Step by Step Design of an Interior-Point Solver in Self-Dual Conic Optimization, With Applications

Jean Charles Gilbert

Step by step design of an interior-point solver in self-dual conic optimization with applications

Jean Charles Gilbert, INRIA

December 18, 2019

These notes present a project in numerical optimization dealing with the implementation of an interior-point method for solving a self-dual conic optimization (SDCO) problem. The cone is the Cartesian product of cones of positive semidefinite matrices of various dimensions (imposing to matrices to be positive semidefinite) and of a positive orthant (imposing to a vector to have nonnegative components). Therefore, the solved problem encompasses semidefinite and linear optimization.

The project was given in a course entitled *Advanced Continuous Optimization II* at the University Paris-Saclay, in 2016-2020. The solver is designed step by step during a series of 5 sessions of 4 hours each. Each session corresponds to a chapter of these notes (or a part of it). The correctness of the SDCO solver is verified during each session on small academic problems, having diverse properties. During the last session, the developed piece of software is used to minimize a univariate polynomial on an interval and to solve a few small size rank relaxations of QCQO (quadratically constrained quadratic optimization) problems, modeling various instances of the OPF (optimal power flow) problem. The student has to master not ony the implementation of the interior-point solver, but is also asked to understand the underliying theory by solving exercises consisting in proving some properties of the implemented algorithms.

The goal of the project is *not* to design an SDCO solver that would beat the best existing solver but to help the students to understand and demystify what there is inside such a piece of software. As a side outcome, this course also shows that a rather performent SDCO solver can be realized in a relatively short time. In addition, the student will be able to improve the developed piece of software, in case this is required by professionel needs. For a review of SDCO codes and more details on their development, we refer the reader to [51, 1], in particular [29, 49], which have guided us in the composition of these notes.

In case you use these notes in a paper, thanks to cite them like in [14].

Table of Contents

1 Directions of displacement

This first session introduces the *self-dual conic optimization* (SDCO) problem to solve. It also shows how to compute a direction of displacement, which will be used at each iteration by the interior point algorithms described in the other sessions.

1.1 The problem to solve

1.1.1 Space structures

We denote by $\mathbb{E}$ the Euclidean vector space of the variables defining the primal form of the convex optimization considered in all this project; see section 1.1.2. It is equipped with a scalar product denoted by $\langle \cdot, \cdot \rangle_{\mathbb{E}}$. Conic optimization is formulated thanks to a cone of $\mathbb{E}$, which is denoted by K. By definition a *cone* K of $\mathbb{E}$ is a set verifying $\mathbb{R}_{++}K \subset K$, where $R_{++} := \{t \in \mathbb{R} : t > 0\}$, meaning that tx must belong to K as soon as $t > 0$ and $x \in K$. The *dual cone* of K is the closed convex cone defined by

$$K^+ := \{y \in \mathbb{E} : \langle y, x \rangle_{\mathbb{E}} \geqslant 0 \text{ for all } x \in K\}.$$

The cone K is said to be *self-dual* if $K^+ = K$. Hence, a self-dual cone is necessarily closed and convex. The unknowns of the optimization problem we consider below are elements of a vector space $\mathbb{E}$ that are constrained to belong to some self-dual cone K and to satisfy some affine constraint.

More specifically, the vector space $\mathbb{E}$ considered in this project is the Cartesian product

$$\mathbb{E} := \mathcal{S}^{n_1} \times \cdots \times \mathcal{S}^{n_s} \times \mathbb{R}^{n_l},$$

where, for $j \in [1\,{:}\,s]$, $\mathcal{S}^{n_j}$ is the *space of symmetric real matrices* of order n_j; it has dimension $n_j(n_j+1)/2$. This means that the unknowns of the optimization problem below are s symmetric matrices of order $n_1, \ldots, n_s$ and a vector of variables of dimension n_l. An element $x \in \mathbb{E}$ will be denoted by

$$x = (x^{s,1}, \ldots, x^{s,s}, x^l),$$

where $x^{s,j} \in \mathcal{S}^{n_j}$, for $j \in [1\,{:}\,s]$, and $x^l \in \mathbb{R}^{n_l}$ (unusually, we have chosen to represent matrices by lower case letters, because in $\mathbb{E}$ they are assembled with a vector denoted by a lower case letter).

The scalar product on some $\mathcal{S}^{n_j}$ is defined and denoted by

$$\langle\cdot,\cdot\rangle_{\mathcal{S}^{n_j}} : (a,b) \in \mathcal{S}^{n_j} \times \mathcal{S}^{n_j} \mapsto \langle a,b\rangle_{\mathcal{S}^{n_j}} := \operatorname{tr} ab = \sum_{k,l=1}^{n_j} a_{kl} b_{kl},$$

where $\operatorname{tr} a := \sum_i a_{ii}$ denotes the *trace* of the matrix $a \in \mathbb{R}^{n_j \times n_j}$. The associated norm is the so-called *Frobenius norm*; its value at $a \in \mathcal{S}^{n_j}$ is denoted by

$$\|a\|_{\mathcal{S}^{n_j}} = \left(\sum_{k,l=1}^{n_j} a_{kl}^2 \right)^{1/2}.$$

The Euclidean scalar product is supposed to equip $\mathbb{R}^{n_l}$. Finally the scalar product on $\mathbb{E}$ takes at $(x,s) \in \mathbb{E} \times \mathbb{E}$ the value

$$\langle x, s\rangle_{\mathbb{E}} := \sum_{j=1}^{s} \langle x^{s,j}, s^{s,j}\rangle_{\mathcal{S}^{n_j}} + (x^l)^{\mathsf{T}}(s^l).$$

The associated norm is

$$\|x\|_{\mathbb{E}} := \left(\sum_{j=1}^{s} \|x^{s,j}\|_{\mathcal{S}^{n_j}}^2 + \|x^l\|_2^2 \right)^{1/2}.$$

A product of two symmetric matrices is not necessary symmetric. Since these products will appear below, it is natural to introduce the vector space

$$\mathbb{F} := \mathbb{R}^{n_1 \times n_1} \times \cdots \times \mathbb{R}^{n_s \times n_s} \times \mathbb{R}^{n_l},$$

where $\mathbb{R}^{n_j \times n_j}$ denotes the set of square real matrices of order n_j. Of course, $\mathbb{E} \subset \mathbb{F}$. The dimension of $\mathbb{F}$ is given by

$$n_{\mathbb{F}} := \sum_{j=1}^{s} n_j^2 + n_l. \tag{1.1}$$

The space $\mathbb{F}$ has its own scalar product, which takes at $(x,s) \in \mathbb{F} \times \mathbb{F}$ the value

$$\langle x, s\rangle_{\mathbb{F}} := \sum_{j=1}^{s} \operatorname{tr}(x^{s,j})^{\mathsf{T}}(s^{s,j}) + (x^l)^{\mathsf{T}}(s^l). \tag{1.2}$$

