

$$\theta''(t) = \sum_{i=1}^n \frac{\lambda_i^2}{(1 + t\lambda_i)^2} \text{ d'où } \theta''(0) = \|\lambda\|^2$$

Posons $\gamma = n\bar{\lambda} - \|\lambda\|^2$, alors

$$\theta = \gamma t - \sum_{i=1}^n \ln(1 + t\lambda_i)$$

3.5.1 Première majorante θ_1

Pour $x_i = 1 + t\lambda_i$, on a alors

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 1 + t\bar{\lambda} \quad \text{et} \quad \sigma_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = t^2 \sigma^2$$

Supposons que les x_i sont strictement positif, en appliquant le théorème (3.8) en annexe on obtient

$$-\sum_{i=1}^n \ln(1 + t\lambda_i) \leq -(n-1) \ln(1 + \alpha_1 t) - \ln(1 + \beta_1 t)$$

avec

$$\alpha_1 = \bar{\lambda} + \frac{\sigma}{\sqrt{n-1}} \quad \text{et} \quad \beta_1 = \bar{\lambda} - \sigma\sqrt{n-1}$$

Les logarithmes sont bien définis dès lorsque $t \leq \hat{t}_1$ avec

$$\hat{t}_1 = \begin{cases} \frac{-1}{\beta_1} & \text{si } \beta_1 < 0 \\ +\infty & \text{sinon} \end{cases}$$

On en déduit la majoration suivante pour tout $t \in [0, \hat{t}_1[$

$$\theta \leq \theta_1(t) = \gamma t - (n-1) \ln(1 + \alpha_1 t) - \ln(1 + \beta_1 t)$$

Rappelons que l'on a

$$\gamma = n\bar{\lambda} - \|\lambda\|^2$$

On a alors les relations suivantes

$$-\gamma + (n-1)\alpha_1 + \beta_1 = (n-1)\alpha_1^2 + \beta_1^2 = \sum_{i=1}^n \lambda_i^2$$

D'où

$$\gamma = (n-1)(\alpha_1 - \alpha_1^2) + (\beta_1 - \beta_1^2)$$

D'autre part pour tout $t \in [0, \hat{t}_1[$

$$\theta_1'(t) = \gamma - \frac{(n-1)\alpha_1}{1 + \alpha_1 t} - \frac{\beta_1}{1 + \beta_1 t}$$

et

$$\theta_1''(t) = \frac{(n-1)\alpha_1^2}{(1+\alpha_1 t)^2} + \frac{\beta_1^2}{(1+\beta_1 t)^2}$$

On sait alors que

$$\theta'(0) = \theta_1'(0) = -n(\bar{\lambda}^2 + \sigma^2) = -\sum_{i=1}^n \lambda_i^2$$

$$\theta''(0) = \theta_1''(0) = n(\bar{\lambda}^2 + \sigma^2) = \sum_{i=1}^n \lambda_i^2 = \mathbf{Tr}(E^2)$$

θ_1 est strictement convexe sur $]0, \infty[$ et $\theta_1'(0) < 0$. Si t tend vers ∞ , alors puisque θ_1 majore θ qui est inf-compacte, θ_1 admet un minimum sur $[0, \hat{t}_1[$. Si $\hat{t}_1 < \infty$, alors $\theta_1(t)$ tend vers ∞ si t tend vers \hat{t}_1 . Donc θ_1 admet un minimum unique sur $[0, \hat{t}_1[$. Ce minimum est obtenu en t_{opt} telle que $\theta_1'(t_{opt}) = 0$.

On est donc ramené à résoudre l'équation du second degré

$$\alpha_1 \beta_1 \gamma t^2 + ((\alpha_1 + \beta_1)\gamma - n\alpha_1 \beta_1)t + (\gamma - (n-1)\alpha_1 - \beta_1) = 0$$

les racines de cette équation sont du type

$$t = \frac{-1}{2\alpha_1 \beta_1 \gamma} \left[\alpha_1 \gamma + \beta_1 \gamma - n\alpha_1 \beta_1 \pm \sqrt{[(\gamma(\alpha_1 - \beta_1) - n\alpha_1 \beta_1)^2 + 4\alpha_1 \beta_1 \gamma(\alpha_1 - \beta_1)]} \right]$$

On prend la seule des deux racines qui appartient à $[0, \hat{t}_1[$.

