



THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Méthodes de points intérieurs en programmation convexe: le cas particulier de la programmation conique

Adams, Virginie

Award date:
2003

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

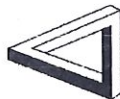
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



FUNDP
Faculté des Sciences
Département de Mathématique

Rempart de la Vierge, 8
B-5000 Namur Belgique

**Méthodes de points intérieurs en
programmation convexe.
Le cas particulier de la programmation
conique.**



Mémoire présenté pour l'obtention
du grade de
Licencié en Sciences Mathématiques
par

Adams Virginie

Promoteur : J-J. Strodiot

Année Académique 2002-2003

Je tiens à remercier tout particulièrement mon promoteur, Monsieur J-J. Strodiot, pour l'aide qu'il m'a apportée. Son soutien, sa disponibilité et ses conseils m'ont été très précieux.

Je remercie également tous les professeurs qui ont contribué à ma formation.

Enfin, je remercie ma famille et mes amis qui m'ont soutenue tout au long de mes études.

Résumé

La programmation convexe consiste à optimiser une fonction convexe sur un ensemble convexe. Ce mémoire s'intéresse aux méthodes de points intérieurs permettant de résoudre ce type de problèmes. Il étudie la théorie de base de ces méthodes. Il examine ensuite le cas de la programmation conique ainsi que la théorie de la dualité qui lui est associée, et en particulier, les cas de la programmation semi-définie et de la programmation du second ordre. Lors de l'étude des méthodes de points intérieurs en programmation conique, nous nous attarderons sur une direction importante, la direction de Nesterov-Todd, et présenterons un algorithme basé sur cette direction.

Abstract

Convex programming consists of optimising a convex function on a convex set. This paper is interested in interior-point methods that allow to resolve this type of problems. It studies basic theory of these methods. Then, the case of conic programming and its duality theory are examined. In particular, cases of semidefinite programming and second-order conic programming will be studied. When studying interior-point methods in conic programming, we will look into an important direction, the Nesterov-Todd direction, and we will present an algorithm based on it.

Table des matières

Introduction	5
1 Théorie de base des méthodes de points intérieurs	7
1.1 Le problème de programmation convexe	7
1.2 La notion de fonction barrière	8
1.3 La notion de fonction self-concordante	9
1.3.1 Quelques notations	9
1.3.2 Définition	12
1.3.3 Propriétés	13
1.3.4 Quelques exemples	16
1.4 Algorithme primal	17
2 Programmation conique et dualité	27
2.1 Introduction	27
2.2 Le problème de programmation conique	28
2.3 Dualité	32
2.3.1 La notion de cône dual	32
2.3.2 Dualité faible	33
2.3.3 Dualité forte	36
2.3.4 Exemples	39
2.4 La fonction conjuguée	40
2.5 Les cônes "self-scaled" (ou symétriques)	48
2.5.1 Définitions et notations	49
2.5.2 Les points scaling	52
2.5.3 Gradients et normes	57
2.5.4 Un théorème utile	65

Introduction

La programmation convexe consiste à optimiser une fonction convexe sur un ensemble convexe. Nous savons déjà qu'il existe des méthodes efficaces qui permettent de résoudre les problèmes de programmation linéaire, et qu'il existe, pour chacune de ces méthodes, une borne supérieure polynomiale sur le nombre d'opérations arithmétiques nécessaires pour atteindre un optimum avec une précision donnée.

Le but de ce mémoire est d'obtenir des méthodes efficaces pour la classe des problèmes convexes. Pour cette très large classe de problèmes, Nesterov et Nemirovsky ont développé en 1994 une théorie complète des méthodes de points intérieurs avec une complexité polynomiale. Ces méthodes sont basées sur l'existence d'une fonction barrière "self-concordante" qui permet de définir une trajectoire centrale menant à l'ensemble solution. Les méthodes de points intérieurs les plus efficaces sont les méthodes primales-duales, qui produisent des solutions aux deux problèmes, primal et dual. Depuis 1994, des extensions de ces méthodes ont vu le jour afin de permettre de résoudre des problèmes convexes non linéaires. Du point de vue pratique, ces algorithmes ont été mis au point pour des problèmes convexes dont la fonction objectif est linéaire et les contraintes sont soit linéaires, soit coniques.

Dans le premier chapitre, nous définissons le problème de programmation convexe et introduisons le concept de fonction barrière "self-concordante", ce qui nous permettra d'obtenir un premier algorithme primal polynomial pour calculer les solutions optimales.

Les chapitres suivants sont consacrés à l'étude de la programmation conique. Dans le chapitre 2, nous commencerons par définir le problème et étudierons la dualité. Nous nous attarderons ensuite sur les notions de fonction conjuguée et de cône self-scaled.

Le troisième chapitre sera consacré à l'étude des problèmes de programmation semi-définie et de programmation conique du second ordre, qui sont deux cas particuliers de la programmation conique, et nous en donnerons quelques exemples.

Dans le dernier chapitre, nous présentons une méthode de points intérieurs primale-duale polynomiale pour résoudre les problèmes de programmation conique. Nous nous pencherons en particulier sur la programmation semi-définie. Nous analyserons plus en détail une des directions de recherche les plus importantes : la direction de Nesterov-Todd. Nous terminerons notre étude par la présentation d'un algorithme basé sur cette direction.

Chapitre 1

Théorie de base des méthodes de points intérieurs

1.1 Le problème de programmation convexe

Un problème d'optimisation convexe peut s'exprimer de plusieurs façons équivalentes.

Considérons le problème suivant :

$$\begin{cases} \min f(x) \\ \text{s.c. } x \in C \end{cases}$$

où $\begin{cases} f : \mathbb{R}^n \rightarrow \mathbb{R} \text{ est une fonction convexe,} \\ C \subseteq \mathbb{R}^n \text{ est un ensemble convexe.} \end{cases}$

Ce problème est équivalent au problème suivant, dont la fonction objectif est linéaire :

$$\begin{cases} \min_{(x,t) \in \mathbb{R}^{n+1}} t \\ \text{s.c. } (x,t) \in \hat{C} \end{cases}$$

où $\hat{C} = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \text{ tq } x \in C \text{ et } f(x) \leq t\}$ est convexe.

Nous constatons aisément que si x^* est une solution du premier problème, alors $(x^*, f(x^*))$ est une solution du second. Inversément, si (x^*, t^*) est une solution du second problème, alors x^* est une solution du premier et $f(x^*) = t^*$.

Dès lors, nous pouvons supposer qu'un problème convexe est de la forme :

$$(P) \begin{cases} \min c^T x \\ \text{s.c. } x \in C \end{cases}$$

où $\begin{cases} c \in \mathbb{R}^n, \\ C \subseteq \mathbb{R}^n \text{ est convexe.} \end{cases}$

Notre but étant d'étudier les méthodes de points intérieurs, nous supposons que l'intérieur de C est non vide. Une manière de représenter un tel ensemble C est d'imposer que l'intérieur de C soit le domaine d'une fonction convexe. Nous introduisons donc une fonction convexe étendue $F : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ et nous définissons C comme la fermeture du domaine de F , i.e.

$$C = \text{cl}(\text{dom } F) = \text{cl} \{x \in \mathbb{R}^n \text{ tq } F(x) < +\infty\}.$$

De plus, afin de générer des points à l'intérieur de C , nous allons également exiger que F soit une fonction barrière pour l'ensemble C .

1.2 La notion de fonction barrière

Définition 1.2.1

Soit C un ensemble convexe fermé dont l'intérieur est non vide. Une fonction $F : \text{int } C \rightarrow \mathbb{R}$ est une fonction barrière pour l'ensemble C si F est strictement convexe et $F(x) \rightarrow +\infty$ lorsque x tend vers la frontière de C .

Quand l'ensemble convexe C est défini à l'aide d'un nombre fini d'inégalités, i.e. lorsque

$$C = \{x \in \mathbb{R}^n \text{ tq } f_i(x) \leq 0, \quad i = 1, \dots, m\}$$

où chacune des m fonctions $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe, il est souvent possible de trouver une fonction barrière qui convient en utilisant la fonction barrière logarithmique

$$F : \text{int } C \rightarrow \mathbb{R}, \quad F(x) = - \sum_{i=1}^m \ln(-f_i(x)).$$

Mais ce n'est pas toujours possible, comme le montre l'exemple qui suit. En effet, si $C = \mathbb{R}^+$ et $f_1(x) = -x^x$, alors $F(x) = -x \ln x$ n'est pas une fonction barrière pour C .

Soit F une fonction barrière pour l'ensemble des contraintes C . Le problème convexe est remplacé par une famille de problèmes de minimisation sans contrainte :

$$(P_\mu) \quad \inf_{x \in \mathbb{R}^n} \frac{c^T x}{\mu} + F(x)$$

où $\mu \in \mathbb{R}_0^+$ ($\equiv \mathbb{R}_{++}$) et est appelé paramètre barrière.

Le terme barrière a pour but d'éviter que les itérés sortent de l'ensemble admissible C .

Si nous supposons l'existence d'une solution $x(\mu)$ pour chacun de ces problèmes, nous appelons l'ensemble $\{x(\mu) \text{ tq } \mu > 0\} \subseteq C$ la trajectoire centrale primale associée au problème (P) . Remarquons que si une telle solution existe, alors elle est unique car F est strictement convexe.

Lorsque μ tend vers 0, le terme $\frac{c^T x}{\mu}$ devient prépondérant dans la fonction objectif. Dès lors, nous pouvons penser que la trajectoire centrale va converger vers la solution optimale du problème (P) .

La stratégie de la plupart des méthodes de points intérieurs est alors de suivre la trajectoire centrale jusqu'à obtenir un point suffisamment proche de la solution optimale.

Les questions sur lesquelles nous allons maintenant nous attarder sont les suivantes : Comment calculer $x(\mu)$ et comment choisir une fonction barrière valable ? Pour tenter de répondre à ces questions, nous allons introduire la notion de fonction self-concordante.

1.3 La notion de fonction self-concordante

Avant de présenter le concept de fonction self-concordante, nous avons besoin de rappeler quelques notations.

1.3.1 Quelques notations

Les dérivées première, seconde et troisième d'une fonction $F : \mathbb{R}^n \rightarrow \mathbb{R}$ évaluées au point x seront notées par $DF(x)$, $D^2F(x)$ et $D^3F(x)$ respectivement. Ces différentielles sont des applications linéaires.

Nous avons :

$$\begin{aligned} DF(x) &: \mathbb{R}^n \rightarrow \mathbb{R} : h_1 \rightarrow DF(x)[h_1], \\ D^2F(x) &: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} : (h_1, h_2) \rightarrow D^2F(x)[h_1, h_2], \\ D^3F(x) &: \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} : (h_1, h_2, h_3) \rightarrow D^3F(x)[h_1, h_2, h_3]. \end{aligned}$$

Lorsque nous choisissons un produit scalaire $\langle \cdot, \cdot \rangle$ sur \mathbb{R}^n , les première et deuxième différentielles peuvent être représentées par un vecteur et une matrice appelés respectivement le vecteur gradient, noté $g(x)$, et la matrice hessienne, notée $H(x)$, de F en x .

$$\text{Nous avons } DF(x)[h] = \langle g(x), h \rangle \text{ et } D^2F(x)[h_1, h_2] = \langle h_1, H(x)h_2 \rangle .$$

Il est clair que le vecteur gradient et la matrice hessienne dépendent du produit scalaire utilisé sur \mathbb{R}^n . Notre but étant d'utiliser la méthode Newton, nous introduisons le pas de Newton en x , noté $n(x)$, ainsi que le produit scalaire intrinsèque, noté $\langle \cdot, \cdot \rangle_x$

$$n(x) \equiv -H(x)^{-1}g(x) \quad \text{et} \quad \langle u, v \rangle_x \equiv \langle u, H(x)v \rangle,$$

qui sont tous deux invariants par transformation affine.

En effet :

- $n(x) \equiv -H(x)^{-1}g(x)$.
Appliquons la transformation linéaire $x = Zy$ où Z est une matrice non singulière.
Alors, si $\bar{f}(y) \equiv f(Zy)$, nous avons :

$$\begin{aligned} \bar{n}(y) &= -\bar{H}(y)^{-1}\bar{g}(y), \\ \bar{g}(y) &= Z^T g(x), \\ \text{et } \bar{H}(y) &= Z^T H(x)Z. \end{aligned}$$

$$\begin{aligned} \text{D'où } \bar{n}(y) &= -\bar{H}(y)^{-1}\bar{g}(y) \\ &= -Z^{-1}H(x)^{-1} \underbrace{Z^{-T}Z^T}_I g(x) \\ &= Z^{-1}n(x) \end{aligned}$$

$$\text{et donc, } n(x) = Z\bar{n}(y).$$

- $\|n(x)\|_x^2 = \langle n(x), H(x)n(x) \rangle = \langle n(x), -g(x) \rangle$.

Appliquons à nouveau la transformation linéaire $x = Zy$ où Z est une matrice non singulière.

Nous avons alors :

$$\begin{aligned} \|n(x)\|_x^2 &= -\langle Z\bar{n}(y), g(x) \rangle && \text{par le point précédent} \\ &= -\langle \bar{n}(y), Z^T g(x) \rangle \\ &= -\langle \bar{n}(y), \bar{g}(y) \rangle && \text{par l'expression de } \bar{g}(y) \\ &= \langle \bar{n}(y), \bar{H}(y)\bar{n}(y) \rangle && \text{par l'expression de } \bar{n}(y) \\ &= \|\bar{n}(y)\|_x^2 \end{aligned}$$

qui est la norme associée à y et \bar{f} .

Remarquons que $\langle u, v \rangle_x \equiv \langle u, H(x)v \rangle$ est un produit scalaire à condition que $H(x)$ soit symétrique et définie positive. Par la suite, nous supposons donc que ces hypothèses sont vérifiées.

Si nous notons respectivement par g_x et H_x le gradient et le hessien de F relativement au produit scalaire intrinsèque $\langle \cdot, \cdot \rangle_x$, nous avons immédiatement que pour tout y , $g_x(y) = H(x)^{-1}g(y)$.

En effet, F est différentiable en y si il existe $g(y) \in \mathbb{R}^n$ tel que

$$\lim_{\|\Delta y\| \rightarrow 0} \frac{f(y + \Delta y) - f(y) - \langle g(y), \Delta y \rangle}{\|\Delta y\|} = 0.$$

Le gradient dépend donc du produit scalaire choisi. Nous devons donc avoir $\langle g(y), \Delta y \rangle = \langle g_x(y), \Delta y \rangle_x$, c'est-à-dire $\langle g(y), \Delta y \rangle = \langle g_x(y), H(x)\Delta y \rangle$, et par conséquent, $g_x(y) = H(x)^{-1}g(y)$.

De manière similaire, nous avons $H_x(y) = H(x)^{-1}H(y)$.

De plus, nous avons :

$$\begin{aligned} n(y) &= -H(y)^{-1}g(y) && \text{par définition} \\ &= -H(y)^{-1}H(x)g_x(y) && \text{par l'expression de } g_x(y) \\ &= -H_x(y)^{-1}g_x(y) && \text{par l'expression de } H_x(y). \end{aligned}$$

Enfin, nous notons par $B_x(y, r)$ la boule ouverte de rayon r centrée en y , où le rayon est mesuré selon la norme $\|\cdot\|_x$.

1.3.2 Définition

Définition 1.3.1

Soit C un sous-ensemble convexe de \mathbb{R}^n dont l'intérieur est non vide et soit $F : \text{int } C \rightarrow \mathbb{R}$ une fonction de classe C^2 dont la matrice hessienne est définie positive sur l'intérieur de C . F est dite self-concordante si $\forall x \in \text{int } C$, on a $B_x(x, 1) \subseteq \text{int } C$ et si, quel que soit $y \in B_x(x, 1)$, on a $\forall h$

$$(1 - \|y - x\|_x)^2 D^2 F(x)[h, h] \leq D^2 F(y)[h, h] \leq \frac{1}{(1 - \|y - x\|_x)^2} D^2 F(x)[h, h]. \quad (1.1)$$

Nous notons SC la famille des fonctions self-concordantes.

La condition (1.1) signifie que le hessien en x peut être utilisé pour estimer le hessien en n'importe quel point $y \in B_x(x, 1)$.

Signalons encore que lorsque F est de classe C^3 , F est self-concordante si et seulement si F est une fonction barrière telle que

$$D^3 F(x)[h, h, h] \leq 2 [D^2 F(x)[h, h]]^{3/2} \quad (1.2)$$

$\forall x \in \text{int } C$ et $\forall h \in \mathbb{R}^n$.

Les fonctions linéaires et quadratiques satisfont évidemment cette condition. Celle-ci signifie que la matrice hessienne change lentement, c'est-à-dire que la fonction F est presque quadratique.

Si nous introduisons la fonction $F_{x,h} : \mathbb{R} \rightarrow \mathbb{R}$, $F_{x,h}(t) = F(x + th)$, où $x \in \text{int } C$ et $h \in \mathbb{R}^n$, nous avons alors :

$$\begin{aligned} F_{x,h}''(t) &= D^2 F(x + th)[h, h] \\ \text{et } F_{x,h}'''(t) &= D^3 F(x + th)[h, h, h]. \end{aligned}$$

La condition (1.2) peut dès lors s'écrire

$$F_{x,h}'''(t) \leq 2F_{x,h}''(t)^{3/2}, \quad \forall x + th \in \text{int } C, \forall h \in \mathbb{R}^n.$$

La raison pour laquelle nous avons introduit la notion de fonction self-concordante est la suivante. Bien que la méthode de Newton soit invariante par transformation affine, l'analyse de complexité dépend du système de coordonnées utilisé. En utilisant une fonction self-concordante, la complexité sera indépendante des changements de coordonnées.

1.3.3 Propriétés

Dans cette section, nous allons présenter quelques propriétés des fonctions self-concordantes.

Proposition 1.3.1

1. Si F_1 et F_2 sont self-concordantes sur l'intérieur de C_1 et C_2 respectivement, et si $\text{int } C_1 \cap \text{int } C_2 \neq \emptyset$, alors $F_1 + F_2$ est self-concordante sur $\text{int } C_1 \cap \text{int } C_2$.
2. Si F est self-concordante sur $\text{int } C \subseteq \mathbb{R}^n$, $b \in \mathbb{R}^m$ et $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est un opérateur linéaire injectif, alors $x \rightarrow F(ax - b)$ est self-concordante, en supposant que le domaine $\{x \text{ tq } Ax - b \in \text{int } C\}$ est non vide. (propriété d'invariance affine)

Preuve :

Nous démontrons les résultats uniquement pour les fonctions à une dimension de classe C^3 .

1. Soient F_1 et F_2 deux fonctions self-concordantes.
 $\forall x \in \text{dom } (f_1 + f_2)$, nous avons $f_1''(x) \geq 0$ et $f_2''(x) \geq 0$.
 De plus, $|f_1'''(x) + f_2'''(x)| \leq 2(f_1''(x)^{3/2} + f_2''(x)^{3/2})$
 $\leq 2(f_1''(x) + f_2''(x))^{3/2}$,
 où la dernière inégalité vient du fait que $(u^{3/2} + v^{3/2})^{2/3} \leq u + v$,
 $\forall u, v \geq 0$.
2. Pour prouver l'invariance affine de la self-concordance, nous posons $\tilde{f}(y) = f(ay + b)$ où $a \neq 0$. Alors \tilde{f} est self-concordante si et seulement si f l'est.
 En effet, $\tilde{f}''(y) = a^2 f''(x)$ et $\tilde{f}'''(y) = a^3 f'''(x)$, où $x = ay + b$.
 Nous avons alors

$$\begin{aligned} |\tilde{f}'''(y)| &\leq 2\tilde{f}''(y)^{3/2} \\ \iff |a^3 f'''(x)| &\leq 2(a^2 f''(x))^{3/2} \\ \iff |f'''(x)| &\leq 2f''(x)^{3/2}. \end{aligned}$$

□

Etant donné le rôle important joué par la direction de Newton dans la théorie des méthodes de points intérieurs, nous présentons maintenant une proposition qui montre comment varie la norme intrinsèque lorsqu'on passe de x à $x_+ \equiv x + n(x)$. Pour établir cette proposition, nous avons besoin du lemme technique suivant, dont nous ne ferons pas la preuve ici.

Lemme 1.3.1

Soit $F : \text{int } C \rightarrow \mathbb{R}$ tel que $B_x(x, 1) \subseteq \text{int } C \ \forall x \in \text{int } C$. Alors les assertions suivantes sont équivalentes :

1. F est self-concordante
2. $\|H_x(y)\|_x, \|H_x(y)^{-1}\|_x \leq \frac{1}{(1 - \|y - x\|_x)^2}, \forall x \in \text{int } C, \forall y \in B_x(x, 1)$.
3. $\|I - H_x(y)\|_x, \|I - H_x(y)^{-1}\|_x \leq \frac{1}{(1 - \|y - x\|_x)^2} - 1, \forall x \in \text{int } C, \forall y \in B_x(x, 1)$.

Proposition 1.3.2

Soit F une fonction self-concordante et soit $x_+ = x + n(x)$.

Si $\|n(x)\|_x < 1$, alors $\|n(x_+)\|_{x_+} \leq \left(\frac{\|n(x)\|_x}{1 - \|n(x)\|_x} \right)^2$.

Preuve :

Nous savons que

$$g_x(x_+) = H(x)^{-1}g(x_+),$$

$$H_x(x_+) = H(x)^{-1}H(x_+),$$

$$\text{et } n(x_+) = -H_x(x_+)^{-1}g_x(x_+).$$

D'où

$$\begin{aligned} \|n(x_+)\|_{x_+}^2 &= \|H_x(x_+)^{-1}g_x(x_+)\|_{x_+}^2 \\ &= \langle H_x(x_+)^{-1}g_x(x_+), H(x_+)H_x(x_+)^{-1}g_x(x_+) \rangle \\ &= \langle H_x(x_+)^{-1}g_x(x_+), H(x) \underbrace{H_x(x_+)H_x(x_+)^{-1}}_I g_x(x_+) \rangle \\ &= \langle H_x(x_+)^{-1}g_x(x_+), g_x(x_+) \rangle_x \\ &\leq \|H_x(x_+)^{-1}\|_x \|g_x(x_+)\|_x^2 \\ &\leq \frac{\|g_x(x_+)\|_x^2}{(1 - \|n(x)\|_x)^2} \quad \text{par le lemme.} \end{aligned}$$

Comme $g_x(x) = -n(x)$, nous avons :

$$\begin{aligned}
 \|g_x(x_+)\|_x &= \|g_x(x_+) - g_x(x) - n(x)\|_x \\
 &= \left\| \int_0^1 [H_x(x + tn(x)) - I]n(x)dt \right\|_x \\
 &\leq \|n(x)\|_x \int_0^1 \|I - H_x(x + tn(x))\|_x dt \\
 &\leq \|n(x)\|_x \int_0^1 \left(\frac{1}{(1 - t\|n(x)\|_x)^2} - 1 \right) dt \quad \text{par le lemme} \\
 &= \|n(x)\|_x \left(\frac{\|n(x)\|_x}{1 - \|n(x)\|_x} \right) \\
 &= \frac{\|n(x)\|_x^2}{1 - \|n(x)\|_x}.
 \end{aligned}$$

D'où, en combinant les deux, il suit que

$$\|n(x_+)\|_{x_+} \leq \frac{\|g_x(x_+)\|_x}{(1 - \|n(x)\|_x)} \leq \left(\frac{\|n(x)\|_x}{1 - \|n(x)\|_x} \right)^2.$$

□

Afin d'étudier la complexité des méthodes de points intérieurs, nous avons besoin de la définition suivante.

Définition 1.3.2

Une fonction $F : \text{int } C \rightarrow \mathbb{R}$ est dite fonction barrière fortement self-concordante si $F \in SC$ et

$$\theta_F \equiv \sup_{x \in \text{int } C} \|g_x(x)\|_x^2 < +\infty.$$

Le paramètre θ_F est appelé la valeur de complexité de F . Nous notons SCB la famille des fonctions barrières fortement self-concordantes. Par la suite, nous omettrons le mot fortement.

La proposition suivante sur les fonctions barrières self-concordantes nous sera utile plus tard.

Proposition 1.3.3

Supposons que $F \in SCB$. Si $x, y \in \text{int } C$, alors $\langle g(x), y - x \rangle < \theta_F$.

Bien qu'il soit souhaitable d'avoir des fonctions barrières self-concordantes avec de petites valeurs de complexité, il existe une valeur minimale en dessous de laquelle θ_F ne peut pas descendre. Nesterov et Nemirovsky ont montré que $\theta_F \geq 1 \forall F \in SCB$.

Nous pouvons montrer que si F_1 et $F_2 \in SCB$ et $\text{int } C_1 \cap \text{int } C_2 \neq \emptyset$, alors $F = F_1 + F_2 \in SCB$ et $\theta_F \leq \theta_{F_1} + \theta_{F_2}$.

1.3.4 Quelques exemples

1. $F : \mathbb{R} \rightarrow \mathbb{R}$, $F(x) = x^4$ n'est pas self-concordante.

En effet, $F'(x) = 4x^3$, $F''(x) = 12x^2$ et $F'''(x) = 24x$.

Par la définition de fonction self-concordante, nous devrions avoir

$$\frac{F'''(x)^2}{4F''(x)^3} \leq 1.$$

Or,

$$\frac{F'''(x)^2}{4F''(x)^3} = \frac{(24x)^2}{4(12x^2)^3} = \frac{1}{12x^4},$$

qui tend vers l'infini lorsque x tend vers zéro.

2. La fonction barrière logarithmique $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ définie par

$f(x) = -\ln x$ est self-concordante.

En effet, en utilisant $f''(x) = \frac{1}{x^2}$ et $f'''(x) = -\frac{2}{x^3}$, nous obtenons

$$\frac{|f'''(x)|}{2f''(x)^{3/2}} = \frac{2/x^3}{2(1/x^2)^{3/2}} = 1.$$

3. La fonction $f(x) = -\sum_{i=1}^m \ln(b_i - a_i^T x)$,

où int $C = \{x \text{ tq } a_i^T x < b_i, i = 1, \dots, m\}$, est self-concordante.

En effet, $-\ln y$ est une fonction self-concordante par l'exemple qui précède.

Posons $y = b_i - a_i^T x$.

Par la proposition 1.3.1, $-\ln(b_i - a_i^T x)$ est une fonction self-concordante,

de même que $-\sum_{i=1}^m \ln(b_i - a_i^T x)$.

