A *Short Course* on Semidefinite Programming

(in order of appearance)

Henry Wolkowicz, Etienne de Klerk, Levent Tunçel, Franz Rendl

hwolkowicz@uwaterloo.ca

Department of Combinatorics and Optimization

University of Waterloo

CONCORDIA CUM VERITATE

About these Notes:

Semidefinite Programming, SDP, refers to optimization problems where the vector variable is a symmetric matrix which is required to be positive semidefinite. Though SDPs (under various names) have been studied as far back as the 1940s, the interest has grown tremendously during the last fifteen years. This is partly due to the many diverse applications in e.g. engineering, combinatorial optimization, and statistics. Part of the interest is due to the great advances in efficient solutions for these types of problems.

These notes summarize the theory, algorithms, and applications for semidefinite programming. They are prepared for a shortcourse given on the first day of the **SIAM Conference on Optimization**, May 15-19, 2005, at City Conference Centre, Stockholm, Sweden, URL: www.siam.org/meetings/op05/index.htm

Preface

The purpose of these notes is three fold: first, they provide a comprehensive treatment of the area of Semidefinite Programming, a new and exciting area in Optimization; second, the notes illustrate the strength of convex analysis in Optimization; third, they emphasize the interaction between theory and algorithms and solutions of practical problems.

1 Introduction and Motivation

1.1 Outline

- Basic Properties and Notation
- Examples/Applications
- Historical Notes

1.2.1 Basic Properties and Notation

Basic <u>linear</u> ***Semidefinite Programming*** looks just like *Linear Programming*

$$
(\mathbf{PSDP}) \qquad \begin{array}{rlll} p^* = & \max & \operatorname{trace} CX & (\langle C, X\rangle) \\ & \text{s.t.} & \mathcal{A}(X) \cong \mathcal{A}X = b & \text{(linear)} \\ & & X \succeq 0, \quad (X \in \mathcal{P}) & \text{(nonneg)} \end{array}
$$

$C, X \in \mathcal{S}^n, \quad b \in \mathbb{R}^m$

$\mathcal{S}^n$:= space of $n \times n$ symmetric matrices

$\mathcal{A}$:= linear operator

Space of Symmetric Matrices

$A = A^T \in \mathcal{S}^n$:= space of $n \times n$ (real) symmetric matrices

A is positive semidefinite (positive definite), $\quad (A \succeq 0$ $(A \succ 0))$
if $x^T A x \geq 0 (> 0), \; \forall x \neq 0$.

$\preceq$ denotes the Löwner partial order, [54]
$A \preceq B$ if $B - A \succeq 0$ (positive semidefinite)

TFAE:

1. $A \succeq 0 \quad (A \succ 0)$
2. the vector of eigenvalues $\lambda(A) \geq 0 \quad (\lambda(A) > 0)$
3. all principal minors ≥ 0 $\quad$ (all leading principal minors > 0)

Some Properties/Equivalences of $\mathcal{S}^n, \mathcal{P}$

1. $A \in \mathcal{S}^n$:
   (a) All eigenvalues are real
   (b) $A = P\Lambda P^T, \; P^T P = I, \; \Lambda = \mathrm{Diag}\,(\lambda)$, i.e. A can be orthogonally diagonalized
2. $A \succeq 0$:
   (a) All eigenvalues are real nonnegative
   (b) $A = S^2, \; S \succeq 0$, i.e. A has a square root in $\mathcal{P}$
3. $A \succ 0$:
   (a) All eigenvalues are real positive
   (b) $A = S^2, \; S \succ 0$, i.e. A has a square root in $\operatorname{int} \mathcal{P}$
   (c) $A = LL^T$, L lower triangular with positive diagonal (Cholesky factorization)

Linear transformation $\mathcal{A}$; Adjoint:

$$\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m, \qquad \mathcal{A}^* : \mathbb{R}^m \to \mathcal{S}^n$$

$$\mathcal{A} \text{ defined by associated set } \{A_i \in \mathcal{S}^n, i = 1, \ldots m\} :$$
$$(\mathcal{A}X)_i = \operatorname{trace}(A_i X); \quad \mathcal{A}^* y = \sum_{i=1}^m y_i A_i,$$

where the adjoint of $\mathcal{A}$ is defined by

$$\langle \mathcal{A}X, y \rangle = \langle X, \mathcal{A}^* y \rangle, \quad \forall X \in \mathcal{S}^n, \forall y \in \mathbb{R}^m$$

$\mathcal{P} = \mathcal{S}^n_+$ - cone of positive semidefinite matrices
replaces
$\mathbb{R}^n_+$ - nonnegative orthant

SDP or LMI

$\mathrm{trace}\, CX = \langle C, X \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} = (\mathrm{vec}\, C)^T (\mathrm{vec}\, C)$

(primal) SDP is equivalent to (Linear Matrix Inequalities):

$$
\begin{array}{llll}
 & p^* = & \max & \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} & \\
(\mathbf{PSDP}) & & \text{s.t.} & \langle A_i, X \rangle = b_i, & i = 1, \ldots m \\
 & & & X \succeq 0, \quad (X \in \mathcal{P}) &
\end{array}
$$

1.2.2 Application - Max-Cut Problem, MC

(MC) is a combinatorial optimization problem on undirected graphs with weights on the edges.

Problem 1.2.2.1 *Find a partition of the set of vertices into two parts that maximizes the sum of the weights on the edges that have one end in each part of the partition.*

One Formulation of MC

vector $v \in \{\pm 1\}^n$, $n = |V|$, represents cut in graph G

$$S = \{i : v_i = +1\} \qquad V \backslash S = \{i : v_i = -1\}.$$

$$(MC) \qquad \begin{array}{rll} \mu^* = & \max & \sum_{1 \le i < j \le n} w_{ij} \left(\frac{1 - v_i v_j}{2} \right) \\ & \text{s.t.} & v \in \{\pm 1\}^n \end{array}$$

Equivalently,

$$(\text{MC1}) \qquad \begin{array}{rll} \mu^* = & \max & v^T Q\, v \\ & \text{s.t.} & v_i^2 = 1, \qquad i = 1, \ldots, n, \end{array}$$

where $Q = \frac{1}{4}(\text{Diag}\,(A\,e) - A)$ is $\frac{1}{4}$ Laplacian matrix, L, $A = (w_{ij})$ is weighted adjacency matrix of G.

Equivalent Formulation of MC

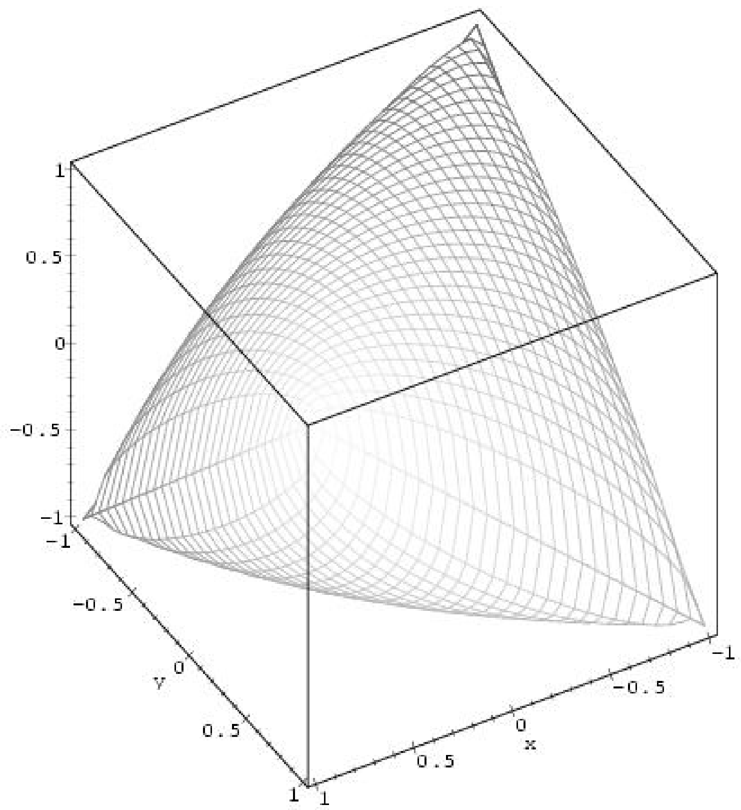
With $Q := \frac{1}{4}L$, $X := vv^T, v \in \{\pm 1\}^n$,
then $v^T Qv = \operatorname{trace} QX$ and equivalent formulation is:

$$
\text{(MC1)} \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & \operatorname{rank}(X) = 1 \\ & & X \succeq 0, X \in \mathcal{S}^n, \end{array}
$$

relax by deleting the ***hard*** constraint $\operatorname{rank}(X) = 1$ (get *elliptope*)

Note $X \succeq 0$, rank r iff $X = VV^T$ with $V\, n \times r$, full rank

Elliptope for $n = 3$, [50]



GW Approximation Result

Goemans & Williamson proved that
if $w_{ij} \geq 0 \ \forall \, i, j$, then

$$\mu^* \geq \alpha \, \nu_1^*,$$

where $\alpha = \min_{0 \leq \theta \leq \pi} \frac{2}{\pi} \frac{\theta}{1 - \cos\theta} \approx 0.87856$.

Since $\frac{1}{\alpha} \approx 1.13823 \leq 1.14$, this implies

$$\mu^* \leq \nu_1^* \leq 1.14 \, \mu^*.$$

Other Approximations

Other such convex relaxations have been studied before.