3.5.2 Une deuxième majorante θ_2

La fonction θ_2 est donnée par

$$\theta_2(t) = -\|\lambda\|(1 + \|\lambda\|)t - \ln(1 + t\|\lambda\|)$$

définie sur $[0, \hat{t}_2[$ avec $\hat{t}_2 = \|\lambda\|^{-1}$.

Proposition 3.4.

Pour tout $t \in [0, \hat{t}_2[$ on a

a) $\theta_2(0) = \theta(0) = 0$ et $\theta_2'(0) = \theta'(0) = -\|\lambda^2\| < 0$.

b) $\theta_2''(0) = \theta''(0) = \|\lambda^2\| > 0$.

c) $\theta(t) \leq \theta_2(t)$.

Preuve :

a) A partir de la définition on a $\theta_2(0) = \theta(0) = 0$, et

$$\theta_2'(t) = \frac{\|\lambda\|}{1 - t\|\lambda\|} - \|\lambda\|(1 + \|\lambda\|)$$

ce qui donne $\theta_2'(0) = -\|\lambda^2\| < 0$, d'où $\theta'(0) = \theta_2'(0) < 0$.

D'autre part

$$\theta_2''(t) = \frac{\|\lambda\|^2}{(1 - t\|\lambda\|)^2} > 0$$

d'où $\theta_2''(0) = \theta''(0) = \|\lambda^2\| > 0$.

b) On considère la fonction :

$$h(t) = \theta_2(t) - \theta(t) = (-\|\lambda\| - \sum_{i=1}^n \lambda_i)t - \ln(1 - t\|\lambda\|) + \sum_{i=1}^n \ln(1 + t\lambda_i)$$

On a par définition $h(0) = 0$ et pour étudier le signe de la fonction h on distingue deux cas :

1) S'il existe i tel que $\|\lambda\| = -\lambda_i$, alors $h(t) = 0$ pour tout $t \in [0, \hat{t}_2[$.

2) Dans le cas contraire, on sait que $-\|\lambda\| < \lambda_i \leq \|\lambda\|$. En outre

$$h'(t) = t \sum_{i=1}^n \lambda_i^2 [(1 - t\|\lambda\|)^{-1} - (1 + t\lambda_i)^{-1}]$$

Puisque $1 - t\|\lambda\| < 1 + t\lambda_i$, pour tout i , alors $h'(t) > 0$ pour tout $t \in [0, \hat{t}_2[$, d'où la fonction $h(t)$ est strictement croissante et comme $h(0) = 0$, alors $h(t) \geq 0$ pour tout $t \in [0, \hat{t}_2[$, ce qui donne

$$\theta \leq \theta_2(t) \quad \text{pour tout } t \in [0, \hat{t}_2[\quad \square$$

Théorème 3.7.

La fonction θ_2 est strictement convexe. Elle atteint son minimum sur $[0, \hat{t}_2[$ au point $\hat{t} = (1 + \|\lambda\|)^{-1}$ et on a

$$\theta_2(\hat{t}) = \ln(1 + \|\lambda\|) - \|\lambda\|$$

Preuve :

On note que l'on a

$$\theta_2'(t) = \frac{\|\lambda\|}{1 - t\|\lambda\|} - \|\lambda\|(1 + \|\lambda\|)$$

$$\theta_2''(t) = \frac{\|\lambda\|^2}{(1 - t\|\lambda\|)^2} > 0$$

D'où θ_2 est strictement convexe et alors, θ_2 atteint son minimum sur $[0, \hat{t}_2[$ au point \hat{t} si et seulement si $\theta_2'(\hat{t}) = 0$, ce qui donne $\hat{t} = (1 + \|\lambda\|)^{-1}$ et

$$\theta_2(\hat{t}) = \ln(1 + \|\lambda\|) - \|\lambda\| \quad \square$$

3.5.3 La troisième fonction majorante θ_3

En outre, L'introduction Les fonctions de type

$$\theta_3(t) = \gamma_3 t - \delta_3 \ln(1 + \beta_3 t), t \in [0, \hat{t}_3[$$

permettent d'avoir une comparaison entre $\theta_1, \theta_2, \theta_3$, on admet que :

$$\beta_3 = \beta_1 = \bar{\lambda} - \sigma_\lambda \sqrt{n-1}$$

et

$$\|\lambda\|^2 = \delta_3 \beta_3^2 = \delta_3 \beta_3 - \gamma_3, \quad \hat{t}_3 = \sup[t : 1 + t\beta_3 > 0]$$

Les fonctions θ_3 ainsi définies sont convexes.