4. Nous savons grâce aux deux exemples précédents que $F(x) = -\sum_{i=1}^n \ln x_i$

est self-concordante sur \mathbb{R}_{++}^n . C'est aussi une fonction barrière self-concordante.

En effet, $g(x)$ est le vecteur à n dimensions $\frac{-1}{x_i}$ et le hessien $H(x)$ est la matrice diagonale des $\frac{1}{x_i^2}$.

D'où

$$\begin{aligned} \|g_x(x)\|_x^2 &= g(x)^T H(x)^{-1} g(x) \\ &= \left(\frac{-1}{x_1} \dots \frac{-1}{x_n} \right) \begin{pmatrix} x_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & x_n^2 \end{pmatrix} \begin{pmatrix} \frac{-1}{x_1} \\ \vdots \\ \frac{-1}{x_n} \end{pmatrix} \\ &= n \quad \forall x \in \mathbb{R}_{++}^n. \end{aligned}$$

Donc $\theta_F = \sup \|g_x(x)\|_x^2 = n$.

1.4 Algorithme primal

Considérons le problème convexe suivant :

$$(P) \begin{cases} \min & \langle c, x \rangle \\ \text{s.c.} & x \in C \end{cases}$$

où $\begin{cases} c \in \mathbb{R}^n, \\ C \subseteq \mathbb{R}^n \text{ est un ensemble fermé et convexe dont l'intérieur est} \\ \text{non vide.} \end{cases}$

Soit aussi F une fonction self-concordante.

Nous savons que les méthodes de points intérieurs de suivi de chemin résolvent (P) en suivant la trajectoire centrale. Celle-ci est composée des minimisants $x(\mu)$ de la fonction self-concordante

$$F_\mu(x) = \frac{1}{\mu} \langle c, x \rangle + F(x) \text{ pour } \mu \rightarrow 0^+.$$

Nous démontrerons l'existence et l'unicité de $x(\mu)$ dans le chapitre suivant, pour un cône convexe fermé C.

La proposition qui suit compare la valeur $\langle c, x(\mu) \rangle$ avec la valeur optimale p^* du problème (P). De façon plus générale, compte tenu du fait que nous suivons approximativement la trajectoire centrale, nous comparons la valeur de $\langle c, y \rangle$ avec p^* , pour $y \in C$.

Proposition 1.4.1

Soit F une fonction barrière self-concordante sur l'intérieur de C et soit $x(\mu)$ l'unique minimum de F_μ .

Alors

1. $\langle c, x(\mu) \rangle \leq p^* + \mu\theta_F$.
2. $\langle c, y \rangle \leq p^* + \mu\theta_F[1 + \|y - x(\mu)\|_{x(\mu)}] \quad \forall y \in C$.

Preuve :

1. Par hypothèse, $x(\mu)$ est l'unique minimum de $F_\mu(x)$.

Donc,

$$g(F_\mu(x)) = 0,$$

c'est-à-dire

$$\frac{c}{\mu} + g(x(\mu)) = 0,$$

ou encore

$$c = -\mu g(x(\mu)).$$

De plus, nous avons $\forall y \in \text{int } C$

$$\begin{aligned} \langle c, x(\mu) \rangle - \langle c, y \rangle &= \langle -c, y - x(\mu) \rangle \\ &= \langle \mu g(x(\mu)), y - x(\mu) \rangle \\ &< \mu\theta_F \quad \text{par la proposition 1.3.3} \end{aligned}$$

D'où, par passage à la limite, $\langle c, x(\mu) \rangle \leq p^* + \mu\theta_F$.

En particulier, pour $x = x(\mu)$, nous avons

$$\begin{aligned}\langle c, y \rangle - p^* &\leq [\langle c, x(\mu) \rangle - p^*][1 + \|y - x(\mu)\|_{x(\mu)}] \\ &\leq \mu\theta_F[1 + \|y - x(\mu)\|_{x(\mu)}]\end{aligned}$$

où la dernière inégalité est obtenue par la première partie de la proposition.

□

Remarquons que lorsque $\mu \rightarrow 0^+$, nous avons $\langle c, x(\mu) \rangle \rightarrow p^*$, ce qui signifie que la valeur optimale du problème (P) est atteinte en suivant la trajectoire centrale.

Etant donné qu'il est trop coûteux de calculer les points se trouvant sur la trajectoire centrale, nous allons imposer que les itérés restent dans un voisinage imposé de la trajectoire centrale et nous ferons tendre le paramètre barrière μ_k vers 0.

Soit x_k le $k^{\text{ième}}$ itéré. Nous pouvons définir une mesure de proximité en utilisant le pas de Newton en x_k en essayant de minimiser la fonction objectif $F_{\mu_k} \equiv \frac{\langle c, x \rangle}{\mu_k} + F(x)$, i.e., de trouver $x(\mu_k)$.

Nous avons donc (en omettant l'indice k) :

$$\begin{aligned}n_\mu(x) &= -(\nabla^2 F_\mu(x))^{-1} \nabla F_\mu(x) \\ &= -H(x)^{-1} \left(\frac{c}{\mu} + g(x) \right) \\ &= -\frac{1}{\mu} H(x)^{-1} c + n(x).\end{aligned}\tag{1.4}$$

Nous pouvons alors définir une mesure de proximité $\delta(x, \mu)$ de x au point central $x(\mu)$, comme la norme intrinsèque du pas de Newton $n_\mu(x)$, i.e. $\delta(x, \mu) \equiv \|n_\mu(x)\|_x$.

Ce choix peut être justifié par le fait que le pas de Newton en x visant $x(\mu)$ est supposé être une bonne approximation de la distance $x - x(\mu)$.

Le but d'une méthode au pas court est de suivre approximativement la trajectoire centrale. Nous allons donc imposer que les itérés restent assez proches de celle-ci.

La méthode au pas court pour résoudre le problème (P) a donc la forme suivante.

Algorithme 1.4.1 (La méthode barrière au pas court)

Etant donné une fonction barrière self-concordante F sur C , une borne supérieure sur la mesure de proximité $\tau > 0$, un paramètre de décroissance $0 < \gamma < 1$, un paramètre de tolérance $\epsilon > 0$ et un itéré initial x_0 tel que $\delta(x_0, \mu_0) < \tau$, poser $k = 0$ et répéter les étapes suivantes jusqu'à ce que $\mu_k < \epsilon$:

1. Diminution du paramètre barrière

$$\mu_{k+1} = (1 - \gamma)\mu_k$$

2. Prendre un pas de Newton pour $F_{\mu_{k+1}}$

$$x_{k+1} = x_k + n_{\mu_{k+1}}(x_k)$$

3. $k := k + 1$

Il reste maintenant à résoudre le problème suivant : comment choisir les paramètres τ et γ de manière à ce que la proximité de la trajectoire centrale soit préservée et tels que $\delta(x_k, \mu_k) < \tau \Rightarrow \delta(x_{k+1}, \mu_{k+1}) < \tau$? Grâce à la self-concordance de la fonction barrière, nous pouvons déjà affirmer qu'il est possible de trouver des valeurs de τ et γ qui vérifient ces conditions.

Afin d'établir un rapport entre les deux proximités $\delta(x_k, \mu_k)$ et $\delta(x_{k+1}, \mu_{k+1})$, nous allons considérer une mesure intermédiaire $\delta(x_k, \mu_{k+1})$, la proximité d'un itéré à la cible suivante sur la trajectoire centrale.

Si nous notons x l'itéré en cours et x_+ l'itéré suivant, μ le paramètre barrière associé à l'itéré x et μ_+ le paramètre barrière associé à l'itéré x_+ , nous allons considérer $\delta(x, \mu_+) = \|n_{\mu_+}(x)\|_x$.

Nous allons montrer que

$$1. \delta(x, \mu) < \tau < 1 \Rightarrow \delta(x, \mu_+) \leq \frac{\tau + \gamma\sqrt{\theta_F}}{1 - \gamma} \equiv \alpha.$$

$$2. \delta(x, \mu_+) < 1 \Rightarrow \delta(x_+, \mu_+) \leq \left(\frac{\delta(x, \mu_+)}{1 - \delta(x, \mu_+)} \right)^2.$$

Dès lors, si $\alpha < 1$, nous aurons $\delta(x_+, \mu_+) \leq \left(\frac{\alpha}{1-\alpha}\right)^2$. Il suffit donc de choisir τ et γ tels que $0 < \tau < 1$, $0 < \gamma < 1$, $\left(\frac{\alpha}{1-\alpha}\right)^2 \leq \tau$. Par exemple, $\tau = \frac{1}{9}$ et $\gamma = \frac{1}{1+8\sqrt{\theta_F}}$ conviennent.

Nous pouvons établir les relations suivantes entre ces trois mesures.

Proposition 1.4.2

Soit F une fonction barrière telle que $\theta_F < +\infty$.

Soit $x \in \text{int } C$ et $\mu_+ = (1 - \gamma)\mu$.

Alors

$$\delta(x, \mu_+) \leq \frac{\delta(x, \mu) + \gamma\sqrt{\theta_F}}{1 - \gamma}.$$

Preuve :

Par (1.4), nous savons que

$$\begin{aligned} n_\mu(x) &= -\frac{1}{\mu}H(x)^{-1}c + n(x) \\ \text{et } n_{\mu_+}(x) &= -\frac{1}{\mu_+}H(x)^{-1}c + n(x). \end{aligned}$$

En combinant les deux égalités, nous obtenons :

$$\mu_+n_{\mu_+}(x) - \mu_+n(x) = -H(x)^{-1}c = \mu n_\mu(x) - \mu n(x).$$

En divisant par μ et en utilisant $\mu_+ = (1 - \gamma)\mu$, nous avons :

$$(1 - \gamma)n_{\mu_+}(x) - (1 - \gamma)n(x) = n_\mu(x) - n(x),$$

ce qui est équivalent à

$$(1 - \gamma)n_{\mu_+}(x) = n_\mu(x) - \gamma n(x).$$

D'où

$$(1 - \gamma)\|n_{\mu_+}(x)\|_x \leq \|n_\mu(x)\|_x + \gamma\|n(x)\|_x,$$

ou encore, par la définition de la mesure de proximité et la définition de fonction barrière self-concordante,

$$(1 - \gamma)\delta(x, \mu_+) \leq \delta(x, \mu) + \gamma\sqrt{\theta_F}.$$

□

Proposition 1.4.3

Soit F une fonction self-concordante et $x \in \text{int } C$. Soit $\mu_+ = (1 - \gamma)\mu$ et $x_+ = x + n_{\mu_+}(x)$. On suppose aussi que $\delta(x, \mu_+) < 1$.

Alors $x_+ \in \text{int } C$ et $\delta(x_+, \mu_+) \leq \frac{\delta(x, \mu_+)^2}{(1 - \delta(x, \mu_+))^2}$.

Preuve :

- $x_+ \in \text{int } C$. En effet, $\|x_+ - x\|_x \equiv \|n_{\mu_+}(x)\|_x = \delta(x, \mu_+) < 1$ par hypothèse.
Donc $x_+ \in B_x(x, 1)$ et $B_x(x, 1) \subseteq \text{int } C$ car F est self-concordante.
- Comme la fonction F_{μ_+} est self-concordante, en appliquant la proposition 1.3.2, nous avons :

$$\begin{aligned} \|n_{\mu_+}(x)\|_x < 1 \Rightarrow \|n_{\mu_+}(x_+)\|_{x_+} &\leq \left(\frac{\|n_{\mu_+}(x)\|_x}{1 - \|n_{\mu_+}(x)\|_x} \right)^2 \\ &= \left(\frac{\delta(x, \mu_+)}{1 - \delta(x, \mu_+)} \right)^2. \end{aligned}$$

□

En utilisant les deux propositions que nous venons de démontrer, nous obtenons que tous les itérés restent dans un voisinage de la trajectoire centrale.

Il reste maintenant à montrer que la seconde condition, à savoir $\delta(x, \mu) < \tau \Rightarrow \delta(x_+, \mu_+) < \tau$, est également vérifiée lorsque les paramètres τ et γ sont bien choisis. C'est le but de la proposition suivante.

Proposition 1.4.4

Soit F une fonction barrière self-concordante et soit $x \in \text{int } C$.

Si $0 < \tau < 1$ et $0 < \gamma < 1$ sont tels que

$$\alpha \equiv \frac{\tau + \gamma\sqrt{\theta_F}}{1 - \gamma} < 1 \quad \text{et} \quad \left(\frac{\alpha}{1 - \alpha} \right)^2 \leq \tau,$$

alors $\delta(x, \mu) < \tau \Rightarrow \delta(x_+, \mu_+) < \tau$.

En particulier, $\tau = \frac{1}{9}$ et $\gamma = \frac{1}{1+8\sqrt{\theta_F}}$ satisfont ces conditions.

Preuve :

Soit $\delta(x, \mu) < \tau$. Comme $\tau < 1$ par hypothèse, il suit que

$$\begin{aligned} \delta(x, \mu_+) &\leq \frac{\delta(x, \mu) + \gamma\sqrt{\theta_F}}{1 - \gamma} && \text{par la proposition 1.4.2} \\ &< \frac{\tau + \gamma\sqrt{\theta_F}}{1 - \gamma} && \text{par hypothèse} \\ &< 1 && \text{par hypothèse} \end{aligned}$$

$$\begin{aligned} \text{et } \delta(x_+, \mu_+) &\leq \frac{\delta(x, \mu_+)^2}{(1 - \delta(x, \mu_+))^2} && \text{par la proposition 1.4.3} \\ &< \frac{\left(\frac{\tau + \gamma\sqrt{\theta_F}}{1 - \gamma}\right)^2}{\left(1 - \frac{\tau + \gamma\sqrt{\theta_F}}{1 - \gamma}\right)^2} \\ &= \left(\frac{\alpha}{1 - \alpha}\right)^2 && \text{par définition} \\ &\leq \tau && \text{par hypothèse.} \end{aligned}$$

□

La dernière proposition s'intéresse à la complexité de la méthode barrière au pas court.

Proposition 1.4.5

1. Soit $\{\mu_k\}$ la suite générée par l'algorithme barrière au pas court avec $\gamma = \frac{1}{1+8\sqrt{\theta_F}}$ et soit $\mu < \mu_0$.

Alors, $\mu_k \leq \mu \quad \forall k \geq K$, où $K = \mathcal{O}\left(\sqrt{\theta_F} \ln \frac{\mu_0}{\mu}\right)$.

2. Soit $\epsilon > 0$ donné.

Alors

$$\mathcal{O}\left(\sqrt{\theta_F} \ln \left(\frac{\theta_F \mu_0}{\epsilon}\right)\right) \tag{1.5}$$

itérations suffisent pour avoir x qui satisfait $\langle c, x \rangle \leq p^* + \epsilon$.

Preuve :

1. Nous avons $\mu_k = (1 - \gamma)\mu_{k-1} = \dots = (1 - \gamma)^k \mu_0$. Nous obtenons $\mu_k \leq \mu$ si $(1 - \gamma)^k \mu_0 \leq \mu$ c'est-à-dire si $k \ln(1 - \gamma) \leq \ln \frac{\mu}{\mu_0}$. Comme $\ln(1 - \gamma) < 0$, ceci est équivalent à $k \geq \frac{\ln \frac{\mu}{\mu_0}}{\ln(1 - \gamma)}$.

En utilisant la définition de γ et l'inégalité $\ln(1 + x) > \frac{x}{1 + x} \quad \forall x > 0$, il suit que

$$\begin{aligned} -\ln(1 - \gamma) &= \ln\left(1 + \frac{1}{8\sqrt{\theta_F}}\right) \\ &> \frac{1}{8\sqrt{\theta_F}} \frac{1}{1 + \frac{1}{8\sqrt{\theta_F}}} \\ &= \frac{1}{8\sqrt{\theta_F} + 1} \\ &\geq \frac{1}{10\sqrt{\theta_F}} \quad \text{car } \theta_F \geq 1. \end{aligned}$$

Finalement, nous obtenons $\frac{\ln \frac{\mu_0}{\mu}}{-\ln(1 - \gamma)} \leq 10\sqrt{\theta_F} \ln \frac{\mu_0}{\mu}$.

D'où $\forall k \geq 10\sqrt{\theta_F} \ln \frac{\mu_0}{\mu}$, nous avons $\mu_k \leq \mu$.

2. Soit $\epsilon > 0$. Par la deuxième partie de la proposition 1.4.1, nous avons que $\forall k$,

$$\langle c, x_k \rangle - p^* \leq \mu_k \theta_F (1 + \|x_k - x(\mu_k)\|_{x(\mu_k)}).$$

En utilisant la définition de fonction self-concordante, nous pouvons montrer que chaque x généré par l'algorithme est tel que $\|x - x(\mu)\|_x \leq \frac{1}{6}$ et $\|x - x(\mu)\|_{x(\mu)} \leq \frac{1}{5}$.

Nous obtenons alors

$$\langle c, x_k \rangle - p^* \leq \frac{6}{5} \mu_k \theta_F. \quad (1.6)$$

Dès lors, si $\mu_k \leq \frac{5\epsilon}{6\theta_F}$, $\langle c, x_k \rangle - p^* \leq \epsilon$.

En utilisant la première partie de la preuve avec $\mu = \frac{5\epsilon}{6\theta_F}$, nous obtenons

$$k \geq 10\sqrt{\theta_F} \ln\left(\frac{6\theta_F}{5\epsilon} \mu_0\right) \Rightarrow \mu_k \leq \mu = \frac{5\epsilon}{6\theta_F},$$

ce qui est équivalent à $\frac{6}{5} \mu_k \theta_F \leq \epsilon$, ou encore, par (1.6), à $\langle c, x_k \rangle - p^* \leq \epsilon$.

□

Nous terminons ce chapitre par quelques remarques. Signalons tout d'abord qu'aucune méthode de points intérieurs n'a atteint de meilleures bornes de complexité que (1.5), même dans le cas restreint de la programmation linéaire où $\theta_F = n$. Cependant, la méthode barrière au pas court n'est pas efficace en pratique, comparé à d'autres méthodes de points intérieurs, comme les méthodes primales-duales par exemple.

Mentionnons enfin le fait que nous avons supposé tout au long du chapitre que l'ensemble convexe C avait un intérieur non vide, ce qui nous a permis de définir les fonctions barrières sur $\text{int } C$. Cette hypothèse n'est plus valable lorsqu'il y a des contraintes linéaires d'égalité comme en programmation linéaire.

Chapitre 2

Programmation conique et dualité

2.1 Introduction

La méthode développée au chapitre précédent a deux inconvénients. Tout d'abord, elle est inefficace pour la résolution de problèmes pratiques, bien que sa complexité soit la meilleure obtenue jusqu'à ce jour. Ensuite, elle nécessite l'emploi d'une fonction barrière self-concordante qui est souvent difficile à trouver lorsque C est un ensemble convexe arbitraire. Cependant, cette tâche peut se révéler assez simple lorsque C est un cône convexe.

Dans ce chapitre, nous allons donc nous intéresser aux problèmes dont l'ensemble des contraintes C a une structure particulière, à savoir les problèmes dont l'ensemble des contraintes C est un cône convexe fermé pointé dont l'intérieur est non vide. Ce genre de problèmes sera appelé problème de programmation conique.

De nombreux problèmes pratiques peuvent être modélisés sous cette forme et il est même possible de montrer que tout problème convexe est équivalent à un problème de programmation conique. Nous distinguerons trois cônes : le cône du second ordre, le cône des matrices symétriques semi-définies positives et l'orthant non négatif \mathbb{R}_+^n , donnant lieu respectivement à la programmation conique du second ordre, à la programmation semi-définie et à la programmation linéaire. Pour ces trois cônes, il est possible de construire une fonction barrière self-concordante et d'établir un algorithme primal de suivi de chemin au pas court avec la même complexité qu'en programmation linéaire. Un autre avantage de la programmation conique est que la théorie de la dualité, que nous étudierons dans ce chapitre, est très riche et assez semblable à la dualité en programmation linéaire.

De plus, il existe des fonctions barrières ayant des propriétés intéressantes qui nous permettront de définir une trajectoire centrale et d'établir des méthodes de suivi de chemin primales-duales pour calculer les solutions optimales. Ces méthodes seront efficaces, non seulement du point de vue de la complexité, mais aussi pour résoudre des problèmes pratiques. Ces méthodes sont en fait la généralisation à la programmation conique des méthodes de points intérieurs pour la programmation linéaire.

2.2 Le problème de programmation conique

Afin d'introduire le problème de programmation conique, considérons le problème convexe sous sa forme la plus générale :

$$(P) \begin{cases} \min & c^T x \\ \text{s.c.} & x \in C \end{cases}$$

où $c \in \mathbb{R}^n$ et $C \subseteq \mathbb{R}^n$.

Nous supposons ici que C est défini par des égalités et des inégalités linéaires, avec un ordre qui tient compte de la non linéarité de C . Par exemple, nous considérons des contraintes telles que $Ax \geq b$, mais avec un autre ordre que l'ordre habituel. Afin de définir correctement ce nouvel ordre, rappelons pour commencer les propriétés principales de l'ordre classique.

L'ordre partiel des vecteurs de \mathbb{R}^n est défini par :

$$a \geq b \iff a_i \geq b_i, \quad i = 1, \dots, n \iff a - b \in \mathbb{R}_+^n.$$

Cet ordre satisfait les propriétés suivantes :

- réflexivité : $a \geq a$.
- antisymétrie : $a \geq b$ et $b \geq a \Rightarrow a = b$.
- transitivité : $a \geq b$ et $b \geq c \Rightarrow a \geq c$.

De plus, cet ordre est compatible avec les opérateurs linéaires dans le sens où :

- $a \geq b$ et $\lambda \in \mathbb{R}_+ \Rightarrow \lambda a \geq \lambda b$.
- $a \geq b$ et $c \geq d \Rightarrow a + c \geq b + d$.

Remarquons enfin que l'ordre est complètement identifié par l'orthant non négatif \mathbb{R}_+^n .

L'idée est donc de remplacer cet ensemble \mathbb{R}_+^n par un autre sous-ensemble de \mathbb{R}^n , noté K . Le nouvel ordre (associé à K) est défini par :

$$a \geq_K b \iff a - b \in K.$$

L'ensemble K sera l'ensemble des vecteurs non négatifs pour le nouvel ordre.

Afin que les propriétés citées ci-dessus soient vérifiées, K ne peut pas être arbitraire.

Définition 2.2.1

Un ensemble K est un cône s'il est fermé pour la multiplication scalaire non négative, i.e. $x \in K$ et $\lambda \in \mathbb{R}_+ \Rightarrow \lambda x \in K$.

Un cône K est pointé si $K \cap (-K) = \{0\}$, où $-K = \{x \text{ tq } -x \in K\}$.

De ces définitions, il est aisé de voir qu'un cône pointé n'est jamais vide et ne contient pas de ligne droite. De plus, on vérifie facilement qu'un cône est convexe si et seulement si il est fermé pour l'addition,

$$\text{i.e. } x, y \in K \Rightarrow x + y \in K.$$

En effet, si K est convexe et $x, y \in K$, nous avons :

$$x + y = 2 \left(\frac{1}{2}x + \frac{1}{2}y \right) \in K.$$

Inversément, si $x, y \in K$ et $x + y \in K$, nous avons :

$$\forall t \in [0, 1] : tx + (1 - t)y \in K,$$

et par conséquent, K est convexe.

Nous allons montrer que lorsque K est un cône convexe pointé, l'ordre associé à K est bien un ordre.

En effet,

- réflexivité :

$$a \geq_K a \iff a - a \in K \iff 0 \in K$$

où $0 \in K$ car K est pointé.

- antisymétrie :

$$\begin{aligned}
 a \geq_K b \text{ et } b \geq_K a &\iff a - b \in K \text{ et } b - a \in K \\
 &\iff a - b \in K \text{ et } a - b \in (-K) \\
 &\implies a - b = 0 \text{ car } K \text{ est pointé} \\
 &\iff a = b.
 \end{aligned}$$

- transitivité :

$$\begin{aligned}
 a \geq_K b \text{ et } b \geq_K c &\iff a - b \in K \text{ et } b - c \in K \\
 &\implies a - b + b - c \in K \text{ car } K \text{ est convexe} \\
 &\iff a - c \in K \\
 &\iff a \geq_K c.
 \end{aligned}$$

- compatibilité avec les opérateurs linéaires :

1.

$$\begin{aligned}
 a \geq_K b &\iff a - b \in K \\
 &\implies \lambda(a - b) \in K \text{ car } \lambda \in \mathbb{R}_+ \\
 &\iff \lambda a \geq_K \lambda b.
 \end{aligned}$$

2.

$$\begin{aligned}
 a \geq_K b \text{ et } c \geq_K d &\iff a - b \in K \text{ et } c - d \in K \\
 &\implies a - b + c - d \in K \text{ car } K \text{ est convexe} \\
 &\iff (a + c) - (b + d) \in K \\
 &\iff a + c \geq_K b + d.
 \end{aligned}$$

Remarquons enfin que l'orthant non négatif n'est pas simplement un cône convexe pointé, mais aussi un cône fermé dont l'intérieur est non vide. Le fait qu'il soit fermé nous permet de garder le sens d'une inégalité lors du passage à la limite, i.e. si $a_k \geq b_k$ avec $a_k \rightarrow a$ et $b_k \rightarrow b$, alors $a \geq b$; le fait d'avoir un intérieur non vide est important pour définir l'inégalité stricte. En effet, $a > b \iff a - b \in \mathbb{R}_{++}^n$.

Dès lors, à partir de maintenant, nous supposons que l'ensemble K est un cône convexe fermé pointé dont l'intérieur est non vide.

Nous pouvons maintenant définir un problème d'optimisation conique.

Soit K un cône convexe fermé pointé dont l'intérieur est non vide. Soit $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ un opérateur linéaire, et soit $c \in \mathbb{R}^n$ et $b \in \mathbb{R}^m$. Alors, le problème primal d'optimisation conique s'écrit :

$$(CP) \begin{cases} \min & c^T x \\ \text{s.c.} & Ax = b \\ & x \in K. \end{cases}$$

Lorsque $K = \mathbb{R}_+^n$, le problème (CP) est exactement un problème de programmation linéaire.