The smallest convex set containing all the rank-one matrices X corresponding to cuts is the *cut polytope*:

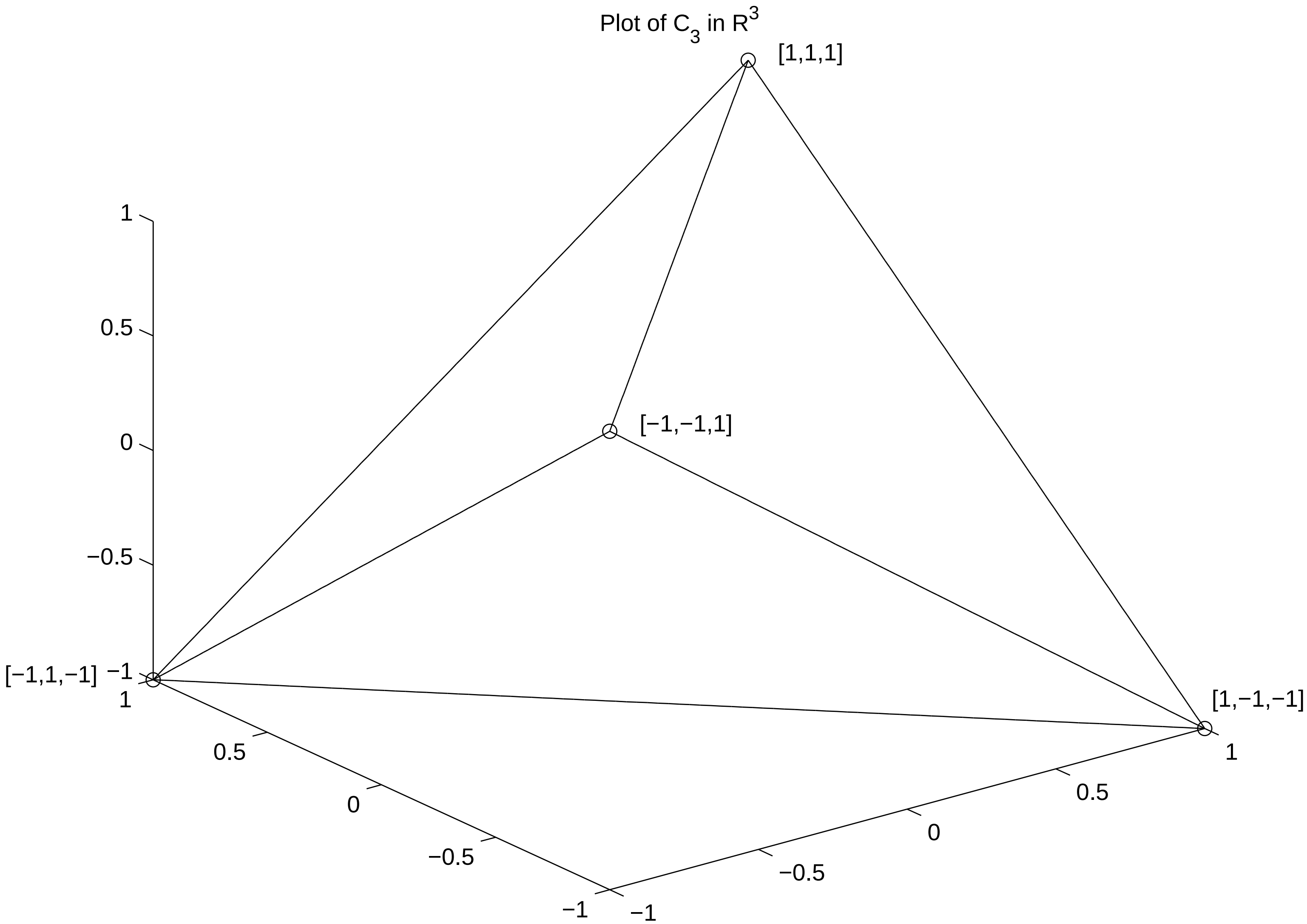
$$C_n := \text{Conv}\{X : X = vv^T, v \in \{\pm 1\}^n\}.$$

In fact,

$$\begin{aligned} \mu^* = \quad & \max \quad \operatorname{trace} QX \\ & \text{s.t.} \quad X \in C_n. \end{aligned}$$

However, it is not known how to optimize in polynomial-time over C_n.

Cut polytope for $n = 3$



Metric Polytope

Another convex relaxation is the *metric polytope* M_n, defined by

$$M_n := \{X \in \mathcal{S}^n : \mathrm{diag}\,(X) = e, \text{ and}$$
$$X_{ij} + X_{ik} + X_{jk} \geq -1, X_{ij} - X_{ik} - X_{jk} \geq -1,$$
$$- X_{ij} + X_{ik} - X_{jk} \geq -1, -X_{ij} - X_{ik} + X_{jk} \geq -1,$$
$$\forall\, 1 \leq i < j < k \leq n\}.$$

These inequalities model the fact that for any three mutually connected vertices in the graph, it is only possible to cut either zero or two of the edges joining them.

It is known that:

$C_n = M_n$ for $n \leq 4$, but $C_n \subsetneq M_n$ for $n \geq 5$.

Primal-Dual Pair

$$\text{(MCSDP)} \quad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & X \succeq 0, X \in \mathcal{S}^n, \end{array}$$

$$\text{(DMCSDP)} \quad \begin{array}{rll} \mu^* = & \min & e^T y \\ & \text{s.t.} & \operatorname{Diag}(y) \succeq Q \\ & \text{equivalently:} & \operatorname{Diag}(y) - Z = Q \\ & & Z \succeq 0, Z \in \mathcal{S}^n \end{array}$$

See Appendix B, Page 69-19, for information and examples on solving relaxations of Max-Cut problems using NEOS or MATLAB.

Rewrite Lagrangian/payoff

payoff function; player Y to player X (Lagrangian)

$$L(X,y) := \operatorname{trace}(CX) + y^T(b - \mathcal{A}X) \;\; (= \langle C, X\rangle + \langle y, b - \mathcal{A}X\rangle)$$

Optimal (worst case) strategy for player X:

$$p^* = \max_{X \succeq 0} \min_{y} L(X,y)$$

For each fixed $X \succeq 0$: y free yields hidden constraint $b - \mathcal{A}X = 0$; recovers primal problem (PSDP).

Lagrangian Duality

$$\begin{aligned} L(X,y) &= \operatorname{trace}(CX) + y^T(b - \mathcal{A}X) \\ &= b^T y + \operatorname{trace}\,(C - \mathcal{A}^* y)\, X \end{aligned}$$

using adjoint operator, $\mathcal{A}^* y = \sum_i y_i A_i$

satisfies $\langle \mathcal{A}^* y, X \rangle = \langle y, \mathcal{A}X \rangle, \quad \forall X, y$

Then, ***WEAK DUALITY (Lagrangian)*** holds:

$$p^* = \max_{X \succeq 0} \min_y L(X,y) \le d^* := \min_y \max_{X \succeq 0} L(X,y)$$

Derive dual: for each <u>fixed</u> y, $X \succeq 0$ yields <u>hidden constraint</u>

$$g(y) := C - \mathcal{A}^* y \preceq 0$$

Hidden Constraint

$$p^* = \max_{X \succeq 0} \min_{y} L(X,y) \leq d^* := \min_{y} \max_{X \succeq 0} L(X,y)$$

dual obtained from optimal strategy of competing player Y;
use *hidden constraint* $g(y) = C - \mathcal{A}^* y \preceq 0$

$$(\textbf{DSDP}) \qquad \begin{array}{rcc} d^* = & \min & b^T y \\ & \text{s.t.} & \mathcal{A}^* y \succeq C \end{array}$$

for the primal

$$(\textbf{PSDP}) \qquad \begin{array}{rcc} p^* = & \max & \operatorname{trace} CX \\ & \text{s.t.} & \mathcal{A} X = b \\ & & X \succeq 0 \end{array}$$

1.2.3 Weak Duality - Optimality

Proposition 1.2.3.1 *(Weak Duality) If X feas. in (PSDP), y feas. in (DSDP), $Z = \mathcal{A}^* y - C \succeq 0$ is slack variable, then*

$$\operatorname{trace} CX - b^T y = -\operatorname{trace} XZ \le 0.$$

Proof. (Direct - using: $\operatorname{trace} ZX = \operatorname{trace} X^{1/2} X^{1/2} Z = \operatorname{trace} X^{1/2} Z^{1/2} Z^{1/2} X^{1/2} = \|X^{1/2} Z^{1/2}\|^2 \ge 0$;
(Note: $XZ = 0$ iff $\operatorname{trace} XZ = 0$)

$$\begin{aligned} \operatorname{trace} CX - b^T y &= \operatorname{trace} (\mathcal{A}^* y - Z) X - b^T y \\ &= \operatorname{trace} y^T \mathcal{A} X - \operatorname{trace} ZX - b^T y \\ &= y^T (\mathcal{A} X - b) - \operatorname{trace} ZX = -\operatorname{trace} ZX, \end{aligned}$$

■

Characterization of optimality

primal-dual pair $X \succeq 0, y$ (and slack $Z \succeq 0$) are optimal iff

$\mathcal{A}^* y - Z = C$	dual feasibility	$(OC1)$
$\mathcal{A}X = b$	primal feasibility	$(OC2)$
$ZX = 0$	complementary slackness	$(OC3)$

And

$ZX = \mu I$ perturbed complementary slackness (POC3)

Forms the basis for:
interior point methods
(central path: X_μ, y_μ, Z_μ solutions of (OC1),(OC2),(OC3));
primal simplex method
dual simplex method

