國立交通大學

應用數學系

碩士論文

半正定規劃

Semidefinite Programming Problems

研究生：葉彬

指導教授：翁志文 教授

中華民國九十九年六月

半正定規劃

Semidefinite Programming Problems

研究生：葉彬　　Student: Bin Yeh
指導教授：翁志文 教授　　Advisor: Chih-Wen Weng

國立交通大學
應用數學系
碩士論文

A Thesis
Submitted to Department of Applied Mathematics
College of Science
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master
in

Applied Mathematics

June 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年六月

半正定規劃

研究生：葉彬　　指導教授：翁志文

國立交通大學應用數學系碩士班

摘要

在半正定規劃的問題中，我們要求一些對稱矩陣的仿射組合必須是半正定，在這樣的限制下試圖將目標線性函數最小化。這些限制未必是線性，但它們具有中凸的性質故半正定規劃是一種中凸規劃。在這篇論文中我們探討了一些半正定規劃的基本性質與基礎理論並給出證明。

Semidefinite Programming Problems

Student: Bin Yeh Advisor: Chih-Wen Weng

Department of Applied Mathematics
National Chiao Tung University

Abstract

In semidefinite programming problems one minimizes a linear function subject to some constraints which requires an affine combination of symmetric matrices to be positive semidefinite. The constraints may not be linear but it is convex so semidefinite programming problems are convex optimization problems. In this paper we give some basic properties and fundamental theorems with their proofs regrading semidefinite programming problems.

誌謝

這一路上最感謝我的指導教授翁志文老師，因為我的不成熟而給他添了很多麻煩，實在是非常的對不起老師，在這邊要跟他說一聲對不起與謝謝您。

兩年多來還有其他很多很照顧我的教授們，包括傅恆霖老師、白啟光老師、黃大原老師、陳秋媛老師、王夏聲老師，真的非常感謝他們的指導與照顧。

我還要感謝我同師門的學長鎬文和喻培，當然還有其他給我許多幫助的學長姐，像是芳婷學姐、敏筠學姐、惠蘭學姐、國元學長、智懷學長，他們給了我很多的幫助。

也非常謝謝眾多好同學們，永潔、光祥、文昱、易萱、宜君、舜婷、玉雯、慧棻，以及研究室的夥伴們：育生、建文、思綸、健峰、思賢，還有我在求學過程中認識的許多學弟妹：侖欣、彥伶、權益、湧昇、昇華、育慈、定國、冠緯、閔豪、瑩晏、瑜堯、紹鵬、昌翰，謝謝你們陪我渡過快樂的時光。

最後我要感謝我的家人，爸爸，媽媽與弟弟。無論是什麼情況，他們總是支持著我、鼓勵著我，萬分的感謝你們。

要感謝的人很多，很幸運我能得到這麼多人的幫助。感恩之心難以言喻，就僅在此聊表感激之情，沒有各位就沒有今日的我。

Contents

1 Introduction

This paper is based on the survey paper by Vandenberghe and Boyd [1] and some details, mostly proofs, which are omitted in that survey. In this paper we focus on fundamental theory about semidefinite programming problem, rather than its application.

Semidefinite programming (SDP) is an optimization problem concerned with the optimization of a linear objective function over the intersection of the cone of positive semidefinite matrices. Linear programming, a well-used mathematical model which is used to obtain the best outcome with some certain restriction in the form of linear equations, is actually a special case of semidefinite programming problem. There are many different types of algorithms to solve semidefinite programming problem and these algorithms are capable of getting the result in polynomial time.

There are a lot of applications of semidefinite programming. In operations research and combinatorial optimization many problems are modeled as semidefinite programming problem so they can be well approximated. For example, Goemans and Williamson found an algorithm to obtain maximum cut using semidefinite programming [3].

In section 2 we give some basic definition about semidefinite programming problem and an example for better understanding. In section 3 the maybe most important property of semidefinite programming problem, duality, is discussed. In the further sections some more lemmas and propositions are given and eventually in section 5 a critical theorem is proved.

Some proofs in this paper are based on the lecture note written by László Lovász [2] and the book by Abraham Berman and Naomi Shaked-Monderer [4] and the note by Konstantin Aslanidi [5]. The main idea of proof of lemma 4.13 is provided by Renato Paes Leme.

2 Semidefinite programing problems

Throughout, we use $\mathbb{R}^n$ to denote the set of column vectors of size n and $\mathbb{R}^{n\times n}$ to denote the set of $n \times n$ symmetric matrices over $\mathbb{R}$. For a real matrix M, we use $M \geq 0$ when M has nonnegative entries, and use $M \geq_p 0$ when M is symmetric and positive semidefinite, i.e. $z^T M z \geq 0$ for all $z \in \mathbb{R}^n$. Similarly, we will denote positive definite by $M >_p 0$. The following is the general form of a **semidefinite programming problem**.

