

A GUIDE TO CONIC OPTIMISATION AND ITS APPLICATIONS

ADAM N. LETCHFORD^{1,*} AND ANDREW J. PARKES²

Abstract. Most OR academics and practitioners are familiar with linear programming (LP) and its applications. Many are however unaware of *conic optimisation*, which is a powerful generalisation of LP, with a prodigious array of important real-life applications. In this invited paper, we give a gentle introduction to conic optimisation, followed by a survey of applications in OR and related areas. Along the way, we try to help the reader develop insight into the strengths and limitations of conic optimisation as a tool for solving real-life problems.

Mathematics Subject Classification. 90-01, 90C22, 90C90

Received 14 September 2017. Accepted 27 April 2018.

1. INTRODUCTION

Most OR students, academics and practitioners are familiar with linear programming (LP). For many problems arising in industry and elsewhere, LP is an attractive option, due to its simplicity, the ease of doing sensitivity analysis, the existence of effective algorithms, and, perhaps most importantly, the availability of good software. With modern software packages such as CPLEX, Gurobi or Xpress, it is now possible to routinely solve LP instances with thousands of variables and/or constraints to proven optimality. These software packages can also cope with integer-constrained variables.

Of course, in real-world applications, one often encounters problems that have *non-linear* aspects. Sometimes, one can construct good linear approximations to such problems, perhaps with the help of binary variables, and thereby make them amenable to solution with the above-mentioned packages. In other cases, however, the nonlinearity is significant and must be faced head-on. For such cases, various non-linear programming (NLP) algorithms have been developed (see, *e.g.*, [13, 14]).

The purpose of this paper is to introduce a very important special case of NLP called conic optimisation (CO). Despite having a very special structure, CO is remarkably powerful. It has a prodigious array of important real-life applications, not only in OR, but also in related areas, such as statistics, computer science, engineering and finance. Moreover, CO inherits some of the nice features enjoyed by LP, such as the existence of efficient (polynomial-time) algorithms, a well-developed duality theory, and the availability of good software.

Apart from LP, the most important special cases of CO are second order cone programming (SOCP) and semidefinite programming (SDP). There already exist several good surveys on both SOCP (*e.g.*, [2, 53, 62]) and SDP (*e.g.*, [44, 93, 94, 98]). There also exists an excellent monograph on CO in general (Nesterov and

Keywords and phrases: Conic optimisation, second-order cone programming, semidefinite programming.

¹ Department of Management Science, Lancaster University, Lancaster, UK.

² School of Computer Science, University of Nottingham, Nottingham, UK.

* Corresponding author: a.n.letchford@lancaster.ac.uk

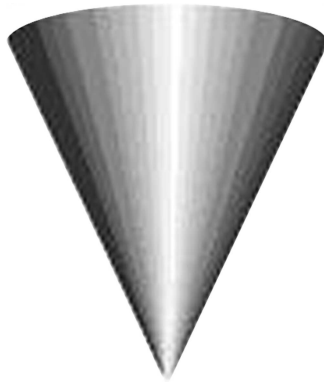


FIGURE 1. An idealised ice-cream cone.

Nemirovsky [73]). It is fair to say, however, that most of these works assume that the reader is a nonlinear programming expert. This can make them rather inaccessible to the OR generalist.

In this work, we assume only minimal knowledge of LP, geometry and linear algebra. Moreover, we place an emphasis on applications in OR and related areas, covering, *e.g.*, inventory control, facility location, portfolio optimisation, problems involving binary variables, and various methods for optimising under uncertainty, such as mean-variance, chance-constrained and robust optimisation. Along the way, we try to help the reader develop insight into the strengths and limitations of CO as a tool for modelling and solving real-life problems.

The paper is structured as follows. In Section 2, we define CO and present some basic theory. In Section 3, we cover algorithms and complexity. In Section 4, we review the main applications of SOCP in OR and related fields. In Section 5, we do the same for SDP. In Section 6, we list some of the available software packages, and in Section 7, we make some concluding remarks.

Throughout the paper, \mathbb{R}_+ , \mathbb{R}_+^n and \mathcal{S}^n denote the non-negative reals, the real vectors with n non-negative components, and the real symmetric matrices of order n , respectively. Moreover, given a vector $x \in \mathbb{R}^n$, we let $\|x\|_1$ and $\|x\|_2$ denote $\sum_{i=1}^n |x_i|$ and $\sqrt{\sum_{i=1}^n x_i^2}$, respectively.

2. DEFINITIONS AND BASIC THEORY

In this section, we define CO formally and present the minimal amount of theory and notation needed to understand how it works. We start by defining *cones* in Section 2.1. We define CO itself in Section 2.2. We then present elementary duality theory in Section 2.3.

2.1. Cones

If a non-mathematician hears the word *cone*, the thing most likely to come to mind is either an ice-cream cone or a traffic cone. An idealised version of an ice-cream cone is displayed in Figure 1. The idealised version is supposed to extend upward forever. One can check that it corresponds to the following set:

$$\left\{ x \in \mathbb{R}^3 : x_3 \geq \sqrt{x_1^2 + x_2^2} \right\}.$$

The ancient Greek astronomer Apollonius of Perga (c. 262 BC–c. 190 BC) discovered that, if one “slices” the idealised ice-cream cone at various angles, one can obtain interesting convex shapes; see Figure 2. This already gives a hint as to why cones could be of relevance to optimisation: whereas LP forces us to optimise over polyhedra, we can optimise over various non-polyhedral convex sets by “slicing” cones in various ways. This enables us to model and solve a variety of important nonlinear problems, as we will see in Sections 4 and 5.

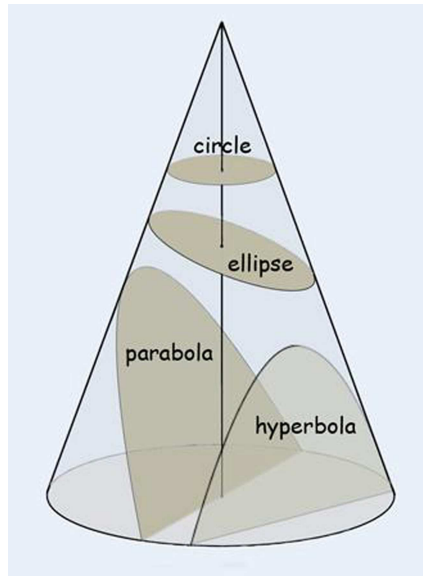


FIGURE 2. The conic sections (taken from Seligman [88]).

Mathematically speaking, a cone is a set of points (in some underlying space, such as Euclidean space) with the following property: if a point x belongs to the cone, then so does the point λx for any $\lambda \in \mathbb{R}_+$. Note that cones can be extremely complicated. Indeed, consider an *arbitrary* set $S \subset \mathbb{R}^n$. The set

$$C(S) = \left\{ (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+ : x = \lambda x' \text{ for some } x' \in S \right\}$$

is a cone, and it is “just as complicated” as S itself. Indeed, if we intersect $C(S)$ with the hyperplane defined by the equation $\lambda = 1$, the resulting “slice” is S (or, more precisely, an embedding of S into a space of higher dimension).

If one wishes to say anything useful about cones, then, one must restrict attention to cones with a special structure. In the optimisation context, one is typically interested in so-called *proper* cones. A cone C is called proper if it is:

- convex: if $x^1, x^2 \in C$, then $\mu x^1 + (1 - \mu)x^2 \in C$ for all $\mu \in [0, 1]$;
- closed: C contains all of its limit points;
- full-dimensional: there is no hyperplane containing C ;
- pointed: if $x \in C$, then $-x \notin C$.

One can check that the idealised ice-cream cone is proper. The following five proper cones arise frequently in various contexts:

- the *non-negative* cone, which is simply \mathbb{R}_+ ;
- the *second-order cone* of order n , which is

$$\{(x, t) \in \mathbb{R}^n \times \mathbb{R}_+ : t \geq \|x\|_2\};$$

- the *positive semidefinite* (psd) cone of order n , which is

$$\left\{ X \in \mathcal{S}^n : v^T X v \geq 0 \ (\forall v \in \mathbb{R}^n) \right\}; \tag{2.1}$$

the *copositive* cone of order n (first defined by Motzkin [71]), which is

$$\left\{ X \in \mathcal{S}^n : v^T X v \geq 0 \ (\forall v \in \mathbb{R}_+^n) \right\};$$

- the *completely positive* cone of order n (first defined by Hall [40]), which is

$$\left\{ X \in \mathcal{S}^n : X = A^T A \text{ for some real non-negative matrix } A \right\}.$$

(There are several other cones of interest to optimisation, such as the correlation cone [27], the p th-order cone [33], the exponential cone [32] and the relative entropy cone [21], but we do not give details, for the sake of brevity.)

