

Aspects of Semidefinite Programming

Applied Optimization

Volume 65

Series Editors:

Panos M. Pardalos
University of Florida, U.S.A.

Donald Hearn
University of Florida, U.S.A.

The titles published in this series are listed at the end of this volume.

Aspects of Semidefinite Programming

Interior Point Algorithms and Selected Applications

by

Etienne de Klerk

*Delft University of Technology,
Delft, The Netherlands*

KLUWER ACADEMIC PUBLISHERS

NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 0-306-47819-6
Print ISBN: 1-4020-0547-4

©2004 Kluwer Academic Publishers
New York, Boston, Dordrecht, London, Moscow

Print ©2002 Kluwer Academic Publishers
Dordrecht

All rights reserved

No part of this eBook may be reproduced or transmitted in any form or by any means, electronic, mechanical, recording, or otherwise, without written consent from the Publisher

Created in the United States of America

Visit Kluwer Online at: <http://kluweronline.com>
and Kluwer's eBookstore at: <http://ebooks.kluweronline.com>

Contents

Acknowledgments	ix
Foreword	xi
List of Notation	xiii
1. INTRODUCTION	1
1.1 Problem statement	1
1.2 The importance of semidefinite programming	3
1.3 Special cases of semidefinite programming	3
1.4 Applications in combinatorial optimization	4
1.5 Applications in approximation theory	8
1.6 Engineering applications	10
1.7 Interior point methods	11
1.8 Other algorithms for SDP	16
1.9 The complexity of SDP	17
1.10 Review literature and internet resources	18
Part I Theory and Algorithms	
2. DUALITY, OPTIMALITY, AND DEGENERACY	21
2.1 Problems in standard form	22
2.2 Weak and strong duality	25
2.3 Feasibility issues	29
2.4 Optimality and complementarity	32
2.5 Degeneracy	36
3. THE CENTRAL PATH	41
3.1 Existence and uniqueness of the central path	41
3.2 Analyticity of the central path	46
3.3 Limit points of the central path	48
3.4 Convergence in the case of strict complementarity	51
3.5 Convergence proof in the absence of strict complementarity	56
3.6 Convergence rate of the central path	58

4. SELF-DUAL EMBEDDINGS	61
4.1 Introduction	61
4.2 The embedding strategy	64
4.3 Solving the embedding problem	67
4.4 Interpreting the solution of the embedding problem	68
4.5 Separating small and large variables	70
4.6 Remaining duality and feasibility issues	72
5. THE PRIMAL LOGARITHMIC BARRIER METHOD	75
5.1 Introduction	75
5.2 A centrality function	77
5.3 The projected Newton direction for the primal barrier function	78
5.4 The affine-scaling direction	81
5.5 Behaviour near the central path	83
5.6 Updating the centering parameter	85
5.7 Complexity analysis	87
5.8 The dual algorithm	89
5.9 Large update methods	90
5.10 The dual method and combinatorial relaxations	93
6. PRIMAL-DUAL AFFINE-SCALING METHODS	95
6.1 Introduction	95
6.2 The Nesterov-Todd (NT) scaling	97
6.3 The primal-dual affine-scaling and Dikin-type methods	99
6.4 Functions of centrality	102
6.5 Feasibility of the Dikin-type step	104
6.6 Complexity analysis for the Dikin-type method	107
6.7 Analysis of the primal-dual affine-scaling method	109
7. PRIMAL-DUAL PATH-FOLLOWING METHODS	115
7.1 The path-following approach	115
7.2 Feasibility of the full NT step	118
7.3 Quadratic convergence to the central path	120
7.4 Updating the barrierparameter μ	123
7.5 A long step path-following method	124
7.6 Predictor-corrector methods	125
8. PRIMAL-DUAL POTENTIAL REDUCTION METHODS	133
8.1 Introduction	134
8.2 Determining step lengths via plane searches of the potential	135
8.3 The centrality function Ψ	136
8.4 Complexity analysis in a potential reduction framework	137
8.5 A bound on the potential reduction	139
8.6 The potential reduction method of Nesterov and Todd	142