Problem 2.1. ($\mathrm{SDP}_n(c, F_0, F_1, \ldots, F_m)$)

Given a column vector $c \in \mathbb{R}^m$ and $m+1$ $n \times n$ symmetric matrices $F_0, F_1, \ldots, F_m \in \mathbb{R}^{n\times n}$. Find

$$\min_x \quad c^T x$$

for $x = [x_1, x_2, \ldots, x_m]^T \in \mathbb{R}^m$ subject to

$$F_0 + \sum_{i=1}^{m} x_i F_i \geq_p 0.$$

Give a $\mathrm{SDP}_n(c, F_0, F_1, \ldots, F_m)$ and let $F(x) = F_0 + \sum_{i=1}^{m} x_i F_i$. A vector $x \in \mathbb{R}^m$ is said to be **feasible** if $F(x) \geq_p 0$ and the set $\{x \in \mathbb{R}^m \mid F(x) \geq_p 0\}$ is called the **feasible region** of the problem. A vector x_{opt} in the feasible region is called an **optimal point** if $c^T x_{\text{opt}}$ reaches the minimum of $c^T x$ among all feasible points x. In this case the value $c^T x_{\text{opt}}$ is called the **minimum** of the problem. The number

$$\inf_x c^T x \in \mathbb{R} \cup \{-\infty\},$$

where the infimum is taking for feasible x, is called the **infimum** of the problem. If the infimum is not equal to the minimum, then $\text{SDP}_n(c, F_0, F_1, \ldots, F_m)$ is said to be **infeasible**. The following is a simple example of semidefinite programming problem.

Example 2.2. Consider

$$\text{SDP}_2(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}).$$

The feasible region is

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + x_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \geq_p 0 \right\}$$
$$= \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix} \geq_p 0 \right\}.$$

Note that $\begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix}$ is semidefinite if and only if both x_1 and $x_1x_2 - 1$ are nonnegative, i.e. $x_1 \geq 0$ and $x_1x_2 - 1 \geq 0$. Thus the feasible region is

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} | x_1 \geq 0, x_1x_2 \geq 1 \right\}. \tag{1}$$

The minimum of $(0,1)(x_1,x_2)^T = x_2$ does not exist in the feasible region since x_2 tends to zero as x_1 tends to infinity. However the infimum of this problem, obtained by above discussion, is zero. Thus, this example is infeasible.

The following is the general form of a **linear programming problem**, which we will show to be a special case of semidefinite programming problem.

Problem 2.3. ($\mathrm{LPP}_m(c,b,A)$)

Given $b, c \in \mathbb{R}^m$ and a symmetric matrix $A \in \mathbb{R}^{m\times m}$. Find

$$\min_x \quad c^T x$$

subject to all $x \in \mathbb{R}^m$ satisfying

$$Ax + b \geq 0.$$

Note that a linear programming problem is a special case of semidefinite programming problem. In fact by setting $F_0 = \mathrm{diag}(b)$, $F_i = \mathrm{diag}(A_i)$ for $i = 1, \ldots, m$, where A_i is the ith column of A, the $\mathrm{LPP}_m(c,b,A)$ becomes the $\mathrm{SDP}_n(c, F_0, F_1, \ldots, F_m)$.

Similar to the semidefinite programming problem, in $\mathrm{LPP}_m(c,b,A)$, a vector $x \in \mathbb{R}^m$ is said to be **feasible** if $Ax + b \geq 0$ and the set $\{x \in \mathbb{R}^m \mid Ax + b \geq 0\}$ is called the **feasible region**. A vector x_{opt} in the feasible region is called an **optimal point** if $c^T x_{\mathrm{opt}}$ reaches the minimum of $c^T x$ for all feasible points x.

We now show Example 2.2 is no way to be interpreted as a linear programming problem. To the contrary assume Example 2.2 is also a $LP_m(c,b,A)$.

Then from the definition $m = 2$ and A is a 2×2 matrix. Let V denote the nullspace of A. Then referring to (1) as the feasible region $\Omega := \{x \in \mathbb{R}^2 \mid Ax + b \geq 0\}$ also as a feasible region of $LP_m(c, b, A)$ we must have $\Omega + V \subseteq \Omega$. This only happens when $V = 0$, i.e. A is invertible. But then the feasible region of $LP_m(c, b, A)$ is $\{x \in \mathbb{R}^2 \mid x \geq -A^{-1}b\}$, which is clearly not to be Ω, a contradiction.

3 Dual problem

To solve a classical optimization problem, the problem and its dual problem play an important role. The following problem is the **dual** of $\mathrm{SDP}_n(c, F_0, F_1, \dots, F_m)$.

Problem 3.1. ($\mathrm{SDP}_n^*(c, F_0, F_1, \dots, F_m)$)

Given $c \in \mathbb{R}^m$ and $m+1$ symmetric matrices $F_0, F_1, \dots, F_m \in \mathbb{R}^{n\times n}$. Find

$$\max_Z -\mathrm{tr}(F_0 Z)$$

subject to all symmetric $n \times n$ matrices Z with

$$\begin{aligned} Z &\geq_p 0, \\ \mathrm{tr}(F_i Z) &= c_i \qquad \text{for} \quad 1 \leq i \leq m, \end{aligned} \tag{2}$$

where $\mathrm{tr}(M)$ is the trace of M.