Strong Duality

$$\begin{aligned}
\text{primal value} &= \langle C, X\rangle \\
&= \langle \mathcal{A}^* y - Z, X\rangle && \text{by dual feasibility} \\
&= \langle y, \mathcal{A}X\rangle - \langle Z, X\rangle && \text{by adjoint} \\
&= \langle y, b\rangle - \langle Z, X\rangle && \text{by primal feasibility} \\
&= \langle y, b\rangle && \text{by complementary slackness} \\
&= \text{dual value}
\end{aligned}$$

1.2.4 Preliminary Examples

Example 1.2.4.1 Minimizing the Maximum Eigenvalue
Arises in e.g. stability of differential equations

- *The mathematical problem:*
  - ◇ *given* $A(x) \in \mathcal{S}^n$ *depending linearly on vector* $x \in \mathbb{R}^m$
  - ◇ *Find* x *to minimize the maximum eigenvalue of* $A(x)$
- *SDP Model:*
  - ◇ *the largest eigenvalue* $\lambda_{\max}(A(x)) \leq \alpha$ ***iff*** $\lambda_{\max}(A(x) - \alpha I) \leq 0$ ***iff*** $A(x) - \alpha I \preceq 0$.
  - ◇ *The DSDP (in dual form) is:*

$$\max -\alpha \qquad \textit{s.t.} \quad A(x) - \alpha I \preceq 0.$$

Pseudoconvex Optimization

Example 1.2.4.2 Pseudoconvex (Nonlinear) Optimization

$$(\mathbf{PCP}) \qquad \begin{array}{rll} d^* = & \min & \frac{(c^T x)^2}{d^T x} \\ & \textit{s.t.} & Ax + b \geq 0, \end{array}$$

(given $Ax + b \geq 0 \Rightarrow d^T x > 0$*)*
Then (using 2×2 *determinant), (PCP) is equivalent to*

$$\begin{array}{rll} d^* = & \min & t \\ & \textit{s.t.} & \begin{bmatrix} \mathrm{Diag}\,(Ax+b) & 0 & 0 \\ 0 & t & c^T x \\ 0 & c^T x & d^T x \end{bmatrix} \succeq 0 \end{array}$$

1.2.5 Historical Notes

- See URL: orion.math.uwaterloo.ca/˜hwolkowi/henry/software/sdpbibliog.pdf for a regularly updated annotated bibliograhy.
- Arguably most active area in optimization (see HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications, 2000, [96], for comprehensive results, history, references, ... and e.g. books [16, 98, 99, 7], [93, 69, 13]).
- Lyapunov over 100 years ago on stability analysis of differential equations
- Bohnenblust 1948 on geometry of the cone of SDPs, [11]
- Yakubovitch in the 1960's and Boyd and others on convex optimization in control in the 1980's, e.g. solving Ricatti Equations (called LMIs), e.g. [14],[94, 93]

More: Historical Notes

• matrix completion problems (another name for SDP) started early 1980's, continues to be a very active area of research, [20],[29], and e.g.: [48, 47, 46, 45, 49, 41, 33, 21, 42]. (More recently,it is being used to solve large scale SDPs.)
• combinatorial optimization applications 1980's: Lovász *theta function* [53]; the strong approximation results for the max-cut problem by Goemans-Williamson, e.g. [28], survey papers: [26, 27],[79].
• linear complementarity problems can be extended to problems over the cone of semidefinite matrices, e.g. [23, 40, 44, 43, 60].

More: Historical Notes

• Complexity, Distance to Ill-Posedness, and Condition Numbers SDP is a convex program and it can be solved to any desired accuracy in polynomial time, see seminal work of Nesterov and Nemirovski e.g. [64, 65, 69, 67, 63, 66, 68]. Another measure of complexity is the distance to ill-posedness: e.g. work by Renegar [81, 85, 84, 83, 82].

• Cone Programming this is a generalization of SDP, also called *generalized linear programming*, in paper by Bellman and Fan 1963, [9]. Other books deal with problems over cones date back to 60s, e.g. [71],[35, 55, 39, 38, 77]. More recently, generalization of SDP to more general cones, e.g. Güler and Tuncel, [31, 92] and also Hauser [32]

More: Historical Notes

- Other Related Areas e.g.: Eigenvalue Functions, e.g. [15], [73, 74]; Financial Applications; Generalized Convexity, e.g. [86, 57]; Statistics; Nonlinear Programming; ...

Further historical notes thanks Didier Henrion

LMI Terminology coined by Jan Willems in 1971

Historically, first LMIs 1890; Lyapunov - differential equation

$\frac{d}{dt}x(t) = \dot{x}(t) = Ax(t)$ is stable (all trajectories converge to zero)

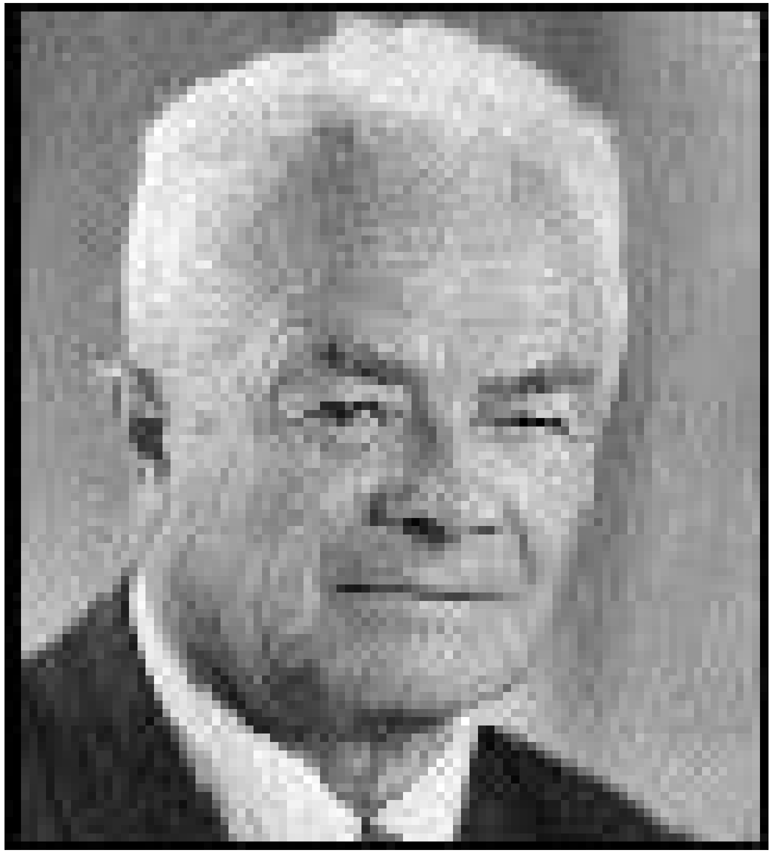
iff there exists a solution to the matrix inequalities

$$A^T P + PA \prec 0 \quad P = P^T \succ 0$$

Aleksandr Mikhailovich Lyapunov (1857 - 1918)

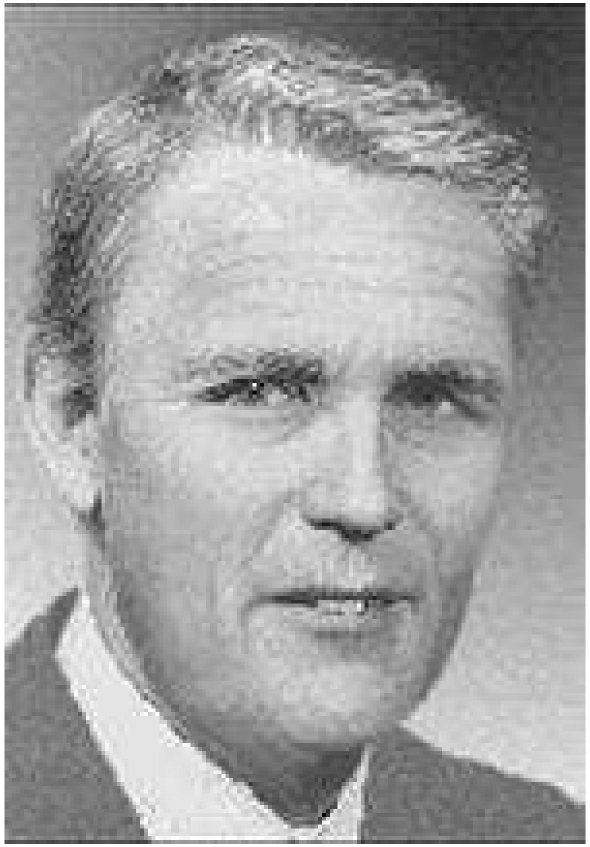
Some history (2)

1960s: Yakubovich, Popov, Kalman, Anderson et al obtained the positive real lemma
Reduces solution of LMI to a simple graphical criterion (Popov, circle and Tsypkin criteria)



Vladimir Andreevich Yakubovich

(1926 Novosibirsk)



Rudolf Emil Kalman

(1930 Budapest)

1970s: Willems focused on solving algebraic equations such as Lyapunov’s or Riccati’s
equations (AREs), rather than LMIs
Then most of the work dedicated to numerical algebra, development of Matlab (1984), focus on solving control AREs, until..

Mathematical programming

1979: ellipsoid algorithm of Khachiyan:
polynomial bound on worst case interation count for LP
1984: Karmarkar introduces interior-point (IP) methods for LP:
improved complexity bound and efficiency
1988: Nesterov, Nemirovski, Alizadeh extend IP methods for SDP
1994: Goemans and Williamson prove than LMI relaxations of MAXCUT (a problem of combinatorial optimization) provide solutions at least 88% the optimal value
1994: Research effort in control culminated in the LMI book by Boyd, El Ghaoui, Feron, Balakrishnan - Contributions also by Apkarian, Bernussou, Gahinet, Geromel, Peres and many others..