The associated norm is $\|x\|_{\mathbb{F}} := \langle x, x\rangle_{\mathbb{F}}^{1/2}$. Clearly, when x and $s \in \mathbb{E}$, $\langle x, s\rangle_{\mathbb{F}} = \langle x, s\rangle_{\mathbb{E}}$. The vector space $\mathbb{F}$ appears for example, when one defines the *Hadamard product* of two vectors x and $s \in \mathbb{E}$, which is the vector

$$x \bullet s = (x^{s,1}s^{s,1}, \ldots, x^{s,s}s^{s,s}, x^l \cdot s^l) \in \mathbb{F},$$

where $x^{s,j}s^{s,j}$ is the product of the matrices $x^{s,j}$ and $s^{s,j}$ and $x^l \cdot s^l$ is the standard Hadamard product of two vectors, which is $(x^l \cdot s^l)_i = x_i^l s_i^l$ for all $i \in [1\,{:}\,n_l]$. The fact

that $x \bullet s$ may not be in $\mathbb{E}$ when x and $s \in \mathbb{E}$ will be a source of trouble in the sequel. The *double Hadamard product*

$$x \bullet s \bullet v$$

is equivalently defined by $(x \bullet s) \bullet v$ or $x \bullet (s \bullet v)$. For $x = (x^{s,1}, \ldots, x^{s,s}, x^l) \in \mathbb{F}$, one defines

$$x^{\mathsf{T}} = ((x^{s,1})^{\mathsf{T}}, \ldots, (x^{s,s})^{\mathsf{T}}, x^l) \in \mathbb{F}. \tag{1.3}$$

Note the calculus rules

$$(a \bullet b)^{\mathsf{T}} = b^{\mathsf{T}} \bullet a^{\mathsf{T}}, \tag{1.4a}$$

$$\langle a \bullet b, c \rangle_{\mathbb{F}} = \langle a, c \bullet b^{\mathsf{T}} \rangle_{\mathbb{F}} = \langle b, a^{\mathsf{T}} \bullet c \rangle_{\mathbb{F}}. \tag{1.4b}$$

Note also that, when $a \in \mathbb{F}$, $a + a^{\mathsf{T}} \in \mathbb{E}$.

The self-dual cone K of $\mathbb{E}$ considered in this project is a Cartesian product of cones, namely

$$K := \mathcal{S}_+^{n_1} \times \cdots \times \mathcal{S}_+^{n_s} \times \mathbb{R}_+^{n_l},$$

where $\mathcal{S}_+^{n_j}$ is the *cone of positive semidefinite matrices* of $\mathcal{S}^{n_j}$ and $\mathbb{R}_+^{n_l} := \{x \in \mathbb{R}^{n_l} : x \geqslant 0\}$ is the *nonnegative orthant* of $\mathbb{R}^{n_l}$ (vector inequalities have to be understood componentwise: $x \geqslant 0$ iff $x_i \geqslant 0$ for all i). The cone K is self-dual since this is the case for $\mathcal{S}_+^{n_j}$ and $\mathbb{R}_+^{n_l}$. This choice of K implies that the unknown matrices in $\mathcal{S}^{n_j}$ of the optimization problem are forced to be positive semidefinite and that the unknown vector in $\mathbb{R}^{n_l}$ is forced to have nonnegative components. We associate with K the operator $\min_K : \mathbb{E} \to \mathbb{R}$ defined at $x \in \mathbb{E}$ by

$$\min_K(x) = \min \left(\lambda_{\min}(x^{s,1}), \ldots, \lambda_{\min}(x^{s,s}), \min_{i \in [1\,:\,n_l]} x_i^l \right), \tag{1.5}$$

where $\lambda_{\min}(x^{s,j})$ is the smallest eigenvalue of the symmetric matrix $x^{s,j}$. Therefore $x \in K$ if and only if $\min_K(x) \geqslant 0$. The *strict feasible sets* of the optimization problems (P) and (D) defined below make use of the *strict cone*

$$K^s := \mathcal{S}_{++}^{n_1} \times \cdots \times \mathcal{S}_{++}^{n_s} \times \mathbb{R}_{++}^{n_l},$$

where $\mathcal{S}_{++}^{n_j}$ is the cone of positive definite matrices of $\mathcal{S}^{n_j}$ and $\mathbb{R}_{++}^{n_l} := \{x \in \mathbb{R}^{n_l} : x > 0\}$ is the *positive orthant* (strict inequalities also act componentwise: $x > 0$ iff $x_i > 0$ for all i). Clearly, $x \in K^s$ if and only if $\min_K(x) > 0$. For $x = (x^{s,1}, \ldots, x^{s,s}, x^l)$ in K^s, one can define

$$\begin{array}{rcl} x^{-1} & := & ((x^{s,1})^{-1}, \ldots, (x^{s,s})^{-1}, (x^l)^{-1}) \in K^s, \\ x^{-\mathsf{T}} & := & (x^{\mathsf{T}})^{-1} = (x^{-1})^{\mathsf{T}} \in K^s, \\ x^{1/2} & := & ((x^{s,1})^{1/2}, \ldots, (x^{s,s})^{1/2}, (x^l)^{1/2}) \in K^s, \end{array}$$

where the exponent -1 (resp. $1/2$) refers to the inverse (resp. square root) of a positive definite matrix or the componentwise inverse (resp. square root) of a positive vector.

A particular element of K^s, which is used continually below, is

$$e := (I^{n_1}, \ldots, I^{n_s}, e^l),$$

where I^{n_j} is the identity matrix of order n_j and e^l is the vector of all ones in $\mathbb{R}^{n_l}$. Observe that $x \bullet e = e \bullet x = x$,

$$\langle x \bullet s, e\rangle_{\mathbb{E}} = \langle x, s\rangle_{\mathbb{E}}, \qquad \text{and} \qquad n_c := \|e\|_{\mathbb{E}}^2 = \sum_{j=1}^{s} n_j + n_l. \tag{1.6}$$

The number n_c is called the *complexity module* of the problem (hence the index c). Note the difference with the dimension $n_{\mathbb{F}}$ of the space $\mathbb{F}$, defined in (1.1), in which the square of the dimensions n_j appears instead. We will see with (2.9) that the number of iterations to reach optimality at a given precision depends on the square root of n_c.

1.1.2 The primal and dual SDCO problems

The primal (P) and dual (D) *self-dual conic optimization* (SDCO) problems read

$$\boxed{(P) \quad \left\{ \begin{array}{l} \inf \langle c, x\rangle_{\mathbb{E}} \\ A(x) = b \\ x \in K \end{array} \right.} \qquad \text{and} \qquad \boxed{(D) \quad \left\{ \begin{array}{l} \sup\ b^{\mathsf{T}} y \\ A^*(y) + s = c \\ s \in K, \end{array} \right.} \tag{1.7}$$

where $c \in \mathbb{E}$, $A : \mathbb{E} \to \mathbb{R}^m$ is a linear map, $b \in \mathbb{R}^m$, and $A^* : \mathbb{R}^m \to \mathbb{E}$ is the adjoint of A when $\mathbb{R}^m$ is equipped with the Euclidean scalar product. These problems are Lagrangian dual to each other. They are convex in the sense that their objective is convex and their feasible set is also convex. When $s = 0$ and $n_l \neq 0$, one recovers the *linear optimization* (LO) problem; when $s = 1$ and $n_l = 0$, one recovers the standard *semidefinite optimization* (SDO) problem.