Nous considérons deux cas particuliers :

1. Le cône de Lorentz, défini par :

$$L^n = \left\{ x = (x_1, \dots, x_n)^T \in \mathbb{R}^n \text{ tq } x_n \geq \sqrt{\sum_{i=1}^{n-1} x_i^2} \right\}.$$

Ce qui peut également s'écrire :

$$L^{n+1} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \text{ tq } \|x\|_2 \leq t\},$$

ce qui montre que L^{n+1} est l'épigraphe de la norme euclidienne.

Lorsque ce cône est utilisé, le problème d'optimisation conique correspondant est appelé problème de programmation conique du second ordre.

Un exemple typique de ce type de problème est le suivant :

$$(SOCP) \left\{ \begin{array}{l} \min \sum_{i=1}^m c_i^T x_i + d_i t_i \\ \text{s.c.} \sum_{i=1}^m A_i x_i + a_i t_i = b \\ \|x_i\|_2 \leq t_i, \quad i = 1, \dots, m \end{array} \right.$$

où $x_i \in \mathbb{R}^{n_i}$, $t_i \in \mathbb{R}$, $i = 1, \dots, m$ sont les variables d'optimisation et $c_i \in \mathbb{R}^{n_i}$, $d_i \in \mathbb{R}$, $A_i \in \mathbb{R}^{n \times n_i}$, $a_i \in \mathbb{R}^n$ et $b \in \mathbb{R}^n$ sont les paramètres du problème.

Ici, $K = K_1 \times \dots \times K_m$ où $K_i = L^{n_i+1} = \{(x_i, t_i) \text{ tq } \|x_i\|_2 \leq t_i\}$.

2. Le cône semi-défini positif S_+^n . Ce cône est un sous-ensemble de l'espace S^n des matrices $n \times n$ symétriques et est constitué de toutes les matrices $n \times n$ symétriques semi-définies positives, c'est-à-dire les matrices telles que $A^T = A$ et $x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n$.

Lorsque le cône S_+^n est utilisé, le problème est appelé problème de programmation semi-définie. Dans ce cas, l'ordre est noté par le symbole \succeq .

Un exemple typique de problème de programmation semi-définie est :

$$(SDP) \begin{cases} \min & \langle C, X \rangle \\ \text{s.c.} & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \\ & X = X^T \succeq 0 \end{cases}$$

$$\text{où} \begin{cases} X, A_i, C \in S^n, b_i \in \mathbb{R}, \quad i = 1, \dots, m, \\ \langle C, X \rangle = \text{tr}(CX) = \sum_{j,k=1}^n C_{jk} X_{jk}, \\ A : S^n \rightarrow \mathbb{R}^m, \quad A(X) = (\langle A_1, X \rangle, \dots, \langle A_m, X \rangle). \end{cases}$$

Ces deux cas particuliers seront étudiés plus en détails dans le chapitre qui suit.

Nous pouvons montrer que la formulation conique nous permet d'exprimer tous les problèmes d'optimisation convexe sans aucune perte de généralité.

2.3 Dualité

2.3.1 La notion de cône dual

L'avantage de la formulation conique est que le dual Lagrangien d'un problème conique peut s'exprimer comme un problème conique, grâce à la notion de cône dual.

Définition 2.3.1

Le cône dual d'un cône K dans \mathbb{R}^n est défini par

$$K^* = \{x^* \in \mathbb{R}^n \text{ tq } x^T x^* \geq 0, \quad \forall x \in K\}.$$

Un cône K est dit auto-dual si $K^* = K$.

L'orthant non négatif \mathbb{R}_+^n est auto-dual. Le cône dual d'un sous-espace linéaire est le sous-espace vectoriel qui lui est orthogonal.

La proposition qui suit montre ce que deviennent les propriétés d'un cône lorsqu'on passe à son dual.

Proposition 2.3.1

1. Le dual d'un cône convexe fermé est un cône convexe fermé.
2. Le dual d'un cône convexe fermé pointé dont l'intérieur est non vide a les mêmes propriétés.
3. $(K^*)^* = K$ est un cône convexe fermé.

2.3.2 Dualité faible

Le problème dual associé au problème primal

$$(CP) \begin{cases} \min & c^T x \\ \text{s.c.} & Ax = b \\ & x \in K \end{cases}$$

est défini par

$$(CD) \begin{cases} \max & b^T y \\ \text{s.c.} & c - A^T y \in K^*. \end{cases}$$

Très souvent, ce problème s'écrit sous la forme équivalente suivante :

$$(CD') \begin{cases} \max & b^T y \\ \text{s.c.} & A^T y + s = c \\ & s \in K^*. \end{cases}$$

Lorsque $K \equiv \mathbb{R}_+^n$, nous retrouvons le problème dual classique d'un problème de programmation linéaire.

La proposition suivante caractérise le problème dual associé au dual.

Proposition 2.3.2

Le problème dual associé au problème dual (CD) est équivalent au problème primal (CP).

Preuve :

Tout d'abord, remarquons que le problème dual (CD) peut être écrit sous la forme du problème primal

$$\begin{cases} \min & -b^T y + 0^T s \\ \text{s.c.} & (A^T \ I) \begin{pmatrix} y \\ s \end{pmatrix}^T = c \\ & (y, s) \in \mathbb{R}^m \times K^*. \end{cases}$$

Comme $(K^*)^* = K$ et $(\mathbb{R}^m)^* = \{0\}$, le cône dual de $\mathbb{R}^m \times K^*$ est $\{0\} \times K$ et le dual de (CD) est :

$$\begin{cases} \max & c^T z \\ \text{s.c.} & (-b, 0) - (Az, z) \in \{0\} \times K \end{cases}$$

$$\iff \begin{cases} \max & c^T z \\ \text{s.c.} & Az + b = 0 \\ & -z \in K. \end{cases}$$

En posant $x = -z$, nous avons :

$$\begin{cases} \min & c^T x \\ \text{s.c.} & Ax = b \\ & x \in K, \end{cases}$$

ce qui est exactement le problème primal (CP) . □

Les identités qui suivent constituent une relation très importante entre les points admissibles des problèmes primal et dual.

$$x^T s = x^T (c - A^T y) = c^T x - (Ax)^T y = c^T x - b^T y.$$

Cette valeur est appelée saut de dualité. Comme $x \in K$ et $s \in K^*$, il suit que $x^T s \geq 0$ et par conséquent, $b^T y \leq c^T x$. Cela signifie que la valeur optimale du problème dual est toujours plus petite que celle du problème primal. C'est la propriété de dualité faible.

Pour comprendre comment les problèmes primal et dual sont liés d'un point de vue géométrique, nous fixons \hat{x} satisfaisant $A\hat{x} = b$ et (\hat{y}, \hat{s}) satisfaisant $A^T \hat{y} + \hat{s} = c$. Soit L le noyau de A , i.e. $L = \{x \text{ tq } Ax = 0\}$.

Nous avons

$$c^T x = (A^T \hat{y} + \hat{s})^T x = \hat{y}^T Ax + \hat{s}^T x = \underbrace{\hat{y}^T b}_{\text{constante}} + \hat{s}^T x, \quad \forall x \text{ tq } Ax = b$$

et

$$\begin{aligned} Ax = b & \iff A(x - \hat{x}) = 0 \\ & \iff x - \hat{x} \in L \\ & \iff x \in L + \hat{x}. \end{aligned}$$

En effet, si $Ax = b$, nous avons $A(x - \hat{x}) = Ax - A\hat{x} = b - b = 0$. Inversément, si $A(x - \hat{x}) = 0$, en prenant $x = x - \hat{x} + \hat{x}$, nous obtenons $Ax = A(x - \hat{x}) + A\hat{x} = b$.

Le problème primal (CP) est donc équivalent à une constante près à :

$$\begin{cases} \min & \hat{s}^T x \\ \text{s.c.} & x \in L + \hat{x} \\ & x \in K. \end{cases} \quad (2.1)$$

D'un autre côté, si (y, s) est admissible pour le dual, alors y est déterminé de manière unique par s .

Nous avons

$$b^T y = (A\hat{x})^T y = \hat{x}^T A^T y = \hat{x}^T (c - s) = \underbrace{c^T \hat{x}}_{\text{constante}} - \hat{x}^T s, \quad \forall (y, s) \text{ admissible}$$

et

$$\begin{aligned} A^T y + s = c & \iff (s - \hat{s})^T x = 0 \\ & \iff s - \hat{s} \in L^\perp \\ & \iff s \in L^\perp + \hat{s}. \end{aligned}$$

En effet, si $A^T y + s = c$, nous avons $A^T(y - \hat{y}) + (s - \hat{s}) = 0$ car $A^T \hat{y} + \hat{s} = c$ par hypothèse. Par conséquent, $(s - \hat{s})^T x = (A^T(\hat{y} - y))^T x = (\hat{y} - y)^T Ax = 0$. Inversément, si $(s - \hat{s})^T x = 0$, en prenant $s = s - \hat{s} + \hat{s}$ et $y = y - \hat{y} + \hat{y}$, nous obtenons :

$$A^T y + s = A^T(y - \hat{y}) + A^T \hat{y} + (s - \hat{s}) + \hat{s} = \underbrace{(A^T(y - \hat{y}) + (s - \hat{s}))}_{=0} + A^T \hat{y} + \hat{s} = c.$$

Dès lors, le problème dual (CD) est équivalent, à une constante près à :

$$\begin{cases} \min & \hat{x}^T s \\ \text{s.c.} & x \in L^\perp + \hat{s} \\ & s \in K^*. \end{cases} \quad (2.2)$$

Par les systèmes (2.1) et (2.2), les relations géométriques entre les problèmes primal et dual sont assez visibles. De plus, comme $(L^\perp)^\perp = L$ et $(K^*)^* = K$, on voit clairement que le dual du problème dual est le problème primal, comme nous l'avons démontré dans la proposition 2.3.2.

2.3.3 Dualité forte

Par la propriété de dualité faible, nous savons que si p^* est la valeur optimale du problème primal (CP) et d^* est la valeur optimale du problème dual (CD), alors $d^* \leq p^*$. Nous allons montrer que sous certaines hypothèses, nous avons l'égalité entre ces deux valeurs. C'est la propriété de dualité forte.

En programmation linéaire, la propriété de dualité forte était satisfaite lorsque l'un des deux problèmes, primal ou dual, était admissible. En programmation conique, il faut imposer une hypothèse plus forte pour obtenir l'égalité entre les solutions optimales des deux problèmes. Il faut que l'un des deux problèmes au moins soit strictement admissible.

Définition 2.3.2

Un point x (respectivement (y, s)) est strictement admissible pour le primal (respectivement dual) si $Ax = b$ et $x \in \text{int } K$ (respectivement $A^T y + s = c$ et $s \in \text{int } K^*$). Le problème (CP) (respectivement (CD)) est non borné si $p^* = -\infty$ (respectivement $d^* = +\infty$).

Par la propriété de dualité faible, nous avons immédiatement que si (CP) est non borné, alors $p^* = d^* = -\infty$. De la même façon, si (CD) est non borné, alors $p^* = d^* = +\infty$.

Théorème 2.3.1 (Dualité forte)

Si le problème dual (CD) est borné et strictement admissible, alors le problème primal (CP) est soluble et $p^* = d^*$.

Preuve :

Comme $d^* \leq p^*$, il nous suffit de montrer qu'il existe une solution primale admissible x^* tq $c^T x^* \leq d^*$.

Considérons l'ensemble convexe suivant :

$$M = \{c - A^T y \text{ tq } y \in \mathbb{R}^m, b^T y \geq d^*\}.$$

La preuve se fera en plusieurs parties.

1. Supposons que $b \neq 0$ et montrons que M est non vide et $M \cap \text{int } K^*$ est vide.

Supposons par l'absurde que M est vide. Dans ce cas, $b^T y < d^* \in \mathbb{R}$, $\forall y \in \mathbb{R}^m$. En particulier, $\forall \lambda > 0$, nous avons $b^T(\lambda b) = \lambda \|b\|^2 < d^*$, ce qui est impossible car $\sup_{\lambda > 0} b^T(\lambda b) = \sup_{\lambda > 0} \lambda \|b\|^2 = +\infty$.

A nouveau par l'absurde, supposons que $M \cap \text{int } K^* \neq \emptyset$. Alors $\exists \bar{y} \in \mathbb{R}^m$ tq $b^T \bar{y} \geq d^*$ et tq $c - A^T \bar{y} \in \text{int } K^*$. Par définition de l'intérieur, il existe un petit voisinage de \bar{y} , soit $\mathcal{N}_\epsilon(\bar{y})$, dont les points sont admissibles pour (CD). Comme $b \neq 0$, il y a des points y dans ce voisinage tq $b^T y > b^T \bar{y}$. En effet :

- Si $b^T \bar{y} > 0$, $y = \bar{y} + \frac{\epsilon}{2} \bar{y} = (1 + \frac{\epsilon}{2}) \bar{y}$ convient car

$$b^T y = \left(1 + \frac{\epsilon}{2}\right) b^T \bar{y} > b^T \bar{y}.$$

- Si $b^T \bar{y} < 0$, $y = (1 - \frac{\epsilon}{2}) \bar{y}$ convient car

$$b^T y = \left(1 - \frac{\epsilon}{2}\right) b^T \bar{y} > b^T \bar{y}.$$

- Si $b^T \bar{y} = 0$, $y = \bar{y} + \frac{\epsilon}{2} b$ convient car

$$b^T y = \underbrace{b^T \bar{y}}_{=0} + \frac{\epsilon}{2} \|b\|^2 > b^T \bar{y} = 0.$$

Comme $b^T \bar{y} \geq d^*$, nous obtenons $b^T y > b^T \bar{y} \geq d^*$, ce qui est impossible car d^* est la valeur optimale de (CD).

2. En appliquant le théorème de séparation pour les ensembles convexes à M et $\text{int } K^*$, nous obtenons :

$$\exists x \in \mathbb{R}^n, x \neq 0 \text{ tq } \sup_{z \in M} x^T z \leq \inf_{z \in \text{int } K^*} x^T z. \quad (2.3)$$

3. Montrons que $x \in K$ et $\inf_{z \in \text{int } K^*} x^T z = 0$.

Par (2.3), $x^T z$ est borné inférieurement sur $\text{int } K^*$. Cette borne inférieure est positive. Sinon $\exists \bar{z} \in \text{int } K^*$ tq $x^T \bar{z} < 0$. Mais alors $\lim_{\mu \rightarrow +\infty} x^T(\mu \bar{z}) = -\infty$, ce qui est impossible car $\mu \bar{z} \in \text{int } K^*$, $\forall \mu > 0$ et $x^T(\mu \bar{z})$ doit donc être borné inférieurement. Donc, $x^T z \geq 0$ pour tout z appartenant à la fermeture de $\text{int } K^*$, c'est-à-dire K^* . Dès lors, nous concluons que $x \in K^{**} = K$ et que $\inf_{z \in \text{int } K^*} x^T z = 0$ car

$$0 \leq \inf_{z \in \text{int } K^*} x^T z \leq \inf_{\mu > 0} x^T(\mu \bar{z}) = 0, \text{ où } \bar{z} \in \text{int } K^*.$$

4. Montrons que $\exists \mu \geq 0$ tq $Ax = \mu b$.

Par le point précédent et par (2.3), nous avons $\sup_{z \in M} x^T z \leq 0$, c'est-à-dire, en utilisant la définition de M et en prenant $z = c - A^T y \in M$:

$$\forall y \in \mathbb{R}^m \text{ du demi-espace tq } b^T y \geq d^*, \text{ nous avons } x^T c \leq (Ax)^T y. \quad (2.4)$$

Donc, le problème de programmation linéaire

$$\min\{(Ax)^T y \text{ tq } b^T y \geq d^*\}$$

est borné inférieurement et par conséquent, par le théorème de dualité en programmation linéaire, son dual

$$\begin{cases} \max & d^{*T} \mu \\ \text{s.c.} & b\mu = Ax \\ & \mu \geq 0 \end{cases}$$

est admissible, c'est-à-dire $\exists \mu \geq 0$ tq $Ax = b\mu$.

5. Montrons que $\mu > 0$.

Si $\mu = 0$, alors $Ax = 0$ par le point précédent, et par (2.4), $x^T c \leq 0$. Comme le dual (CD) est strictement admissible, $\exists \bar{y} \in \mathbb{R}^m$ tq $c - A^T \bar{y} \in \text{int } K^*$. De plus, si $x \in K$, $x \neq 0$, nous avons $x^T (c - A^T \bar{y}) > 0$ (sinon le produit est nul par définition de K^*).

D'où $x^T (c - A^T \bar{y}) = x^T c - (Ax)^T \bar{y} = x^T c > 0$, ce qui contredit le fait que $x^T c \leq 0$.

6. Montrons que $x^* = \frac{x}{\mu}$ est admissible pour le primal et $c^T x^* \leq d^*$.

- $x \in K \implies x^* = \frac{x}{\mu} \in K$ car $\mu > 0$.
- $Ax = \mu b \implies Ax^* = b$.

Par conséquent, x^* est admissible.

Par (2.4), $\forall y \in \mathbb{R}^m$ tq $b^T y \geq d^*$, $x^{*T} c \leq (Ax^*)^T y = b^T y$. Par conséquent, $c^T x^* \leq d^*$. Sinon $d^* < c^T x^*$. Nous obtenons alors

$$\forall y \in \mathbb{R}^m \text{ tq } b^T y \geq d^*, d^* < c^T x^* \leq b^T y,$$

ce qui est impossible.

7. Pour terminer la preuve, nous considérons le cas où $b = 0$.

Dans ce cas, la valeur optimale de (CD) est $d^* = 0$ et $x^* = 0$ est admissible pour (CP) et nous avons $c^T x^* \leq d^*$.

□

Ce théorème nous permet de savoir quand les deux problèmes ont la même valeur optimale et quand cette valeur est atteinte par un des deux problèmes.

Nous obtenons le corollaire suivant.

Corollaire 2.3.1

Si les deux problèmes, primal et dual, admettent un point strictement admissible, alors le saut de dualité est nul, la valeur optimale est finie et chaque problème admet une solution.

2.3.4 Exemples

Quand le problème primal n'est pas strictement admissible, il peut arriver que la propriété de dualité forte ne soit pas vérifiée, comme le montre l'exemple qui suit.

Considérons le problème :

$$(CP_1) \begin{cases} \min & -x_1 \\ \text{s.c.} & x_2 + x_3 = 0 \\ & x \in K \end{cases}$$

où $K = L^3 = \{x \in \mathbb{R}^3 \text{ tq } \sqrt{x_1^2 + x_2^2} \leq x_3\}$.

Comme K est auto-dual, le problème dual est :

$$(CD_1) \begin{cases} \max & 0y \\ \text{s.c.} & -(1, y, y) \in K. \end{cases}$$

En écrivant $x_1^2 + x_2^2 \leq x_3^2$ comme $x_1^2 \leq (x_2 + x_3)(x_2 - x_3)$, on voit aisément que $p^* = 0$.

D'un autre côté, pour être admissible pour le dual, y doit satisfaire $1 + y^2 \leq y^2$, ce qui est impossible.

Donc $d^* = -\infty$ et $p^* \neq d^*$, même si une des valeurs optimales est finie.

Dans le second exemple, nous illustrons le fait que même si les deux valeurs optimales sont finies, la propriété de dualité forte n'est pas nécessairement vérifiée.

Considérons le problème suivant :

$$(CP_2) \begin{cases} \min & -x_1 \\ \text{s.c.} & x_1 + x_4 = 1 \\ & x_2 + x_3 = 0 \\ & x \in K \end{cases}$$

où K est le cône $L^3 \times \mathbb{R}_+$.

Comme K est auto-dual, le dual est :

$$(CD_2) \begin{cases} \max & y_1 \\ \text{s.c.} & -(1 + y_1, y_2, y_2, y_1) \in K. \end{cases}$$

Comme pour l'exemple précédent, nous obtenons $p^* = 0$. Pour que y soit admissible pour le dual, il faut que

$$\begin{cases} (1 + y_1)^2 + y_2^2 \leq y_2^2 \\ -y_1 \geq 0. \end{cases}$$

Tous les points admissibles pour le dual satisfont donc $y_1 = -1$, et par conséquent $d^* = -1$.

2.4 La fonction conjuguée

Pour établir des liens intéressants entre le problème de programmation conique (CP) et son dual (CD), nous devons supposer que l'intérieur de K est le domaine d'une fonction barrière.

Afin de comprendre comment une fonction barrière $f : \text{int } K \rightarrow \mathbb{R}$ est en relation avec la théorie de la dualité, nous considérons l'application gradient négative $x \rightsquigarrow -g(x)$. Cette application transforme l'intérieur de K en K^* . En effet, par la proposition 1.3.3, si $x, x' \in \text{int } K$, alors $\langle g(x), x' - x \rangle < \theta_f$. En appliquant cela à tx' au lieu de x' et en faisant tendre $t \rightarrow \infty$, nous obtenons

$$\begin{aligned} \langle g(x), tx' - x \rangle &< \theta_f \\ \iff t\langle g(x), x' \rangle - \langle g(x), x \rangle &< \theta_f \\ \iff \langle g(x), x' \rangle &< \frac{\theta_f + \langle g(x), x \rangle}{t}. \end{aligned}$$

A la limite,

$$\langle g(x), x' \rangle < 0$$

$$\text{i.e. } \langle -g(x), x' \rangle \geq 0 \quad \forall x' \in K$$

$$\text{i.e. } -g(x) \in K^*.$$

De plus, cette application est injective.

En effet, supposons que $g(x) = g(x')$.

Soit l'application $\bar{x} \rightsquigarrow -\langle g(x), \bar{x} \rangle + f(\bar{x})$. Cette application est strictement convexe car f l'est. L'ensemble des minimums de cette application est donné par $\{\bar{x}^* \text{ tq } -g(x) + g(\bar{x}^*) = 0\}$. x est donc un minimum. De la même façon, x' est un minimum. Comme la fonction est strictement convexe, le minimum est unique et par conséquent, $x = x'$.

Nous verrons que l'application $x \rightsquigarrow -g(x)$ n'est pas seulement injective, mais qu'elle est également surjective, et par conséquent qu'elle est une bijection entre l'intérieur de K et l'intérieur de K^* . De plus, nous montrerons que l'application inverse est également la fonction gradient négative d'une fonction barrière, à savoir la fonction conjuguée de f :

$$f^*(s) = - \inf_{x \in \text{dom } f} (\langle x, s \rangle + f(x)).$$

Cette fonction a toujours eu un rôle prépondérant en optimisation, et ce bien avant que la recherche sur les méthodes de points intérieurs ne se développe et que la notion de fonction self-concordante ne soit introduite.

Remarque 2.4.1

Cette définition est un peu différente de la définition standard de la fonction conjuguée : $s \rightsquigarrow f^*(s) = \sup_{x \in \text{dom } f} \langle x, s \rangle - f(x)$. Le but de ce changement est de réduire le nombre de signes moins qui apparaissent et d'avoir le cône dual comme domaine de cette fonction plutôt que le cône polaire.

Par la suite, il sera important de se rappeler que la définition de fonction conjuguée, tout comme celle de cône dual, est basée sur le produit scalaire sous-jacent. Par conséquent, si le produit scalaire change, il en est de même pour la fonction conjuguée.

La définition de la fonction conjuguée f^* s'applique à chaque fonction f , pas seulement aux fonctions barrières. C'est une fonction convexe, car c'est le supremum de fonctions convexes, ou plutôt de fonctions linéaires :

$$s \rightsquigarrow -\langle x, s \rangle - f(x).$$

Le domaine de f^* est défini comme l'ensemble constitué de tous les s pour lesquels $f^*(s)$ est fini, i.e. $\{s \text{ tq } f^*(s) < +\infty\}$.

Dans cette section, nous supposons que f satisfait uniquement les propriétés suivantes, à moins qu'autre chose ne soit spécifié :

- $f \in C^2$.
- $\text{dom } f$ est ouvert et convexe.
- $H(x)$ est définie positive $\forall x \in \text{dom } f$.

Le théorème qui suit donne différentes caractérisations équivalentes des fonctions self-concordantes. Ce résultat ne sera pas démontré mais nous l'utiliserons dans les démonstrations qui suivent.

Théorème 2.4.1

Pour une fonction f , les conditions suivantes sont équivalentes :

(1a) $\forall x, y \in \text{dom } f$, si $\|y - x\|_x < 1$, alors

$$\frac{\|v\|_y}{\|v\|_x} \leq \frac{1}{1 - \|y - x\|_x} \quad \forall v \neq 0.$$

(1b) $\forall x \in \text{dom } f, \forall y \in B(x, \epsilon)$,

$$\frac{\|v\|_y}{\|v\|_x} \leq \frac{1}{1 - \|y - x\|_x} \quad \forall v \neq 0.$$

(1c) $\forall x \in \text{dom } f$,

$$\limsup_{y \rightarrow x} \frac{\|I - H_x(y)\|_x}{\|y - x\|_x} \leq 2.$$

De plus, les conditions suivantes sont aussi équivalentes :

(2a) $\forall x \in \text{dom } f, B_x(x, 1) \subseteq \text{dom } f$.

(2b) $\exists 0 < r \leq 1$ tq $\forall x \in \text{dom } f, B_x(x, r) \subseteq \text{dom } f$.

(2c) Si une suite $\{x_k\}$ converge vers un point de la frontière de $\text{dom } f$, alors $f(x_k) \rightarrow \infty$.

D'où, comme l'ensemble SC est précisément constitué de toutes les fonctions satisfaisant les conditions (1a) et (2a), en choisissant une condition du premier ensemble et une condition du second, on définit l'ensemble SC comme l'ensemble des fonctions satisfaisant les deux conditions choisies.

Afin d'établir le résultat principal de cette section, nous avons besoin de deux propositions et un théorème.

Proposition 2.4.1

L'application gradient $g : \text{dom } f \rightarrow \mathbb{R}^n$ est injective. Si $f \in SC$, alors $\text{dom } f^* = \{-g(x) : x \in \text{dom } f\}$ et $\text{dom } f^*$ est ouvert.

Preuve :

- Si $-g(x_1) = s = -g(x_2)$, alors x_1 et x_2 minimisent la fonction strictement convexe $x \rightsquigarrow \langle x, s \rangle + f(x)$.
En effet, le gradient de cette fonction est $s + g(x)$, qui s'annule en x_1 et x_2 . Comme la fonction est strictement convexe (car f l'est), le minimum est unique. Par conséquent, $x_1 = x_2$ et g est une application injective.

- La fonction $x \rightsquigarrow \langle x, s \rangle + f(x)$ a un minimum, soit $s \in \text{dom } f^* = \{s \text{ tq } \inf_{x \in \text{dom } f} \{\langle x, s \rangle + f(x)\} > -\infty\}$.