1.3 Motivation/Examples/Applications

- Quadratic constrained quadratic programs
- Lovász theta function
- Statistics
- minimizing the L_2-operator norm of a matrix
- linear programming
- robust mathematical programming
- engineering, e.g. control theory
- Combinatorial Problems

1.3.1 Quadratic Constrained Quadratic Programs

What is SEMIDEFINITE PROGRAMMING?
Why use it?

- Quadratic approximations are better than linear approximations.

(For example, model $x \in \{0, 1\}$ using $x^2 - x = 0$.)

And, we can solve relaxations of quadratic approximations efficiently using semidefinite programming.

HOW DOES SDP arise from quadratic approximations?

Let $\quad q_i(y) = \frac{1}{2} y^T Q_i y + y^T b_i + c_i, \; y \in \mathbb{R}^n, \; i = 0, 1, \ldots, m$

$$
\textbf{(QQP)} \quad \begin{cases} q^* = & \min & q_0(y) \\ & \text{s.t.} & q_i(y) = 0 \\ & & i = 1, \ldots m \end{cases}
$$

- Lagrangian: $\quad L(y, x) = \quad$ (with x as Lagrange multipliers)

quadratic in y | linear in y | constant in y

$$\frac{1}{2} y^T (Q_0 - \textstyle\sum_{i=1}^m x_i Q_i) y \;+\; y^T (b_0 - \textstyle\sum_{i=1}^m x_i b_i) \;+\; (c_0 - \textstyle\sum_{i=1}^m x_i c_i)$$

- Primal-Dual pair: $\quad q^* = \min\limits_y \max\limits_x L(y, x) \geq d^* = \max\limits_x \min\limits_y L(y, x)$

Homogenization

- homogenize (add y_0): $\quad y_0 y^T(b_0 - \sum_{i=1}^m x_i b_i), \;\; y_0^2 = 1.$

$$
\begin{aligned}
d^* &= \max_x \min_y && L(y,x) \\
&= \max_x \min_{y_0^2=1} && \tfrac{1}{2} y^T (Q_0 - \textstyle\sum_{i=1}^m x_i Q_i) y \quad (+t y_0^2) \\
& && + y_0 y^T (b_0 - \textstyle\sum_{i=1}^m x_i b_i) \\
& && + (c_0 - \textstyle\sum_{i=1}^m x_i c_i) \quad (-t)
\end{aligned}
$$

Hidden Constraint

with t as Lagrange mulitplier for $y_0^2 = 1$ constraint;

use *hidden semidefinite constraint* to yield SDP constraint

$$(\mathcal{A} : \mathbb{R}^{m+1} \rightarrow \mathcal{S}_{n+1}) \qquad B - \mathcal{A}\begin{pmatrix} t \\ x \end{pmatrix} \succeq 0.$$

$$B = \begin{pmatrix} 0 & b_0^T \\ b_0 & Q_0 \end{pmatrix}, \quad \mathcal{A}\begin{pmatrix} t \\ x \end{pmatrix} = \begin{bmatrix} -t & \sum_{i=1}^m x_i b_i^T \\ \sum_{i=1}^m x_i b_i & \sum_{i=1}^m x_i Q_i \end{bmatrix}$$

Dual Program

The dual program is equivalent to the SDP (with $c_0 = 0$)

$$(\mathbf{D}) \qquad \begin{aligned} d^* = \quad & \sup \quad -t - \textstyle\sum_{i=1}^m x_i c_i \\ & \text{s.t.} \quad \mathcal{A}\begin{pmatrix} t \\ x \end{pmatrix} \preceq B \\ & \qquad x \in \mathbb{R}^m, t \in \mathbb{R} \end{aligned}$$

As in LP: dual obtained from
optimal strategy of competing player:

$$(\mathbf{DD}) \qquad \begin{aligned} p^* = \quad & \inf \quad \operatorname{trace} BU \\ & \text{s.t.} \quad \mathcal{A}^* U = \begin{pmatrix} -1 \\ -c \end{pmatrix} \\ & \qquad U \succeq 0. \end{aligned}$$

1.3.2 Generalized Eigenvalue Problems for $X = X^T$

1. **The Generalized Eigenvalue Problem**
   Let M, A be two $n \times n$ symmetric matrices, $M \succ 0$.
   The set of eigenvalues of the *matrix pencil*, denoted $[M, A]$, is $\{\lambda \in \mathbb{R} : \lambda M - A \text{ is singular}\}$.

$$\lambda_{\max}([M, X]) \le t \quad \text{is equivalent to} \quad tM - X \succeq 0$$

2. **Spectral Norm of Symmetric X**

$$\{|\lambda_i(X)| \le t,\ \forall i\} \text{ iff } \{tI - X \succeq 0,\ tI + X \succeq 0\}$$

3. **Sum of k largest eigenvalues of Symmetric X**
   Let $S_k(X)$ denote the sum of the largest k eigenvalues of X.

$$S_k(X) \le t \ \underline{\text{iff}}\ \ t - ks - \operatorname{trace} Z \succeq 0, Z \succeq 0, Z - X + sI \succeq 0.$$

1.3.3 SDP Application in Statistics

If $m_0, m_1, m_2, \ldots, m_{2n}$ are moments of some distribution, then

$$H(m_0, m_1, \ldots, m_{2n}) = \begin{bmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & \cdots & \cdots & m_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ m_n & m_{n+1} & \cdots & m_{2n} \end{bmatrix} \succeq 0$$

Find distribution with maximal variance and $l_i \leq m_i \leq u_i$:

$$\begin{array}{ll} \max & y \\ \text{s.t.} & \begin{bmatrix} m_2 - y & m_1 \\ m_1 & 1 \end{bmatrix} \succeq 0 \\ & l_i \leq m_i \leq u_i \; (i = 1, \ldots 2n) \end{array}$$

2 Theory Of Cone Programming

2.1 Outline

- Convex cones and Partial Orders
- Convex Cone Programs
- Strong Duality
- Log Barrier and the Central Path

2.2 Convex Cones; Löwner Partial Order

Definition 2.2.0.1 *Let* $\alpha \in \mathbb{R}$ *and* $S, T \subset \mathbb{R}^n$. *Then* $\alpha S = \{y : y = \alpha s,$ *for some* $s \in S\}$ *and* $S + T = \{y : y = s + t,$ *for some* $s \in S, t \in T\}$

Definition 2.2.0.2 $\mathcal{K} \subset \mathbb{R}^n$ *is a* *cone* *if* $\alpha\mathcal{K} \subset K, \quad \forall \alpha > 0$.

Definition 2.2.0.3 *the cone* $\mathcal{K}$ *is a* *convex cone* *if* $\mathcal{K} + \mathcal{K} \subset \mathcal{K}$.

Definition 2.2.0.4 *A cone* $\mathcal{K}$ *is a* *pointed cone* *if* $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$.

Definition 2.2.0.5 *A cone* $\mathcal{K} \subset \mathbb{R}^n$ *is a* *proper cone* *if it is closed, pointed, and convex and has nonempty interior.*

Examples of Cones

Example 2.2.0.1 *open half line:* $\{x \in \mathbb{R} : x > 0\}$.

Example 2.2.0.2 *closed half line:* $\{x \in \mathbb{R} : x \geq 0\}$.

Example 2.2.0.3 *psd matrices, $\mathcal{P}$:* $\{X \in \mathcal{S}^n : X \succeq 0\}$.

Example 2.2.0.4 *Lorentz cone*
(ice-cream cone, second-order cone):
$L^m = \{x = (x_1, \ldots, x_m) \in \mathbb{R}^m : x_m \geq \sqrt{x_1^2 + \ldots + x_{m-1}^2}$.

Polar Cone

And, an important closed convex cone is:

Definition 2.2.0.6 *Polar (dual, conjugate) cone of set* $\mathcal{S}$*:*

$\mathcal{S}^+ := \{z : \langle x, z \rangle \geq 0, \forall x \in \mathcal{S}\}$.

Example 2.2.0.5 *nonnegative orthant, psd cone* $\mathcal{P}$*, and Lorentz cone* L^m*, are all self-polar, i.e.* $\mathcal{K} = \mathcal{K}^+$.

Further Examples

Definition 2.2.0.7 *Direct sum of two cones*

$\mathcal{K} \oplus \mathcal{L} := \{(k\ l) : k \in \mathcal{K}, l \in \mathcal{L}\}$.

Example 2.2.0.6 *Let L denote the half line in $\mathbb{R}$, then $\underbrace{L \oplus L \cdots \oplus L}_{n}$ is the n-dimension nonnegative orthant.*

Properties of Cones

Suppose $\mathcal{K}$, $\mathcal{K}_1$, $\mathcal{K}_2$ are proper cones, then:

- $\mathcal{K}^+$ is proper.
- $(\mathcal{K}^+)^+ = \mathcal{K}$.
- $(\mathcal{K}_1 \cap \mathcal{K}_2)^+ = \overline{\mathcal{K}_1^+ + \mathcal{K}_2^+}$.
- $(\mathcal{K}_1 + \mathcal{K}_2)^+ = \mathcal{K}_1^+ \cap \mathcal{K}_2^+$.
- $\mathcal{K}_1 \oplus \mathcal{K}_2$ is proper.
- $(\mathcal{K}_1 \oplus \mathcal{K}_2)^+ = \mathcal{K}_1^+ \oplus \mathcal{K}_2^+$.

