



**Polynômes positifs, sommes de carrés
Programmation semi-définie positive
Application à la minimisation**

Richard Leroy

Mémoire de DEA

Sous la direction de Marie-Françoise Roy

Année 2003-2004



Table des matières

1	Préliminaires matriciels	1
1.1	Matrices symétriques ; matrices (semi-)définies positives . . .	1
1.1.1	Produit scalaire ; normes	1
1.1.2	Matrices (semi-)définies positives	2
1.1.3	Le cône $\mathcal{S}_n^+(\mathbb{R})$	5
1.2	Un peu de calcul matriciel	7
2	Programmation semi-définie positive	9
2.1	Formulation du problème	9
2.1.1	Problème primal standard	9
2.1.2	Formulations non standard	11
2.2	Deux cas particuliers	13
2.2.1	Programmation linéaire	13
2.2.2	Programmation quadratique	13
2.3	Dualité	14
2.3.1	Formulation du problème dual	14
2.3.2	Dualité forte	16
2.4	Quelques applications	20
2.4.1	Problèmes aux valeurs propres	20
2.4.2	Problèmes géométriques	21
2.4.3	Une application en statistiques	23
2.4.4	Une application en théorie des graphes	24
2.5	Un algorithme de point intérieur	27

2.5.1	Le principe : l'introduction d'une fonction potentiel	28
2.5.2	Un algorithme de réduction du potentiel	33
3	Polynômes positifs, sommes de carrés ; minimisation	45
3.1	Reconnaître les sommes de carrés : un algorithme	45
3.1.1	L'espace \mathcal{L}_f	45
3.1.2	Matrice de Gram	48
3.1.3	Application : exemple de Mötzkin	52
3.2	Minimisation des polynômes à l'aide de la programmation semi-définie	53
3.2.1	La méthode de Parrilo et Sturmfels	53
3.2.2	Une alternative : GloptiPoly	58
3.3	L'ensemble des sommes de carrés dans l'ensemble des poly- nômes positifs : aspects quantitatifs	61



Introduction

Le problème de la minimisation des polynômes réels fait l'objet de nombreuses techniques de résolutions : utilisation de bases de Gröbner, calcul de valeurs propres, de discriminants ou de résultants... Ces techniques se révèlent difficiles à mettre en oeuvre en pratique. Ceci est dû à la complexité des algorithmes calculant le minimum d'un polynôme réel (problème NP).

Ces dernières années se sont développées de nouvelles méthodes, dites de relaxation, calculant un minorant du minimum recherché. Parmi elles, les techniques de programmation semi-définie positive tiennent une bonne place, comme l'ont mis en avant P. Parrilo et B. Sturmfels dans leur article [\[PS\]](#).

Nous introduirons ici la programmation semi-définie positive, qui trouve une utilité dans de nombreux domaines et donnerons quelques exemples d'applications. Puis nous étudierons les sommes de carrés de polynômes réels, qui jouent un rôle primordial dans l'application de la programmation semi-définie à la minimisation de polynômes. Divers aspects seront étudiés, à la fois qualitatifs (savoir reconnaître une somme de carrés de polynômes) et quantitatifs (quel volume occupent les sommes de carrés parmi les polynômes positifs?). Enfin, la méthode de minimisation proposée par Parrilo et Sturmfels sera exposée. Un exemple d'étude sera donné, en utilisant un autre logiciel que celui utilisé par Parrilo et Sturmfels et basé sur une méthode différente, mais présentant les mêmes avantages.

Préliminaires matriciels

Dans ce chapitre, nous rappelons quelques propriétés des matrices symétriques (en particulier semi-définies positives), et effectuons quelques rappels de calcul matriciel. Ceci nous sera utile dans l'étude ultérieure de la programmation semi-définie positive.

Les propriétés non démontrées dans ce chapitre sont pour la plupart classiques, et relevant d'un cours d'algèbre linéaire ou bilinéaire de base. On pourra consulter avec profit les ouvrages de Horn et Johnson ([HJ1], [HJ2]) et le premier chapitre de [Hel]

1.1 Matrices symétriques ; matrices (semi-)définies positives

1.1.1 Produit scalaire ; normes

On définit tout d'abord un produit scalaire sur l'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées de taille $n \geq 1$, à coefficients dans \mathbb{R} :

Si $A, B \in \mathcal{M}_n(\mathbb{R})$, le produit scalaire de A et B , noté $A \bullet B$ est défini par :

$$A \bullet B = \text{Tr}({}^t AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij}$$

où la trace TrA d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est la somme des éléments diagonaux de A . On la propriété importante suivante :

$$\forall A, B \in \mathcal{M}_n(\mathbb{R}), Tr(AB) = Tr(BA).$$

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

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

On utilisera également la norme spectrale :

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

Intéressons-nous maintenant aux matrices symétriques, ie les matrices $A \in \mathcal{M}_n(\mathbb{R})$ vérifiant

$$A = {}^tA.$$

On note $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques réelles de taille n . Pour une matrice symétrique A , on obtient facilement les résultats suivants :

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

où $\lambda_1(A), \dots, \lambda_n(A)$ désignent les valeurs propres (toutes réelles) de A . La plus grande valeur propre d'une matrice symétrique A est caractérisée par :

Théorème 1.1.1 [Rayleigh-Ritz]

Soit $A \in \mathcal{S}_n(\mathbb{R})$. Alors

$$\lambda_{\max}(A) = \max_{\substack{v \in \mathbb{R}^n \\ \|v\|=1}} {}^tvAv$$

où $\|v\|$ désigne la norme euclidienne du vecteur v .

Corollaire 1.1.2

Soit $A \in \mathcal{S}_n(\mathbb{R})$. Alors

$$\lambda_{\max}(A) \geq \max_{i \in [1,n]} a_{ii}$$

1.1.2 Matrices (semi-)définies positives

Etudions maintenant les matrices (semi-)définies positives.

1.1 Matrices symétriques; matrices (semi-)définies positives

Définition 1.1.3

Soit $A \in \mathcal{S}_n(\mathbb{R})$.

○ A est dite semi-définie positive, et l'on notera $A \in \mathcal{S}_n^+(\mathbb{R})$ ou $A \succeq 0$ si

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

○ A est dite définie positive, et l'on notera $A \in \mathcal{S}_n^{++}(\mathbb{R})$ ou $A \succ 0$ si

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

Dressons une liste de propriétés de ces matrices :

Lemme 1.1.4

○ Toute sous-matrice principale d'une matrice définie positive (resp. semi-définie positive) est encore définie positive (resp. semi-définie positive).

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

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

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

Lemme 1.1.5

Soit $A \in \mathcal{S}_n^+(\mathbb{R})$. Alors

$$\exists i \in \llbracket 1, n \rrbracket, a_{ii} = \max_{j,k \in \llbracket 1, n \rrbracket} |a_{jk}|.$$

Preuve :

Soient $i, j \in \llbracket 1, n \rrbracket$.

On considère le vecteur x de \mathbb{R}^n défini par :

$$x_i = \begin{cases} 0 & \text{si } i \neq j \text{ et } i \neq k \\ 1 & \text{si } i = j \\ -1 & \text{si } i = k \end{cases}$$

L'inégalité ${}^t x A x \geq 0$ fournit alors :

$$2a_{jk} \leq a_{jj} + a_{kk}.$$

En considérant le vecteur $y \in \mathbb{R}^n$ défini par

$$y_i = \begin{cases} 0 & \text{si } i \neq j \text{ et } i \neq k \\ 1 & \text{si } i = j \\ 1 & \text{si } i = k \end{cases},$$

on montre de même que

$$-2a_{jk} \leq a_{jj} + a_{kk}.$$

On obtient alors :

$$|a_{jk}| \leq \frac{a_{jj} + a_{kk}}{2} \leq \max |a_{ii}|,$$

ce qui suffit pour conclure. ■

Le lemme suivant est utile en pratique, pour étudier quelques exemples en petite dimension :

Lemme 1.1.6

Soit $A \in \mathcal{S}_n^+(\mathbb{R})$. On suppose qu'il existe $i \in \llbracket 1, n \rrbracket$ tel que $a_{ii} = 0$. Alors

$$\forall j \in \llbracket 1, n \rrbracket, a_{ij} = 0.$$

Proposition 1.1.7

Soit $B \in \mathcal{M}_n(\mathbb{R})$ inversible. Alors

$$A \in \mathcal{S}_n^+(\mathbb{R}) \Leftrightarrow {}^tBAB \in \mathcal{S}_n^+(\mathbb{R})$$

et

$$A \in \mathcal{S}_n^{++}(\mathbb{R}) \Leftrightarrow {}^tBAB \in \mathcal{S}_n^{++}(\mathbb{R}).$$

On se référera aux ouvrages classiques pour diverses caractérisations des matrices (semi-)définies positives. On retiendra notamment les caractérisations suivantes :

Théorème 1.1.8

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

1. $A \in \mathcal{S}_n^+(\mathbb{R})$ (resp. $A \in \mathcal{S}_n^{++}(\mathbb{R})$)
2. $\lambda_{\min}(A) \geq 0$ (resp. $\lambda_{\min}(A) > 0$)
3. $\exists P \in \mathcal{M}_n(\mathbb{R}), A = {}^tPP$ (resp. $\exists P \in \mathcal{M}_n(\mathbb{R}), \text{rg}(P) = n$ tel que $A = {}^tPP$)

On a aussi, pour une matrice semi-définie positive, l'existence d'une racine carrée :

1.1 Matrices symétriques; matrices (semi-)définies positives

Proposition 1.1.9

Soit $A \in \mathcal{S}_n^+(\mathbb{R})$. Alors il existe une unique matrice $B \in \mathcal{S}_n^+(\mathbb{R})$ telle que $A = B^2$.

De plus, B commute avec A , et vérifie $\text{rg}(A) = \text{rg}(B)$.

On note souvent $B = A^{\frac{1}{2}}$.

Nous terminons cette section par le théorème du complément de Schur :

Théorème 1.1.10 [Complément de Schur]

Soient $A \in \mathcal{S}_m^{++}(\mathbb{R})$, $C \in \mathcal{S}_n(\mathbb{R})$ et $B \in \mathcal{M}_{m,n}(\mathbb{R})$. On a alors les équivalences suivantes :

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

et

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

1.1.3 Le cône $\mathcal{S}_n^+(\mathbb{R})$

Nous étudions dans cette section quelques propriétés spécifiques à $\mathcal{S}_n^+(\mathbb{R})$ et à sa structure conique.

Définition 1.1.11

Un ensemble $C \subset \mathbb{R}^m$ est un cône si

$$x, y \in C \Rightarrow \forall \lambda \geq 0, \lambda(x + y) \in C.$$

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

Remarque :

Par définition, un cône est convexe.

Proposition 1.1.12

$\mathcal{S}_n^+(\mathbb{R})$ est un cône fermé pointé dans $\mathcal{S}_n(\mathbb{R})$.

Remarque :

L'ensemble $\mathcal{S}_n^{++}(\mathbb{R})$ des matrices définies positives n'est pas un cône car $0 \notin \mathcal{S}_n^{++}(\mathbb{R})$.

Il est facile de voir que $\mathcal{S}_n^{++}(\mathbb{R})$ est l'intérieur de $\mathcal{S}_n^+(\mathbb{R})$.

Remarque :

La structure de cône de $\mathcal{S}_n^+(\mathbb{R})$ induit un ordre partiel sur $\mathcal{S}_n(\mathbb{R})$, dit ordre de Löwner :

$$A \succeq B \Leftrightarrow A - B \in \mathcal{S}_n^+(\mathbb{R}).$$

On notera de la même manière :

$$A \succ B \Leftrightarrow A - B \in \mathcal{S}_n^{++}(\mathbb{R}).$$

On a les propriétés suivantes ([Zha]) :

Proposition 1.1.13

Soient $A, B \in \mathcal{S}_n^+(\mathbb{R})$. Alors :

1. $A + B \succeq B$
2. $A^{\frac{1}{2}}BA^{\frac{1}{2}} \succeq 0$
3. $Tr(AB) \leq Tr(A)Tr(B)$
4. Les valeurs propres de AB sont positives.

Remarque :

Attention, le produit AB n'est pas forcément symétrique. Mais il est facile de voir que la condition sur AB d'être symétrique est équivalente à la condition de commutativité de A et B .

Le lemme suivant aura de nombreuses conséquences dans la suite de ce document :

Lemme 1.1.14

Soient $A, B \in \mathcal{S}_n^+(\mathbb{R})$. Alors $A \bullet B \geq 0$.

De plus, on a l'équivalence suivante :

$$A \bullet B = 0 \Leftrightarrow AB = 0.$$

Le lemme suivant donne un encadrement du produit scalaire de deux matrices symétriques définies positives.

Lemme 1.1.15

Soient $A, B \in \mathcal{S}_n^+(\mathbb{R})$. Alors :

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

Nous allons maintenant énoncer une propriété-clé du cône $\mathcal{S}_n^+(\mathbb{R})$, à la base de la théorie de la dualité en programmation semi-définie positive, que nous évoquerons par la suite.

Définition 1.1.16

Le cône polaire C^* d'un cône C est l'ensemble

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

1.2 Un peu de calcul matriciel

Lemme 1.1.17

$$(\mathcal{S}_n^+(\mathbb{R}))^* = \mathcal{S}_n^+(\mathbb{R}).$$

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

Corollaire 1.1.18

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

1.2 Un peu de calcul matriciel

Le but de cette section est d'introduire les notations facilitant la manipulation ultérieure des expressions matricielles.

Notation :

• Soit $A \in \mathcal{M}_{p,q}(\mathbb{R})$.

$\mathbf{vec}A \in \mathbb{R}^{pq}$ est le vecteur-colonne formé des colonnes de A les unes après les autres.

• Soit $v \in \mathbb{R}^{pq}$.

$Mat(v) \in \mathcal{M}_{p,q}(\mathbb{R})$ est la matrice dont la i^{e} colonne est formée des éléments $(i-1)p+1, \dots, ip$ de v .

Mat et \mathbf{vec} sont donc des applications réciproques l'une de l'autre.

Nous introduisons maintenant le produit de Kronecker de deux matrices :

Définition 1.2.1

Soient $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{k,l}(\mathbb{R})$.

Le produit de Kronecker de A par B est défini par :

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathcal{M}_{mk,nl}(\mathbb{R}).$$

Établissons une liste de quelques propriétés du produit de Kronecker, que nous utiliserons beaucoup par la suite :

Proposition 1.2.2

Soient A, B, C, D quatre matrices avec des tailles adaptées. Alors :

$$\begin{aligned} {}^t(A \otimes B) &= {}^tA \otimes {}^tB \\ (A \otimes B)(C \otimes D) &= (AC) \otimes (BD) \\ \mathbf{vec}(ABC) &= ({}^tC \otimes A) \mathbf{vec}(B) \end{aligned}$$

CHAPITRE 1 : *Préliminaires matriciels*

Programmation semi-définie positive

Dans ce chapitre, nous aborderons l'étude des problèmes de minimisation de fonctions linéaires en la variable $X \in \mathcal{M}_n(\mathbb{R})$ sur un domaine défini par des contraintes linéaires et de semi-définie positivité. On verra qu'il existe des méthodes efficaces pour résoudre ce genre de problèmes.

2.1 Formulation du problème

2.1.1 Problème primal standard

On considèrera le problème d'optimisation suivant :

Etant donnés deux entiers $n, m \geq 1$, un vecteur $\mathbf{b} \in \mathbb{R}^m$ et des matrices $A_i \in \mathcal{M}_n(\mathbb{R})$, $i \in \llbracket 1, m \rrbracket$ et $C \in \mathcal{M}_n(\mathbb{R})$, on cherche à

$$\begin{array}{ll}
 \text{(P)} & \text{minimiser} & C \bullet X \\
 & \text{sous les contraintes} & \forall i \in \llbracket 1, m \rrbracket, A_i \bullet X = \mathbf{b}_i \\
 & & X \in \mathcal{S}_n^+(\mathbb{R})
 \end{array}$$

où la variable est $X \in \mathcal{S}_n(\mathbb{R})$.

Remarque :

On peut supposer, sans perte de généralité, que les matrices C et A_i sont symétriques, ce que nous ferons dans la suite de ce chapitre.

En effet, puisque

$$\forall M \in \mathcal{M}_n(\mathbb{R}), \quad {}^t M \bullet X = M \bullet X,$$

il suffit pour cela de remplacer C par $\frac{1}{2} ({}^t C + C)$ (respectivement A_i par $\frac{1}{2} ({}^t A_i + A_i)$).

On peut de plus supposer, sans perte de généralité, que les matrices A_i sont linéairement indépendantes, ce que nous ferons également dans la suite de ce chapitre.

Définition 2.1.1

Avec les hypothèses précédentes, le problème (P) est appelé problème primal standard.

On l'écrira par la suite sous la forme concise suivante :

$$\begin{aligned} (P) \quad & \min \quad C \bullet X \\ & \text{contraintes} \quad \forall i \in \llbracket 1, m \rrbracket, A_i \bullet X = \mathbf{b}_i \\ & \quad \quad \quad X \in \mathcal{S}_n^+(\mathbb{R}) \end{aligned}$$

ou encore sous la forme compacte (équivalente) :

$$\begin{aligned} (P) \quad & \min \quad C \bullet X \\ & \text{contraintes} \quad \mathcal{A} \text{ vec} X = \mathbf{b} \\ & \quad \quad \quad X \in \mathcal{S}_n^+(\mathbb{R}) \end{aligned}$$

où $\mathcal{A} \in \mathcal{M}_{m,n^2}(\mathbb{R})$ a pour i^e ligne ${}^t \text{vec} A_i$.

Nous aurons besoin par la suite des définitions suivantes :

Définition 2.1.2

Une solution du problème primal (P) (ou solution primale) est une matrice $X \in \mathcal{S}_n^+(\mathbb{R})$ telle que $\mathcal{A} \text{ vec} X = \mathbf{b}$.

On note \mathcal{F}_P l'ensemble des solutions du problème primal (P) .

Une solution primale stricte (ou intérieure) est une solution primale X vérifiant de plus $X \in \mathcal{S}_n^{++}(\mathbb{R})$.

On note \mathcal{F}_P^+ l'ensemble des solutions primales intérieures.

Définition 2.1.3

La valeur optimale primale est définie par :

$$p^* = \inf_{\substack{\mathcal{A} \text{ vec} X = \mathbf{b} \\ X \in \mathcal{S}_n^+(\mathbb{R})}} C \bullet X$$

2.1 Formulation du problème

On dit que X^* est une solution primale optimale si $X^* \in \mathcal{F}_P$ et si $C \bullet X^* = p^*$.

2.1.2 Formulations non standard

Une autre formulation fréquente

Il existe plusieurs formulations différentes (mais équivalentes) du problème de programmation semi-définie positive.

Ainsi, on rencontre souvent dans la littérature le problème

$$(\mathbf{P}_2) \quad \min \quad {}^t \mathbf{b} \mathbf{y}$$

$$\text{contraintes} \quad \sum_{i=1}^m \mathbf{y}_i A_i \preceq C$$

où la variable est $\mathbf{y} \in \mathbb{R}^m$, et les matrices C et A_i jouissent des mêmes propriétés que dans la définition 2.1.1. La contrainte $\sum_{i=1}^m \mathbf{y}_i A_i \preceq C$ est appelée inégalité matricielle linéaire (LMI, pour *linear matrix inequality*).

