Programmation Semidéfinie et Optimisation Combinatoire

Frédéric Roupin

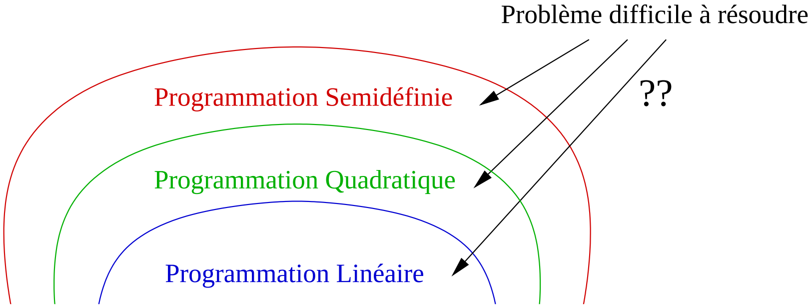
roupin@lipn.univ-paris13.fr

MPRO 2017-2018

Plan du cours

Problématique Générale

- Objectif : résoudre des problèmes combinatoires ou continus non-convexes
- Problèmes généralement NP-difficiles
- Démarche : établir des relaxations convexes pour obtenir des bornes (inférieures ou supérieures)



- algorithmes approchés
- méthodes de séparation/évaluation (Branch&Bound)
- méthodes de Branch&Cut

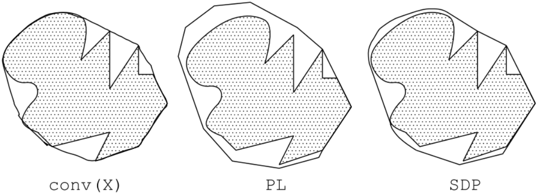
Relaxation d'un problème d'optimisation difficile

Un problème P : on cherche à optimiser $f(x)$ pour x dans une région X non convexe (souvent définie par un ensemble de contraintes)

Objectif. Obtenir une borne de la valeur optimale

Méthode. Relaxation convexe du problème. Trouver X' convexe tel que :

- $X \subset X'$.
- Optimiser $f(x)$ pour $x \in X'$ est "facile".
- Le saut $\sigma = |val_{opt}(P) - val_{opt}(P')|$ dû à la relaxation est petit.
- Cas "idéal" : enveloppe convexe de X.
- Méthodes réalistes : programmation linéaire, programmation semidéfinie (cas non-linéaire).



Mixed Integer Non-Linear Programming

Modèle Mathématique

- Mixed Integer Non-Linear Problems

$$\begin{array}{lll} \min & g_0(x) & \\ \text{sous contraintes :} & g_k(x) \leqslant 0 & k = 1, \ldots, m \\ & l \leqslant x \leqslant u & \\ & x_j \in \mathbb{Z} & i = 1, \ldots, p \end{array}$$

- Certaines fonctions $g_k(x)$ sont non-linéaires.
- Sont combinées ici les difficultés de deux sous-classes de problèmes : contraintes d'intégrité et non-convexité.
- Très souvent, les variables entières sont utilisées :
  - pour faire un choix entre plusieurs options : variables de décision
  - Pour activer/désactiver des variables continues ou des contraintes.

Mixed Integer Non-Linear Programming

Modèle mathématique

- Mixed Integer Non-Linear Problems

$$
\begin{array}{lll}
\min & g_0(x) & \\
\text{sous contraintes :} & g_k(x) \leqslant 0 & k = 1, \ldots, m \\
 & l \leqslant x \leqslant u & \\
 & x_j \in \mathbb{Z} & i = 1, \ldots, p
\end{array}
$$

- Peu de solvers peuvent attaquer cette classe générale de problèmes. Méthodes: (spatial)Branch-and-Bound (relaxations convexes), Cutting Planes, décomposition de Bender généralisée,...
- Le choix du modèle est déterminant pour la qualité des relaxations sous-jacentes. L'approche "black box" peut être catastrophique (preprocessing hasardeux, instabilité numérique,...).
- Convexifications : LP, CQP, SDP. En particulier, les fonctions ou contraintes non-convexes et/ou non quadratiques sont souvent approximées.