Proposition 3.5.

Les fonctions $\theta_i, i = 1, 2, 3$ sont strictement convexes sur $[0, \hat{t}_i[$, $\theta_i(t) \rightarrow +\infty$ lorsque $t \rightarrow \hat{t}_i$, et on a

$$\theta(t) \leq \theta_1(t) \leq \theta_3(t) \leq \theta_2(t) \leq +\infty \quad \text{pour tout } t > 0$$

Preuve :

L'inégalité $\theta(t) \leq \theta_1(t)$ est une conséquence directe du théorème (3.8).

Soit $\nu(t) = \theta_3 - \theta_1$, comme $\beta_1 = \beta_3$ et $\alpha_1 \geq \beta_1$ alors on a pour $t > 0$

$$\nu''(t) = \frac{\delta_3 \beta_3^2 - \beta_1^2}{(1 + \beta_1 t)^2} - \frac{(n-1)\alpha_1^2}{(1 + \alpha_1 t)^2} = \frac{(n-1)\alpha_1^2}{(1 + \beta_1 t)^2} - \frac{(n-1)\alpha_1^2}{(1 + \alpha_1 t)^2} \geq 0$$

Puisque $\nu(0) = \nu'(0) = 0$, on déduit que $\nu(t) \geq 0$ pour $t > 0$.

Puis, on pose $\mu(t) = \theta_2 - \theta_3$. Alors, $\mu(0) = \mu'(0) = 0$ et

$$\mu''(t) = \|\lambda\|^2 \left[\frac{1}{(1 + \beta_2 t)^2} - \frac{1}{(1 + \beta_3 t)^2} \right] \geq 0$$

avec $\beta_2 = -\|\lambda\|$.

D'où $\mu(t) \geq 0$ pour tout $t > 0$.

On déduit que $\theta_i, i = 1, 2, 3$ atteint son minimum en un point unique \bar{t}_i , qui est la racine de

$$\theta'_i(\bar{t}_i) = 0$$

On a

$$\bar{t}_i = \frac{\delta_i}{\gamma_i} - \frac{1}{\beta_i} \quad \text{et} \quad \theta_i(\bar{t}_i) = \frac{\|\lambda\|^2}{\beta_i} + \frac{\|\lambda\|^2}{\beta_i^2} \ln(1 - \beta_i)$$

En particulier

$$\bar{t} = \frac{1}{1 + \|\lambda\|} \quad \text{et} \quad \theta_2(\bar{t}) = -\|\lambda\| + \ln(1 + \|\lambda\|)$$

Les trois racines \bar{t}_1, \bar{t}_2 et \bar{t}_3 sont calculés explicitement.

Il est clair que

$$\theta(\bar{t}_2) \leq \theta_2(\bar{t}_2), \quad \theta(\bar{t}_3) \leq \theta_3(\bar{t}_3) \leq \theta_3(\bar{t}_2) \leq \theta_2(\bar{t}_2)$$

et

$$\theta(\bar{t}_1) \leq \theta_1(\bar{t}_1) \leq \theta_1(\bar{t}_3) \leq \theta_3(\bar{t}_3) \leq \theta_2(\bar{t}_2)$$

□

3.5.4 Algorithme :

- Initialisation $k = 0$, ϵ : précision donnée, y_0 vecteur donné, $r = 0,001$.
- Début itération
 - Tant que $\|\nabla f_r(y^k)\| > \epsilon$ faire
 - $d^k = -(\nabla^2 f_r(y^k))^{-1} \nabla f_r(y^k)$
 - $H_k = \sum_{i=1}^m d_i^k A_i$
 - $E^k = L_k^{-1} H_k (L_k^{-1})^t$
 - calculer la trace de E^k et $(E^k)^2$
 - Déterminer le pas de déplacement t^k dans la direction de d^k
 - $y^{k+1} = y^k + t^k d^k$
 - Fin tant que.
- Fin itération.

La détermination du pas de déplacement t^k se fait en effectuant :

- **Stratégie Ls** : une recherche linéaire de type Goldstein-Price sur $[0, \tilde{t}]$.
- **Stratégie S_i** , $i = 1, 2, 3$: prendre le \hat{t}_i qui minimise la majorante θ_i , $i = 1, 2, 3$.