Il s'agit en fait de ce qu'on appellera plus loin le problème dual associé au problème primal (\mathbf{P}) .

Montrons par exemple comment exprimer le problème (\mathbf{P}_2) sous la forme standard (\mathbf{P}) .

Introduisons pour cela une nouvelle variable Z :

$$(\mathbf{P}_2) \quad \min \quad {}^t \mathbf{b} \mathbf{y}$$

$$\text{contraintes} \quad \sum_{i=1}^m \mathbf{y}_i A_i + S = C$$

$$Z \succeq 0$$

Les variables sont désormais \mathbf{y} et S .

On considère l'ensemble $\mathcal{V} = \left\{ S \in \mathcal{S}_n(\mathbb{R}) / S = C - \sum_{i=1}^m \mathbf{y}_i A_i, \mathbf{y} \in \mathbb{R}^m \right\}$. Puisque

les A_i sont linéairement indépendantes, \mathcal{V} est un sous-espace affine de $\mathcal{S}_n(\mathbb{R})$ de dimension m . Par conséquent, il existe des matrices $G_i \in \mathcal{S}_n(\mathbb{R})$, $i \in \llbracket 1, \frac{n(n+1)}{2} - m \rrbracket$, et un vecteur $d \in \mathbb{R}^{\frac{n(n+1)}{2} - m}$ tels que

$$\mathcal{V} = \left\{ S \in \mathcal{S}_n(\mathbb{R}) / \forall i \in \llbracket 1, \frac{n(n+1)}{2} - m \rrbracket, G_i \bullet S = d_i \right\}.$$

En résolvant un système linéaire de m équations à $\frac{n(n+1)}{2}$ variables (sous-déterminé dans les cas qui nous occupent), on peut calculer une matrice G_0 vérifiant :

$$\forall i \in \llbracket 1, m \rrbracket, G_0 \bullet A_i = -\mathbf{b}_i.$$

Ainsi, pour $S = C - \sum_{i=1}^m \mathbf{y}_i A_i$, on a :

$${}^t \mathbf{b} \mathbf{y} = G_0 \bullet S - G_0 \bullet C.$$

Le problème (\mathbf{P}_2) s'exprime alors sous la forme :

$$\begin{aligned} (\mathbf{P}_2) \quad & \min \quad G_0 \bullet S \\ & \text{contraintes} \quad \forall i \in \llbracket 1, \frac{n(n+1)}{2} - m \rrbracket, G_i \bullet S = d_i \\ & \quad \quad \quad S \succeq 0, \end{aligned}$$

ce qui est bien la forme d'un problème primal standard.

Le problème primal à plusieurs inconnues

On peut également considérer le problème suivant, dont les inconnues sont les $X_j \in \mathcal{S}_{n_j}(\mathbb{R})$:

$$\begin{aligned} (\mathbf{P}_3) \quad & \min \quad \sum_{j=1}^L C_j \bullet X_j \\ & \text{contraintes} \quad \forall i \in \llbracket 1, m \rrbracket, \sum_{j=1}^L A_{i,j} \bullet X_j = \mathbf{b}_i \\ & \quad \quad \quad \forall j \in \llbracket 1, L \rrbracket, X_j \succeq 0 \end{aligned}$$

En fait, il ne s'agit pas d'une généralisation du problème (\mathbf{P}) , puisqu'il suffit, pour s'y ramener, de considérer les matrices diagonales par blocs :

$$\begin{aligned} X &= \text{diag}(X_1, \dots, X_L) \\ C &= \text{diag}(C_1, \dots, C_L) \\ A_i &= \text{diag}(A_{i,1}, \dots, A_{i,L}) \end{aligned}$$

Le problème primal à plusieurs inégalités matricielles linéaires (LMI)

De même, on peut considérer le problème

$$\begin{aligned} (\mathbf{P}_4) \quad & \min \quad {}^t \mathbf{b} \mathbf{y} \\ & \text{contraintes} \quad \forall j \in \llbracket 1, L \rrbracket, \sum_{i=1}^m \mathbf{y}_i A_{i,j} \preceq C_j \end{aligned}$$

qui est équivalent au problème (\mathbf{P}_2) , en définissant les matrices diagonales par blocs :

$$\begin{aligned} C &= \text{diag}(C_1, \dots, C_L) \\ A_i &= \text{diag}(A_{i,1}, \dots, A_{i,L}) \end{aligned}$$

2.2 Deux cas particuliers

2.2.1 Programmation linéaire

La programmation semi-définie positive est une généralisation de la programmation linéaire, dont le problème primal peut s'exprimer sous la forme standard suivante :

$$\begin{aligned}
 \text{(LP)} \quad & \min \quad {}^t c x \\
 & \text{contraintes} \quad \forall i \in \llbracket 1, m \rrbracket, \quad {}^t a_i x = b_i \\
 & \quad \quad \quad x \geq 0
 \end{aligned}$$

Il suffit, pour le voir, de considérer le problème (\mathbf{P}_3) avec tous les blocs de dimension 1, ie $\forall j, n_j = 1$.

Nous verrons par la suite qu'il existe des liens étroits entre programmation linéaire (LP) et programmation semi-définie positive (SDP). Par exemple, une théorie de dualité est également développée en SDP. De plus, certains algorithmes de résolution de problèmes LP peuvent être étendu au cas SDP.

2.2.2 Programmation quadratique

On s'intéresse maintenant au problème suivant (dit de programmation quadratique à contraintes quadratiques) :

$$\begin{aligned}
 \text{(QCQP)} \quad & \min \quad f_0(x) \\
 & \text{contraintes} \quad \forall i \in \llbracket 1, L \rrbracket, \quad f_i(x) \leq 0
 \end{aligned}$$

où chaque f_i est une fonction quadratique convexe :

$$f_i(x) = {}^t(A_i x + b)(A_i x + b) - {}^t c_i x - d_i$$

On peut alors exprimer, à l'aide du complément de Schur (1.1.10), le problème (QCQP) sous la forme du problème SDP suivant :

$$\begin{aligned}
 \text{(QCQP-SDP)} \quad & \min \quad t \\
 & \text{contraintes} \quad \begin{bmatrix} I & A_0 x + b_0 \\ {}^t(A_0 x + b_0) & {}^t c_0 x + d_0 + t \end{bmatrix} \succeq 0 \\
 & \quad \quad \quad \forall i \in \llbracket 1, L \rrbracket, \quad \begin{bmatrix} I & A_i x + b_i \\ {}^t(A_i x + b_i) & {}^t c_i x + d_i \end{bmatrix} \succeq 0
 \end{aligned}$$

Ainsi, la programmation SDP est une généralisation de la programmation QCQP. Cependant, des algorithmes spécifiques aux problèmes QCQP sont à préférer du point de vue algorithmique [Boy].

2.3 Dualité

A l'image de ce qui est fait en programmation linéaire, une théorie SDP de dualité peut être développée.

2.3.1 Formulation du problème dual

On rappelle tout d'abord l'énoncé du problème primal :

Etant donnés deux entiers $n, m \geq 1$, un vecteur $\mathbf{b} \in \mathbb{R}^m$, des matrices linéairement indépendantes $A_i \in \mathcal{S}_n(\mathbb{R})$, $i \in \llbracket 1, m \rrbracket$ et une matrice $C \in \mathcal{S}_n(\mathbb{R})$, le problème primal s'énonce ainsi :

$$\begin{aligned}
 (\mathbf{P}) \quad & \min && C \bullet X \\
 & \text{contraintes} && \forall i \in \llbracket 1, m \rrbracket, A_i \bullet X = \mathbf{b}_i \\
 & && X \in \mathcal{S}_n^+(\mathbb{R})
 \end{aligned}$$

où la variable est $X \in \mathcal{S}_n(\mathbb{R})$.

Le problème dual peut quant à lui être formulé ainsi :

$$\begin{aligned}
 (\mathbf{D}) \quad & \max && {}^t \mathbf{b} \mathbf{y} \\
 & \text{contraintes} && C - \sum_{i=1}^m y_i A_i \in \mathcal{S}_n^+(\mathbb{R})
 \end{aligned}$$

On peut également formuler le problème dual de la manière qui suit :

$$\begin{aligned}
 (\mathbf{D}) \quad & \max && {}^t \mathbf{b} \mathbf{y} \\
 & \text{contraintes} && C - \text{Mat}({}^t \mathbf{A} \mathbf{y}) \in \mathcal{S}_n^+(\mathbb{R})
 \end{aligned}$$

Remarque :

Il s'agit en fait du problème (\mathbf{P}_2) (page 11). On a déjà vu que ce problème pouvait se mettre sous la forme d'un problème SDP. Le problème dual est donc un problème de programmation semi-définie positive.

Exemple :

Dans le cas d'un problème de programmation linéaire, on retrouve le dual standard.

On a de manière évidente les définitions duales suivantes :

Définition 2.3.1

Une solution du problème dual (\mathbf{D}) (ou solution duale) est un couple $(\mathbf{y}, S) \in$

2.3 Dualité

$\mathbb{R}^m \times \mathcal{S}_n^+(\mathbb{R})$ telle que $S = C - \text{Mat}({}^t \mathbf{A} \mathbf{y}) \in \mathcal{S}_n^+(\mathbb{R})$.

On note \mathcal{F}_D l'ensemble des solutions du problème dual (D).

Une solution duale stricte (ou intérieure) est une solution duale (\mathbf{y}, S) vérifiant de plus $S = C - \text{Mat}({}^t \mathbf{A} \mathbf{y}) \in \mathcal{S}_n^{++}(\mathbb{R})$.

On note \mathcal{F}_D^+ l'ensemble des solutions duales intérieures.

Définition 2.3.2

La valeur optimale primale est définie par :

$$d^* = \sup \{ {}^t \mathbf{b} \mathbf{y} / C - \text{Mat}({}^t \mathbf{A} \mathbf{y}) \in \mathcal{S}_n^+(\mathbb{R}) \}$$

On dit que (\mathbf{y}^*, S^*) est une solution duale optimale si $(\mathbf{y}^*, S^*) \in \mathcal{F}_D$ et si ${}^t \mathbf{b} \mathbf{y}^* = d^*$.

On a le résultat suivant :

Proposition 2.3.3 (dualité faible)

Soient $X \in \mathcal{F}_P$ et $(\mathbf{y}, S) \in \mathcal{F}_D$. Alors :

$$C \bullet X - {}^t \mathbf{b} \mathbf{y} = X \bullet S \geq 0.$$

Preuve :

On a :

$$\begin{aligned} C \bullet X - {}^t \mathbf{b} \mathbf{y} &= \left(\sum y_i A_i + S \right) \bullet X - {}^t \mathbf{b} \mathbf{y} \\ &= \sum (A_i \bullet X) y_i + S \bullet X - {}^t \mathbf{b} \mathbf{y} \\ &= \sum \mathbf{b}_i y_i + S \bullet X - {}^t \mathbf{b} \mathbf{y} \\ &= S \bullet X = X \bullet S \end{aligned}$$

D'autre part, puisque $X \in \mathcal{S}_n^+(\mathbb{R})$, X admet une racine carrée $X^{\frac{1}{2}} \in \mathcal{S}_n^+(\mathbb{R})$, et on a alors :

$$\begin{aligned} X \bullet S &= \text{Tr}({}^t X S) \\ &= \text{Tr}(X S) \\ &= \text{Tr}\left(X^{\frac{1}{2}} X^{\frac{1}{2}} S\right) \\ &= \text{Tr}\left(X^{\frac{1}{2}} S X^{\frac{1}{2}}\right) \geq 0, \end{aligned}$$

la dernière inégalité ayant lieu puisque $X^{\frac{1}{2}} S X^{\frac{1}{2}} \in \mathcal{S}_n^+(\mathbb{R})$, chacun des facteurs étant semi-défini positif. ■

Définition 2.3.4

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

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

2.3.2 Dualité forte

Voyons tout d'abord que l'écart de dualité de solutions optimales n'est pas toujours nul; donnons un exemple où il est infini (d'autres exemples sont étudiés dans [Tod]) :

Exemple :

On considère le problème primal

$$\begin{aligned}
 (\mathbf{P}) \quad & \min && \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \bullet X \\
 & \text{contraintes} && \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \bullet X = 0 \\
 & && \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet X = 2 \\
 & && X \succeq 0
 \end{aligned}$$

Toute solution a pour première ligne 0 1, et une telle matrice ne peut pas être semi-définie positive (si une matrice est semi-définie positive et admet un élément diagonal (i, i) nul, alors toute la ligne i et toute la colonne i est nulle). Par conséquent, $\mathcal{F}_P = \emptyset$, et donc p^* est infini.

Le problème dual de (\mathbf{P}) est le suivant :

$$\begin{aligned}
 (\mathbf{D}) \quad & \max && 2\mathbf{y}_2 \\
 & \text{contraintes} && \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{y}_1 + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{y}_2 \preceq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}
 \end{aligned}$$

On peut réécrire la contrainte sous la forme

$$S = \begin{pmatrix} -\mathbf{y}_1 & -\mathbf{y}_2 \\ -\mathbf{y}_2 & 0 \end{pmatrix} \succeq 0$$

ce qui équivaut à $\begin{cases} \mathbf{y}_2 = 0 \\ \mathbf{y}_1 \leq 0 \end{cases}$

Par conséquent, $\mathbf{y} = (0; 0)$ est optimal, avec comme valeur optimale 0. L'écart de dualité est ainsi infini dans cet exemple.

Lemme de Farkas généralisé

On développe dans cette section les outils qui nous serviront pour l'étude de la dualité forte.

2.3 Dualité

On considère l'ensemble

$$K_1 = \mathcal{A}(\mathcal{S}_n^+(\mathbb{R})) = \{\mathcal{A} \text{ vec} X / X \in \mathcal{S}_n^+(\mathbb{R})\}.$$

Lemme 2.3.5

S'il existe $\mathbf{y} \in \mathbb{R}^m$ tel que $\text{Mat}({}^t \mathcal{A} \mathbf{y}) \succ 0$, alors K_1 est fermé.

Preuve :

On considère une suite $(X_i)_{i \in \mathbb{N}}$ de matrices semi-définies positives telle que $\mathcal{A} \text{ vec} X_i \rightarrow c$.

• La suite $(X_i)_{i \in \mathbb{N}}$ est bornée :

$$\mathbf{y} \bullet \mathcal{A} \text{ vec} X_i = {}^t \mathcal{A} \mathbf{y} \bullet \text{ vec} X_i = \text{Mat}({}^t \mathcal{A} \mathbf{y}) \bullet X_i \geq \lambda_{\min}(\text{Mat}({}^t \mathcal{A} \mathbf{y})) \lambda_{\max}(X_i),$$

la dernière inégalité provenant du lemme 1.1.15.

Mais par continuité du produit scalaire, on a :

$$\mathbf{y} \bullet \mathcal{A} \text{ vec} X_i \rightarrow \mathbf{y} \bullet c$$

et par conséquent, la suite $(\lambda_{\max}(X_i))_i$ est bornée. Les résultats 1.1.5 et 1.1.2 permettent alors de conclure.

• Puisque la suite $(X_i)_{i \in \mathbb{N}}$ est bornée, et puisque $\mathcal{S}_n^+(\mathbb{R})$ est un fermé, il existe une sous-suite $(X_{i_k})_{k \in \mathbb{N}}$ telle que $X_{i_k} \rightarrow \hat{X} \succeq 0$ et $\mathcal{A} \text{ vec} \hat{X} = c$, CQFD. ■

Enonçons maintenant le lemme de Farkas généralisé :

Lemme 2.3.6

Supposons qu'il existe $\mathbf{y}_0 \in \mathbb{R}^m$ tel que $\text{Mat}({}^t \mathcal{A} \mathbf{y}_0) \succ 0$. On a alors l'équivalence suivante :

$$(\exists X \succeq 0, \mathcal{A} \text{ vec} X = \mathbf{b}) \Leftrightarrow (\forall \mathbf{y}, \text{Mat}({}^t \mathcal{A} \mathbf{y}) \succeq 0 \Rightarrow {}^t \mathbf{b} \mathbf{y} \geq 0)$$

Preuve :

\Rightarrow : clair puisque ${}^t \mathbf{b} \mathbf{y} = {}^t (\mathcal{A} \text{ vec} X) \mathbf{y} = \text{Mat}({}^t \mathcal{A} \mathbf{y}) \bullet X \geq 0$

\Leftarrow : Supposons que le système $\mathcal{A} \text{ vec} X = \mathbf{b}$, $X \succeq 0$ n'ait pas de solution. Alors $\mathbf{b} \notin K_1$. D'après le lemme précédent, K_1 est un cône fermé, et donc, par un théorème de séparation (cf [Roc] th 11.7 p 100), il existe un vecteur \mathbf{y} tel que

$${}^t \mathbf{b} \mathbf{y} < 0 \text{ et } \forall X \succeq 0, {}^t (\mathcal{A} \text{ vec} X) \mathbf{y} \geq 0.$$

Mais cette dernière condition signifie

$$\forall X \succeq 0, X \bullet \text{Mat}({}^t \mathcal{A} \mathbf{y}) \geq 0,$$

ce qui est équivalent à

$$\text{Mat}({}^t\mathbf{A}\mathbf{y}) \succeq 0$$

(théorème de la trace de Fejer 1.1.18), et le lemme est démontré. ■

Dualité forte

Nous pouvons maintenant énoncer le théorème de dualité forte :

Théorème 2.3.7 [Dualité forte]

Supposons qu'il existe $\mathbf{y} \in \mathbb{R}^m$ tel que $\text{Mat}({}^t\mathbf{A}\mathbf{y}) \succ 0$.

Alors $p^* = d^*$.

Preuve :

- Si $p^* = -\infty$, alors, par dualité faible, $d^* = -\infty$.
- De même, si $d^* = +\infty$, alors par dualité faible $p^* = +\infty$.
- Si $p^* = +\infty$, alors $\mathcal{F}_P = \emptyset$, et le lemme de Farkas généralisé implique l'existence d'un vecteur \mathbf{y}_1 tel que

$${}^t\mathbf{b}\mathbf{y}_1 > 0 \text{ et } \text{Mat}({}^t\mathbf{A}\mathbf{y}_1) \preceq 0.$$

Mais cette dernière contrainte implique que $d^* = +\infty$. En effet, soit $\mathbf{y} \in \mathcal{F}_D$.

Alors

$$\forall \alpha \geq 0, \mathbf{y} + \alpha\mathbf{y}_1 \in \mathcal{F}_D$$

et

$${}^t\mathbf{b}(\mathbf{y} + \alpha\mathbf{y}_1) \xrightarrow{\alpha \rightarrow +\infty} +\infty$$

et par conséquent, $p^* = d^* = +\infty$.