Part II Selected Applications

9. CONVEX QUADRATIC APPROXIMATION	149
9.1 Preliminaries	149
9.2 Quadratic approximation in the univariate case	150
9.3 Quadratic approximation for the multivariate case	152
10. THE LOVÁSZ ϑ -FUNCTION	157
10.1 Introduction	157
10.2 The sandwich theorem	158
10.3 Other formulations of the ϑ -function	162
10.4 The Shannon capacity of a graph	164
11. GRAPH COLOURING AND THE MAX- k -CUT PROBLEM	169
11.1 Introduction	171
11.2 The ϑ -approximation of the chromatic number	172
11.3 An upper bound for the optimal value of MAX- k -CUT	172
11.4 Approximation guarantees	174
11.5 A randomized MAX- k -CUT algorithm	175
11.6 Analysis of the algorithm	178
11.7 Approximation results for MAX- k -CUT	179
11.8 Approximate colouring of κ -colourable graphs	183
12. THE STABILITY NUMBER OF A GRAPH	187
12.1 Preliminaries	188
12.2 The stability number via copositive programming	189
12.3 Approximations of the copositive cone	193
12.4 Application to the maximum stable set problem	198
12.5 The strength of low-order approximations	201
12.6 Related literature	204
12.7 Extensions: standard quadratic optimization	204
13. THE SATISFIABILITY PROBLEM	211
13.1 Introduction	212
13.2 Boolean quadratic representations of clauses	214
13.3 From Boolean quadratic inequality to SDP	215
13.4 Detecting unsatisfiability of some classes of formulae	218
13.5 Rounding procedures	223
13.6 Approximation guarantees for the rounding schemes	224
Appendices	229
A– Properties of positive (semi)definite matrices	229
A.1 Characterizations of positive (semi)definiteness	229
A.2 Spectral properties	230
A.3 The trace operator and the Frobenius norm	232

A.3	The trace operator and the Frobenius norm	232
A.4	The Löwner partial order and the Schur complement theorem	235
B–	Background material on convex optimization	237
B.1	Convex analysis	237
B.2	Duality in convex optimization	240
B.3	The KKT optimality conditions	242
C–	The function $\log \det(X)$	243
D–	Real analytic functions	247
E–	The (symmetric) Kronecker product	249
E.1	The Kronecker product	249
E.2	The symmetric Kronecker product	251
F–	Search directions for the embedding problem	257
G–	Regularized duals	261
	References	265
	Index	279

Acknowledgments

I would like to thank Tamas Terlaky for his encouragement to publish a revised version of my PhD thesis in the Kluwer Applied Optimization series. Several people have assisted me by proof-reading preliminary versions of the manuscript. In particular, I am indebted to Aharon Ben-Tal, Maria Gonzalez, Margareta Halická, Dorota Kurowicka, Hans van Maaren, Renato Monteiro, Dima Pasechnik and Kees Roos in this regard.

I also thank Jos Sturm for patiently answering many questions about the convergence behaviour of the central path.

I thank Lucie Aarts and Margareta Halická for sharing their notes on which I based Appendix E and Appendix D respectively. Florin Barb I thank for moral and technical support.

Much of the material presented here is based on joint research, and I gratefully acknowledge the contributions of my co-authors: Immanuel Bomze, Margareta Halická, Bingsheng He, Dick den Hertog, Hans van Maaren, Dima Pasechik, Kees Roos, Tamas Terlaky, and Joost Warners.

Finally, I thank John Martindale, Angela Quilici and Edwin Beschler at Kluwer Academic Publishers for their patience and assistance.

Foreword

This monograph has grown from my PhD thesis *Interior point Methods for Semidefinite Programming* [39] which was published in December 1997. Since that time, Semidefinite Programming (SDP) has remained a popular research topic and the associated body of literature has grown considerably. As SDP has proved such a useful tool in many applications, like systems and control theory and combinatorial optimization, there is a growing number of people who would like to learn more about this field.

My goal with this monograph is to provide a personal view on the theory and applications of SDP in such a way that the reader will be equipped to read the relevant research literature and explore new avenues of research. Thus I treat a selected number of topics in depth, and provide references for further reading. The chapters are structured in such a way that the monograph can be used for a graduate course on SDP.

With regard to algorithms, I have focused mainly on methods involving the so-called Nesterov–Todd (NT) direction in some way. As for applications, I have selected interesting ones — mainly in combinatorial optimization — that are not extensively covered in the existing review literature.