The second-order cone is a natural generalisation of the ice-cream cone to higher dimensions. It is sometimes called the *Lorentz* cone, after the Dutch physicist Hendrik Lorentz (1853–1928). (Indeed, for those familiar with special relativity, the second-order cone with $n = 3$ is the forward light cone of the origin, where x represents space and t represents time.)

The psd cone of order n is usually denoted by \mathcal{S}_+^n . It can be defined in many different ways. It is known (see, e.g., Horn [48]) that a matrix $X \in \mathcal{S}^n$ belongs to \mathcal{S}_+^n if and only if any of the following (equivalent) conditions hold:

- the quadratic function $f(v) = v^T X v$ is non-negative for all $v \in \mathbb{R}^n$ (this is just definition (2.1));
- the same function $f(v)$ is convex;
- the region $\{v \in \mathbb{R}^n : v^T X v \leq 1\}$ is an ellipsoid;
- all eigenvalues of X are non-negative;
- all principal submatrices of X have non-negative determinants;
- there exists a lower-triangular matrix $A \in \mathbb{R}^{n \times n}$ such that $X = A^T A$ (Cholesky factorisation);
- there exist vectors $u^1, \dots, u^n \in \mathbb{R}^n$ such that $X_{ij} = u^i \cdot u^j$ for all i, j (Gram representation);
- X can be written as a non-negative linear combination of symmetric rank-1 real matrices of the form vv^T .

From these definitions, one can see that, for a given value of n , the completely positive cone is contained in \mathcal{S}_+^n , which in turn is contained in the copositive cone.

To make the above ideas more concrete, we give some examples. The matrix $\begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}$ is completely positive, since it has the factorisation

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

The matrix $\begin{pmatrix} 5 & -1 \\ -1 & 1 \end{pmatrix}$ is psd, since it has the factorisation:

$$\begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}.$$

It is not however completely positive, since it contains negative entries. Finally, the matrix $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is copositive, since $v^T X v = 2v_1 v_2 \geq 0$ for all $v \in \mathbb{R}_+^2$. It is however not psd, since

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

and therefore it has -1 as an eigenvalue.

2.2. Conic optimisation

A *conic programme* or *conic optimisation problem* is a problem that can be modelled as the problem of optimising a linear function over the intersection of a hyperplane and a proper cone. That is, a CO problem can be written in the form

$$\sup \{c^T x : Ax = b, x \in C\}, \quad (2.2)$$

where $c \in \mathbb{Z}^n$, $b \in \mathbb{Z}^m$, $A \in \mathbb{Z}^{m \times n}$ and $C \subset \mathbb{R}^n$ is a proper cone. (We have to use supremum rather than maximum here, for technical reasons. See the last example in this subsection.)

Actually, as it stands, the definition (2.2) is too general, since it includes \mathcal{NP} -hard problems as special cases (see the next section). Nesterov and Nemirovsky [73] showed that particularly effective solution methods could be devised for CO when the proper cones in question are *symmetric*. For brevity, we do not define symmetric cones here. Instead, we just remark that most symmetric cones of interest can be constructed using three basic building blocks; namely, the non-negative, second order and psd cones mentioned in the previous subsection. More specifically:

- The product of n non-negatives cones is \mathbb{R}_+^n . This is a symmetric cone, and it is the cone used in LP.
- The product of a finite number of second-order cones is symmetric. Cones of this type are used in second-order cone programming or SOCP (sometimes also called conic quadratic programming or CQP).
- The product of a finite number of psd cones is also symmetric. Cones of this type are used in semidefinite programming (SDP).

It might seem at first that restricting attention to products of cones is a big limitation, since it implies that each cone must involve a different set of variables. However, this apparent limitation can be overcome by “splitting” variables into two or more copies. For example, if we wish to impose the second-order conic constraints $t \geq \|(x/y)\|_2$ and $s \geq \|(x/z)\|_2$ simultaneously, where the vector x is involved in both constraints, we replace the vector x with two vectors, say x' and x'' , add the constraints $t \geq \|(x'/y)\|_2$ and $s \geq \|(x''/z)\|_2$, and add the (linear) constraint $x' = x''$.

It turns out that SDP generalises SOCP, which in turn generalises LP. Indeed, the second-order conic constraint $t \geq \|x\|_2$ is equivalent to

$$\begin{pmatrix} t & x^T \\ x & tI_n \end{pmatrix} \in \mathcal{S}_+^{n+1},$$

where I_n is the identity matrix of order n ; and the non-negativity constraint $t \geq 0$ is implied by $t \geq \|x\|_2$ together with $x_i = 0$ for $i = 1, \dots, n$. (Geometrically speaking, the non-negative cone is a “slice” of the second-order cone, which is in turn a “slice” of the psd cone.)

In the case of LP and SOCP, it is usual to view the variables as being arranged in a vector, which is usually denoted by x . In the case of SDP, however, it is usual to view them as being arranged in a square symmetric matrix, which is usually denoted by X . Accordingly, people sometimes refer to *vector variables* and *matrix variables*. When X is a matrix variable, the cone C is to be thought of as a subset of \mathcal{S}^n rather than \mathbb{R}^n .

To help the reader, we now give a couple of examples.

Example 2.1. Consider the following SOCP:

$$\sup \left\{ x_1 : x_1 - x_2 = 0, x_3 + x_4 = 2, x_3 \geq \sqrt{x_1^2 + x_2^2}, x_4 \geq 0 \right\}.$$

The largest value that x_3 can take is 2, which means that $x_1^2 + x_2^2$ cannot exceed 2. The optimal solution is therefore $x_1^* = x_2^* = \sqrt{2}$, $x_3^* = 2$ and $x_4^* = 0$, with a profit of $\sqrt{2}$.

Example 2.2. Consider the following SDP:

$$\sup \left\{ X_{12} + X_{21} : X_{11} = 1, X_{22} = 3, X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \in \mathcal{S}_+^2 \right\}.$$

Note that the determinant of X is $X_{11}X_{22} - X_{12}X_{21}$, and this must be non-negative. Since $X_{11}X_{22}$ must equal 3, and X_{12} must equal X_{21} , the profit is maximised by setting X_{12} and X_{21} to $\sqrt{3}$. This yields a profit of $2\sqrt{3}$.

The above examples show that an SOCP or SDP can have a unique optimal solution in which one or more variables take *irrational* values. The following example shows that it is also possible for the supremum not to be attainable in an SOCP:

Example 2.3. Consider the following SOCP:

$$\sup \left\{ x_1 - x_3 : x_2 = 1, x_3 \geq \sqrt{x_1^2 + x_2^2} \right\}.$$

This is equivalent to:

$$\sup \left\{ x_1 - \sqrt{x_1^2 + 1} : x_1 \in \mathbb{R} \right\}.$$

We can bring the profit arbitrarily close to 0 (by making x_1 arbitrarily large), but we cannot actually reach 0.

In a similar way, it can be shown that the supremum may not be attainable in an SDP.

2.3. Duality

There is an elegant duality theory for CO, which can be viewed as a generalisation of LP duality. First, assume that we are working with a vector variable $x \in \mathbb{R}^n$. Associated with any cone $C \subset \mathbb{R}^n$ we can define the *dual* (a.k.a. *polar*) cone

$$C^* := \{x \in \mathbb{R}^n : y^T x \geq 0 (\forall y \in C)\}.$$

If we are dealing with a matrix variable $X \in \mathcal{S}^n$ instead, and a cone $C \subset \mathcal{S}^n$, then the definition of the polar cone must be modified slightly to:

$$C^* := \{X \in \mathcal{S}^n : X \bullet Y \geq 0 (\forall Y \in C)\},$$

where $X \bullet Y$ denotes the matrix inner product $\sum_{i=1}^n \sum_{j=1}^n X_{ij}Y_{ij}$, which is also equal to $\text{tr}(Y^T X)$.