- On peut donc supposer que p^* et d^* sont finis. Supposons que $d^* < p^*$. Alors le système

$$\begin{aligned} C \bullet X &= d^* \\ \mathbf{A} \text{vec} X &= \mathbf{b} \\ X &\succeq 0 \end{aligned}$$

n'admet pas de solution. Par conséquent, d'après le lemme de Farkas généralisé, il existe un scalaire y_0 et un vecteur \mathbf{y} tel que

$$y_0 C + \sum_{i=1}^m y_i A_i \succeq 0 \text{ et } d^* y_0 + {}^t\mathbf{b}\mathbf{y} < 0 \quad (*)$$

Trois cas se présentent :

- Si $y_0 = 0$, alors (*) est équivalent à

$$\text{Mat}({}^t\mathbf{A}\mathbf{y}) \succeq 0 \text{ et } {}^t\mathbf{b}\mathbf{y} < 0$$

2.3 Dualité

ce qui implique, d'après le lemme de Farkas généralisé, que le système

$$\begin{aligned} \mathcal{A} \text{vec} X &= \mathbf{b} \\ X &\succeq 0 \end{aligned}$$

n'admet pas de solution, et donc que $p^* = +\infty$, contradiction.

◦ Si $y_0 > 0$, alors, en divisant les deux membres de (*) par y_0 , on obtient :

$$C - \text{Mat} \left[{}^t \mathcal{A} \left(-\frac{\mathbf{y}}{y_0} \right) \right] \succeq 0 \text{ et } d^* - {}^t \mathbf{b} \left(-\frac{\mathbf{y}}{y_0} \right) < 0$$

ce qui signifie que d^* n'est pas la valeur optimale du problème dual, contradiction.

◦ Si $y_0 < 0$, alors, divisant les deux membres de (*) par $-y_0$, on obtient :

$$-C + \text{Mat} \left[{}^t \mathcal{A} \left(-\frac{\mathbf{y}}{y_0} \right) \right] \succeq 0 \text{ et } -d^* + {}^t \mathbf{b} \left(-\frac{\mathbf{y}}{y_0} \right) < 0$$

et on a même, pour un certain $\varepsilon > 0$:

$$-C + \text{Mat} \left[{}^t \mathcal{A} \left(-\frac{\mathbf{y}}{y_0} \right) \right] \succeq 0 \text{ et } -d^* + {}^t \mathbf{b} \left(-\frac{\mathbf{y}}{y_0} \right) < -\varepsilon \quad (**)$$

Mais par optimalité de d^* , il existe un \mathbf{y}^* tel que

$$C - \text{Mat} \left[{}^t \mathcal{A} \mathbf{y}^* \right] \succeq 0 \text{ et } d^* - {}^t \mathbf{b} \mathbf{y}^* < \varepsilon \quad (***)$$

En sommant (**) et (***), on obtient :

$$\text{Mat} \left[{}^t \mathcal{A} \left(-\frac{\mathbf{y}}{y_0} - \mathbf{y}^* \right) \right] \succeq 0 \text{ et } {}^t \mathbf{b} \left(-\frac{\mathbf{y}}{y_0} - \mathbf{y}^* \right) < 0$$

ce qui, toujours d'après le lemme de Farkas généralisé, implique que $\mathcal{F}_P = \emptyset$, et donc que $p^* = +\infty$, contradiction.

Ainsi, l'hypothèse $d^* < p^*$ conduit à une contradiction. Puisque d'après le lemme de dualité faible, on a $d^* \leq p^*$, on obtient donc finalement $d^* = p^*$, ce qu'il fallait démontrer. ■

D'autres conditions peuvent assurer l'égalité $p^* = d^*$. Par exemple, on peut citer :

1. $\mathcal{F}_P \neq \emptyset$ et $\mathcal{F}_D^+ \neq \emptyset$
2. $\mathcal{F}_P^+ \neq \emptyset$ et $\mathcal{F}_D \neq \emptyset$

On pourra consulter [Boy] ou [Hel].

Complémentarité

A l'image de ce qui se passe en programmation linéaire, on peut exprimer la condition pour X^* et \mathbf{y}^* d'être des solutions optimales de (\mathbf{P}) et (\mathbf{D}) sous la forme d'une condition dite de complémentarité :

Théorème 2.3.8

Soient $X^ \in \mathcal{F}_P$ et $(\mathbf{y}^*, S^*) \in \mathcal{F}_D$ d'écart de dualité nul ($C \bullet X^* - {}^t \mathbf{b} \mathbf{y}^* = 0$). Alors X^* et (\mathbf{y}^*, S^*) sont des valeurs optimales pour (\mathbf{P}) et (\mathbf{D}) respectivement **si et seulement si** $X^* S^* = 0$.*

Preuve :

Il suffit de se rappeler que l'écart de dualité vaut

$$X^* \bullet S^*$$

et que

$$X^* \bullet S^* = 0 \Leftrightarrow X^* S^* = 0$$

la dernière équivalence provenant du lemme 1.1.14 ■

2.4 Quelques applications

2.4.1 Problèmes aux valeurs propres

Il s'agit des plus anciens problèmes traités à l'aide de la programmation semi-définie positive.

Minimiser la plus grande valeur propre d'une matrice dépendant affinement d'un paramètre

On se donne k matrices symétriques $A_i \in \mathcal{S}_n(\mathbb{R})$.

On suppose que l'on étudie les matrices symétriques

$$A(x) = A_0 + x_1 A_1 + \cdots + x_k A_k,$$

où $x \in \mathbb{R}^k$. Plus précisément, notre problème est de minimiser la plus grande valeur propre de $A(x)$, pour x décrivant \mathbb{R}^k . Ce problème est alors équivalent au problème SDP suivant :

$$\lambda_{max}(A(x)) = \min \quad \lambda$$

contraintes $\lambda I - A(x) \succeq 0$

où les variables sont $x \in \mathbb{R}^k$ et $\lambda \in \mathbb{R}$. On verra plus loin un exemple d'application en théorie des graphes.

2.4 Quelques applications

Sommes des k premières valeurs propres d'une matrice

Soit $A \in \mathcal{S}_n(\mathbb{R})$ une matrice symétrique réelle. On considère ses valeurs propres $\lambda_1 \geq \dots \geq \lambda_n$. On cherche à exprimer $\lambda_1 + \dots + \lambda_k$ pour $k \in \llbracket 1, n \rrbracket$ sous la forme d'un problème SDP. On a le résultat suivant [Ali] :

Théorème 2.4.1

Sous les mêmes hypothèses, on a :

$$\begin{aligned} \lambda_1 + \dots + \lambda_k = \quad & \min && kz + \text{Tr}(V) \\ & \text{contraintes} && zI + V \succeq A \\ & && V \succeq 0 \end{aligned}$$

Quelques variantes

On peut également s'intéresser à des sommes pondérées de valeurs propres, ou aux sommes $|\mu_1| + \dots + |\mu_k|$ où les μ_i sont les valeurs propres de A , rangées de façon à ce que l'on ait :

$$|\mu_1| \geq \dots \geq |\mu_n|.$$

On pourra consulter pour cela [Ali].

Norme spectrale d'une matrice

Soit $A \in \mathcal{M}_n(\mathbb{R})$ (pas forcément symétrique). On considère sa norme spectrale $\max_{\lambda \in Sp({}^tMM)} \sqrt{\lambda}$. On peut la calculer à l'aide du problème SDP suivant :

$$\begin{aligned} \|A\|_2 = \quad & \min && t \\ & \text{contraintes} && \begin{bmatrix} tI & A \\ {}^tA & tI \end{bmatrix} \succeq 0 \end{aligned}$$

On utilise ici aussi le théorème du complément de Schur(1.1.10).

2.4.2 Problèmes géométriques

Séparation de points par une ellipsoïde

On se donne deux ensembles de points $\{x_1, \dots, x_k\}$ et $\{y_1, \dots, y_l\}$. On cherche à séparer ces deux ensembles de points par une ellipsoïde. Cela revient à trouver

$$f(x) = {}^txAx + {}^tbx + c$$

telle que

$$\begin{aligned} \forall i = 1, \dots, k, f(x_i) &\leq 0 \\ \forall j = 1, \dots, l, f(y_j) &\geq 0 \\ A &\succ 0 \end{aligned}$$

On peut même vouloir que cette ellipse soit la plus sphérique possible. Pour cela, A doit être proche de l'identité (le rapport des longueurs du grand axe et du petit axe vaut $\|A\|_2 \|A^{-1}\|_2$). Ainsi, pour résoudre ce problème, on peut considérer le problème SDP suivant :

$$\begin{aligned} (\mathcal{E}) \quad \min \quad & \gamma \\ \text{contraintes} \quad & \forall i = 1, \dots, k, f(x_i) \leq 0 \\ & \forall j = 1, \dots, l, f(y_j) \geq 0 \\ & I \preceq A \preceq \gamma I \end{aligned}$$

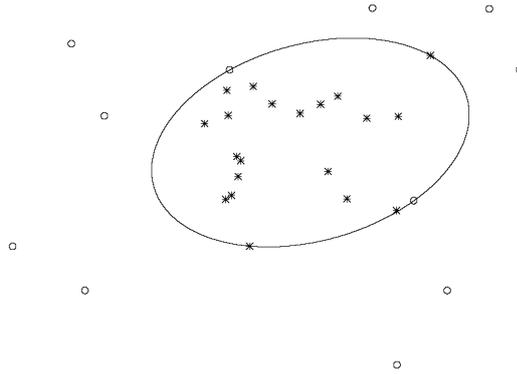


FIG. 2.1 – Un exemple d'ellipse la plus sphérique possible séparant deux ensembles de points.

Plus petite sphère contenant k ellipsoïdes

On se donne k ellipsoïdes $\mathcal{E}_1, \dots, \mathcal{E}_k$, sous la forme de leur équation $f_i(x) = {}^t x A_i x + 2 {}^t b_i x + c_i$, ie :

$$\mathcal{E}_i = \{x / f_i(x) \leq 0\}.$$

Une boule \mathcal{B} de centre x_c et de rayon $r = \sqrt{{}^t x_c x_c - \gamma}$, d'équation ${}^t x x - 2 {}^t x_c x + \gamma \leq 0$, contient les ellipsoïdes $\mathcal{E}_1, \dots, \mathcal{E}_k$, si et seulement si il existe $\tau_1, \dots, \tau_k \geq 0$ tels que :

$$\forall i = 1, \dots, k, \begin{bmatrix} I & -x_c \\ -{}^t x_c & \gamma \end{bmatrix} \preceq \tau_i \begin{bmatrix} A_i & b_i \\ b_i & c_i \end{bmatrix}$$

2.4 Quelques applications

On veut donc minimiser le rayon $r\sqrt{{}^t x_c x_c - \gamma}$. On va exprimer la condition $r^2 \leq t$ sous la forme matricielle :

$$\begin{bmatrix} I & x_c \\ {}^t x_c & t + \gamma \end{bmatrix} \succeq 0$$

et minimiser la variable t .

On obtient alors le problème SDP suivant :

$$\begin{aligned} (\mathcal{B}) \quad & \min \quad t \\ \text{contraintes} \quad & \forall i = 1, \dots, k, \begin{bmatrix} I & -x_c \\ -{}^t x_c & \gamma \end{bmatrix} \preceq \tau_i \begin{bmatrix} A_i & b_i \\ {}^t b_i & c_i \end{bmatrix} \\ & \forall i = 1, \dots, k, \tau_i \geq 0 \\ & \begin{bmatrix} I & x_c \\ {}^t x_c & t + \gamma \end{bmatrix} \succeq 0 \end{aligned}$$

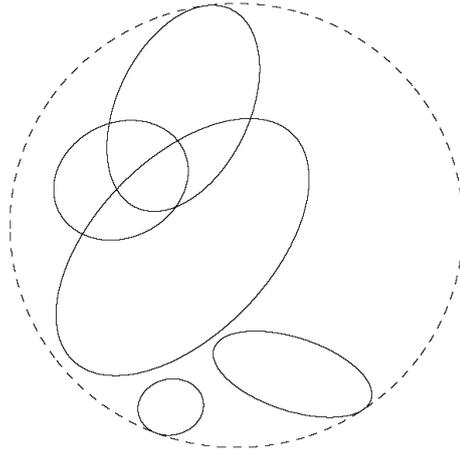


FIG. 2.2 – Plus petite boule contenant cinq ellipses.

2.4.3 Une application en statistiques

On se donne des réels m_0, \dots, m_{2n} . On se demande s'il existe une variable aléatoire X telle que chaque m_i soit le i^{e} moment de X . On peut montrer que la suite m_i , $0 \leq i \leq 2n$, est la suite des moments d'une variable aléatoire si et seulement si l'on a :

$$H(m_0, \dots, m_{2n}) = \begin{pmatrix} m_0 & \dots & m_n \\ \vdots & & \vdots \\ m_n & \dots & m_{2n} \end{pmatrix} \succeq 0$$

Supposons que l'on souhaite déterminer la variance maximale sur toutes les variables aléatoires, avec les contraintes sur les moments suivantes :

$$\forall i, l_i \leq m_i \leq u_i.$$

Ce problème peut se formuler sous la forme du problème SDP suivant :

$$V_{max}(X) = \begin{array}{ll} \max & y \\ \text{contraintes} & \forall i = 0, \dots, 2n, l_i \leq x_i \leq u_i \\ & H(x_0, \dots, x_{2n}) \succeq 0 \\ & \begin{pmatrix} x_2 - y & x_1 \\ x_1 & 1 \end{pmatrix} \succeq 0 \end{array}$$

2.4.4 Une application en théorie des graphes

Introduction à la théorie des graphes

Nous donnons ici les définitions de base en théorie des graphes.

Définition 2.4.2

Un graphe G (fini, non orienté, simple) est un couple $G = (V, E)$ où V est de la forme

$$V = \{1, \dots, n\}$$

et où E un ensemble fini d'éléments distincts $\{i, j\}$ où $i, j \in V, i \neq j$. V est appelé l'ensemble des sommets de G , et E désigne l'ensemble des arêtes de G .

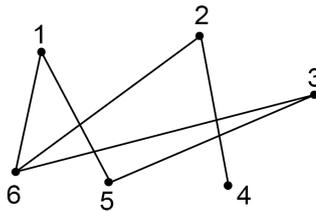


FIG. 2.3 – Un exemple de graphe.

Exemple :

Les graphes complets : un graphe G est dit complet sous toute paire de sommets est une arête.

2.4 Quelques applications

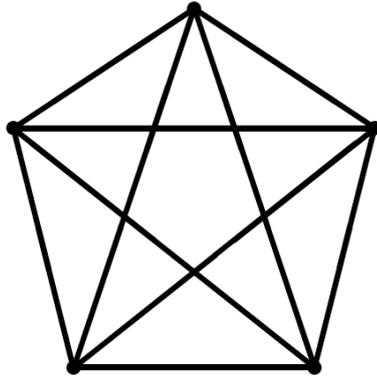


FIG. 2.4 – Graphe complet à 5 sommets.

Définition 2.4.3

Deux sommets (distincts) i, j de G sont dits adjacents si ce sont les extrémités d'une arête de G , ie $\{i, j\} \in E$

Une sommet $u \in V$ et une arête $e \in E$ sont dits incidents si l'arête e admet le sommet u comme extrémité, ie $u \in e$.

Définition 2.4.4

- $G' = (V', E')$ est un sous-graphe du graphe $G = (V, E)$ si

$$\left\{ \begin{array}{l} V' \subseteq V \\ E' \subseteq E \\ \text{chaque arête } e \in E' \text{ a les mêmes extrémités dans } G' \text{ et dans } G. \end{array} \right.$$

- Une clique de G est un sous-graphe complet de G .
- Le nombre de clique d'un graphe G est le nombre de sommets d'une clique maximale de G . On le note $\omega(G)$.

On s'intéresse maintenant à la coloration des graphes.

Définition 2.4.5

- Une k -coloration d'un graphe $G = (V, E)$ est une application $\gamma : V \rightarrow C$ de l'ensemble des sommets dans un ensemble C à k éléments (appelé ensemble des couleurs), telle que si deux sommets x et y sont adjacents, alors $\gamma(x) \neq \gamma(y)$.
- Le nombre chromatique de G est le nombre minimal de couleurs permettant de colorier G . On le note $\chi(G)$.

La fonction Θ de Lovasz

Soit $G = (V, E)$ un graphe à n sommets. On s'intéresse à la détermination du nombre de clique et du nombre chromatique de G . Ce problème est

en général NP-complet. Cependant, on peut parfois le résoudre de manière efficace, en introduisant la fonction Θ de Lovasz.

Définition 2.4.6

On considère l'ensemble S_G des matrices $A \in \mathcal{S}_n(\mathbb{R})$ telles que $A_{ij} = 1$ si $i = j$ ou si $\{i, j\} \in E$.

Etant donnée une telle matrice A , on note $\lambda_{\max}(A)$ sa plus grande valeur propre.

La fonction Θ de Lovasz est définie par :

$$\Theta(G) = \inf_{A \in S_G} \lambda_{\max}(A).$$

Proposition 2.4.7

On a la double inégalité suivante :

$$\omega(G) \leq \Theta(G) \leq \chi(G).$$

Preuve :

$\omega(G) \leq \Theta(G)$: Si $\omega(G) = k$, alors quitte à réordonner les sommets de G , toute matrice de S_G est de la forme :

$$\begin{pmatrix} 1 & \cdots & 1 & \\ \vdots & & \vdots & B \\ 1 & \cdots & 1 & \\ & C & & D \end{pmatrix}$$

Mais la plus grande valeur propre d'une sous-matrice principale d'une matrice symétrique est plus petite que la plus grande valeur propre de la matrice tout entière, et donc $k \leq \Theta(G)$.

$\Theta(G) \leq \chi(G)$: On a déjà étudié le problème de la minimisation de la plus grande valeur propre d'un ensemble de matrices. En appliquant ce qui avait été vu alors, on a la formule suivante, qui exprime la détermination de $\Theta(G)$ sous la forme d'un problème SDP :

$$\begin{aligned} \Theta(G) = \quad & \inf \quad \lambda \\ & \text{contraintes} \quad A \preceq \lambda I \\ & \quad \quad \quad A_{ij} = 1 \text{ si } i = j \text{ ou } \{i, j\} \in E \end{aligned}$$

ou encore :

$$\begin{aligned} \Theta(G) = \quad & \inf \quad \lambda \\ & \text{contraintes} \quad Y + J \preceq \lambda I \\ & \quad \quad \quad Y_{ij} = 0 \text{ si } i = j \text{ ou } \{i, j\} \in E \end{aligned}$$

2.5 Un algorithme de point intérieur

où $J \in \mathcal{M}_n(\mathbb{R})$ est la matrice dont tous les éléments valent 1.

Soit γ une k -coloration de G . On regroupe les sommets de G selon leur couleur, ce qui donne une partition C_1, \dots, C_k de V . On pose $c_i = |C_i|$ pour tout i .

Pour $i = 1, \dots, k$, on pose $M_i = k(J_{c_i} - I_{c_i})$, où J_{c_i} désigne la matrice carrée de taille c_i remplie de 1, et I_{c_i} la matrice identité de taille c_i .