3.6 Tests Numériques

Exemple 3.1.

$$m_d = \min_y [b^t y : \sum_{i=1}^m y_i A_i - C \in K, y \in \mathbb{R}^m] \quad (SDP)$$

avec $m = 2, n = 3$ et

$$C = \begin{pmatrix} -1 & 1 & -1 \\ 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix}, A_1 = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Exemple 3.2.

$$m_d = \min_y [b^t y : \sum_{i=1}^m y_i A_i - C \in K, y \in \mathbb{R}^m] \quad (SDP)$$

avec $m = 3, n = 4$ et

$$C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 3 \end{pmatrix}.$$

Exemple cube

$$m_d = \min_y [b^t y : \sum_{i=1}^m y_i A_i - C \in K, y \in \mathbb{R}^m] \quad (SDP)$$

Dans cet exemple on prend $n = 2m, C = I, b_k = 2, k = 1, 2, \dots, m$, les matrices des contraintes sont données par :

$$A_k[i, j] = \begin{cases} 1 & \text{si } i = j = k \quad \text{ou } i = j = k + m \\ \lambda^2 & \text{si } i = j = k + 1 \quad \text{ou } i = j = k + m + 1 \\ -\lambda & \text{si } i = k, j = k + 1 \quad \text{ou } i = k + m, j = k + m + 1 \\ -\lambda & \text{si } i = k + 1, j = k \quad \text{ou } i = k + m + 1, j = k + m \\ 0 & \text{ailleurs} \end{cases}$$

Avec $k + m + 1 \leq n$.

Pour $\lambda = 0$, la solution optimale est : $y_i = 1$, pour tout $i = 1, \dots, m$.

La programmation de cet algorithme a été effectuée sur une station Pentium III au LIMOS (Laboratoire d'informatique, de modélisation et d'optimisation des systèmes), Université de Blaise Pascal, Clermont Ferrand.

❖ CHAPITRE 3. RÉOLUTION D'UN PROGRAMME SEMI-DÉFINI PAR UNE APPROCHE BARRIÈRE LOGARITHMIQUE

Les tests numériques sont résumés dans les tableaux suivants où on donne la valeur de $b^t y(r)$, le nombre des itérations, la valeur optimale et le temps d'exécution pour $r = 0.3, r = 0.01$, en appliquant les méthodes

R.L : Recherche linéaire

M1 : désigne la méthode utilisant la première majorante

M2 : désigne la méthode utilisant la deuxième majorante.

Il est rappelé que l'on a

$$m_d \leq b^t y(r) \leq m_d + nr$$

1) $r = 0.3$

Exemples	$b^t y(r)$			m_d	$m_d + nr$	Nb Itér			Temps		
	RL	M1	M2			RL	M1	M2	RL	M1	M2
Ex 1	0.4	0.4	0.4	0	0.9	55	6	6	6^{ct}	0^{ct}	0^{ct}
Ex 2	div	-8.38	-8.38	-9	-7.8	div	20	17	div	6^{ct}	5^{ct}

2) $r = 0.01$

Exemples	$b^t y(r)$			m_d	$m_d + nr$	Nb Itér			Temps		
	RL	M1	M2			RL	M1	M2	RL	M1	M2
Ex 1	div	0.01	0.01	0	0.03	div	8	7	div	0^{ct}	0^{ct}
Ex 2	div	-8.99	-8.99	-9	-8.96	div	79	81	div	22^{ct}	17^{ct}

Example cube

1) $r = 0.3$

(m,n)	$b^t y(r)$			m_d	$m_d + nr$	Nb Itér			Temps		
	RL	M1	M2			RL	M1	M2	RL	M1	M2
(9,18)	23.4	23.4	div	18	23.4	3	8	div	2^s	2^s	div
(50,100)	130	130	div	100	130	3	12	div	7^{mn}	27^{mn}	div
(60,120)	156	156	156	120	156	3	12	3	18^{mn}	$1^h 25^{mn}$	9^{mn}

❖ CHAPITRE 3. RÉOLUTION D'UN PROGRAMME SEMI-DÉFINI PAR UNE APPROCHE BARRIÈRE LOGARITHMIQUE

2) $r = 0.01$

(m,n)	$b^t y(r)$			m_d	$m_d + nr$	Nb Itér			Temps		
	RL	M1	M2			RL	M1	M2	RL	M1	M2
(9,18)	div	18.18	18.18	18	18.18	div	21	3	div	7^s	2^s
(50,100)	div	101	101	100	101	div	45	3	div	$1^h 51^{mn}$	3^{mn}
(60,120)	div	121.2	121.2	120	121.2	div	49	3	div	$5^h 18^{mn}$	9^{mn}

Dans les tableaux le signe div indique que la méthode n'a pas permis d'obtenir la solution.