D'où

$$\{-g(x) : x \in \text{dom } f\} \subseteq \text{dom } f^*.$$

Pour montrer l'inclusion inverse, supposons que $s \in \text{dom } f^*$ et $f \in SC$. Comme le domaine de f^* est constitué des s tq $f^*(s)$ est fini, la fonction $\langle x, s \rangle + f(x)$ est bornée inférieurement.

En effet $\inf_{x \in \text{dom } f} \langle x, s \rangle + f(x) = -\langle x, g(x) \rangle + f(x) \in \mathbb{R}$ car f est strictement convexe.

Nous pouvons montrer que sous ces conditions ([3], p. 34), la fonction admet un minimum x tq $-g(x) = s$.

Par conséquent,

$$\text{dom } f^* \subseteq \{-g(x) : x \in \text{dom } f\}.$$

- Nous pouvons montrer ([3], p. 34) que le domaine de f^* est ouvert.

□

Proposition 2.4.2

Si $f \in SCB$ et $\text{dom } f = \text{int } K$, alors $\text{dom } f^* = \text{int } K^*$.

Preuve :

Nous avons vu au début de cette section que

$$\{-g(x) : x \in \text{int } K\} \subseteq K^*.$$

Par la proposition qui précède, nous avons alors $\text{dom } f^* \subseteq K^*$ et $\text{dom } f^*$ est ouvert. Dès lors, $\text{dom } f^* \subseteq \text{int } K^*$.

Il nous reste à montrer l'inclusion inverse. Pour cela, grâce à la proposition précédente, il suffit de montrer que si $s \in \text{int } K^*$, alors $\exists x \in \text{int } K$ tq $-g(x) = s$, c'est-à-dire il suffit de montrer que la fonction self-concordante $\bar{f}(x) \equiv \langle x, s \rangle + f(x)$ admet un minimum.

Comme $s \in \text{int } K^*$, $\exists \alpha > 0$ tq $\langle x, s \rangle \geq \alpha \|x\| \quad \forall x \in K$.

Sinon $\forall \alpha > 0 \exists y \in K, \|y\| = 1$ tq $\langle y, s \rangle < \alpha$.

Comme $s \in \text{int } K^*$, $\exists \epsilon > 0$ tq $B(s, \epsilon) \subseteq K^*$.

En particulier, pour $0 < \alpha < \epsilon$, $\exists y \in K, \|y\| = 1$ tq $\langle y, s \rangle < \alpha$,

i.e. $s - \epsilon y \in B(s, \epsilon) \subseteq K^*$

i.e. $\langle s - \epsilon y, y \rangle = \langle s, y \rangle - \epsilon \|y\|^2 < \alpha - \epsilon < 0$,

ce qui est impossible car $y \in K$ et $s - \epsilon y \in K^*$.

Par conséquent, $\exists \alpha > 0$ tq $\langle x, s \rangle \geq \alpha \|x\| \quad \forall x \in K$.

En particulier, si $x \in \text{int } K$, alors

$$\begin{aligned} \langle x, s \rangle + \langle g(x), x \rangle &= \langle x, s \rangle - \langle g(x), 0 - x \rangle \\ &\geq \alpha \|x\| - \theta_f \quad \text{par la proposition 1.3.3.} \end{aligned}$$

Si $\|x\| \geq \frac{\theta_f}{\alpha}$, \bar{f} est strictement croissante à chaque itération le long de $\{x + tx : t \geq 0\}$.

D'où $\inf\{\bar{f}(x) : x \in \text{int } K\} = \inf\{\bar{f}(x) : x \in \text{int } K \text{ et } \|x\| \leq \frac{\theta_f}{\alpha}\}$.

Soit l'ensemble niveau suivant :

$$L \equiv \{x \in \text{int } K \text{ tq } \|x\| \leq \frac{\theta_f}{\alpha}, \bar{f}(x) \leq \bar{f}(\bar{x}) \in \mathbb{R}, \text{ où } \bar{x} \in \text{int } K\}.$$

Si nous montrons que cet ensemble est compact, comme \bar{f} est continue, nous aurons que \bar{f} admet un minimum.

L est borné par définition. Il suffit donc de montrer que L est fermé, i.e. si $x_j \rightarrow x$ et $x_j \in L \quad \forall j$, alors $x \in L$.

- $x_j \in L \Rightarrow \|x_j\| \leq \frac{\theta_f}{\alpha} \quad \forall j$.

Par passage à la limite, nous obtenons $\|x\| \leq \frac{\theta_f}{\alpha}$.

- $x \in \text{int } K$ sinon $x \in K \setminus \text{int } K = \partial K$.
Or, par le théorème 2.4.1, si une suite $\{x_j\}$ converge vers un point de la frontière de K , $\bar{f}(x_j) \rightarrow +\infty$. Ce qui est impossible car $\bar{f}(x_j) \leq \bar{f}(\bar{x}) \in \mathbb{R}$.
- Par hypothèse et par le point qui précède, $x_j \rightarrow x \in \text{int } K$, \bar{f} est continue sur $\text{int } K$ et $\bar{f}(x_j) \leq \bar{f}(\bar{x})$. D'où, par passage à la limite, $\bar{f}(x) \leq \bar{f}(\bar{x})$.

□

Les identités que nous allons établir dans le théorème suivant sont absolument fondamentales dans le développement des méthodes de points intérieurs primales-duales.

Théorème 2.4.2

Supposons que $f \in SC$. Alors $f^* \in C^2$. De plus, si x et s satisfont $s = -g(x)$, alors

$$-g^*(s) = x \text{ et } H^*(s) = H(x)^{-1},$$

où g^* et H^* sont respectivement le gradient et le hessien de f^* .

Preuve :

Par la proposition 2.4.1, l'application $-g : \text{dom } f \rightarrow \text{dom } f^*$ est inversible. $\forall s \in \text{dom } f^*$, soit $x(s)$ le point de $\text{dom } f$ tq $-g(x(s)) = s$.

Comme g est de classe C^1 et que la dérivée première de g est inversible (c'est-à-dire le hessien $H(x)$ est inversible), le théorème des fonctions inverses implique que $s \rightsquigarrow x(s)$ est une application de classe C^1 .

En dérivant les deux membres de $-g(x(s)) = s$ par rapport à s , nous obtenons

$$\begin{aligned} -H(x(s))D_s x(s) &= I, \\ \text{i.e. } D_s x(s) &= -H(x(s))^{-1}, \end{aligned} \tag{2.5}$$

où $D_s x(s)$ représente la dérivée de $x(s)$ par rapport à s .

Soit $[D_s x(s)]^T$ l'adjointe de $D_s x(s) : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Nous avons $f^*(s) = -\langle x(s), s \rangle - f(x(s))$.

Comme l'application $s \rightsquigarrow x(s)$ est de classe C^1 , il en est de même pour la fonction f^* . En dérivant les deux membres par rapport à s , nous avons

$$\begin{aligned} g^*(s) &= -x(s) - [D_s x(s)]^T s - [D_s x(s)]^T \underbrace{g(x(s))}_{=-s} \\ &= -x(s). \end{aligned} \tag{2.6}$$

Donc, lorsque x et s satisfont $-g(x) = s$, nous avons $-g^*(s) = x$.
De plus, comme l'application $s \rightsquigarrow x(s)$ est de classe C^1 , par (2.6), $f^* \in C^2$.

Finalement, en combinant (2.5) et (2.6), nous obtenons

$$H^*(s) = -D_s x(s) = H(x(s))^{-1}$$

c'est-à-dire, si $-g(x) = s$,

$$H^*(s) = H(x)^{-1}.$$

□

Nous obtenons alors le résultat suivant.

Théorème 2.4.3

Si $f \in SC$, alors $f^* \in SC$. Si $f \in SCB$ et $\text{dom } f = \text{int } K$, alors $f^* \in SCB$ et $\text{dom } f^* = \{-g(x) : x \in \text{int } K\} = \text{int } K^*$.

Preuve :

Supposons que $f \in SC$. Supposons aussi que x et s satisfont $-g(x) = s$. Par le théorème précédent, $\forall w_1, w_2$,

$$\begin{aligned} \langle w_1, w_2 \rangle_s^* &= \langle w_1, H^*(s)w_2 \rangle \\ &= \langle w_1, H(x)^{-1}w_2 \rangle \\ &= \langle H(x)^{-1}w_1, H(x)^{-1}w_2 \rangle_x. \end{aligned}$$

Cette identité sera utilisée plus tard dans la preuve. Nous avons également la relation

$$\|w\|_s^* = \|H(x)^{-1}w\|_x.$$

- Montrons que $f^* \in SC$. Pour cela, il suffit de montrer, par le théorème 2.4.1, que si $s \in \text{dom } f^*$, alors

$$B_s^*(s, \frac{1}{4}) \subseteq \text{dom } f^* \tag{2.7}$$

et

$$\limsup_{\Delta s \rightarrow 0} \frac{\|I - H_s^*(s + \Delta s)\|_s^*}{\|\Delta s\|_s^*} \leq 2. \tag{2.8}$$

Supposons que Δs est un vecteur satisfaisant $r = \|\Delta s\|_s^* < \frac{1}{4}$, c'est-à-dire $\|v\|_x < \frac{1}{4}$, où $v \equiv -H(x)^{-1}\Delta s$. Nous avons $\frac{3r^2}{(1-r)^3} \leq 9r^2$.

Sous ces conditions, nous pouvons montrer que ([3], p. 34)

$$\exists u \in B_x(v, 9\|v\|_x^2) \text{ tq } g_x(x+u) = g_x(x) + v.$$

Or,

$$\begin{aligned} -g(x+u) &= -H(x)g_x(x+u) \\ &= -H(x)(g_x(x) + v) \\ &= -g(x) - H(x)v \\ &= s + \Delta s. \end{aligned}$$

En conséquence, par la proposition 2.4.1, nous avons que $s + \Delta s \in \text{dom } f^*$, ce qui établit (2.7).

De plus, nous avons

$$\|u\|_x \leq \|v\|_x + 9\|v\|_x^2 = \|\Delta s\|_s^* + 9(\|\Delta s\|_s^*)^2. \quad (2.9)$$

Afin d'établir (2.8), observons que

$$\begin{aligned} \|I - H_s^*(s + \Delta s)\|_s^* &= \max_z \frac{|\langle z, [I - H_s^*(s + \Delta s)]z \rangle_s^*|}{(\|z\|_s^*)^2} \\ &= \max_w \frac{|\langle H(x)w, [I - H_s^*(s + \Delta s)]H(x)w \rangle_s^*|}{(\|H(x)w\|_s^*)^2} \\ &\quad \text{en posant } z = H(x)w \\ &= \max_w \frac{|\langle w, H(x)^{-1}[I - H_s^*(s + \Delta s)]H(x)w \rangle_x|}{(\|w\|_x)^2} \\ &= \max_w \frac{|\langle w, [I - H(x)^{-1}H_s^*(s + \Delta s)]w \rangle_x|}{(\|w\|_x)^2} \\ &= \|I - H(x)^{-1}H_s^*(s + \Delta s)H(x)\|_x. \end{aligned} \quad (2.10)$$

Or,

$$H_s^*(s + \Delta s) = H^*(s)^{-1}H^*(s + \Delta s) = H(x)H(x+u)^{-1},$$

ce qui implique

$$H(x)^{-1}H_s^*(s + \Delta s)H(x) = H(x+u)^{-1}H(x) = H_x(x+u)^{-1}. \quad (2.11)$$

Nous obtenons alors

$$\begin{aligned} \|I - H_s^*(s + \Delta s)\|_s^* &= \|I - H_x(x+u)^{-1}\|_x \text{ par (2.10) et (2.11)} \\ &\leq \frac{1}{(1 - \|u\|_x)^2} - 1 \text{ par le lemme 1.3.1} \\ &\leq \frac{1}{(1 - \|\Delta s\|_s^* - 9(\|\Delta s\|_s^*)^2)^2} - 1 \text{ par (2.9)}. \end{aligned}$$

Dès lors,

$$\begin{aligned}
& \limsup_{\Delta s \rightarrow 0} \frac{\|I - H_s^*(s + \Delta s)\|_s^*}{\|\Delta s\|_s^*} \\
& \leq \limsup_{\Delta s \rightarrow 0} \frac{1}{\|\Delta s\|_s^*} \left(\frac{1}{(1 - \|\Delta s\|_s^* - 9\|\Delta s\|_s^{*2})^2} - 1 \right) \\
& = \limsup_{\Delta s \rightarrow 0} \frac{1}{\|\Delta s\|_s^*} \frac{(1 - (1 - \|\Delta s\|_s^* - 9\|\Delta s\|_s^{*2})) (1 + (1 - \|\Delta s\|_s^* - 9\|\Delta s\|_s^{*2}))}{(1 - \|\Delta s\|_s^* - 9\|\Delta s\|_s^{*2})^2} \\
& = \limsup_{\Delta s \rightarrow 0} (1 + 9\|\Delta s\|_s^*) \frac{(1 + (1 - \|\Delta s\|_s^* - 9\|\Delta s\|_s^{*2}))}{(1 - \|\Delta s\|_s^* - 9\|\Delta s\|_s^{*2})^2} \\
& \leq 2, \text{ ce qui est exactement (2.8).}
\end{aligned}$$

Nous avons donc montré que $f^* \in SC$.

- Supposons que $f \in SCB$ et $\text{dom } f = \text{int } K$. Par la proposition 2.4.2, $\text{dom } f^* = \text{int } K^*$. Il reste à montrer que $f^* \in SCB$.

Observons tout d'abord que

$$\|g_s^*(s)\|_s^* = \|H^*(s)^{-1}g^*(s)\|_s^* = \|H(x)x\|_s^* = \|x\|_x. \quad (2.12)$$

Nous avons $\langle g(x), 0 - x \rangle \geq 0$ car $-g(x) \in K^*$. Comme $0 \in K$, nous pouvons montrer que ([3], p. 38) $\|x\|_x \leq 4\theta_f + 1$.

Par conséquent, en utilisant (2.12),

$$\theta_{f^*} \equiv \sup_{s \in \text{dom } f^*} \|g_s^*(s)\|_s^{*2} = \sup_{s \in \text{dom } f^*} \|x\|_x^2 \leq (4\theta_f + 1)^2.$$

□

2.5 Les cônes "self-scaled" (ou symétriques)

Un cône self-scaled (ou symétrique) est un cône K dont l'intérieur est le domaine d'une fonction barrière f ayant certaines propriétés. Afin d'établir ces propriétés, nous allons introduire certaines notations ainsi que quelques définitions.

2.5.1 Définitions et notations

Définition 2.5.1

Pour $z \in \text{int } K$, soit K_z^* le cône dual de K par rapport au produit scalaire intrinsèque $\langle \cdot, \cdot \rangle_z$,

i.e.

$$K_z^* = \{s \text{ tq } \langle x, s \rangle_z \geq 0, \quad \forall x \in K\}.$$

K est intrinsèquement auto-dual si $K_z^* = K, \quad \forall z \in \text{int } K$.

Soit f_z^* la fonction conjuguée de f par rapport au produit scalaire $\langle \cdot, \cdot \rangle_z$, i.e.

$$f_z^*(s) \equiv - \inf_{x \in \text{int } K} (\langle x, s \rangle_z + f(x)).$$

Si K est intrinsèquement auto-dual, alors, par le théorème 2.4.3, f et f_z^* ont l'intérieur de K comme domaine.

Définition 2.5.2

Une fonction barrière $f : \text{int } K \rightarrow \mathbb{R}$ est logarithmiquement homogène si $\forall x \in \text{int } K$ et $\forall t > 0$, $f(tx) = f(x) - \theta_f \ln t$.

Le théorème qui suit fournit une caractérisation importante de l'homogénéité logarithmique. Il ne sera pas démontré mais les résultats seront utilisés par la suite.

Théorème 2.5.1

Une fonction self-concordante $f : \text{int } K \rightarrow \mathbb{R}$ est une fonction barrière logarithmiquement homogène si et seulement si $\forall x \in \text{int } K$ et $\forall t > 0$, $g(tx) = \frac{1}{t}g(x)$.

De plus, une telle fonction satisfait les identités suivantes :

$$H(tx) = \frac{1}{t^2}H(x), \quad g_x(x) = -x \quad \text{et} \quad \|g_x(x)\|_x = \sqrt{\theta_f}.$$

Nous pouvons alors définir la notion de fonction intrinsèquement auto-conjuguée.

Définition 2.5.3

f est intrinsèquement auto-conjuguée si f est logarithmiquement homogène, si K est intrinsèquement auto-dual et si $\forall z \in \text{int } K$,

$$\exists \text{ une constante } C_z \text{ tq } f_z^* \equiv f + C_z,$$

i.e.,

$$\forall s \in \text{int } K, \quad f_z^*(s) = f(s) + C_z.$$

L'intérêt des fonctions intrinsèquement auto-conjuguées est que pour chaque produit scalaire $\langle \cdot, \cdot \rangle_z$, les gradient g_z et hessien H_z de f pour le primal sont identiques aux gradient g_z^* et hessien H_z^* de f_z^* pour le dual.

Définition 2.5.4

Un cône K est self-scaled (ou symétrique) si l'intérieur de K est le domaine d'une fonction barrière intrinsèquement auto-conjuguée.

Dans la définition de fonction intrinsèquement auto-conjuguée, il n'y a pas de restriction sur le choix de la valeur de la constante C_z . Mais sous certaines conditions, la constante prendra une valeur spécifique, comme le montre la proposition suivante.

Proposition 2.5.1

Si $f : \text{int } K \rightarrow \mathbb{R}$ est une fonction barrière intrinsèquement auto-conjuguée, alors $\forall z \in \text{int } K, \quad C_z = -(\theta_f + 2f(z))$.

Preuve :

Le théorème 2.4.2 montre que sans se soucier du produit scalaire utilisé, la fonction conjuguée satisfait

$$\begin{aligned} f^*(s) &= -\langle x, s \rangle - f(x) \\ &= \langle g^*(s), s \rangle - f(-g^*(s)). \end{aligned}$$

Comme f est intrinsèquement auto-conjuguée, nous avons

$$\begin{aligned} f(x) &= f_z^*(x) - C_z \\ &= \langle g_z^*(x), x \rangle_z - f(-g_z^*(x)) - C_z \\ &= \langle g_z(x), x \rangle_z - f(-g_z(x)) - C_z. \end{aligned}$$

En particulier, pour $x = z$,

$$\begin{aligned} f(z) &= \langle g_z(z), z \rangle_z - f(-g_z(z)) - C_z \\ &= -\theta_f - f(z) - C_z \quad \text{par le théorème 2.5.1.} \end{aligned}$$

D'où

$$C_z = -(\theta_f + 2f(z)).$$

□

Le théorème suivant fournit une caractérisation utile des fonctions intrinsèquement auto-conjuguées.

Théorème 2.5.2

Supposons que K est un cône et soit f une fonction barrière logarithmiquement homogène dont le domaine est l'intérieur de K . Alors f est intrinsèquement auto-conjuguée si et seulement si $\forall z \in \text{int } K$, l'application $x \rightsquigarrow -g_z(x)$ est une involution, i.e. le domaine de l'application est l'intérieur de K et $-g_z(-g_z(x)) = x \quad \forall x \in \text{int } K$.

Preuve :

Si f est intrinsèquement auto-conjuguée, nous avons

$$g_z^* = g_z \quad \text{et} \quad H_z^* = H_z.$$

Dès lors, le théorème 2.4.2 implique que l'application $x \rightsquigarrow -g_z(x)$ est une involution.

Inversément, supposons que $\forall z \in \text{int } K, -g_z(-g_z(x)) = x \quad \forall x \in \text{int } K$. Comme, par le théorème 2.4.2, nous avons $-g_z^*(-g_z(x)) = x \quad \forall x \in \text{int } K$, il suit que $g_z = g_z^*$. Dans ce cas, f diffère de f_z^* par au plus une constante, i.e. f est intrinsèquement auto-conjuguée.

□

Ce théorème fournit un moyen assez simple de vérifier que la fonction barrière logarithmique pour l'orthant non négatif est intrinsèquement auto-conjuguée.

Soit e le vecteur composé de 1. Le produit scalaire associé, $\langle \cdot, \cdot \rangle_e$ est le produit scalaire habituel et la fonction barrière est la fonction barrière logarithmique. Nous avons

$$g_z(x) = H_e(z)^{-1} g_e(x) = - \left(\frac{z_1^2}{x_1}, \dots, \frac{z_n^2}{x_n} \right)^T.$$

En conséquence, $-g_z(-g_z(x)) = x$.

La famille de tous les cônes self-scaled est limitée et est identique à celle des cônes symétriques. Il existe cinq cônes symétriques de base, tous les autres étant des produits cartésiens de ceux-ci :

- le cône des matrices semi-définies positives.
- le cône du seconde ordre.
- le cône des matrices hermitiennes semi-définies positives.
- le cône des matrices quaternion hermitiennes semi-définies positives.
- un cône exceptionnel à 27 dimensions .

Pour chacun de ces cônes, on peut développer une théorie des méthodes primales-duales.

Par la suite, afin d'éviter les notations excessives, nous fixons $e \in \text{int } K$ et nous posons $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_e$. Le choix de e est arbitraire, ce qui signifie que n'importe quel point de l'intérieur de K convient. Jusqu'à la fin de cette section, lorsque K est supposé self-scaled, nous noterons f une fonction barrière intrinsèquement auto-conjuguée dont le domaine est l'intérieur de K . De plus, lorsque nous ferons référence à l'orthant non négatif, au cône du second ordre ou au cône des matrices semi-définies positives, nous prendrons comme fonction barrière intrinsèquement auto-conjuguée la fonction barrière logarithmique.

2.5.2 Les points scaling

Supposons que K est self-scaled et soit $z \in \text{int } K$. En nous rappelant que $\langle u, v \rangle_z = \langle u, H(z)v \rangle$, où $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_e$ avec e un point arbitraire appartenant à l'intérieur de K , nous avons

$$K = K_z^* = H(z)^{-1} K^* = H(z)^{-1} K.$$

La dernière égalité vient du fait que $K_z^* = K \ \forall z \in \text{int } K$. En particulier, pour $z = e$, $K_e^* \equiv K^* = K$.

En conséquence, $H(z)$ est un opérateur linéaire bijectif de K dans lui-même, i.e. $H(z)$ est un automorphisme linéaire de K , un scaling de K .

Définition 2.5.5

w est un point scaling par rapport au produit scalaire $\langle \cdot, \cdot \rangle_e$ pour la paire ordonnée x, s si $H(w)x = s$.

Remarquons que si w est le point scaling pour la paire ordonnée x, s , alors $-g(w)$ est le point scaling pour la paire s, x .

En effet,

$$\begin{aligned} H(-g(w))s &= H^*(-g(w))s \text{ car } H = H^* \\ &= H(w)^{-1}s \text{ par le théorème 2.4.2} \\ &= x. \end{aligned}$$

Si K est l'orthant non négatif et $e = (1, \dots, 1)$, nous avons $f(w) = -\sum_{i=1}^n \ln w_i$.

Dès lors, $g(w) = -\left(\frac{1}{w_1}, \dots, \frac{1}{w_n}\right)^T$

$$\text{et } H(w) = \begin{pmatrix} \frac{1}{w_1^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{w_n^2} \end{pmatrix}$$

et par conséquent, w est le vecteur dont la $j^{\text{ème}}$ composante est $\sqrt{\frac{x_j}{s_j}}$.

Afin de comprendre le rôle joué par le point scaling dans la théorie des méthodes primales-duales, il faut se rappeler la dépendance des instances primales et duales par rapport au produit scalaire utilisé.

Si pour le produit scalaire initial $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_e$, le problème primal et son dual sont donnés par

$$\begin{cases} \min & c^T x \\ \text{s.c.} & Ax = b \\ & x \in K \end{cases}$$

et

$$\begin{cases} \max & b^T y \\ \text{s.c.} & A^T y + s = c \\ & s \in K(\equiv K^*), \end{cases}$$

alors, en termes du produit scalaire local $\langle \cdot, \cdot \rangle_w$, le problème primal et son dual s'écrivent :

$$\begin{cases} \min & \langle c_w, x \rangle_w \\ \text{s.c.} & Ax = b \\ & x \in K \end{cases}$$

et

$$\begin{cases} \max & b^T y \\ \text{s.c.} & A_w^T y + s_w = c_w \\ & s_w \in K (\equiv K^*), \end{cases}$$

où $\begin{cases} c_w \equiv H(w)^{-1}c, \\ A_w^T \equiv H(w)^{-1}A^T. \end{cases}$

En conséquence, si nous supposons que w est le point scaling d'une paire x, s , en réécrivant les instances en termes du produit scalaire $\langle \cdot, \cdot \rangle_w$, la paire est transformée en la paire $x, s_w \equiv H(w)^{-1}s$ qui satisfait $x = s_w$. Cela signifie que chaque point de la paire est transformé en le point d'intersection des régions admissibles des problèmes primal et dual réécrits.

Les méthodes primales-duales que nous considérons sont invariantes sous la transformation considérée, dans le sens où les itérés qu'elles génèrent pour les instances réécrites sont les transformations des itérés générés pour les instances originales. Pour les nouvelles instances, les directions de recherche primale et duale sont obtenues en projetant le gradient par rapport au produit scalaire $\langle \cdot, \cdot \rangle_w$ d'une certaine fonction sur le noyau de A et sur l'image de A_w^T respectivement. Comme ces espaces sont des compléments orthogonaux par rapport au produit scalaire local, la somme des directions de recherche est le gradient lui-même.

Proposition 2.5.2

Supposons que f est une fonction barrière dont le domaine est l'intérieur d'un cône K (non nécessairement self-scaled). Soit K^ le cône dual de K par rapport à un produit scalaire arbitraire $\langle \cdot, \cdot \rangle$. Si $x \in \text{int } K$ et $s \in \text{int } K^*$, alors $\exists w \in \text{int } K$ tq $H(w)x = s$.*

Preuve :

Considérons la fonction $\psi(w) \equiv \langle x, -g(w) \rangle + \langle s, w \rangle$ dont le domaine est l'intérieur du cône K .

Pour avoir la thèse, il suffit de montrer que ψ admet un minimum.

En effet, si ψ admet un minimum, le gradient de la fonction en ce point sera nul, i.e. $-H(w)x + s = 0$.