Soit alors

$$Y = \begin{pmatrix} -M_1 & 0 & \dots & 0 \\ 0 & -M_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -M_k \end{pmatrix} \in \mathcal{S}_n(\mathbb{R})$$

Y vérifie bien les contraintes $Y_{ij} = 0$ si $i = j$ ou $\{i, j\} \in E$, et on a en outre

$$Y + J \preceq kI.$$

Par conséquent, $\Theta(G) \leq k$. ■

Ainsi, puisque l'on a exprimé notre problème sous la forme d'un problème SDP, la détermination de Θ peut se faire de manière efficace, alors que dans le cas général, celle de $\omega(G)$ ou de $\chi(G)$ est délicate.

Le cas le plus favorable est celui d'un graphe parfait, pour lequel $\omega(G) = \chi(G)$. Cependant, la difficulté réside dans la reconnaissance de tels graphes.

2.5 Un algorithme de point intérieur

Dans cette section, on étudiera un algorithme qui résoud, de manière efficace (on précisera le sens de cette efficacité par la suite), les problèmes de programmation semi-définie. Rappelons l'énoncé des problèmes primal et dual :

$$\begin{aligned} \text{(P)} \quad & \min && C \bullet X \\ & \text{contraintes} && \mathcal{A} \text{vec} X = \mathbf{b} \\ & && X \in \mathcal{S}_n^+(\mathbb{R}) \end{aligned}$$

$$\begin{aligned} \text{(D)} \quad & \max && {}^t \mathbf{b} \mathbf{y} \\ & \text{contraintes} && C - \text{Mat}({}^t \mathcal{A} \mathbf{y}) \in \mathcal{S}_n^+(\mathbb{R}) \end{aligned}$$

2.5.1 Le principe : l'introduction d'une fonction potentiel

On suppose dans toute cette partie qu'il existe une solution primale-duale intérieure (X_0, \mathbf{y}_0, S_0) , c'est-à-dire un triplet vérifiant :

$$\begin{cases} X_0 \in \mathcal{F}_P \cap \mathcal{S}_n^{++}(\mathbb{R}) \\ (\mathbf{y}_0, S_0) \in \mathcal{F}_D \\ S_0 \in \mathcal{S}_n^{++}(\mathbb{R}) \end{cases}$$

Rappelons que par définition, on a : $S_0 = C - \text{Mat}({}^t \mathbf{A} \mathbf{y}_0)$.

Dans cette partie, q désignera un réel strictement positif.

Définition 2.5.1

On définit la fonction-potentiel primale ϕ par :

$$\begin{aligned} \phi : \mathcal{S}_n(\mathbb{R}) \times \mathbb{R} &\rightarrow \mathbb{R} \\ (X, z) &\longmapsto q \ln(C \bullet X - z) - \ln \det X \end{aligned}$$

La fonction-potentiel primale duale ψ est définie par :

$$\begin{aligned} \psi : \mathcal{S}_n(\mathbb{R}) \times \mathcal{S}_n(\mathbb{R}) &\rightarrow \mathbb{R} \\ (X, S) &\longmapsto q \ln(X \bullet S) - \ln \det(XS) \end{aligned}$$

Stratégie de l'algorithme

La stratégie consiste à générer une suite (X_k, \mathbf{y}_k, S_k) de solutions primales-duales intérieures telles que $\psi(X_k, S_k)$ décroisse d'au moins une constante $\delta > 0$ à chaque itération. En choisissant alors convenablement q , on aboutira à la majoration de l'écart de dualité $C \bullet X_k - {}^t \mathbf{b} \mathbf{y}_k < cste \times \epsilon$, où ϵ est la précision recherchée. Plus précisément, le résultat fondamental est le suivant :

Théorème 2.5.2 *Soit $0 < \epsilon \ll 1$.*

Soit (X_0, \mathbf{y}_0, S_0) une solution primale-duale intérieure.

On pose $q = n + \lceil \sqrt{n} \rceil$ dans les fonctions-potentiels.

On suppose d'une part qu'il existe une constante E telle que

$$\psi(X_0, S_0) < E\sqrt{n}$$

et d'autre part qu'il existe une constante $\delta > 0$ et un algorithme générant une suite de solutions primales-duales intérieures (X_k, \mathbf{y}_k, S_k) telles que

$$\forall j \geq 0, \psi(X_{j+1}, S_{j+1}) \leq \psi(X_j, S_j) - \delta.$$

2.5 Un algorithme de point intérieur

On a alors :

$$\forall k \geq \frac{\sqrt{n} |\ln \varepsilon|}{\delta}, C \bullet X_k - {}^t \mathbf{b} \mathbf{y}_k < e^E \varepsilon$$

(e désignant la constante $\exp(1)$).

Preuve :

Soit $k \geq \frac{\sqrt{n} |\ln \varepsilon|}{\delta}$. On a :

$$\begin{aligned} \psi(X_k, S_k) &< \psi(X_0, S_0) - k\delta \\ &\leq \sqrt{n}E - \sqrt{n} |\ln \varepsilon| \\ &\leq \sqrt{n} (E - |\ln \varepsilon|) \\ &\leq \sqrt{n} (\ln e^E - |\ln \varepsilon|) \\ &= \sqrt{n} \ln (e^E \varepsilon) \end{aligned}$$

Par conséquent,

$$\begin{aligned} \sqrt{n} \ln (X_k \bullet S_k) &< -n \ln (X_k \bullet S_k) + \ln \det(X_k S_k) + \sqrt{n} \ln (e^E \varepsilon) \\ &\leq -n \ln n + \sqrt{n} \ln (e^E \varepsilon) \end{aligned}$$

La dernière inégalité est obtenue en vertu d'un lemme qui restera à démontrer.

On obtient dès lors facilement la majoration

$$\ln(X_k \bullet S_k) < \ln(e^E \varepsilon)$$

ce qui conduit, puisque $X_k \bullet S_k = C \bullet X_k - {}^t \mathbf{b} \mathbf{y}_k$, à la majoration annoncée. ■

Afin de terminer la démonstration précédente, il nous reste à prouver le lemme suivant :

Lemme 2.5.3

Soient $M, N \in \mathcal{S}_n^{++}(\mathbb{R})$. On a l'inégalité suivante :

$$\ln \det(MN) - n \ln(M \bullet N) \leq -n \ln n$$

Preuve :

On considère les valeurs propres μ_1, \dots, μ_n de MN . Ces valeurs propres sont toutes réelles et strictement positives : ce sont les mêmes que celles de $M^{\frac{1}{2}} N M^{\frac{1}{2}} = {}^t M^{\frac{1}{2}} N M^{\frac{1}{2}}$ (qui est symétrique réelle définie positive). En

effet, ces deux matrices ont le même polynôme caractéristique :

$$\begin{aligned}
 \chi_{MN}(X) &= \det(XI - MN) \\
 &= \det\left(M^{\frac{1}{2}}XIM^{-\frac{1}{2}} - M^{\frac{1}{2}}M^{\frac{1}{2}}NM^{\frac{1}{2}}M^{-\frac{1}{2}}\right) \\
 &= \det\left(M^{\frac{1}{2}}\left(XI - M^{\frac{1}{2}}NM^{\frac{1}{2}}\right)M^{-\frac{1}{2}}\right) \\
 &= \det\left(M^{\frac{1}{2}}\right)\chi_{M^{\frac{1}{2}}NM^{\frac{1}{2}}}(X)\det\left(M^{-\frac{1}{2}}\right) \\
 &= \chi_{M^{\frac{1}{2}}NM^{\frac{1}{2}}}(X)
 \end{aligned}$$

On peut alors écrire :

$$\begin{aligned}
 \ln \det(MN) - n \ln(M \bullet N) &= n \left(\ln \left[(\det MN)^{\frac{1}{n}} \right] - \ln(\text{Tr}(MN)) \right) \\
 &= n \left(\ln \left[(\mu_1 \dots \mu_n)^{\frac{1}{n}} \right] - \ln(\mu_1 + \dots + \mu_n) \right)
 \end{aligned}$$

L'inégalité arithmético-géométrique et la croissance du logarithme permettent alors de conclure. ■

Un nouveau problème primal-dual associé au problème original

On va maintenant associer au problème primal (**P**) un nouveau problème (**P'**) qui va recentrer la solution initiale (nous expliciterons ce point plus loin). On lui associera à son tour une fonction-potential primale qui permettra de contrôler celle de (**P**).

Soit $X_0 \in \mathcal{S}_n^{++}(\mathbb{R})$ une solution initiale intérieure de (**P**). Soit $L_0 \in \mathcal{M}_n(\mathbb{R})$ une matrice telle que

$$X_0 = L_0 {}^t L_0$$

Par exemple, on peut prendre pour L_0 un facteur dans une décomposition de Choleski de X_0 , on bien prendre pour L_0 sa racine carrée.

Soit $r \in \mathbb{N}$ fixé. On considère l'application suivante :

$$\begin{aligned}
 \mathcal{T} : \mathcal{S}_n(\mathbb{R}) &\rightarrow \mathcal{S}_n(\mathbb{R}) \times \mathbb{R}^r \\
 X &\mapsto \left(\frac{(n+r)L_0^{-1}X {}^t L_0^{-1}}{r + X_0^{-1} \bullet X}, \left(\frac{(n+r)}{r + X_0^{-1} \bullet X} \right) \mathbf{1}_r \right)
 \end{aligned}$$

\mathcal{T} admet une inverse à gauche \mathcal{S} :

$$\begin{aligned}
 \mathcal{S} : \mathcal{S}_n(\mathbb{R}) \times \mathbb{R}^r &\rightarrow \mathcal{S}_n(\mathbb{R}) \\
 (A, a) &\mapsto \frac{L_0 A {}^t L_0}{\sum_{j=1}^r a_j / r}
 \end{aligned}$$

2.5 Un algorithme de point intérieur

Un calcul évident permet effectivement de montrer que $\mathcal{S}(\mathcal{T}(X)) = X$ pour $X \in \mathcal{S}_n(\mathbb{R})$. On introduit maintenant notre nouveau problème de minimisation :

$$\begin{aligned}
 (\mathbf{P}') \quad & \min \quad \bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}} \\
 & \text{contraintes} \quad \bar{A} \text{vec} \bar{X} + \bar{A} \bar{\mathbf{x}} = 0 \\
 & \quad \text{Tr}(\bar{X}) + {}^t \mathbf{1}_r \bar{\mathbf{x}} = n + r \\
 & \quad \bar{X} \in \mathcal{S}_n^+(\mathbb{R}) \\
 & \quad \bar{\mathbf{x}} \in \mathbb{R}^r, \bar{\mathbf{x}} \geq 0
 \end{aligned}$$

où l'on a noté :

$$\begin{cases} \bar{C} = {}^t L_0 C L_0 \\ \bar{\mathbf{c}}(z) = -(z/r) \mathbf{1}_r \\ \bar{A} = \mathcal{A}(L_0 \otimes L_0) \\ \bar{A} = (-1/r) \mathbf{b} {}^t \mathbf{1}_r \end{cases}$$

On associe à (\mathbf{P}') une fonction-potential primale $\bar{\phi}$:

$$\bar{\phi}(\bar{X}, \bar{\mathbf{x}}, z) = q \ln [\bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}] - \ln \det \bar{X} - \sum_{j=1}^r \ln \bar{x}_j$$

On peut alors contrôler l'accroissement de ϕ grâce au lemme important suivant :

Lemme 2.5.4

Soit $(\bar{X}, \bar{\mathbf{x}}) \in \mathcal{S}_n(\mathbb{R}) \times \mathbb{R}^r$, tel que $\bar{x}_1 = \dots = \bar{x}_r$. Soit $X = \mathcal{S}(\bar{X}, \bar{\mathbf{x}})$.

On pose $q = n + r$.

Alors :

$$\phi(X, z) - \phi(X_0, z) = \bar{\phi}(\bar{X}, \bar{\mathbf{x}}, z) - \bar{\phi}(I, \mathbf{1}_r, z)$$

Preuve :

D'une part, on a :

$$\begin{aligned}
 \phi(X, z) &= q \ln(C \bullet X - z) - \ln \det X \\
 &= q \ln \left[\text{Tr} \left(\frac{{}^t L_0^{-1} \bar{C} L_0^{-1} L_0 \bar{X} {}^t L_0}{\sum x_j/r} \right) - z \right] - \ln \det \frac{L_0 \bar{X} {}^t L_0}{\sum x_j/r} \\
 &= q \ln \left[\frac{\bar{C} \bullet \bar{X}}{\sum x_j/r} - z \right] - \ln \left[\left(\frac{1}{\sum x_j/r} \right)^n \right] - \ln \det \bar{X} - \ln \det X_0 \\
 &= q \ln(\bar{C} \bullet \bar{X} - z \sum x_j/r) - q \left(\ln \sum x_j/r \right) + n \left(\ln \sum x_j/r \right) - \ln \det \bar{X} - \ln \det X_0 \\
 &= q \ln(\bar{C} \bullet \bar{X} - {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}) - r \ln \bar{x}_1 - \ln \det \bar{X} - \ln \det X_0 \\
 &= q \ln(\bar{C} \bullet \bar{X} - {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}) - \sum \ln x_j - \ln \det \bar{X} - \ln \det X_0
 \end{aligned}$$

D'autre part, on a :

$$\begin{aligned}
 \phi(X_0, z) &= q \ln(C \bullet X_0 - z) - \ln \det X_0 \\
 &= q \ln \left(\text{Tr} \left[{}^t L_0^{-1} \bar{C} L_0^{-1} L_0 {}^t L_0 \right] - z \right) \\
 &= q \ln \left(\text{Tr}(\bar{C}) - z \right) - \ln \det X_0 \\
 &= q \ln(\bar{C} \bullet I - z) - \ln \det X_0
 \end{aligned}$$

Ces deux calculs mènent immédiatement au résultat annoncé. ■

On va maintenant majorer $\bar{\phi}(\bar{X}, \bar{x}, z) - \bar{\phi}(I, \mathbf{1}_r, z)$. On utilisera pour cela le lemme suivant :

Lemme 2.5.5

Soit $M \in \mathcal{S}_n(\mathbb{R})$ vérifiant : $0 \prec M \prec I$. Alors :

$$\ln \det M \geq \text{Tr}(M) - n - \frac{\text{Tr} \left[(I - M)^2 \right]}{2[1 - \rho(I - M)]}$$

Preuve :

Soient μ_1, \dots, μ_n les valeurs propres de M (remarquer qu'elles sont toutes comprises entre 0 et 1).

On peut ainsi développer $\ln \mu_j$ en série entière, pour obtenir :

$$\begin{aligned}
 \ln \mu_j &= \ln [1 - (1 - \mu_j)] \\
 &= \mu_j - 1 - \sum_{k=2}^{+\infty} \frac{(1 - \mu_j)^k}{k} \\
 &\geq \mu_j - 1 - \sum_{k=2}^{+\infty} \frac{(1 - \mu_j)^k}{2} \\
 &= \mu_j - 1 - \frac{(1 - \mu_j)^2}{2[1 - (1 - \mu_j)]} \\
 &\geq \mu_j - 1 - \frac{(1 - \mu_j)^2}{2[1 - \rho(I - M)]}
 \end{aligned}$$

En sommant pour j allant de 1 à n , on obtient :

$$\underbrace{\sum_{j=1}^n \ln \mu_j}_{=\ln \det M} \geq \underbrace{\sum_{j=1}^n \mu_j}_{=\text{Tr}(M)} - n - \frac{\overbrace{\sum_{j=1}^n (1 - \mu_j)^2}^{=\text{Tr}[(I - M)^2]}}{2[1 - \rho(I - M)]}$$

ce qui est le résultat attendu. ■

2.5 Un algorithme de point intérieur

Corollaire 2.5.6

Soit $(\bar{X}, \bar{\mathbf{x}}) \in \mathcal{S}_n(\mathbb{R}) \times \mathbb{R}^r$ vérifiant les trois conditions suivantes :

$$\text{Tr}(\bar{X}) + {}^t \mathbf{1}_r \bar{\mathbf{x}} = n + r$$

$$0 \prec \bar{X} \prec I$$

$$\forall i, 0 < \bar{x}_i < 1$$

On pose encore $q = n + r$. On a alors :

$$\bar{\phi}(\bar{X}, \bar{\mathbf{x}}, z) - \bar{\phi}(I, \mathbf{1}_r, z) \leq (n+r) \ln \left(\frac{\bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}}{\text{Tr}(\bar{C}) + {}^t \bar{\mathbf{c}}(z) \mathbf{1}_r} \right) + \frac{\rho(\bar{X} - I)^2 + \|\bar{\mathbf{x}} - \mathbf{1}_r\|_\infty^2}{2[1 - \rho(I - \bar{X}) - \|\mathbf{1}_r - \bar{\mathbf{x}}\|_\infty]}$$

Preuve :

$$\begin{aligned} & \bar{\phi}(\bar{X}, \bar{\mathbf{x}}, z) - \bar{\phi}(I, \mathbf{1}_r, z) \\ &= q \ln [\bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}] - q \ln [\text{Tr}(\bar{C}) + {}^t \bar{\mathbf{c}}(z) \mathbf{1}_r] - \ln \det \bar{X} - \sum \ln \bar{x}_j \\ &= q \ln \left[\frac{\bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}}{\text{Tr}(\bar{C}) + {}^t \bar{\mathbf{c}}(z) \mathbf{1}_r} \right] - \ln \det \begin{pmatrix} \bar{X} & (0) \\ (0) & \begin{bmatrix} \bar{x}_1 & & \\ & \ddots & \\ & & \bar{x}_r \end{bmatrix} \end{pmatrix} \\ &\leq q \ln \left[\frac{\bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}}{\text{Tr}(\bar{C}) + {}^t \bar{\mathbf{c}}(z) \mathbf{1}_r} \right] - \text{Tr} \begin{pmatrix} \bar{X} & (0) \\ (0) & \begin{bmatrix} \bar{x}_1 & & \\ & \ddots & \\ & & \bar{x}_r \end{bmatrix} \end{pmatrix} + (n+r) + \frac{\text{Tr}((\bar{X} - I)^2) + \sum (\bar{x}_j - 1)^2}{2[1 - \max(\rho(I - \bar{X}), \|\mathbf{1}_r - \bar{\mathbf{x}}\|_\infty)]} \\ &= q \ln \left[\frac{\bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}}}{\text{Tr}(\bar{C}) + {}^t \bar{\mathbf{c}}(z) \mathbf{1}_r} \right] + \frac{\text{Tr}((\bar{X} - I)^2) + \sum (\bar{x}_j - 1)^2}{2[1 - \max(\rho(I - \bar{X}), \|\mathbf{1}_r - \bar{\mathbf{x}}\|_\infty)]}, \end{aligned}$$

la dernière égalité découlant de la première hypothèse.

