A short survey on semidefinite programming

de Klerk, E.; Roos, C.; Terlaky, T.

Published in:
Ten years LNMB

Publication date:
1997

Citation for published version (APA):
de Klerk, E., Roos, C., & Terlaky, T. (1997). A short survey on semidefinite programming. In W. K. Klein Haneveld, O. J. Vrieze, & L. C. M. Kallenberg (Eds.), *Ten years LNMB: Ph.D. research and graduate courses of the Dutch Network of Operations Research* (pp. 323-339). (CWI tract; No. 122). Centrum voor Wiskunde en Informatica.

A Short Survey on Semidefinite Programming*

E. de Klerk, C. Roos and T. Terlaky

Faculty of Technical Mathematics and Informatics,

Delft University of Technology,

E-mail: e.deklerk@twi.tudelft.nl, c.roos@twi.tudelft.nl,

t.terlaky@twi.tudelft.nl

Abstract

Semidefinite programming (SDP) is one of the fastest developing branches of mathematical programming. The reason is twofold: efficient solution algorithms for SDP have come to light in the past few years, and SDP finds applications in combinatorial optimization and engineering. In this short survey we show how SDP duality theory can be used to prove classical results, and review the development of interior point algorithms for SDP.

1 Introduction

One could easily be led to believe that the field of semidefinite programming (SDP) originated in this decade. A glance at a bibliography of SDP papers indeed indicates an explosion of research effort, starting around 1991. A closer look reveals that interest in this class of problems is somewhat older, and dates back to the 1960's (see e.g. [6]). A paper on SDP from 1981 is descriptively named *Linear Programming with Matrix Variables* [11], and this apt title may be the best way to introduce the problem.

The goal is to minimize the inner product

$$\langle C, X\rangle := \mathbf{Tr}\,(CX),$$

of two $n \times n$ symmetric matrices, a constant matrix C and a variable matrix X, subject to a set of constraints, where '$\mathbf{Tr}$ ' denotes the trace (sum of diagonal elements) of a matrix.[1] The first of the constraints are linear:

$$\langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m,$$

*Some authors prefer to use the term 'optimization' instead of 'programming'.

[1] This inner product corresponds to the familiar Euclidean inner product of two vectors – if the columns of the two matrices C and X are stacked to form vectors $\mathrm{vec}(X)$ and $\mathrm{vec}(C)$, then $\mathrm{vec}(C)^T\mathrm{vec}(X) = \mathbf{Tr}\ (CX)$.

where the A_i's are given symmetrix matrices, and the b_i's given scalars. Up to this point, the stated problem is merely a linear programming (LP) problem with the entries of X as variables. We now add the convex, nonlinear constraint that X must be symmetrix positive semidefinite, denoted by $X \succeq 0$.[2] The convexity follows from the convexity of the cone of positive semidefinite matrices. The problem under consideration is therefore

$$\min_X \{\mathrm{Tr}\ (CX)\ : \mathrm{Tr}\ (A_i X) = b_i\ (i = 1, \ldots, m),\ X \succeq 0\}$$

The Lagrangian dual of our problem takes the form

$$\max_{y,S} \left\{ b^T y\ :\ \sum_{i=1}^{m} y_i A_i + S = C,\ S \succeq 0,\ y \in \mathbb{R}^m \right\}.$$

The *weak duality theorem* therefore implies that $\mathrm{Tr}\ (CX) - b^T y \geq 0$ for all feasible solutions. Equality holds at optimality if both problems have feasible sets with nonempty interiors (*strong duality*) (see e.g. [5]). The duality theory for SDP is weaker than that of LP. Difficulties associated with general convex programming can occur if the strict feasibility condition is not met; thus a problem can be solvable although its Lagrangian dual is infeasible, or both problems can be solvable but with a positive duality gap at optimality.

SDP problems are of interest for a number of reasons, including

- SDP contains important classes of problems as special cases;
- important applications exist in combinatorial optimization and engineering;
- efficient solution strategies have emerged in the past few years (explaining the resurgence in research interest).

Each of these considerations will be discussed briefly. The discussion will be such as to limit the overlap with previous surveys. We will mainly focus on recent developments of interior point methods for SDP, and on the usefulness of SDP duality theory as a technique of proof.

An excellent survey by Vandenberghe and Boyd [58] deals with basic theory, diverse applications, and potential reduction algorithms (up to 1995). Two more recent surveys which focus more on applications of SDP in combinatorial optimization are by Alizadeh [2] and Ramana and Pardalos [48]. The former also deals with interior point methodology, whilst the latter contains surveys of geometric properties of the SDP feasible set (so-called spectrahedra), as well as complexity and duality theory.