Pour montrer que ψ admet un minimum, il suffit de montrer que si $\{w_i\} \subset \text{int } K$ est une suite tq, soit $\|w_i\| \rightarrow \infty$, soit $w_i \rightarrow \bar{w} \in \partial K$, alors $\psi(w_i) \rightarrow \infty$. En effet, ceci revient à dire que les ensembles niveaux sont compacts. De plus, la fonction ψ est continue et nous savons que toute fonction continue sur un compact admet un minimum.

Supposons tout d'abord que $\|w_i\| \rightarrow \infty$. Comme $s \in \text{int } K^*$ et $\{w_i\} \subset K$, il suit que $\langle s, w_i \rangle \rightarrow \infty$. De plus, $\langle x, -g(w_i) \rangle \geq 0$ car $-g(w_i) \in K^*$. Par conséquent, $\psi(w_i) \rightarrow \infty$.

Supposons maintenant que $w_i \rightarrow \bar{w} \in \partial K$. Alors, nous pouvons montrer que ([3], pg 33) $\|g(w_i)\| \rightarrow \infty$. Comme $-g(w_i) \in K^*$ et $x \in \text{int } K$, nous avons donc $\langle x, -g(w_i) \rangle \rightarrow \infty$. De plus, $\langle s, w_i \rangle \geq 0$ car $s \in K^*$ et $w_i \in K$. Nous concluons que $\psi(w_i) \rightarrow \infty$.

□

En appliquant cette proposition à un cône K self-scaled, nous avons pour chaque paire ordonnée $x, s \in \text{int } K$ l'existence d'un point scaling w . Afin de démontrer l'unicité de ce point scaling, nous nous appuyons sur le théorème suivant.

Théorème 2.5.3

Supposons que K est self-scaled. Si $x, w \in K$, alors

$$f(H(w)x) = f(x) + 2(f(e) - f(w)) \quad (2.13)$$

$$g(H(w)x) = H(x)^{-1}g(x) \quad (2.14)$$

$$H(H(w)x) = H(w)^{-1}H(x)H(w)^{-1} \quad (2.15)$$

Preuve :

Posons $C \equiv C_e$ où la valeur de C_e est donné par la proposition 2.5.1, i.e. $C = -(\theta_f + 2f(e))$.

Nous avons

$$\begin{aligned}
 f(H(w)x) &= f^*(H(w)x) - C \\
 &= -\inf_{v \in K} (\langle v, H(w)x \rangle + f(v)) - C \\
 &= -\inf_{v \in K} (\langle v, x \rangle_w + f(v)) - C \\
 &= f_w^*(x) - C \\
 &= f(x) + C_w - C \\
 &= f(x) - (\theta_f + 2f(w)) + (\theta_f + 2f(e)) \\
 &= f(x) + 2(f(e) - f(w)),
 \end{aligned}$$

ce qui est exactement (2.13).

Pour obtenir (2.14), il suffit de dériver les deux membres de (2.13) par rapport à x .

En effet, $H(w)g(H(w)x) = g(x)$ et donc $g(H(w)x) = H(w)^{-1}g(x)$.

De la même façon, en dérivant les deux membres de (2.14) par rapport à x , nous obtenons (2.15).

En effet, $H(H(w)x)H(w) = H(w)^{-1}H(x)$ et par conséquent, $H(H(w)x) = H(w)^{-1}H(x)H(w)^{-1}$.

□

Le théorème qui suit montre que l'ensemble des automorphismes linéaires $\{H(z) : z \in \text{int } K\}$ forme un ensemble complet de scaling, dans le sens où $\forall x \in \text{int } K, \{H(z)x : z \in \text{int } K\} = \text{int } K$, i.e. pour tout point de l'intérieur de K , il existe un automorphisme unique $H(z)$ qui associe x à ce point.

Théorème 2.5.4

Si K est self-scaled, alors $H(z)$ est un automorphisme linéaire de K , $\forall z \in \text{int } K$. De plus, si $x, s \in \text{int } K$, alors $\exists! w \in \text{int } K$ tq $H(w)x = s$.

Preuve :

Grâce à la proposition 2.5.2, il nous reste seulement à montrer l'unicité du point scaling pour la paire ordonnée x, s .

Supposons que $H(w_1)x = s$ et $H(w_2)x = s$.

Par (2.15),

$$\begin{aligned} H(s)H(x) &= H(H(w_1)x)H(x) = H(H(w_2)x)H(x) \\ &= (H(w_1)^{-1}H(x))^2 = (H(w_2)^{-1}H(x))^2. \end{aligned}$$

En prenant les matrices inverses, nous avons

$$\begin{aligned} (H(x)^{-1}H(w_1))^2 &= (H(x)^{-1}H(w_2))^2 \\ \iff H_x(w_1)^2 &= H_x(w_2)^2. \end{aligned}$$

Comme la racine carrée d'un opérateur linéaire défini positif est unique, il suit que $H_x(w_1) = H_x(w_2)$.

D'où $H_{w_1}(w_2) \equiv H_x(w_1)^{-1}H_x(w_2) = I$.

Donc $g_{w_2} \equiv H_{w_1}(w_2)^{-1}g_{w_1} = g_{w_1}$.

En particulier, si f est logarithmiquement homogène,

$$\begin{aligned} w_2 &= -g_{w_2}(w_2) = -g_{w_1}(w_2) \\ \text{et } w_1 &= -g_{w_1}(w_1). \end{aligned}$$

Si nous considérons la fonction $w \rightsquigarrow \frac{1}{2}\|w\|_{w_1}^2 + f(w)$, le gradient selon $\langle \cdot, \cdot \rangle_{w_1}$ en w est $w + g_{w_1}(w)$. Comme $w_1 = -g_{w_1}(w_1)$ et $w_2 = -g_{w_1}(w_2)$, le gradient est nul en w_1 et w_2 . Comme la fonction est strictement convexe, elle admet au plus un minimum et par conséquent, $w_1 = w_2$. □

2.5.3 Gradients et normes

Nous continuons à fixer un point arbitraire $e \in \text{int } K$ et $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_e$.

Le théorème qui suit concerne l'erreur de l'application linéaire $x \rightsquigarrow x - e$ faite en approximant l'application gradient $x \rightsquigarrow g(x) - g(e)$. En terme de la norme $\|\cdot\| = \|\cdot\|_e$, pour les fonctions self-concordantes, l'erreur est petite lorsque x est proche de e . Ce théorème donne un aperçu de l'erreur pour tout x , pas simplement pour x proche de e .

Théorème 2.5.5

Supposons que K est self-scaled. Si $x \in \text{int } K$, alors

$$(x - e) - (g(x) - g(e)) \in K,$$

i.e.

$$-g(x) \in e - (x - e) + K.$$

Preuve :

Comme $K = K^*$, il suffit de montrer que $\forall v \in \text{int } K$, $\langle (x - e) - (g(x) - g(e)), v \rangle \geq 0$, ce qui est le même que $\psi(x) \geq \psi(e)$, où $\psi(x) \equiv \langle x - g(x), v \rangle$. Cela revient encore à montrer que e minimise la fonction ψ .

Remarquons tout d'abord que ψ admet un minimum. En effet, ψ est continue. De plus, si $\{x_i\} \subset \text{int } K$ est tq, soit $\|x_i\| \rightarrow \infty$, soit $x_i \rightarrow \bar{x} \in \partial K$, alors $\psi(x_i) \rightarrow \infty$, ce qui montre que les ensemble niveaux sont fermés et bornés.

Soit x' le minimum de ψ . En ce point, le gradient de la fonction est nul, i.e. $v - H(x')v = 0$, ou encore $H(x')v = v$.

Comme $H(e)v = Iv = v$ et que le minimum est unique par le théorème 2.5.4, nous obtenons $x' = e$.

□

Par définition, les normes locales induites par une fonction self-concordante f changent petit à petit. Si $\|x - e\| < 1$, alors $\forall v$,

$$1 - \|x - e\| \leq \frac{\|v\|_x}{\|v\|} \leq \frac{1}{1 - \|x - e\|} .$$

De meilleures bornes peuvent être trouvées si f est intrinsèquement auto-conjuguée, comme nous allons le montrer.

Soit \mathcal{B} le plus grand ensemble symétrique centré en 0 et qui satisfait $e + \mathcal{B} \subset \text{int } K$, i.e.

$$\mathcal{B} = \{v : e \pm v \in \text{int } K\}.$$

Nous définissons la norme $|\cdot|$ sur \mathbb{R}^n par

$$|v| \equiv \inf\{t \geq 0 : \frac{1}{t}v \in \mathcal{B}\}.$$

Cette norme est appelée la jauge de Minkowski de \mathcal{B} .

Comme $B(e, 1) \equiv \{x : \|x - e\| < 1\}$ est un ensemble centré de manière symétrique autour de e et qui satisfait $B(e, 1) \subseteq \text{int } K$, nous avons $B(e, 1) \subseteq e + \mathcal{B}$, ce qui peut également s'écrire $|v| \leq \|v\| \quad \forall v$.

Nous avons le résultat suivant.

Théorème 2.5.6

Supposons que K est self-scaled.

Si x satisfait $x - e \in K$, alors $\|v\|_x \leq \|v\| \quad \forall v$.

Preuve :

Nous établissons d'abord l'inégalité moins restrictive

$$\|v\|_x \leq (4\theta_f + 1)\|v\|. \quad (2.16)$$

Observons que $x - e \in K$ et $B(e, 1) \subseteq K$ (par la self-concordance) impliquent que $B(x, 1) \subseteq K$.

Comme $\forall v$, nous avons soit $\langle g(x), v \rangle \geq 0$, soit $\langle g(x), -v \rangle \geq 0$, nous pouvons montrer que ([3], pg 38) $B(x, 1) \subseteq B_x(x, 4\theta_f + 1)$, ce qui donne (2.16).

Comme $\|H(x)\|^{1/2} = \sup_v \frac{\|v\|_x}{\|v\|}$, l'inégalité à prouver est équivalente à affirmer que $\|H(x)\| \leq 1$.

Soit $x_0 \equiv x$ et $x_1 \equiv H(x_0)^{-1}e$. Comme $g(e) = -e$ par l'homogénéité logarithmique, nous avons

$$x_1 = H(x_0)^{-1}e = -H(x_0)^{-1}g(e) = -g_{x_0}(e).$$

Donc, en appliquant le théorème précédent avec x_0 au lieu de e et e au lieu de x , nous obtenons

$$\begin{aligned} e - x_0 - (g_{x_0}(e) - g_{x_0}(x_0)) &\in K \\ \iff e + x_1 - 2x_0 &\in K. \end{aligned}$$

Comme $x_0 - e \in K$, il suit que $x_1 - x_0 \in K$ et par conséquent, $x_1 - e \in K$. De plus, en utilisant (2.15), nous avons

$$H(x_1) = H(H(x_0)^{-1}e) = H(x_0)H(e)H(x_0) = H(x_0)^2.$$

En raisonnant par récurrence, nous trouvons $x_k \equiv H(x_{k-1})^{-1}e$ qui satisfait $x_k - e \in K$ et $H(x_k) = H(x)^{2^k}$.

Donc, $\|H(x_k)\| = \|H(x)\|^{2^k}$.

En appliquant (2.16) avec x_k au lieu de x , nous obtenons

$$\|v\|_{x_k} \leq (4\theta_f + 1)\|v\|,$$

ce qui entraîne

$$\|H(x_k)\| = \sup \frac{\|v\|_{x_k}^2}{\|v\|^2} \leq (4\theta_f + 1)^2.$$

Par conséquent, $\|H(x)\| \leq 1$. Sinon $\|H(x)\| > 1$, ce qui entraîne $\|H(x_k)\| = \|H(x)\|^{2^k} \rightarrow +\infty$ lorsque k tend vers l'infini, ce qui est impossible car $H(x_k)$ est borné.

□

Nous avons alors

Théorème 2.5.7

Supposons que K est self-scaled. Si x satisfait $|x - e| < 1$, alors $\forall v$,

$$\frac{1}{1 + |x - e|} \leq \frac{\|v\|_x}{\|v\|} \leq \frac{1}{1 - |x - e|} \quad (2.17)$$

$$1 - |x - e| \leq \frac{\|v\|_{-g(x)}}{\|v\|} \leq 1 + |x - e| \quad (2.18)$$

Preuve :

Soit $x_1 = \frac{1}{1 + |x - e|} x$ et $x_2 = e - \frac{1}{|x - e|}(x - e)$.

Comme $|x_2 - e| = 1$, nous avons $x_2 \in K$.

Observons que

$$\begin{aligned} x_1 + \frac{|x - e|}{1 + |x - e|} x_2 &= \frac{x}{1 + |x - e|} + \frac{|x - e|}{1 + |x - e|} \left(e - \frac{x - e}{|x - e|} \right) \\ &= \frac{x + e|x - e| - (x - e)}{1 + |x - e|} \\ &= \frac{e(1 + |x - e|)}{1 + |x - e|} \\ &= e. \end{aligned}$$

D'où $e - x_1 \in K$. En conséquence, en appliquant le théorème précédent avec x_1 au lieu de e et e au lieu de x , $\|v\| \leq \|v\|_{x_1} \quad \forall v$.

De plus, par l'homogénéité logarithmique de f , le hessien est homogène de degré -2, i.e. $H(tx) = \frac{1}{t^2}H(x)$, ce qui nous permet d'écrire :

$$\begin{aligned} \|v\|_{x_1}^2 &= v^T H(x_1)v \\ &= v^T H\left(\frac{1}{1 + |x - e|}x\right)v \\ &= (1 + |x - e|)^2 v^T H(x)v \\ &= (1 + |x - e|)^2 \|v\|_x^2. \end{aligned}$$

Par conséquent, $\|v\| \leq (1 + |x - e|)\|v\|_x \quad \forall v$, ce qui donne la première partie de (2.17).

De manière similaire, pour établir la seconde partie de (2.17), nous définissons

$$x'_1 = (1 - |x - e|)e \quad \text{et} \quad x'_2 = e + \frac{1}{x - e}(x - e).$$

Nous remarquons que $x'_2 \in K$ et $x = x'_1 + |x - e|x'_2$. D'où $x - x'_1 \in K$.

Le théorème précédent implique donc $\forall v$, $\|v\|_x \leq \|v\|_{x_1}$ et l'homogénéité logarithmique donne

$$\|v\|_x \leq \frac{1}{1 - |x - e|}\|v\| \quad \forall v,$$

ce qui établit l'inégalité voulue.

Pour établir (2.18), il suffit d'utiliser (2.17) ainsi que la relation $H(-g(x)) = H(x)^{-1}$ établie dans le théorème 2.4.2. □

Les bornes (2.17) impliquent que le hessien de f varie petit à petit. Tout comme nous avons établi les bornes dans le lemme 1.3.1, en appliquant le même raisonnement à la nouvelle norme, nous obtenons $\forall x$ tq $|x - e| < 1$,

$$\|I - H(x)\| \leq \frac{1}{(1 - |x - e|)^2} - 1.$$

Nous savons déjà que si f est une fonction barrière et si $x, y \in \text{dom } f$ tq $\langle g(x), y - x \rangle \geq 0$, alors $y \in B_x(x, 4\theta_f + 1)$. Lorsque f est intrinsèquement auto-conjugué, nous avons un résultat plus fort.

Théorème 2.5.8

Supposons que K est un cône self-scaled.

Si $x \in \text{int } K$ satisfait $\langle g(e), x - e \rangle \geq 0$, alors $x \in B(e, \theta_f)$.

Preuve :

Supposons $x \in \text{int } K$ tq $\langle g(e), x - e \rangle \geq 0$. Comme $g(e) = -e$, nous avons

$$\begin{aligned} & \langle g(e), x - e \rangle \geq 0 \\ \iff & \langle -e, x - e \rangle \geq 0 \\ \iff & \langle x, e \rangle - \|e\|^2 \leq 0 \\ \iff & \langle x, e \rangle \leq \|e\|^2. \end{aligned}$$

De plus, comme x et $e \in \text{int } K (= \text{int } K^*)$, $\langle x, e \rangle > 0$. Par conséquent, nous avons

$$0 < \langle x, e \rangle \leq \|e\|^2.$$

Donc le point $\bar{x} = \frac{\|e\|^2}{\langle x, e \rangle} x$ est un multiple scalaire de x , où le scalaire n'est pas plus petit que 1. Nous pouvons également dire que x est une combinaison convexe de 0 et \bar{x} .

D'où

$$\begin{aligned} \|x - e\| &\leq \max\{\|0 - e\|, \|\bar{x} - e\|\} \\ &= \max\{\sqrt{\theta_f}, \|\bar{x} - e\|\} \text{ par le théorème 2.5.1.} \end{aligned}$$

$$\text{Soit } v \equiv \frac{1}{\|\bar{x} - e\|} (\bar{x} - e).$$

Comme $\|v\| = 1$ et $K^* = K$, nous avons $e - v \in K$ et $\langle \bar{x}, e - v \rangle \geq 0$, i.e. $\langle e + \|\bar{x} - e\|v, e - v \rangle \geq 0$.

De plus,

$$\begin{aligned} \langle v, e \rangle &= \frac{1}{\|\bar{x} - e\|} \langle \bar{x} - e, e \rangle \\ &= \frac{1}{\|\bar{x} - e\|} \left(\underbrace{\langle \bar{x} - e, e \rangle}_{\|e\|^2 = \theta_f} - \underbrace{\langle e, e \rangle}_{\theta_f} \right) \\ &= 0. \end{aligned}$$

Nous obtenons alors

$$\begin{aligned} \langle e + \|\bar{x} - e\|v, e - v \rangle &\geq 0 \\ \iff \underbrace{\langle e, e \rangle}_{\theta_f} - \underbrace{\langle e, v \rangle}_0 + \|\bar{x} - e\| \underbrace{\langle v, e \rangle}_0 - \|\bar{x} - e\| \underbrace{\langle v, v \rangle}_1 &\geq 0 \\ \iff \|\bar{x} - e\| &\leq \theta_f. \end{aligned}$$

Par conséquent, $\|x - e\| \leq \theta_f$.

Par ouverture, l'inégalité est stricte, i.e. $x \in B(e, \theta_f)$.

□

Théorème 2.5.9

Supposons que K est self-scaled et $x \in \text{int } K$.

Alors $\forall t \in \mathbb{R}, \exists ! x_t \in \text{int } K$ tq $H(x_t) = H(x)^t$.

Preuve :

L'unicité découle du théorème 2.5.4.

Définissons $T \equiv \{t \in \mathbb{R} : \exists x_t \in \text{int } K \text{ tq } H(x_t) = H(x)^t\}$.

Il suffit de montrer que T est un sous-ensemble de \mathbb{R} fermé et dense.

- Pour prouver que T est fermé, supposons que t est dans la fermeture de T et montrons que $t \in T$.

Soit $\{t_i\} \subset T$ une suite satisfaisant $t_i \rightarrow t$. Si la suite $\{x_{t_i}\}$ admet une sous-suite convergeant vers un point $x' \in \text{int } K$, par continuité du hessien, nous aurons $H(x') = H(x)^t$.

En effet, $\{t_{i_k}\} \subset T$. D'où $H(x_{t_{i_k}}) = H(x)^{t_{i_k}}$. De plus, $x_{t_{i_k}} \rightarrow x'$, $t_{i_k} \rightarrow t$, et le hessien est continu. Par conséquent, $H(x') = H(x)^t$. Dans ce cas, $t \in T$ et $x_t \equiv x'$, ce qui établit la fermeture.

Il reste à montrer que $\{x_{t_i}\}$ admet une sous-suite convergeant vers un point $x' \in \text{int } K$. Cela revient à montrer que la suite est bornée et n'admet aucune sous-suite convergeant vers un point de la frontière de K .

Si la suite (respectivement une sous-suite) satisfait $\|x_{t_i}\| \rightarrow \infty$, en tenant compte du fait que $H(x_{t_i}) \rightarrow H(x)^t$, nous obtenons

$$\|x_{t_i}\|_{x_{t_i}}^2 = \langle x_{t_i}, H(x)^t x_{t_i} \rangle \rightarrow \infty$$

lorsque $t_i \rightarrow \infty$, ce qui contredit le fait que $\|x_{t_i}\|_{x_{t_i}}^2 = \theta_f$ par le théorème 2.5.1.

Par conséquent, la suite $\{x_{t_i}\}$ est bornée.

Si la suite (respectivement une sous-suite) converge vers un point de la frontière de K , i.e. $x_{t_i} \rightarrow x \in \partial K$, et comme $B_{x_{t_i}}(x_{t_i}, 1) \subseteq \text{int } K$, nous en concluons que

$$\begin{aligned} & x \notin \text{int } K \\ \Rightarrow & x \notin B_{x_{t_i}}(x_{t_i}, 1) \\ \Rightarrow & \|x - x_{t_i}\|_{x_{t_i}}^2 > 1 \\ \Rightarrow & \underbrace{\langle x - x_{t_i}, H(x_{t_i})}_{\rightarrow 0} \underbrace{(x - x_{t_i})}_{\rightarrow 0} > 1 \\ \Rightarrow & \|H(x_{t_i})\| \text{ doit être non borné, i.e. } \|H(x_{t_i})\| \rightarrow \infty. \end{aligned}$$

Ceci contredit la fait que $H(x_{t_i}) \rightarrow H(x)^t$.

En conséquence, $\{x_{t_i}\}$ n'admet pas de sous-suite convergeant vers un point $x' \in \text{int } K$, ce qui complète la preuve que T est fermé.

- Montrons que T est dense.

Remarquons tout d'abord que 0 et $1 \in T$ car $H(e) = I = H(x)^0$ et $H(x) = H(x)^1$. Il suffit alors d'établir :

$$t \in T \Rightarrow -t, 2t \in T$$

et

$$t_1 \text{ et } t_2 \in T \Rightarrow \frac{1}{2}(t_1 + t_2) \in T.$$

1. Supposons que $t \in T$. Nous avons

$$\begin{aligned} H(-g(x_t)) &= H(x_t)^{-1} \text{ par le théorème 2.4.2} \\ &= H(x)^{-t} \text{ car } t \in T. \end{aligned}$$

D'où $-t \in T$.

En définissant $x_{2t} \equiv H(-g(x_t))e = H(x_t)^{-1}e$ et par (2.15), nous avons

$$H(x_{2t}) = H(H(x_t)^{-1}e) = H(x_t)H(e)H(x_t) = H(x_t)^2 = H(x)^{2t}.$$

D'où $2t \in T$.

2. Nous admettons que pour chaque produit scalaire $\langle \cdot, \cdot \rangle$ et pour chaque opérateur linéaire défini positif H , si t_1 et $t_2 \in \mathbb{R}$, H^{t_1} est défini positif par rapport au produit scalaire défini par $u, v \rightsquigarrow \langle u, H^{t_2}v \rangle$.

Supposons que t_1 et $t_2 \in T$. Par le point précédent, nous savons que $-t_2 \in T$. Soit w le point scaling pour la paire x_{t_1}, x_{-t_2} , i.e. $H(w)x_{t_1} = x_{-t_2}$.

D'où

$$\begin{aligned} H(H(w)x_{t_1}) &= H(x_{-t_2}) \\ \iff H(w)^{-1}H(x_{t_1})H(w)^{-1} &= H(x_{-t_2}) \text{ par (2.15)} \\ \iff (H(w)^{-1}H(x_{t_1}))^2 &= H(x_{-t_2})H(x_{t_1}) \\ \iff (H(x_{t_1})^{-1}H(w))^{-2} &= H(x)^{t_1-t_2} \\ \iff H_{x_{t_1}}(w)^{-2} &= H(x)^{t_1-t_2}. \end{aligned}$$

Comme la racine carrée d'un opérateur défini positif est unique, nous avons

$$H_{x_{t_1}}(w) = H(x)^{\frac{1}{2}(t_2-t_1)}.$$

En conséquence,

$$\begin{aligned} H(w) &= H(x_{t_1})H_{x_{t_1}}(w) \\ &= H(x)^{t_1}H(x)^{\frac{1}{2}(t_2-t_1)} \\ &= H(x)^{\frac{1}{2}(t_1+t_2)}. \end{aligned}$$

D'où $\frac{1}{2}(t_1 + t_2) \in T$.

□

2.5.4 Un théorème utile

Nous terminons ce chapitre par un théorème qui joue un rôle central dans l'analyse des méthodes primales-duales que nous considérons.

La preuve de ce théorème s'appuie sur le théorème et la proposition qui suivent.

Théorème 2.5.10

Supposons que K est un cône self-scaled.

Si $x \in \text{int } K$ et $\lambda \equiv \max\{\|H(x), \|H(x)^{-1}\|\}$, alors

$$\|x + g(x)\| \geq \frac{\lambda - 1}{\sqrt{\lambda}}.$$

Preuve :

Si $\lambda \leq 1$, alors $\frac{\lambda-1}{\sqrt{\lambda}} \leq 0$. Comme $\|x + g(x)\| \geq 0$, l'inégalité est toujours vraie.

Supposons maintenant que $\lambda > 1$. Considérons le cas où $\lambda = \|H(x)\|$, i.e. λ est la plus grande valeur propre de $H(x)$. Pour $0 \leq \epsilon < \lambda$, nous posons $z_\epsilon = \frac{-1}{\lambda-\epsilon} g(x)$.

Comme

$$\begin{aligned} H_{z_\epsilon}(x) &= H(z_\epsilon)^{-1}H(x) \\ &= \frac{1}{(\lambda - \epsilon)^2} \underbrace{H(-g(x))^{-1}}_{H(x)} H(x) \\ &= \frac{1}{(\lambda - \epsilon)^2} H(x)^2, \end{aligned}$$

la plus grande valeur propre de $H_{z_\epsilon}(x)$ est $\left(\frac{\lambda}{\lambda-\epsilon}\right)^2$.

En particulier, si $0 < \epsilon < \lambda$, la plus grande valeur propre est plus grande que 1, i.e. $\|H_{z_\epsilon}\| > \|H\| > 1$.

Par le théorème 2.5.6, nous avons $x - z_\epsilon \notin K$, i.e. $x + \frac{1}{\lambda - \epsilon} g(x) \notin K$.

De manière équivalente, en supposant $\epsilon \neq \lambda - 1$,

$$\frac{\lambda - 1 - \epsilon}{\lambda - \epsilon} \left(x + \frac{1}{\lambda - 1 - \epsilon} (x + g(x)) \right) \notin K.$$

Comme K est un cône, si $0 < \epsilon < \lambda - 1$, nous avons alors

$$x + \frac{1}{\lambda - 1 - \epsilon} (x + g(x)) \notin K.$$

En conséquence, comme $B_x(x, 1) \subseteq \text{int } K$ et que ϵ peut être choisi proche de 0, $\|x + g(x)\|_x \geq \lambda - 1$.