CONCLUSION GÉNÉRALE

Dans notre étude, nous avons abordé des questions ouvertes concernant la programmation semi-définie (SDP) linéaire.

Les réponses que nous avons apporté sont d'une valeur algorithmique et numérique très intéressante et ouvrent plusieurs perspectives.

L'approche barrière logarithmique pour (SDP), est un véritable exploit, sur l'aspect théorique et algorithmique. L'aspect numérique pourra être poussé à un niveau de performance très appréciable pour la pratique.

La technique de fonctions majorantes pour déterminer le pas du déplacement suivant la direction de descente, est une alternative très fiable qui devra se confirmer comme la technique de choix aussi bien pour (SDP) que pour d'autres classes de problèmes d'optimisation.

ANNEXE

On suppose connus la moyenne \bar{x} et l'écart type σ d'une série statistique $\{x_1, x_2, \dots, x_n\}$ de n nombres réels et on cherche certaines informations sur les x_i à partir de ces données.

On rappelle que

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{et} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

Le premier résultat concerne le maximum et le minimum des x_i . Il est donné par H.Wolkowicz et G.P.H.Styan dans [5], voir aussi J.-P.Crouzeix et Alberto Seeger dans [6] pour des résultats complémentaires.

Proposition 3.6.

$$\begin{aligned} \bar{x} - (n-1)^{\frac{1}{2}}\sigma &\leq \min_i x_i \leq \bar{x} - (n-1)^{-\frac{1}{2}}\sigma \\ \bar{x} - (n-1)^{-\frac{1}{2}}\sigma &\leq \max_i x_i \leq \bar{x} - (n-1)^{\frac{1}{2}}\sigma \end{aligned}$$

Nous allons montrer un second résultat, dans le cas où les x_i sont strictements positifs. Il concerne leur produit.

Théorème 3.8.

Supposons que les x_i sont strictements positifs, alors

$$A \leq \sum_{i=1}^n \ln(x_i) \leq B$$

avec

$$A = (n-1) \ln\left(\bar{x} + \frac{\sigma}{\sqrt{n-1}}\right) + \ln(\bar{x} - \sigma\sqrt{n-1})$$

et

$$B = \ln(\bar{x} + \sigma\sqrt{n-1}) + (n-1)\ln\left(\bar{x} - \frac{\sigma}{\sqrt{n-1}}\right)$$

Preuve :

- a) Posons $s_1 = \sum_{i=1}^n x_i = n\bar{x}$ et $s_2 = \sum_{i=1}^n x_i^2 = n(\sigma^2 + \bar{x}^2)$. Si tous les x_i sont égaux alors $\sigma = 0$ et $x_i = \bar{x}$ pour tout i : le résultat est trivialement vrai. On supposera donc que les x_i ne sont pas tous égaux entre eux et donc $\sigma > 0$.

En fait on a nécessairement $s_2^2 < s_1^2$ car sinon tous les x_i seraient nuls sauf un, or on a supposé que les x_i étaient strictement positifs.

- b) Puisque on s'intéresse au problème de minimiser un produit de termes positifs x_i dont on connaît la somme s_1 et la somme des carrés s_2 , on est donc ramené à considérer les deux problèmes suivants

$$A = \min \left[f(x) = \sum_{i=1}^n \ln(x_i) : \sum_{i=1}^n x_i = s_1 \quad \text{et} \quad \sum_{i=1}^n x_i^2 = s_2^2 \right]$$

$$B = \max \left[f(x) = \sum_{i=1}^n \ln(x_i) : \sum_{i=1}^n x_i = s_1 \quad \text{et} \quad \sum_{i=1}^n x_i^2 = s_2^2 \right]$$

En appliquant les conditions d'optimalité du premier ordre, on voit que si x est solution d'un de ces problèmes, il existe α et β tels que pour tout i on a

$$x_i^2 - \alpha x_i + \beta = 0$$

L'équation du second degré

$$\omega^2 - \alpha\omega + \beta = 0$$

admet deux racines ω et ω' . Les x_i se partagent en p pour lesquels $x_i = \omega$ et $(n-p)$ pour lesquels $x_i = \omega'$.