In making these choices I hasten to acknowledge that much of the algorithmic analysis can be done in a more general setting (*i.e.*, working with self-concordant barriers, self-dual cones and Euclidean Jordan algebras). I only consider real symmetric positive semidefinite matrix variables in this book; this already allows a wealth of applications.

Required background

The reader is expected to have some background on the following topics:

- linear algebra and multivariate calculus;
- linear and non-linear programming;
- basic convex analysis;
- combinatorial optimization or complexity theory.

I have provided appendices on the necessary matrix analysis (Appendix A), matrix calculus (Appendix C) and convex analysis & optimization (Appendix B), in an attempt to make this volume self-contained to a large degree. Nevertheless, when reading this book it will be handy to have access to the following textbooks, as I refer to them frequently.

- *Convex Analysis* by Rockafellar [160];
- *Matrix Analysis* and *Topics in Matrix Analysis* by Horn and Johnson [85, 86];
- *Nonlinear Programming: Theory and Algorithms* by Bazarraa, Sherali and Shetty [16];
- *Theory and Algorithms for Linear Optimization: An interior point approach*, by Roos, Terlaky and Vial [161];
- *Randomized Algorithms* by Motwani and Raghavan [128].

In particular, the analysis of interior point algorithms presented in this monograph owes much to the analysis in the book by Roos, Terlaky and Vial [161].

Moreover, I do not give an introduction to complexity theory or to randomized algorithms; the reader is referred to the excellent text by Motwani and Raghavan [128] for such an introduction.

List of Notation

Matrix notation

- A^T : transpose of $A \in \mathbf{R}^{m \times n}$;
 A^{-T} : $(A^T)^{-1}$;
 a_{ij} : ij th entry of $A \in \mathbf{R}^{m \times n}$;
 $A \sim B$: $A = T^{-1}BT$ for some nonsingular T
 \equiv the matrices A and B are similar;
 $A \succeq B$ ($A \succ B$) : $A - B$ is symmetric positive semidefinite (positive definite);
 $A \preceq B$ ($A \prec B$) : $A - B$ is symmetric negative semidefinite (negative definite);
 $\mathcal{R}(A)$: range (column space) of $A \in \mathbf{R}^{n \times n}$;

Special vectors and matrices

- I_r : $r \times r$ identity matrix;
 I : identity matrix of size depending on the context;
 $0_{m \times n}$: $m \times n$ zero matrix;
 0_n : zero vector in \mathbf{R}^n ;
 0 : zero vector/matrix of size depending on the context;
 e_n : vector of all-ones in \mathbf{R}^n ;
 e : vector of all-ones of size depending on the context;

Sets of matrices and vectors

- \mathbf{R}^n : n -dimensional real Euclidian vector space;
 \mathbf{R}_+^n : positive orthant of \mathbf{R}^n ;
 \mathbb{Z}^n : n -tuples of integers
 \mathbb{Z}_+^n : n -tuples of *nonnegative* integers
 $\mathbf{R}^{n \times n}$: space of real ($n \times n$) matrices;
 $\mathcal{S}_n = \{X \mid X \in \mathbf{R}^{n \times n}, X = X^T\}$;
 $\mathcal{S}_n^+ = \{X \mid X \in \mathcal{S}_n, X \succeq 0\}$;
 $\mathcal{S}_n^{++} = \{X \mid X \in \mathcal{S}_n, X \succ 0\}$;
 $\mathcal{C}_n = \{A \in \mathcal{S}_n \mid x^T A x \geq 0 \quad \forall x \in \mathbf{R}_+^n\}$ (Copositive matrices);
 $\mathcal{N}_n = \{A \in \mathcal{S}_n \mid a_{ij} \geq 0 \quad \forall i, j = 1, \dots, n\}$ (Nonnegative matrices);
 $\text{svec}(\mathcal{S}_n^+) = \left\{x \in \mathbf{R}^{\frac{1}{2}n(n+1)} \mid \text{smat}(x) \in \mathcal{S}_n^+\right\}$;
 $\{-1, 1\}^n = \{x \in \mathbf{R}^n \mid x_i \in \{-1, 1\}, i = 1, \dots, n\}$;