Cas particuliers Quadratiques

Mixed Integer Quadratically Constrained Quadratic Programming

$$(MIQCQP) \quad \boxed{\begin{array}{lll} \min & x^T Q_0 x + c_0^T x & \\ \text{s.c.} & x^T Q_i x + c_i^T x \leqslant a_i & i \in \{1, \ldots, m\} \\ & l \leqslant x \leqslant u & \\ & x_j \in \mathbb{Z} & j = 1, \ldots, p \end{array}}$$

→ Toutes les fonctions $g_k(x)$ sont quadratiques (ou linéaires).
→ Peut malgré tout modéliser beaucoup de problèmes combinatoires !
→ Peut apparaître comme un sous-problème lors de la résolution d'un MINLP plus général.

Mixed Integer Quadratic Programming

$$(MIQP) \quad \boxed{\begin{array}{lll} \min & x^T Q x + c^T x & \\ \text{s.c.} & Ax = b & \\ & l \leqslant x \leqslant u & \\ & x_j \in \mathbb{Z} & j = 1, \ldots, p \end{array}}$$

Généralités sur la programmation semidéfinie

- Utilisation assez récente en optimisation combinatoire
- Cas particulier de programmation convexe non-linéaire (fonction et contraintes convexes)
- Généralisation de la programmation linéaire (tout PL est un SDP particulier).
- *Avantage* : relaxation plus “fine” que la PL, bornes de très bonne qualité
- *Inconvénient* : résolution numérique coûteuse
  - => généralement réservée aux problèmes difficiles (lorsque l'approche par PL échoue)

Slides du cours/Polycopié :

http://www-lipn.univ-paris13.fr/ ~ roupin/

Matrices (Semidéfinies) Positives

S_n : espace des matrices réelles symétriques $n \times n$

Produit scalaire dans S_n :

$$(A, B) \in S_n^2 \quad A \bullet B = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij}$$

Definition

$A \in S_n$ est positive, $A \succcurlyeq 0$, si et seulement si pour tout $x \in \Re^n$
$x^T A x = \sum_{i=1}^n \sum_{j=1}^n A_{ij} x_i x_j = A \bullet xx^T \geqslant 0$

$$xx^T = \begin{bmatrix} x_1^2 & \cdots & x_1 x_i & \cdots & x_1 x_n \\ \vdots & \ddots & & & \vdots \\ x_i x_1 & & x_i^2 & & x_i x_n \\ \vdots & & & \ddots & \vdots \\ x_n x_1 & \cdots & x_n x_i & \cdots & x_n^2 \end{bmatrix}$$

Exemple. $x_1^2 + 2x_2^2 - x_1 x_2 + x_1 = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 2 \end{bmatrix} \bullet xx^T + \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Matrices (Semidéfinies) Positives

Proposition

$A \in S_n$ est positive, $A \succcurlyeq 0$ si et seulement si $A = M^T \begin{bmatrix} \lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & & 0 \\ 0 & \dots & \lambda_n \end{bmatrix} M$

avec $\lambda \geqslant 0$ et $M^T M = I$

Definition

Mineur symétrique d'une matrice : déterminant d'une sous-matrice obtenue en choisissant un sous-ensemble de lignes et de colonnes de mêmes indices

Proposition

Une matrice est (semidéfinie) positive si et seulement si tous ses mineurs symétriques sont positifs ou nuls

Matrices (définies) positives : Définitions équivalentes

Matrices définies positives :

$A \succ 0$
$\forall z \neq 0 \in R^n \quad z^T A z > 0$
Les valeurs propres de A sont > 0
Les mineurs principaux de A sont > 0

Matrices positives (semidéfinies positives)

$A \succcurlyeq 0$
$\forall z \in R^n \quad z^T A z \geqslant 0$
Les valeurs propres de A sont $\geqslant 0$
Les mineurs symétriques de A sont $\geqslant 0$

$\succcurlyeq$ définit un ordre partiel : $A \succcurlyeq B \Leftrightarrow A - B \succcurlyeq 0$.

Definition

Mineur symétrique d'une matrice :
déterminant d'une sous-matrice obtenue en choisissant un sous-ensemble de lignes et de colonnes de mêmes indices

Proposition

Une matrice est (semidéfinie) positive si et seulement si tous ses mineurs symétriques sont positifs ou nuls

Exemple : $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 4 \\ 1 & 4 & 3 \end{bmatrix}$

- *mineurs symétriques* 1×1 : $1 \geqslant 0$, $6 \geqslant 0$, $3 \geqslant 0$
- *mineurs symétriques* 2×2 : $6 - 4 \geqslant 0$, $18 - 16 \geqslant 0$, $3 - 1 \geqslant 0$
- *mineur symétrique* 3×3 : $\det(A) = 0 \geqslant 0$

Décomposition de Cholesky

Théorème

Une matrice A symétrique est définie positive si et seulement si elle est factorisable sous la forme $A = VV^T$ où V est une matrice triangulaire inférieure à diagonale positive. Cette factorisation est unique.