Il ne reste qu'à remarquer que l'on a les inégalités suivantes pour conclure :

$$\begin{cases} \text{Tr}((I - \bar{X})^2) \leq \|I - \bar{X}\|^2 \\ \sum (\bar{x}_j - 1)^2 \leq \|\mathbf{1}_r - \bar{\mathbf{x}}\|^2 \\ \max(\rho(I - \bar{X}), \|\mathbf{1}_r - \bar{\mathbf{x}}\|) \leq \|I - \bar{X}\| + \|\mathbf{1}_r - \bar{\mathbf{x}}\| \end{cases} \quad \blacksquare$$

2.5.2 Un algorithme de réduction du potentiel

A l'image de ce que l'on fait classiquement en programmation linéaire, on va s'intéresser à la version modifiée de (\mathbf{P}') suivante :

$$\begin{aligned} (\mathbf{P}'') \quad & \min \quad \bar{C} \bullet \bar{X} + {}^t \bar{\mathbf{c}}(z) \bar{\mathbf{x}} \\ & \text{contraintes} \quad \bar{A} \text{vec} \bar{X} + \bar{A} \bar{\mathbf{x}} = 0 \\ & \quad \text{Tr}(\bar{X}) + {}^t \mathbf{1}_r \bar{\mathbf{x}} = n + r \\ & \quad \|\bar{X} - I\|^2 + \|\bar{\mathbf{x}} - \mathbf{1}_r\|^2 \leq \beta^2 < 1 \end{aligned}$$

Proposition 2.5.7

On peut reformuler (P') de la manière suivante :

$$\begin{aligned}
 (P')_{bis} \quad & \min \quad \begin{pmatrix} \text{vec}\bar{C} \\ \bar{c}(z) \end{pmatrix} \bullet \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} \\
 \text{contraintes} \quad & \mathcal{A}' \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_r \\ n+r \end{pmatrix} \\
 & \left\| \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} - \begin{pmatrix} \text{vec}I \\ \mathbf{1}_r \end{pmatrix} \right\| \leq \beta < 1
 \end{aligned}$$

où l'on a noté

$$\mathcal{A}' = \begin{pmatrix} \bar{A} & \bar{A} \\ {}^t \text{vec}I & {}^t \mathbf{1}_r \end{pmatrix}$$

Preuve :

Il suffit de remarquer successivement que :

$$\begin{aligned}
 \odot \begin{pmatrix} \text{vec}\bar{C} \\ \bar{c}(z) \end{pmatrix} \bullet \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} &= \begin{pmatrix} {}^t \text{vec}\bar{C} & {}^t \bar{c}(z) \end{pmatrix} \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} \\
 &= {}^t \text{vec}\bar{C} \text{vec}\bar{X} + {}^t \bar{c}(z) \bar{x} \\
 &= (\text{vec}\bar{C}) \bullet (\text{vec}\bar{X}) + {}^t \bar{c}(z) \bar{x} \\
 &= \bar{C} \bullet \bar{X} + {}^t \bar{c}(z) \bar{x}
 \end{aligned}$$

$$\begin{aligned}
 \odot \mathcal{A}' \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} &= \begin{pmatrix} \bar{A} & \bar{A} \\ {}^t \text{vec}I & {}^t \mathbf{1}_r \end{pmatrix} \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} \\
 &= \begin{pmatrix} \bar{A} \text{vec}\bar{X} + \bar{A}\bar{x} \\ {}^t \text{vec}I \text{vec}\bar{X} + {}^t \mathbf{1}_r \bar{x} \end{pmatrix} \\
 &= \begin{pmatrix} \bar{A} \text{vec}\bar{X} + \bar{A}\bar{x} \\ ({}^t \text{vec}I) \bullet (\text{vec}\bar{X}) + {}^t \mathbf{1}_r \bar{x} \end{pmatrix} \\
 &= \begin{pmatrix} \bar{A} \text{vec}\bar{X} + \bar{A}\bar{x} \\ I \bullet \bar{X} + {}^t \mathbf{1}_r \bar{x} \end{pmatrix} \\
 &= \begin{pmatrix} \bar{A} \text{vec}\bar{X} + \bar{A}\bar{x} \\ Tr(\bar{X}) + {}^t \mathbf{1}_r \bar{x} \end{pmatrix}
 \end{aligned}$$

$$\begin{aligned}
 \odot \left\| \begin{pmatrix} \text{vec}\bar{X} \\ \bar{x} \end{pmatrix} - \begin{pmatrix} \text{vec}I \\ \mathbf{1}_r \end{pmatrix} \right\|^2 &= \|\text{vec}\bar{X} - \text{vec}I\|^2 + \|\bar{x} - \mathbf{1}_r\|^2 \\
 &= \|\bar{X} - I\|^2 + \|\bar{x} - \mathbf{1}_r\|^2
 \end{aligned}$$

2.5 Un algorithme de point intérieur

■

Pour la suite de l'algorithme, une étude de la projection orthogonale sur $\ker \mathcal{A}'$ est nécessaire :

Lemme 2.5.8

Soit

$$P_{\mathcal{A}'} = \left(I - \frac{\begin{pmatrix} \text{vec} I \\ \mathbf{1}_r \end{pmatrix} \begin{pmatrix} {}^t \text{vec} I & {}^t \mathbf{1}_r \end{pmatrix}}{n+r} \right) \left(I - \begin{pmatrix} {}^t \bar{\mathcal{A}} \\ {}^t \bar{A} \end{pmatrix} \left[\begin{pmatrix} \bar{\mathcal{A}} & \bar{A} \end{pmatrix} \begin{pmatrix} {}^t \bar{\mathcal{A}} \\ {}^t \bar{A} \end{pmatrix} \right]^{-1} \begin{pmatrix} \bar{\mathcal{A}} & \bar{A} \end{pmatrix} \right)$$

Alors $P_{\mathcal{A}'}$ est la matrice de la projection orthogonale sur $\ker \mathcal{A}'$.

Preuve :

Remarquons que $P_{\mathcal{A}'}$ s'écrit :

$$P_{\mathcal{A}'} = \left[I - {}^t M (M {}^t M)^{-1} M \right] \left[I - {}^t N (N {}^t N)^{-1} N \right] = P_M * P_N$$

$$\text{avec } \begin{cases} M = \begin{pmatrix} {}^t \text{vec} I & {}^t \mathbf{1}_r \end{pmatrix} \\ N = \begin{pmatrix} \bar{\mathcal{A}} & \bar{A} \end{pmatrix} \end{cases}$$

⊙ P_M et P_N sont les matrices des projections orthogonales sur $\ker M$ et $\ker N$ respectivement.

Faisons la démonstration pour P_M :

un calcul simple montre que $P_M^2 = P_M = {}^t P_M$, ce qui montre que P_M est la matrice de la projection orthogonale sur $\text{Im}(P_M)$.

Il reste à montrer que $\text{Im}(P_M) = \ker M$.

D'une part, $M P_M = 0$, ce qui montre que $\text{Im}(P_M) \subset \ker M$.

D'autre part,

$$\begin{aligned} v \in \ker M &\Rightarrow Mv = 0 \\ &\Rightarrow P_M v = v - {}^t M (M {}^t M)^{-1} \underbrace{Mv}_{=0} = v \\ &\Rightarrow v \in \text{Im}(P_M) \end{aligned}$$

⊙ P_M et P_N commutent : il suffit de remarquer que l'on a :

$$N {}^t M = \bar{\mathcal{A}} \text{vec} I + \bar{A} \mathbf{1}_r = 0.$$

On peut alors facilement en conclure que

$$P_{\mathcal{A}'}^2 = P_{\mathcal{A}'} = {}^t P_{\mathcal{A}'}$$

et que

$$\text{Im}(P_{\mathcal{A}'}) = \ker \mathcal{A}' ,$$

ce qu'il fallait démontrer. ■

On a remarqué dans la démonstration précédente que $\begin{pmatrix} \text{vec} I \\ \mathbf{1}_r \end{pmatrix}$ était une solution particulière du système $\mathcal{A}' \begin{pmatrix} \text{vec} \bar{X} \\ \bar{\mathbf{x}} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_r \\ n+r \end{pmatrix}$. On peut dès lors facilement résoudre le problème (\mathbf{P}'') *bis* : il admet une unique solution $\begin{pmatrix} \text{vec} \bar{X}_1 \\ \bar{\mathbf{x}}_1 \end{pmatrix}$, donnée par :

$$\begin{pmatrix} \text{vec} \bar{X}_1 \\ \bar{\mathbf{x}}_1 \end{pmatrix} = \begin{pmatrix} \text{vec} I \\ \mathbf{1}_r \end{pmatrix} - \beta P_{\mathcal{A}'} \begin{pmatrix} \text{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix}$$

On pose alors :

$$\begin{cases} X(z) = \mathcal{S}(\bar{X}_1, \bar{\mathbf{x}}_1) \\ \mathbf{y}(z) = (N^t N)^{-1} N \begin{pmatrix} \text{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \\ S(z) = C - \text{Mat} [{}^t \mathcal{A} \mathbf{y}(z)] \end{cases}$$

où l'on a encore noté $N = (\bar{\mathcal{A}} \bar{A})$.

Etudions quelques propriétés de $X(z)$:

Proposition 2.5.9

$X(z)$ est une solution primale intérieure, c'est-à-dire :

$$\begin{cases} \mathcal{A} \text{vec} X(z) = \mathbf{b} \\ X(z) \in \mathcal{S}_n^{++}(\mathbb{R}) \end{cases}$$

Preuve :

Le fait que $X(z) \in \mathcal{S}_n^{++}(\mathbb{R})$ découle de la définition de $X(z)$ et du fait que l'on ait :

$$\begin{cases} \bar{X}_1 \in \mathcal{S}_n^{++}(\mathbb{R}) \\ \forall j, \bar{\mathbf{x}}_{1,j} > 0 \end{cases}$$

ces deux propriétés découlant immédiatement de la contrainte

$$\|\bar{X} - I\|^2 + \|\bar{\mathbf{x}} - \mathbf{1}_r\|^2 \leq \beta^2 < 1.$$

D'autre part,

$$\bar{A} \text{vec} \bar{X}_1 + \bar{A} \bar{\mathbf{x}}_1 = 0$$

2.5 Un algorithme de point intérieur

$$\begin{aligned}
&\Leftrightarrow \mathcal{A}(L_0 \otimes L_0) \mathbf{vec} \bar{X}_1 + \left(-\frac{1}{r}\right) \mathbf{b} {}^t \mathbf{1}_r \bar{\mathbf{x}}_1 = 0 \\
&\Leftrightarrow \mathcal{A} \mathbf{vec} \left(L_0 \bar{X}_1 {}^t L_0 \right) - \frac{1}{r} \mathbf{b} {}^t \mathbf{1}_r \bar{\mathbf{x}}_1 = 0 \\
&\Leftrightarrow \mathcal{A} \mathbf{vec} \left[\left(\sum \frac{\bar{x}_{1,j}}{r} \right) X(z) \right] = \frac{1}{r} \mathbf{b} {}^t \mathbf{1}_r \bar{\mathbf{x}}_1 \\
&\Leftrightarrow \left(\sum \frac{\bar{x}_{1,j}}{r} \right) \mathcal{A} \mathbf{vec} X(z) = \left(\sum \frac{\bar{x}_{1,j}}{r} \right) \mathbf{b} \\
&\Leftrightarrow \mathcal{A} \mathbf{vec} X(z) = \mathbf{b}, \text{ la dernière équivalence étant vraie car } \left(\sum \frac{\bar{x}_{1,j}}{r} \right) \neq 0, \\
&\text{puisque } \beta < 1. \blacksquare
\end{aligned}$$

On va maintenant étudier $S(z)$, pour laquelle la situation est moins claire. On introduit pour cela les quantités suivantes :

$$\Delta_0 = C \bullet X_0 - z$$

et

$$\Delta_1 = S(z) \bullet X_0 = C \bullet X_0 - {}^t \mathbf{b} \mathbf{y}(z).$$

Proposition 2.5.10

On pose $P(z) = P_{\mathcal{A}'} \left(\begin{array}{c} \mathbf{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{array} \right)$.

On suppose qu'il existe $\alpha \in]0, 1[$ tel que $\|P(z)\| \leq \alpha \frac{\Delta_0}{n+r}$.

Alors $S(z) \in \mathcal{S}_n^{++}(\mathbb{R})$ et ${}^t \mathbf{b} \mathbf{y}(z) > z$.

De plus, on a :

$$\left\| {}^t L_0 S(z) L_0 - \frac{\Delta_1}{n} I \right\| \leq \frac{\Delta_1}{n} \alpha \sqrt{\frac{n + n^2/r}{n + n^2/r - \alpha^2}} \quad (*)$$

et

$$\left| \frac{n+r}{n} \frac{\Delta_1}{\Delta_0} - 1 \right| \leq \frac{\alpha}{\sqrt{n + n^2/r}} \quad (**)$$

Preuve :

⊙ Remarquons tout d'abord que l'on peut exprimer $P(z)$ à l'aide de $X(z)$, $\mathbf{y}(z)$ et $S(z)$:

$$\begin{aligned}
P(z) &= \left(I - \frac{{}^t M M}{n+r} \right) \left(I - {}^t N (N {}^t N)^{-1} N \right) \left(\begin{array}{c} \mathbf{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{array} \right) \\
&= \left(I - {}^t N (N {}^t N)^{-1} - \frac{{}^t M M}{n+r} \right) \left(\begin{array}{c} \mathbf{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{array} \right) \text{ car } M {}^t N = 0 \\
&= \left(\begin{array}{c} \mathbf{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{array} \right) - {}^t N \mathbf{y}(z) - \frac{{}^t M M}{n+r} \left(\begin{array}{c} \mathbf{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{array} \right) \text{ par définition de } \mathbf{y}(z) \\
&= \left(\begin{array}{c} \mathbf{vec} \bar{C} \\ \bar{\mathbf{c}}(z) \end{array} \right) - \left(\begin{array}{c} {}^t \bar{A} \\ {}^t \bar{A} \end{array} \right) \mathbf{y}(z) - \frac{1}{n+r} \left(\begin{array}{c} \mathbf{vec} I \\ \mathbf{1}_r \end{array} \right) [{}^t \mathbf{vec} I \mathbf{vec} \bar{C} + {}^t \mathbf{1}_r \bar{\mathbf{c}}(z)]
\end{aligned}$$

Or :

$${}^t \text{vec} I \text{ vec} \bar{C} = I \bullet \bar{C} = \text{Tr}(\bar{C}) = \text{Tr}(L_0 C {}^t L_0) = \text{Tr}({}^t L_0 L_0 C) = \text{Tr}({}^t X_0 C) = C \bullet X_0$$

et

$${}^t \mathbf{1}_r \bar{\mathbf{c}}(z) = -\frac{z}{r} \underbrace{{}^t \mathbf{1}_r \mathbf{1}_r}_{=r} = -z$$

et on obtient donc :

$$P(z) = \begin{pmatrix} \text{vec} \bar{C} - {}^t \bar{\mathcal{A}} \mathbf{y}(z) \\ \bar{\mathbf{c}}(z) - {}^t \bar{\mathcal{A}} \mathbf{y}(z) \end{pmatrix} - \frac{C \bullet X_0 - z}{n+r} \begin{pmatrix} \text{vec} I \\ \mathbf{1}_r \end{pmatrix}$$

De plus,

$$\begin{aligned} \text{vec} \bar{C} - {}^t \bar{\mathcal{A}} \mathbf{y}(z) &= \text{vec} ({}^t L_0 C L_0) - {}^t [\mathcal{A}(L_0 \otimes L_0)] \mathbf{y}(z) \\ &= \text{vec} ({}^t L_0 C L_0) - ({}^t L_0 \otimes {}^t L_0) {}^t \mathcal{A} \mathbf{y}(z) \\ &= \text{vec} ({}^t L_0 C L_0) - ({}^t L_0 \otimes {}^t L_0) \text{vec} [\text{Mat} ({}^t \mathcal{A} \mathbf{y}(z))] \\ &= \text{vec} ({}^t L_0 C L_0) - \text{vec} ({}^t L_0 \text{Mat} ({}^t \mathcal{A} \mathbf{y}(z)) L_0) \\ &= \text{vec} ({}^t L_0 [C - \text{Mat} ({}^t \mathcal{A} \mathbf{y}(z))] L_0) \\ &= \text{vec} ({}^t L_0 S(z) L_0) \end{aligned}$$

et d'autre part

$$\bar{\mathbf{c}}(z) - {}^t \bar{\mathcal{A}} \mathbf{y}(z) = -\frac{z}{r} \mathbf{1}_r - \left(-\frac{1}{r}\right) \mathbf{1}_r {}^t \mathbf{b} \mathbf{y}(z) = \frac{{}^t \mathbf{b} \mathbf{y}(z) - z}{r} \mathbf{1}_r$$

et l'on obtient donc finalement :

$$P(z) = \begin{pmatrix} \text{vec} ({}^t L_0 S(z) L_0) \\ \frac{{}^t \mathbf{b} \mathbf{y}(z) - z}{r} \mathbf{1}_r \end{pmatrix} - \frac{C \bullet X_0 - z}{n+r} \begin{pmatrix} \text{vec} I \\ \mathbf{1}_r \end{pmatrix} \quad (***)$$

⊙ Supposons que $S(z) \notin \mathcal{S}_n^{++}(\mathbb{R})$. Alors ${}^t L_0 S(z) L_0$ n'est pas non plus définie positive, et admet donc une valeur propre $\mu \leq 0$.

Alors (***) implique :

$$\|P(z)\| \geq \rho \left(\frac{\Delta_0}{n+r} I - {}^t L_0 S(z) L_0 \right) \geq \frac{\Delta_0}{n+r} - \mu \geq \frac{\Delta_0}{n+r},$$

ce qui est en contradiction avec l'hypothèse.

De même, si ${}^t \mathbf{b} \mathbf{y}(z) \leq z$, (***) implique

$$\|P(z)\| \geq \frac{\Delta_0}{n+r} - \frac{{}^t \mathbf{b} \mathbf{y}(z) - z}{r} \geq \frac{\Delta_0}{n+r},$$

2.5 Un algorithme de point intérieur

ce qui est aussi une contradiction.

⊙ En outre, (***) implique :

$$P(z) = \begin{pmatrix} \text{vec} \left[{}^t L_0 S(z) L_0 - \frac{\Delta_1}{n} I \right] - \begin{bmatrix} \frac{\Delta_0}{n+r} - \frac{\Delta_1}{n} \\ \frac{\Delta_0 - \Delta_1}{r} - \frac{\Delta_0}{n+r} \end{bmatrix} \text{vec} I \\ \begin{bmatrix} \frac{\Delta_0 - \Delta_1}{r} - \frac{\Delta_0}{n+r} \\ \mathbf{1}_r \end{bmatrix} \end{pmatrix}$$

Mais comme

$$\begin{aligned} I \bullet \left[{}^t L_0 S(z) L_0 - \frac{\Delta_1}{n} I \right] &= I \bullet ({}^t L_0 S(z) L_0) - \frac{\Delta_1}{n} I \bullet I \\ &= \text{Tr} ({}^t L_0 S(z) L_0) - \frac{\Delta_1}{n} n \\ &= \text{Tr} (L_0 {}^t L_0 S(z)) - \Delta_1 \\ &= X_0 \bullet S(z) - \Delta_1 = 0, \end{aligned}$$

on a (Pythagore) :

$$\begin{aligned} \|P(z)\|^2 &= \left\| {}^t L_0 S(z) L_0 - \frac{\Delta_1}{n} I \right\|^2 + n \left(\frac{\Delta_0}{n+r} - \frac{\Delta_1}{n} \right)^2 + r \left(\frac{\Delta_0 - \Delta_1}{r} - \frac{\Delta_0}{n+r} \right)^2 \\ &= \left\| {}^t L_0 S(z) L_0 - \frac{\Delta_1}{n} I \right\|^2 + \left(n + \frac{n^2}{r} \right) \left(\frac{\Delta_1}{n} - \frac{\Delta_0}{n+r} \right)^2. \end{aligned}$$

Supposons que l'on n'ait pas (*). La dernière égalité implique alors :

$$\|P(z)\|^2 > \left(\frac{\Delta_1}{n} \right)^2 \alpha^2 \frac{n + n^2/r}{n + n^2/r - \alpha^2} + (n + n^2/r) \left(\frac{\Delta_1}{n} - \frac{\Delta_0}{n+r} \right)^2$$

En considérant cette dernière expression comme un trinôme en $\frac{\Delta_1}{n}$ et en la minimisant, on obtient la contradiction suivante :

$$\|P(z)\|^2 > \alpha^2 \left(\frac{\Delta_0}{n+r} \right)^2.$$

Par conséquent, on a bien montré (*).