Give a $\mathrm{SDP}_n^*(c, F_0, F_1, \dots, F_m)$, a symmetric matrix Z is said to be **feasible** if $Z \geq_p 0$ and $\mathrm{tr}(F_i Z) = c_i$ for $1 \leq i \leq m$. The set $\{Z \in \mathbb{R}^{n\times n} | Z \geq_p 0 \text{ and } \mathrm{tr}(F_i Z) = c_i \text{ for } 1 \leq i \leq m\}$ is called the **feasible region** of the problem. A symmetric matrix Z_{opt} in the feasible region is called an **optimal point** if $-\mathrm{tr}(F_0 Z)$ reaches the maximum of $-\mathrm{tr}(F_0 Z)$ among all feasible points Z. In this case the value $c^T x_{\mathrm{opt}}$ is called the **maximum** of the problem. The number

$$\sup_Z -\mathrm{tr}(F_0 Z) \in \mathbb{R} \cup \{\infty\},$$

where the supremum is taking for feasible Z, is called the **supremum** of the problem. If the supremum is not equal to the maximum, then $\mathrm{SDP}_n^*(c, F_0, F_1, \ldots, F_m)$ is said to be **infeasible**. The original semidefinite programming problem will be referred as the **primal problem**.

In Section 5, we will show that the supremum of $\mathrm{SDP}_n^*(c, F_0, F_1, \ldots, F_m)$ is no larger than the infimum of $\mathrm{SDP}_n(c, F_0, F_1, \ldots, F_m)$. This explains their dual relation. Here we give an example of dual semidefinite programming problem.

Example 3.2. The dual problem of

$$\mathrm{SDP}_2(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$$

in Example 2.2 is

$$\mathrm{SDP}_2^*(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}).$$

That is to maximize

$$-\mathrm{tr}(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} Z)$$

for all 2×2 symmetric matrices Z subject to

$$\operatorname{tr}(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} Z) = 0, \quad \operatorname{tr}(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} Z) = 1, \quad \text{and } Z \geq_p 0.$$

Let $Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}$. The first condition says

$$0 = tr(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}) = \operatorname{tr}(\begin{pmatrix} Z_{11} & Z_{12} \\ 0 & 0 \end{pmatrix}) = Z_{11},$$

and the second condition implies

$$1 = tr(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}) = Tr(\begin{pmatrix} 0 & 0 \\ Z_{21} & Z_{22} \end{pmatrix}) = Z_{22}.$$

Thus we only need to consider matrices of the form $\begin{pmatrix} 0 & Z_{12} \\ Z_{21} & 1 \end{pmatrix}$. The third condition $Z \geq_p 0$ is equivalent to

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} Z \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \geq 0$$

for all $x_1, x_2 \in \mathbb{R}$. As

$$0 \leq \begin{pmatrix} x_1 & x_2 \end{pmatrix} Z \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 Z_{21} & x_1 Z_{12} + x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= x_1 x_2 (Z_{21} + Z_{12}) + x_2^2 = 2 x_1 x_2 Z_{12} + x_2^2,$$

$Z \geq_p 0$ is equivalent to $Z_{12} = 0$. Then our feasible region is

$$\left\{ \begin{pmatrix} 0 & Z_{12} \\ Z_{21} & 1 \end{pmatrix} | Z_{21} = Z_{12} = 0 \right\}.$$

The goal is to maximize

$$-tr(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}) = -tr(\begin{pmatrix} Z_{21} & Z_{22} \\ Z_{11} & Z_{12} \end{pmatrix}) = -(Z_{21} + Z_{12}),$$

which is always zero in the feasible region. Then the maximum of

$$\text{SDP}_2^*(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$$

is 0 which, as we showed before, is also the infimum but not the minimum of

$$\text{SDP}_2(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}).$$

The general theory of relation between a primal problem and its dual problem will be given the section 5.

4 A few lemmas

Before proceeding, we introduce some basic properties of positive semidefinite matrices. Positive semidefinite matrices have some great properties in the way of geometry space. To see this we need some definition first. A **convex cone** C is a set of vectors in a vector space such that (i) for any vector v in C, rv is also in C for any $r > 0$, and (ii) for any two vectors v, $v' \in C$, $v + v' \in C$.

For any convex cone $C \subseteq \mathbb{R}^{n\times n}$, the **polar cone** C^* is defined by

$$C^* = \{A | A \in \mathbb{R}^{n\times n}, A \cdot B \geq 0 \text{ for all } B \in C\}$$

where $A \cdot B$ is the inner product for matrices, defined as

$$A \cdot B = tr(A^t B) = \sum_{1 \leq i,j \leq n} A_{i,j} B_{i,j}.$$

Note that a polar cone is also a convex cone, which is trivial to prove by applying the definition of convex cone.