Functions of matrices

- $\lambda_i(A)$: i th largest eigenvalue of A , if $\lambda_j(A) \in \mathbf{R} \quad \forall j$;
 $\lambda_{\max}(A) = \max_i \lambda_i(A)$, if $\lambda_i(A) \in \mathbf{R} \quad \forall i$;
 $\lambda_{\min}(A) = \min_i \lambda_i(A)$, if $\lambda_i(A) \in \mathbf{R} \quad \forall i$;
 $\mathbf{Tr}(A) = \sum_i a_{ii} = \sum_i \lambda_i(A)$ (trace of $A \in \mathbf{R}^{n \times n}$);
 $\langle A, B \rangle = \mathbf{Tr}(AB^T)$;
 $\det(A)$: determinant of $A \in \mathbf{R}^{n \times n} = \prod_i \lambda_i(A)$;
 $\|A\|^2 = \mathbf{Tr}(AA^T) = \sum_i \sum_j a_{ij}^2$ (Frobenius norm)
 $= \sum_i \lambda_i^2(A)$ if $A \in \mathcal{S}_n$;
 $\|A\|_2 = (\lambda_{\max}(A^T A))^{\frac{1}{2}}$ (spectral norm)
 $= \lambda_{\max}(A)$ if $A \succeq 0$;
 $\rho(A) = \max_i |\lambda_i(A)|$ (spectral radius of A);
 $\kappa(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$ if $\lambda_i(A) > 0 \quad \forall i$
 $=$ condition number of A if $A \succ 0$;
 $A^{\frac{1}{2}}$: unique symmetric square root factor of $A \succeq 0$;
 $\text{Diag}(x)$: $n \times n$ diagonal matrix with components of $x \in \mathbf{R}^n$ on diagonal;
 $\text{diag}(X)$: n -vector obtained by extracting diagonal of $X \in \mathbf{R}^{n \times n}$;
 $\mathbf{vec}(A) = [a_{11}, a_{21}, \dots, a_{n1}, a_{12}, a_{22}, \dots, a_{nn}]^T$ for $A \in \mathbf{R}^{n \times n}$;
 $\mathbf{svec}(A) = [a_{11}, \sqrt{2}a_{12}, \dots, \sqrt{2}a_{1n}, a_{22}, \sqrt{2}a_{23}, \dots, a_{nn}]^T$ for $A \in \mathcal{S}_n$;
 \mathbf{smat} : inverse operator of \mathbf{svec} ;

Set notation

- $\text{ri}(\mathcal{C})$: relative interior of a convex set \mathcal{C} ;
 $\dim(\mathcal{L})$: dimension of a subspace \mathcal{L} ;
 \mathcal{C}^* : Dual cone of a cone $\mathcal{C} \subset \mathbf{R}^n$
 $= \{x \in \mathbf{R}^n \mid x^T y \geq 0 \quad \forall y \in \mathcal{C}\}$;

SDP problems in standard form

- (P) : primal problem in standard form;
 (D) : Lagrangian dual problem of (P) ;
 \mathcal{P} : feasible set of problem (P) ;
 \mathcal{D} : feasible set of problem (D) ;
 \mathcal{P}^* : optimal set of problem (P) ;
 \mathcal{D}^* : optimal set of problem (D) ;
 (P_{gf}) : ELSD dual of (D) ;
 (D_{cor}) : Lagrangian dual of (P_{gf}) ;
 $\mathcal{L} = \text{span}\{A_1, \dots, A_m\}$;