Pour finir, la minimisation du trinôme en $\frac{\Delta_1}{n}$ implique directement que

$$(n + n^2/r) \left(\frac{\Delta_1}{n} - \frac{\Delta_0}{n+r} \right)^2 \leq \alpha^2 \left(\frac{\Delta_0}{n+r} \right)^2,$$

d'où l'on déduit facilement (**). ■

Il reste encore une étape avant de passer au théorème principal de cette section. Cela concerne la majoration de la quantité $\bar{\phi}(\bar{X}_1, \bar{\mathbf{x}}_1, z) - \bar{\phi}(I, \mathbf{1}_r, 1)$:

Lemme 2.5.11

On pose $q = n + r$. Soit $(\bar{X}_1, \bar{\mathbf{x}}_1)$ la solution du problème (P'') bis. Alors :

$$\bar{\phi}(\bar{X}_1, \bar{\mathbf{x}}_1, z) - \bar{\phi}(I, \mathbf{1}_r, 1) \leq -(n+r)\beta \frac{\|P(z)\|}{{}^t\bar{\mathbf{c}}(z)\mathbf{1}_r + \text{Tr}(\bar{C})} + \frac{\beta^2}{2(1-\beta)}$$

Preuve :

⊙ On rappelle que

$$\begin{pmatrix} \text{vec}\bar{X}_1 \\ \bar{\mathbf{x}}_1 \end{pmatrix} = \begin{pmatrix} \text{vec}I \\ \mathbf{1}_r \end{pmatrix} - \beta P_{\mathcal{A}'} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix}.$$

On obtient alors :

$$\begin{aligned} \bar{C} \bullet (\bar{X}_1 - I) + {}^t\bar{\mathbf{c}}(z)(\bar{\mathbf{x}}_1 - \mathbf{1}_r) &= \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \bullet \begin{pmatrix} \text{vec}\bar{X}_1 - \text{vec}I \\ \bar{\mathbf{x}}_1 - \mathbf{1}_r \end{pmatrix} \\ &= \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \bullet -\beta \begin{pmatrix} P(z) \\ \|P(z)\| \end{pmatrix} \\ &= -\frac{\beta}{\|P(z)\|} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \bullet P(z) \\ &= -\frac{\beta}{\|P(z)\|} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \bullet P_{\mathcal{A}'} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \\ &= -\frac{\beta}{\|P(z)\|} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \bullet {}^tP_{\mathcal{A}'}P_{\mathcal{A}'} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \quad \text{car } P_{\mathcal{A}'}^2 = P_{\mathcal{A}'} \\ &= -\frac{\beta}{\|P(z)\|} \left[P_{\mathcal{A}'} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \right] \bullet \left[P_{\mathcal{A}'} \begin{pmatrix} \text{vec}\bar{C} \\ \bar{\mathbf{c}}(z) \end{pmatrix} \right] \\ &= -\frac{\beta}{\|P(z)\|} \|P(z)\|^2 \\ &= -\beta \|P(z)\| \end{aligned}$$

⊙ Il suffit maintenant d'utiliser l'inégalité : $\forall x > 0, \ln(1+x) \leq x$ et d'appliquer le corollaire 2.5.6 pour conclure. ■

Conclusion

Rappelons l'énoncé des problèmes primal et dual de départ :

$$\begin{aligned} (\mathbf{P}) \quad \min \quad & C \bullet X \\ \text{contraintes} \quad & \mathcal{A} \text{vec}X = \mathbf{b} \\ & X \in \mathcal{S}_n^+(\mathbb{R}) \end{aligned}$$

2.5 Un algorithme de point intérieur

$$(D) \quad \max \quad {}^t \mathbf{b} \mathbf{y}$$

$$\text{contraintes} \quad C - \text{Mat}({}^t \mathbf{A} \mathbf{y}) \in \mathcal{S}_n^+(\mathbb{R})$$

Notre but était de réduire le potentiel (2.5.2). C'est ce que l'on fait dans le théorème suivant, où l'on montre l'existence d'un tel algorithme de réduction du potentiel :

Théorème 2.5.12

On considère une solution primale-duale intérieure (X_0, \mathbf{y}_0, S_0) .

On pose $r = \lceil \sqrt{n} \rceil$, $q = n + r$ et $z_0 = {}^t \mathbf{b} \mathbf{y}_0$.

On pose enfin $X_1 = X(z_0)$, $\mathbf{y}_1 = \mathbf{y}(z_0)$ et $S_1 = S(z_0)$.

Il existe alors une constante δ telle que l'on ait l'alternative suivante :

- soit $X_1 \in \mathcal{S}_n^{++}(\mathbb{R})$ ET $\psi(X_1, S_0) \leq \psi(X_0, S_0) - \delta$
- soit $S_1 \in \mathcal{S}_n^{++}(\mathbb{R})$ ET $\psi(X_0, S_1) \leq \psi(X_0, S_0) - \delta$.

Preuve :

S'il existe une constante $0 < \alpha < 1$ telle que

$$\|P(z)\| \geq \alpha \frac{\Delta_0}{n+r}$$

alors :

$$\begin{aligned} \psi(X_1, S_0) - \psi(X_0, S_0) &= \phi(X_1, z_0) - \phi(X_0, z_0) \\ &= \bar{\phi}(\bar{X}_1(z_0), \bar{\mathbf{x}}_1(z_0), z_0) - \bar{\phi}(I, \mathbf{1}_r, z_0) \\ &\leq -\beta\alpha + \frac{\beta^2}{2(1-\beta)}, \end{aligned}$$

la dernière inégalité étant obtenue grâce au lemme 2.5.11.

Sinon, l'hypothèse de la proposition 2.5.10 est satisfaite. En appliquant également le lemme 2.5.5, on obtient :

$$\begin{aligned} n \ln X_0 \bullet S_1 - \ln \det X_0 S_1 &= n \ln \left(\frac{n X_0 \bullet S_1}{\Delta_1} \right) - \ln \det \frac{n X_0 S_1}{\Delta_1} \\ &= n \ln n - \ln \det \frac{n X_0 S_1}{\Delta_1} \quad \text{car } \Delta_1 = X_0 \bullet S_1 \\ &\leq n \ln n + \frac{\|n {}^t L_0 S_1 L_0 / \Delta_1 - I\|^2}{2(1 - \|n {}^t L_0 S_1 L_0 / \Delta_1 - I\|)} \\ &\leq n \ln n + \frac{\gamma^2}{2(1-\gamma)} \quad \text{d'après le (*) de la prop 2.5.10} \\ &\leq n \ln X_0 \bullet S_0 - \ln \det X_0 S_0 + \frac{\gamma^2}{2(1-\gamma)} \quad \text{en vertu du lemme 2.5.3} \end{aligned}$$

D'autre part, le (**) de la proposition 2.5.10 implique :

$$\Delta_1 < \left(1 - \frac{r}{n+r} - \frac{n}{n+r} \frac{\alpha}{\sqrt{n+n^2/r}} \right) \Delta_0$$

d'où

$$r \ln \frac{X_0 \bullet S_1}{X_0 \bullet S_0} = r \ln \frac{\Delta_1}{\Delta_0} \leq \frac{r^2}{n+r} \left(-1 + \frac{\alpha}{\sqrt{r+r^2/n}} \right).$$

On en déduit finalement :

$$\psi(X_0, S_1) - \psi(X_0, S_0) \leq \frac{r^2}{n+r} \left(-1 + \frac{\alpha}{\sqrt{r+r^2/n}} \right) + \frac{\gamma^2}{2(1-\gamma)}.$$

En posant $\delta = \min \left(-\beta\alpha + \frac{\beta^2}{2(1-\beta)}, \frac{r^2}{n+r} \left(-1 + \frac{\alpha}{\sqrt{r+r^2/n}} \right) + \frac{\gamma^2}{2(1-\gamma)} \right)$, on obtient la réduction du potentiel attendue. ■

Remarque :

Posons $\alpha = 0.55$ et $\beta = 0.3$. Alors

$$-\beta\alpha + \frac{\beta^2}{2(1-\beta)} = -0.1007142857.$$

De plus, on étudie la suite $u_n = \frac{r^2}{n+r} \left(-1 + \frac{\alpha}{\sqrt{r+r^2/n}} \right) + \frac{\gamma^2}{2(1-\gamma)}$ où $r = \lceil \sqrt{n} \rceil$, on peut montrer, par une méthode technique mais élémentaire, que

$$\forall n \geq 2, u_n \leq u_4 = -0.1046678683.$$

Pour $n \geq 2$ et pour ce choix de α, β , on a donc $\delta > 0.1$.

On en déduit donc l'algorithme suivant :

Rappel : \mathcal{F}_P désigne l'ensemble des solutions du problème primal (**P**), et \mathcal{F}_D l'ensemble des solutions du problème dual (**D**). Rappelons à nouveau l'énoncé de ces deux problèmes :

$$\begin{aligned} (\mathbf{P}) \quad & \min && C \bullet X \\ & \text{contraintes} && \mathcal{A} \text{vec} X = \mathbf{b} \\ & && X \in \mathcal{S}_n^+(\mathbb{R}) \end{aligned}$$

$$\begin{aligned} (\mathbf{D}) \quad & \max && {}^t \mathbf{b} \mathbf{y} \\ & \text{contraintes} && C - \text{Mat}({}^t \mathcal{A} \mathbf{y}) \in \mathcal{S}_n^+(\mathbb{R}) \end{aligned}$$

2.5 Un algorithme de point intérieur

<p>Entrée :</p> <ul style="list-style-type: none"> Une matrice $X_0 \in \mathcal{F}_P \cap \mathcal{S}_n^{++}(\mathbb{R})$ Un vecteur $\mathbf{y}_0 \in \mathcal{F}_D$ intérieur (ie $S_0 := C - \text{Mat}({}^t \mathcal{A} \mathbf{y}_0) \in \mathcal{S}_n^{++}(\mathbb{R})$) Une constante $\varepsilon > 0$. <p>Sortie :</p> <ul style="list-style-type: none"> Une solution primale-duale $X \in \mathcal{F}_P, \mathbf{y} \in \mathcal{F}_D$ telle que $C \bullet X - {}^t \mathbf{b} \mathbf{y} < \varepsilon$. <p>Procédure :</p> <ol style="list-style-type: none"> 1) $\alpha := 0.55, \beta := 0.3$ (pour avoir $d > 0.1$, cf remarque précédente) 2) $X := X_0, S := C - \text{Mat}({}^t \mathcal{A} \mathbf{y}_0), \mathbf{y} := \mathbf{y}_0$ et $z := {}^t \mathbf{b} \mathbf{y}_0, r = \lceil \sqrt{n} \rceil$ 3) Tant que $C \bullet X - {}^t \mathbf{b} \mathbf{y} \geq \varepsilon$ faire <ul style="list-style-type: none"> Calculer $\mathbf{y}(z), S(z)$ et $P(z)$ (formules pages 36 et 38) Si $\ P(z)\ \geq \alpha (C \bullet X - z) / (n + r)$ Alors <ul style="list-style-type: none"> $(\bar{X}, \bar{\mathbf{x}}) := (I, \mathbf{1}_r) - \beta P(z)$ $X := \mathcal{S}(\bar{X}, \bar{\mathbf{x}})$ Sinon <ul style="list-style-type: none"> $S := S(z)$ $\mathbf{y} := \mathbf{y}(z)$ $z := {}^t \mathbf{b} \mathbf{y}$ Fin Si <p>Fait</p>

Algorithme 1: Algorithme naïf de réduction de potentiel

Remarque :

Il s'agit ici d'un algorithme naïf, que l'on peut bien sûr améliorer dans la pratique. On pourra par exemple consulter [Ali] pour un choix de α et β plus efficace, et quelques autres pistes d'amélioration.

Correction et complexité de l'algorithme

La correction de l'algorithme précédent découle directement du théorème 2.5.12.

Quant à sa complexité, elle fait l'objet du théorème 2.5.2, qui donne un majorant du nombre d'itérations nécessaires.

CHAPITRE 2 : *Programmation semi-définie positive*

Polynômes positifs, sommes de carrés ; minimisation

3.1 Reconnaître les sommes de carrés : un algorithme

Dans cette partie, on montre le lien entre l'écriture d'un polynôme comme somme de carrés de polynômes et l'écriture à l'aide de matrices semi-définies positives. On en déduira un algorithme pour décider si un polynôme est une somme de carrés ou non. On consultera l'article [PW] qui donne une méthode pour donner, le cas échéant, une représentation.

3.1.1 L'espace \mathcal{L}_f

Dans cette section, nous serons amenés à étudier les polynômes en n variables. Quelques notations sont nécessaires à la lisibilité de l'ensemble.

Notation :

$\mathbb{R}[X]$ désignera l'anneau $\mathbb{R}[x_1, \dots, x_n]$, et on notera $\sum \mathbb{R}[X]^2$ les sommes de carrés de polynômes.

On étudiera particulièrement les monômes de $\mathbb{R}[X]$, ce qui nous amène à la notation et à la définition suivantes :

Notation :

Pour $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, x^α désignera le monôme $x_1^{\alpha_1} \dots x_n^{\alpha_n}$

Définition 3.1.1

Pour $d \in \mathbb{N}$, on pose

$$\Lambda_d := \{(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n / \alpha_1 + \dots + \alpha_n \leq d\}$$

On a alors

$$|\Lambda_d| = C_{n+d}^n := N$$

et on notera par conséquent

$$\Lambda_d = \{\beta_1, \dots, \beta_N\}.$$

On en vient maintenant à l'objet de cette partie. Dans ce qui suit, on se fixe un polynôme $f \in \mathbb{R}[X]$, et on pose $m = (x^{\beta_1}, \dots, x^{\beta_N})$.

Définition 3.1.2

On note \mathcal{L}_f l'ensemble des matrices symétriques réelles B de taille N telles que

$$f = mB^t m$$

Afin d'étudier \mathcal{L}_f , on aura besoin du lemme suivant :

Lemme 3.1.3

On munit les monômes de $\mathbb{R}[X]$ de l'ordre (total) du degré lexicographique $<_{deglex}$. Soient

$$A = \left\{ (\alpha, \beta) \in \Lambda_d \times \Lambda_d / x^\beta \leq_{deglex} x^\alpha \right\}$$

et

$$\pi : \begin{array}{ccc} A & \rightarrow & \Lambda_{2d} \\ (\alpha, \beta) & \mapsto & \alpha + \beta \end{array}.$$

Alors π est surjective, et par conséquent :

$$\frac{1}{2} C_{n+d}^d (C_{n+d}^d + 1) \geq C_{n+2d}^{2d}$$

Démonstration :

- Montrer la surjectivité de π revient à montrer que tout monôme x^γ (avec $\gamma \in \Lambda_{2d}$) peut s'écrire comme produit de deux monômes x^α, x^β avec $\alpha, \beta \in \Lambda_d$ et $\beta <_{deglex} \alpha$ ou $\alpha = \beta$. Il suffit alors de prendre pour α l'exposant du plus grand monôme (pour $<_{deglex}$) de degré total $\leq d$ qui divise x^γ , et de poser alors $\beta = \gamma - \alpha$.
- Puisque π est surjective, on a :

$$|A| \geq |\Lambda_{2d}|$$

ce qui entraîne l'inégalité annoncée dans le lemme. ■

3.1 Reconnaître les sommes de carrés : un algorithme

On peut maintenant démontrer la proposition suivante :

Proposition 3.1.4

\mathcal{L}_f est un sous-espace affine de $\mathcal{S}_N(\mathbb{R})$, de dimension

$$\frac{1}{2}C_{n+d}^d \left(C_{n+d}^d + 1 \right) - C_{n+2d}^{2d}.$$

Démonstration :

On identifie $\mathcal{S}_N(\mathbb{R})$ à $\mathbb{R}^{\frac{N(N+1)}{2}}$.

L'identité

$$f = mB^t m$$

peut se réécrire sous la forme du système linéaire

$$(S) \quad \sum_{\beta_i + \beta_j = \alpha} b_{ij} = a_\alpha$$

Ce système (S) est équivalent au système (S') où les variables b_{ij} avec $j > i$ sont éliminées en utilisant la symétrie de B : $b_{ij} = b_{ji}$.

On obtient alors un système équivalent (S') en les variables $(b_{ij})_{1 \leq j \leq i \leq N}$, et dont la matrice associée A est de taille $\left(|\Lambda_{2d}|, \frac{N(N+1)}{2} \right)$.

De ce fait, \mathcal{L}_f est un sous-espace affine de $\mathcal{S}_N(\mathbb{R})$, de dimension

$$\frac{N(N+1)}{2} - \text{rg}(A)$$

A est une matrice qui ne contient aucune ligne nulle (cf lemme précédent), et telle que chaque colonne contienne exactement un élément non nul (égal à 1 ou à 2) : chaque variable b_{ij} n'apparaît que dans une seule équation.

Soit $\mathcal{B} = (e_k)_{1 \leq k \leq \frac{N(N+1)}{2}}$ la base canonique de $\mathbb{R}^{\frac{N(N+1)}{2}}$.

A multiplication par un scalaire non nul près, chaque e_k apparaît dans les colonnes de A .

Par conséquent, après permutation des lignes et des colonnes, on peut se ramener au cas où A est une matrice du type

$$\left(\begin{array}{cccc|ccc} \lambda_1 & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & \lambda_{|\Lambda_{2d}|} & * & \cdots & * \end{array} \right)$$

où tous les λ_i sont non nuls.

Par conséquent, le rang de A est maximal, égal à $|\Lambda_{2d}| = C_{n+2d}^{2d}$, CQFD. ■

Ainsi, pour déterminer \mathcal{L}_f , il suffit de résoudre un système linéaire. On a implémenté ceci en Maple, dans la procédure Lf .