Lemma 4.1. *A positive multiple of a positive semidefinite symmetric matrix is still positive semidefinite, and sum of two positive semidefinite symmetric matrices is still positive semidefinite.*

Proof. Let symmetric matrices $A, B \in \mathbb{R}^{n \times n}$ be positive semidefinite and $r > 0$. Then for any vector $x \in \mathbb{R}^n$, we have

$$x^T (rA) x = r(x^T Ax) \geq 0$$

and

$$x^T (A + B) x = (x^T Ax) + (x^T Bx) \geq 0,$$

which finished the proof. $\square$

The above lemma shows that the set of all symmetric positive semidefinite matrices P_n is actually a convex cone in $\mathbb{R}^{n \times n}$. For next lemma, we will use Hadamard product of matrices; that is, for two matrices $A, B \in \mathbb{R}^{n \times n}$, the Hadamard product $A \circ B$ is defined as

$$(A \circ B)_{ij} = A_{ij} B_{ij}$$

for all $1 \leq i, j \leq n$.

A property of Hadamard product will be given later. To prove the property we first show the following proposition.

Proposition 4.2. *Let A be a symmetric $n \times n$ real matrix. Then A is positive semidefinite if and only if there exists an $n \times n$ matrix B such that $A = BB^T$.*

Proof. ($\Leftarrow$)
For all $x \in \mathbb{R}^n$,

$$x^T A x = x^T (BB^T) x = (B^T x)^T (B^T x) = (B^T x) \cdot (B^T x) \geq 0.$$

($\Rightarrow$)
Since A is symmetric, it follows that A is normal and therefore A is unitary similar to a diagonal matrix, i.e. $PAP^T = D$ for some $n \times n$ matrix P with $P^T P = I$ and some diagonal matrix D. Then $A = P^T DP$. Note since A and D are similar, they have same eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and they are all nonnegative because A is positive semidefinite. Then we may assume $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Now let

$$D' = \begin{pmatrix} \sqrt{\lambda_1} & & & 0 \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \sqrt{\lambda_n} \end{pmatrix},$$

thus

$$A = P^T DP = P^T D' D' P = (P^T D')(P^T D')^T.$$

$\square$

Corollary 4.3. *Let A be symmetric $n \times n$ real matrix. Then A is positive semidefinite if and only if there exist vectors $v_1, v_2, ..., v_n \in \mathbb{R}^n$ such that $A = \sum_{i=1}^{n} v_i v_i^T$.*

Proof. $(\Rightarrow)$

By last lemma $A = BB^T$ for some $n \times n$ matrix B. Let v_i be the ith column vector of B for $1 \leq i \leq n$. Then

$$(A)_{ij} = (BB^T)_{ij} = \sum_{h=1}^{n} (v_h)_i (v_h)_j = \sum_{h=1}^{n} (v_h v_h^T)_{ij} = (\sum_{h=1}^{n} v_h v_h^T)_{ij}.$$

$(\Leftarrow)$

Let $A = \sum_{i=1}^{n} v_i v_i^T$. Let $B = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}$. Then

$$(A)_{ij} = (\sum_{h=1}^{n} v_h v_h^T)_{ij} = \sum_{h=1}^{n} (v_h v_h^T)_{ij} = \sum_{h=1}^{n} (v_h)_i (v_h)_j = (BB^T)_{ij}.$$

Then by last proposition, A is positive semidefinite.

□

Corollary 4.4. *The Hadamard product of two positive semidefinite matrices is still positive semidefinite.*

Proof. Let $A = \sum_{i=1}^{k} v_i v_i^T$ and $B = \sum_{j=1}^{l} w_j w_j^T$, then

$$A \circ B = \sum_{i=1,j=1}^{k,l} (v_i \circ w_j)(v_i \circ w_j)^T.$$

Therefore by last corollary $A \circ B$ is positive semidefinite. □

Now the lemma mentioned earlier can be proved.

Lemma 4.5. *The polar cone of P_n is itself, i.e. $P_n^* = P_n$.*

Proof. Let symmetric matrices $A, B \in \mathbb{R}^{n \times n}$ be positive semidefinite. Note

$$tr(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \sum_{1 \leq j \leq n} A_{ij} B_{ji} = \sum_{i=1}^{n} \sum_{1 \leq j \leq n} A_{ji} B_{ji} = e^T (A \circ B) e$$

where $e \in \mathbb{R}^n$ is the all-1 vector. Then by corollary 4.4, $A \circ B$ is still positive semidefinite and then $e^T(A \circ B)e \geq 0$. Therefore, $P_n^* \supseteq P_n$.

On the other hand, suppose $A \in P_n^*$. For any column vector $x \in \mathbb{R}^n$, the $n \times n$ symmetric matrix xx^T is positive semidefinite since

$$y^T(xx^T)y = (x^Ty)^T(x^Ty) = (\sum_{i=1}^{n} x_iy_i)(\sum_{i=1}^{n} x_iy_i) = (\sum_{i=1}^{n} x_iy_i)^2 \geq 0$$

for any $y \in \mathbb{R}^n$. Thus

$$xAx^T = tr(x(Ax^T)) = tr(Ax^Tx) \geq 0$$

for any $x \in \mathbb{R}^n$. Then $A \in P_n$.

□

So the polar cone of P_n is itself, which leads to the next proposition.

Proposition 4.6. *A symmetric matrix A is positive semidefinite if and only if $tr(AB) \geq 0$ for all symmetric positive semidefinite matrix B.*

We immediately have the following corollary.