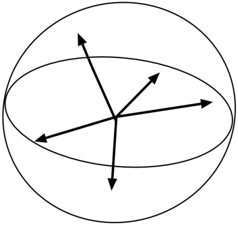
- A est la matrice de *Gram* de n vecteurs $v_1, ..., v_n$ de $\Re^n$: $\forall (i,j)\ a_{ij} = v_i^T v_j$. i.e. $A = VV^T$. En particulier les éléments diagonaux de A sont les normes au carré des vecteurs lignes de V.
- Décomposition de Cholesky incomplète : si A n'est pas définie et de rang r, elle est également factorisable en VV^T avec $rang(V) = r$. Mais il n'y a plus unicité :

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 5 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sqrt{5} & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sqrt{5} & \frac{2}{\sqrt{5}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 5 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Cas particulier important pour la modélisation en SDP

$\forall i \in \{1, ..., n\}$, $A_{ii} = 1 \Leftrightarrow$ vecteurs $v_1, \ldots, v_n \in S^n$, sphère unité.



Conséquence : se donner une matrice positive $n \times n$ dont les éléments diagonaux valent tous 1 équivaut à se donner un champ de vecteurs de la sphère unité de $\Re^n$ $\longrightarrow$ Changement de variables

Inégalités matricielles linéaires

Definition

On appelle **inégalité matricielle linéaire** toute expression de la forme $\sum_{i=1}^{m} x_i A_i \succcurlyeq A_0$, où $A_i \in S_n$ $(i = 0, \ldots, m)$ et les x_i sont des variables réelles

Exemple 1. Soit un réel t tel que $\begin{bmatrix} 1 & t & 1 \\ t & 2 & 1 \\ 1 & 1 & 4-t \end{bmatrix} \succcurlyeq 0$

ce qui est équivalent à l'*infinité* d'inégalités linéaires simples :

$\forall z \in R^n, \; z^T X z \geqslant 0.$ avec $X = \begin{bmatrix} 1 & t & 1 \\ t & 2 & 1 \\ 1 & 1 & 4-t \end{bmatrix}$

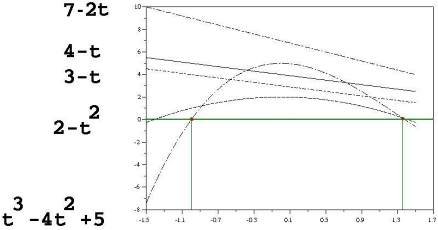
ex. $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & t & 1 \\ t & 2 & 1 \\ 1 & 1 & 4-t \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 7 - t \geqslant 0$

$$\begin{bmatrix} 1 & t & 1 \\ t & 2 & 1 \\ 1 & 1 & 4-t \end{bmatrix} \succcurlyeq 0$$

Cherchons t vérifiant l'inégalité matricielle,i.e. t vérifiant :

$4 - t \geqslant 0$; $2 - t^2 \geqslant 0$; $3 - t \geqslant 0$;

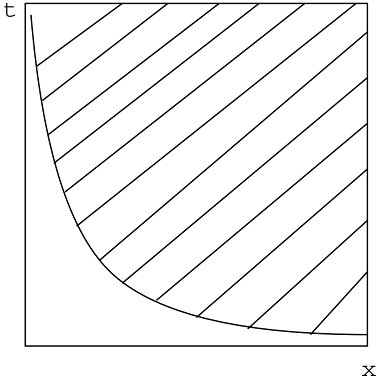
$7 - 2t \geqslant 0$; $t^3 - 4t^2 + 5 \geqslant 0$.



=> deux inégalités linéaires simples : $-1 \leqslant t \leqslant 1,381966$

Exemple 2. Soit x et t réels vérifiant $\begin{bmatrix} x & 1 \\ 1 & t \end{bmatrix} \succcurlyeq 0$

$\iff x \geqslant 0,\ t \geqslant 0,\ xt \geqslant 1$ (mineurs symétriques).



La région du plan obtenue est convexe mais n'est pas un polytope.