Useful Lemma, e.g. to Prove Farkas' Lemma

Lemma 2.2.0.1

$$\mathcal{K} \textit{ is a closed convex cone} \iff \mathcal{K} = \mathcal{K}^{++}$$

Partial Orders

Definition 2.2.0.8 $x \succeq_{\mathcal{K}} y$ *(respectively* $x \succ_{\mathcal{K}} y$*) if* $x - y \in \mathcal{K}$ *(respectively* $x - y \in \operatorname{int} \mathcal{K}$*).*

Remark 2.2.0.1 *If* $\mathcal{K}$ *is a pointed, convex cone, then "*$\succeq_{\mathcal{K}}$*" is a* *(linear) partial order* *(reflexive, transitive, and antisymmetric):*

- $0 \in \mathcal{K} \Rightarrow x \succeq_{\mathcal{K}} x$ *(reflexive);*
- $\mathcal{K}$ *is convex* $\Rightarrow$ *if* $x \succeq_{\mathcal{K}} y$ *and* $y \succeq_{\mathcal{K}} z$*, then* $x \succeq_{\mathcal{K}} z$ *(transitive);*
- $\mathcal{K}$ *is pointed* $\Rightarrow$ *if* $x \succeq_{\mathcal{K}} y$ *and* $y \succeq_{\mathcal{K}} x$*, then* $x = y$ *(antisymmetric);*
- $\mathcal{K}$ *is convex cone* $\Rightarrow$
  if $a, b \geq 0, u \succeq_{\mathcal{K}} x$ *and* $v \succeq_{\mathcal{K}} y$*, then* $au + bv \succeq_{\mathcal{K}} ax + by$ *(linear - homogeneous, additive);*

2.3 Convex Cone Program

Let: $\mathcal{K}, \mathcal{L}$ be convex cones;
f real valued convex function, $\mathcal{G}$ is $\mathcal{K}$-convex, i.e.
$\mathcal{G}(\alpha u + (1-\alpha)v) \preceq_{\mathcal{K}} \alpha\mathcal{G}(u) + (1-\alpha)\mathcal{G}(v), \quad \forall 0 \le \alpha \le 1, \forall u, v$

$$(CP) \qquad \begin{array}{rl} \mu^* = \min & f(x) \\ \text{s.t.} & \mathcal{G}(x) \preceq_{\mathcal{K}} 0 \\ & x \succeq_{\mathcal{L}} 0 \end{array}$$

$$\mu^* = \min_{x \in \mathcal{L}} \max_{y \in \mathcal{K}^+} f(x) + \langle y, \mathcal{G}(x) \rangle$$

-

with Lagrangian dual (weak duality)

$$(DCP) \qquad \mu^* \ge \nu^* = \max_{y \in \mathcal{K}^+} \min_{x \in \mathcal{L}} \; f(x) + \langle y, \mathcal{G}(x) \rangle$$

Linear Cone Optimization

Example 2.3.0.7 *If* $f(x) = \langle c, x \rangle$, $g(x) = b - \mathcal{A}x$, *then we have a linear cone programming problem, (LCP). The dual*

$$(DLCP) \qquad \begin{aligned} \mu^* \geq \nu^* &= \max_{y \in \mathcal{K}^+} \min_{x \in \mathcal{L}} \langle c, x \rangle + \langle y, b - \mathcal{A}x \rangle \\ &= \max_{y \in \mathcal{K}^+} \min_{x \in \mathcal{L}} \langle c - \mathcal{A}^* y, x \rangle + \langle y, b \rangle \\ &= \max_{\substack{y \in \mathcal{K}^+ \\ c - \mathcal{A}^* y \in \mathcal{L}^+}} \min_{x \in \mathcal{L}} \langle c - \mathcal{A}^* y, x \rangle + \langle y, b \rangle \end{aligned}$$

reduces to elegant (LP or SDP type)

$$(DLCP) \qquad \begin{array}{rc} \mu^* \geq \nu^* = \max & \langle y, b \rangle \\ \textit{s.t.} & \mathcal{A}^* y \preceq_{\mathcal{L}^+} c \\ & y \succeq_{\mathcal{K}^+} 0 \end{array}$$

Linear Cone Optimization cont...

elegant (LP or SDP type) symmetric dual pair

$$(PLCP) \qquad \begin{array}{rcl} \mu^* = \min & & \langle c, x\rangle \\ \text{s.t.} & & \mathcal{A}x \succeq_{\mathcal{K}} b \\ & & x \succeq_{\mathcal{L}} 0 \end{array}$$

$$(DLCP) \qquad \begin{array}{rcl} \mu^* \geq \nu^* = \max & & \langle y, b\rangle \\ \text{s.t.} & & \mathcal{A}^* y \preceq_{\mathcal{L}^+} c \\ & & y \succeq_{\mathcal{K}^+} 0 \end{array}$$

Abstract Convex Program

general abstract convex program is

$$(CP) \qquad \begin{array}{rll} \mu^* := & \inf & f(x) \\ & \text{subject to} & \mathcal{G}(x) \preceq_L 0 \\ & & x \succeq_K 0, \end{array}$$

$f : X \rightarrow$ real valued convex function on K;
X, Y Banach spaces;
$x \succeq_K y$ partial order induced convex cone K, i.e. $y - x \in K$;
$\mathcal{G} : X \rightarrow Y$ is L-convex function on K, i.e.

$$\mathcal{G}(\lambda x + (1 - \lambda y) \preceq_L \lambda \mathcal{G}(x) + (1 - \lambda)\mathcal{G}(y), \quad \forall x, y \in K, \forall 0 < \lambda < 1$$

Definitions

1. $K \subset \mathbb{R}^n$ is a <u>cone</u> if $\lambda K \subset K, \forall \lambda > 0$.
2. The cone $K \subset \mathbb{R}^n$ is a <u>convex cone</u> if $K + K \subset K$.
3. The <u>polar cone</u> of the set Ω is
   $\Omega^+ := \{\phi \in X^* : \langle \phi, x \rangle \geq 0, \ \forall x \in \Omega\}$, where X^* is the (topological) dual space of X.

Weak/Strong Duality

Lagrangian: $\mathcal{L}(x,\phi) := f(x) + \langle \phi, \mathcal{G}(x) \rangle$

$$\mu^* = \min_{x \in K} \max_{\phi \in L^+} \mathcal{L}(x,\phi) \geq \nu^* := \max_{\phi \in L^+} \min_{x \in K} \mathcal{L}(x,\phi) \qquad \text{Weak Duality}$$

$$\mu^* = \nu^* = \min_{x \in K} \mathcal{L}(x,\phi^*), \ \text{for some } \phi^* \in L^* \qquad \text{Strong Duality}$$

i.e. Strong Duality means equality and dual attainment for some Lagrange multiplier vector ϕ^*.

CQ and CS

$$\mu^* = \nu^* = \min_{x \in K} \mathcal{L}(x, \phi^*), \text{ for some } \phi^* \in L^* \qquad \text{Strong Duality}$$

Strong duality holds under a constraint qualification, e.g.

$$\exists \hat{x} \in K \text{ with } \mathcal{G}(\hat{x}) \prec 0 \qquad \text{Slater's Condition - strict feasibility}$$

primal attainment, $f(x^*) = \mu^*$, implies

$$\langle \phi^*, \mathcal{G}(x^*) \rangle = 0 \qquad \text{Complementary Slackness}$$

Optimality Conditions

x^* optimal implies

$$\nabla \mathcal{L}(x^*, \lambda^*) \in (K - x^*)^+$$

equivalently,

$$\begin{array}{rclllr} \nabla \mathcal{L}(x^*, \phi^*) - Z & = & 0, & Z \succeq_{K^+} 0 & \phi^* \succeq_{L^+} 0 & \text{D.F} \\ \mathcal{G}(x^*) + W & = & 0, & X^* \succeq_K 0 & W \succeq_L 0 & \text{D.F} \end{array}$$

$$\langle Z, X^* \rangle = 0, \quad \langle \phi^*, W \rangle = 0 \qquad \text{C.S.}$$

LP Example

$$LP \qquad \begin{array}{rrcl} \min & c^T x & & \\ \text{s.t.} & a - Ax & = & 0 \\ & b - Bx & \leq & 0 \\ & x & \geq & 0 \end{array}$$

$$\nabla \mathcal{L}(x^*, \psi^*, \phi^*) = c - A^T \psi^* - B^T \phi^* = w \in (\Re^n_+ - x^*)^+$$

$w^T(0 - x^*) \geq 0$ and $w^T(\alpha x^* - x^*) \geq 0$ implies $w^T x^* = 0$.
$w^T(\alpha y^* - x^*) \geq 0, y \geq 0$ implies $w \geq 0$.
(and strong duality implies $(\psi^*)^T(b - Bx) = 0$)

Failures

$$(CP) \qquad \begin{array}{rll} \mu^* := & \inf & f(x) \\ & \text{subject to} & b - \mathcal{A}(x) = 0 \\ & & x \succeq_K 0, \end{array}$$

$\mathcal{F} = K \cap \{\hat{x} + \mathcal{N}(A)\}$ Feasible Set

$$\begin{array}{rcll} \nabla f(x^*) \in (\mathcal{F} - x^*)^+ & = & \mathrm{cl}\,\{(K - x^*)^+ + \mathbb{R}(\mathcal{A}^*)\} & \text{Opt. Cond.} \\ & \stackrel{?}{=} & (K - x^*)^+ + \mathbb{R}(\mathcal{A}^*) & \text{sum of 2 ccc} \end{array}$$

equivalently

$$\nabla f(x^*) + \mathcal{A}^* \phi \stackrel{?}{\in} (K - x^*)^+ \quad \text{Lagr. Mult. } \phi$$