Interior point analysis

- $\log(t)$: natural logarithm of $t > 0$;
 $\psi(t) = t - \log(1+t) \quad t > -1$;
 $f_p^\mu(X) = \frac{1}{\mu} \text{Tr}(CX) - \log \det X$;
 $f_d^\mu(S, y) = \frac{1}{\mu} b^T y + \log \det(S)$;
 $f_{pd}^\mu(X, S) = \text{Tr}\left(\frac{XS}{\mu}\right) - \log \det\left(\frac{XS}{\mu}\right) - n$;
 $D = \left[X^{\frac{1}{2}} \left(X^{\frac{1}{2}} S X^{\frac{1}{2}} \right)^{-\frac{1}{2}} X^{\frac{1}{2}} \right]^{\frac{1}{2}} \quad \text{where } (X, S) \in \text{ri}(\mathcal{P} \times \mathcal{D})$
 \equiv Nesterov–Todd scaling matrix;
 $V = D^{-\frac{1}{2}} X D^{-\frac{1}{2}} = D^{\frac{1}{2}} S D^{\frac{1}{2}}$;
 $\delta(X, S, \mu) = \frac{1}{2} \left\| \sqrt{\mu} V^{-1} - \frac{1}{\sqrt{\mu}} V \right\|$;
 $D_X = D^{-\frac{1}{2}} \Delta X D^{-\frac{1}{2}} \quad (\Delta X \in \mathcal{L}^\perp)$;
 $D_S = D^{\frac{1}{2}} \Delta S D^{\frac{1}{2}} \quad (\Delta S \in \mathcal{L})$;
 $\Psi(X, S) = -\log \det(XS) + n \log \text{Tr}(XS) - n \log n$ (Centrality function);
 $\Phi(X, S) = (n + \nu\sqrt{n}) \log \text{Tr}(XS) - \log \det(XS) - n \log n$
 \equiv Tanabe–Todd–Ye potential function;

Graph theory

- $G = (V, E)$: simple undirected graph with vertex set V and edge set E ;
 $\alpha(G)$: stability number of G ;

- $\chi(G)$: chromatic number of G ;
- $\omega(G)$: clique number of G ;
- \bar{G} : complementary graph (complement) of G ;
- $\vartheta(G)$: Lovász ϑ -function of G ;
- $\Theta(G)$: Shannon capacity of G ;

Notation for asymptotic analysis Let $f(n), g(n) : \mathbf{R} \mapsto \mathbf{R}_+$. We say that

$f(n) = O(g(n))$: $f(n)/g(n)$ is bounded from above;

$f(n) = \tilde{O}(g(n))$: $f(n) = O([\log(n)]^p g(n))$ for some $p > 0$ (independent of n);

$f(n) = \Omega(g(n))$: $g(n)/f(n)$ is bounded from above;

$f(n) \sim g(n)$ $\lim_{n \rightarrow \infty} f(n)/g(n) = 1$.

1

INTRODUCTION

Preamble

This monograph deals with algorithms for a subclass of nonlinear, convex optimization problems, namely semidefinite programs, as well as selected applications of these problems. To place the topics which are dealt with in perspective, a short survey of the field of semidefinite programming¹ is presented in this chapter. Applications in combinatorial optimization and engineering are reviewed, after which interior point algorithms for this problem class are surveyed.

1.1 PROBLEM STATEMENT

Semidefinite programming (SDP) is a relatively new field of mathematical programming, and most papers on SDP were written in the 1990's, although its roots can be traced back a few decades further (see *e.g.* Bellman and Fan [19]). A paper on semidefinite programming from 1981 is descriptively named *Linear Programming with Matrix Variables* (Craven and Mond [37]), and this apt title may be the best way to introduce the problem.

The goal is to minimize the inner product

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

¹Some authors prefer the more descriptive term 'optimization' to the historically entrenched 'programming'.

of two $n \times n$ symmetric matrices, namely a constant matrix C and a variable matrix X , subject to a set of constraints, where "Tr" denotes the trace (sum of diagonal elements) of a matrix.² The first of the constraints are linear:

$$\mathbf{Tr}(A_i X) = b_i, \quad i = 1, \dots, m,$$

where the A_i 's are given symmetric 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 symmetric positive semidefinite³, denoted by $X \succeq 0$.⁴

The convexity follows from the convexity of the cone of positive semidefinite matrices. (We recommend that the reader briefly review the properties of positive semidefinite matrices in Appendix A.)

The problem under consideration is therefore

$$(P) : p^* := \inf_X \{ \mathbf{Tr}(CX) : \mathbf{Tr}(A_i X) = b_i \ (i = 1, \dots, m), X \succeq 0 \},$$

which has an associated Lagrangian dual problem:

$$(D) : d^* := \sup_{y, S} \left\{ b^T y : \sum_{i=1}^m y_i A_i + S = C, S \succeq 0, y \in \mathbf{R}^m \right\}.$$

The duality theory of SDP is weaker than that of LP. One still has the familiar *weak duality* property: Feasible X, y, S satisfy

$$\mathbf{Tr}(CX) - b^T y = \mathbf{Tr} \left(\left(S + \sum_{i=1}^m y_i A_i \right) X \right) - \sum_{i=1}^m y_i \mathbf{Tr}(A_i X) = \mathbf{Tr}(SX) \geq 0,$$

where the inequality follows from $X \succeq 0$ and $S \succeq 0$ (see Theorem A.2 in Appendix A). In other words, the *duality gap* is nonnegative for feasible solutions.