De plus,

$$\begin{aligned} \|x + g(x)\|_x^2 &= \langle x + g(x), H(x)(x + g(x)) \rangle \\ &\leq \lambda \langle x + g(x), x + g(x) \rangle \\ &= \lambda \|x + g(x)\|^2. \end{aligned}$$

D'où

$$\|x + g(x)\| \geq \frac{\|x + g(x)\|_x}{\sqrt{\lambda}} \geq \frac{\lambda - 1}{\sqrt{\lambda}}.$$

Dans le cas où $\lambda = \|H(x)^{-1}\| = \|H(-g(x))\|$, il suffit d'échanger les rôles de x et de $-g(x)$.

□

Proposition 2.5.3

Supposons que f est une fonction barrière logarithmiquement homogène dont le domaine est l'intérieur de K .

Si $x \in \text{int } K$, alors $\|x + g(x)\| \geq \min\{\frac{1}{4}, \|x - e\|\}$.

Preuve :

Considérons la fonction self-concordante

$$f'(x) \equiv \frac{1}{2} \|x\|^2 + f(x)$$

où $\|\cdot\|$ est la norme locale en e donnée par f . Le produit scalaire en e donné par f' est défini par $\langle \cdot, \cdot \rangle' = 2 \langle \cdot, \cdot \rangle$. Si g' désigne le gradient de f' par rapport au produit scalaire $\langle \cdot, \cdot \rangle'$, il suit que $g'(x) = \frac{1}{2}(x + g(x))$.

Par hypothèse, f est logarithmiquement homogène, ce qui implique que $g(e) = -e$ et par conséquent, $g'(e) = 0$.

Soit $x \in \bar{B}'(0, r)$ tq $\|x\|' \leq r$ où $r \leq \frac{1}{4}$.
Nous pouvons alors montrer que ([3], pg 34)

$$\exists u \in \bar{B}'\left(x, \frac{3r^2}{(1-r)^3}\right) \text{ tq } g'(e+u) = g'(e) + x = x.$$

Soit $\bar{u} = e + u$. Alors

$$\|\bar{u} - e\|' = \|u\|' \leq \|u - x\|' + \|x\|' \leq \frac{3r^2}{(1-r)^3} + r.$$

et $g'(\bar{u}) = g'(e+u) = x$.

Par conséquent, l'image de la boule fermée $\bar{B}'\left(e, r + \frac{3r^2}{(1-r)^3}\right)$ par l'application $x \rightsquigarrow g'(x)$ contient la boule fermée $\bar{B}'(0, r)$.

En particulier, si $r \leq \frac{1}{4\sqrt{2}}$, alors $\frac{r}{(1-r)^3} < 1$ et l'image de $\bar{B}'(e, 2r)$ contient $\bar{B}'(0, r)$.

Comme l'application $x \rightsquigarrow g'(x)$ est injective, il suit que $\forall x \in \text{int } K$,

$$\|g'(x)\|' \geq \min \left\{ \frac{1}{4\sqrt{2}}, \frac{1}{2}\|x - e\|' \right\}$$

i.e.

$$\frac{1}{\sqrt{2}}\|x + g(x)\| \geq \min \left\{ \frac{1}{4\sqrt{2}}, \frac{1}{\sqrt{2}}\|x - e\| \right\}.$$

□

Théorème 2.5.11

Supposons que K est un cône self-scaled. Si $x \in \text{int } K$, alors

$$\frac{\|x + g(x)\|}{\max \{\|H(x)^{1/2}\|, \|H(x)^{-1/2}\|\}} \geq \min \left\{ \frac{1}{5}, \frac{4}{5}\|x - e\| \right\}.$$

Preuve :

Par le théorème et la proposition qui précèdent, nous avons

$$\frac{\|x + g(x)\|}{\max \{\|H(x)^{1/2}\|, \|H(x)^{-1/2}\|\}} \geq \max \left\{ \frac{\lambda - 1}{\lambda}, \frac{\beta}{\sqrt{\lambda}} \right\}$$

où $\beta \equiv \min \left\{ \frac{1}{4}, \|x - e\| \right\}$.

Soit la fonction $\delta \rightsquigarrow \frac{\delta - 1}{\delta}$. Cette fonction est croissante lorsque $\delta > 0$ tandis que pour $\gamma > 0$ fixé, la fonction $\delta \rightsquigarrow \frac{\gamma}{\sqrt{\delta}}$ est décroissante.

En conséquence, la fonction $\delta \rightsquigarrow \max \left\{ \frac{\delta - 1}{\delta}, \frac{\gamma}{\sqrt{\delta}} \right\}$ atteint son minimum

lorsque $\frac{\delta - 1}{\delta} = \frac{\gamma}{\sqrt{\delta}}$, i.e. lorsque $\delta = \left(\frac{\gamma + \sqrt{\gamma^2 + 4}}{2} \right)^2$.

La valeur de la fonction en ce point vaut $\frac{2\gamma}{\gamma + \sqrt{\gamma^2 + 4}}$.

Dès lors,

$$\begin{aligned} \frac{\|x + g(x)\|}{\max \{\|H(x)^{1/2}\|, \|H(x)^{-1/2}\|\}} &\geq \frac{2\beta}{\beta + \sqrt{\beta^2 + 4}} \\ &\geq \frac{\beta}{1 + \beta} \\ &\geq \min \frac{\left\{ \frac{1}{4}, \|x - e\| \right\}}{1 + \frac{1}{4}} \\ &= \min \left\{ \frac{1}{5}, \frac{4}{5} \|x - e\| \right\}. \end{aligned}$$

□

Chapitre 3

La programmation semi-définie et la programmation conique du second ordre

3.1 Les fonctions barrières self-concordantes pour le cône des matrices semi-définies positives et le cône du second ordre

Nous avons vu dans le premier chapitre que la fonction barrière logarithmique $F(x) = -\sum_{i=1}^n \ln x_i$ était une fonction barrière self-concordante pour l'orthant positif \mathbb{R}_{++}^n .

Nous considérons maintenant le cas du cône des matrices semi-définies positives ainsi que le cône du second ordre.

Proposition 3.1.1

La fonction $F(x) = -\ln(\det X)$, où $\det X = \prod_{j=1}^n \lambda_j(X)$, définie sur le cône \mathcal{S}_{++}^n des matrices définies positives est une fonction barrière self-concordante avec une valeur de complexité $\theta_F = n$.

Preuve :

Montrons tout d'abord que F est une fonction barrière.

$$\begin{aligned} F(x) &= -\ln(\det X) \\ &= -\ln\left(\prod_{j=1}^n \lambda_j(X)\right) \\ &= -\sum_{j=1}^n \ln \lambda_j(X). \end{aligned}$$

Cette fonction est strictement convexe. Il reste donc à montrer que $F(x)$ tend vers l'infini lorsque X tend vers la frontière de \mathcal{S}_+^n .

Pour toute suite de matrices symétriques définies positives approchant la frontière de \mathcal{S}_+^n , au moins une des valeurs propres va tendre vers 0. Dès lors, $F(x) \rightarrow +\infty$ et $F(x)$ est bien une fonction barrière pour \mathcal{S}_{++}^n .

Montrons que F est self-concordante, i.e.

$$D^3F(X)[H, H, H] \leq 2[D^2F(x)[H, H]]^{3/2}.$$

$\forall X \in \mathcal{S}_{++}^n$ et $\forall H \in \mathcal{S}^n$, grâce aux notations introduites au paragraphe 1.3.1, nous avons

$$\begin{aligned} DF(X)[H] &= \langle -X^{-1}, H \rangle, \\ D^2F(X)[H, H] &= \langle X^{-1}HX^{-1}, H \rangle, \\ D^3F(X)[H, H, H] &= -2\langle X^{-1}HX^{-1}, HX^{-1}H \rangle. \end{aligned}$$

où $\langle A, B \rangle = \text{tr}(AB)$ avec A et B deux matrices symétriques.

Comme $\text{tr}(AB) = \text{tr}(BA)$, nous avons

$$\begin{aligned} D^2F(X)[H, H] &= \text{tr}(X^{-1}HX^{-1}H) \\ &= \text{tr}(X^{-1/2}X^{-1/2}HX^{-1/2}X^{-1/2}H) \\ &= \text{tr}(X^{-1/2}HX^{-1/2}X^{-1/2}HX^{-1/2}) \\ &= \text{tr}(X^{-1/2}HX^{-1/2})^2 \\ &= \|X^{-1/2}HX^{-1/2}\|_{Fr}^2 \\ &\geq 0. \end{aligned}$$

où $\|\cdot\|_{Fr}$ est la norme de Frobenius d'une matrice.

De la même façon, nous obtenons

$$D^3 F(X)[H, H, H] = -2\text{tr}(X^{-1/2} H X^{-1/2})^3.$$

D'où

$$|-2\text{tr}(X^{-1/2} H X^{-1/2})^3| \leq 2 (\text{tr}(X^{-1/2} H X^{-1/2})^2)^{3/2},$$

ce qui est équivalent à dire que F est self-concordante.

Pour terminer la preuve, il reste à montrer que $\theta_F = n$.

Par définition, $\theta_F = \sup_{X \in S_{++}^n} \|g_X(X)\|_X^2$,

où $g_X(X)$ est le gradient de F en X ,

$$\langle A, B \rangle_X = \langle A, H(X)B \rangle = \text{tr}(AH(X)B),$$

$H(X)$ est le hessien de F en X .

Nous savons que

$$g_X(X) = H(X)^{-1}g(X) \quad \text{où } g(X) = -X^{-1},$$

et $H(X)Z = X^{-1}ZX^{-1}$ i.e. $H(x)^{-1}Z = XZX$.

D'où

$$\begin{aligned} \|g_X(X)\|_X^2 &= \langle H(X)^{-1}g(X), H(X)^{-1}g(X) \rangle_X \\ &= \langle H(X)^{-1}g(X), g(X) \rangle \\ &= \langle XX^{-1}X, X^{-1} \rangle \\ &= \text{tr}(XX^{-1}XX^{-1}) \\ &= \text{tr}(I) \\ &= n. \end{aligned}$$

□

Afin de considérer une fonction barrière self-concordante pour le cône du second ordre, nous avons besoin du lemme suivant, qui est une caractérisation des fonctions logarithmiquement homogènes définies au chapitre précédent.

Lemme 3.1.1

Soit F une fonction self-concordante sur l'intérieur d'un cône convexe fermé K . Alors les assertions suivantes sont équivalentes :

1. $g(\alpha x) = \frac{1}{\alpha}g(x) \quad \forall \alpha > 0, \forall x \in \text{int } K$.
2. $\theta_F \in \mathbb{R}$ et $F(\alpha x) = F(x) - \theta_F \ln \alpha \quad \forall \alpha > 0, \forall x \in \text{int } K$.

En particulier, lorsque la première assertion est vérifiée, la fonction est une fonction barrière self-concordante avec $\theta_F = -\langle x, g(x) \rangle, \forall x \in \text{int } K$.

Preuve :

En dérivant les deux membres de la seconde assertion par rapport à x , nous obtenons

$$\begin{aligned}\alpha g(\alpha x) &= g(x) \\ \text{i.e. } g(\alpha x) &= \frac{1}{\alpha} g(x),\end{aligned}$$

ce qui est exactement la première assertion.

Pour démontrer l'implication inverse, supposons que f est une fonction self-concordante satisfaisant $g(\alpha x) = \frac{1}{\alpha} g(x)$, $\forall x \in \text{int } K$ et $\forall \alpha > 0$.

Montrons tout d'abord que $\|x\|_x^2 = \|g_x(x)\|_x^2 = -\langle x, g(x) \rangle$ et que ces quantités sont indépendantes de $x \in \text{int } K$.

En dérivant la première assertion par rapport à α , nous obtenons

$$H(\alpha x)x = -\frac{1}{\alpha^2} g(x).$$

En particulier, pour $\alpha = 1$,

$$\begin{aligned}H(x)x &= -g(x) \\ \text{i.e. } x &= -H(x)^{-1}g(x) \\ \text{ou encore } x &= -g_x(x).\end{aligned}$$

D'où

$$\|x\|_x^2 = \langle x, x \rangle_x = \langle x, H(x)x \rangle = -\langle x, g(x) \rangle = \|g_x(x)\|_x^2.$$

De plus, en dérivant $-\langle x, g(x) \rangle$ par rapport à x , nous avons

$$-g(x) - H(x)x = -H(x)\underbrace{[H(x)^{-1}g(x) + x]}_{-x} = 0,$$

ce qui montre bien que les quantités énoncées sont indépendantes de x .

Comme $\theta_F = \sup_{x \in \text{int } K} \|g_x(x)\|_x^2$ et $\|g_x(x)\|_x^2$ est indépendant de x , nous avons

$$\theta_F = -\langle x, g(x) \rangle.$$

Enfin, nous avons

$$\begin{aligned}
 f(\alpha x) &= f(x) + \int_1^\alpha \langle g(tx), x \rangle dt \quad \text{par la formule de la moyenne} \\
 &= f(x) + \int_1^\alpha \frac{1}{t} \langle g(x), x \rangle dt \quad \text{par hypothèse} \\
 &= f(x) - \theta_F \int_1^\alpha \frac{1}{t} dt \\
 &= f(x) - \theta_F \ln \alpha,
 \end{aligned}$$

ce qui termine la preuve. □

Nous avons alors la proposition suivante.

Proposition 3.1.2

La fonction $F(x, t) = -\ln(t^2 - \|x\|_2^2)$ définie sur le cône du second ordre

$$K = \{(x, t) \in \mathbb{R}^{n+1} \mid t > 0, \|x\|_2 \leq t\}$$

est une fonction barrière self-concordante avec une valeur de complexité $\theta_F = 2$. De plus, elle est logarithmiquement homogène.

Preuve :

Lorsque $\|x\|_2 \rightarrow t$, $F(x, t) \rightarrow -\ln 0 = +\infty$, ce qui montre que $F(x, t)$ est une fonction barrière.

Nous admettons que cette fonction est self-concordante (preuve technique).

Il reste à montrer que F est logarithmiquement homogène avec $\theta_F = 2$.

Pour démontrer l'homogénéité logarithmique, il nous suffit de montrer que la première condition du lemme est vérifiée.

Comme $g(x, t) = \frac{1}{t^2 - \|x\|_2^2} (2x, -2t)$,

nous avons

$$\begin{aligned}
 g(\alpha x, \alpha t) &= \frac{1}{\alpha^2 t^2 - \alpha^2 \|x\|_2^2} (2\alpha x, -2\alpha t) \\
 &= \frac{1}{\alpha} \frac{1}{t^2 - \|x\|_2^2} (2x, -2t) \\
 &= \frac{1}{\alpha} g(x, t),
 \end{aligned}$$

ce qui montre que F est logarithmiquement homogène.

Enfin,

$$\begin{aligned}\theta_F &= -\langle (x, t), g(x, t) \rangle \\ &= -\frac{2\|x\|_2^2 - 2t^2}{t^2 - \|x\|_2^2} \\ &= 2.\end{aligned}$$

□

3.2 La programmation semi-définie

3.2.1 Le problème de programmation semi-définie

Un problème de programmation semi-définie est un problème conique où le cône K est le cône \mathcal{S}_+^n de toutes les matrices symétriques semi-définies positives. Un tel problème peut s'exprimer de la façon suivante :

$$(SDP) \begin{cases} \min & \langle C, X \rangle \\ \text{s.c.} & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \\ & X = X^T \succeq 0 \end{cases}$$

$$\text{où } \begin{cases} X, A_i, C \in \mathcal{S}^n, \quad \forall i = 1, \dots, m \\ b_i \in \mathbb{R}, \quad \forall i = 1, \dots, m \\ \langle D, E \rangle = \text{tr}(DE). \end{cases}$$

Avant de donner des exemples de problèmes de programmation semi-définie, nous allons étudier son problème dual. Pour cela, nous avons besoin de la proposition suivante.

Proposition 3.2.1

Le cône \mathcal{S}_+^n est un cône convexe fermé pointé auto-dual dont l'intérieur est non vide.

Preuve :

Nous devons seulement prouver que le cône \mathcal{S}_+^n est auto-dual, i.e.

$$\mathcal{S}_+^n = (\mathcal{S}_+^n)^* \equiv \{S \in \mathcal{S}^n \text{ tq } \text{tr}(SX) \geq 0 \quad \forall X \in \mathcal{S}_+^n\}.$$

La preuve se fait en deux parties. Dans la première, nous montrons que $\mathcal{S}_+^n \subseteq (\mathcal{S}_+^n)^*$. Nous démontrons l'inclusion inverse dans la seconde partie.

Afin d'introduire d'autres exemples, nous avons besoin du résultat suivant.

Proposition 3.2.2 (Complément de Schur)

Si $A \succ 0$, alors

$$\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0 \text{ (respectivement } \succ 0)$$

si et seulement si $C - B^T A^{-1} B \succeq 0$ (respectivement $\succ 0$).

La matrice $C - B^T A^{-1} B$ est appelée complément de Schur de A .

Pour obtenir le résultat, il suffit d'observer que la matrice

$$\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} = \begin{pmatrix} I & 0 \\ B^T A^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{pmatrix} \begin{pmatrix} I & A^{-1} B \\ 0 & I \end{pmatrix}$$

est semi-définie positive (respectivement définie positive) si et seulement si $C - B^T A^{-1} B$ a la même propriété.

Exemple 3.2.2 : Minimiser la norme- L_2 d'une matrice

Considérons le problème suivant :

$$\begin{cases} \min & \|P(z)\|_2 \\ \text{s.c.} & z \in \mathbb{R}^m. \end{cases}$$

Pour écrire ce problème sous la forme d'un problème de programmation semi-définie, nous devons montrer que si P est une matrice $m \times n$, alors $\lambda \geq \|P\|_2$ si et seulement si la matrice $\begin{pmatrix} \lambda I & P \\ P^T & \lambda I \end{pmatrix} \succeq 0$ où $\|P\|_2^2 = \lambda_{\max}(P^T P)$.

En effet, si $\lambda = 0$, nous avons

$$\|P\|_2 \leq 0 \iff P = 0 \iff \begin{pmatrix} 0 & P \\ P^T & 0 \end{pmatrix} \succeq 0.$$

Nous montrons uniquement que $\begin{pmatrix} 0 & P \\ P^T & 0 \end{pmatrix} \succeq 0 \Rightarrow P = 0$ car toutes les autres implications sont évidentes.

Soit $P = [P_{ij}]$, $1 \leq i \leq m$, $1 \leq j \leq n$ tq $\begin{pmatrix} 0 & P \\ P^T & 0 \end{pmatrix} \succeq 0$.

Alors $\forall x \in \mathbb{R}^m, \forall y \in \mathbb{R}^n$,

$$\begin{aligned} & \begin{pmatrix} x^T & y^T \end{pmatrix} \begin{pmatrix} 0 & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \geq 0 \\ \iff & \begin{pmatrix} y^T P^T & x^T P \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \geq 0 \\ \iff & y^T P^T x + x^T P y \geq 0 \\ \iff & 2x^T P y \geq 0. \end{aligned}$$

Donc, $x^T P y \geq 0, \forall x \in \mathbb{R}^m, \forall y \in \mathbb{R}^n$.

En particulier, si $x_{[i]}$ est le vecteur de \mathbb{R}^m composé de 0 sauf la $i^{\text{ème}}$ composante qui vaut 1 et $y_{[j]}$ le vecteur de \mathbb{R}^n composé de 0 sauf la $j^{\text{ème}}$ composante qui vaut 1, nous obtenons

$$\begin{aligned} x_{[i]}^T P y_{[j]} &= P_{ij} \geq 0 \\ \text{et } x_{[i]}^T P (-y_{[j]}) &= -P_{ij} \geq 0. \end{aligned}$$

Dès lors, $P_{ij} = 0, \forall i = 1, \dots, m, \forall j = 1, \dots, n$. D'où $P = 0$.

Supposons maintenant que $\lambda > 0$. Par la proposition 3.2.2, nous avons alors

$$\begin{aligned} \begin{pmatrix} \lambda I & P \\ P^T & \lambda I \end{pmatrix} \succeq 0 & \iff \lambda I - P^T (\lambda I)^{-1} P \geq 0 \\ & \iff \lambda I - \frac{1}{\lambda} P^T P \geq 0 \\ & \iff \lambda^2 I - P^T P \geq 0 \\ & \iff \lambda^2 \geq \lambda_{\max}(P^T P) = \|P\|_2^2 \\ & \iff \lambda \geq \|P\|_2. \end{aligned}$$

Nous pouvons alors résoudre le problème de minimiser $\|P(z)\|_2$ en résolvant le problème

$$\begin{cases} \max & -\lambda \\ \text{s.c.} & \begin{pmatrix} \lambda I & P(z) \\ P(z)^T & \lambda I \end{pmatrix} \succeq 0. \end{cases}$$

Exemple 3.2.3 : Programmation linéaire

Considérons le problème de programmation linéaire

$$(PL) \begin{cases} \max & b^T y \\ \text{s.c.} & A^T y \leq c \end{cases}$$

où $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ et $y \in \mathbb{R}^m$.

Ce problème peut être écrit comme le problème de programmation semi-définie sous forme duale

$$\begin{cases} \max & b^T y \\ \text{s.c.} & \text{diag}(c - A^T y) \succeq 0 \end{cases} \\ \iff \begin{cases} \max & b^T y \\ \text{s.c.} & C - \sum_{i=1}^m y_i A_i \succeq 0 \end{cases}$$

où $C = \text{diag}(c)$ et $A_i = \text{diag}(a_i)$ avec $a_i = i^{\text{ème}}$ colonne de A^T .

Exemple 3.2.4 : Un problème de programmation non linéaire quasi-convexe

Considérons le problème

$$\begin{cases} \min & \frac{(b^T y)^2}{d^T y} \\ \text{s.c.} & A^T y \leq c \end{cases}$$

où nous supposons que $d^T y > 0 \forall y$ admissible.

Comme la fonction objectif peut être écrite comme $(b^T y)(d^T y)^{-1}(b^T y)$, nous voyons la ressemblance avec le complément de Schur, et nous pouvons alors vérifier assez facilement que ($\forall y$ admissible),

$$\begin{aligned} \lambda \geq \frac{(b^T y)^2}{d^T y} &\iff \lambda d^T y - (b^T y)^2 \geq 0 \\ &\iff \begin{pmatrix} \lambda & b^T y \\ b^T y & d^T y \end{pmatrix} \succeq 0. \end{aligned}$$

Notre problème non linéaire peut alors être écrit sous la forme

$$\begin{cases} \min & \lambda \\ \text{s.c.} & \text{diag}(c - A^T y) \succeq 0 \\ & \begin{pmatrix} \lambda & b^T y \\ b^T y & d^T y \end{pmatrix} \succeq 0. \end{cases}$$

Ce problème a deux contraintes semi-définies positives. Celles-ci peuvent être combinées en une seule :

$$\text{diag} \left(\text{diag}(c - A^T y), \begin{pmatrix} \lambda & b^T y \\ b^T y & d^T y \end{pmatrix} \right) \succeq 0$$

où C et les A_i sont des diagonales par blocs, avec m blocs de dimension 1×1 et un bloc de dimension 2×2 .

Exemple 3.2.5 : Programmation avec contraintes convexes quadratiques

Nous considérons l'optimisation d'une fonction linéaire soumise à des contraintes convexes quadratiques, i.e. le problème

$$\begin{cases} \max & b^T y \\ \text{s.c.} & f_i(y) \leq 0, \quad i = 1, \dots, m \end{cases}$$

où $f_i(y) \equiv y^T C_i y - d_i^T y - \epsilon_i$, avec $C_i \succeq 0$, $i = 1, \dots, m$.

Soit $C_i = G_i^T G_i$, $i = 1, \dots, m$.

Alors, en utilisant le complément de Schur, $f_i(y) \leq 0$ peut s'écrire

$$\begin{pmatrix} I & G_i y \\ y^T G_i^T & d_i^T y + \epsilon_i \end{pmatrix} \succeq 0.$$

Dès lors, le problème est équivalent à

$$\begin{cases} \max & b^T y \\ \text{s.c.} & \begin{pmatrix} I & G_i y \\ y^T G_i^T & d_i^T y + \epsilon_i \end{pmatrix} \succeq 0, \quad i = 1, \dots, m. \end{cases}$$

La contrainte peut également s'écrire

$$A \equiv \begin{pmatrix} (1 + d_i^T y + \epsilon_i)I & \begin{pmatrix} 1 - d_i^T y - \epsilon_i \\ 2G_i y \end{pmatrix} \\ \begin{pmatrix} 1 - d_i^T y - \epsilon_i \\ 2G_i y \end{pmatrix}^T & 1 + d_i^T y + \epsilon_i \end{pmatrix} \succeq 0.$$

En effet, si $1 + d_i^T y + \epsilon_i > 0$, alors

$$\begin{aligned}
 & A \succeq 0 \\
 \Leftrightarrow & 1 + d_i^T y + \epsilon_i - \begin{pmatrix} 1 - d_i^T y - \epsilon_i \\ 2G_i y \end{pmatrix}^T (1 + d_i^T y + \epsilon_i)^{-1} I \begin{pmatrix} 1 - d_i^T y - \epsilon_i \\ 2G_i y \end{pmatrix} \succeq 0 \\
 & \text{par le complément de Schur} \\
 \Leftrightarrow & 1 + d_i^T y + \epsilon_i - \frac{1}{1 + d_i^T y + \epsilon_i} [(1 - d_i^T y - \epsilon_i)^2 + 4y^T G_i^T G_i y] \geq 0 \\
 \Leftrightarrow & (1 + d_i^T y + \epsilon_i)^2 - (1 - d_i^T y - \epsilon_i)^2 - 4y^T G_i^T G_i y \geq 0 \\
 \Leftrightarrow & 4(d_i^T y + \epsilon_i) - 4y^T C_i y \geq 0 \\
 \Leftrightarrow & (d_i^T y + \epsilon_i) - y^T C_i y \geq 0 \\
 \Leftrightarrow & f_i(y) \leq 0.
 \end{aligned}$$

De plus, nous avons

$$\begin{aligned}
 f_i(y) \leq 0 & \Rightarrow d_i^T y + \epsilon_i \geq 0 \\
 & \Rightarrow 1 + d_i^T y + \epsilon_i > 0
 \end{aligned}$$

et $A \succeq 0 \Rightarrow 1 + d_i^T y + \epsilon_i > 0$.