Specific Failure

$$\begin{pmatrix} \frac{1}{i} & 1 \\ 1 & i \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -i \end{pmatrix} \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \notin \mathbb{S}^n_+ + \operatorname{span}\left\{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

Let $C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Then

$$0 = \mu^* := \min \quad \operatorname{trace} CX \text{ s.t. } \operatorname{trace} A_1 X = 0, \qquad X \succeq 0.$$

But

$$C + \phi A_1 \notin (\mathbb{S}^n_+ - 0)^+ = \mathbb{S}^n_+, \qquad \forall \phi \in \Re$$

Fix: Replace $\mathbb{S}^n_+$ with the Minimal Cone

$K = \mathbb{S}^n_+ \cap \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^\perp = \operatorname{cone}\{A_1\}$; so $K^+ = \{A_1\}^+$

2.4 Strong Duality

2.4.1 Outline

- Faces and minimal cones
- Optimality Conditions without constraint qualifcations

2.4.2 Facial Structure

The cone $K \subset T$ is a *face* of the cone T, denoted $K \lhd T$, if

$$x, y \in T, \; x + y \in K \Rightarrow x, y \in K.$$

The faces of $\mathcal{P}$ have a very special structure. Each face, $K \lhd \mathcal{P}$, is characterized by a unique subspace, $S \subset \Re^n$:

$$K = \{X \in \mathcal{P} : \mathcal{N}(X) \supset S\}.$$

Moreover,

$$\operatorname{relint}(K) = \{X \in \mathcal{P} : \mathcal{N}(X) = S\}.$$

(Or use $\mathcal{R}(X)$ and $S^{\perp}$.)

Complementary (or conjugate) Faces

The *complementary* (or conjugate) face of K is $K^c = K^\perp \cap \mathcal{P}$ and

$$K^c = \{X \in \mathcal{P} : \mathcal{N}(X) \supset S^\perp\}.$$

Moreover,

$$\operatorname{relint}(K^c) = \{X \in \mathcal{P} : \mathcal{N}(X) = S^\perp\}.$$

Two additional facts:

1. Each face $\mathcal{K}$ (resp., $\mathcal{K}^c$) is exposed, i.e.

$$\mathcal{K} = \mathcal{P} \cap \phi^{\perp}, \qquad \text{for some } \phi \in \mathcal{P}^{+}$$

   i.e. it equals the intersection of $\mathcal{P}$ with a supporting hyperplane; the supporting hyperplane corresponds to any $X \in \operatorname{relint}(K^c)$ (resp., $\operatorname{relint}(K)$). Also, complementary faces are orthogonal and satisfy $XY = 0, \forall X \in K, Y \in K^c$.

2. The cone $\mathcal{P}$ is projectionally exposed, i.e. every face of $\mathcal{P}$ is the image of $\mathcal{P}$ under some projection. In fact, if $Q \in \mathcal{S}_n$ is the projection onto the subspace S, the null space of matrices in $\operatorname{relint}(K)$, then the face K satisfies

$$K = (I - Q)\mathcal{P}(I - Q).$$

2.4.3 Primal-Dual Pair

Consider the semidefinite linear programming problem

$$\textbf{(P)} \qquad \begin{array}{rcl} p^* = & \sup & c^t x \\ & \text{subject to} & Ax \preceq b \\ & & x \in \Re^m, \end{array}$$

If Slater's CQ holds

$$\exists \hat{x} : A\hat{x} \prec b$$

then strong duality holds

$$\textbf{(D)} \qquad \begin{array}{rcl} p^* = d^* := & \boxed{\min} & \operatorname{trace} bU \\ & \text{subject to} & A^*U = c \\ & & U \succeq 0 \end{array}$$

2.4.4 Minimal Cone

The minimal cone of $(\mathbf{P})$ is defined as

$$\mathcal{P}^f = \cap\{K \lhd \mathcal{P} : K \supset (b - A(F))\},$$

i.e. the minimal cone is the intersection of all faces of $\mathcal{P}$ containing the feasible slacks.

2.4.5 Strong Duality Without CQ

$$(\mathbf{P}) \qquad \begin{array}{rcc} p^* = & \sup & c^t x \\ & \text{subject to} & Ax \preceq_{\mathcal{P}^f} b \\ & & x \in \Re^m, \end{array}$$

$$(\mathbf{D}) \qquad \begin{array}{rcc} p^* = d^* := & \boxed{\min} & \operatorname{trace} bU \\ & \text{subject to} & A^*U = c \\ & & U \succeq_{(\mathcal{P}^f)^+} 0 \end{array}$$

Applications in e.g. QAP,GP,MC ...: [102, 97, 3, 91]

A Related Papers

A.1 Outline

References

[1] F. ALIZADEH. Lecture notes on semidefinite programming. Technical report, Rutgers University, Piscataway, NJ, 2000. URL: http://karush.rutgers.edu/~alizadeh/.

[2] F. ALIZADEH, J-P.A. HAEBERLY, and M.L. OVERTON. A new primal-dual interior-point method for semidefinite programming. In J.G. Lewis, editor, *Proceedings of the Fifth SIAM Conference on Applied Linear Algebra*, pages 113–117. SIAM, 1994.

[3] M.F. ANJOS and H. WOLKOWICZ. A strengthened SDP relaxation via a second lifting for the Max-Cut problem. *Discrete Appl. Math.*, 119:79–106, 2002.

[4] K.M. ANSTREICHER and N.W. BRIXIUS. Solving quadratic assignment problems using convex quadratic programming relaxations. Technical report, University of Iowa, Iowa City, IA, 2000.

[5] K.M. ANSTREICHER, X. CHEN, H. WOLKOWICZ, and Y. YUAN. Strong duality for a trust-

region type relaxation of the quadratic assignment problem. *Linear Algebra Appl.*, 301(1-3):121–136, 1999.

[6] K.M. ANSTREICHER and H. WOLKOWICZ. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.*, 22(1):41–55, 2000.

[7] V. BALAKRISHNAN and E. FERON, editors. *Linear matrix inequalities in control theory and applications*. John Wiley & Sons Ltd., Chichester, 1996. Internat. J. Robust Nonlinear Control **6** (1996), no. 9-10.

[8] R. BELLMAN. *Introduction to Matrix Analysis*. McGraw Hill, New York, 1957.

[9] R. BELLMAN and K. FAN. On systems of linear inequalities in Hermitian matrix variables. In *Proceedings of Symposia in Pure Mathematics, Vol 7, AMS*, 1963.

[10] A. BEN-TAL and A.S. NEMIROVSKI. Convex Optimization in Engineering: Modeling, Analysis, Algorithms. Technical Report on notes from

course given by A. Nemirovski, Delft University of Technology, Delft, The Netherlands, 1998.

[11] F. BOHNENBLUST. Joint positiveness of matrices. Technical report, UCLA, 1948. Manuscript.

[12] J.M. BORWEIN and H. WOLKOWICZ. Characterization of optimality for the abstract convex program with finite-dimensional range. *J. Austral. Math. Soc. Ser. A*, 30(4):390–411, 1980/81.

[13] S. BOYD, V. BALAKRISHNAN, E. FERON, and L. El GHAOUI. Control system analysis and synthesis via linear matrix inequalities. *Proc. ACC*, pages 2147–2154, 1993.

[14] S. BOYD and C. H. BARRATT. *Linear Controller Design. Limits of Performance*. Prentice-Hall, 1991.

[15] J. CULLUM, W.E. DONATH, and P. WOLFE. The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices. *Mathematical Programming Study*, 3:35–55, 1975.

[16] E. de KLERK. *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Ap-*

plications. Applied Optimization Series. Kluwer Academic, Boston, MA, 2002.

[17] J.E. DENNIS Jr. and R.B. SCHNABEL. *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996. Corrected reprint of the 1983 original.

[18] M.M. DEZA and M. LAURENT. *Geometry of cuts and metrics*. Springer-Verlag, Berlin, 1997.

[19] L.C.W. DIXON, S.E. HERSOM, and Z.A. MAANY. Initial experience obtained solving the low thrust satellite trajectory optimisation problem. Technical Report T.R. 152, The Hatfield Polytechnic Numerical Optimization Centre, 1984.

[20] H. DYM and I. GOHBERG. Extensions of band matrices with band inverses. *Linear Algebra Appl.*, 36:1–24, 1981.

[21] S.M. FALLAT, C.R. JOHNSON, and R.L. SMITH. The general totally positive matrix completion problem with few unspecified entries.

Electron. J. Linear Algebra, 7:1–20 (electronic), 2000.

[22] A.V. FIACCO and G.P. McCORMICK. *Nonlinear programming*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second (first 1968) edition, 1990. Sequential unconstrained minimization techniques.

[23] A. FISCHER. A Newton-type method for positive-semidefinite linear complementarity problems. *J. Optim. Theory Appl.*, 86(3):585–608, 1995.

[24] T. FUJIE and M. KOJIMA. Semidefinite programming relaxation for nonconvex quadratic programs. *J. Global Optim.*, 10(4):367–380, 1997.

[25] D.M. GAY. Computing optimal locally constrained steps. *SIAM J. Sci. Statist. Comput.*, 2:186–197, 1981.

[26] M.X. GOEMANS. Semidefinite programming in combinatorial optimization. *Math. Programming*, 79:143–162, 1997.

[27] M.X. GOEMANS. Semidefinite programming and combinatorial optimization. *Documenta*

Mathematica, Extra Volume ICM 1998:657–666, 1998. Invited talk at the International Congress of Mathematicians, Berlin, 1998.