Corollary 4.7. *Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric matrices. If $A, B \geq_p 0$, then $tr(AB) \geq 0$.*

Proposition 4.8. *Let $A \in P_n$ and $B \in \mathbb{R}^{n\times n}$ be a positive definite matrix. Then $tr(AB) = 0$ if and only if $A = 0$.*

Proof. It's clearly $tr(0B) = 0$. On the other hand, since $A \geq_p 0$ by taking e_i, i.e. the vector with 1 at the ith entry and zero for the rest, we have

$$A_{ii} = e_i^TAe_i \geq 0.$$

Furthermore there exists a $n \times n$ matrix P such that $P^tP = PP^t = I$ and $A' = P^tAP = diag(\lambda_1, ..., \lambda_n)$ is diagonal. Note A' is still positive semidefinite. Then

$$0 = tr(AB) = tr(P^tPAB) = tr(P^tABP) = tr(P^tAPP^tBP)$$

$$= tr(A'P^tBP) = \sum_{1 \le i \le n} \lambda_i (P^tBP)_{ii} = \sum_{1 \le i \le n} \lambda_i (e_i^t P^tBPe_i)$$

$$= \sum_{1 \le i \le n} \lambda_i ((Pe_i)^t B(Pe_i)) \ge 0$$

Since $(Pe_i)^tB(Pe_i) > 0$ for all $1 \le i \le n$, it follows that $\lambda_i = 0$ for all $1 \le i \le n$, $A' = 0$. Therefore $A = 0$.

□

Now we need something else to prove next lemma. A set $C \subseteq \mathbb{R}^n$ is convex if

$$(1-t)x + ty \in C \text{ for all } x, y \in C, t \in [0,1].$$

Let $|x_0| = \sqrt{x \cdot x}$ for all $x \in \mathbb{R}^n$, where $x \cdot x$ is the inner product of x and x. For any $C \subseteq \mathbb{R}^n$, let $cl(C) = \{x | x \in \mathbb{R}^n$ and for any $r > 0,$ there is a $y \in C$ such that $|x - y| < r\}$. A subset M of $\mathbb{R}^n$ is called an affine set if $M = \{x | Bx = b\}$ for some $B \neq 0, B \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$

A hyperplane in $\mathbb{R}^n$ is a subset which can be written as $\{x | x \in \mathbb{R}^n, x \cdot b = \beta\}$ for some $\beta \in \mathbb{R}$ and $b \neq 0, b \in \mathbb{R}^n$. A hyperplane separates $\mathbb{R}^n$ into two parts $\{x | x \in \mathbb{R}^n, x \cdot b \ge \beta\}$ and $\{x | x \in \mathbb{R}^n, x \cdot b < \beta\}$. The sets C_1 and C_2 are separated if there exist a hyperplane H such that C_1 and C_2 are in different parts which H separates $\mathbb{R}^n$ into.

Proposition 4.9. *Let $C \subset \mathbb{R}^n$ be a convex set and $x_0 \in \mathbb{R}^n$ with $x_0 \notin cl(C)$. Then there is a hyperplane H separates C and x_0.*

Proof. Since $\mathbb{R}^n$ is compact, we have

$$|y^* - x_0| = \inf_{y \in C} |y - x_0|$$

for some $y^* \in cl(C)$. Note $|y^* - x_0| \neq 0$ since $x_0 \notin cl(C)$. Let

$$h(x) = \frac{1}{|y^* - x_0|^2}(2(x \cdot (y^* - x_0)) + |x_0|^2 - |y^*|^2)$$

Then

$$\begin{aligned} h(x_0) &= \frac{1}{|y^* - x_0|^2}(2(x_0 \cdot y^*) - 2(x_0 \cdot x_0) + |x_0|^2 - |y^*|^2) \\ &= \frac{1}{|y^* - x_0|^2}(2(x_0 \cdot y^*) - (x_0 \cdot x_0) - |y^*|^2) \\ &= \frac{1}{|y^* - x_0|^2}(2(x_0 \cdot y^*) - (x_0 \cdot x_0) - (y^* \cdot y^*)) \\ &= \frac{-1}{|y^* - x_0|^2}(x_0 - y^*) \cdot (x_0 - y^*) = -1 \end{aligned}$$

And

$$\begin{aligned} h(y^*) &= \frac{1}{|y^* - x_0|^2}(2(y^* \cdot (y^* - x_0)) + |x_0|^2 - |y^*|^2) \\ &= \frac{1}{|y^* - x_0|^2}(2(y^* \cdot y^*) - 2(y^* \cdot x_0) + |x_0|^2 - |y^*|^2) \\ &= \frac{1}{|y^* - x_0|^2}((y^* \cdot y^*) - 2(y^* \cdot x_0) + (x_0 \cdot x_0)) \\ &= \frac{1}{|y^* - x_0|^2}(y^* - x_0) \cdot (y^* - x_0) = 1 \end{aligned}$$

Thus the hyperplane $\{x | h(x) = 0\}$ separates C and x_0 since C is convex. □

Proposition 4.10. *Let $C \subseteq \mathbb{R}^n$ be a convex set. Let $M \subseteq \mathbb{R}^n$ be a nonempty affine set with $C \cap M = \emptyset$. Then there is an hyperplane H containing M such that C is contained in one of the two parts that H separates $\mathbb{R}^n$ into.*

Proof. By induction on $dim(M)$. First let S be a subspace with $M = S + a$ for some $a \in \mathbb{R}^n$. If $dim(S) = n - 1$ then M itself is such a hyperplane, so we are done.