Inégalités matricielles linéaires (2)

- Soient $A_0, A_1, \ldots, A_m$ dans S_n, espace des matrices réelles symétriques $n \times n$
- Pour $x \in \Re^m$ on pose $F(x) = \sum_{i=1}^m x_i A_i - A_0$

$F(x)$ est une matrice appartenant
à l'espace affine défini par le repère :

- $-A_0$: origine
- $A_1, \ldots, A_m$ base
- x : coordonnées de $F(x)$ dans ce repère

Inégalités matricielles linéaires (3)

Etudions $S = \{x \in \Re^m \ : \ F(x) = \sum_{i=1}^m x_i A_i - A_0 \succcurlyeq 0\}$
S est défini par une inégalité matricielle linéaire

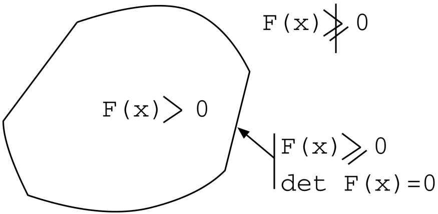
- $F(x)$ est dans l'intersection d'un espace affine et du cône des matrices (semidéfinies) positives
- La région recherchée est **convexe** car F est linéaire :
  $F(\lambda x + (1-\lambda) y) = \lambda F(x) + (1-\lambda) F(y) \succcurlyeq 0$
  $\forall \ \lambda \in [0,1], \ x \in S, \ \ y \in S$
- mais n'est généralement pas un polytope pour $m \geqslant 2$:
  x et y réels vérifiant $\begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix} \succcurlyeq 0$
  $\iff x \geqslant 0, \ y \geqslant 0, \ xy \geqslant 1$ (mineurs symétriques)

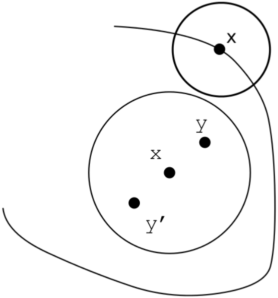
Propriétés topologiques

$S = \{x \in \Re^m \ : \ \sum_{i=1}^m A_i x_i - A_0 \succcurlyeq 0\}$

Théorème

$\overset{\circ}{S} = \{x \mid F(x) \succ 0\}$ *et* $\partial S = \{x \mid F(x) \succcurlyeq 0 \text{ et } \det F(x) = 0\}$



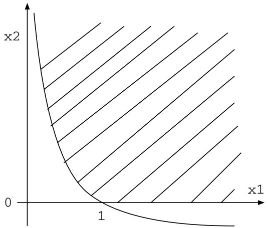


Nature de la frontière

det $F(x) = 0 \Leftrightarrow x$ annule un certain nombre de polynômes. En effet certains mineurs de $F(x)$ sont nuls (le nombre dépend du rang de $F(x)$).
$\longrightarrow$ On obtient des surfaces algébriques. En PL seulement des hyperplans !

Exemple. $F(x) = \begin{bmatrix} x_1 & 1 & 0 \\ 1 & x_2+1 & 0 \\ 0 & 0 & x_2 \end{bmatrix} \succcurlyeq 0$

On cherche x tel que det $F(x) = 0$.



Frontière: $\{x : x_1(x_2+1) - 1 = 0,\ 0 < x_1 < 1\} \cup \{x : x_2 = 0,\ x_1 \geqslant 1\}$

Programmes Semidéfinis

Definition

Un programme semidéfini est la minimisation d'une forme linéaire de $\Re^m$ soumise à une inégalité matricielle linéaire

$$(SDP) \begin{cases} \min & c^T x \\ \text{s.c.} & F(x) = \sum_{i=1}^{m} x_i A_i - A_0 \succcurlyeq 0 \end{cases}$$

Proposition

La programmation linéaire est un cas particulier de la programmation semidéfinie:

$$(PL) \begin{cases} \min & c^T x \\ \text{s.c.} & F(x) = \mathbf{diag}(Ax - b) \succcurlyeq 0 \Leftrightarrow Ax - b \geqslant 0 \end{cases}$$

$A_i^T x - b_i$ est valeur propre (positive) de la matrice diagonale $F(x)$

Formulation d'un PL en PSD

$$(PL)\begin{cases}\text{Minimiser} & c^T x\\ \text{Sous contrainte} & F(x) = \mathbf{diag}(Ax - b) \succcurlyeq 0\end{cases}$$

Exemple.

$$(P)\begin{cases}\text{Minimiser} & x_1 + 3x_2\\ \text{s.}c.\ :\ 2x_1 - x_2 \geqslant 1 & \\ \quad 5x_1 + 7x_2 \geqslant -6 & \end{cases}$$