Sinon, si $1 + d_i^T y + \epsilon_i = 0$, alors

$$A = \begin{pmatrix} 0 & \begin{pmatrix} 2 \\ 2G_i y \end{pmatrix} \\ \begin{pmatrix} 2 \\ 2G_i y \end{pmatrix}^T & 0 \end{pmatrix} \succeq 0$$

avec un mineur principal négatif, ce qui est impossible.

L'avantage de la seconde formulation est que la contrainte semi-définie

$$\begin{pmatrix} \alpha I & v \\ v^T & \alpha \end{pmatrix} \succeq 0$$

peut être exprimée comme

$$\begin{pmatrix} \alpha \\ v \end{pmatrix} \in L^2 \equiv \left\{ \begin{pmatrix} \beta \\ w \end{pmatrix} \text{ tq } \beta \geq \|w\|_2 \right\}$$

où L^2 est le cône de Lorentz.

C'est une manière plus efficace de résoudre le problème car les méthodes de points intérieurs qui résolvent les problèmes de programmation conique du second ordre ont, dans le pire des cas, une complexité meilleure qu'une méthode de programmation semi-définie.

3.3 La programmation conique du second ordre

3.3.1 Le problème de programmation conique du second ordre

Considérons le problème

$$(SOCP) \left\{ \begin{array}{l} \min \sum_{i=1}^m c_i^T x_i + d_i t_i \\ \text{s.c.} \sum_{i=1}^m A_i x_i + a_i t_i = b \\ \|x_i\|_2 \leq t_i, \quad i = 1, \dots, m \end{array} \right.$$

où $\left\{ \begin{array}{l} x_i \in \mathbb{R}^{n_i} \text{ et } t_i \in \mathbb{R}, \quad i = 1, \dots, m \text{ sont les variables d'optimisation,} \\ c_i \in \mathbb{R}^{n_i}, d_i \in \mathbb{R}, A_i \in \mathbb{R}^{n \times n_i}, a_i \in \mathbb{R}^n \text{ et } b \in \mathbb{R}^n \text{ sont les paramètres} \\ \text{du problème.} \end{array} \right.$

Ici, le cône K est le produit des cônes $K_1 \times \dots \times K_m$, où K_i est le cône de Lorentz

$$L^{n_i+1} \equiv \{(x_i, t_i) \text{ tq } \|x_i\| \leq t_i\}.$$

Tout comme pour la programmation semi-définie, avant de donner des exemples de problèmes de programmation conique du second ordre, nous étudions son dual. Pour cela, nous avons besoin de la proposition suivante.

Proposition 3.3.1

Le cône L^{n+1} est un cône convexe fermé pointé auto-dual dont l'intérieur est non vide.

Preuve :

Nous démontrons seulement que le cône de Lorentz L^{n+1} est auto-dual, i.e.

$$L^{n+1} = (L^{n+1})^* \equiv \{(s, \rho) \in \mathbb{R}^n \times \mathbb{R} \text{ tq } x^T s + t\rho \geq 0 \quad \forall (x, t) \in L^{n+1}\}.$$

Soit $(s, \rho) \in L^{n+1}$. Alors, $\forall (x, t) \in L^{n+1}$, nous avons $t \geq \|x\|$ et $\rho \geq \|s\|$.

D'où $x^T s + t\rho \geq x^T s + \|x\| \|s\| \geq 0$ par Cauchy-Schwarz, et par conséquent, $(s, \rho) \in (L^{n+1})^*$.

Inversément, soit $(s, \rho) \in (L^{n+1})^*$. Nous devons montrer que $(s, \rho) \in L^{n+1}$, i.e. $\rho \geq \|s\|$.

Si $s \neq 0$, alors $(-\frac{s}{\|s\|}, 1) \in L^{n+1}$.

D'où $-\frac{s^T}{\|s\|}s + 1 \cdot \rho \geq 0$, i.e. $\rho \geq \|s\|$ et donc $(s, \rho) \in L^{n+1}$.

Si $s = 0$, comme $(0, 1) \in L^{n+1}$ et $(s, \rho) \in (L^{n+1})^*$, nous avons $\rho \geq 0$.

D'où $x^T s + t\rho = \rho \geq 0$, i.e. $(s, \rho) \in L^{n+1}$.

□

Comme le cône dual d'un produit de cônes est le produit des cônes duals et comme les cônes de Lorentz sont auto-duals, nous avons

$$(L^{n_1+1} \times \dots \times L^{n_m+1})^* = L^{n_1+1} \times \dots \times L^{n_m+1}.$$

Nous pouvons maintenant écrire le dual du problème (SOCP).

$$(SOCD) \begin{cases} \max & b^T y \\ \text{s.c.} & \begin{pmatrix} c_i \\ d_i \end{pmatrix} - \begin{pmatrix} A_i^T \\ a_i^T \end{pmatrix} y \in L^{n_i+1}, \quad i = 1, \dots, m \end{cases}$$

$$\iff (SOCD) \begin{cases} \max & b^T y \\ \text{s.c.} & \|c_i - A_i^T y\| \leq d_i - a_i^T y, \quad i = 1, \dots, m. \end{cases}$$

Lorsque $A_i = 0$ et $c_i = 0$, $\forall i = 1, \dots, m$, le problème (SOCD) devient un problème de programmation linéaire. Si $a_i = 0$, $\forall i = 1, \dots, m$, la $i^{\text{ème}}$ contrainte de (SOCD) se réduit à $\|c_i - A_i^T y\| \leq d_i$, ce qui est équivalent (en supposant $d_i \geq 0$) à la contrainte quadratique convexe $\|c_i - A_i^T y\|^2 \leq d_i^2$.

Nous avons vu dans la section précédente qu'un problème de programmation quadratique convexe avec des contraintes quadratiques pouvait s'exprimer sous la forme d'un problème de programmation conique du second ordre. Donc, la programmation conique du second ordre inclut des classes importantes de problèmes d'optimisation convexe telles que la programmation linéaire, la programmation quadratique et la programmation quadratique avec contraintes quadratiques.

Remarquons cependant que la programmation conique du second ordre est moins générale que la programmation semi-définie.

En effet, le cône du second ordre peut être imbriqué dans le cône des matrices semi-définies positives car

$$\|u\| \leq t \iff \begin{pmatrix} tI & u \\ u^T & t \end{pmatrix} \succeq 0$$

i.e. une contrainte du cône du second ordre est équivalente à une inégalité matricielle linéaire.

En utilisant ces propriétés, le problème (SOCD) peut s'exprimer comme un problème de programmation semi-définie de la façon suivante :

$$\begin{cases} \max & b^T y \\ \text{s.c.} & \begin{pmatrix} (d_i - a_i^T y)I & c_i - A_i^T y \\ (c_i - A_i^T y)^T & d_i - a_i^T y \end{pmatrix} \succeq 0, \quad i = 1, \dots, m. \end{cases}$$

Cependant, résoudre un problème de programmation conique du second ordre via un problème de programmation semi-définie n'est pas intéressant car, comme nous l'avons déjà signalé, les méthodes de points intérieurs qui résolvent le problème de programmation conique du second ordre sont plus efficaces que celles qui résolvent le problème de programmation semi-définie.

3.3.2 Exemples

Exemple 3.3.1

Beaucoup de problèmes autres que les problèmes de programmation quadratique avec contraintes quadratiques peuvent être modélisés comme des problèmes de programmation conique du second ordre. C'est le cas par exemple des problèmes impliquant des sommes et le maximum de normes.

Soit $A_i \in \mathbb{R}^{n_i \times n}$ et $b_i \in \mathbb{R}^{n_i}$, $i = 1, \dots, p$, donnés. Le problème sans contrainte

$$\min \sum_{i=1}^p \|A_i x + b_i\|$$

peut être exprimé comme un problème de programmation conique du second ordre en introduisant des variables auxiliaires t_1, \dots, t_p .

Nous obtenons alors

$$\begin{cases} \min & \sum_{i=1}^p t_i \\ \text{s.c.} & \|A_i x + b_i\| \leq t_i, \quad i = 1, \dots, p. \end{cases}$$

De façon similaire, les problèmes impliquant un maximum de normes peuvent être exprimés comme un problème de programmation conique du second ordre. Le problème

$$\min \max_{1 \leq i \leq p} \|A_i x + b_i\|$$

est équivalent au problème de programmation conique du second ordre

$$\begin{cases} \min & t \\ \text{s.c.} & \|A_i x + b_i\| \leq t, \quad i = 1, \dots, p. \end{cases}$$

Exemple 3.3.2

La programmation conique du second ordre peut aussi être utilisée pour résoudre certains problèmes robustes simples d'optimisation convexe dans lesquels il y a une incertitude sur les données.

Considérons le problème de programmation linéaire suivant :

$$\begin{cases} \min & c^T x \\ \text{s.c.} & a_i^T x \leq b_i, \quad i = 1, \dots, m \end{cases}$$

dans lequel il y a une certaine incertitude ou variation des paramètres c, a_i, b_i . Nous supposons ici que c et b_i sont fixés et que les a_i sont connus et se trouvent dans des ellipsoïdes :

$$a_i \in \epsilon_i \equiv \{\bar{a}_i + P_i u \text{ tq } \|u\| \leq 1\}$$

où $\forall i = 1, \dots, m, P_i$ est une matrice symétrique semi-définie positive.

Dans le pire des cas, nous devons satisfaire les contraintes pour toutes les valeurs possibles des paramètres a_i . Nous avons alors le problème de programmation linéaire robuste suivant :

$$(PLR) \begin{cases} \min & c^T x \\ \text{s.c.} & a_i^T x \leq b_i, \quad \forall a_i \in \epsilon_i, \quad i = 1, \dots, m. \end{cases}$$

La contrainte peut également s'exprimer comme

$$\max\{a_i^T x \text{ tq } a_i \in \epsilon_i\} = \bar{a}_i^T x + \|P_i x\| \leq b_i,$$

ce qui est une contrainte sur le cône du second ordre.

Le problème de programmation linéaire robuste peut alors être exprimé comme un problème de programmation conique du second ordre :

$$\begin{cases} \min & c^T x \\ \text{s.c.} & \bar{a}_i^T x + \|P_i x\| \leq b_i \quad i = 1, \dots, m. \end{cases}$$

Chapitre 4

Méthodes de points intérieurs en programmation conique

4.1 La trajectoire centrale primale-duale

Soit K un cône convexe pointé fermé auto-dual dont l'intérieur est non vide. Nous considérons les problèmes coniques primal et dual :

$$(CP) \quad \begin{cases} \min & \langle c, x \rangle \\ \text{s.c.} & Ax = b \\ & x \in K, \end{cases}$$
$$(CD) \quad \begin{cases} \max & \langle b, y \rangle \\ \text{s.c.} & A^*y + s = c \\ & s \in K. \end{cases}$$

où A est un opérateur linéaire de \mathbb{R}^n dans \mathbb{R}^m , $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ et A^* est l'adjointe de A .

Dans le but de mettre au point des méthodes de points intérieurs primales-duales pour résoudre cette paire de problèmes, nous supposons qu'il existe une fonction barrière self-concordante logarithmiquement homogène F sur l'intérieur de K et $\forall \mu > 0$, nous considérons les problèmes barrières primal et dual correspondants :

$$(CP_\mu) \quad \begin{cases} \min & \langle c, x \rangle + \mu F(x) \\ \text{s.c.} & Ax = b \\ & (x \in \text{int}K), \end{cases}$$

et

$$(CD_\mu) \begin{cases} \max & \langle b, y \rangle - \mu F(s) \\ \text{s.c.} & A^*y + s = c \\ & (s \in \text{int}K). \end{cases}$$

Théorème 4.1.1

Si (CD) est strictement admissible, alors (CP_μ) a une solution unique $\forall \mu > 0$.

Si (CP) est strictement admissible, alors (CD_μ) a une solution unique $\forall \mu > 0$.

Preuve :

Etant donné que le problème dual de (CD) est le problème (CP) , la deuxième partie est une conséquence directe de la première. Nous nous limitons donc à démontrer le premier point.

- Soit (\hat{y}, \hat{s}) un point strictement admissible pour le problème (CD) . Alors, $A^*\hat{y} + \hat{s} = c$ et $\hat{s} \in \text{int}K$. De plus, $\forall x \in \text{int}K$, nous avons

$$\begin{aligned} \langle c, x \rangle &= \langle A^*\hat{y} + \hat{s}, x \rangle \\ &= \langle A^*\hat{y}, x \rangle + \langle \hat{s}, x \rangle \\ &= \langle \hat{y}, Ax \rangle + \langle \hat{s}, x \rangle \\ &= \underbrace{\langle \hat{y}, b \rangle}_{\text{constante}} + \langle \hat{s}, x \rangle. \end{aligned}$$

D'où le problème (CP_μ) est équivalent à

$$(CP'_\mu) \begin{cases} \min & \langle \hat{s}, x \rangle + \mu F(x) \\ \text{s.c.} & Ax = b \\ & (x \in \text{int}K). \end{cases}$$

- Comme $\hat{s} \in \text{int}K$,

$$\exists \alpha > 0 \text{ tq } \langle \hat{s}, x \rangle \geq \alpha \|x\| \quad \forall x \in K. \quad (4.1)$$

En effet, si ce n'est pas le cas,

$$\forall \alpha > 0, \exists y \in K, \|y\| = 1 \text{ tq } \langle \hat{s}, y \rangle < \alpha.$$

Comme $\hat{s} \in \text{int}K$,

$$\exists \epsilon > 0 \text{ tq } B(\hat{s}, \epsilon) \subseteq K = K^* = \{x^* \text{ tq } \langle x^*, x \rangle \geq 0 \quad \forall x \in K\}.$$

En particulier, pour $0 < \alpha < \epsilon$, $\exists y \in K$, $\|y\| = 1$, $\langle \hat{s}, y \rangle < \alpha$.
De plus, $\hat{s} - \epsilon y \in B(\hat{s}, \epsilon) \subseteq K^*$.

D'où

$$\langle \hat{s} - \epsilon y, y \rangle = \langle \hat{s}, y \rangle - \epsilon \|y\|^2 < \alpha - \epsilon < 0,$$

ce qui est impossible car $y \in K$ et $\hat{s} - \epsilon y \in K^*$.

- $\forall x \in \text{int}K$ tq $\|x\| > \frac{\mu\theta_F}{\alpha}$, la fonction objectif de (CP'_μ) est strictement croissante sur $\{x + tx \text{ tq } t \geq 0\}$. En d'autres mots, la fonction $v(t) \equiv \langle \hat{s}, x + tx \rangle + \mu F(x + tx)$ est strictement croissante pour $t > 0$.

En effet,

$$\begin{aligned} v'(t) &= \langle \hat{s}, x \rangle + \mu \langle g((1+t)x), x \rangle \\ &= \langle \hat{s}, x \rangle + \frac{\mu}{1+t} \langle g(x), x \rangle \text{ par le lemme 3.1.1} \\ &\geq \alpha \|x\| - \frac{\mu}{1+t} \langle g(x), 0-x \rangle \text{ par (4.1)} \\ &\geq \alpha \|x\| - \frac{\mu}{1+t} \theta_F \text{ par la proposition 1.3.3} \\ &\geq \alpha \|x\| - \mu \theta_F \text{ car } 1+t > 1 \\ &> 0 \text{ car } \alpha \|x\| > \frac{\mu\theta_F}{\alpha}. \end{aligned}$$

Dès lors, le problème (CP_μ) est équivalent au problème

$$\left\{ \begin{array}{l} \min \langle \hat{s}, x \rangle + \mu F(x) \\ \text{s.c. } Ax = b \\ \|x\| \leq \frac{\mu\theta_F}{\alpha} \\ (x \in \text{int}K). \end{array} \right.$$

Comme F est une fonction barrière, tous les ensembles niveaux de ce problème sont compacts. D'où la fonction F est continue et le problème admet au moins une solution. L'unicité découle de la convexité stricte de la fonction objectif.

□

Définition 4.1.1

La trajectoire centrale primale-duale associée aux problèmes (CP) et (CD) est l'ensemble

$$\mathcal{C} \equiv \left\{ (x(\mu), s(\mu)) \text{ tq } \begin{array}{l} x(\mu) \text{ est solution de } (CP_\mu) \\ s(\mu) \text{ est solution de } (CD_\mu) \end{array}, \text{ avec } \mu > 0 \right\}.$$

$$\begin{aligned}
(CP'_\mu) \quad & \begin{cases} \min & \langle \hat{s}, x \rangle + \mu F(x) \\ \text{s.c.} & x \in L + \hat{x} \\ & (x \in \text{int}K), \end{cases} \\
(CD'_\mu) \quad & \begin{cases} \min & \langle \hat{x}, s \rangle + \mu F(s) \equiv f(s) \\ \text{s.c.} & s \in L^\perp + \hat{s} \\ & (s \in \text{int}K), \end{cases}
\end{aligned}$$

où L est le noyau de A .

Pour montrer que s est une solution du problème barrière dual, il suffit de montrer que $s \in \text{int}K$, $\hat{s} - s \in L^\perp$ et $\hat{x} + \mu g(s) \in L$ (condition d'optimalité : $\nabla f(s) \perp (L^\perp)$).

Par le lemme qui précède, $-g(x(\mu)) \in \text{int}K$ car $x(\mu) \in \text{int}K$. D'où, K étant un cône, $s = -\mu g(x(\mu)) \in \text{int}K$. Par hypothèse, $x(\mu)$ est solution de (CP_μ) . Cela implique que $x(\mu) \in L + \hat{x}$ et $\hat{s} + \underbrace{\mu g(x(\mu))}_{=-s} \in L^\perp$.

Pour prouver que $\hat{x} + \mu g(s) \in L$, remarquons tout d'abord que

$$\begin{aligned}
g(s) &= g(-\mu g(x(\mu))) \\
&= \frac{1}{\mu} g(-g(x(\mu))) \quad \text{par le lemme 3.1.1} \\
&= -\frac{1}{\mu} x(\mu) \quad \text{par le lemme qui précède.}
\end{aligned}$$

Par conséquent, $\hat{x} + \mu g(s) = \hat{x} - x(\mu) \in L$.

□

Une conséquence importante des relations $s(\mu) = -\mu g(x(\mu))$ et $x(\mu) = -\mu g(s(\mu))$ est qu'en suivant une des trajectoires centrales $\{x(\mu) \text{ tq } \mu > 0\}$ et $\{s(\mu) \text{ tq } \mu > 0\}$, nous pouvons générer l'autre. Le problème dual peut être résolu à l'aide du problème primal et inversement. Le saut de dualité le long de la trajectoire centrale est donné par la proposition suivante.

Proposition 4.1.2

Soit K un cône convexe fermé auto-dual dont l'intérieur est non vide et F une fonction barrière self-concordante logarithmiquement homogène sur l'intérieur de K . Le saut de dualité associé au point $(x(\mu), s(\mu))$ sur la trajectoire centrale est

$$\langle c, x(\mu) \rangle - \langle b, y(\mu) \rangle = \langle x(\mu), s(\mu) \rangle = \mu \theta_F.$$

En particulier, le saut de dualité tend vers 0 lorsque $\mu \rightarrow 0^+$.

Preuve :

Comme $x(\mu)$ et $s(\mu)$ sont admissibles pour (CP_μ) et (CD_μ) respectivement, nous avons :

$$\begin{aligned} \langle c, x(\mu) \rangle - \langle b, y(\mu) \rangle &= \langle A^*y(\mu), x(\mu) \rangle + \langle s(\mu), x(\mu) \rangle - \langle b, y(\mu) \rangle \\ &= \langle y(\mu), Ax(\mu) \rangle + \langle s(\mu), x(\mu) \rangle - \langle b, y(\mu) \rangle \\ &= \langle s(\mu), x(\mu) \rangle \\ &= -\mu \langle g(x(\mu)), x(\mu) \rangle, \end{aligned}$$

où $s(\mu) = -\mu g(x(\mu))$.

Comme F est logarithmiquement homogène, par le lemme 3.1.1,

$$-\langle g(x(\mu)), x(\mu) \rangle = \theta_F.$$

□

Dans la proposition qui suit, nous caractérisons la trajectoire centrale au moyen d'un système d'équations.

Proposition 4.1.3

Soit K un cône convexe fermé auto-dual dont l'intérieur est non vide et F une fonction barrière self-concordante logarithmiquement homogène sur l'intérieur de K . Supposons que les problèmes (CP) et (CD) sont strictement admissibles. Alors $(x(\mu), s(\mu)) \in \mathcal{C}$ si et seulement si $(x(\mu), s(\mu))$ est solution du système

1. $Ax = b$, $x \in \text{int}K$ (primal strictement admissible),
2. $A^*y + s = c$, $s \in \text{int}K$ (dual strictement admissible),
3. $s = -\mu g(x)$ ($\Leftrightarrow x = -\mu g(s)$),

avec la convention que y est défini de manière unique lorsque s est connu (car A est surjectif).

Preuve :

La condition nécessaire a été démontrée à la proposition 4.1.1.

Supposons donc que les conditions 1. et 3. sont satisfaites par certains points x et s . En particulier, x et s sont strictement admissibles pour les problèmes (CP'_μ) et (CD'_μ) , où (\hat{x}, \hat{s}) est le point admissible de (CP) et (CD) donné dans l'hypothèse.

Par la dualité, il suffit de montrer que x est solution du problème (CP'_μ) , i.e. que $\hat{s} + \mu g(x) \in L^\perp$ (par la condition d'optimalité). Mais cela est immédiat par le fait que $s = -\mu g(x)$ est admissible pour (CD'_μ) , i.e. $-\mu g(x) \in L^\perp + \hat{s}$. □

Une fois la trajectoire centrale caractérisée, il reste à définir une sorte de distance de n'importe quel point de K à cette trajectoire centrale. Afin d'obtenir l'optimum, nous suivrons approximativement cette trajectoire centrale, exactement comme en programmation linéaire. Comme la matrice hessienne $H(x)$ de la fonction barrière est supposée définie positive sur K , nous pouvons utiliser son inverse pour mesurer les distances entre les points de K . Plus précisément, nous définissons la norme d'un vecteur $y \in K$ par

$$\|y\|_x^{*2} = \langle H(x)^{-1}y, y \rangle.$$

Pour un point $z = (x, s)$ strictement admissible pour (CP) et (CD) ,

$$z \in \mathcal{C} \iff s + \mu g(x) = 0, \mu > 0 \iff \frac{1}{\mu}s + g(x) = 0, \mu > 0.$$

Une bonne mesure de proximité est

$$\begin{aligned} \text{dist}(z, z(\mu)) &= \left\| \frac{1}{\mu}s + g(x) \right\|_x^* \geq 0 \\ &= 0 \text{ ssi } s = -\mu g(x). \end{aligned}$$

Bien qu'écrite sous forme non symétrique selon x, s , cette distance est symétrique en x, s , comme le montre le théorème qui suit.

Théorème 4.1.2

$\forall \mu > 0$ et $x, s \in \text{int}K$, nous avons

$$\left\| \frac{1}{\mu}s + g(x) \right\|_x^* = \left\| \frac{1}{\mu}x + g(s) \right\|_s^*.$$

Preuve :

Nous faisons la preuve uniquement pour le cône des matrices semi-définies positives \mathcal{S}_+^n .

Soient X et S deux matrices carrées symétriques de dimension n et soit $F(x) = -\ln(\det X)$. Nous savons que dans ce cas, $g(X) = -X^{-1}$ et $H(X)^{-1}U = XUX$.

Nous avons donc

$$\begin{aligned}
 \left(\left\| \frac{1}{\mu} S + g(X) \right\|_X^* \right)^2 &= \langle H(X)^{-1} \left(\frac{1}{\mu} S + g(X) \right), \frac{1}{\mu} S + g(X) \rangle \\
 &= \langle X \left(\frac{1}{\mu} S - X^{-1} \right) X, \frac{1}{\mu} S - X^{-1} \rangle \\
 &= \text{tr} \left(X \left(\frac{1}{\mu} S - X^{-1} \right) X \left(\frac{1}{\mu} S - X^{-1} \right) \right) \\
 &= \text{tr} \left(\left(\frac{1}{\mu} X^{1/2} S X^{1/2} - I \right)^2 \right) \\
 &= \left\| \frac{1}{\mu} X^{1/2} S X^{1/2} - I \right\|_{Fr}^2.
 \end{aligned}$$

Donc,

$$(\text{dist}(z, z(\mu)))^2 = \left\| \frac{1}{\mu} X^{1/2} S X^{1/2} - I \right\|_{Fr}^2. \quad (4.2)$$

Pour prouver la symétrie, observons que

$$\begin{aligned}
 \left\| \frac{1}{\mu} X^{1/2} S X^{1/2} - I \right\|_{Fr}^2 &= \text{tr} \left(\left(\frac{1}{\mu} X^{1/2} S X^{1/2} - I \right)^2 \right) \\
 &= \text{tr} \left(\frac{1}{\mu^2} X^{1/2} S X^{1/2} X^{1/2} S X^{1/2} - \frac{2}{\mu} X^{1/2} S X^{1/2} + I \right) \\
 &= \text{tr} \left(\frac{1}{\mu^2} X^{1/2} S X^{1/2} X^{1/2} S X^{1/2} \right) - \frac{2}{\mu} \text{tr} \left(X^{1/2} S X^{1/2} \right) + \text{tr}(I) \\
 &= \text{tr} \left(\frac{1}{\mu^2} X S X S - \frac{2}{\mu} X S + I \right).
 \end{aligned}$$

De manière similaire, en remplaçant X par S , nous obtenons

$$\left(\left\| \frac{1}{\mu} X + g(S) \right\|_S^* \right)^2 = \left(\frac{1}{\mu^2} S X S X - \frac{2}{\mu} S X + I \right).$$

Comme $\text{tr}(AB) = \text{tr}(BA)$ pour des matrices symétriques, il suit que

$$\left\| \frac{1}{\mu} X^{1/2} S X^{1/2} - I \right\|_{Fr}^2 = \left\| \frac{1}{\mu} S^{1/2} X S^{1/2} - I \right\|_{Fr}^2.$$

□

La dernière proposition de cette section caractérise le saut de dualité de points proches de la trajectoire centrale.