[28] M.X. GOEMANS and D.P. WILLIAMSON. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995.

[29] B. GRONE, C.R. JOHNSON, E. MARQUES de SA, and H. WOLKOWICZ. Positive definite completions of partial Hermitian matrices. *Linear Algebra Appl.*, 58:109–124, 1984.

[30] Ming Gu. Primal-dual interior-point methods for semidefinite programming in finite precision. *SIAM J. Optim.*, 10(2):462–502 (electronic), 2000.

[31] O. GÜLER and L. TUNÇEL. Characterization of the barrier parameter of homogeneous convex cones. *Math. Programming*, 81:55–76, 1998.

[32] R. HAUSER. Self-scaled barrier functions: Decomposition and classification. Technical Report DAMTP 1999/NA13, Department of Applied

Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England, 1998.

[33] T.L. HAYDEN, J. WELLS, W-M. LIU, and P. TARAZAGA. The cone of distance matrices. *Linear Algebra Appl.*, 144:153–169, 1991.

[34] C. HELMBERG, F. RENDL, R. J. VANDERBEI, and H. WOLKOWICZ. An interior-point method for semidefinite programming. *SIAM J. Optim.*, 6(2):342–361, 1996.

[35] R.B. HOLMES. *Geometric Functional Analysis and its Applications*. Springer-Verlag, Berlin, 1975.

[36] R.A. HORN and C.R. JOHNSON. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[37] R.A. HORN and C.R. JOHNSON. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[38] J. JAHN. *Mathematical Vector Optimization in Partially Ordered Linear Spaces*. Peter Lang, Frankfurt am Main, 1986.

[39] G. JAMESON. *Ordered linear spaces*. Springer-Verlag, New York, 1970.

[40] B. JANSEN, C. ROOS, and T. TERLAKY. A family of polynomial affine scaling algorithms for positive semidefinite linear complementarity problems. *SIAM J. Optim.*, 7(1):126–140, 1997.

[41] C.R. JOHNSON. Matrix completion problems: a survey. In *Matrix theory and applications (Phoenix, AZ, 1989)*, pages 171–198. Amer. Math. Soc., Providence, RI, 1990.

[42] C.R. JOHNSON and R.L. SMITH. The symmetric inverse M-matrix completion problem. *Linear Algebra Appl.*, 290(1-3):193–212, 1999.

[43] M. KOJIMA, M. SHIDA, and S. SHINDOH. Reduction of monotone linear complementarity problems over cones to linear programs over cones. Technical report, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1995.

[44] M. KOJIMA, S. SHINDOH, and S. HARA. Interior-point methods for the monotone semidefi-

nite linear complementarity problem in symmetric matrices. *SIAM J. Optim.*, 7(1):86–125, 1997.

[45] M. LAURENT. Cuts, matrix completions and graph rigidity. *Math. Programming*, 79:255–284, 1997.

[46] M. LAURENT. The real positive semidefinite completion problem for series-parallel graphs. *Linear Algebra Appl.*, 252:347–366, 1997.

[47] M. LAURENT. A connection between positive semidefinite and Euclidean distance matrix completion problems. *Linear Algebra Appl.*, 273:9–22, 1998.

[48] M. LAURENT. A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems. In *Topics in Semidefinite and Interior-Point Methods*, volume 18 of *The Fields Institute for Research in Mathematical Sciences, Communications Series*, pages 51–76, Providence, Rhode Island, 1998. American Mathematical Society.

[49] M. LAURENT. Polynomial instances of the positive semidefinite and Euclidean distance matrix

completion problems. *SIAM J. Matrix Anal. Appl.*, 22:874–894, 2000.

[50] M. LAURENT and S. POLJAK. On a positive semidefinite relaxation of the cut polytope. *Linear Algebra Appl.*, 223/224:439–461, 1995.

[51] A.S. LEWIS. *Take-home final exam, Course CO663 in Convex Analysis*. University of Waterloo, Ontario, Canada, 1994.

[52] A.S. LEWIS. Derivatives of spectral functions. *Math. Oper. Res.*, 21(3):576–588, 1996.

[53] L. LOVÁSZ. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:1–7, 1979.

[54] K. LÖWNER. Über monotone matrixfunktionen. *Math. Z.*, 38:177–216, 1934.

[55] D.G. LUENBERGER. *Optimization by Vector Space Methods*. John Wiley, 1969.

[56] Z.A. MAANY. A new algorithm for highly curved constrained optimization. Technical Report T.R. 161, The Hatfield Polytechnic Numerical Optimization Centre, 1985.

[57] A.W. MARSHALL and I. OLKIN. *Inequalities: Theory of Majorization and its Applications*. Academic Press, 1979.

[58] M. MESBAHI and G.P. PAPAVASSILOPOULOS. On the rank minimization problem over a positive semidefinite linear matrix inequality. *IEEE Trans. Automat. Control*, 42(2):239–243, 1997.

[59] R.D.C. MONTEIRO. Primal-dual path-following algorithms for semidefinite programming. *SIAM J. Optim.*, 7(3):663–678, 1997.

[60] R.D.C. MONTEIRO and T. TSUCHIYA. Polynomiality of primal-dual algorithms for semidefinite linear complementarity problems based on the Kojima-Shindoh-Hara family of directions. *Math. Programming*, 84:39–53, 1999.

[61] R.D.C. MONTEIRO and Y. ZHANG. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Math. Programming*, 81(3, Ser. A):281–299, 1998.

[62] J.J. MORÉ and D.C. SORENSEN. Computing a trust region step. *SIAM J. Sci. Statist. Comput.*, 4:553–572, 1983.

[63] Y.E. NESTEROV and A.S. NEMIROVSKI. A general approach to polynomial-time algorithms design for convex programming. Technical report, Centr. Econ. & Math. Inst., USSR Acad. Sci., Moscow, USSR, 1988.

[64] Y.E. NESTEROV and A.S. NEMIROVSKI. Polynomial barrier methods in convex programming. *Èkonom. i Mat. Metody*, 24(6):1084–1091, 1988.

[65] Y.E. NESTEROV and A.S. NEMIROVSKI. Self–concordant functions and polynomial–time methods in convex programming. Book–Preprint, Central Economic and Mathematical Institute, USSR Academy of Science, Moscow, USSR, 1989. Published in Nesterov and Nemirovsky [69].

[66] Y.E. NESTEROV and A.S. NEMIROVSKI. *Optimization over positive semidefinite matrices: Mathematical background and user's manual*. USSR Acad. Sci. Centr. Econ. & Math. Inst., 32 Krasikova St., Moscow 117418 USSR, 1990.

[67] Y.E. NESTEROV and A.S. NEMIROVSKI. Self-concordant functions and polynomial time methods in convex programming. Technical report, Centr. Econ. & Math. Inst., USSR Acad. Sci., Moscow, USSR, April 1990.

[68] Y.E. NESTEROV and A.S. NEMIROVSKI. Conic formulation of a convex programming problem and duality. Technical report, Centr. Econ. & Math. Inst., USSR Academy of Sciences, Moscow USSR, 1991.

[69] Y.E. NESTEROV and A.S. NEMIROVSKI. *Interior Point Polynomial Algorithms in Convex Programming*. SIAM Publications. SIAM, Philadelphia, USA, 1994.

[70] Y.E. NESTEROV and M.J. TODD. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22(1):1–42, 1997.

[71] L.W. NEUSTADT. *Optimization*. Princeton University Press, Princeton, N. J., 1976. A Theory of Necessary Conditions, With a chapter by H. T. Banks.

[72] J.M. ORTEGA and W.C. RHEINBOLDT. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York, NY, 1970.

[73] M.L. OVERTON. On minimizing the maximum eigenvalue of a symmetric matrix. *SIAM J. Matrix Anal. Appl.*, 9:256–268, 1988.

[74] M.L. OVERTON. Large-scale optimization of eigenvalues. *SIAM J. Optim.*, 2:88–120, 1992.

[75] G. PATAKI. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23(2):339–358, 1998.

[76] G. PATAKI and L. TUNÇEL. On the generic properties of convex optimization problems in conic form. *Math. Programming*, to appear.

[77] A.L. PERESSINI. *Ordered Topological Vector Spaces*. Harper & Row Publishers, New York, 1967.

[78] S. POLJAK, F. RENDL, and H. WOLKOWICZ. A recipe for semidefinite relaxation for

$(0, 1)$-quadratic programming. *J. Global Optim.*, 7(1):51–73, 1995.

[79] F. RENDL. Semidefinite programming and combinatorial optimization. *Appl. Numer. Math.*, 29:255–281, 1999.

[80] F. RENDL and H. WOLKOWICZ. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Math. Programming*, 77(2, Ser. B):273–299, 1997.

[81] J. RENEGAR. On the computational complexity and geometry of the first-order theory of the reals. III. Quantifier elimination. *J. Symbolic Comput.*, 13(3):329–352, 1992.

[82] J. RENEGAR. Is it possible to know a problem instance is ill-posed? Some foundations for a general theory of condition numbers. *J. Complexity*, 10(1):1–56, 1994.

[83] J. RENEGAR. Incorporating condition measures into the complexity theory of linear programming. *SIAM J. Optim.*, 5(3):506–524, 1995.

[84] J. RENEGAR. Linear programming, complexity theory and elementary functional analysis. *Math. Programming*, 70(3, Ser. A):279–351, 1995.

[85] J. RENEGAR. Condition numbers, the barrier method, and the conjugate-gradient method. *SIAM J. Optim.*, 6(4):879–912, 1996.