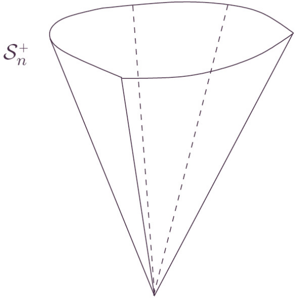
$$\Leftrightarrow\quad (SDP)\begin{cases}\text{Minimiser} & x_1 + 3x_2\\ \text{s.}c.\ : & \begin{bmatrix} 2x_1 - x_2 - 1 & 0\\ 0 & 5x_1 + 7x_2 + 6\end{bmatrix} \succcurlyeq 0\end{cases}$$

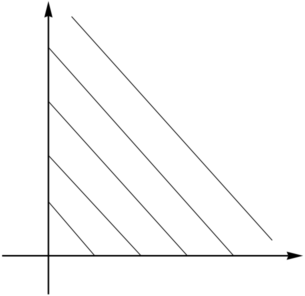
Optimisation Semidéfinie vs Optimisation Linéaire

Produit scalaire et positivité

PL : $c^T x = \sum_i c_i x_i$ Cône $\mathbb{R}^{n+} = \{x \,:\, x^T y \geqslant 0 \,\forall y \in \mathbb{R}^{n+}\}$

SDP : $C \bullet X = \sum_{ij} C_{ij} X_{ij}$ Cône $\mathcal{S}_+^n = \{X \,:\, X \bullet Y \geqslant 0 \,\forall Y \in \mathcal{S}_+^n\}$





$S_n^+ = \left\{X \in S_n \,:\, X \bullet zz^T \geqslant 0 \,\forall z \in \Re^n\right\}$

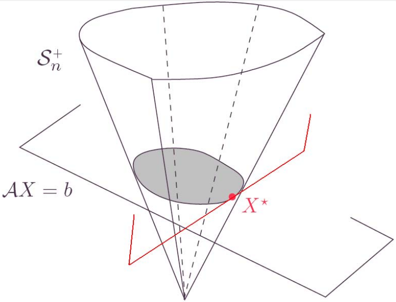
$\mathbb{R}^{n+}$

Optimisation Semidéfinie vs Optimisation Linéaire

Produit scalaire et positivité

PL : $c^T x = \sum_i c_i x_i$ Cône $\mathbb{R}^{n+} = \{x : x^T y \geqslant 0 \, \forall y \in \mathbb{R}^{n+}\}$

SDP : $C \bullet X = \sum_{ij} C_{ij} X_{ij}$ Cône $\mathcal{S}^n_+ = \{X : X \bullet Y \geqslant 0 \, \forall Y \in \mathcal{S}^n_+\}$



$$S_n^+ = \left\{X \in S_n : X \bullet zz^T \geqslant 0 \, \forall z \in \Re^n\right\} \Leftrightarrow X \succcurlyeq 0$$

$$(SDP) \begin{cases} \min & C \bullet X & = \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \\ \text{s.c.} & A_i \bullet X = b_i & \forall i \in \{1, \ldots, m\} \\ & X \succcurlyeq 0 & \end{cases}$$

Outil fondamental pour la modélisation : les compléments de Schur

Théorème

Si A matrice $p \times p$ définie positive, $C \in S_n$, et B matrice $p \times n$, alors $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succcurlyeq 0$ *est équivalent à* $C - B^T A^{-1} B \succcurlyeq 0$.

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

Cas particuliers importants

- $A = I_p$. On a $\begin{bmatrix} I_p & B \\ B^T & C \end{bmatrix} \succcurlyeq 0 \Leftrightarrow C - B^T B \succcurlyeq 0$
- $p = 1$, $C = X$, $B = x^T$. Soit $(X, x) \in S_n \times \Re^n$, on a

$$X - xx^T \succcurlyeq 0 \Leftrightarrow \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succcurlyeq 0$$

Exemple de PSD non linéaire non formulable en PL

$$(P)\left\{\begin{array}{l}\text{Minimiser} \quad \dfrac{(c^T x)^2}{d^T x} \\ \text{s.c.} \;\; Ax + b \geqslant 0\end{array}\right.$$

Hypothèse: $Ax + b \geqslant 0$ implique $d^T x > 0$.