Suppose it holds for all S with dimension larger than $k < n - 1$ for some k. When $dim(S) = k - 1$, we have $dim(S^{\perp}) \geq 2$ and therefore contains a subspace T of dimension 2. The set $C - M = \{x - y | x \in C, y \in M\}$ does not contain 0, so we can find a subset $L \subset T$ of dimension 1 such that $0 \in L$ and $L \cap (C - M) = \emptyset$. We now can add the basis of L into the basis of S obtaining a new subspace S'. Then by induction hypothesis, there is a hyperplane H containing S' and C is contained in one of two parts that H separates $\mathbb{R}^n$ into. Thus we are done. □

Proposition 4.11. *Let $C_1, C_2 \subseteq \mathbb{R}^n$ be nonempty convex sets with $C_1 \cap C_2 = \emptyset$. Then they are separated.*

Proof. Taking $C = C_1 - C_2 = \{x - y | x \in C_1, y \in C_2\}$ and $M = \{0\}$ in last proposition. Thus there is a hyperplane H contains 0 and $H \cap C = \emptyset$. Let $H = \{x | x \in \mathbb{R}^n, x \cdot b = 0\}$ for some $b \in \mathbb{R}^n$. By using $-b$ to replace b if necessary, we have

$$\inf_{x \in C} x \cdot b \geq 0,$$

$$\sup_{x \in C} x \cdot b > 0.$$

Then

$$0 \leq \inf_{x \in C} x \cdot b = \inf_{x_1 \in C_1} x_1 \cdot b - \sup_{x_2 \in C_1} x_2 \cdot b,$$

that is

$$\inf_{x_1 \in C_1} x_1 \cdot b \geq \sup_{x_2 \in C_1} x_2 \cdot b.$$

Then C_1 and C_2 are separated by hyperplane

$$H' = \{x | x \cdot b = \inf_{x_1 \in C_1} x_1 \cdot b\}.$$

$\square$

The next lemma is well-known as the semidefinite version of Farkas' Lemma, which is a similar theorem regarding linear programming.

Lemma 4.12 (Homogenous Version)**.** *Let A_1, A_2, ..., A_m be symmetric matrices in $\mathbb{R}^{n \times n}$. Then the system*

$$x_1 A_1 + \ldots + x_m A_m >_p 0$$

has no solution in x_1, x_2, ..., x_m if and only if there exists a symmetric matrix $Y \neq 0$ such that $A_i \cdot Y = tr(A_i Y) = 0$ for all $1 \leq i \leq m$ and $Y \geq_p 0$.

Proof. $(\Rightarrow)$

Since the system has no solution, we have that

$$\{\sum_i x_i A_i | x_i \in \mathbb{R}\} \cap int(P_n) = \emptyset$$

where $int(P_n)$ is the interior of P_n. Moreover, by directly checking the definition, $\{\sum_i x_i A_i | x_i \in \mathbb{R}\}$ is a convex cone and therefore a convex set. We can consider $\mathbb{R}^{n \times n}$ as $\mathbb{R}^{n^2}$, thus by last lemma, there is a hyperplane

$$H = \{x | x \in \mathbb{R}^{n \times n}, x \cdot Y = \beta\}$$

separates $\{\sum_i x_i A_i | x_i \in \mathbb{R}\}$ and P_n for some $Y \in \mathbb{R}^{n \times n}$ and $\beta \in \mathbb{R}$, which we may assume

$$\beta = \inf_{P \in P_n} P \cdot Y$$

by the definition of hyperplane. Furthermore, we may assume

$$\inf_{P \in P_n} P \cdot Y \geq 0$$

by replacing Y with $-Y$ if necessary. But since P_n is a convex cone, for any $x \in P_n$, $rx \in P_n$ for all $r > 0$. By taking r small enough we have

$$\inf_{P \in P_n} P \cdot Y = 0$$

Then our hyperplane

$$H = \{x | x \cdot Y = 0\}$$

implies that $(\sum_i x_i A_i) \cdot Y \leq 0$ and $P \cdot Y > 0$ for all $P \in int(P_n)$. Now for a fixed $1 \leq i \leq m$, let $x_i = 1$ and $x_j = 0$ for all $i \neq j$. Then

$$0 \geq (\sum_i x_i A_i) \cdot Y = A_i \cdot Y.$$

We can also let $x_i = -1$ and $x_j = 0$ for all $i \neq j$. Then

$$0 \geq (\sum_i x_i A_i) \cdot Y = -A_i \cdot Y.$$

These two inequalities show that $A_i \cdot Y = 0$ for all $1 \leq i \leq m$. On the other hand, $P \cdot Y \geq 0$ for all $P \in P_n$. Then by proposition 4.6, $Y \geq_p 0$.