2 Special cases of SDP

If the matrix X is restricted to be diagonal, then the requirement $X \succeq 0$ reduces to the requirement that the diagonal elements of X must be nonnegative. In other words, we once

[2] By definition, for symmetric X one has $X \succeq 0$ if $z^T X z \geq 0$, $\forall z \in \mathbb{R}^n$, or equivalently, if all eigenvalues of X are nonnegative.

again have an LP problem. Optimization problems with convex quadratic constraints are likewise special cases of SDP.[3] This follows from the well-known *Shur complement* trick: if

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

for an invertible A, then the matrix

$$S_{hur} := C - B^T A^{-1} B$$

is called the Shur complement of A in X. One has

$$\begin{aligned} X \succ 0 \quad &\text{if and only if} \quad A \succ 0 \text{ and } S_{hur} \succ 0 \\ \text{If } A \succ 0, \text{ then } X \succeq 0 \quad &\text{if and only if} \quad S_{hur} \succeq 0. \end{aligned}$$

It follows that we can represent the quadratic constraint

$$(Az+b)^T(Az+b) - (c^T z + d) \leq 0, \quad z \in \mathbb{R}^n$$

by the semidefinite constraint

$$\begin{bmatrix} I & Az+b \\ (Az+b)^T & c^T z + d \end{bmatrix} \succeq 0.$$

In the same way, we can represent the *second order cone*, $\|x\|^2 \leq \left(\sum_{i=1}^n x_i\right)^2$, by

$$\begin{bmatrix} \left(\sum_{i=1}^n x_i\right) I & x \\ x^T & \sum_{i=1}^n x_i \end{bmatrix} \succeq 0.$$

Another nonlinear example which arises frequently is

$$\min \left\{ \frac{(c^T x)^2}{d^T x} \; : \; Ax \geq b \right\},$$

where it is known that $d^T x > 0$ if $Ax \geq b$. An equivalent SDP problem is:[4]

$$\min \left\{ t \; : \; \begin{bmatrix} t & c^T x & 0 \\ c^T x & d^T x & 0 \\ 0 & 0 & \operatorname{diag}(Ax - b) \end{bmatrix} \succeq 0 \right\}.$$

Several problems involving matrix norm or eigenvalue minimization may be stated as SDP's. An extensive list of such problems may be found in [58]. A simple example is the classical problem of finding the largest eigenvalue $\lambda_{\max}(A)$ of a symmetric matrix A. The key observation here is that $t \geq \lambda_{\max}(A)$ if and only if $tI - A \succeq 0$. The SDP problem therefore becomes

$$\min \{ t \; : \; tI - A \succeq 0, \quad t \in \mathbb{R} \}.$$

An SDP algorithm for this problem is described in [26].

[3]This includes the well-known convex quadratic programming (QP) problem.

[4]We use the notation 'diag' as follows: for a matrix X, $\operatorname{diag}(X)$ is the vector obtained by extracting the diagonal of X; for a vector x, $\operatorname{diag}(x)$ is the diagonal matrix with the coordinates of x as diagonal elements.

3 Applications in combinatorial optimization

General quadratic optimization problems allow SDP relaxations. The key observation is that

$$x^T Q x = \mathrm{Tr}\,(Q x x^T),$$

for a given matrix Q and vector x. The rank one matrix $X = xx^T$ is positive semidefinite. We can therefore relax the condition $X := xx^T$ to $X \succeq 0$. This relaxation is originally due to Shor [53].

Combinatorial optimization problems can in turn be written as quadratic optimization problems. The condition $x_i \in \{-1, 1\}$ is equivalent to $x_i^2 = 1$, for example.

Lovász and Schrijver [35] considered the generic combinatorial problem

$$q^{\max} = \max\left\{x^T Q x \ : \ x_i \in \{-1, 1\} \ (\forall i)\right\} \tag{1}$$

and suggested the relaxation

$$\bar{q} = \max\left\{\mathrm{Tr}\,(QX) \ : \ \mathrm{diag}\,(X) = e, \ X \succeq 0\right\}. \tag{2}$$

For this general relaxation Nesterov [42] recently proved that

$$\bar{q} - \underline{q} \geq q^{\max} - q^{\min} \geq \frac{4-\pi}{\pi}\left(\bar{q} - \underline{q}\right)$$

where $(q^{\min}, q^{\max})$ is the range of feasible objective values in (1), and $(\underline{q}, \bar{q})$ is the range of feasible values in the relaxation problem (2). Moreover, a random feasible solution x to (1) can be computed from the solution to the relaxation. The expected objective value of x, say $E(x)$, satisfies[5]

$$\frac{q^{\max} - E(x)}{q^{\max} - q^{\min}} < \frac{4}{7}.$$

For specific problems this bound can be improved. The showcase example is the *maximal cut problem*, i.e. the problem of finding a cut of maximal weight through a graph with weighted edges. In a pioneering article, Goemans and Williamson [21] proved that $\bar{q} \leq 1.14 q^{\max}$ in this case. They moreover devised a randomized algorithm which produces a cut with expected value greater than $0.878 q^{\max}$. Similar improvements were also reported in [21] for satisfiability problems.

The SDP relaxations are not always useful, though. Cases where the SDP relaxation is no stronger than the usual LP relaxation are reviewed in [48].

SDP offers more than just a numerical tool to generate lower and upper bounds on optimal values. It also provides a technique of proof via duality theory. We consider the classical *sandwich theorem*, and give a proof (which is new to the best of our knowledge) using strong duality theory. The theorem relates three characterizing numbers of a graph: the colouring number[6] $\chi(G)$, the maximal clique number[7] $\omega(G)$, and the Lovász number $\vartheta(G)$, which will be defined presently.