Proposition 4.1.4

Si $Z = (X, S)$ est une paire de points strictement admissible pour (CP) et (CD) telle que $\text{dist}(Z, Z(\mu)) \leq 1$, alors :

$$\begin{aligned} \text{saut de dualité de } (X, S) &= \langle S, X \rangle \\ &\leq 2 \cdot \text{saut de dualité de } (X(\mu), S(\mu)) = 2\mu\theta_F. \end{aligned}$$

Preuve :

A nouveau, nous ne démontrons le résultat que pour le cône des matrices semi-définies positives \mathcal{S}_+^n .

Par (4.2), la relation $\text{dist}(Z, Z(\mu)) \leq 1$ signifie que

$$\left\| \frac{1}{\mu} X^{1/2} S X^{1/2} - I \right\|_{Fr}^2 \leq 1$$

i.e.

$$\|X^{1/2} S X^{1/2} - \mu I\|_{Fr}^2 \leq \mu^2.$$

En notant δ le vecteur des valeurs propres de la matrice symétrique $X^{1/2} S X^{1/2}$, nous concluons grâce à l'inégalité précédente que

$$\sum_{i=1}^n (\delta_i - \mu)^2 \leq \mu^2.$$

D'où

$$\begin{aligned} \text{le saut de dualité de } (X, S) &= \text{tr}(X, S) = \text{tr}(X^{1/2} S X^{1/2}) \\ &= \sum_{i=1}^n \delta_i = \sum_{i=1}^n (\mu + (\delta_i - \mu)) \\ &\leq n\mu + \sum_{i=1}^n |\delta_i - \mu| \quad \text{car } \delta_i \leq \mu + |\delta_i - \mu| \\ &\leq n\mu + \sqrt{n} \sqrt{\sum_{i=1}^n (\delta_i - \mu)^2} \quad \text{car } \|\cdot\|_1 \leq \sqrt{n} \|\cdot\|_2 \\ &\leq n\mu + \sqrt{n}\mu \\ &\leq \theta_F \mu + \theta_F \mu \quad \text{car } \theta_F = n \\ &= 2\theta_F \mu. \end{aligned}$$

□

4.2 Un cadre de travail primal-dual

Notre but est maintenant de dessiner la trajectoire centrale primale-duale associée aux problèmes (CP) et (CD) pour obtenir une solution (approchée) à ces problèmes. Supposons que (μ, x, s) , avec $\mu > 0$, est l'itéré actuel, (x, s) est un point strictement admissible primal-dual qui est une bonne approximation du point $(x(\mu), s(\mu))$ sur la trajectoire centrale. La question est : comment mettre à jour ce point pour obtenir un nouvel itéré (μ_+, x_+, s_+) avec des propriétés similaires et une valeur plus petite du paramètre barrière ($\mu_+ < \mu$).

Afin de répondre à cette question, rappelons que le point $(x(\mu), s(\mu))$ est l'unique solution du système

$$\begin{cases} A(x) = b, A^*(y) + s = c, \\ x, s \in \text{int}K, \\ G_\mu(x, s) \equiv s + \mu g(x) = 0. \end{cases} \quad (4.3)$$

où $g(x)$ est le gradient de la fonction barrière F en x .

Nous supposons que A est surjectif (i.e. A^* est injectif). Dans ce cas, y est déterminé dès que s est connu. La première partie du système (4.3) étant linéaire et la seconde non linéaire, nous linéarisons la deuxième partie autour de (x, s) pour obtenir

$$G_\mu(x, s) + \frac{\partial G_\mu}{\partial x}(x, s)(x_+ - x) + \frac{\partial G_\mu}{\partial s}(x, s)(s_+ - s) = 0$$

i.e.

$$s + \mu g(x) + \mu H(x)(x_+ - x) + (s_+ - s) = 0. \quad (4.4)$$

Les corrections $\Delta x = x_+ - x$ et $\Delta s = s_+ - s$ doivent être calculées de telle sorte que x_+ et s_+ satisfassent la première partie du système (4.3) et l'équation linéarisée (4.4).

En d'autres mots, les corrections résoudre le système

$$\begin{cases} A(\Delta x) = 0, A^*(\Delta y) + \Delta s = 0 \\ s + \Delta s + \mu g(x) + \mu H(x)\Delta x = 0. \end{cases} \quad (4.5)$$

Remarquons que si nous écrivons la condition $G_\mu(x, s) = 0$ sous une forme équivalente, par exemple $x + \mu g(s) = 0$, la dernière équation du système (4.5) change. Dès lors, les corrections Δx et Δs sont différentes, ce qui donne lieu à un autre algorithme. C'est le cas en programmation semi-définie où il existe d'autres équations équivalentes à $G_\mu(x, s) = 0$ qui permettent d'obtenir des algorithmes très efficaces.

A partir de maintenant, nous concentrons donc notre attention sur les problèmes de programmation semi-définie. En premier lieu, rappelons les formes primale et duale de ce type de problème :

$$(CP) \begin{cases} \min & \langle C, X \rangle \\ \text{s.c.} & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \\ & X \in \mathcal{S}_+^n, \end{cases}$$

$$(CD) \begin{cases} \max & b^T y \\ \text{s.c.} & \sum_{i=1}^m y_i A_i + S = C \\ & S \in \mathcal{S}_+^n. \end{cases}$$

Dans notre cas, $g(X) = -X^{-1}$ et le système d'équations non linéaires $G_\mu(X, S) = S + \mu g(X) = 0$ peut s'écrire sous les formes équivalentes suivantes :

$$XS = \mu I, \quad SX = \mu I.$$

En choisissant la première forme, nous obtenons que la trajectoire centrale associée à (CP) et (CD) est définie par le système

$$\begin{cases} \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \\ X, S \in \text{int}\mathcal{S}_+^n \\ \sum_{i=1}^m y_i A_i + S = C \\ XS = \tau I, \quad \forall \tau > 0. \end{cases} \quad (4.6)$$

Tout comme en programmation linéaire, nous utiliserons $\tau = \sigma\mu$ où $\sigma \in [0, 1]$ et μ est la mesure de dualité $\mu = \frac{\langle X, S \rangle}{n}$.

Avec cette forme pour la trajectoire centrale, les corrections $(\Delta X, \Delta y, \Delta S)$ sont solutions du système linéaire

$$\begin{cases} \langle A_i, \Delta X \rangle = 0, \quad i = 1, \dots, m \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S = 0 \\ \Delta X S + X \Delta S = \sigma \mu I - X S. \end{cases} \quad (4.7)$$

Une fois que ces corrections ont été calculées, nous pouvons calculer le nouvel itéré $X_+ = X + \Delta X$, $y_+ = y + \Delta y$, $S_+ = S + \Delta S$.

Si $P = I$, en général X et S ne commutent pas, avec pour conséquence que l'unicité de la direction AHO n'est pas garantie. Nous devons imposer une condition plus forte pour obtenir l'unicité, par exemple $\|S^{1/2}XS^{1/2} - \nu I\|_{Fr} \leq \frac{\nu}{2}$, $\forall \nu > 0$.

Nous obtenons finalement l'algorithme de base des méthodes de points intérieurs pour les problèmes de programmation semi-définie.

Algorithme 4.2.1 (Un cadre de travail primal-dual)

Soit (X^0, y^0, S^0) strictement admissible pour le primal et le dual.

Pour $k = 0, 1, 2, \dots$

1. Résoudre le système

$$\begin{cases} \langle A_i, \Delta X^k \rangle = 0, & i = 1, \dots, m \\ \sum_{i=1}^m \Delta y_i^k A_i + \Delta S^k = 0 \\ H_P(\Delta X^k S^k + X^k \Delta S^k) = \sigma_k \mu_k I - H_P(X^k S^k) \end{cases}$$

où $\sigma_k \in [0, 1]$ et $\mu_k = \frac{1}{n} \langle X^k, S^k \rangle$.

2. $(X^{k+1}, y^{k+1}, S^{k+1}) = (X^k, y^k, S^k) + \alpha_k (\Delta X^k, \Delta y^k, \Delta S^k)$ en choisissant α_k tel que $(X^{k+1}, S^{k+1}) \succ 0$.

4.3 La direction de Nesterov-Todd

La méthode utilisant la direction NT est plus robuste que les autres méthodes, dans le sens où des problèmes de stagnation, qui surviennent lorsque nous prenons de très petits pas, se produisent moins souvent.

4.3.1 Définition et calcul de W

La direction NT est obtenue en choisissant $P = W^{-1/2}$, où

$$\begin{aligned} W &= X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} \\ &= S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2}. \end{aligned}$$

La proposition qui suit établit les principales propriétés de la matrice W , ce qui nous permettra de justifier le choix de P .

Proposition 4.3.1

La matrice $P = W^{-1/2}$ est l'unique matrice P symétrique définie positive telle que $PXP = P^{-1}SP^{-1}$. En particulier, $W^{-1}XW^{-1} = S$ et $WSW = X$.

Preuve :

Comme X et S sont symétriques et définies positives, W l'est également et par conséquent, P est symétrique et définie positive.

De plus, $PXP = P^{-1}SP^{-1}$,
i.e. $W^{-1/2}XW^{-1/2} = W^{1/2}SW^{1/2}$,
i.e. $WSW = X$ ou encore $W^{-1}XW^{-1} = S$,
ce qui est immédiat car

$$\begin{aligned} WSW &= X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2}SX^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} \\ &= X^{1/2}IX^{1/2} = X. \end{aligned}$$

D'un autre côté, soit P une matrice symétrique définie positive telle que $PXP = P^{-1}SP^{-1}$ et posons $W = P^{-2}$. La matrice W est aussi symétrique et définie positive.

Montrons que

$$\begin{aligned} W &= X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} \\ &= S^{-1/2}(S^{1/2}XS^{1/2})^{-1/2}S^{-1/2}. \end{aligned}$$

Comme $PXP = P^{-1}SP^{-1}$, nous avons $W^{-1/2}XW^{-1/2} = W^{1/2}SW^{1/2}$, i.e. $WSW = X$ et $W^{-1}XW^{-1} = S$. En prémultipliant $WSW = X$ par $X^{-1/2}$ et en postmultipliant par $SX^{1/2}$, nous obtenons

$$\begin{aligned} X^{-1/2}WSWSX^{1/2} &= X^{-1/2}XSX^{1/2} \\ \Leftrightarrow X^{-1/2}WSX^{1/2}X^{-1/2}WSX^{1/2} &= X^{1/2}SX^{1/2} \quad \text{car } X^{1/2}X^{-1/2} = I \\ \Leftrightarrow X^{-1/2}WSX^{1/2} &= (X^{1/2}SX^{1/2})^{1/2} \\ \Leftrightarrow X^{-1/2}S^{-1}W^{-1}X^{1/2} &= (X^{1/2}SX^{1/2})^{-1/2} \quad \text{en prenant l'inverse,} \end{aligned}$$

et en prémultipliant et postmultipliant par $X^{1/2}$ cette dernière équation, nous avons

$$\begin{aligned} X^{1/2}X^{-1/2}S^{-1}W^{-1}X^{1/2}X^{1/2} &= X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} \\ \Leftrightarrow S^{-1}W^{-1}X &= X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} \\ \Leftrightarrow S^{-1}SW &= X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} \quad \text{car } W^{-1}XW^{-1} = S \\ \Leftrightarrow W &= X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2}. \end{aligned}$$

Montrons maintenant que $W = S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2}$. En prémultipliant l'équation $W^{-1}XW^{-1} = S$ par $S^{-1/2}$ et en la postmultipliant par $XS^{1/2}$, nous obtenons

$$\begin{aligned} S^{-1/2}W^{-1}XW^{-1}XS^{1/2} &= S^{-1/2}SXS^{1/2} \\ \Leftrightarrow S^{-1/2}W^{-1}XS^{1/2}S^{-1/2}W^{-1}XS^{1/2} &= S^{-1/2}SXS^{1/2} \\ \Leftrightarrow S^{-1/2}W^{-1}XS^{1/2} &= (S^{1/2}XS^{1/2})^{1/2}, \end{aligned}$$

et en prémultipliant et postmultipliant cette dernière relation par $S^{-1/2}$, nous avons

$$\begin{aligned} S^{-1/2}S^{-1/2}W^{-1}XS^{1/2}S^{-1/2} &= S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2} \\ \Leftrightarrow S^{-1}W^{-1}X &= S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2} \\ \Leftrightarrow S^{-1}SW &= S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2} \quad \text{car } W^{-1}XW^{-1} = S \\ \Leftrightarrow W &= S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2}. \end{aligned}$$

□

Comme les deux matrices PXP et $P^{-1}SP^{-1}$ sont égales, elles commutent et par la proposition 4.2.2, la direction NT est définie uniquement par le système

$$\begin{cases} \langle A_i, \Delta X \rangle = 0, & i = 1, \dots, m \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S = 0 \\ H_P(\Delta XS + X\Delta S) = \sigma\mu I - H_P(XS) \end{cases} \quad (4.10)$$

où $P = W^{-1/2}$.

En utilisant la matrice W , il est possible d'écrire la dernière équation de ce système en termes de W au lieu de P . Ce système devient

$$\begin{cases} \langle A_i, \Delta X \rangle = 0, & i = 1, \dots, m \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S = 0 \\ W^{-1}\Delta XW^{-1} + \Delta S = \sigma\mu X^{-1} - S \end{cases} \quad (4.11)$$

où $\mu = \frac{1}{n}\langle X, S \rangle$.

Proposition 4.3.2

Les solutions des systèmes (4.10) et (4.11) sont les mêmes.

Ceci prouve que la solution de (4.11) est aussi une solution de (4.10). De plus, comme par la proposition 4.2.2 la solution est unique, les systèmes sont équivalents. □

La proposition qui suit indique comment calculer la matrice W .

Proposition 4.3.3

Soit $X = LL^T$ et $S = RR^T$ les factorisations de Cholesky des matrices symétriques définies positives X et S . Soit aussi $R^T L = UDV^T$ la décomposition de $R^T L$ en valeurs singulières (U et V sont des matrices orthogonales et D est une matrice diagonale). Alors $W = GG^T$ où $G = LVD^{-1/2}$.

Preuve :

- $Q = L^{-1}X^{1/2}$ est une matrice orthogonale.
En effet, $QQ^T = L^{-1}X^{1/2}X^{1/2}L^{-T} = L^{-1}(LL^T)L^{-T} = I$.
- $X^{1/2}SX^{1/2} = (Q^T V)D^2(V^T Q)$.
Comme $LQ = X^{1/2} = (X^{1/2})^T = Q^T L^T$, nous avons

$$\begin{aligned} X^{1/2}SX^{1/2} &= Q^T L^T (RR^T) LQ \\ &= Q^T (R^T L)^T (R^T L) Q \\ &= Q^T (UDV^T)^T (UDV^T) Q \\ &= Q^T V D U^T U D V^T Q \\ &= (Q^T V) D^2 (V^T Q). \end{aligned}$$

- $(X^{1/2}SX^{1/2})^{-1/2} = Q^T V D^{-1} V^T Q$.
 Q et V sont orthogonales. Nous avons donc $Q^{-1} = Q^T$, $V^{-1} = V^T$ et $Q^T V$ est orthogonale.
Par le point précédent, nous avons

$$(X^{1/2}SX^{1/2})^{-1} = (V^T Q)^{-1} D^{-2} (Q^T V)^{-1} = Q^T V D^{-2} V^T Q.$$

La racine carrée de $Q^T V D^{-2} V^T Q$ est $Q^T V D^{-1} V^T Q$ car

$$Q^T V D^{-1} V^T Q Q^T V D^{-1} V^T Q = Q^T V D^{-2} V^T Q.$$

Soit U n'importe quelle matrice $n \times n$ inversible et considérons le changement d'échelle où X est remplacé par $\hat{X} \equiv U^{-1}XU^{-T}$. Nous avons également $\hat{A}_i \equiv U^T A_i U$, $\hat{C} \equiv U^T C U$ et $\hat{S} \equiv U^T S U$. Ces nouvelles matrices mises à échelle définissent le problèmes mis à échelle

$$(S\hat{D}P) \begin{cases} \min & \langle \hat{C}, \hat{X} \rangle \\ \text{s.c.} & \langle \hat{A}_i, \hat{X} \rangle = b_i, \quad i = 1, \dots, m \\ & \hat{X} \succeq 0, \end{cases}$$

$$(S\hat{D}D) \begin{cases} \max & b^T y \\ \text{s.c.} & \sum_{i=1}^m y_i \hat{A}_i + \hat{S} = \hat{C} \\ & \hat{S} \succeq 0. \end{cases}$$

Si (X, y, S) est admissible pour le problème original et son dual, (\hat{X}, y, \hat{S}) est admissible pour le problème mis à échelle. De même, si $(\Delta X, \Delta y, \Delta S)$ est la direction de recherche associée au point (X, y, S) pour le problème original, alors la direction de recherche en un point (\hat{X}, y, \hat{S}) mis à échelle est $(\Delta \hat{X}, \Delta y, \Delta \hat{S})$.

La direction $(\Delta X, \Delta y, \Delta S)$ est invariante par changement d'échelle si

$$(\Delta \hat{X}, \Delta y, \Delta \hat{S}) = (U^{-1} \Delta X U^{-T}, \Delta y, U^T \Delta S U) \quad (4.15)$$

La direction NT associée au problème original vérifie le système d'équations (4.11), tandis que celle associée au problème mis à échelle vérifie

$$(\widehat{NT}) \begin{cases} \langle \hat{A}_i, \Delta \hat{X} \rangle = 0, \quad i = 1, \dots, m \\ \sum_{i=1}^m \Delta y_i \hat{A}_i + \Delta \hat{S} = 0 \\ \hat{W}^{-1} \Delta \hat{X} \hat{W}^{-1} + \Delta \hat{S} = \sigma \mu \hat{X}^{-1} - \hat{S} \end{cases}$$

où $\hat{W} = U^{-1} W U^{-T}$.

Montrons que $(\Delta \hat{X}, \Delta y, \Delta \hat{S})$ défini par (4.15) est solution du système ci-dessus.

En effet,

- $\langle \hat{A}_i, \Delta \hat{X} \rangle = \text{tr}(U^T A_i U U^{-1} \Delta X U^{-T})$
 $= \text{tr}(U^T A_i \Delta X U^{-T})$
 $= \text{tr}(A_i \Delta X)$
 $= \langle A_i, \Delta X \rangle$
 $= 0.$
- $\sum_{i=1}^m \Delta y_i \hat{A}_i + \Delta \hat{S} = \sum_{i=1}^m \Delta y_i U^T A_i U + U^T \Delta S U$
 $= U^T \left(\sum_{i=1}^m \Delta y_i A_i + \Delta S \right) U$
 $= 0.$
- $\hat{W}^{-1} \Delta \hat{X} \hat{W}^{-1} + \Delta \hat{S} = U^T W^{-1} U U^{-1} \Delta X U^{-T} U^T W^{-1} U + U^T \Delta S U$
 $= U^T W^{-1} \Delta X W^{-1} U + U^T \Delta S U$
 $= U^T \underbrace{(W^{-1} \Delta X W^{-1} + \Delta S)}_{\sigma \mu \Delta X^{-1} - S} U$
 $= \sigma \mu \hat{X}^{-1} - \hat{S}.$

Nous pouvons montrer que la direction HKM est également invariante par changement d'échelle, ce qui n'est pas le cas de la direction AHO.

4.4 Un algorithme basé sur la direction NT : l'algorithme prédicteur-correcteur de Mehrotra

Dans les sections précédentes, nous avons défini la trajectoire centrale primale-duale pour les problèmes de programmation semi-définie ainsi qu'une mesure de proximité par rapport à cette trajectoire. Nous avons également défini des directions de recherche qui permettent de passer d'un itéré au suivant. Nous obtenons alors des algorithmes analogues à ceux établis en programmation linéaire, et de même complexité. Tout comme en programmation linéaire, ces algorithmes peuvent devenir encore plus efficaces en choisissant soigneusement le paramètre de centrage σ .

Conclusion

A travers ce mémoire, nous avons montré qu'il existe des méthodes efficaces pour résoudre des problèmes de programmation convexe, et en particulier des problèmes de programmation conique. Ces méthodes sont connues sous le nom de méthodes de points intérieurs primales-duales. Elles sont basées sur l'existence d'une fonction barrière self-concordante, qui permet de définir une trajectoire centrale menant à l'ensemble solution. Les algorithmes existants permettent de résoudre des problèmes convexes dont la fonction objectif est linéaire et les contraintes linéaires ou coniques.

Différentes directions de recherche existent. Nous n'avons cité ici que les trois plus courantes, et une seule d'entre elles a été étudiée un peu plus en détail. D'autres directions pourraient être étudiées et une comparaison de celles-ci pourrait être effectuée.

Actuellement, des méthodes existent pour résoudre des problèmes convexes non linéaires. Ces méthodes pourraient également faire l'objet d'une étude plus approfondie.

La procédure peut se résumer à l'aide de l'algorithme suivant

Algorithme A.2.1 (Factorisation de Cholesky)

Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique et définie positive.

```

pour  $i = 1, \dots, n$ 
     $l_{ii} \leftarrow \sqrt{a_{ii}}$ 
    pour  $j = i + 1, \dots, n$ 
         $l_{ji} \leftarrow a_{ji} / l_{ii}$ 
        pour  $k = i + 1, \dots, j$ 
             $l_{jk} \leftarrow a_{jk} - l_{ji} l_{ki}$ 
        fin (pour)
    fin (pour)
fin (pour)
    
```

Le coût d'une telle factorisation est de $n^3/3$ opérations. Cette factorisation est utilisée pour résoudre des systèmes linéaires de la forme $Ax = b$ où A est une matrice symétrique définie positive et $b \in \mathbb{R}^n$. En utilisant le facteur de Cholesky L , le système peut être écrit comme $LL^T x = b$ et peut être résolu en deux étapes :

Résoudre $Ly = b$ pour obtenir y par substitution progressive ;
 Résoudre $L^T x = y$ pour obtenir x par substitution régressive.

A.3 Décomposition en valeurs singulières

Soit $A \in \mathbb{R}^{m \times n}$ et soit $k = \min\{m, n\}$. La décomposition en valeurs singulières de A est $A = UDV^T$, où $U \in \mathbb{R}^{m \times m}$ et $V \in \mathbb{R}^{n \times n}$ sont des matrices orthogonales, $D \in \mathbb{R}^{m \times n}$ est définie par $d_{ii} = \sigma_i \geq 0$, $i = 1, \dots, k$, $d_{ij} = 0$, $i \neq j$. Les quantités $\sigma_1, \dots, \sigma_k$ sont appelées les valeurs singulières de A .

Théorème A.3.1

Soit $A \in \mathbb{R}^{m \times n}$. Alors la décomposition en valeurs singulières de A existe, les éléments diagonaux σ_i de D sont les racines carrées non négatives des valeurs propres de AA^T si $m \leq n$, ou de $A^T A$ si $m \geq n$, et les colonnes de U et V sont les vecteurs propres de AA^T et $A^T A$ respectivement. Le nombre de valeurs singulières non nulles est égal au rang de A .

Annexe B

Calcul différentiel

B.1 Gradient

Soit $f : D_f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction dont le domaine est D_f . Nous supposons que D_f est un sous-ensemble ouvert de \mathbb{R}^n . Soit $\langle \cdot, \cdot \rangle$ un produit scalaire arbitraire sur \mathbb{R}^n et soit $\|\cdot\|$ la norme induite par ce produit scalaire. La fonction f est dite différentiable en $x \in D_f$ si il existe un vecteur $g(x)$ satisfaisant

$$\lim_{\|\Delta x\| \rightarrow 0} \frac{f(x + \Delta x) - f(x) - \langle g(x), \Delta x \rangle}{\|\Delta x\|} = 0.$$

Le vecteur $g(x)$ est le gradient de f en x par rapport au produit scalaire $\langle \cdot, \cdot \rangle$. La fonction f est dite différentiable si elle est différentiable en chaque point $x \in D_f$. Elle est dite de classe C^1 si la fonction $x \mapsto g(x)$ est continue sur D_f . Le gradient change lorsque le produit scalaire change.

Proposition B.1.1

Si S est défini positif et si f est différentiable en x , alors le gradient de f en x par rapport au produit scalaire $\langle \cdot, \cdot \rangle_S$ est $S^{-1}g(x)$.

La propriété d'être différentiable et d'être de classe C^1 est indépendante du produit scalaire. Le gradient dépend du produit scalaire, mais pas la différentiabilité.

B.2 Hessien

La fonction f est dite deux fois différentiable en $x \in D_f$ si f est de classe C^1 et si il existe un opérateur linéaire $H(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfaisant

$$\lim_{\|\Delta x\| \rightarrow 0} \frac{g(x + \Delta x) - g(x) - H(x)\Delta x}{\|\Delta x\|} = 0.$$

Si l'opérateur $H(x)$ existe, il est appelé le hessien de f en x par rapport au produit scalaire $\langle \cdot, \cdot \rangle$.

La fonction est dite deux fois différentiable si elle est deux fois différentiable en chaque $x \in D_f$. Si la fonction $y \rightarrow H(y)$ est continue en x , alors $H(x)$ est auto-adjointe. Si la fonction $x \rightarrow H(x)$ est continue sur D_f , alors f est dite de classe C^2 . Si ce produit scalaire est le produit scalaire habituel et que le hessien s'exprime comme une matrice, le fait d'être auto-adjoint est équivalent à être symétrique.

La propriété d'être de classe C^2 ne dépend pas du produit scalaire choisi, tandis que le hessien en dépend.

Proposition B.2.1

Si S est une matrice symétrique définie positive et si f est deux fois différentiable en x , alors le hessien de f en x par rapport au produit scalaire $\langle \cdot, \cdot \rangle_S$ est $S^{-1}H(x)$.

Bibliographie

- [1] Y. Nesterov and A. Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, 1994.
- [2] Y. Nesterov and M.J. Todd. *Self-Scaled Barriers and Interior-Point Methods in Convex Programming*. Math. Oper. Res., 22 (1997), p. 1-42.
- [3] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. MPS-SIAM Series on Optimization, Philadelphia, 2001.
- [4] J.J. Strodiot. *Interior-Point Methods in Convex Optimization*. FUNDP, 2002.
- [5] M.J. Todd, K.C. Toh and R.H. Tutuncu. *On the Nesterov-Todd Direction in Semidefinite Programming*. SIAM J. Optimization, Vol. 8, No. 3, p. 769-796, August 1998.
- [6] S.Wright. *Primal-dual Interior-Point Methods*. SIAM, Philadelphia, 1997.