Exemple : On a testé notre procédure sur l'exemple donné dans l'article de Parrilo et Sturmfels (on a pris $m = (1, x, y, z, x^2, xy, y^2, xz, yz, z^2)$ comme dans leur article) :

> $P := x^4 + y^4 + z^4 - 4xyz + x + y + z - \lambda;$
 $P := x^4 + y^4 + z^4 - 4xyz + x + y + z - \lambda$
 > $\text{Lf}(P, m);$

l'espace affine est de dimension :

$$\begin{pmatrix} -\lambda & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & b_{1,5} & b_{1,6} & b_{1,7} & b_{1,8} & b_{1,9} & b_{1,10} \\ \frac{1}{2} & -2b_{1,5} & -b_{1,6} & -b_{1,8} & 0 & -b_{3,5} & -b_{3,6} & b_{2,8} & b_{2,9} & -b_{4,8} \\ \frac{1}{2} & -b_{1,6} & -2b_{1,7} & -b_{1,9} & b_{3,5} & b_{3,6} & 0 & b_{3,8} & -b_{4,7} & b_{3,10} \\ \frac{1}{2} & -b_{1,8} & -b_{1,9} & -2b_{1,10} & -b_{2,8} & -b_{2,9} - b_{3,8} - 2 & b_{4,7} & b_{4,8} & -b_{3,10} & 0 \\ b_{1,5} & 0 & b_{3,5} & -b_{2,8} & 1 & 0 & b_{5,7} & 0 & b_{5,9} & b_{5,10} \\ b_{1,6} & -b_{3,5} & b_{3,6} & -b_{2,9} - b_{3,8} - 2 & 0 & -2b_{5,7} & 0 & -b_{5,9} & b_{6,9} & -b_{8,9} \\ b_{1,7} & -b_{3,6} & 0 & b_{4,7} & b_{5,7} & 0 & 1 & -b_{6,9} & 0 & b_{7,10} \\ b_{1,8} & b_{2,8} & b_{3,8} & b_{4,8} & 0 & -b_{5,9} & -b_{6,9} & -2b_{5,10} & b_{8,9} & 0 \\ b_{1,9} & b_{2,9} & -b_{4,7} & -b_{3,10} & b_{5,9} & b_{6,9} & 0 & b_{8,9} & -2b_{7,10} & 0 \\ b_{1,10} & -b_{4,8} & b_{3,10} & 0 & b_{5,10} & -b_{8,9} & b_{7,10} & 0 & 0 & 1 \end{pmatrix}$$

3.1.2 Matrice de Gram

On garde les mêmes notations. On sait maintenant associer à un polynôme $f \in \mathbb{R}[X]$ un espace affine \mathcal{L}_f qui le représente, et dont on connaît la dimension.

Le théorème

Dans cette section, on va expliciter la condition pour f d'être une somme de carrés en une condition sur \mathcal{L}_f , à savoir :

Théorème 3.1.5

Soit $f \in \mathbb{R}[X]$ de degré $2d$.

Alors :

$$f \in \sum \mathbb{R}[X]^2 \text{ ssi } \mathcal{L}_f \cap \mathcal{S}_n^+(\mathbb{R}) \neq \emptyset$$

Définition 3.1.6

Une matrice $B \in \mathcal{L}_f \cap \mathcal{S}_n^+(\mathbb{R})$ est appelée matrice de Gram associée à f .

Démonstration :

\Rightarrow Supposons que $f \in \sum \mathbb{R}[X]^2$:

$$f = \sum_{i=1}^t h_i^2.$$

3.1 Reconnaître les sommes de carrés : un algorithme

Soit $A \in M_{N \times N}(\mathbb{R})$ la matrice dont la i^e colonne est composée des coefficients de h_i . Alors

$$f = m (A^t A)^t m.$$

Posant $B = A^t A$, on obtient la matrice attendue.

◀ Supposons qu'il existe une matrice $B \in \mathcal{L}_f \cap \mathcal{S}_n^+(\mathbb{R})$ de rang t . B est alors diagonalisable dans une base orthonormée, et on peut donc écrire

$$B = P D^t P$$

où $P \in GL_N(\mathbb{R})$ et $D = \text{diag}(d_1, \dots, d_t, 0, \dots, 0)$ avec $d_i > 0$ pour tout i . On note $V = (v_{ij})$. On pose alors, pour $i \in \llbracket 1, t \rrbracket$,

$$h_i = \sqrt{d_i} \sum_{j=1}^N v_{ji} x^{\beta_j} \in \mathbb{R}[X].$$

De l'égalité $f = m(PD^tP)^t m$, on déduit finalement $f = h_1^2 + \dots + h_t^2$. ■

Remarque :

La preuve ci-dessus construit explicitement l'écriture de f comme somme de carrés à partir de la matrice B , à condition de connaître les valeurs propres de B . Ces valeurs propres sont réelles, mais on ne reste pas forcément dans le corps des coefficients de B .

Réduction de la dimension

Le problème qui se pose en pratique est la dimension de \mathcal{L}_f qui croît rapidement avec le nombre de variables et le degré du polynôme. Mais on peut simplifier ce problème en éliminant de la liste m des monômes ceux dont on sait qu'ils ne peuvent pas apparaître dans une décomposition de f en somme de carrés. Par exemple, on a implémenté sous Maple la règle suivante : Soit $\gamma \in \Lambda_{2d}$. Supposons que γ s'écrive $\gamma = 2\alpha$ et qu'il ne s'écrive d'aucune autre manière comme somme d'éléments de Λ_d . Si $a_\gamma = 0$, alors x^α ne peut apparaître dans l'écriture éventuelle de f en somme de carrés.

Exemple 1 : Voici un premier exemple en Maple :

```
> P:=x^2*y^2+x^2+y^2+1;
      P := x^2 y^2 + x^2 + y^2 + 1
> elim(P);
      [[1, 1], [1, 0], [0, 1], [0, 0]]
```

Ainsi, seuls les monômes $xy, x, y, 1$ peuvent intervenir dans une décomposition de f en somme de carrés, en suivant notre règle.

Exemple 2 : L'exemple de Parrilo et Sturmfels :

> P:=x**4+y**4+z**4-4*x*y*z+x+y+z-lambda;

$$P := x^4 + y^4 + z^4 - 4xyz + x + y + z - \lambda$$

> elim(P);

[[2, 0, 0], [1, 1, 0], [1, 0, 1], [1, 0, 0], [0, 2, 0], [0, 1, 1], [0, 1, 0], [0, 0, 2], [0, 0, 1], [0, 0, 0]]

On va voir que notre méthode ne nous a pas permis d'enlever de monômes ; à l'aide de la procédure lexico(nbvar,deg), qui génère tous les monômes en nbvar variables de degré d, on retrouve bien tous les monômes fournis précédemment par la procédure elim(P) :

> lexico(3,2);

[[2, 0, 0], [1, 1, 0], [1, 0, 1], [1, 0, 0], [0, 2, 0], [0, 1, 1], [0, 1, 0], [0, 0, 2], [0, 0, 1], [0, 0, 0]]

Intérêt : On peut ainsi réduire la dimension de l'espace \mathcal{L}_f , qui devient, si le nombre de monômes pouvant intervenir est M :

$$\dim(\mathcal{L}_f) = \frac{M(M+1)}{2} - \text{rg}(A)$$

où A est la matrice associée au nouveau système linéaire. A est de taille $\left(|\Lambda_{2d}|, \frac{M(M+1)}{2}\right)$. Pour calculer son rang, on peut remarquer qu'il peut désormais apparaître dans A des lignes nulles, disons au nombre de l . Par un raisonnement semblable à ce qui a été fait auparavant, on obtient :

$$\text{rg}(A) = |\Lambda_{2d}| - l$$

Matrices semi-définies positives et inégalités polynômiales

Dans un premier temps, on a donné un algorithme qui détermine \mathcal{L}_f , ainsi qu'implémenté une procédure sous Maple qui renvoie la forme générale d'une matrice de \mathcal{L}_f , matrice à $p = \dim(\mathcal{L}_f)$ paramètres, notés $\theta_1, \dots, \theta_p$.

On a ensuite réduit cette dimension, et donné une condition nécessaire et suffisante pour que f soit une somme de carrés : il faut et il suffit que \mathcal{L}_f contienne une matrice semi-définie positive.

Dans cette partie, nous allons exprimer cette dernière condition à l'aide d'inégalités polynômiales sur les paramètres $\theta_1, \dots, \theta_p$. Pour cela l'outil essentiel sera la règle de Descartes. Il nous faut donc introduire la définition suivante :

Définition 3.1.7

Le nombre de changements de signes $V(\underline{a})$ d'une suite finie $\underline{a} \in (\mathbb{R}^*)^k$ est

3.1 Reconnaître les sommes de carrés : un algorithme

défini par induction sur $k \in \mathbb{N}^*$ par :

$$V(a_1) = 0$$

$$V(a_1, \dots, a_k) = \begin{cases} V(a_1, \dots, a_{k-1}) + 1 & \text{si } \text{sign}(a_{k-1}a_k) = -1 \\ V(a_1, \dots, a_{k-1}) & \text{sinon} \end{cases}$$

Cette définition s'étend à une suite finie \underline{a} d'éléments de \mathbb{R} en considérant la suite finie \underline{b} obtenue en supprimant les zéros de \underline{a} et en définissant alors $V(\underline{a}) = V(\underline{b})$.

Proposition 3.1.8

Soit $M \in \mathcal{S}_n(\mathbb{R})$. On note $\chi_M = T^n + a_{n-1}T^{n-1} + \dots + a_0$ son polynôme caractéristique. Soit $k \in \llbracket 0, n \rrbracket$ la valuation de χ_M .

On a alors l'équivalence suivante, en posant $a_n = 1$:

$$M \in \mathcal{S}_n^+(\mathbb{R}) \Leftrightarrow \forall i \in \llbracket k, n \rrbracket, (-1)^{i+n}a_i > 0.$$

Démonstration :

χ_M n'admet que des racines réelles. De plus, $M \in \mathcal{S}_n^+(\mathbb{R})$ si et seulement si χ_M n'admet que des racines positives. La règle de Descartes, exacte dans le cas d'un polynôme n'admettant que des racines réelles, nous donne :

$$\text{card}\{z \geq 0 / \chi_M(z) = 0\} = V(\chi_M)$$

Factorisons χ_M de la façon suivante :

$$\chi_M(T) = T^k Q(T)$$

Alors $Q(0) \neq 0$, et par conséquent :

$$\begin{aligned} \text{card}\{z \geq 0 / \chi_M(z) = 0\} &= k + \text{card}\{z \geq 0 / Q(z) = 0\} \\ &= k + V(Q) \end{aligned}$$

On a finalement :

$$\begin{aligned} M \in \mathcal{S}_n^+(\mathbb{R}) \\ \Leftrightarrow \text{card}\{z \geq 0 / \chi_M(z) = 0\} &= n \\ \Leftrightarrow k + V(Q) &= n \end{aligned}$$

Mais, puisque $\deg(Q) = n - k$, $V(Q) \leq n - k$, et donc $M \in \mathcal{S}_n^+(\mathbb{R})$ si et seulement si $V(Q)$ est maximal, ie les coefficients de Q sont tous non nuls et alternés, CQFD. ■

On peut exprimer ces conditions sous la forme d'inégalités larges :

Corollaire 3.1.9

Avec les mêmes notations :

$$M \in \mathcal{S}_n^+(\mathbb{R}) \Leftrightarrow \forall i \in \llbracket 0, n \rrbracket, (-1)^{i+n} a_i \geq 0$$

Démonstration :

\Rightarrow Conséquence immédiate de la proposition précédente.

\Leftarrow Supposons que $\forall i \in \llbracket 0, n \rrbracket, (-1)^{i+n} a_i \geq 0$. Montrons que χ_M ne peut pas avoir de racine strictement négative, ce qui permet de conclure puisque χ_M n'admet que des racines réelles.

Soit donc $z < 0$.

Supposons que χ_M soit de degré pair :

$$\chi_M = T^{2d} + a_{2d-1}T^{2d-1} + \dots + a_0$$

On a alors :

$$\chi_M(z) = z^{2d} + \underbrace{\sum_{i=0}^{d-1} a_{2i} (z^2)^i}_{\geq 0} + z \underbrace{\sum_{i=0}^{d-1} \underbrace{a_{2i+1}}_{\leq 0} (z^2)^i}_{\leq 0} \geq z^{2d} > 0$$

et donc z n'est pas racine de χ_M .

Un argument analogue permet de conclure dans le cas où χ_M est de degré impair. ■

3.1.3 Application : exemple de Mötzkin

Mötzkin, en 1967, fut le premier à donner un exemple d'un polynôme positif qui ne soit pas somme de carrés de polynômes. Il s'agit du polynôme

$$S = z^6 + x^4 y^2 + x^2 y^4 - 3x^2 y^2 z^2$$

La positivité de S résulte de l'inégalité arithmético-géométrique, en remarquant que $x^2 y^2 z^2$ est la moyenne géométrique de $z^6, x^4 y^2$ et $x^2 y^4$.

On se propose d'utiliser ce qu'on a fait dans les parties précédentes pour retrouver le fait que $S \notin \sum \mathbb{R}[X]^2$.

> $S := z^6 + x^4 y^2 + x^2 y^4 - 3x^2 y^2 z^2;$

$$S := z^6 + x^4 y^2 + x^2 y^4 - 3x^2 y^2 z^2$$

> $m := \text{elim}(S);$

$$m := \llbracket [2, 1, 0], [1, 2, 0], [1, 1, 1], [0, 0, 3] \rrbracket$$

Ainsi, seuls les monômes $x^2 y, xy^2, xyz$ et z^3 peuvent intervenir dans une éventuelle décomposition de S en somme de carrés.

3.2 Minimisation des polynômes à l'aide de la programmation semi-définie

> Lf(S,m) ;

l'espace affine est de dimension :

$$0 \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

\mathcal{L}_S est de dimension 0 : il ne contient qu'une seule matrice. Evidemment, on n'a pas besoin des conditions obtenues dans la partie 1.3.3 (inégalités sur les coefficients du polynôme caractéristique) pour voir que cette matrice n'est pas semi-définie positive, -3 étant visiblement valeur propre. Par conséquent, on peut directement conclure que $S \notin \sum \mathbb{R}[X]^2$.

3.2 Minimisation des polynômes à l'aide de la programmation semi-définie

Soit $P \in \mathbb{R}[X] = \mathbb{R}[X_1, \dots, X_n]$ de degré $\deg P = 2d$. Le problème est de minimiser P :

$$\begin{aligned} \text{(P1)} \quad & \text{minimiser} && P(x) \\ & \text{sous la contrainte} && x \in \mathbb{R}^n \end{aligned}$$

Si le minimum est fini, on veut aussi déterminer un point qui réalise ce minimum. Ce problème est en pratique très difficile à résoudre (pour un article de synthèse concernant les différentes méthodes algébriques exactes, on pourra consulter [\[PS\]](#)).

3.2.1 La méthode de Parrilo et Sturmfels

La première idée qui peut venir à l'esprit est celle-ci : Soit p^* le minimum de P sur \mathbb{R}^n . Une condition suffisante pour que $\lambda \leq p^*$ est que $P - \lambda$ soit une somme de carrés de $\mathbb{R}[X]$. On va donc considérer le problème suivant :

$$\begin{aligned} \text{(P2)} \quad & \text{maximiser} && \lambda \\ & \text{sous la contrainte} && P(x) - \lambda \in \sum \mathbb{R}[X]^2 \end{aligned}$$

En considérant l'espace $\mathcal{L}_{P-\lambda}$ (voir la section précédente), ce problème est équivalent au problème SDP suivant :

$$\begin{aligned}
 \text{(P3)} \quad & \min && C \bullet X \\
 & \text{contraintes} && X \in \mathcal{L}_{P-\lambda} \\
 & && X \succeq 0
 \end{aligned}$$

où $C = E_{11}$, la matrice dont le seul coefficient non nul est le coefficient $(1, 1)$, qui vaut 1.

On note p^{sos} la valeur optimale de ce problème SDP.

Ce problème est une relaxation du problème (P1), c'est-à-dire qu'on a toujours $p^{sos} \leq p^*$. L'idée de Parrilo et Sturmfels est de reformuler ce problème en termes de formes quadratiques. L'intérêt est de pouvoir produire un test pour vérifier si $p^{sos} = p^*$.

Formulation en termes de formes quadratiques

On rappelle que $N = C_{n+d}^n$ désigne le nombre de monômes en n variables de degré inférieur ou égal à d .

On pose $m = N - 1$, et on identifie $\mathcal{S}_n(\mathbb{R})$ avec le \mathbb{R} -espace vectoriel $\mathbb{R}[x]_2 = \mathbb{R}[x_0, \dots, x_m]_2$ des formes quadratiques réelles en $m + 1$ variables. Le dual de $\mathcal{S}_n(\mathbb{R})$ est noté $\mathbb{R}[\partial]_2 = \mathbb{R}[\partial_0, \dots, \partial_m]_2$, et le crochet de dualité est donné par dérivation, et sera noté \bullet . On a par exemple l'identité suivante, valable pour tout $f \in \mathbb{R}[\partial]_2$ et $p = (p_0, \dots, p_m) \in \mathbb{R}^{m+1}$:

$$f(\partial_0, \dots, \partial_m) \bullet \frac{1}{2} \left(\sum_{i=0}^m p_i x_i \right)^2 = f(p_0, \dots, p_m). \quad (*)$$

Preuve :

On a :

$$\begin{aligned}
 \partial_i \partial_j \left[\frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2 \right] &= \partial_i \left[\sum_{k=0}^m p_k x_k \partial_j \left(\sum_{k=0}^m p_k x_k \right) \right] \\
 &= \partial_i \left[p_j \sum_{k=0}^m p_k x_k \right] \\
 &= p_j p_i
 \end{aligned}$$

On écrit $f(x) = \sum \alpha_{ij} x_i x_j$. On obtient alors :

$$\begin{aligned}
 f(\partial) \bullet \left[\frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2 \right] &= \sum \alpha_{ij} \partial_i \partial_j \left[\frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2 \right] \\
 &= \sum \alpha_{ij} p_i p_j \\
 &= f(p),
 \end{aligned}$$

3.2 Minimisation des polynômes à l'aide de la programmation semi-définie

ce qu'il fallait démontrer. ■

Considérons le problème suivant :

$$\begin{aligned}
 (\mathbf{P4}) \quad & \min && f(p) \\
 & \text{contraintes} && g_0(p) = 1 \\
 & && g_1(p) = \dots = g_r(p) = 0
 \end{aligned}$$

où $f, g_0, \dots, g_r \in \mathbb{R}[\partial]_2$ sont donnés, et où l'on cherche une solution optimale $p \in \mathbb{R}^{m+1}$.

Ce problème peut se réécrire de la façon suivante :

$$\begin{aligned}
 (\mathbf{P4}) \quad & \min && f(\partial) \bullet q(x) \\
 & \text{contraintes} && g_0(\partial) \bullet q(x) = 1 \\
 & && g_1(\partial) \bullet q(x) = \dots = g_r(\partial) \bullet q(x) = 0 \\
 & && q(x) = \frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2
 \end{aligned}$$

Ce problème admet la relaxation SDP suivante :

$$\begin{aligned}
 (\mathbf{P}) \quad & \min && f(\partial) \bullet q(x) \\
 & \text{contraintes} && g_0(\partial) \bullet q(x) = 1 \\
 & && g_1(\partial) \bullet q(x) = \dots = g_r(\partial) \bullet q(x) = 0 \\
 & && q(x) \succeq 0
 \end{aligned}$$

où $q(x) \succeq 0$ signifie que q est positive sur \mathbb{R}^{m+1} .