Solutions (X, y, S) with zero duality gap

$$\mathbf{Tr}(CX) - b^T y = \mathbf{Tr}(SX) = 0$$

are optimal. For LP, if either the primal or the dual problem has an optimal solution, then both have optimal solutions, and the duality gap at optimality is zero. This is the

²This 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 $\mathbf{vec}(X)$ and $\mathbf{vec}(C)$, then $\mathbf{vec}(C)^T \mathbf{vec}(X) = \mathbf{Tr}(CX)$. The inner product induces the so-called *Frobenius norm*:

$$\|A\|^2 := \langle A, A \rangle = \mathbf{Tr}(AA^T) = \sum_{i,j} A_{ij}^2.$$

See Appendix A for more details.

³By definition, a symmetric matrix X is positive semidefinite if $z^T X z \geq 0 \ \forall z \in \mathbf{R}^n$, or equivalently, if all eigenvalues of X are nonnegative.

⁴The symbol ' \succeq ' denotes the so-called Löwner partial order on the symmetric matrices: $A \succeq B$ means $A - B$ is positive semidefinite; see also Appendix A.

strong duality property. The SDP case is more subtle: One problem may be solvable and its dual infeasible, or the duality gap may be positive at optimality, *etc.* The existence of primal and dual optimal solutions is guaranteed if both (P) and (D) allow positive definite solutions, *i.e.* feasible $X \succ 0$ and $S \succ 0$. This is called the *Slater constraint qualification* (or Slater regularity condition). These duality issues will be discussed in detail in Chapter 2.

1.2 THE IMPORTANCE OF SEMIDEFINITE PROGRAMMING

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

- SDP contains important classes of problems as special cases, such as linear and quadratic programming (LP and QP);
- important applications exist in combinatorial optimization, approximation theory, system and control theory, and mechanical and electrical engineering;
- Loosely speaking, SDP problems can be solved to ϵ -optimality in polynomial time by *interior point algorithms* (see Section 1.9 for a more precise discussion of the computational complexity of SDP); interior point algorithms for SDP have been studied intensively in the 1990's, explaining the resurgence in research interest in SDP.

Each of these considerations will be discussed briefly in the remainder of this chapter.

1.3 SPECIAL CASES OF SEMIDEFINITE PROGRAMMING

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 have an LP problem. Optimization problems with convex quadratic constraints are likewise special cases of SDP.⁵ This follows from the well-known *Schur complement* theorem (Theorem A.9 in Appendix A). Thus we can represent the quadratic constraint

$$(Ax + b)^T(Ax + b) - (c^T x + d) \leq 0, \quad x \in \mathbf{R}^n,$$

by the semidefinite constraint

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

In the same way, we can represent the *second order cone* (or ‘ice cream cone’):

$$\left\{ (t, x) \mid t \geq \sqrt{\sum_{i=1}^n x_i^2} \right\},$$

⁵This includes the well-known convex quadratic programming (QP) problem.

by requiring that a suitable arrow matrix be positive semidefinite:

$$\begin{bmatrix} tI & x \\ x^T & t \end{bmatrix} \succeq 0.$$

Another nonlinear example is

$$\min_x \left\{ \frac{(c^T x)^2}{d^T x} \mid Ax \geq b \right\},$$

where it is known that $d^T x > 0$ if $Ax \geq b$. An equivalent SDP problem is:⁶

$$\min_{t,x} \left\{ t \mid \begin{bmatrix} t & c^T x & 0 \\ c^T x & d^T x & 0 \\ 0 & 0 & \text{Diag}(Ax - b) \end{bmatrix} \succeq 0 \right\}.$$

Several problems involving matrix norm or eigenvalue minimization may be stated as SDP's. A list of such problems may be found in Vandenberghe and Boyd [181]. 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

$$\min_t \{t \mid tI - A \succeq 0, t \in \mathbf{R}\}.$$

An SDP algorithm for this problem is described by Jansen *et al.* [88, 90].