[5] The same bounds were obtained by Ye [60] for the 'box-constrained' problem where $x_i \in \{-1, 1\}$ is replaced by $-1 \leq x_i \leq 1$ in problem (1).

[6] Number of colours needed to colour all vertices so that no two adjacent vertices share the same colour.

[7] The cardinality of the maximal clique (connected subgraph).

For a graph $G = (V, E)$, a maximal clique is a subset $C \subset V$ with

$$\forall i, j \in C (i \neq j) \ : \{i, j\} \in E,$$

such that $|C|$ is maximal. The Lovász number $\theta(G)$ can be defined[8] as the optimal value of the SDP relaxation (see [34, 23]):

$$\theta(G) := \max \mathbf{Tr} \left(ee^T X \right) = e^T X e \tag{3}$$

subject to

$$\left. \begin{array}{rcl} X_{ij} & = & 0, \ \{i,j\} \notin E \ (i \neq j) \\ \mathbf{Tr}\,(X) & = & 1 \\ X & \succeq & 0. \end{array} \right\} \tag{4}$$

The sandwich theorem states the following.

Theorem 3.1 (Lovász's Sandwich Theorem) *For any graph $G = (V, E)$ one has*

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

Proof:

In order to prove the first inequality of the theorem, let x_C denote a 0-1 vector which defines a clique C of size k in G, i.e:

$$(x_C)_i = \begin{cases} 1 & \text{if } i \in C \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that the rank one matrix

$$X := \frac{1}{k} x_C x_C^T$$

is feasible in (4) with objective value

$$e^T X e = \frac{1}{k} \left(e^T x_C \right)^2 = \frac{k^2}{k} = k.$$

We therefore have $\omega(G) \leq \theta(G)$, which is the first part of the sandwich theorem. The second part is to prove $\theta(G) \leq \chi(G)$. To this end, we write down the Lagrangian dual of the SDP relaxation (4) to obtain

$$\theta(G) = \min \lambda \tag{5}$$

subject to

$$\left. \begin{array}{rcl} Y + ee^T & \preceq & \lambda I \\ Y_{ij} & = & 0, \ \{i,j\} \in E \ (i \neq j) \\ Y_{ii} & = & 0, \ i \in V. \end{array} \right\} \tag{6}$$

[8]Strictly speaking, the definition given here is of the Lovász number of the complement of G (nodes in the complement of G are connected if and only if they are not connected in G).

Given a colouring of G with k colours, we must constuct a feasible solution for (6) with $\lambda \le k$. Such a colouring defines a partition $V = \cup_{i=1}^{k} C_i$ where the C_i's are subsets of nodes sharing the same colour. In other words, the C_i's must be disjoint stable sets (co-cliques). Now let $\gamma_i = |C_i|$ and define

$$M_i := k\left(I_{\gamma_i} - J_{\gamma_i}\right), \quad i = 1, \dots, k,$$

where I_{γ_i} is the $(\gamma_i \times \gamma_i)$ identity matrix, and J_{γ_i} the all-one matrix of the same size.

We will show that the block diagonal matrix

$$Y = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & M_k \end{pmatrix} \tag{7}$$

is feasible in (6) if $\lambda = k$. By construction, Y satisfies the last two constraints in (6). We must still show that $Y + ee^T \preceq kI$, i.e. the largest eigenvalue of $Y + ee^T$ must be at most k.

The Raleigh-Ritz theorem states that for any symmetric matrix A, one has:

$$\lambda_{\max}(A) = \max\left\{x^T A x \;:\; \|x\| = 1\right\}. \tag{8}$$

It follows that the maximal eigenvalue of Y is given by

$$\lambda_{\max}(Y) = \max\left\{\sum_{i=1}^{k} \alpha_i \lambda_{\max}(M_i),\ \sum_{i=1}^{k} \alpha_i = 1,\ \alpha_i \ge 0\ (\forall i)\right\}. \tag{9}$$

Moreover one has $\lambda_{\max}(M_i) = k$, so that (9) yields $\lambda_{\max}(Y) = k$. The eigenvector corresponding to k is orthogonal to the all-one vector e. To see this, note that $Yx = \lambda x$ implies

$$-k(\gamma_i - 1)\sum_{j \in C_i} x_j = \lambda \sum_{j \in C_i} x_j, \quad i = 1, \dots, k,$$

so that $\sum_{j \in C_i} x_j = 0$ $(i = 1, \dots, k)$ if $\lambda > 0$. In particular, $e^T x = 0$ from which it follows that k is also an eigenvalue of $Y + ee^T$. Assuming that k is not the largest eigenvalue of $Y + ee^T$, then the largest eigenvalue must have an eigenspace orthogonal to the eigenspace of k. The orthogonal complement of the eigenspace of k is spanned by the vectors

$$\left(x_{C_i}\right)_j := \begin{cases} 1 & \text{if } j \in C_i \\ 0 & \text{otherwise,} \end{cases}$$

where $i = 1, \dots, k$. The maximal eigenvalue of $Y + ee^T$ can therefore be computed from (8):

$$\begin{aligned} \lambda_{\max}\left(Y + ee^T\right) &= \max_x \left\{x^T\left(Y + ee^T\right)x \;:\; x \in \operatorname{span}\{x_{C_1}, \dots, x_{C_k}\},\ \|x\| = 1\right\} \\ &= \max_\alpha \left\{x^T Y x + \left(e^T x\right)^2 \;:\; x = \sum_{i=1}^{k} \alpha_i x_{C_i},\ \sum_{i=1}^{k} \gamma_i \alpha_i^2 = 1\right\} \end{aligned}$$