It is easy to show that (a) the dual of a proper cone is proper, (b) if a cone is proper, then it is the dual of its dual, (c) the dual of the completely positive cone is the copositive cone (and vice-versa), and (d) the non-negative, second-order and psd cones are *self-dual*. (This is a property shared by all symmetric cones.)

The following fact can be shown, *e.g.*, by an application of Lagrangian relaxation:

Theorem 2.4 (Weak Duality for CO). *For any proper cone $C \subset \mathbb{R}^n$, and any $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$, we have:*

$$\sup \{c^T x : Ax = b, x \in C\} \leq \inf \{b^T y : A^T y - c \in C^*, y \in \mathbb{R}^m\}.$$

Specialised to the case of LP, Theorem 2.4 reduces to the well-known fact that

$$\max \{c^T x : Ax = b, x \in \mathbb{R}_+^n\} \leq \min \{b^T y : A^T y \geq c, y \in \mathbb{R}^m\}.$$

In that case, we have strong duality, *i.e.*, the inequality can be changed to an equation. For CO in general, however, strong duality is not guaranteed. Fortunately, it is guaranteed to hold under certain conditions that commonly occur in practice. For example, it holds if the *Slater condition* is satisfied, *i.e.*, if there is a feasible x that lies in the interior of C (see, *e.g.*, [14, 73]).

Even when duality is strong, however, some care must be taken when interpreting the dual. For example, although the optimal dual solution y^* provides meaningful dual prices, the components of the vector $A^T y^* - c$ are not always meaningful “reduced costs” (see, *e.g.*, Helmborg [43]). Moreover, concepts such as degeneracy, dual degeneracy and complementary slackness must be handled carefully (*e.g.*, Alizadeh *et al.* [3]).

3. ALGORITHMS AND COMPLEXITY

Now we turn our attention to algorithms for solving CO problems. To begin, it is helpful to recall the following facts about LP:

1. The *simplex method* (Dantzig [22, 23]) is typically very fast in practice, but takes exponential time in the worst case.
2. The *ellipsoid method* (Khachiyan [51]) runs in polynomial time in the worst case, but is too slow to be of practical use.
3. There exist various *interior-point methods* (IPMs) that run in polynomial time in the worst case, and tend to be reasonably fast in practice (see, *e.g.*, Gondzio [37]).

The situation for CO is a bit more complicated:

1. There is no known effective analogue of the simplex method for general CO, nor indeed for SOCP or SDP in particular.
2. The ellipsoid method can solve CO in polynomial time (Grötschel *et al.* [39]), but only to *fixed precision*, and only when two technical conditions are met: (i) it must be possible to test in polynomial time whether a given rational point lies in the given cone, and (ii) the feasible region must be explicitly bounded (*e.g.*, by the addition of a constraint which forces the solution to lie inside a ball of known centre and radius). Unfortunately, as in the case of LP, this result is of little use in practice.
3. There exist IPMs that can solve SOCPs and SDPs in polynomial time, but again, only to fixed precision and only when the feasible region is bounded (Nesterov and Nemirovsky [73]). The best of these IPMs are now fast enough to be of practical use.

To see why COs can only be solved to fixed precision, recall from Section 2.2 that even SOCPs and SDPs can have irrational solutions. Such solutions cannot be represented exactly with a finite number of bits. (For a discussion on how one might represent solutions exactly as *algebraic numbers*, see Nie *et al.* [75].) One consequence of this limited precision is that even testing feasibility of an SOCP or SDP is a non-trivial problem, if one wants a precise answer. In fact, it is not known whether it can be done in polynomial time; see Ramana [84] for a discussion.

The other technical condition mentioned above is that the feasible region must be explicitly bounded. To see why, consider the following constraints:

$$x_3 - x_2 = 1, x_4 - 2x_2 = 1, x_3 \geq \sqrt{x_1^2 + x_2^2}.$$

These can be shown to imply $x_4 \geq x_1^2$. Chaining such constraints together, we can get $x_7 \geq x_4^2$, $x_{10} \geq x_7^2$, and so on. At the end, we get $x_{3p+1} \geq x_1^{(2^p)}$. If we add the constraint $x_1 = 2$, we find that $x_{3p+1} \geq 2^{(2^p)}$. So there

exist SOCPs such that it takes an exponential number of bits to represent any feasible solution. An analogous example for SDPs is given in Alizadeh [1]. Requiring the feasible region to lie inside a ball eliminates such “pathological” SOCP and SDP instances.

As for the question of testing membership of a cone, testing whether a given $x^* \in \mathbb{Q}$ lies in the non-negative cone is trivial, and so is testing whether a given rational point (x^*, t^*) lies in the second-order cone. Less obviously, one can also check whether a given rational matrix $X^* \in \mathcal{S}^n$ belongs to \mathcal{S}_+^n in polynomial time, *via* a modified form of Gaussian elimination [39]. (In practice, one can just compute the minimum eigenvalue of X^* , to some desired precision, and check whether it is non-negative.)

On the other hand, testing whether a rational matrix is completely positive is \mathcal{NP} -hard [72]. (By duality, testing copositivity is $\text{co-}\mathcal{NP}$ -hard.) In fact, Burer [16] showed that any mixed 0–1 linear or quadratic program can be transformed into a completely positive program, *i.e.*, a conic optimisation problem over the completely positive cone. As a result, completely positive programming is \mathcal{NP} -hard in the strong sense. (By duality, copositive programming is $\text{co-}\mathcal{NP}$ -hard in the strong sense.) For a survey of results on completely positive and copositive programming, see Burer [17].

At the end of the previous subsection, we mentioned that dual information for SOCP or SDP, such as reduced costs and dual prices, must be interpreted with care. The lack of a simplex method leads to some other issues. For example:

- It can be hard to exploit sparsity in the constraint matrix (*e.g.*, Benson *et al.* [10]).
- It is not trivial to re-optimize SOCPs and SDPs efficiently after, *e.g.*, adding a constraint or variable (*e.g.*, [69, 92]). In other words, it is difficult to do “warm starts”.

Work on these issues is ongoing within the CO community.

We now wish to mention two other remarkable papers. Ben-Tal and Nemirovski [12] proved that one can “simulate” an SOCP with n variables to arbitrary precision ϵ using an LP with $\mathcal{O}(n \log(1/\epsilon))$ variables. This means that one can solve an SOCP approximately by solving a single LP of reasonable size. Conversely, Braun *et al.* [15] proved that one cannot simulate SDPs by LPs or SOCPs in an analogous way, thereby proving that SDP is, in some sense, fundamentally more powerful than LP or SOCP.

Finally, we mention that there exist several other algorithms for solving SDPs, such as the spectral bundle method (Helmberg and Rendl [46]), the boundary point method (Povh *et al.* [82]), augmented Lagrangian methods [67, 96], cutting-plane methods (Krishnan and Mitchell [52]), and a method that works directly with the Cholesky factorisation (Burer and Monteiro [19]).

4. APPLICATIONS OF SOCP

In this section, we give some examples of the kinds of problems that can be tackled using SOCP. For additional examples, see, *e.g.*, [2, 53, 62, 73].

4.1. Problems involving convex quadratic functions

SOCP includes as a special case all convex nonlinear programs involving quadratic functions. For brevity, we give just two examples, one from finance and one from statistics.

In the famous portfolio selection model of Markowitz [68], there are n stocks, for which we have a vector $r \in \mathbb{R}^n$ of *expected returns* and a psd matrix $Q \in \mathbb{R}^{n \times n}$ of *covariances*. We have a vector $x \in \mathbb{R}_+^n$ of decision variables, where x_i represents the proportion invested in stock i . The expected return of a portfolio x is then $r^T x$ and the variance (a measure of risk) is $x^T Q x$. Minimising risk subject to a lower bound L on the expected return is then equivalent to the following convex quadratic program:

$$\min \left\{ x^T Q x : r^T x \geq L, \|x\|_1 = 1, x \in \mathbb{R}_+^n \right\}.$$

If we compute the Cholesky factorisation $Q = A^T A$, then the objective function is equivalent to $\|Ax\|_2^2$. So we can reformulate the problem as:

$$\min \left\{ t : t \geq \|y\|_2, y = Ax, r^T x \geq L, \|x\|_1 = 1, x \in \mathbb{R}_+^n, y \in \mathbb{R}^n, t \in \mathbb{R}_+ \right\}.$$

This is an SOCP. Similarly, the problem of maximising expected return subject to an upper bound U on the variance can be modelled as the SOCP:

$$\max \left\{ r^T x : t \geq \|y\|_2, y = Ax, t \leq \sqrt{U}, \|x\|_1 = 1, x \in \mathbb{R}_+^n, y \in \mathbb{R}^n, t \in \mathbb{R}_+ \right\}.$$

In statistical estimation, one often encounters problems of the form:

$$\min \left\{ \sum_{j=1}^m w_j \|y^j - x\|_2^2 : x \in S \right\},$$

where

- y^1, \dots, y^m are observations of a random n -vector with unknown mean;
- x is an estimate of the mean (to be determined);
- $S \subset \mathbb{R}^n$ is a convex set representing some prior information on x ;
- w_j is the weight (importance) given to the j th observation;
- the objective is to minimise a weighted sum of squared residuals.