The *feasible sets* of (P) and (D) are respectively denoted by

$$\mathcal{F}_P := \{x \in K : A(x) = b\} \quad \text{and} \quad \mathcal{F}_D := \{(y, s) \in \mathbb{R}^m \times K : A^*(y) + s = c\},$$

and their *strictly feasible sets* are denoted by

$$\mathcal{F}_P^s := \{x \in K^s : A(x) = b\} \quad \text{and} \quad \mathcal{F}_D^s := \{(y, s) \in \mathbb{R}^m \times K^s : A^*(y) + s = c\}.$$

Accordingly, a point $x \in \mathcal{F}_P^s$ is said to be *strictly feasible for* (P) and a pair $(y, s) \in \mathcal{F}_D^s$ is said to be *strictly feasible for* (D). We also introduce the following Cartesian products

$$\mathcal{F} := \mathcal{F}_P \times \mathcal{F}_D \quad \text{and} \quad \mathcal{F}^s := \mathcal{F}_P^s \times \mathcal{F}_D^s.$$

The solution sets of these problems are denoted by

$$\mathrm{Sol}(P) \qquad \text{and} \qquad \mathrm{Sol}(D)$$

and their optimal values by

$$\mathrm{val}(P) \qquad \text{and} \qquad \mathrm{val}(D).$$

The *duality gap* is zero if $\mathrm{val}(D) = \mathrm{val}(P)$ (with possible infinite values) and, otherwise, is the positive difference $\mathrm{val}(P) - \mathrm{val}(D) > 0$.

By the Riesz-Fréchet theorem, there are elements $a_i \in \mathbb{E}$ such that for all $x \in \mathbb{E}$, there holds

$$A(x) = \begin{pmatrix} \langle a_1, x\rangle_{\mathbb{E}} \\ \vdots \\ \langle a_m, x\rangle_{\mathbb{E}} \end{pmatrix}. \tag{1.8}$$

Each $a_i \in \mathbb{E}$ is therefore an assembling of s matrices $a_i^{s,j} \in \mathcal{S}^{n_j}$, for $j \in [1\,{:}\,s]$, and a vector a_i^l in $\mathbb{R}^{n_l}$. It will be also assumed that the Euclidean scalar product is used on $\mathbb{R}^m$, in which case $A^*(y)$ is a weighted sum of the vectors a_i (see question 1.1):

$$A^*(y) = \sum_{i=1}^{m} y_i a_i \in \mathbb{E}, \tag{1.9}$$

where the y_i's are the components of $y \in \mathbb{R}^m$.

1.1.3 The central path

It is known that when $\mathcal{F}^s \neq \varnothing$, $z := (x, y, s)$ is a primal-dual solution to (P) and (D) if and only if

$$\begin{cases} A^*(y) + s = c, \quad s \in K \\ A(x) = b, \quad x \in K \\ \langle x, s\rangle_{\mathbb{E}} = 0. \end{cases} \tag{1.10}$$

These may be considered as the necessary and sufficient optimality conditions of the convex problems (P) or (D), under the "constraint qualification condition" $\mathcal{F}^s \neq \varnothing$ (it is known, indeed, that "$\mathcal{F}^s \neq \varnothing$ and the surjectivity of A" is equivalent to Robinson's constraint qualification [42] at one or any point of the feasible sets of the problems (P) and (D) [13]). It can be shown (see question 1.2) that for x and $s \in K$, there holds

$$\langle x, s\rangle_{\mathbb{E}} = 0 \qquad \Longleftrightarrow \qquad x \bullet s = 0. \tag{1.11}$$

Therefore, (1.10) also reads

$$\begin{cases} A^*(y) + s = c, \quad s \in K \\ A(x) = b, \quad x \in K \\ x \bullet s = 0. \end{cases} \tag{1.12}$$

We prefer (1.12) to (1.10), since then the "number of equations" in (1.12) is equal to the "number of unknowns". Indeed, the triple (x, y, s) lies in $\mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ and the equation values are in the same space when $x \bullet s \in \mathbb{E}$ (this is certainly the case when $x \bullet s = 0$).

When A is surjective, the central path is the (image of the) map $\mu \in \mathbb{R}_{++} \mapsto z(\mu) \in \mathbb{E}\times\mathbb{R}^m\times\mathbb{E}$, where $z = z(\mu)$ is the unique solution to the perturbed optimality conditions (see question 1.3)

$$\begin{cases} A^*(y) + s = c, \quad s \in K^s \\ A(x) = b, \quad x \in K^s \\ x \bullet s = \mu e. \end{cases} \tag{1.13}$$

When $\mathcal{F}^s \neq \varnothing$, the central path converges to a primal-dual solution, when $\mu \downarrow 0$. The goal of the *path-following* algorithms that we shall consider is to follow the central path to reach a solution asymptotically.

1.2 Two directions

1.2.1 Overview

Given a point $z = (x, y, s) \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$, the goal of one iteration of the primal-dual path-following interior-point (IP) algorithm we consider is to make a displacement d in $\mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ towards a central point $z(\mu)$, for some $\mu \geqslant 0$, that is "closer" to the solution than the current iterate z. This displacement $d = (d_x, d_y, d_s) \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ is a Newton-like step on the perturbed system (1.13) without its conic conditions $x \in K^s$ and $s \in K^s$ (these will be enforced at each iteration), namely

$$\begin{cases} A^*(y) + s = c \\ A(x) = b \\ x \bullet s = \mu e. \end{cases} \tag{1.14}$$

Computing the standard Newton direction on this system raises a serious difficulty, however. When z is not on the central path, which is usually the case, the point $x \bullet s$ may not be in $\mathbb{E}$, since the matrix components of $x \bullet s$ may not be symmetric matrices. In that case, one should work with more equations than unknowns in (1.14), which a situation we wish to avoid, since the linearized system may have no solution. To tell it otherwise, one would like $z + d$ to satisfy the system (1.14), which reads

$$\begin{cases} A^*(d_y) + d_s = r_{\mathrm{D}} \\ A(d_x) = r_{\mathrm{P}} \\ d_x \bullet s + x \bullet d_s = \mu e - (x \bullet s + d_x \bullet d_s), \end{cases} \tag{1.15}$$

where we have introduced the primal r_{P} and dual r_{D} residuals

$$r_{\mathrm{P}} := b - A(x) \in \mathbb{R}^m \qquad \text{and} \qquad r_{\mathrm{D}} := c - A^*(y) - s \in \mathbb{E}. \tag{1.16}$$

The last equation of (1.15) is nonlinear in d. Its linearization is obtained by dropping the term $d_x \bullet d_s$, resulting in the system

$$\begin{cases} A^*(d_y) + d_s = r_{\mathrm{D}} \\ A(d_x) = r_{\mathrm{P}} \\ d_x \bullet s + x \bullet d_s = \mu e - x \bullet s. \end{cases} \tag{1.17}$$

If d_s is in $\mathbb{E}$ by the first equation of (1.17) (since A^* takes its values in $\mathbb{E}$, and c and $s \in \mathbb{E}$), there are examples [26] showing that d_x might not be in $\mathbb{E}$ (because $x \bullet d_s \bullet s^{-1}$ in the third equation is not necessarily in $\mathbb{E}$, by lack of symmetry of its matrix components). For this reason a "symmetrization" operation must be introduced in addition to the linearization of the system, in order to keep $x + d_x$ in $\mathbb{E}$. This can be done using a large number of methods, leading to many different IP directions; not all of them have the desired properties. We present below two of the most often implemented techniques: the NT and HKM directions. These directions have the following features:

- they preserve the primal and dual displacements d_x and d_s in $\mathbb{E}$,
- they are uniquely defined at any strictly feasible iterate $z \in \mathcal{F}^s$,
- they are less computationally demanding than most other directions.