Substituting the expression for x, and using the construction of Y simplifies this to

$$\begin{aligned}\lambda_{\max}\left(Y+ee^T\right) &= \max_{\alpha}\left\{-k\sum_{i=1}^{k}\alpha_i^2\left(\gamma_i^2-\gamma_i\right)+\left(\sum_{i=1}^{k}\gamma_i\alpha_i\right)^2 \ : \ \sum_{i=1}^{k}\gamma_i\alpha_i^2=1\right\}\\ &= k+\max_{\alpha}\left\{-k\sum_{i=1}^{k}(\alpha_i\gamma_i)^2+\left(\sum_{i=1}^{k}\gamma_i\alpha_i\right)^2 \ : \ \sum_{i=1}^{k}\gamma_i\alpha_i^2=1\right\}.\end{aligned}$$

The expression in brackets is nonpositive, since it is of the form

$$-kz^Tz+(e^Tz)^2 \leq -kz^Tz+(\|e\|\|z\|)^2=-kz^Tz+k\|z\|^2=0,$$

where $z_i=\alpha_i\gamma_i$, $(i=1,\ldots,k)$. This leads to the contradiction $\lambda_{\max}\left(Y+ee^T\right)\leq k$.

We conclude that $\lambda_{\max}\left(Y+ee^T\right)=k$, as required. □

Moreover, we have given a proof of the equivalence of two different definitions of $\theta(G)$ via (3) and (5).[9]

4 Engineering applications

The richest field of application of SDP is currently *system and control theory.* The standard reference for these problems is Boyd et al. [10]. Introductory examples are given in [58] and [45].

An application which receives less attention is *structural design*, where the best known SDP problem involves optimal truss[10] design. Two variants are:

1. minimize the weight of the structure such that its fundamental frequency remains above a critical value;
2. minimize the worst-case compliance ('stored energy') of the truss given a set of forces which the structure has to withstand.

The second of these problems allows another nice application of SDP duality theory. The problem may be stated as

Displacement formulation

$$\min_{t,x_1,\ldots,x_k}\max_j x_j^Tf_j, \quad j=1,\ldots,k$$

subject to

$$\begin{aligned}\left(\sum_{i=1}^{m}t_ib_ib_i^T\right)x_j &= f_j, \quad j=1,\ldots,k\\ \sum_{i=1}^{m}t_i &= V, \quad t\geq 0,\end{aligned}$$

[9]These and other equivalent definitions of $\theta(G)$ are discussed in [23].

[10]A truss is here defined as a structure of bars which connect a fixed ground structure of nodes. The design is fixed once the sizes of the bars have been decided.

where the t_i's are the bar volumes (design variables), and the f_j's are the set of forces which the truss has to withstand. The displacement of the nodes subject to force f_j is given by the vector x_j. The fixed vectors b_i depend only on the layout of the nodes and on the material properties (Young moduli) of the bars. The first constraint requires equilibrium of the structure and the second fixes its total volume. The objective is to minimize the worst-case compliance.

The name 'displacement' formulation stems from the displacement variables x_j. From engineering considerations, the problem may also be stated by using the forces in the bars as variables.

Bar forces formulation

$$\min_{\beta,t} \max_{j=1,\ldots,k} \sum_{\substack{i=1 \\ t_i>0}}^{m} \frac{\beta_{ij}^2}{t_i}$$

subject to

$$\begin{aligned}
\sum_{i=1}^{m} t_i &= V \\
f_j &= \sum_{i=1}^{m} \beta_{ij} b_i, \quad j=1,\ldots,k \\
t_i &\geq 0, \quad i=1,\ldots,m \\
\beta_{ij}=0 \quad \text{if } t_i &= 0, \quad i=1,\ldots,m, \quad j=1,\ldots,k,
\end{aligned}$$

where β_{ij} is proportional to the reaction force in bar i due to f_j. The second constraint simply requires static equilibrium, i.e. a 'balance of forces'.

From a purely mathematical point of view it is far from obvious that the two formulations are equivalent. This equivalence can be shown using SDP duality. We will sketch the proof here. Using the Shur complement trick, the displacement formulation can be written as an SDP problem (for details, see [12]).

SDP reformulation of the displacement formulation

$$\min \tau$$

subject to

$$\begin{aligned}
\begin{bmatrix} \tau & f_j^T \\ f_j & \left(\sum_{i=1}^{m} t_i b_i b_i^T\right) \end{bmatrix} &\succeq 0, \quad j=1,\ldots,k \\
\sum_{i=1}^{m} t_i &= V, \quad t \geq 0.
\end{aligned}$$

The equivalence proof is now done in three steps:

1: write down the dual of the SDP reformulation and simplify it;

2: obtain the dual of the resulting problem from Step 1;

3: reduce the problem obtained in Step 2 to the 'bar forces' formulation.