$$\left\{\begin{array}{l}\text{Minimiser} \quad t \\ \text{s.c.} \;\; Ax + b \geqslant 0 \\ \quad \dfrac{(c^T x)^2}{d^T x} \leqslant t\end{array}\right.$$

t : *complément de Schur*

$$\left\{\begin{array}{l}\text{Minimiser} \quad t \\ \text{s.c.} \quad \begin{bmatrix} \mathbf{diag}(Ax+b) & 0 & 0 \\ 0 & t & c^T x \\ 0 & c^T x & d^T x \end{bmatrix} \succcurlyeq 0\end{array}\right.$$

Exemple d'application : Résolution approchée d'un système linéaire

Soit $A \in \mathbb{R}^{p \times n}$, et $Ax = b$ n'admettant pas de solution. On décide de minimiser :

$$\max_{i \in \{1,\ldots,p\}} \left| A_i^T x - b_i \right|$$

avec A_i : ième ligne ($i \in \{1,\ldots,p\}$) de A

1. Modéliser ce problème comme un programme linéaire.
2. Dans certaines applications les quantités b_i sont exprimées à l'aide d'une échelle logarithmique (car elles représentent par exemple une puissance ou une intensité). On souhaite alors minimiser

$$\max_{i \in \{1,...,p\}} \left| log\left(A_i^T x\right) - log\left(b_i\right) \right|$$

   modéliser ce nouveau problème comme un programme semidéfini.

Programmation Quadratique Convexe

Proposition

La programmation Quadratique Convexe est un cas particulier de la Programmation Semidéfinie

$\forall i \in \{0, \ldots, m\}$ $f_i(x) = x^T A_i^T A_i x + c_i^T x + d_i$ est convexe

$$(Q) \begin{cases} \min & f_0(x) \\ \text{s.c.} & f_i(x) \leqslant 0 \quad i = 1, \ldots, m \end{cases}$$

Schur : $f_i(x) \leqslant 0 \Leftrightarrow \begin{bmatrix} I_n & A_i x \\ x^T A_i^T & -c_i^T x - d_i \end{bmatrix} \succcurlyeq 0$

$$(Q) \Leftrightarrow \begin{cases} \min & t \\ \text{s.c.} & \begin{bmatrix} I_n & A_0 x \\ x^T A_0^T & -c_0^T x - d_0 + t \end{bmatrix} \succcurlyeq 0 \qquad f_0(x) - t \leqslant 0 \\ & \begin{bmatrix} I_n & A_i x \\ x^T A_i^T & -c_i^T x - d_i \end{bmatrix} \succcurlyeq 0 \ \forall i \neq 0 \end{cases}$$

Programmation Quadratique Convexe (2)

$$(Q)\left\{\begin{array}{lll}\min & x^TA_0^TA_0x+c_0^Tx+d_0(\leqslant t) & \\ \text{s.c.} & x^TA_i^TA_ix+c_i^Tx+d_i\leqslant 0 & i=1,\dots,m\end{array}\right.$$

Utilisation d'une unique inégalité matricielle linéaire:

$$(Q)\Leftrightarrow\left\{\begin{array}{ll}\min & t \\ \text{s.c.} & \begin{bmatrix} d_0-t & \frac{1}{2}c_0^T \\ \frac{1}{2}c_0 & A_0^TA_0\end{bmatrix}\bullet\begin{bmatrix}1 & x^T \\ x & X\end{bmatrix}\leqslant 0 \\ & \begin{bmatrix} d_i & \frac{1}{2}c_i^T \\ \frac{1}{2}c_i & A_i^TA_i\end{bmatrix}\bullet\begin{bmatrix}1 & x^T \\ x & X\end{bmatrix}\leqslant 0 \\ & \begin{bmatrix}1 & x^T \\ x & X\end{bmatrix}\succcurlyeq 0\end{array}\right.$$

$$\underbrace{A_i^TA_i\bullet(X-xx^T)}_{\geqslant 0}+x^TA_i^TA_ix+c_i^Tx+d_i\leqslant 0$$