1.2.2 The NT direction

This section presents the *Nesterov-Todd* (NT) direction, called that way because of [35, 36, 47; 1997-1998]. This direction has the additional property of being scale invariant.

Derivation of the NT direction

The NT direction $d = (d_x, d_y, d_s) \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ at $z = (x, y, s) \in K^s \times \mathbb{R}^m \times K^s$ can be shortly presented as the result of three operations [8].

1. *Scaling.* A *weight* $w \in K^s$ is introduced, which is the unique element in K^s that satisfies the identity (see question 1.4)

$$w \bullet s \bullet w = x. \tag{1.18}$$

It is given by the formulas

$$w := x^{1/2} \bullet \left(x^{1/2} \bullet s \bullet x^{1/2}\right)^{-1/2} \bullet x^{1/2} = s^{-1/2} \bullet \left(s^{1/2} \bullet x \bullet s^{1/2}\right)^{1/2} \bullet s^{-1/2}. \tag{1.19}$$

Let us also introduce the scaled variable

$$v := w^{-1/2} \bullet x \bullet w^{-1/2} = w^{1/2} \bullet s \bullet w^{1/2} \in K^s. \tag{1.20}$$

Note that the vector component of w and v simply read

$$w^l = (x^l)^{1/2} \cdot (s^l)^{-1/2} \qquad \text{and} \qquad v^l = (x^l)^{1/2} \cdot (s^l)^{1/2}. \tag{1.21}$$

2. *Symmetrization.* If $z = (x, y, s)$ is the current iterate, one would like to find a displacement $d = (d_x, d_y, d_s)$ such that the third equation in (1.14) is satisfied at $z + d$, namely

$$(x + d_x) \bullet (s + d_s) = \mu e,$$

for some $\mu \geqslant 0$ to be defined. After a left-Hadamard-multiplication by $w^{-1/2}$ and a right-Hadamard-multiplication by $w^{1/2}$, this equation takes the equivalent form

$$(v + \tilde{d}_x) \bullet (v + \tilde{d}_s) = \mu e,$$

where

$$\tilde{d}_x = w^{-1/2} \bullet d_x \bullet w^{-1/2} \qquad \text{and} \qquad \tilde{d}_s = w^{1/2} \bullet d_s \bullet w^{1/2}.$$

We now *symmetrize* the left-hand side of this equation as follows

$$\frac{1}{2}\left[(v + \tilde{d}_x) \bullet (v + \tilde{d}_s) + (v + \tilde{d}_s) \bullet (v + \tilde{d}_x)\right] = \mu e.$$

The left-hand side is now an element in $\mathbb{E}$.

3. *Pseudo-linearization* (the term "pseudo" is used because the weight w, which depends on the current iterate, is not linearized). The linearization consists in dropping the nonlinear terms from the last identity, namely the Hadamard products $\tilde{d}_x \bullet \tilde{d}_s$ and $\tilde{d}_s \bullet \tilde{d}_x$. The identity then becomes

$$\frac{1}{2}\left[v \bullet (\tilde{d}_x + \tilde{d}_s) + (\tilde{d}_x + \tilde{d}_s) \bullet v\right] = \mu e - v \bullet v.$$

This is a *Lyapunov equation* in the unknown $\tilde{d}_x + \tilde{d}_s$. Since $v \in K^s$, it has a unique solution, which is

$$\tilde{d}_x + \tilde{d}_s = \mu v^{-1} - v.$$

Left and right-Hadamard-multiplication of the two sides of this last equation by $w^{1/2}$ yield the symmetric pseudo-linearization at z of the third equation in (1.14):

$$d_x + w \bullet d_s \bullet w = \mu s^{-1} - x.$$

Therefore, the NT direction is obtained by solving the following linear system in $d = (d_x, d_y, d_s) \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$:

$$\boxed{\begin{array}{c} A^*(d_y) + d_s = r_{\text{D}} \\ A(d_x) = r_{\text{P}} \\ d_x + w \bullet d_s \bullet w = \mu s^{-1} - x, \end{array}} \tag{1.22}$$

in which the value of μ still needs to be specified. It is a consequence of the calulation made in section 1.2.2, that, when A is surjective, the system (1.22) has a unique solution, whatever the right-hand side is (in particular, whatever $\mu > 0$ is).

Computation of the weight w

The weight w given by (1.19) intervenes in the computation of the NT direction and must be computed explicitly. The vector component w^l of w is easily computed by (1.21), so that we only consider a particular matrix component $w^{s,j}$ of w below.

Each matrix components $w^{s,j}$ of w can be computed using two Cholesky factorizations and one singular value decompositions (SVD) [47; § 4.1]. The Cholesky factorizations are those of the matrix components in $x \in K^s$ and $s \in K^s$:

$$\boxed{x^{s,j} = L_{x^{s,j}} L_{x^{s,j}}^{\mathsf{T}}} \qquad \text{and} \qquad \boxed{s^{s,j} = L_{s^{s,j}} L_{s^{s,j}}^{\mathsf{T}},}$$

where $L_{x^{s,j}}$ and $L_{s^{s,j}}$ are lower triangular nonsingular matrices of order n_j. The SVD is the one of $L_{s^{s,j}}^{\mathsf{T}} L_{x^{s,j}}$:

$$\boxed{L_{s^{s,j}}^{\mathsf{T}} L_{x^{s,j}} = U_j \Sigma_j V_j^{\mathsf{T}}.}$$

where U_j and V_j are orthogonal matrices and Σ_j is the diagonal matrix formed of the positive singular values of the nonsingular matrix $L_{s^{s,j}}^{\mathsf{T}} L_{x^{s,j}}$. Define

$$Q_j := L_{x^{s,j}}^{-1} (x^{s,j})^{1/2},$$

which is an orthogonal matrix. It needs not be computed. Now, there holds

$$\begin{aligned}
(x^{s,j})^{1/2} s^{s,j} (x^{s,j})^{1/2} &= (x^{s,j})^{1/2} \underbrace{L_{x^{s,j}}^{-\mathsf{T}} L_{x^{s,j}}^{\mathsf{T}}}_{I^{n_j}} \underbrace{L_{s^{s,j}} L_{s^{s,j}}^{\mathsf{T}}}_{s^{s,j}} \underbrace{L_{x^{s,j}} L_{x^{s,j}}^{-1}}_{I^{n_j}} (x^{s,j})^{1/2} \\
&= Q_j^{\mathsf{T}} \underbrace{L_{x^{s,j}}^{\mathsf{T}} L_{s^{s,j}}}_{V_j \Sigma_j U_j^{\mathsf{T}}} \underbrace{L_{s^{s,j}}^{\mathsf{T}} L_{x^{s,j}}}_{U_j \Sigma_j V_j^{\mathsf{T}}} Q_j \\
&= Q_j^{\mathsf{T}} V_j \Sigma_j^2 V_j^{\mathsf{T}} Q_j.
\end{aligned}$$