This sequence of steps is described detail in [41]. A survey of these and related formulations is given in [12], with emphasis on SDP formulations. Other structural design problems which may be formulated as SDP's include sandwich plate design [8], optimization of variable thickness sheets [51], and minimal compliance design with optimized materials [49]. Good reviews of interior point methods in truss topology design are Bendsøe et al. [9], and Jarre et al. [28] (see also [7]).

Other engineering applications of SDP include: VLSI transistor sizing, pattern recognition using ellipsoids, and logarithmic Chebychev approximation (see [58]).

5 Efficient solution strategies

Bearing the links between LP and SDP in mind, it may come as little surprise that interior point algorithms for LP have been successfully extended to SDP.

The field of interior point methods for LP more or less started with the famous paper by Karmarkar [30] in 1984, and in the following decade more that a thousand papers appeared on this topic. Some recent review papers include [20] and [26]. Several new books on the subject have also appeared recently, including [50] and [59].

The first extension of interior point algorithms from LP to SDP was by Nesterov and Nemirovskii [43], and independently by Alizadeh [1] in 1991. Nesterov and Nemirovskii actually considered a more general class of convex optimization problems, where the nonlinearity is 'banished' to a convex cone, like $X \succeq 0$. They show that such conic optimization problems can be solved by sequential minimization techniques, where the conic constraint is discarded and a barrier term is added to the objective. Suitable barriers are called *self-concordant.* These barriers go to infinity as the boundary of the cone is approached, an can be minimized efficiently by Newton's method.[11] The function

$$f_{bar}(X) = -\log\det(X)$$

is such a barrier for the cone of semidefinite matrices. Using this barrier, several classes of algorithms may be formulated which have polynomial worst-case iteration bounds for the computation of ϵ-optimal solutions.

5.1 Logarithmic barrier methods

Primal log-barrier methods use Newton's method to solve a sequence of problems of the form

$$\min_X \{\mathbf{Tr}\,(CX) - \mu\log\det(X)\ :\ \mathbf{Tr}\,(A_iX) = b_i\ (i = 1,\ldots,m)\}$$

where the parameter μ is sequentially decreased to zero. Such algorithms were analysed by Faybusovich in [17, 18] and later by other authors in [24] and [4]. Note that the condition $X \succeq 0$ has been replaced by adding a 'barrier term' to the objective.[12] The condition $X \succeq 0$

[11]The definition of self-concordant barriers is omitted here; a well-written introductory text is [27].

[12]This idea actually dates back to the 1960's and the work of Fiacco and McCormick [19]; the implications for complexity theory only became clear almost three decades later.

is maintained by controlling the Newton process carefully – large decreases of μ necessitate damped Newton steps (see e.g. [4]), while small updates allow full Newton steps (see e.g. [24]).

Following the trend in LP, so-called *primal-dual* methods soon became more popular. These methods minimize the duality gap

$$\mathbf{Tr}\,(CX) - b^T y = \mathbf{Tr}\,(XS)$$

and employ the combined primal-dual barrier function

$$f_{pd} := -(\log\det(X) + \log\det(S)) = -\log\det(XS).$$

This means that a sequence of problems of the following form are solved

$$\min_{X,y,S}\left\{\mathbf{Tr}\,(XS) - \mu\log\det(XS)\ :\ \mathbf{Tr}\,(A_iX) = b_i\ (i = 1,\ldots,m),\ \sum_{i=1}^{m} y_iA_i + S = C\right\}. \quad (10)$$

The first order optimality conditions for (10) are

$$\left.\begin{array}{rcl} \mathbf{Tr}\,(A_iX) &=& b_i,\quad i = 1,\ldots,m \\ \sum_{i=1}^{m} y_iA_i + S &=& C \\ XS &=& \mu I \\ X,S &\succ& 0. \end{array}\right\} \quad (11)$$

This system has a unique positive definite solution pair, denoted by $X(\mu) \succ 0$ and $S(\mu) \succ 0$.[13] Primal-dual log-barrier methods solve the system (11) approximately, followed by a reduction in μ. Ideally, the goal is to obtain primal and dual steps ΔX and ΔS, respectively, which satisfy $X + \Delta X \succeq 0$, $S + \Delta S \succeq 0$ and

$$\begin{array}{rcl} \mathbf{Tr}\,(A_i\Delta X) &=& 0,\quad i = 1,\ldots,m \\ \sum_{i=1}^{m} \Delta y_iA_i + \Delta S &=& 0 \\ (X + \Delta X)(S + \Delta S) &=& \mu I. \end{array} \quad (12)$$

The last equation is nonlinear, and primal-dual methods differ with regard to how it is linearized. Moreover, care must be taken that the solution matrices ΔX and ΔS are symmetrical. Zhang [61] suggested to replace the nonlinear equation by

$$L\,(\Delta XS + X\Delta S)\,L^{-1} + \left[L\,(\Delta XS + X\Delta S)\,L^{-1}\right]^T \;=\; 2\mu I - L\,(XS)\,L^{-1} + \left[L(XS)L^{-1}\right]^T,$$

where the matrix L determines the symmetrization strategy. Some popular choices for L are listed in Table 1. The proof of the existence and uniqueness of each of the resulting search directions was done by Shidah et al. in [52].[14] Other properties (such as scale-invariance) are compared by Todd et al. in [56].