Puisque les x_i ne sont pas tous égaux, on a nécessairement $1 \leq p \leq n-1$.

On a donc $s_1 = p\omega + (n-p)\omega'$ et $s_2^2 = p\omega^2 + (n-p)(\omega')^2$. On en déduit

$$s_2^2 = p\omega^2 + \frac{(s_1 - p\omega)^2}{n-p}$$

On obtient alors l'équation

$$\omega^2 - 2\omega\bar{x} + \bar{x}^2 - \frac{n-p}{p}\sigma^2 = 0$$

dont les racines sont

$$\omega_1 = \bar{x} + \sigma \sqrt{\frac{n-p}{p}} \quad \text{et} \quad \omega_2 = \bar{x} - \sigma \sqrt{\frac{n-p}{p}}$$

Les ω' correspondants sont

$$\omega'_1 = \bar{x} - \sigma \sqrt{\frac{p}{n-p}} \quad \text{et} \quad \omega'_2 = \bar{x} + \sigma \sqrt{\frac{p}{n-p}}$$

On a alors,

$$A = \min_p [\min[h(p), k(p)]] \quad \text{et} \quad B = \max_p [\max[h(p), k(p)]]$$

avec

$$h(p) = p \ln \left[\bar{x} + \sigma \sqrt{\frac{n-p}{p}} \right] + (n-p) \ln \left[\bar{x} - \sigma \sqrt{\frac{p}{n-p}} \right]$$

$$k(p) = p \ln \left[\bar{x} - \sigma \sqrt{\frac{n-p}{p}} \right] + (n-p) \ln \left[\bar{x} + \sigma \sqrt{\frac{p}{n-p}} \right]$$

On remarque que l'on a pour $p = 1, \dots, n-1$

$$k(p) = k(n-p)$$

d'où par symétrie

$$A = \min_p [h(p)] \quad \text{et} \quad B = \max_p [h(p)]$$

Dans cette partie de la preuve, nous allons montrer que la fonction h est décroissante sur l'intervalle $[1, n-1]$. On aura donc

$$A = h(n-1) \quad \text{et} \quad B = h(1)$$

Pour simplifier les calculs, on fait le changement de variable $z = \sqrt{\frac{p}{n-p}}$, et on considère la fonction $\tilde{h}(z) = h(p)$, on a

$$\tilde{h}(z) = \frac{nz^2}{z^2+1} \ln[\bar{x} + \sigma z^{-1}] + \frac{n}{(z^2+1)} \ln[\bar{x} - \sigma z]$$

et

$$\tilde{h}'(z) = \frac{2nz}{(z^2+1)^2} \left[\ln(\bar{x} + \sigma z^{-1}) - \ln(\bar{x} - \sigma z) \right] - \frac{n\sigma}{(z^2+1)^2} \left[\frac{1}{\bar{x} + \sigma z^{-1}} + \frac{1}{\bar{x} - \sigma z} \right]$$

Puisque la fonction $x \rightarrow x^{-1}$ est convexe sur $]0, +\infty[$, on a l'inégalité

$$\int_{\bar{x}-\sigma z}^{\bar{x}+\frac{\sigma}{z}} x^{-1} dx \leq \frac{1}{2} \left[\left(\bar{x} + \frac{\sigma}{z} \right) - \left(\bar{x} - \sigma z \right) \right] \left[\frac{1}{\bar{x} + \sigma z^{-1}} + \frac{1}{\bar{x} - \sigma z} \right]$$

Reportons dans l'expression de $\tilde{h}'(z)$, on voit que $\tilde{h}'(z) \leq 0$. Par conséquent \tilde{h} est décroissante sur l'intervalle $[(n-1)^{\frac{-1}{2}}, \sqrt{n-1}]$, puis h est décroissante sur $[1, n-1]$.