From this identity and the orthogonality of $Q_j^{\mathsf{T}} V_j$, it results

$$((x^{s,j})^{1/2} s^{s,j} (x^{s,j})^{1/2})^{-1/2} = Q_j^{\mathsf{T}} V_j \Sigma_j^{-1} V_j^{\mathsf{T}} Q_j.$$

From the first identity in (1.18), we get

$$w^{s,j} = \underbrace{(x^{s,j})^{1/2} Q_j^{\mathsf{T}}}_{L_{x^{s,j}}} V_j \Sigma_j^{-1} V_j^{\mathsf{T}} \underbrace{Q_j (x^{s,j})^{1/2}}_{L_{x^{s,j}}^{\mathsf{T}}} = L_{x^{s,j}} V_j \Sigma_j^{-1} V_j^{\mathsf{T}} L_{x^{s,j}}^{\mathsf{T}}.$$

Finally

$$\boxed{w^{s,j} = G_j G_j^{\mathsf{T}}, \qquad \text{where } G_j := L_{x^{s,j}} V_j \Sigma_j^{-1/2} = L_{s^{s,j}}^{-\mathsf{T}} U_j \Sigma_j^{1/2}.} \tag{1.23}$$

The matrices G_j and their transpose intervene below in (1.29), for which it is useful to introduce the vectors in $\mathbb{F}$:

$$g := (G_1, \ldots, G_s, (w^l)^{1/2}) \in \mathbb{F}, \tag{1.24}$$

allowing us to write (recall the definition of g^{T} in (1.3))

$$w = g \bullet g^{\mathsf{T}} \in \mathbb{E}. \tag{1.25}$$

It will also be useful to note that

$$g^{\mathsf{T}} \bullet s \bullet g = g^{-1} \bullet x \bullet g^{-\mathsf{T}} = \sigma \qquad \text{and} \qquad g^{-1} \bullet x \bullet s \bullet g = \sigma^2, \tag{1.26}$$

where the matrix components of $\sigma \in \mathbb{E}$ are the diagonal positive definite matrices Σ_j.

Computation of the NT direction

A standard way of solving the system (1.22) defining the NT direction d consists in eliminating first d_x using the second and third equation, to get

$$A(w \bullet d_s \bullet w) = -r_{\text{P}} + A(\mu s^{-1} - x).$$

Next, eliminating d_s using the first equation, we get

$$A(w \bullet A^*(d_y) \bullet w) = r_{\text{P}} + A(w \bullet r_{\text{D}} \bullet w) + A(x - \mu s^{-1}). \tag{1.27}$$

Once d_y is known as a solution to (1.27), d_s can be computed by the first equation in (1.22) and then d_x by the third equation in (1.22). So let us concentrate on ways of computing d_y by (1.27).

We assume that A is surjective, which implies that (1.27) is a linear system in d_y with a positive definite linear operator, hence having a unique solution. Here are two ways of computing d_y by (1.27).

- A first way of computing the solution d_y to (1.27) is to write this equation as the linear system in d_y:

$$\boxed{\mathcal{M}(d_y) = r_{\text{P}} + A(w \bullet r_{\text{D}} \bullet w) + A(x - \mu s^{-1}),} \tag{1.28}$$

  where $\mathcal{M} : d_y \in \mathbb{R}^m \mapsto A(w \bullet A^*(d_y) \bullet w) \in \mathbb{R}^m$ is a symmetric positive definite (hence nonsingular) linear map and to form the "compact" (when m is small) $m \times m$ coefficient matrix $\mathcal{M}$. For i and $j \in [1:m]$, there holds

$$\begin{aligned} \mathcal{M}_{kl} &= [A(w \bullet A^*(e^l) \bullet w)]_k \\ &= \langle a_k, w \bullet a_l \bullet w \rangle_{\mathbb{E}} \qquad [(1.8) \text{ and } (1.9)] \\ &= \langle a_k, g \bullet g^{\mathsf{T}} \bullet a_l \bullet g \bullet g^{\mathsf{T}} \rangle_{\mathbb{E}} \qquad [(1.25)] \\ &= \langle g^{\mathsf{T}} \bullet a_k \bullet g, g^{\mathsf{T}} \bullet a_l \bullet g \rangle_{\mathbb{E}}. \end{aligned} \tag{1.29}$$

  Then, one takes the Cholesky factorization $\mathcal{M} = L_{\mathcal{M}} L_{\mathcal{M}}^{\mathsf{T}}$ and solve (1.28). Forming $\mathcal{M}$ mainly requires $2mn^3$ operations (for getting the products $g^{\mathsf{T}} \bullet a_k \bullet g$, with $i \in [1:m]$). When n is large, this is the most computationally expensive part of the solver, limiting its use to problems with n not exceeding a few hundreds. Problem with higher dimension must be sparse to be considered by SCDO, and this one must be able to consider such sparse problems, which is not the case of the solver presented in these notes.

- Another way of computing d_y is to observe that (1.27) or ... ▲

Most implementations use the first approach.

1.2.3 The HKM direction ▲

Another efficient direction is the so-called HKM direction (obtained independently in [16, 21], see also [40]). It is also implemented in SDPT3, for instance. ...

1.3 Implementation

1.3.1 Data representation

We propose to follow the apparent data structure used in SeDuMi.

- An element $x \in \mathbb{E}$ is represented by a large *vector* of dimension $n_{\mathbb{F}}$. This large vector is formed of the matrices $x^{s,j} \in \mathcal{S}^{n_j}$ represented by a vector containing its $n_j \times n_j$ elements (not only its symmetric part having $n_j(n_j+1)/2$ elements) and the vector $x^l \in \mathbb{R}^{n_l}$.

  This representation has the advantage of making many computations easy. For instance the scalar product $\langle x, s\rangle$, for x and $s \in \mathbb{E}$, is obtained in Matlab by `x'*s` if `x` is the $n_{\mathbb{F}}$-vector representing x and `s` is the $n_{\mathbb{F}}$-vector representing s. This makes the design of the piece of software easier and also makes it more efficient computationally (since there is no loop).

  The above representation of the elements of $\mathbb{E}$ makes it necessary to often switch between the vector and matrix representations of matrices. This can be done with the CVX functions `mat` and `vec` (see below), probably inexpensively if the result is not safeguarded.