[86] S. SCHAIBLE and W.T. ZIEMBA, editors. *Generalized Concavity in Optimization and Economics*, New York, 1981. Academic Press Inc. [Harcourt Brace Jovanovich Publishers].

[87] N.Z. SHOR. Quadratic optimization problems. *Izv. Akad. Nauk SSSR Tekhn. Kibernet.*, 222(1):128–139, 222, 1987.

[88] R. STERN and H. WOLKOWICZ. Trust region problems and nonsymmetric eigenvalue perturbations. *SIAM J. Matrix Anal. Appl.*, 15(3):755–778, 1994.

[89] R. STERN and H. WOLKOWICZ. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optim.*, 5(2):286–313, 1995.

[90] M.J. TODD. Semidefinite programming. Technical report, School of OR and IE, Cornell University, Ithaca, NY, 2000.

[91] L. TUNÇEL. On the Slater condition for the SDP relaxations of nonconvex sets. Technical Report 4, 2001.

[92] L. TUNÇEL and S. XU. On homogeneous convex cones, Caratheodory number, and duality mapping semidefinite liftings. Technical Report CORR 99-21, Department of Combinatorics and Optimization, Waterloo, Ont, 1999.

[93] L. VANDENBERGHE and S. BOYD. Positive definite programming. In *Mathematical Programming: State of the Art, 1994*, pages 276–308. The University of Michigan, 1994.

[94] L. VANDENBERGHE and S. BOYD. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, 1996.

[95] H. WOLKOWICZ. Some applications of optimization in matrix theory. *Linear Algebra Appl.*, 40:101–118, 1981.

[96] H. WOLKOWICZ, R. SAIGAL, and L. VANDENBERGHE, editors. *HANDBOOK OF*

SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications. Kluwer Academic Publishers, Boston, MA, 2000. xxvi+654 pages.

[97] H. WOLKOWICZ and Q. ZHAO. Semidefinite programming relaxations for the graph partitioning problem. *Discrete Appl. Math.*, 96/97:461–479, 1999. Selected for the special Editors' Choice, Edition 1999.

[98] S. WRIGHT. *Primal-Dual Interior-Point Methods*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, 1996.

[99] Y. YE. *Interior Point Algorithms: Theory and Analysis*. Wiley-Interscience series in Discrete Mathematics and Optimization. John Wiley & Sons, New York, 1997.

[100] Y. Yuan. On a subproblem of trust region algorithms for constrained optimization. *Math. Programming*, 47(1 (Ser. A)):53–63, 1990.

[101] Y. YUAN. A dual algorithm for minimizing a quadratic function with two quadratic constraints. *Journal of Computational Mathematics*, 9:348–359, 1991.

[102] Q. ZHAO, S.E. KARISCH, F. RENDL, and H. WOLKOWICZ. Semidefinite programming relaxations for the quadratic assignment problem. *J. Comb. Optim.*, 2(1):71–109, 1998. Semidefinite programming and interior-point approaches for combinatorial optimization problems (Toronto, ON, 1996).

B Solving SDPs with NEOS and MATLAB

B.1 SDP at NEOS

Solvers for SDP are available at e.g. NEOS: (URL www-neos.mcs.anl.gov/neos/server-solvers.html#SDP) The list of solvers and type of input is:
CSDP [Matlab Binary Input] [Sparse SDPA Input]
DSDP [Sparse SDPA Input]
PENBMI [Matlab Binary Input] [Matlab input]
PENSDP [Matlab Binary Input] [Sparse SDPA Input]
SDPA [Matlab Binary Input] [Sparse SDPA Input]
SDPLR [Matlab Binary Input] [Sparse SDPA Input] [SDPLR input]
SDPT3 [Matlab Binary Input] [Sparse SDPA Input]

SeDuMi [Matlab Binary Input] [Sparse SDPA Input]
You can solve SDP problems with various NEOS interfaces including email.

The Sparse SDPA Input format is described at ftp://plato.asu. and included below. (The notation differs from that we use above. In particular, the primal and dual problems are interchanged.)

We work with a semidefinite programming problem that has been written in the following standard form:

```
(P)     min c1*x1+c2*x2+...+cm*xm
        st  F1*x1+F2*x2+...+Fm*xn-F0=X
                                   X >= 0

The dual of the problem is:

(D)     max tr(F0*Y)
        st  tr(Fi*Y)=ci    i=1,2,...,m
                  Y >= 0
```

Here all of the matrices F0, F1, ..., Fm, X, and Y are assumed to be symmetric of size n by n. The constraints X GE 0 and Y GE 0 mean that X and Y must be positive semidefinite.
Note that several other standard forms for SDP have been

used by a number of authors- these can generally be translated into the SDPA standard form with little effort.
File Format:
The file consists of six sections:
1. Comments. The file can begin with arbitrarily many lines of comments. Each line of comments must begin with '"' or ''.*
2. The first line after the comments contains m, the number of constraint matrices. Additional text on this line after m is ignored.
3. The second line after the comments contains nblocks, the number of blocks in the block diagonal structure of the matrices. Additional text on this line after nblocks is ignored.
4. The third line after the comments contains a vector of numbers that give the sizes of the individual blocks. Negative numbers may be used to indicate that a block is actually a diagonal submatrix. Thus a block size of "-5" indicates a 5 by 5 block in which only the diagonal elements are nonzero.
5. The fourth line after the comments contains the objective function vector c.
6. The remaining lines of the file contain entries in the constraint matrices, with one entry per line. The format for each line is

```
<matno> <blkno> <i> <j> <entry>
```

Here matno *is the number of the matrix to which this entry belongs,* blkno *specifies the block within this matrix, i and j specify a location within the block, and* entry *gives the value of the entry in the matrix. Note that since all matrices are assumed to be symmetric, only entries in the upper triangle of a matrix are given.*
Example:

```
          min 10x1+20x2

          st  X=F1x1+F2x2-F0

              X >= 0

where


  F0=[1 0 0 0
      0 2 0 0
      0 0 3 0
      0 0 0 4]

  F1=[1 0 0 0
```

```
      0 1 0 0
      0 0 0 0
      0 0 0 0]

  F2=[0 0 0 0
      0 1 0 0
      0 0 5 2
      0 0 2 6]
```

In SDPA sparse format,
this problem can be written as:

```
"A sample problem.
2 =mdim
2 =nblocks
2 2
10.0 20.0
0 1 1 1 1.0
0 1 2 2 2.0
0 2 1 1 3.0
0 2 2 2 4.0
1 1 1 1 1.0
1 1 2 2 1.0
2 1 2 2 1.0
```

```
2 2 1 1 5.0
2 2 1 2 2.0
2 2 2 2 6.0
```

B.2 SDP with MATLAB

Following is a MATLAB program that solves the relaxation of the MAXCUT problem taken from [34]. The input is the Laplacian matrix L; the optimal primal matrix X satisfies $\mathrm{diag}\,(X) = \frac{1}{4}e$.

```
function [phi, X, y] = maxcut( L);
% solves: max trace(LX) s.t. X psd, diag(X) = b;  b = ones(n,1)/4
%         min b'y       s.t. Diag(y) - L psd, y unconstrained,
% input:  L ... symmetric matrix
% output: phi ... optimal value of primal, phi =trace(LX)
%         X   ... optimal primal matrix
%         y   ... optimal dual vector
% call:   [phi, X, y] = psd_ip( L);

digits = 6;                        % 6 significant digits of phi
[n, n1] = size( L);                % problem size
b = ones( n,1 ) / 4;               % any b>0 works just as well
X = diag( b);                      % initial primal matrix is pos. def.
y = sum( abs( L))' * 1.1;          % initial y is chosen so that
Z = diag( y) - L;                  % initial dual slack Z is pos. def.
phi = b'*y;                        % initial dual
psi = L(:)' * X( :);               % and primal costs
mu = Z( :)' * X( :)/( 2*n);        % initial complementarity
iter=0;                            % iteration count

disp(['      iter    alphap    alphad      gap     lower      upper']);

while phi-psi > max([1,abs(phi)]) * 10^(-digits)

      iter = iter + 1;             % start a new iteration
      Zi = inv( Z);                % inv(Z) is needed explicitly
      Zi = (Zi + Zi')/2;
      dy =  (Zi.*X) \ (mu * diag(Zi) - b);      % solve for dy
      dX = - Zi * diag( dy) * X + mu * Zi - X;  % back substitute for dX
      dX = ( dX + dX')/2;          % symmetrize
```

```
% line search on primal
      alphap = 2;               % initial steplength changed to 2
%     alphap = 1;               % initial steplength
      [dummy,posdef] = chol( X + alphap * dX );  % test if pos.def
      while posdef > 0,
              alphap = alphap * .8;
              [dummy,posdef] = chol( X + alphap * dX );
              end;
      alphap = alphap * .95;  % stay away from boundary
% line search on dual; dZ is handled implicitly: dZ = diag( dy);
%     alphad = 1;
      alphad = 2;               % initial steplength changed to 2
      [dummy,posdef] = chol( Z + alphad * diag(dy) );
      while posdef > 0;
              alphad = alphad * .8;
              [dummy,posdef] = chol( Z + alphad * diag(dy) );
              end;
      alphad = alphad * .95;
% update
      X = X + alphap * dX;
      y = y + alphad * dy;
      Z = Z + alphad * diag(dy);
      mu = X( :)’ * Z( :) / (2*n);
      if alphap + alphad > 1.8, mu = mu/2; end; % speed up for long steps
      phi = b’ * y; psi = L( :)’ * X( :);
% display current iteration
        disp([ iter alphap alphad (phi-psi) psi  phi ]);

        end;            % end of main loop
```