($\Leftarrow$)

Suppose there is a solution of $x_1 A_1 + \ldots + x_m A_m >_p 0$. Then

$$(\sum_i x_i A_i) \cdot Y = \sum_i x_i (A_i \cdot Y) = 0$$

By proposition 4.8, $Y = 0$, a contradiction.

$\square$

Lemma 4.13 (Nonhomogeneous Version)**.** *Let A_1, A_2, …, A_m, B be symmetric matrices in $\mathbb{R}^{n\times n}$. Then the system*

$$x_1A_1 + \ldots + x_mA_m - B >_p 0$$

has no solution in x_1, x_2, …, x_m if and only if there exists a symmetric matrix $Y \neq 0$ such that $tr(A_iY) = 0$ for all $1 \leq i \leq m$, $tr(BY) \geq 0$ and $Y \geq_p 0$.

Proof. This is done by applying last lemma to following matrices

$$\begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} A_k & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix}.$$

We claim that if the system

$$x_1\begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} + x_2\begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix} + \ldots + x_k\begin{pmatrix} A_k & 0 \\ 0 & 0 \end{pmatrix} + x_{k+1}\begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} >_p 0$$

has a solution, then so does the system

$$x_1A_1 + \ldots + x_mA_m - B >_p 0.$$

To prove this claim, suppose $v_1, \ldots, v_k, v_{k+1}$ is a solution of

$$x_1\begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} + x_2\begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix} + \ldots + x_k\begin{pmatrix} A_k & 0 \\ 0 & 0 \end{pmatrix} + x_{k+1}\begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} >_p 0.$$

Let

$$M = (\sum_{1\leq i\leq k} v_i \begin{pmatrix} A_i & 0 \\ 0 & 0 \end{pmatrix}) + v_{k+1}\begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix}.$$

Then $M >_p 0$. Thus

$$0 < e_{k+1}^t M e_{k+1} = v_{k+1}.$$

Let $u_i = \frac{v_i}{v_{k+1}}$ for all $1 \le i \le k+1$, and

$$M' = (\sum_{1\le i\le k} u_i \begin{pmatrix} A_i & 0 \\ 0 & 0 \end{pmatrix}) + \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{v_{k+1}} M$$

which is still positive definite. Note that all principal submatrices of a positive definite matrix is again positive definite. Therefore the following principal submatrix of M'

$$(\sum_{1\le i\le k} u_i A_i) - B >_p 0.$$

Thus $u_1, u_2, ..., u_k$ is a solution of

$$x_1 A_1 + \ldots + x_m A_m - B >_p 0.$$

($\Rightarrow$)

The claim we proved is equivalent to that if

$$x_1 A_1 + \ldots + x_m A_m - B >_p 0$$

has no solution, then either the system

$$x_1 \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix} + \ldots + x_k \begin{pmatrix} A_k & 0 \\ 0 & 0 \end{pmatrix} + x_{k+1} \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} >_p 0.$$

So by previous lemma, there exists $Y \in \mathbb{R}^{(n+1)\times(n+1)}$ such that $Y \ge_p 0$ and

$$\begin{pmatrix} A_i & 0 \\ 0 & 0 \end{pmatrix} \cdot Y = 0 \text{ for all } 1 \le i \le k, \text{ and } \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} \cdot Y = 0.$$

Let

$$Y = \begin{pmatrix} Y' & y \\ y^T & y_0 \end{pmatrix}$$

for some $y \in \mathbb{R}^n$, $y_0 \in \mathbb{R}$, then

$$A_i \cdot Y' = 0 \text{ for all } 1 \leq i \leq k.$$

Since $Y \geq_p 0$, we take e_{n+1} and we have

$$0 \leq e_{n+1}^T Y e_{n+1} = y_0.$$

Together with

$$0 = \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} \cdot Y = -BY' + y_0,$$

we conclude that

$$BY' = y_0 \geq 0.$$

Note $Y' \geq_p 0$ since Y' is a principal submatrix of $Y \geq_p 0$.

($\Leftarrow$)

By assumption there exists $Y \neq 0$ such that $A_i \cdot Y = tr(A_i Y) = 0$ for all $1 \leq i \leq m$, $Y \geq_p 0$ and $tr(BY) \geq 0$. Let

$$Z = \begin{pmatrix} Y & 0 \\ 0 & B \cdot Y \end{pmatrix}.$$

Suppose there is a solution of $x_1 A_1 + \ldots + x_m A_m - B >_p 0$. Then so does the system

$$x_1 \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix} + \ldots + x_k \begin{pmatrix} A_k & 0 \\ 0 & 0 \end{pmatrix} + x_{k+1} \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} >_p 0.$$

But

$$Z \cdot \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} Y & 0 \\ 0 & B \cdot Y \end{pmatrix} \cdot \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} = 0 \text{ for all } 1 \leq i \leq m,$$

and

$$Z \cdot \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} Y & 0 \\ 0 & B \cdot Y \end{pmatrix} \cdot \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} = -B \cdot Y + B \cdot Y = 0.$$

Note $Z \geq_p 0$. However, since the system

$$x_1 \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix} + \ldots + x_k \begin{pmatrix} A_k & 0 \\ 0 & 0 \end{pmatrix} + x_{k+1} \begin{pmatrix} -B & 0 \\ 0 & 1 \end{pmatrix} >_p 0$$

has solution, such Z should not exist by last lemma, a contradiction.