- The linear operator $A : \mathbb{E} \to \mathbb{R}^m$ introduced in the primal and dual problems in (1.7) is first represented by the vectors $a_i \in \mathbb{E}$, for $i \in [1:m]$, using (1.8). Then each vector $a_i \in \mathbb{E}$ is represented as an $n_{\mathbb{F}}$-vector, as described in the first point, whose transpose forms the ith row of a matrix `A`.

  With these representations of A and x, the value $A(x)$ is obtained in Matlab by `A*x` and the value of $A^*(y)$ is obtained by `A'*y`, which are elegant expressions, easy to program and debug, and computationally efficient.

1.3.2 Calling statement

We propose to give to the sdco solver, the following Matlab function structure

```
[x,y,s,info] = sdco (A,b,c,K,x0,y0,options)
```

where

- `A` is an $m \times n_{\mathbb{F}}$ representation of the linear operator A, as described in section 1.3.1;
- `b` is an m real vector, representing the right hand side of the equality constraint in (P);
- `c` is an $n_{\mathbb{F}}$-vector representing the vector $c \in \mathbb{E}$, which determines the linear objective in (P);
- `K` is a Matlab structure describing the structure of the vectors $x \in \mathbb{E}$:
  - `K.s` is the Matlab vector `[n1; n2; ...; ns]`, providing the orders `nj` $= n_j$, for $j \in [1:s]$, of the matrices $x^{s,j}$;
  - `K.l` is the length n_l of the vector x^l;

- `x0` is an $n_{\mathbb{F}}$-vector representing the vector $x_0 \in \mathbb{E}$, which is the initial guess of the primal solution x;
- `y0` is an m-vector, which is the initial guess of the dual solution $y \in \mathbb{R}^m$; s_0 is then deduced form y_0 by $s_0 := c - A^*(y_0)$;
- `options` is a structure of possible options, which will be described in other session; this one can be used to select some features of the algorithm, like the threshold to decide optimality;
- `x` is an $n_{\mathbb{F}}$-vector representing a vector $x \in \mathbb{E}$, which gives the primal solution x if any;
- `y` is an m vector, giving the dual solution y if any;
- `s` is an $n_{\mathbb{F}}$-vector, giving the dual solution s if any; it is linked to y by $s = c - A^*(y)$;
- `info` is a structure providing information of the run (success, failure, primal/dual infeasibility, primal/dual unboundedness, duality gap, *etc*).

1.3.3 Matlab functions

Here are a few useful Matlab functions.

- `L = chol(A,'lower')` computes the Cholesky factorization of $A = LL^{\mathsf{T}}$, where L is lower triangular (only the lower triangular part of A is used).
- `[U,S,V] = svd(A)` computes the singular value decomposition (SVD) of a real $n_{\mathbb{F}} \times n_{\mathbb{F}}$ matrix $A = USV^{\mathsf{T}}$, where U and V are orthogonal matrices and S is diagonal with nonnegative entries (the singular values of A).
- With CVX, `M = mat(v)` converts an n^2 vector `v` into an $n \times n$ matrix `M`, such that the columns of `M` are stacked below each other in `v`. This operation is frequently used, so that it is appropriate to write a personnal version, using `reshape`, if the function is not available. The reverse operation is realized by `vec`.
- `[Q,R] = qr(A,0)` computes the QR factorization of a real $m \times n$ matrix $A = QR$, where Q is orthogonal and R is upper triangular. The second zero argument is useful when $m > n$ (our case), in which case, only the first n column of Q and the first n rows of R are computed.
- With CVX, `v = vec(M)` converts an $m \times n$ matrix `M` into an mn vector `v`, stacking the columns of `M` below each other in `v`. This operation is frequently used, so that it is appropriate to write a personnal version, using `reshape`, if the function is not available. The reverse operation is realized by `mat`.

It is useful to introduce a Matlab function computing the Hadamard product of two vectors x and s in $\mathbb{E}$:

```
[hdot] = sdco_hdot (x,s,K)
```

Using this function and introducing a matrix B whose kth column is $g^{\mathsf{T}} \bullet a_k \bullet g$, the expression (1.29) of $\mathcal{M}$ simply reads $B^{\mathsf{T}}B$.

1.3.4 Recommendations

In order to get convergence of the algorithm of the next section, it is imperative to have a correctly computed direction. There are very few things that can certify the validity of the realized computation. We can recommend to verify that

- the vector w verifies (1.18),
- the matrix $\mathcal{M}$ defined by (1.29) is symmetric and positive definite,
- $\langle d_x, d_s \rangle = 0$ (this is only true when the current iterate (x, y, s) is feasible, meaning that $r_{\text{P}} = 0$ and $r_{\text{D}} = 0$),
- the residual of the system (1.22) is close to zero.

1.4 Test cases

1.4.1 Test case 1a: an easy SDO problem of dimension 3

Consider the semidefinite optimization (SDO) problem in $y \in \mathbb{R}^3$, written in standard dual form

$$\begin{cases} \sup\ e^{\mathsf{T}} y \\ \begin{pmatrix} 1-y_1 & -y_3 & -y_2 \\ -y_3 & 1-y_2 & 0 \\ -y_2 & 0 & 1-y_3 \end{pmatrix} \succcurlyeq 0, \end{cases}$$

where $e \in \mathbb{R}^3$ is the vector of all ones. It is easily verified that $x_0 = I^3$ and $(y_0, s_0) = (0, I^3)$ are strictly feasible for the primal and dual problems respectively, so that these problems have a solution without duality gap. There holds $b^{\mathsf{T}} y_0 = 0 < 3 = \langle c, x_0 \rangle$, so that $z_0 := (x_0, y_0, s_0)$ is not a primal-dual solution.

1.4.2 Test case 1b: a simple LO problem of dimension 2

Consider the following linear optimization (LO) problem in only 2 variables and a single affine constraint, which reads

$$\begin{cases} \inf\ x_1 + x_2 \\ x_1 + 2x_2 = 1 \\ x \geqslant 0. \end{cases}$$

The initial primal and dual variables are

$$x_0 := \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} \qquad \text{and} \qquad y_0 := 0,$$

which are strictly feasible points.

1.4.3 Test case 1c: two SDO problems and one LO problem in parallel

The goal of this problem is to test the capacity of the developed solver to deal simultaneously with two positive semidefinite matrices and one nonnegative vector as primal variables. The problem consists in solving in parallel two copies of the semidefinite optimization problems of test case 1a, each one starting from a different matrix, and the additional linear optimization (LO) problem of test case 1c.