[13] These solutions give a parametric representation of a smooth curve, called the *central path*, which tends to the *analytic center* of the primal-dual optimal sets as $\mu \to 0$. This was proved by Goldfarb and Scheinberg in [22].

[14] For $L = I$ uniqueness is not always guaranteed; a sufficient condition for uniqueness is $XS + SX \succeq 0$.

L	Reference
$\left[X^{\frac{1}{2}}\left(X^{\frac{1}{2}}SX^{\frac{1}{2}}\right)^{-\frac{1}{2}}X^{\frac{1}{2}}\right]^{\frac{1}{2}}$	Nesterov and Todd [44];
$X^{-\frac{1}{2}}$	Monteiro [38], Kojima et al. [33];
$S^{\frac{1}{2}}$	Monteiro [38], Helmberg et al. [25], Kojima et al. [33];
I	Alizadeh, Haeberley and Overton [3];

Table 1: Choices for the linearization matrix L.

The conspicuous entry $L = \left[X^{\frac{1}{2}}\left(X^{\frac{1}{2}}SX^{\frac{1}{2}}\right)^{-\frac{1}{2}}X^{\frac{1}{2}}\right]^{\frac{1}{2}}$ in Table 1 warrants some comment. Nesterov and Todd [44] showed[15] that for each pair $X \succ 0$, $S \succ 0$ there exists a matrix D such that

$$f''_{bar}(D)X = S.$$

It can be shown that $f''_{bar}(D)$ is the linear operator which satisfies $f''_{bar}(D) : X \mapsto D^{-1}XD^{-1}$. It follows that $X = DSD$, from which it easily follows that $D = L^2$. In this way we obtain the *symmetric primal-dual scaling* $L^{-1}XL^{-1} = LSL$. This symmetry explains the usefulness of D in symmetrization.

Algorithms differ in how μ is updated, and how the symmetrized equations are solved. Methods with use large reductions of μ followed by several damped Newton steps are called long step (or large update) methods. These are analysed in [29], [38], and [55].

Methods which use dynamic updates of μ include the popular *predictor-corrector* methods. References include [3, 31, 32, 46, 54]. Other dynamic μ-updates are described in [14]. Superlinear convergence properties of predictor-corrector schemes are studied in [31, 37].

5.2 Primal-dual potential reduction methods

These algorithms are based on the potential function

$$\phi(X,S) = (n + \nu\sqrt{n})\mathrm{Tr}\,(XS) - \log\det(XS) - n\log n,$$

where $\nu \geq 1$. In order to obtain a polynomial complexity bound it is sufficient to show that ϕ can be reduced by an absolute constant at each iteration [57]. A survey of algorithms which achieve such a reduction is given in [58].

[15]This result was proved in the more general setting of optimization problems where the variable is restricted to a self-dual cone which allows a special type of self-concordant barrier, namely self-scaled barriers. The interested reader is referred to [44].

5.3 Affine–scaling methods

The primal affine–scaling direction for SDP minimizes the primal objective over an ellipsoid which is inscribed in the primal feasible region. Surprisingly, Muramatsu [39] has shown that an algorithm using this search direction may converge to a non-optimal point, regardless of which step length is used. This is in sharp contrast to the LP case, and shows that extension of algorithms from LP to SDP cannot always be taken for granted.

Two primal-dual variants of the affine scaling methods were extended by De Klerk et al. in [15] from LP to SDP. These algorithms minimize the duality gap over ellipsoids in the scaled primal-dual space, where the matrix $L = D^{\frac{1}{2}}$ is used for the scaling. The primal-dual method fails if either of the scalings $L = X^{\frac{1}{2}}$ or $L = S^{\frac{1}{2}}$ from Table 1 is used [40].

5.4 Infeasible start methods

Several infeasible start algorithms have been suggested. A review of traditional big-M initialization strategies may be found in [58]. One of the first infeasible-start predictor-corrector algorithms was by Potra and Sheng [46]. Other references include [31, 37].

The idea of embedding the SDP problem in a self-dual problem with known feasible starting point was investigated for SDP in [13] and [36]. A solution of the self-dual embedding gives information about the solution of the original problem. This analysis was extended in [16] to include pathological cases caused by the weaker duality theory of SDP (as compared to LP). In the latter case the stronger ELSD (extended Lagrange-Slater) dual problem is used in the embedding. These duals have better properties than the usual Lagrangean duals, and were formulated by Ramana [47].

6 Further information

An up-to-date list of publications dealing with SDP may be found in the *semidefinite programming homepage*, maintained by Christolph Helmberg. The address is

`http://www.zib-berlin.de/~bzfhelmb/semidef.html`

Available SDP software can also be accessed via this address.

References

[1] F. Alizadeh. *Combinatorial optimization with interior point methods and semi-definite matrices.* PhD thesis, University of Minnesota, Minneapolis, USA, 1991.

[2] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization* 5, 13–51, 1995.

[3] F. Alizadeh, J.-P.A. Haeberley, and M.L. Overton. Primal-dual methods for semidefinite programming: convergence rates, stability and numerical results. Technical report 721, NYU Computer Science Dept, New York University, New York, NY, 1996.