By introducing a new variable $z \in \mathbb{R}_+$ and new vectors of variables $u \in \mathbb{R}_+^m$ and $t^1, \dots, t^m \in \mathbb{R}^m$, we can reformulate the estimation problem as:

$$\begin{aligned} \min \quad & z \\ \text{s.t.} \quad & z \geq \|u\|_2 \\ & u_j \geq \|t^j\|_2 \quad (j = 1, \dots, m) \\ & t^j = \sqrt{w_j}(y^j - x) \quad (j = 1, \dots, m) \\ & x \in S \\ & (z, u, t) \in \mathbb{R}_+^{1+m+m \times n}. \end{aligned}$$

If S is a polyhedron, the reformulated problem is an SOCP. If S is defined by a mixture of linear and convex quadratic constraints, the problem can be easily converted into an SOCP.

4.2. Problems involving hyperbolic functions

Now recall from Section 2.1 that the *hyperbola* is a cross-section of the ice-cream cone. More precisely, the convex set

$$\{(x, t) \in \mathbb{R}_+^2 \times \mathbb{R} : t \geq \|x\|_2, x_2 = 1\}$$

is easily shown to be an affine image of the convex set

$$\{(x, y) \in \mathbb{R}_+^2 : y \geq 1/x\},$$

the boundary of which is a hyperbola. For this reason, one can easily force the feasible region to be a hyperbola *via* SOCP.

Note that the function $y \geq 1/x$ is equivalent to $xy \geq 1$. More generally, a *hyperbolic function* is any function of the form $t_1 t_2 \geq \|Ax\|^2$, where $t_1, t_2 \in \mathbb{R}_+$ and $x \in \mathbb{R}^n$ are variables, and A is a real matrix with n columns. It is shown in [62] that any convex nonlinear program involving a combination of linear, convex quadratic and hyperbolic functions can be converted to an SOCP. (The feasible region of a constraint involving a hyperbolic function is sometimes called a *hyperboloid*.)

Here is a simple example of an OR problem that involves hyperbolic functions. The classic inventory control model of Harris [41] is

$$\min \{hx/2 + cd/x : x \in \mathbb{R}_+\},$$

where x is the order quantity (to be determined), h is the annual cost of holding one unit in stock, c is the charge for a delivery, and d is the annual demand. Harris solved this problem by calculus, and the optimal value of x is the well-known economic order quantity (EOQ). Ziegler [99] considered the following multi-item extension of this model:

$$\begin{aligned} \min \quad & \sum_{i=1}^n (h_i x_i / 2 + c_i d_i / x_i) \\ \text{s.t.} \quad & \sum_{i=1}^n b_i x_i \leq b_0 \\ & \ell_i \leq x_i \leq u_i \quad (i = 1, \dots, n). \end{aligned}$$

Here, for a given product i , x_i represent the order quantity, h_i is the annual holding cost, c_i is the delivery charge, d_i is the annual demand, b_i is the space occupied by one unit, and ℓ_i and u_i are lower and upper bounds on the order quantity. The constant b_0 represents the storage space available.

Kuo and Mittelman [53] showed that this problem can be transformed into an SOCP as follows. Define new variables s_i and t_i satisfying:

$$\begin{aligned} x_i &= (s_i - t_i) / 2 \\ 1/x_i &= (s_i + t_i) / 2. \end{aligned}$$

Then model the problem as:

$$\begin{aligned} \min \quad & \sum_{i=1}^n ((h_i/2 + c_i d_i) s_i + (h_i/2 - c_i d_i) t_i) \\ \text{s.t.} \quad & \sum_{i=1}^n b_i (s_i - t_i) \leq 2b_0 \\ & 2\ell_i \leq s_i - t_i \leq 2u_i \quad (i = 1, \dots, n) \\ & s_i \geq \sqrt{t_i^2 + 2} \quad (i = 1, \dots, n) \\ & s, t \in \mathbb{R}^n. \end{aligned}$$

4.3. Problems involving norms

SOCP can also be used to model problems involving Euclidean or other norms. For brevity, we give three examples.

The first example concerns facility location. Suppose we have n facilities in the plane and wish to locate a new facility so as to minimise the sum of the weighted distances from the existing facilities. This can be modelled as:

$$\begin{aligned} \min \quad & \sum_{i=1}^n w_i \sqrt{(x_i - \tilde{x})^2 + (y_i - \tilde{y})^2} \\ \text{s.t.} \quad & (\tilde{x}, \tilde{y}) \in \mathbb{R}^2, \end{aligned}$$

where (x_i, y_i) are the co-ordinates of the i th existing facility, and (\tilde{x}, \tilde{y}) are the co-ordinates of the new facility. (This problem is sometimes called the *Fermat-Weber* problem.) Defining new variables d_1, \dots, d_n , representing the Euclidean distance from the new facility to each of the existing facilities, the problem can be converted into the following SOCP:

$$\begin{aligned} \min \quad & \sum_{i=1}^n w_i d_i \\ \text{s.t.} \quad & d_i \geq \sqrt{u_i^2 + v_i^2} \quad (i = 1, \dots, n) \\ & u_i = x_i - \tilde{x} \quad (i = 1, \dots, n) \\ & v_i = y_i - \tilde{y} \quad (i = 1, \dots, n) \\ & (\tilde{x}, \tilde{y}) \in \mathbb{R}^2 \\ & u, v, d \in \mathbb{R}^n. \end{aligned}$$

Moreover, SOCP can also easily handle extensions of this problem in which there are linear, convex quadratic and/or hyperbolic constraints on the co-ordinates \tilde{x} and \tilde{y} and/or the distances d_i .

Our second example is a modified version of the Markowitz portfolio selection model (see Sect. 4.1). Suppose we wish to maximise expected return subject to a *chance constraint*, which states that the probability of the return not exceeding some quantity $\alpha \in \mathbb{R}$ must be less than some small quantity $\beta > 0$. (For example, we may require that there is only a 0.1% chance of losing one million euros or more.) The chance constraint can be written as:

$$r^T x + \Phi^{-1}(\beta) \sqrt{x^T Q x} \geq \alpha,$$

where Φ is the Cumulative Distribution Function of the Normal distribution. Using again the Cholesky factorisation $Q = AA^T$, the constraint can be written as:

$$r^T x + \Phi^{-1}(\beta) \|Ax\|_2 \geq \alpha.$$

To handle this with SOCP, we just add new variables $y \in \mathbb{R}^n$ and $z \in \mathbb{R}$, and add the constraints:

$$\begin{aligned} y &= Ax \\ z &\geq \|y\|_2 \\ r^T x + \Phi^{-1}(\beta) z &\geq \alpha. \end{aligned}$$

(Actually, this transformation only works when $\Phi^{-1}(\beta) < 0$. Fortunately, this is always the case in practice, since β is always less than 1/2.)

Our third and final example comes from *robust optimisation*. Consider the LP

$$\min \{c^T x : Ax \leq b, x \in \mathbb{R}_+^n\},$$

and suppose that the precise values of the components of the matrix A are uncertain. Ben-Tal and Nemirovski [11] suggest writing the LP as:

$$\min \{c^T x : a_i^T x \leq b_i \ (i = 1, \dots, m), x \in \mathbb{R}_+^n\},$$

and then considering the case in which, for $i = 1, \dots, m$, the vector a_i is known to lie inside the *ellipsoid*

$$\{\hat{a}_i + Q_i u : \|u\|_2 \leq 1\},$$

where the vectors \hat{a}_i and the matrices Q_i are known. They then show that the problem of minimising the *worst-case* cost, the so-called *robust counterpart* of the LP, can be formulated as:

$$\min \{c^T x : \hat{a}_i^T x + \|Q_i x\|_2 \leq b_i \ (i = 1, \dots, m), \ x \in \mathbb{R}_+^n\}.$$

This is again easy to handle *via* SOCP.