Denote by c the gradient of the objective of the primal expression of test case 1a and by $A(x) = b$ its affine constraint. Then the primal expression of test case 1c is the following problem in $x = (x^{s,1}, x^{s,2}, x^l) \in \mathcal{S}^3 \times \mathcal{S}^3 \times \mathbb{R}^2$

$$\begin{cases} \inf \langle c, x^{s,1} \rangle + \langle c, x^{s,2} \rangle + (e^2)^\mathsf{T} x^l \\ A(x^{s,1}) = b \\ A(x^{s,2}) = b \\ x^l_1 + 2x^l_2 = 1 \\ x^{s,1} \succcurlyeq 0, \quad x^{s,2} \succcurlyeq 0, \quad x^l \geqslant 0, \end{cases}$$

The initial primal and dual variables are

$$x_0^{s,1} := I^3, \; x_0^{s,1} := \begin{pmatrix} 1 & 0 & 1/4 \\ 0 & 1/2 & 0 \\ 1/4 & 0 & 1 \end{pmatrix}, \; x_0^l := \begin{pmatrix} \frac{5-\sqrt{17}}{2} \\ \frac{-3+\sqrt{17}}{4} \end{pmatrix}, \text{ and } y_0 := (0_{\mathbb{R}^6}, -\tfrac{1+\sqrt{17}}{4}),$$

which are strictly feasible points.

Notes

There are various Matlab SDCO solvers. Having a look at SeDuMi [46, 44] is certainly a good idea, since sdco may be viewed as a reduced feature version of that solver. Another source of inspiration is SDPT3 [49].

Questions

When a question starts with a number within braces, the number gives an *indication* on the difficulty of the question, *with respect to the chapter* (some questions will be easier to answer when subsequent chapters will be known). This indication ranges from 1 to 5: 1 for easy-or-classical and short arguments, 2 for easy-or-classical arguments, 3 for arguments requiring specific knowledge, 4 for more difficult arguments, 5 for very difficult arguments or when a advanced computer program must be written to answer the question.

1.1. {1} *Adjoint of A.* Show that the adjoint of the linear map $A : \mathbb{E} \to \mathbb{R}^m$ defined by (1.8) is given by (1.9) when $\mathbb{R}^m$ is equipped with the Euclidean scalar product.

1.2. {2} *Towards a square optimality system.* Show the equivalence (1.11) when x and s are in the cone K.

1.3. {2} *Unique central point.* Show that (1.13) has a unique solution $z = (x, y, s) \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$, provided A is surjective.

1.4. {2} *Weight w.* Show that, when x and $s \in K^s$, the identity (1.18) has a unique solution $w \in K^s$, which is given by (1.19).

1.5. {5} *On test cases* 1*a,* 1*b, and* 1*c.* Consider test cases 1*a*, 1*b*, and 1*c*. Let $z_0 := (x_0, y_0, s_0)$ and d_0 be the NT direction computed at z_0 for some μ.

1) {1} Is z_0 on the central path?
2) {5} Compute the NT direction for $\mu = 0$ by a computer program.
3) {1} Is $z_1 = z_0 + d_0$ a primal-dual solution to the problem if $\mu = 0$?

2 Moving along the central path

2.1 Algorithmic concepts

In this chapter, we force the iterates $z = (x, y, s)$ to be strictly feasible, meaning that they will be in $\mathcal{F}^s$. For this to work, it will be necessary to have a first iterate that is strictly feasible, and even sufficiently close to the central path.

2.1.1 The closest central point

In a path-following algorithm, the first question that arises deals with the determination of the central point $z(\mu)$ to which the current iterate $z := (x, y, s)$ is the closest. Indeed, in such an algorithm, the next iterate aims a point on the central path $\mathscr{C}$ and gets close to it by a Newton step, provided this central point is not too far from z. If we want the iteration to make a progress towards the solution it is necessary that this aimed central point be closer to the solution than the central point that is the closest to z, hence the necessity to determine the latter.

Finding the closest central point is not a well defined concept, actually, since the central path is not a convex set, making the projection on $\mathscr{C}$ not well defined. To overcome the difficulty, we look instead to the closest central point in a transformed space in which the central path becomes a half-line. The transformation is

$$\tau : z = (x, y, s) \in \mathcal{F} \mapsto x \bullet s \in \mathbb{F}.$$

This transformation is certainly not a bijection, but it turns out to be useful, which is enough to make it appropriate. Since $\mathscr{C} := \{z \in \mathcal{F} : x \bullet s \in \mathbb{R}_{++}e\}$, it follows that $\tau(\mathscr{C}) = \mathbb{R}_{++}e$. Then, by the convexity and closedness of $\overline{\tau(\mathscr{C})}$, it makes sense to look for the point in $\overline{\tau(\mathscr{C})}$ that is the closest to $\tau(z)$. That problem simply reads

$$\min_{\mu \geqslant 0} \|x \bullet s - \mu e\|_{\mathbb{F}}^2. \tag{2.1}$$

It has a unique solution (see question 2.1), which is

$$\overline{\mu}(z) := \frac{\langle x, s \rangle_{\mathbb{E}}}{n_c}, \tag{2.2a}$$

where n_c is given by (1.6). The *closest central point* is therefore "considered" to be the one on the central path with the parameter $\mu = \overline{\mu}(z)$. This is not irrelevant, since when $z \in \mathscr{C}$, it is clear that the closest central point to z in the preceding sense is z itself; reassuring.

2.1.2 Neighborhood of the central path

Experimentation has convinced the algorithmicians that it is not good to let the iterates going too far from the central path $\mathscr{C}$ or, more correctly, too close to the boundary of the feasible set $\mathcal{F}$ (it is not the same thing actually, since the "central" path may be close to the boundary of $\mathcal{F}$ [9]). When z is close to that boundary, the stepsize along the Newton direction may be very small, so much as to prevent any progress to the solution: the iterates get stuck on the boundary of $\mathcal{F}$. For this reason, the iterates are maintained in some neighborhood of the central path, often the one that we now define. Recall that the central path is an analytic concept (not a geometric one, depending only on the feasible set), so that this is a second best solution.

Observe first that for a point $z \in \mathcal{F}^s$:

$$x \bullet s = \mu e \iff v \bullet v = \mu e \iff \mu^{1/2} v^{-1} = \mu^{-1/2} v, \tag{2.3}$$

where $v := w^{-1/2} \bullet x \bullet w^{-1/2} = w^{1/2} \bullet s \bullet w^{1/2} \in K^s$ was already defined in (1.20) and $w \in K^s$ was already defined in (1.19) as the unique solution to (1.18). Then one defines a *proximity measure* of the central point $z(\mu)$ by ([18] and [8; § 7.1]):

$$\delta_\mu : z \in \mathcal{F}^s \mapsto \delta_\mu(z) := \frac{1}{2} \left\| \mu^{1/2} v^{-1} - \mu^{-1/2} v \right\|_{\mathbb{E}} \in \mathbb{R}. \tag{2.4}$$

Observe now that for a point $z \in \mathcal{F}^s$:

$$z \in \mathscr{C} \iff x \bullet s = \overline{\mu}(z) e. \tag{2.5}$$

Therefore, it makes sense to define a *proximity measure* of the central path by

$$\delta : z \in \mathcal{F}^s \mapsto \delta(z) := \delta_{\overline{\mu}(z)}(z). \tag{2.6}$$

Accordingly, for $\theta \in [0, 1)$, one introduces the following neighborhood of the central path

$$\mathcal{V}(\theta) := \{ z \in \mathcal{F}^s : \delta(z) \leqslant \theta \}.$$

Using (1.4b) and $w = g \bullet g^{\mathsf{T}}$ from (1.25), we deduce that

$$\delta(z) = \frac{1}{2} \left\| g^{\mathsf{T}} \bullet \left(\overline{\mu}(z)^{1/2} x^{-1} - \overline{\mu}(z)^{-1/2} s \right) \bullet g \right\|_{\mathbb{E}}, \tag{2.7}$$

so that the distance can be computed efficiently.