[4] K.M. Anstreicher and M. Fampa. A long-step path following algorithm for semidefinite programming problems. Working Paper, Department of Management Sciences, University of Iowa, Iowa City, USA, 1996.

[5] M.S. Bazarraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms.* John Wiley and Sons, New York, 1993.

[6] R. Bellman and K. Fan. On systems of linear inequalities in hermitian matrix variables. In V.L. Klee, editor, *Convexity*, Vol. 7 Proc. Symposia in Pure Mathematics, pages 1–11. Amer. Math. Soc., Providence, RI, 1963.

[7] A. Ben-Tal and A. Nemirovskii. Potential reduction polynomial time method for truss topology design. *SIAM Journal on Optimization* 4(3), 596–612, 1994.

[8] M.P. Bendsøe. *Optimization of Structural Topology, Shape, and Material.* Springer, 1995.

[9] M.P. Bendsøe, A. Ben-Tal, and J. Zowe. Optimization methods for truss geometry and topology design. *Structural Optimization* 7, 141–159, 1994.

[10] S.E. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear matrix inequalities in system and control theory.* SIAM Studies in Applied Mathematics, Vol. 15. SIAM, Philadelphia, USA, 1994.

[11] B. Craven and B. Mond. Linear programming with matrix variables. *Linear Algebra Appl.* **38**, 73–80, 1981.

[12] E. de Klerk, C. Roos, and T. Terlaky. Semi-definite problems in truss topology optimization. Technical Report 95–128, Faculty of Technical Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, 1995.

[13] E. de Klerk, C. Roos, and T. Terlaky. Initialization in semidefinite programming via a self-dual, skew-symmetric embedding. Technical Report 96–10, Faculty of Technical Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, 1996. (To appear in *OR Letters*).

[14] E. de Klerk, C. Roos, and T. Terlaky. On primal-dual path-following algorithms in semidefinite programming. Technical Report 96–102, Faculty of Technical Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, 1996. (To appear *Proceedings of XIIIth Int. Conf. Math. Prog.*, Kluwer).

[15] E. de Klerk, C. Roos, and T. Terlaky. Polynomial primal-dual affine scaling algorithms in semidefinite programming. Technical Report 96–42, Faculty of Technical Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, 1996. (To appear *J. Comb. Opt.*).

[16] E. de Klerk, C. Roos, and T. Terlaky. Infeasible–start semidefinite programming algorithms via self-dual embeddings. Technical Report 97–10, Faculty of Technical Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands,

1997. (To appear in Proceedings of the Workshop on Semidefinite Programming and Interior-Point Approaches for Combinatorial Optimization Problems, FIELDS Institute, May 15-17, 1996).

[17] L. Faybusovich. On a matrix generalization of affine-scaling vector fields. *SIAM J. Matrix Anal. Appl.* **16**, 886–897, 1995.

[18] L. Faybusovich. Semi-definite programming: a path-following algorithm for a linear-quadratic functional. *SIAM Journal on Optimization*, pages 1007–1024, 1996.

[19] A.V. Fiacco and G.P. McCormick. *Nonlinear programming: sequential unconstrained minimization techniques.* John Wiley & Sons, New York, 1968. (Reprint: Volume 4 of *SIAM Classics in Applied Mathematics*, SIAM Publications, Philadelphia, USA, 1990).

[20] R.M. Freund and S. Mizuno. Interior point methods: current status and future directions. *OPTIMA* **51**, 1996.

[21] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of Assoc. Comput. Mach.* **42**, 1115–1145, 1995.

[22] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programming. Working Paper, Dept. of IEOR, Columbia University, New York, NY, 1996.

[23] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization.* Springer–Verlag, Berlin, 1988.

[24] B. He, E. de Klerk, C. Roos, and T. Terlaky. Method of approximate centers for semidefinite programming. Technical Report 96–27, Faculty of Technical Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands, 1996. (To appear in *Optimization Methods and Software*).

[25] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM Journal on Optimization* **6**, 342–361, 1996.

[26] B. Jansen, C. Roos, and T. Terlaky. Interior point methods: a decade after Karmarkar. A survey, with application to the smallest eigenvalue problem. *Statistica Neerlandica* **50**, 1995.

[27] F. Jarre. Interior-point methods for classes of convex programs. In T. Terlaky, editor, *Interior point methods of mathematical programming*, pages 255–296. Kluwer, Dordrecht, The Netherlands, 1996.

[28] F. Jarre, M. Kočvara, and J. Zowe. Optimal truss design by interior point methods. Technical Report 173, Institut für Angewandte Mathematik, Universität Erlangen-Nürnberg, Erlangen, Germany, 1996.

[29] J. Jiang. A long step primal–dual path following method for semidefinite programming. Technical Report, Dept of Applied Mathematics, Tsinghua University, Beijing, China, 1996. (To appear in *OR Letters*).

[30] N.K. Karmarkar. A new polynomial–time algorithm for linear programming. *Combinatorica* **4**, 373–395, 1984.

[31] M. Kojima, M. Shida, and S. Shindoh. Global and local convergence of predictor-corrector infeasible-interior-point algorithms for semidefinite programs. Technical Report B-305, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1995.