Utilisation de la même idée en PL

Montrons que :

$$
(PL)\left\{\begin{array}{ll}
\min & x_1 + 4x_2 \\
\text{Sous les contraintes} & x_1 - x_2 + x_3 = 2 \\
 & 3x_1 - 5x_3 = 6 \\
 & x_1 \geqslant 0\,,\ x_2 \geqslant 0\,,\ x_3 \geqslant 0
\end{array}\right.
$$

est équivalent à

$$
(SDP)\left\{\begin{array}{ll}
\text{Minimiser} & x_1 + 4x_2 \\
\text{Sous les contraintes} & \left(e_1 e_1^T + e_2 e_2^T\right) \bullet X - (4e_1 + 12e_2)^T x + 40 = 0 \\
 & \begin{bmatrix} \mathbf{diag}(x) & 0 & 0 \\ 0 & 1 & x^T \\ 0 & x & X \end{bmatrix} \succcurlyeq 0
\end{array}\right.
$$

où $e_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $e_2 = \begin{pmatrix} 3 \\ 0 \\ -5 \end{pmatrix}$

Dualité

- Le programme dual de (SDP) est
$$\sup_{Z \succcurlyeq 0} \inf_{x \in \Re^m} L(x, Z)$$
avec $L(x,Z) = c^T x + Z \bullet (A_0 - \sum_{i=1}^m x_i A_i)$
  - Le multiplicateur de Lagrange Z associé à l'inégalité matricielle linéaire $F(x) \succcurlyeq 0$ est une matrice positive
  - A comparer avec le lagrangien d'une inégalité linéaire $a^T x - b \geqslant 0$: $L(x, \lambda) = c^T x + \lambda^T (b - a^T x)$ avec $\lambda \geqslant 0$
- Fonction duale :
  - si $A_i \bullet Z = c_i \ \forall i \in \{1, \ldots, m\}$ alors $\inf_{x \in \Re^m} L(x, Z) = A_0 \bullet Z$
  - Sinon : les termes en x_i ne sont pas tous annulés et $\inf_{x \in \Re^m} L(x, Z) = -\infty$

Proposition

Le programme dual de (SDP) est un programme semidéfini :

$$(DSDP) \left\{ \begin{array}{lll} \sup & A_0 \bullet Z & \\ \text{s.c.} & A_i \bullet Z = c_i & i \in \{1, \ldots, m\} \\ & Z \succcurlyeq 0 & \end{array} \right.$$

Formes Duale et Primale

(DSDP) est un programme semidéfini

- Dans le primal : variable $x \in \Re^m$
- Dans le dual : variable $Z \in S_n$
- $F(x)$ et Z appartiennent à l'intersection d'un espace affine (pas le même !) et du cône des matrices positives
- Soit un repère $(-B_0, B_1, \ldots, B_p)$ (origine + base) de $\aleph_0 = \{Z\,;\, Z \in S_n,\ A_i \bullet Z = c_i,\ i = 1, \ldots, p\}$
- $Z = F(y) = -B_0 + \sum_{i=1}^{p} y_i B_i$

Réécriture de $(DSDP)$ sous forme primale

$$(DSDP) \left\{ \begin{array}{ll} \text{sup} & A_0 \bullet Z = A_0 \bullet (-B_0 + \sum_{i=1}^{p} y_i B_i) \\ \text{s.c. :} & A_i \bullet Z = c_i \ \ i = 1, \ldots, m \\ & \Leftrightarrow Z = F(y) = -B_0 + \sum_{i=1}^{p} y_i B_i \succcurlyeq 0 \end{array} \right.$$

Théorèmes de Dualité

Lemme

Soit A et B dans S_n. Si $A \succcurlyeq 0$ et $B \succcurlyeq 0$ alors $A \bullet B \geqslant 0$ et l'égalité est atteinte si et seulement si $AB = 0$

Dualité Faible

Le saut de dualité entre (DSP) et (SDP)
vaut:$c^T x - A_0 \bullet Z = \sum_{i=1}^{m} (A_i \bullet Z) x_i - A_0 \bullet Z = F(x) \bullet Z \geqslant 0$

Dualité Forte (qualification de type "Slater")

- Le problème primal est strictement réalisable
(il existe des points intérieurs : $\exists x \,/\, F(x) \succ 0$).
- Le problème dual est strictement réalisable
($\exists Z \,/\, Z = Z^T \succ 0,\ A_i \bullet Z = c_i,\ i = 1, \ldots, m$)
- Si les deux conditions sont remplies alors il n'y a pas de saut de dualité et il existe des solutions optimales pour le primal et le dual

Exemple d'application des théorèmes de dualité

$Inf \left\{ x : \begin{bmatrix} x & 1 \\ 1 & t \end{bmatrix} \succcurlyeq 0 \right\} = 0$ n'est pas atteint.

Dual de : $\quad Inf\ x\ s.c.\ : \begin{bmatrix} x & 1 \\ 1 & t \end{bmatrix} \succcurlyeq 0$?