1.4 APPLICATIONS IN COMBINATORIAL OPTIMIZATION

In this section we give a short review of some of the most important and successful applications of SDP in combinatorial optimization.

The Lovász ϑ -function

The most celebrated example of application of SDP to combinatorial optimization is probably the Lovász ϑ -function [115].

The Lovász ϑ -function maps a graph $G = (V, E)$ to \mathbf{R}_+ , in such a way that

$$\omega(G) \leq \vartheta(\bar{G}) \leq \chi(G), \quad (1.1)$$

where $\omega(G)$ denotes the clique number⁷ of G , $\chi(G)$ the chromatic number⁸ of G , and \bar{G} the complement of G .⁹

⁶We use the following notation: for a matrix X , $\text{diag}(X)$ is the vector obtained by extracting the diagonal of X ; for a vector x , $\text{Diag}(x)$ is the diagonal matrix with the coordinates of x as diagonal elements.

⁷A maximum clique (or completely connected subgraph) is a subset $C \subset V$ with $\forall i, j \in C (i \neq j) : (i, j) \in E$, such that $|C|$ is as large as possible. The cardinality $|C|$ is called the clique number.

⁸The chromatic number is the number of colours needed to colour all vertices so that no two adjacent vertices share the same colour.

⁹The complementary graph (or complement) of $G = (V, E)$ is the graph $\bar{G} = (V, \bar{E})$ such that for each pair of vertices $i \neq j$ one has $(i, j) \in \bar{E}$ if and only if $(i, j) \notin E$.

The ϑ -function value is given as the optimal value of the following SDP problem:

$$\vartheta(\bar{G}) := \max_X \text{Tr}(ee^T X) = e^T X e,$$

where $e \in \mathbf{R}^{|V|}$ denotes the vector of all-ones, subject to

$$\begin{aligned} x_{ij} &= 0, (i, j) \notin E (i \neq j) \\ \text{Tr}(X) &= 1 \\ X &\succeq 0. \end{aligned}$$

The relation (1.1) is aptly known as the ‘sandwich theorem’ (a name coined by Knuth [103]). The sandwich theorem implies that $\vartheta(\bar{G})$ can be seen as a polynomial time approximation to both $\omega(G)$ and $\chi(G)$, and that the approximation cannot be off by more than a factor $|V|$. This may seem like a rather weak approximation guarantee, but some recent in-approximability results suggest that neither $\omega(G)$ nor $\chi(G)$ can be approximated within a factor $|V|^{1-\epsilon}$ for any $\epsilon > 0$ in polynomial time (see Håstad [81] and Feige and Kilian [57]).

We will review some of the properties of the ϑ -function in Chapter 10, and will give a proof of the sandwich theorem there; moreover, we will look at some alternative definitions of this function. We will also review how the ϑ -function may be used to estimate the Shannon capacity¹⁰ of a graph.

The MAX—CUT problem and extensions

Another celebrated application of SDP to combinatorial optimization is the MAX-CUT problem. Consider a clique $G = (V, E)$ where each edge (i, j) has a given weight $w_{ij} \geq 0$ ($i \neq j$).

The goal is to colour all the vertices of G using two colours (say red and blue), in such a way that the total weight of edges where the endpoints have different colours (called *non—defect edges*) is as large as possible. The non—defect edges define a ‘cut’ in the graph — if one ‘cuts’ all the non—defect edges, then the blue and red vertices are separated. The total weight of the non—defect edges is therefore also called the weight of the cut.

Goemans and Williamson [66] derived a randomized 0.878-approximation algorithm¹¹ for this problem using SDP. The first step was to write the MAX-CUT problem as a Boolean quadratic optimization problem, *i.e.* an optimization problem with quadratic objective function and Boolean variables. For each vertex in V we introduce

¹⁰The Shannon capacity is a graph theoretical property that arises naturally in applications in coding theory; see Chapter 10 for details.

¹¹Consider any class of N P -complete maximization problems. An α -approximation algorithm ($0 < \alpha \leq 1$) for this problem class is then defined as follows. For any problem instance from this class, the algorithm terminates in time bounded by a polynomial in the size of the instance, and produces a feasible solution with objective value at least α times the optimal value for the instance.