□

5 Property of Semidefinite programming Problems

A $\mathrm{SDP}_n(c, F_0, F_1, \ldots, F_m)$ is said to have **feasible solution** if there exists x with $F(x) \geq_p 0$, and is said to be **strictly feasible** if $F(x) >_p 0$ for some x in the feasible region, where as before $F(x) = F_0 + \sum_{i=1}^{m} x_i F_i$. In this case x is called a **strictly feasible solution**. Similarly $\mathrm{SDP}^*{}_n(c, F_0, F_1, \ldots, F_m)$ is said to have **feasible solution** if there is some Z in the feasible region, and Z is said to be **strictly feasible** if $Z >_p 0$. In this case Z is called a **strictly feasible solution**.

Theorem 5.1. *Let p^* and d^* be the infimum of* $\mathrm{SDP}_n(c, F_0, F_1, \ldots, F_m)$ *and supremum of* $\mathrm{SDP}^*{}_n(c, F_0, F_1, \ldots, F_m)$ *respectively, and assume $p^*, d^* < \infty$. Then $p^* \geq d^*$. Moreover, suppose either of the following conditions (i)-(ii) holds.*

(i) The primal problem is strictly feasible.

(ii) The dual problem is strictly feasible.

Then $p^* = d^*$.

Proof. Let x be a vector in the feasible region of primal problem and Z be a symmetric matrix in the feasible region of its dual problem. Then referring to (2) and by corollary 4.7,

$$c^T x + tr(ZF_0) = \sum_{i=1}^{m} tr(ZF_i x_i) + tr(ZF_0) = tr(ZF(x)) \geq 0.$$

Thus $c^T x \geq -tr(ZF_0)$ which shows $p^* \geq d^*$.

Now the system

$$c^T x < p^*$$
$$F(x) \geq_p 0$$

has no solution $x \in \mathbb{R}^m$ by the definition of p^*. Therefore if we define the matrices

$$F_0' = \begin{pmatrix} p^* & 0 \\ 0 & F_0 \end{pmatrix} \text{ and } F_i' = \begin{pmatrix} -c_i & 0 \\ 0 & F_i \end{pmatrix} \text{ for } 1 \leq i \leq m,$$

$F_0' + x_1 F_1' + \ldots + x_m F_m' >_p 0$ has no solution in $\mathbb{R}^n$. Thus by lemma 4.13 there is a positive semidefinite matrix $Y \neq 0$ such that

$$tr(F_0' Y) \geq 0 \text{ and } tr(F_i' Y) = 0 \text{ for } 1 \leq i \leq m.$$

By letting

$$Y = \begin{pmatrix} y_{00} & y \\ y^T & Z \end{pmatrix} \text{ for some } y_{00} \in \mathbb{R} \text{ and } y \in \mathbb{R}^n,$$

we obtain

$$tr(F_0 Z) \geq y_{00} p^* \text{ and } tr(F_i Z) = y_{00} c_i \text{ for } 1 \leq i \leq m.$$

We claim that $y_{00} \neq 0$. Suppose not. Then

$$tr(F_0 Z) \geq 0 \text{ and } tr(F_i Z) = 0 \text{ for } 1 \leq i \leq m.$$

Therefore by lemma 4.13, the existence of Z implies $F_0 + x_1 F_1 + \ldots + x_m F_m >_p 0$ has no solution in $\mathbb{R}^n$, contradicts to the hypothesis that primal problem is strictly feasible. Thus, $y_{00} \neq 0$.

Now $y_{00} \neq 0$ and since Y is positive semidefinite, $y_{00} > 0$. By scaling we may assume $y_{00} = 1$. But then Z satisfies $tr(ZF_0) \geq y_{00} p^* = p^*$ and $d^* \geq tr(ZF_0)$ by definite of d^*, thus $d^* \geq p^*$. Together with first part of the proof we have $p^* = d^*$.

The case that condition (ii) holds is similar to prove. Therefore the proof is complete.

$\square$

6 References

[1] Lieven Vandenberghe and Stephen Boyd, *Semidefinite Programming*, SIAM Review, Vol. 38, No. 1. (Mar., 1996), pp. 49-95.

[2] László Lovász, *Semidefinite programs and combinatorial optimization*, lecture notes, 1995-2001.

[3] Michel X. Goemans and David P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM, Vol. 42, pp. 1115-1145, 1995.

[4] Abraham Berman and Naomi Shaked-Monderer, *Completely positive matrices*, World Scientific Publishing Company, 2003.

[5] Konstantin Aslanidi, *Notes on Quantitative Analysis in Finance*, Available: http://www.opentradingsystem.com, 2007.