$Max\ -2Z_{12}\ s.c.\ : Z_{11} = 1\,;\ Z_{22} = 0\,;\ Z \succcurlyeq 0$

Région admissible du programme dual :

$Z = \begin{bmatrix} 1 & Z_{12} \\ Z_{12} & 0 \end{bmatrix} \succcurlyeq 0 \Rightarrow -Z_{12}^2 \geqslant 0.$

=> Un seul point admissible : $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Application du Théorème :

- Le primal est d'intérieur non vide : pas de saut de dualité.
- Région admissible du dual d'intérieur vide : pas forcément de solution (c'est le cas du primal ici).

Conditions des écarts complémentaires en SDP

Proposition

Supposons que le saut de dualité est nul ($F(x) \bullet Z = 0$). x *et* Z *sont respectivement optimaux pour* (SDP) *et* $(DSDP)$ *si et seulement si:*

$$\begin{array}{l} F(x) \succcurlyeq 0 \\ Z \succcurlyeq 0,\ A_i \bullet Z = c_i,\ \ i = 1, \dots, m \\ ZF(x) = 0 \ \Leftrightarrow\ F(x) \bullet Z = 0 \end{array}$$

Généralisation du cas linéaire :

- En programmation linéaire $F(x) = \mathbf{diag}(Ax - b)$ et Z sont diagonales
- $ZF(x)$ s'écrit $Z_{ii}\mathbf{diag}(Ax - b)_{ii} = Z_{ii}\left(A_i^T x - b_i\right) = 0 \forall i$

Optimisation Semidéfinie vs Optimisation Linéaire

Les Programmes duaux respectifs appartiennent à la même classe de problèmes

$$(\mathrm{PL}) \quad \boxed{\begin{array}{ll} \text{minimiser} & c^T x \\ \text{sous contraintes} & a_i^T x = b_i, \;\; i = 1, \ldots, m \\ & x \geqslant 0 : \mathbb{R}^{n+} \end{array}}$$

$$(\mathrm{SDP}) \quad \boxed{\begin{array}{ll} \text{minimiser} & C \bullet X \\ \text{sous contraintes} & A_i \bullet X = b_i, \;\; i = 1, \ldots, m \\ & X \succcurlyeq 0 : \mathcal{S}_+^n \end{array}}$$

Produit scalaire et positivité

PL : $c^T x = \sum_i c_i x_i$ Cône $\mathbb{R}^{n+} = \{x \, : \, x^T y \geqslant 0 \, \forall y \in \mathbb{R}^{n+}\}$

SDP : $C \bullet X = \sum_{ij} C_{ij} X_{ij}$ Cône $\mathcal{S}_+^n = \{X \, : \, X \bullet Y \geqslant 0 \, \forall Y \in \mathcal{S}_+^n\}$

$\longrightarrow$ Chacun de ces cônes est son propre cône dual : le dual d'un PL est un PL, et le dual d'un SDP est un SDP.

Exemple d'application

$$(SDP)\begin{cases} \min_{z\in\Re} & z \\ \text{s.c.} & \begin{bmatrix} 3z & -z-1 \\ -z-1 & 3z \end{bmatrix} \succcurlyeq 0 \end{cases}$$

1. Résoudre (SDP), i.e. donner sa valeur optimale ainsi que l'ensemble des solutions optimales.
2. Ecrire $(DSDP)$, le programme dual de (SDP).
3. Montrer qu'il n'y a pas de saut de dualité entre $(DSDP)$ et (SDP), et résoudre $(DSDP)$ en utilisant les conditions des écarts complémentaires.

Exemple d'application (suite)

$$(SDP)\left\{\begin{array}{ll} \min_{z\in\Re} & z \\ \text{s.c.} & \begin{bmatrix} 3z & -z-1 \\ -z-1 & 3z \end{bmatrix} \succcurlyeq 0 \end{array}\right.$$

1. Soit le programme quadratique :

$$(P)\left\{\begin{array}{ll} \max & 2x_1x_2 \\ \text{s.c.} & 3x_1^2+3x_2^2-2x_1x_2=1 \\ & (x_1,x_2)\in\Re^2 \end{array}\right.$$

   Etablir que la valeur optimale de $(DSDP)$ est supérieure ou égale à celle de (P), en montrant que $(DSDP)$ est une relaxation de (P).

2. Donner le rang de la solution optimale de $(DSDP)$ trouvée à la question 3. En déduire une solution optimale de (P).