L'identité (*) permet d'affirmer que la valeur optimale de **(P4)** est supérieure ou égale à celle de **(P)**, avec égalité si et seulement si **(P)** admet une solution optimale de la forme

$$q(x) = \frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2 .$$

La détermination du problème dual de **(P)** est immédiate, et fournit le problème :

$$\begin{aligned}
 (\mathbf{D}) \quad & \max && \lambda \\
 & \text{contraintes} && f(\partial) - \sum_{i=1}^r \mu_i g_i(\partial) - \lambda g_0(\partial) \succeq 0
 \end{aligned}$$

où la variable est le $(r+1)$ -uplet $(\lambda, \mu_1, \dots, \mu_r) \in \mathbb{R}^{r+1}$.

On a alors le résultat suivant :

Proposition 3.2.1

Si le problème primal (\mathbf{P}) admet une solution optimale de rang 1, ie de la forme

$$q(x) = \frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2,$$

alors le vecteur (p_0, \dots, p_m) est une solution optimale de $(\mathbf{P1})$.

Variété de Véronèse

On étudie maintenant la minimisation des fonctions quadratiques sur une certaine variété projective, dite de Véronèse. On verra par la suite que ce problème est équivalent au problème $(\mathbf{P1})$.

Définition 3.2.2

On considère la variété, dans l'espace projectif de dimension N , paramétrée par tous les monômes de degré au plus d .

Cette variété est appelée variété de Véronèse, et est notée \mathcal{V} .

Exemple :

On considère le polynôme

$$P(r, s, t) = r^4 + s^4 + t^4 - 4rst + r + s + t.$$

On a alors $n = 3$, $d = 2$, $m = 9$.

La variété de Véronèse est donnée de manière paramétrique par :

$$\mathcal{V} = \{(z_0 : \dots : z_9) = (1 : r : s : t : r^2 : rs : s^2 : rt : st : t^2)\}$$

\mathcal{V} est en fait une variété torique : on peut trouver un idéal généré par des binômes quadratiques $g_1, \dots, g_r \in \mathbb{R}[\partial_0, \dots, \partial_m]$, chacun étant de la forme $\partial_i \partial_j - \partial_k \partial_l$ ([Stu]).

Posons maintenant $g_0(\partial) = \partial_0^2$.

L'équation $g_0(\partial) = 1$ sur \mathcal{V} définit une variété affine $\tilde{\mathcal{V}}$ telle que son complété projectif est \mathcal{V} .

Le problème $(\mathbf{P4})$ est alors équivalent au problème :

$$(\mathbf{P5}) \quad \min \quad f(p) \\ \text{contraintes} \quad p \in \tilde{\mathcal{V}}$$

On peut alors faire le lien entre \mathcal{V} et le problème (\mathbf{D}) :

la valeur optimale de (\mathbf{D}) est le plus grand réel λ tel que $f - \lambda$ est une somme de carrés dans l'anneau des coordonnées de \mathcal{V} .

3.2 Minimisation des polynômes à l'aide de la programmation semi-définie

Retour au problème initial

On revient au problème de départ, celui de minimiser le polynôme P . P peut s'écrire comme une forme quadratique en les monômes de degré $\leq d$. Ce faisant, le problème de minimiser P est exactement le problème **(P5)**, avec $f = P$.

Exemple :

On considère toujours le polynôme

$$P(r, s, t) = r^4 + s^4 + t^4 - 4rst + r + s + t.$$

Alors

$$\begin{aligned} P(r, s, t) &= (r^2)^2 + (s^2)^2 + (t^2)^2 - 4(r)(st) + (1)(r) + (1)(s) + (1)(t) \\ &= (z_4)^2 + (z_6)^2 + (z_9)^2 - 4(z_1)(z_8) + (z_0)(z_1) + (z_0)(z_2) + (z_0)(z_3) \in \mathbb{R}[z_0, \dots, z_9]_2 \end{aligned}$$

Ainsi, pour minimiser le polynôme P , on va résoudre simultanément les relaxations SDP **(P)** et **(D)**.

Si la valeur optimale λ de **(D)** est égale au minimum de P sur $\tilde{\mathcal{V}}$, alors le problème primal **(P)** admet une solution optimale de la forme $q(x) = \frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2$, ce qui fournit un point optimal $p = (p_0 : \dots : p_m) \in \mathcal{V}$ qui minimise P .

On obtient donc ainsi un test pour vérifier si $p^{sos} = p^*$:

On résout les problèmes primal et dual. Si le problème primal admet une solution de rang 1, ie de la forme

$$q(x) = \frac{1}{2} \left(\sum_{k=0}^m p_k x_k \right)^2,$$

on obtient un candidat pour le minimum de P , et un point où il est atteint. En évaluant P en ce point, on obtient une majorant du minimum de P , que l'on peut comparer avec le minorant p^{sos} . Si les deux coïncident, on a donc trouvé le minimum de P .

Intérêt :

L'intérêt de cette méthode est double :

- D'une part, les algorithmes SDP sont efficaces. La dimension qui intervient dans la relaxation est N , qui croît polynomialement en n (à d fixé), alors que pour des résolutions algébriques classiques, la mesure qui intervient est le nombre de Bezout μ , qui augmente de manière exponentielle en n (voir [PS] pour un survey sur ces méthodes, et la comparaison de N et de μ).
- D'autre part, on dispose d'un test pour vérifier si la borne inférieure trouvée à la suite de la relaxation est exactement le minimum.

Une autre voie est également possible, combinant également ces deux avantages : la théorie des moments.

3.2.2 Une alternative : GloptiPoly

GloptiPoly est un add-on de Matlab qui résoud numériquement les relaxations de problèmes de minimisation polynomiale sous des contraintes exprimées sous la forme d'inégalités polynomiales, utilisant la théorie des moments ([HL1]). L'optimalité éventuelle de la solution est détectée, et des points optimaux sont calculés ([HL2]). Voyons tout de suite un exemple d'utilisation.

On reprend l'exemple de l'article de Parrilo et Sturmfels :

On affiche le plus de chiffres significatifs possibles :

```
>> format long e
```

Puis on définit le polynôme à minimiser :

```
>> P{1} = defipoly('min x^4+y^4+z^4-4*x*y*z+x+y+z','x,y,z');
Objective function to minimize: min x^4+y^4+z^4-4*x*y*z+x+y+z
```

Enfin, on lance Gloptipoly :

```
>> output =gloptipoly(P)
GloptiPoly 2.2f - Global Optimization over Polynomials with SeDuMi
Number of variables = 3
Number of constraints = 0
Maximum polynomial degree = 4
Order of LMI relaxation = 2
Building LMI. Please wait..
Number of LMI decision variables = 34
Size of LMI constraints = 100
Sparsity of LMI constraints = 2.9118% of non-zero entries
Norm of perturbation of objective function = 0
Numerical accuracy for SeDuMi = 1e-009
No feasibility radius
Solving LMI problem with SeDuMi..
SeDuMi 1.05R5 by Jos F. Sturm, 1998, 2001-2003.
Alg = 2: xz-corrector, theta = 0.250, beta = 0.500
eqs m = 34, order n = 11, dim = 101, blocks = 2
nnz(A) = 54 + 0, nnz(ADA) = 1156, nnz(L) = 595
it :      b*y      gap   delta rate  t/tP*  t/tD*  feas cg cg
```

3.2 Minimisation des polynômes à l'aide de la programmation semi-définie

```

0 :          3.28E+000 0.000
1 : -8.79E-001 1.06E+000 0.000 0.3240 0.9000 0.9000 1.73 1 1
2 :  8.08E-001 2.84E-001 0.000 0.2672 0.9000 0.9000 1.60 1 1
3 :  1.55E+000 7.22E-002 0.000 0.2541 0.9000 0.9000 0.74 1 1
4 :  1.95E+000 1.63E-002 0.000 0.2260 0.9000 0.9000 0.79 1 1
5 :  2.11E+000 6.52E-004 0.016 0.0400 0.9900 0.9900 0.88 1 1
6 :  2.11E+000 7.69E-005 0.151 0.1179 0.9450 0.9450 0.98 1 1
7 :  2.11E+000 1.48E-005 0.000 0.1921 0.9000 0.9000 0.99 1 1
8 :  2.11E+000 1.67E-006 0.073 0.1133 0.9450 0.9450 1.00 1 1
9 :  2.11E+000 1.63E-007 0.490 0.0975 0.9900 0.9900 1.00 1 1
10 : 2.11E+000 4.94E-008 0.000 0.3031 0.9000 0.9000 1.00 2 2
11 : 2.11E+000 3.36E-009 0.117 0.0681 0.9900 0.9900 1.00 4 4
12 : 2.11E+000 7.17E-010 0.033 0.2131 0.9000 0.9000 1.00 9 9
13 : 2.11E+000 2.34E-010 0.000 0.3264 0.9000 0.9000 1.00 11 11
iter seconds digits      c*x          b*y
13      1.4   9.8 2.1129138790e+000 2.1129138787e+000
|Ax-b| = 8.6e-010, [Ay-c]_+ = 3.8E-011, |x|= 4.2e+000, |y|= 4.4e+000
Max-norms: ||b||=4, ||c|| = 1,
Cholesky |add|=0, |skip| = 2, ||L.L|| = 219.815.
CPU time = 1.359 sec
LMI objective = -2.1129
Checking relaxed LMI vector with threshold = 1e-006
Relaxed vector reaches objective -0.86867
Relaxed vector is feasible
Detecting global optimality (rank shift = 1)..
Relative threshold for rank evaluation = 0.001
Moment matrix of order 1 has size 4 and rank 3
Moment matrix of order 2 has size 10 and rank 3
Rank condition ensures global optimality
Extracting solutions..
Relative threshold for basis detection = 1e-006
Maximum relative error = 4.3942e-010
3 solutions extracted

output =

      status: 1
      obj: -2.112913878652760e+000
      sol: {[3x1 double] [3x1 double] [3x1 double]}

>> output.obj

ans =

```

CHAPITRE 3 : *Polynômes positifs, sommes de carrés; minimisation*

-2.112913878652760e+000

```
>> output.sol{:}
```

ans =

-1.102258570654554e+000
9.881831264135703e-001
-1.102258570668643e+000

ans =

-1.102258570948249e+000
-1.102258570782620e+000
9.881831266139900e-001

ans =

9.881831276140625e-001
-1.102258571472540e+000
-1.102258571594003e+000

Ainsi, Gloptipoly retrouve bien le minimum trouvé par Parrilo et Sturmfels en utilisant leur méthode :

obj: -2.112913878652760e+000

et a extrait les trois points réalisant ce minimum :

```
>> output.sol{:}
```

ans =

-1.102258570654554e+000
9.881831264135703e-001
-1.102258570668643e+000

ans =

-1.102258570948249e+000
-1.102258570782620e+000
9.881831266139900e-001

3.3 L'ensemble des sommes de carrés dans l'ensemble des polynômes positifs : aspects quantitatifs

ans =

```
9.881831276140625e-001  
-1.102258571472540e+000  
-1.102258571594003e+000
```

Gloptipoly assure également que le minimum trouvé à partir de la relaxation est bien le vrai minimum du polynôme (ceci est assuré en vérifiant une condition de rang sur les matrices des moments) :

```
Detecting global optimality (rank shift = 1)..  
Relative threshold for rank evaluation = 0.001  
Moment matrix of order 1 has size 4 and rank 3  
Moment matrix of order 2 has size 10 and rank 3  
Rank condition ensures global optimality
```

On remarquera la facilité d'utilisation de Gloptipoly.

3.3 L'ensemble des sommes de carrés dans l'ensemble des polynômes positifs : aspects quantitatifs

On étudie dans cette partie la place (au sens quantitatif) qu'occupent les sommes de carrés parmi les polynômes positifs. On se référera à l'article de G. Blekherman [Ble] et aux références citées dans cet article pour les démonstrations des résultats énoncés ici.

Intérêt de cette étude

On a vu dans la partie précédente que l'on pouvait résoudre efficacement des relaxations SDP basées sur le principe suivant : remplacer les polynômes positifs par les sommes de carrés de polynômes. Ces relaxations peuvent être résolues à l'aide d'algorithmes performants, alors que dans le cas général, on est confronté à un problème NP. La question naturelle est alors de se demander quelle erreur on commet lorsque l'on remplace les polynômes positifs par les sommes de carrés de polynômes. Autrement dit, trouve-t-on souvent la solution optimale ou, au contraire, donne-t-on souvent un minorant strict de cette solution ?

On a d'abord cru ([PS]) que la méthode exposée précédemment fonctionnait très souvent. Ce n'est qu'après les travaux de G. Blekherman que l'on s'est rendu compte que cette conjecture était fautive.

Dans cette partie nous étudierons donc la place qu'occupent les sommes de

carrés parmi les polynômes positifs. Utilisant la structure conique de ces deux ensembles, on peut quantifier cette place en comparant leur volume. C'est ce qui a été fait dans l'article de G. Blekherman, dont nous donnons les résultats dans ce qui suit.

Préliminaires

Notation :

Soit μ la mesure de Lebesgue sur \mathbb{R}^n . Soit E une partie Lebesgue-mesurable de \mathbb{R}^n .

Le volume de E est noté $vol(E)$.

Définition 3.3.1

Soit V un espace vectoriel normé.

Une partie $E \subset V$ est appelée corps convexe si elle est convexe et d'intérieur non vide.

Notation :

◦ $P = P_{n,2k}$ désigne l'espace vectoriel des polynômes réels homogènes de degré k en n variables. On considère le produit scalaire sur $P_{n,2k}$ suivant :

$$\langle f, g \rangle = \int_{S^{n-1}} fg \, d\sigma$$

où S^{n-1} désigne la sphère unité de \mathbb{R}^n et σ l'unique mesure sur S^{n-1} invariante par rotation.

◦ On considère le cône des polynômes positifs de $P_{n,2k}$:

$$C = C_{n,2k} = \{f \in P_{n,2k} / \forall x \in \mathbb{R}^n, f(x) \geq 0\}$$

◦ On considère également le cône des sommes de carrés dans $P_{n,2k}$:

$$S = S_{n,2k} = \{f \in P_{n,2k} / f = \sum_{finie} f_i^2, f_i \in P_{n,k}\}$$

Afin de comparer quantitativement C et S , on va introduire pour chacun une base compacte, et comparer leur volume.

Notation :

◦ Soit M l'hyperplan formé des polynômes de $P_{n,2k}$ d'intégrale nulle sur la sphère unité :

$$M = \{f \in P_{n,2k} / \int_{S^{n-1}} f \, d\sigma = 0\}$$

3.3 L'ensemble des sommes de carrés dans l'ensemble des polynômes positifs : aspects quantitatifs

On note D_M la dimension de M et B_M la boule unité de M (au sens de la norme induite par le produit scalaire $\langle \cdot, \cdot \rangle$).

On va considérer les sections des cônes C et S par l'hyperplan M , à une translation près :

◦ r^{2k} désignera le polynôme $(x_1^2 + \dots + x_n^2)^k \in P_{n,2k}$.

◦ On introduit maintenant les sections \tilde{C} et \tilde{S} des cônes C et S :

$$\tilde{C} = \{f \in M / f + r^{2k} \in C\}$$

$$\tilde{S} = \{f \in M / f + r^{2k} \in S\}$$

Les théorèmes

On donne maintenant les résultats de Blekherman concernant les volumes de \tilde{C} et \tilde{S} .

Théorème 3.3.2 *On a les deux bornes suivantes :*

$$\left(\frac{\text{Vol}(\tilde{C})}{\text{Vol}(B_M)} \right)^{\frac{1}{D_M}} \geq \frac{1}{2\sqrt{4k+2}} \frac{1}{\sqrt{n}}$$

$$\left(\frac{\text{Vol}(\tilde{S})}{\text{Vol}(B_M)} \right)^{\frac{1}{D_M}} \leq \frac{4^{2k}(2k)!\sqrt{24}}{k!} \frac{1}{\sqrt{n^k}}$$

On peut maintenant donner un minorant du rapport des volumes de \tilde{C} et de \tilde{S} :

Corollaire 3.3.3 *On a :*

$$\left(\frac{\text{Vol}(\tilde{C})}{\text{Vol}(\tilde{S})} \right)^{\frac{1}{D_M}} \geq c(k)\sqrt{n^{k-1}}$$

où

$$c(k) = \frac{k!}{2(2k)!4^{2k}\sqrt{24(4k+2)}}.$$

Ainsi, si l'on fixe le degré $2k$ avec $k \geq 2$, alors le rapport tend vers $+\infty$ avec le nombre n de variables, et en ce sens on peut dire qu'il y a beaucoup plus de polynômes positifs que de sommes de carrés quand le nombre de variables est grand par rapport au degré.

Si $k = 1$, on sait que $\tilde{C} = \tilde{S}$, et en effet le minorant dans le corollaire devient juste $c(1)$.

CHAPITRE 3 : *Polynômes positifs, sommes de carrés ; minimisation*



Bibliographie

- [Ali] Farid Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim. 5 n°1, 1995.
- [Boy] Vanderberghe et Boyd, *Semidefinite programming*, SIAM review, mars 1996.
- [Tod] M. J. Todd, *Semidefinite Optimization*, Cambridge University Press, Acta Numerica 10 (2001) 515-560.
- [Roc] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [HJ1] Horn et Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [HJ2] Horn et Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
- [Hel] C. HELMBERG, *Semidefinite Programming for Combinatorial Optimization*, ZIB-Report 00-34, Octobre 2000.
- [Zha] Fuzhen Zhang, *Matrix Theory, Basics results and techniques*, Springer-Verlag, New York, 1999.
- [PW] V. Powers et T. Wörmann, *An algorithm for sums of squares of real polynomials*, Journal of Pure and Applied Algebra 127 (1998), 99-104.
- [PS] P. Parrilo et B. Sturmfels, *Minimizing Polynomial Functions*, dans Algorithmic and quantitative real algebraic geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 60, pp. 83-99, AMS.

BIBLIOGRAPHIE

- [Par] P. Parrilo, *Structured semidenite programs and semialgebraic geometry methods in robustness and optimization*, Ph.D. thesis, California Institute of Technology, 2000.
- [Stu] B. Sturmfels, *Equations Defining Toric Varieties*, Proceedings of Symposia in Pure Mathematics, Volume 62.2, 1997.
- [HL1] D. Henrion et J.B. Lasserre, *GloptiPoly : Global Optimization over Polynomials with Matlab and SeDuMi*.
- [HL2] D. Henrion et J.B. Lasserre, *Detecting global optimality and extracting solutions in GloptiPoly*.
- [Ble] G. Blekherman, *There are significantly more nonnegative polynomials than sums of squares*.
- [Pis] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge University Press, 1989.
- [CLO] D. Cox, J. Little et D. O'Shea, *Ideals, varieties, and algorithms, 2nd edition*, Springer, 1997.