[32] M. Kojima, M. Shidah, and S. Shindoh. A predictor-corrector interior point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberley-Overton search direction. Technical Report B-311, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1996.

[33] M. Kojima, S. Shindoh, and S. Hara. Interior point methods for the monotone semidefinite linear complementarity problems. Technical Report No. 282, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1994. (To appear in *SIAM Journal on Optimization*).

[34] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. on Information Theory* **25**, 1–7, 1979.

[35] L. Lovász and A. Schrijver. Cones of matrices and set–functions and 0-1 optimization. *SIAM Journal on Optimization* **1**, 166–190, 1991.

[36] Z.-Q. Luo, J.F. Sturm, and S. Zhang. Duality and self-duality for conic convex programming. Technical Report 9620/A, Tinbergen Institute, Erasmus University Rotterdam, 1996.

[37] Z.-Q. Luo, J.F. Sturm, and S. Zhang. Superlinear convergence oa a Symmetric primal–dual path following algorithm for semidefinite programming. Technical Report 9607/A, Tinbergen Institute, Erasmus University Rotterdam, 1996.

[38] R.D.C. Monteiro. Primal-dual algorithms for semidefinite programming. Working Paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, USA, 1995. (To appear in *SIAM Journal on Optimization*).

[39] M. Muramatsu. Affine scaling algorithm fails for semidefinite programming. Working Paper 16, Dept. of Mechanical Engineering, Sophia University, Tokyo, Japan, 1996.

[40] M. Muramatsu and R.J. Vanderbei. Primal-dual affine-scaling algorithm fails for semidefinite programming. Technical Report SOR 97-04, Princeton University, NJ 08544, 1997.

[41] A. Nemirovskii. Interior point polynomial time methods in convex programming. Course notes, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion–Israel Institute of Technology, Haifa, Israel, 1994.

[42] Yu. Nesterov. Quality of semidefinite relaxation for nonconvex quadratic optimization. CORE Discussion paper 9719, Belgium, March 1997.

[43] Yu. Nesterov and A.S. Nemirovskii. *Interior point polynomial algorithms in convex programming.* SIAM Studies in Applied Mathematics, Vol. 13. SIAM, Philadelphia, USA, 1994.

[44] Yu. Nesterov and M.J. Todd. Self-scaled barriers and interior–point methods for convex rogramming. Technical Report 1091, School of OR and IE, Cornell University, Ithaca, New York, USA, 1994. (To appear in *Mathematics of Operations Research*).

[45] J.A. Olkin. Using semi–definite programming for controller design in active noise control. *SIAG/OPT Views and News* **8**, 1–5, Fall 1996.

[46] F.A. Potra and R. Sheng. A superlinearly convergent primal–dual infeasible–interior-point algorithm for semidefinite programming. Reports on Computational Mathematics 78, Dept. of Mathematics, The University of Iowa, Iowa City, USA, 1995.

[47] M. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming* **77**(2), 1997.

[48] M.V. Ramana and P.M. Pardalos. Semidefinite programming. In T. Terlaky, editor, *Interior point methods of mathematical programming*, pages 369–398. Kluwer, Dordrecht, The Netherlands, 1996.

[49] U.T. Ringertz. On finding the optimal distribution of material properties. *Structural Optimization* **5**, 265–267, 1993.

[50] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization: An interior point approach.* John Wiley & Sons, New York, 1997.

[51] M.P. Rossow and J.E. Taylor. A finite element method for the optimal design of variable thickness sheets. *AIAA Journal* **11**(11), 1566–1569, 1973.

[52] M. Shida, S. Shindoh, and M. Kojima. Existence of search directions in interior–point algorithms for the SDP and the monotone SDLCP. Manuscript, 1996. (To appear in *SIAM Journal on Optimization*).

[53] N.Z. Shor. Quadratic optimization problems. *Soviet Journal of Computer and System Sciences* **25**, 1–11, 1987.

[54] J.F. Sturm and S. Zhang. Symmetric primal–dual path following algorithms for semidefinite programming. Technical Report 9554/A, Tinbergen Institute, Erasmus University Rotterdam, 1995.

[55] J.F. Sturm and S. Zhang. On the long step path–following method for semidefinite programming. Technical Report 9638/A, Tinbergen Institute, Erasmus University Rotterdam, 1996.

[56] M.J. Todd, K.C. Toh, and R.H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. Working paper, School of OR and Industrial Engineering, Cornell University, Ithaca, New York 14853-3801, 1996.

[57] L. Vandenberghe and S. Boyd. A primal–dual potential reduction algorithm for problems involving matrix inequalities. *Math. Programming* **69**, 205–236, 1995.

[58] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review* **38**, 49–95, 1996.

[59] S.J. Wright. *Primal-dual interior point methods.* SIAM, Philadelphia, 1997.

[60] Y. Ye. Approximating quadratic programming with bound constraints. Working paper, Dept. of Management Science, University of Iowa, Iowa City, USA, 1997.

[61] Y. Zhang. On extending primal-dual interior point algorithms from linear programming to semidefinite programming. Technical Report, Department of Mathematics and Statistics, University of Maryland at Baltimore County, Baltimore, USA, 1995.