5. APPLICATIONS OF SDP

Although SOCP is a useful modelling tool, SDP is much more powerful. In this section, we give just a few examples of the kinds of problems that can be tackled using SDP. For additional examples, see, *e.g.*, [4, 44, 73, 93, 94, 97, 98].

5.1. Problems involving special types of matrices

Covariance and *correlation* matrices play a fundamental role in statistics, probability and (as we saw in Sects. 4.1 and 4.3) finance. It is well known that a real symmetric matrix is a covariance matrix if and only if it is psd, and a correlation matrix if and only if, in addition, it has 1s on the main diagonal. This means that one can use SDP to solve various optimisation problems involving such matrices, such as:

- The *positive semidefinite matrix completion problem*: given a matrix with missing entries, check if it can be completed to a covariance (or correlation) matrix (see, *e.g.*, Johnson [50]).
- The *nearest correlation matrix problem*: given a matrix that is not a correlation matrix, find a correlation matrix that is as close as possible, where “close” is measured according to, *e.g.*, an L_1 , L_2 or L_∞ norm (see, *e.g.*, Higham [47]).

Somewhat less well known are Euclidean distance (ED) matrices. A matrix $M \in \mathcal{S}^n$ is an ED matrix if and only if there exist points $x^1, \dots, x^n \in \mathbb{R}^n$ such that, for $i, j = 1, \dots, n$, $M_{ij} = \|x_i - x_j\|_2^2$, *i.e.*, the square of the Euclidean distance between x_i and x_j . A classic result of Schoenberg [87] states that a given matrix $M \in \mathcal{S}^n$ is an ED matrix if and only if the symmetric matrix M' belongs to \mathcal{S}_+^{n-1} , where:

$$\begin{aligned} M'_{ii} &= M_{i,n} \quad (i = 1, \dots, n-1) \\ M'_{ij} &= \frac{1}{2}(M_{in} + M_{jn} - M_{ij}) \quad (1 \leq i < j \leq n-1) \end{aligned}$$

Laurent [55] observes that one can therefore also use SDP to solve various optimisation problems involving ED matrices. This includes problems in, *e.g.*, computational biology (such as molecular conformation problems) and engineering (such as wireless sensor network localisation problems); see Liberti *et al.* [61] for a survey.

5.2. Combinatorial optimisation

SDP has proven to be a remarkably useful tool for constructing strong bounds for various *combinatorial optimisation* problems. For brevity, we consider just two examples: the *stable set* and *max-cut* problems. For other examples, see, *e.g.*, [1, 34, 35, 60, 64, 86, 98].

Let $G = (V, E)$ be an undirected graph. A set $S \subset V$ is called *stable* if no two nodes in S are adjacent in G . The *stability number*, denoted by $\alpha(G)$, is the maximum cardinality of a stable set in G . The *stable set problem* calls for a stable set of maximum cardinality. It is not only \mathcal{NP} -hard in the strong sense, but hard to approximate [42].

A simple 0–1 LP formulation of the stable set problem is:

$$\begin{aligned} \max \quad & \sum_{i \in V} x_i \\ \text{s.t.} \quad & x_i + x_j \leq 1 (\{i, j\} \in E) \\ & x \in \{0, 1\}^n. \end{aligned} \tag{5.1}$$

Unfortunately, the LP relaxation of this formulation yields an extremely weak upper bound, since one can just set every variable to 1/2. Padberg [77] noted that one can strengthen the LP relaxation by replacing the constraints (5.1) with *clique inequalities* of the form $\sum_{i \in C} x_i \leq 1$, where C is a maximal clique (set of pairwise-adjacent nodes) in G . Unfortunately, the number of cliques is in general exponential in $|V|$. Even worse, the *separation problem* for the clique inequalities (*i.e.*, the problem of detecting when an LP solution violates a clique inequality) is \mathcal{NP} -hard (*e.g.*, Grötschel *et al.* [39]).

In his seminal paper, Lovász [63] defined a new upper bound for the stable set problem, which he called $\theta(G)$. Grötschel *et al.* [38, 39] showed that $\theta(G)$ can be computed by solving an SDP. One way to do it is as follows. We begin by formulating the stable set problem as the following continuous quadratic optimisation problem:

$$\begin{aligned} \max \quad & \sum_{i \in V} x_i \\ \text{s.t.} \quad & x_i^2 - x_i = 0 (i \in V) \end{aligned} \tag{5.2}$$

$$\begin{aligned} & x_i x_j = 0 (\{i, j\} \in E) \\ & x \in \mathbb{R}^{|V|}. \end{aligned} \tag{5.3}$$

Now we introduce the matrix

$$Y = \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T = \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix},$$

and note that Y should be psd and have rank 1. We then replace xx^T with a *matrix variable* X , and replace the quadratic terms in (5.2) and (5.3) with the corresponding entries in X . This yields the following alternative formulation of the stable set problem:

$$\begin{aligned} \max \quad & \sum_{i \in V} x_i \\ \text{s.t.} \quad & X_{ii} - x_i = 0 (i \in V) \end{aligned} \tag{5.4}$$

$$X_{ij} = 0 (\{i, j\} \in E) \tag{5.5}$$

$$Y = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathcal{S}_+^{n+1} \tag{5.6}$$

$$\text{rank}(Y) = 1.$$

Dropping the rank constraint, which is non-convex, we obtain the desired SDP relaxation. The corresponding upper bound is $\theta(G)$.

The following result is due to Grötschel *et al.* [39]. To help the reader, we give a short proof here.

Proposition 5.1. *If $(x, X) \in \mathbb{R}^n \times \mathcal{S}^n$ satisfies (5.4)–(5.6), then x satisfies all clique inequalities.*

Proof. Given any clique $C \subseteq V$, let $v(C) \in \{0, 1\}^n$ be a vector with a “1” in position i if and only if node i belongs to C . From constraint (5.6) and the first definition of psd-ness in Section 2.1, we have

$$\begin{pmatrix} -1 \\ v(C) \end{pmatrix}^T \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \begin{pmatrix} -1 \\ v(C) \end{pmatrix} \geq 0,$$

or, equivalently,

$$2 \sum_{i \in C} x_i - \sum_{i \in C} \sum_{j \in C} X_{ij} \leq 1. \quad (5.7)$$

Now, due to constraints (5.4) and (5.5), we have $X_{ii} = x_i$ for all $i \in C$ and $X_{ij} = 0$ for all $\{i, j\} \subseteq C$. Thus, (5.7) reduces to $2 \sum_{i \in C} x_i - \sum_{i \in C} x_i \leq 1$, which is equivalent to the clique inequality on C . \square

Although easy to prove, this result is remarkable, given the above-mentioned fact that clique separation is \mathcal{NP} -hard. Indeed, it is a good illustration of the power of SDP relative to LP. Intuitively, the extra power comes from the fact that imposing psd-ness is equivalent to imposing an infinite number of linear inequalities.

Now we move on to max-cut. Given a graph $G = (V, E)$ and an arbitrary set $S \subseteq V$, the set

$$\{\{i, j\} \in E : i \in S, j \in V \setminus S\}$$

is called an *edge cutset* or simply *cut*. Suppose we are also given a vector $w \in \mathbb{R}^{|E|}$ of edge weights. The *max-cut problem* calls for a cut of maximum total weight. The problem is \mathcal{NP} -hard in the strong sense (Garey *et al.* [31]).

A simple 0–1 LP formulation of max-cut is:

$$\max \sum_{e \in E} w_e x_e$$

$$\text{s.t. } x_{ij} + x_{ik} + x_{jk} \leq 2 \quad (\{i, j, k\} \subset V) \quad (5.8)$$

$$x_{ik} + x_{jk} \geq x_{ij} \quad (\{i, j\} \subset V, k \in V \setminus \{i, j\}) \quad (5.9)$$

$$x_{ij} \in \{0, 1\} \quad (\{i, j\} \subset V).$$

Unfortunately, the LP relaxation can be rather weak. Poljak and Tuza [81] showed that, even when $w \geq 0$, the upper bound from the relaxation can approach twice the weight of the optimal cut.