D'où

$$A \leq \sum_{i=1}^n \ln(x_i) \leq B \quad \square$$

BIBLIOGRAPHIE

- [1] AAID DJAMEL. Étude numérique comparative entre des méthodes de résolution d'un problème de transport à quatre indices avec capacités. Mentouri-Constantine. 2010.
- [2] C. Roos and T. Terlaky and J. Ph. Vial. Theory and Algorithms for linear optimization. An interior point approach. Jhon Wiley et Sons, Chister, U.K. (1997).
- [3] S. J. Wright. Primal-dual interior point methods. SIAM, Philadelphia. (1997).
- [4] Y.-E. Nesterov, A. Nemirovski, Optimization over positive semidefinite matrices : Mathematical background and user's manual, Technical report, Central economic and mathematical institute, USSR academy of science, Moscow, USSR (1990)
- [5] H. Wolkowicz, G.-P.-H. Styan, Bounds for eigenvalues using traces, Linear Algebra and Appl. 29 (1980), 471-506.
- [6] J.-P. Crouzeix, A. Seeger, New bounds for the extreme values of a finite sample of real numbers, Journal of Mathematical Analysis and Applications 197 (1996), 411-426.
- [7] R.-T. Rockafellar, Convex analysis, Princeton University Press, New Jersey, (1970).
- [8] HAFSI Narimen, Une recherche linéaire avec des fonctions majorantes dans certains problèmes d'optimisation, Université Mentouri – Constantine 2010.

- [9] Leulmi Assma, Une procedure améliorante d'une méthode projective en programmation linéaire, Université Hadaik- Skikda 2007.
- [10] Benterki Djamel, Résolution des problème de programmation semi-défini par des méthode de réduction du potentiel, Université Ferhat Abbas Setif 2004.
- [11] I.Ekeland - R.Temam, Analyse Convexe et Problèmes Variationnels, Dunod, GauthierVillars, Paris, 1974.
- [12] Les Méthodes de Points Intérieurs.
- [13] Anane Nassima, Méthodes de points intérieurs pour la programmation linéaire basées sur les fonctions noyaux, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2012.
- [14] Extension de Quelques Méthodes de Points Intérieurs pour la Programmation Semi-Définie, 2009.
- [15] Yannick PRIVAT, Introduction à l'optimisation aspects théoriques et numériques, Ecole Nationale Supérieure d'électricité et de Mécanique, Paris, février 2015.
- [16] Alexandre d'aspremont, Quelques applications de la programmation semi définie, CNRS et Centre de Mathématiques Appliquées, Ecole Polytechnique, 91128 Palaiseau, France.
- [17] A. Ben-Tal and A. Nemirovski, Lectures on modern convex optimization : analysis, algorithms, and engineering applications, MPS-SIAM series on optimization, Society for Industrial and Applied Mathematics : Mathematical Programming Society, Philadelphia, PA, 2001.
- [18] Optimisation SDP, MAP 557, 23 Octobre 2015.
- [19] Richard Leroy, Polynômes positifs, sommes de carrés, programmation semi-définie positive, application à la minimisation.UFR Mathématique, Université de Rennes 1, Mémoire de DEA. 2003-2004.

ملخص

نعالج في هذه المذكرة التحليل المحدب، بعض المسائل الرياضية، البرمجة الخطية و الثنائية، المشاكل البرمجة الرياضية نصف المعرفة و تحديد طريقة الحاجز اللوغاريتمي في الجانب النظري والخوارزمي. تم تعريف تقريب جديد من الدوال الحدية العليا لحساب الخطوات نتأجج فعالة في تقليل كلفة العمليات الحسابية الخوارزمية مقارنة بطرق البحث التقليدية.

كلمات مفتاحية : البرمجة الخطية، البرمجة نصف المعرفة، البحث الخطي، طريقة الحاجز اللوغاريتمي، الدوال الحدية.

Résumé :

On traite dans ce mémoire l'analyse convexe, quelques rappels d'optimisation, la programmation linéaire et la dualité, les problèmes de programmation semi-définie et en particulier les méthodes barrières logarithmiques sur le plan théorique et algorithmique. L'introduction de la nouvelle approche des fonctions majorantes a contribué efficacement à la réduction du coût calculatoire de l'algorithme comparativement aux méthodes de recherche linéaire classique.

Mots clés : Programmation Linéaire, Programmation Semi-Définie, Recherche Linéaire, Méthode Barrière Logarithmique, Fonctions barrières.