An SDP relaxation of the max-cut problem was proposed by Schrijver (unpublished), and then studied in, *e.g.*, [26, 36, 57, 79]. For each $i \in V$, let z_i be a variable taking the value 1 if $i \in S$, and -1 otherwise. Then the max-cut problem can be formulated as:

$$\max \frac{1}{2} \sum_{\{i, j\} \in E} (1 - z_i z_j)$$

$$\text{s.t. } z_i^2 = 1 \quad (i \in V)$$

$$z \in \mathbb{R}^{|V|}.$$

The corresponding SDP relaxation is:

$$\begin{aligned} \max \quad & \frac{1}{2} \sum_{\{i,j\} \in E} (1 - Z_{ij}) \\ \text{s.t.} \quad & Z_{ii} = 1 \quad (i \in V) \\ & Z \in \mathcal{S}_+^n. \end{aligned}$$

In a major breakthrough, Goemans and Williamson [36] proved that, when $w \geq 0$, the upper bound from the SDP is no more than 1.131 times the weight of the optimal cut. (They did this by showing how to compute a cut whose weight is at least 0.878 times the SDP bound.)

Observe that the feasible region to the SDP can be projected into x -space *via* the identities $Z_{ij} = 1 - 2x_{ij}$. It turns out that the projection “wraps” rather closely around the convex hull of cut vectors. For example, from a consideration of 3-by-3 principal submatrices of Z , it can be shown that a point in the projection cannot violate any of the triangle inequalities (5.8) and (5.9) by more than 1/4. This provides a partial explanation for the strength of the SDP bound; see, *e.g.*, [8, 30, 57, 58] for more details.

In a similar way, one can construct SDP relaxations of any problem that can be formulated as a 0–1 LP. In fact, it is possible to construct entire *hierarchies* of SDP relaxations; see, *e.g.*, [56, 59, 65, 85, 89]. Moreover, SDP has been successfully applied to many 0–1 *quadratic* programs; see [45, 60, 80, 91] for early work on the subject, and [24, 95] for some recent applications.

5.3. Non-convex quadratic (and polynomial) optimisation

Finally, we consider general quadratic optimisation problems, and their natural generalisation, polynomial optimisation problems.

A quadratically constrained quadratic program (QCQP) is a problem of the form:

$$\begin{aligned} \inf \quad & x^T Q^0 x + c^0 \cdot x \\ \text{s.t.} \quad & x^T Q^k x + c^k \cdot x \leq b_j \quad (k = 1, \dots, m) \\ & x \in \mathbb{R}^n, \end{aligned}$$

where $Q^k \in \mathcal{S}^n$, $c^k \in \mathbb{R}^n$ and $b_k \in \mathbb{R}$ for $k = 0, \dots, m$. If Q^0, \dots, Q^m are all psd, then the QCQP is convex and can be converted into an SOCP (see Sect. 4.1). In general, however, QCQP is \mathcal{NP} -hard in the strong sense. (Indeed, this follows from the fact that the stable set and max-cut problems can be formulated as QCQPs; see the previous subsection). Moreover, non-convex QCQPs arise in many contexts besides OR, including economics and finance (Horst *et al.* [49]) and signal processing (Luo *et al.* [66]).

Ramana [83] proposed the following natural SDP relaxation of QCQP:

$$\begin{aligned} \inf \quad & Q^0 \bullet X + c_0 \cdot x \\ \text{s.t.} \quad & Q^k \bullet X + c_k \cdot x \leq b_k \quad (k = 1, \dots, m) \\ & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathcal{S}_{n+1}^+. \end{aligned}$$

The derivation of this relaxation is similar to the one for the stable set problem described in the previous subsection. (Earlier, Shor [90] derived essentially the same relaxation, but in dual form.)

A hierarchy of SDP relaxations for QCQP, similar to the one of Lovász and Schrijver for 0–1 LPs, was proposed by Fujie and Kojima [29]. Some results on the quality of the SDP bound, similar to the result of Goemans and Williamson [36], are surveyed in Nesterov *et al.* [74].

In practice, the vector x is usually constrained to be non-negative, or even to lie in a hypercube. This fact can be exploited to derive stronger SDP relaxations; see, *e.g.*, [5, 7, 9, 18, 28]. The best such relaxations are shown by Anstreicher [6] to dominate many other known relaxations. There is also evidence that SDP-based algorithms for QCQP can be effective in practice; *e.g.*, [5, 20, 66].

A natural generalisation of QCQP is *polynomial optimisation*, in which the objective and constraint functions can be arbitrary polynomials. Polynomial optimisation is a fascinating inter-disciplinary field, with relevance not only to OR, but also to statistics, computer science, engineering, and branches of pure mathematics such as algebraic geometry, commutative algebra and moment theory [4, 25]. Interesting and powerful SDP relaxations and hierarchies for polynomial optimisation have been proposed by, *e.g.*, Lasserre [54] and Parrilo [78]. However, the complexity of solving the SDPs in the hierarchies is still not fully settled; see O’Donnell [76].

6. SOFTWARE

Finally, we mention some of the available software systems. We cover modelling interfaces in Section 6.1 and solvers in Section 6.2. For brevity, we do not cover all systems, and refer the reader to Mittelmann [70] for a more comprehensive survey. (Readers may also find the “Decision Tree for Optimization Software”¹ useful.)

6.1. Modelling interfaces

We are aware of the following five modelling interface systems:

- CVX² is an influential and long-established system for conic programming based on MatLab. The CVX website states “CVX is a popular modeling framework for disciplined convex programming that... turns MatLab into a modeling language, allowing constraints and objectives to be specified using standard MatLab syntax.”
- CVXOPT³ is similar to CVX but based on Python. It can be used within the open-source mathematics software system SageMath.
- JuliaOpt⁴ now contains a modelling language called JuMP.
- YALMIP⁵ is also based on MatLab, and acts as a convenient front-end to many solvers.
- PICOS⁶ provides a front-end to many solvers and is based on Python. It also does some automatic reformulation (*e.g.*, it has a function that “replaces quadratic constraints by equivalent second order cone constraints”).

We remark that, although using modelling languages makes it much easier to model and solve problems, there is usually a computational overhead, in terms of both time and memory.

6.2. CO solvers

Now we list some of the commonly-used solvers. (A more comprehensive list can be found on the YALMIP site.⁷)

- CSDP⁸ is a C library for solving SDPs using an interior-point method. It is now part of COIN-OR.⁹
- ECOS¹⁰ is an SOCP solver designed for embedded systems. It is written using less than a thousand lines of C code, and it can also handle constraints involving the exponential cone.

¹<http://plato.asu.edu/guide.html>

²<http://cvxr.com/>

³<http://cvxopt.org/>

⁴<http://www.juliaopt.org>

⁵<http://yalmip.github.io/allsolvers/>

⁶<http://picos.zib.de>

⁷<https://yalmip.github.io/allsolvers/>

⁸<https://projects.coin-or.org/Csdp/wiki/CSDPUsed>

⁹<https://projects.coin-or.org/>

¹⁰<https://www.embotech.com/ECOS>

- MOSEK¹¹ has supported both SOCP and SDP since version 7.0. A recent blog entry¹² reports on the solution of an SOCP with over 0.5 million variables (on a powerful parallel processor).
- PENNON¹³ is an implementation of a generalized augmented Lagrangian algorithm for SDPs. It can solve SDPs with general convex objective and constraint functions.
- SDPA¹⁴ is a collection of C++ routines for SDP, based on a primal-dual interior-point method. It is designed to exploit sparsity in the constraints.
- SDPLR¹⁵ is a C library for SDP, based on the augmented Lagrangian method.
- SDPT3¹⁶ is a library of MatLab routines that can solve SOCPs, SDPs, and various other problems. An infeasible path following algorithm is used.
- SeDuMi¹⁷ is another useful MatLab toolbox, based on the concept of “self-dual embedding”.

Finally, we mention two MatLab packages that are specifically designed for solving polynomial optimisation problems: GloptiPoly¹⁸ and SOSTOOLS¹⁹.

7. CONCLUSIONS

In this guide, we hope to have convinced the reader that Conic Optimisation is an elegant and powerful generalisation of standard linear programming, which allows one to capture many forms of non-linearity that arise in practical problems. In particular, both SOCP and SDP enable one to model many *convex* non-linearities arising in practice (such as convex quadratic and hyperbolic functions, and functions involving norms or eigenvalues); and SDP also provides good bounds and approximation algorithms for many *non-convex* (and \mathcal{NP} -hard) problems, including a wide range of combinatorial and global optimisation problems.

Moreover, software for SOCP and SDP, and CO in general, is developing rapidly. This includes not only solvers, but also modelling languages and procedures for automatic reformulation. Hence, although CO is more difficult to master than LP, we believe that it will soon become a standard technique for both practitioners and researchers in optimisation, just as LP is at present.

REFERENCES

- [1] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.* **5** (1995) 13–51.
- [2] F. Alizadeh and D. Goldfarb, Second-order cone programming. *Math. Program.* **95** (2003) 3–51.
- [3] F. Alizadeh, J.-P.A. Haeberly and M.L. Overton, Complementarity and nondegeneracy in semidefinite programming. *Math. Program.* **77** (1997) 111–128.
- [4] M. Anjos and J.B. Lasserre, Handbook on Semidefinite, Conic and Polynomial Optimization. Vol. 166 of *International Series in OR/MS*. Springer, New York (2012).
- [5] K.M. Anstreicher, Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *J. Glob. Optim.* **43** (2009) 471–484.
- [6] K.M. Anstreicher, On convex relaxations for quadratically constrained quadratic programming. *Math. Program.* **136** (2012) 233–251.
- [7] K.M. Anstreicher and S. Burer, Computable representations for convex hulls of low-dimensional quadratic forms. *Math. Program.* **124** (2010) 33–43.
- [8] D. Avis and J. Umemoto, Stronger linear programming relaxations of max-cut. *Math. Program.* (2003) **97** 451–469.
- [9] X. Bao, N.V. Sahinidis and M. Tawarmalani, Semidefinite relaxations for quadratically constrained quadratic programming: a review and comparisons. *Math. Program.* **129** (2011) 129–157.
- [10] S.J. Benson, Y. Ye and X. Zhang, Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM J. Optim.* **10** (2000) 443–461.

¹¹<http://www.mosek.com/>

¹²<http://blog.mosek.com/2017/05/biggest-conic-quadratic-problem-solved.html>

¹³<http://web.mat.bham.ac.uk/kocvara/pennon/>

¹⁴<http://sdpa.sourceforge.net/>

¹⁵<http://sburer.github.io/projects.html>

¹⁶<http://www.math.cmu.edu/~reha/sdpt3.html>

¹⁷<http://sedumi.ie.lehigh.edu/>

¹⁸<http://homepages.laas.fr/henrion/software/gloptipoly3/>

¹⁹<http://www.cds.caltech.edu/sostools/>

- [11] A. Ben-Tal and A. Nemirovski, Robust solutions of uncertain linear programs. *Oper. Res. Lett.* **25** (1999) 1–13.
- [12] A. Ben-Tal and A. Nemirovski, On polyhedral approximations of the second order cone. *Math. Oper. Res.* **26** (2001) 193–205.
- [13] D. Bertsekas, *Nonlinear Programming*, 3rd edn. Athena Scientific, Belmont, MA (2016).
- [14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Cambridge (2004).
- [15] G. Braun, S. Fiorini, S. Pokutta and D. Steurer, Approximation limits of linear programs (beyond hierarchies). *Math. Oper. Res.* **40** (2015) 756–772.
- [16] S. Burer, On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.* **120** (2009) 479–495.
- [17] S. Burer, Copositive programming, in *Handbook on Semidefinite, Conic and Polynomial Optimization*, edited by M. Anjos and J.B. Lasserre. Springer, New York (2012) 201–218.
- [18] S. Burer and A.N. Letchford, On non-convex quadratic programming with box constraints. *SIAM J. Optim.* **20** (2009) 1073–1089.
- [19] S. Burer and R.D.C. Monteiro, Local minima and convergence in low-rank semidefinite programming. *Math. Program.* **103** (2005) 427–444.
- [20] S.A. Burer and D. Vandembussche, A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations. *Math. Program.* **113** (2008) 259–282.
- [21] V. Chandrasekaran and P. Shah, Relative entropy optimization and its applications. *Math. Program.* **161** (2017) 1–32.
- [22] G.B. Dantzig, Maximization of a linear function of variables subject to linear inequalities, in *Activity Analysis of Production and Allocation*, edited by T.C. Koopmans. Wiley, New York (1951) 339–347.
- [23] G.B. Dantzig, *Programming and Extensions*. Princeton University Press, Princeton, NJ (1963).
- [24] E. de Klerk, R. Sotirov and U. Truetsch, A new semidefinite programming relaxation for the quadratic assignment problem and its computational perspectives. *INFORMS J. Comput.* **27** (2015) 378–391.
- [25] J.A. De Loera, R. Hemmecke and M. Köppe, Algebraic and Geometric Ideas in the Theory of Discrete Optimization. Vol. 14 of *SIAM-MOS Series on Optimization*. SIAM, Philadelphia, PA (2013).
- [26] C. Delorme and S. Poljak, Combinatorial properties and the complexity of an eigenvalue approximation of the max-cut problem. *Eur. J. Combin.* **14** (1993) 313–333.
- [27] M.M. Deza and M. Laurent, *Geometry of Cuts and Metrics*. Springer-Verlag, Berlin (1997).
- [28] M. Dür, Copositive programming: a survey, in *Recent Advances in Optimization and its Applications in Engineering*, edited by M. Diehl *et al.* Springer, Berlin (2010) 3–20.
- [29] T. Fujie and M. Kojima, Semidefinite programming relaxation for nonconvex quadratic programs. *J. Glob. Optim.* **10** (1997) 367–380.
- [30] L. Galli, K. Kaparis and A.N. Letchford, Complexity results for the gap inequalities for the max-cut problem. *Oper. Res. Lett.* **40** (2012) 149–152.
- [31] M.R. Garey, D.S. Johnson and L. Stockmeyer, Some simplified NP-complete graph problems. *Theor. Comput. Sci.* **1** (1976) 237–267.
- [32] F. Glineur, An extended conic formulation for geometric optimization. *Found. Comput. Decis. Sci.* **25** (2000) 161–174.
- [33] F. Glineur and T. Terlaky, Conic formulation for ℓ_p -norm optimization. *J. Optim. Th. Appl.* **122** (2004) 285–307.
- [34] M.X. Goemans, Semidefinite programming in combinatorial optimization. *Math. Program.* **79** (1997) 143–161.
- [35] M.X. Goemans and F. Rendl, Combinatorial optimization, in *Handbook of Semidefinite Programming*, edited by H. Wolkowicz, E. Saigal and L. Vandenberghe. Springer, New York (2000) 343–360.
- [36] M.X. Goemans and D. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM* **42** (1995) 1115–1145.
- [37] J. Gondzio, Interior point methods 25 years later. *Eur. J. Oper. Res.* **218** (2012) 587–601.
- [38] M. Grötschel, L. Lovász and A.J. Schrijver, The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica* **1** (1981) 169–197.
- [39] M. Grötschel, L. Lovász and A.J. Schrijver, *Geometric Algorithms and Combinatorial Optimization*. Wiley, New York (1988).
- [40] M. Hall Jr., Discrete problems, in *A Survey of Numerical Analysis*, edited by J. Todd. McGraw-Hill, New York (1962) 518–542.
- [41] F.M. Harris, How many parts to make at once. *Factory* **10** (1913) 135–136, 152.
- [42] J. Hästad, Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Math.* **182** (1999) 105–142.
- [43] C. Helmberg, Fixing variables in semidefinite relaxations. *SIAM J. Matrix Anal. Appl.* **21** (2000) 952–969.
- [44] C. Helmberg, Semidefinite programming. *Eur. J. Oper. Res.* **137** (2002) 461–482.
- [45] C. Helmberg and F. Rendl, Solving quadratic (0, 1)-programs by semidefinite programs and cutting planes. *Math. Program.* **82** (1998) 291–315.
- [46] C. Helmberg and F. Rendl, A spectral bundle method for semidefinite programming. *SIAM J. Optim.* **10** (1999) 673–696.
- [47] N. Higham, Computing the nearest correlation matrix—A problem from finance. *IMA J. Numer. Anal.* **22** (2002) 329–343.
- [48] R.A. Horn, *Matrix Analysis*, 2nd edn. Cambridge University Press, Cambridge (2012).
- [49] R. Horst, P.M. Pardalos and N. Thoai, *Introduction to Global Optimization*, 2nd edn. Kluwer, Dordrecht (2000).
- [50] C.R. Johnson, Matrix completion problems: a survey, in *Matrix Theory and Applications*, edited by C.R. Johnson. American Mathematical Society, Providence, RI (1990) 171–198.
- [51] L.G. Khachiyan, A polynomial algorithm in linear programming. *Soviet Math. Doklady* **20** (1979) 191–194.

- [52] K. Krishnan and J.E. Mitchell, A unifying framework for several cutting plane methods for semidefinite programming. *Optim. Meth. Soft.* **21** (2006) 57–74.
- [53] Y.-J. Kuo and H.D. Mittelmann, Interior-point methods for second-order cone programming and OR applications. *Comput. Optim. Appl.* **28** (2004) 255–285.
- [54] J.B. Lasserre, Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11** (2001) 796–817.
- [55] M. Laurent, A connection between positive semidefinite and Euclidean distance matrix completion problems. *Lin. Alg. Appl.* **273** (1998) 9–22.
- [56] M. Laurent, A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0–1 programming. *Math. Oper. Res.* **28** (2003) 470–496.
- [57] M. Laurent and S. Poljak, On a positive semidefinite relaxation of the cut polytope. *Lin. Alg. Appl.* **223/224** (1995) 439–461.
- [58] M. Laurent and S. Poljak, Gap inequalities for the cut polytope. *Eur. J. Combin.* **17** (1996) 233–254.
- [59] M. Laurent and F. Rendl, Semidefinite programming and integer programming, in *Handbook on Discrete Optimization*, edited by K. Aardal *et al.* Elsevier Amsterdam (2005) 393–514.
- [60] C. L  mar  chal and F. Oustry, SDP relaxations in combinatorial optimization from a Lagrangian viewpoint, in *Advances in Convex Analysis and Global Optimization*, edited by N. Hadjisawas and P.M. Pardalos. Kluwer, Dordrecht (2001) 119–134.
- [61] L. Liberti, C. Lavor, N. Maculan and A. Mucherino, Euclidean distance geometry and applications. *SIAM Rev.* **56** (2014) 3–69.
- [62] M.S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret, Applications of second-order cone programming. *Lin. Alg. Appl.* **284** (1998) 193–228.
- [63] L. Lov  sz, On the Shannon capacity of a graph. *IEEE Trans. Inform. Th.* **IT-25** (1979) 1–7.
- [64] L. Lov  sz, Semidefinite programs and combinatorial optimization, in *Recent Advances in Algorithms and Combinatorics*, edited by B. Reed and C.L. Sales. Springer, New York (2003) 137–194.
- [65] L. Lov  sz and A.J. Schrijver, Cones of matrices and set-functions and 0–1 optimization. *SIAM J. Optim.* **1** (1991) 166–190.
- [66] Z.-Q. Luo, W.-K. Ma, A.M.-C. So, Y. Ye and S. Zhang, Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Process. Mag.* **27** (2010) 20–34.
- [67] J. Malick, J. Povh, F. Rendl and A. Wiegele, Regularization methods for semidefinite programming. *SIAM J. Optim.* **20** (2009) 336–356.
- [68] H.M. Markowitz, Portfolio selection. *J. Finance* **7** (1952) 77–91.
- [69] J.E. Mitchell, Restarting after branching in the SDP approach to MAX-CUT and similar combinatorial optimization problems. *J. Combin. Optim.* **5** (2001) 151–166.
- [70] H.D. Mittelmann, The state-of-the-art in conic optimization software, in *Handbook on Semidefinite, Conic and Polynomial Optimization*, edited by M. Anjos and J.B. Lasserre. Springer, New York (2012) 671–686.
- [71] T.S. Motzkin, Copositive quadratic forms. *Nat. Bur. Stand. Rep.* **1818** (1952) 11–22.
- [72] K.G. Murty and S.N. Kabadi, Some \mathcal{NP} -complete problems in quadratic and nonlinear programming. *Math. Program.* **39** (1987) 117–129.
- [73] Y. Nesterov and A. Nemirovsky, *Interior Point Methods in Convex Programming: Theory and Applications*. SIAM Press, Philadelphia, PA (1994).
- [74] Y. Nesterov, H. Wolkowicz and Y. Ye, Semidefinite Programming Relaxations of Nonconvex Quadratic Optimization, in *Handbook of Semidefinite Programming*, edited by H. Wolkowicz, E. Saigal and L. Vandenberghe. Springer, New York (2000) 361–419.
- [75] J. Nie, K. Ranestad and B. Sturmfels, The algebraic degree of semidefinite programming. *Math. Program.* **122** (2010) 379–405.
- [76] R. O’Donnell, *SOS is not Obviously Automatable*, *Even Approximately*. ECCO Report No. 141 (2016).
- [77] M.W. Padberg, On the facial structure of set packing polyhedra. *Math. Program.* **5** (1973) 199–215.
- [78] P. Parrilo, Semidefinite programming relaxations for semialgebraic problems. *Math. Program.* **96** (2003) 293–320.
- [79] S. Poljak and F. Rendl, Non-polyhedral relaxations of graph-bisection problems. *SIAM J. Optim.* **5** (1995) 467–487.
- [80] S. Poljak, F. Rendl and H. Wolkowicz, A recipe for semidefinite relaxation for (0, 1)-quadratic programming. *J. Glob. Optim.* **7** (1995) 51–73.
- [81] S. Poljak and Zs. Tuza, The expected relative error of the polyhedral approximation of the max-cut problem. *Oper. Res. Lett.* **16** (1994) 191–198.
- [82] J. Povh, F. Rendl and A. Wiegele, A boundary point method to solve semidefinite programs. *Computing*, **78** (2009) 277–286.
- [83] M. Ramana, *An Algorithmic Analysis of Multiquadratic and Semidefinite Programming Problems*. PhD thesis, Johns Hopkins University, Baltimore, MD (1993).
- [84] M.V. Ramana, An exact duality theory for semidefinite programming and its complexity implications. *Math. Program.* **77** (1997) 129–162.
- [85] F. Rendl, Semidefinite relaxations for integer programming, in *50 Years of Integer Programming 1958–2008*, edited by M. J  nger *et al.* Springer, Heidelberg (2010) 687–726.
- [86] F. Rendl, Semidefinite relaxations for partitioning, assignment and ordering problems. *4OR* **10** (2012) 321–346.
- [87] I.J. Schoenberg, Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.* **44** (1938) 522–536.
- [88] C. Seligman, *Online Astronomy Text*. Available at: <http://cseligman.com/text/history/ellipses.htm> (1993).
- [89] H.D. Sherali and W.P. Adams, A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discr. Math.* **3** (1990) 411–430.
- [90] N.Z. Shor, Quadratic optimization problems. *Sov. J. Comput. Syst. Sci.* **25** (1987) 1–11.

- [91] N.Z. Shor, Dual quadratic estimates in polynomial and Boolean programming. *Ann. Oper. Res.* **25** (1990) 163–168.
- [92] A. Skajaa, E.D. Andersen and Y. Ye, Warmstarting the homogeneous and self-dual interior point method for linear and conic quadratic problems. *Math. Program. Comput.* **5**(2013) 1–25.
- [93] M.J. Todd, Semidefinite optimization. *Acta Numerica* **10** (2001) 515–560.
- [94] L. Vanderberghe and S. Boyd, Semidefinite programming. *SIAM Rev.* **38** (1996) 49–95.
- [95] P. Wang, C. Shen, A. van den Hengel and P.H.S. Torr, Large-scale binary quadratic optimization using semidefinite relaxation and applications. *IEEE Trans. Patt. Anal. Mach. Intel.* **39** (2017) 470–485.
- [96] Z. Wen, D. Goldfarb and W. Yin, Alternating direction augmented Lagrangian methods for semidefinite programming. *Math. Program. Comput.* **2** (2010) 203–230.
- [97] H. Wolkowicz, E. Saigal and L. Vandenberghe, Handbook of Semidefinite Programming. Vol. 27 of *International Series in OR/MS*. Springer, New York (2000).
- [98] H. Wolkowicz and M.F. Anjos, Semidefinite programming for discrete optimization and matrix completion problems. *Discr. Appl. Math.* **123** (2002) 513–577
- [99] H. Ziegler, Solving certain singly constrained convex optimization problems in production planning. *Oper. Res. Lett.* **1** (1982) 246–252.