2.2 Two algorithms

2.2.1 A predictor-corrector algorithm

The predictor-corrector algorithms are the most often implemented interior-point methods. We essentially give the description of the *Mizuno-Todd-Ye predictor-corrector method* [30, 8], whose iteration is composed of two phases: a corrector phase, followed by a predictor phase (in reverse order, in a way).

- The *corrector phase* starts with an iterate $z \in \mathcal{V}(1/3)$. Its goal is to compute an intermediate iterate z', close enough to the central path, so that the predictor step that follows will compute a subsequent iterate z_+ again in the neighborhood $\mathcal{V}(1/3)$. This goal is achieved by making a displacement along the NT direction with $\mu = \overline{\mu}(z)$, say d^c (known as the *centering direction*), and a unit stepsize. The intermediate iterate is then
$$z' := z + d^c.$$

- The *predictor phase*, which follows the corrector step, starts with $z' \in \mathcal{F}^s$ and consists in making a displacement along the NT direction with $\mu = 0$, say d^a (known as the *affine-scaling direction*). The stepsize $\alpha > 0$ along that direction is given by the formula (no need of linesearch, provided a condition given below is fulfilled)
$$\alpha = \frac{2}{1 + [1 + 13\|\tilde{d}_x^a \bullet \tilde{d}_s^a + \tilde{d}_s^a \bullet \tilde{d}_x^a\|_{\mathbb{E}}/(2\overline{\mu}(z'))]^{1/2}}, \tag{2.8a}$$
where
$$\tilde{d}_x^a = w^{-1/2} \bullet d_x^a \bullet w^{-1/2} \qquad \text{and} \qquad \tilde{d}_s^a = w^{1/2} \bullet d_s^a \bullet w^{1/2}.$$
Note that the matrix w is computed at z', not at z (hence by (1.19) with (x, s) changed into (x', s')). It can be shown that
$$\|\tilde{d}_x^a \bullet \tilde{d}_s^a + \tilde{d}_s^a \bullet \tilde{d}_x^a\|_{\mathbb{E}} = \|g^{-1} \bullet d_x^a \bullet d_s^a \bullet g + g^{\mathsf{T}} \bullet d_s^a \bullet d_x^a \bullet g^{-\mathsf{T}}\|_{\mathbb{E}}, \tag{2.8b}$$
where g and g^{T} are defined by (1.24), so that the computation of this norm can be done efficiently.

  Note that for the vector parts of d_x^a and d_s^a, there holds
$$\tilde{d}_{x^l}^a \cdot \tilde{d}_{s^l}^a = \tilde{d}_{s^l}^a \cdot \tilde{d}_{x^l}^a = d_{x^l}^a \cdot d_{s^l}^a,$$
so that the weight w does not intervene.

  The stepsize (2.8) is small enough to ensure that the next iterate
$$z_+ := z' + \alpha d^a$$
is in the neighborhood $\mathcal{V}(1/3)$ and large enough to ensure the polynomiality of the algorithm (see below).

Algorithm 2.2.1 (predictor-corrector) A tolerance $\varepsilon > 0$ on $\overline{\mu}(z)$ is given to determine when stopping the iterations. The iteration from z to z_+ starts at some $z \in \mathcal{V}(1/3)$.

1. *Stopping test.* If $\overline{\mu}(z) \leqslant \varepsilon$, stop.

2. *Corrector phase.* Compute the NT direction at z with $\mu = \overline{\mu}(z)$, denote it d^c (centering direction). Take as next iterate

$$z' := z + d^c.$$

3. *Predictor phase.* Compute the NT direction at z' with $\mu = 0$, denote it d^a (affine-scaling direction), and the stepsize α given by (2.8). Set

$$z_+ := z' + \alpha d^a.$$

It can be shown [8] that α given by (2.8) satisfies

$$\alpha \geqslant \frac{2}{1 + [1 + 13n_c/2]^{1/2}},$$

which implies the polynomiality of the algorithm, with the iteration complexity

$$K := \left\lceil \left(1 + \sqrt{1 + \frac{13n_c}{2}}\right) \log \frac{\overline{\mu}(z_0)}{\varepsilon} \right\rceil. \tag{2.9}$$

This means that if the algorithm starts with $z_0 \in \mathcal{V}(1/3)$, it ends up with z_K satisfying $\overline{\mu}(z_K) \leqslant \varepsilon$, where the number K of iterations is given by (2.9).

2.2.2 A large step algorithm ▲

2.3 Implementation

Here are a few recommendations.

- It is important to check numerically that the following formula is verified after each step αd from a point $z \in \mathcal{F}^s$, where $\alpha > 0$ and d is the NT direction computed for a certain μ (see question 2.2):

$$\overline{\mu}(z + \alpha d) = (1 - \alpha)\overline{\mu}(z) + \alpha\mu. \tag{2.10}$$

  In particular, $\overline{\mu}(z') = \overline{\mu}(z)$ and $\overline{\mu}(z_+) = (1 - \alpha)\overline{\mu}(z)$, showing that the progress to the solution is made by the prediction step. Observe indeed that z_* is a solution if and only if $z_* \in \mathcal{F}^s$ and $\overline{\mu}(z_*) = 0$, so that the goal of the algorithm is to force the decrease of $\overline{\mu}(z)$.

  Note finally that if $\mu = \overline{\mu}(z)$, then (2.10) shows that $\overline{\mu}(z + \alpha d)$ is the constant $\overline{\mu}(z)$, whatever $\alpha \geqslant 0$ is.

- It is important to check that the iterates z remains in the neighborhood $\mathcal{V}(1/3)$. Even if the distance to the central path does not intervene in the predictor-corrector algorithm, in the development phase, it is useful to have a function that can compute the distance to the central path and to verify that this one remains below 1/3.

- The value of α given by (2.8) may be very close to one, or even equal to one, due to rounding errors. This usually yields vectors x and/or s that are no longer in the cone K^s. In this case, the algorithm gets stuck in a Cholesky factorization. Limiting this computed value of α to

  $$\texttt{options.max\_stepsize} = 0.999$$

  or so, may help the algorithm to go further and compute a more precise solution. Of course, taking a too small value for this parameter will prevent the algorithm from converging rapidly.

2.4 Test case

2.4.1 Test case 2: minimum matrix norm

We consider the *minimum matrix norm* problem which reads

$$\min_{v \in \mathbb{R}^p} \|\mathcal{B}(v)\|_2 , \tag{2.11}$$

where $\mathcal{B}(v) := B_0 + \sum_{i=1}^p v_i B_i$, the matrices $B_i \in \mathbb{R}^{q \times r}$ and $\|\cdot\|_2$ denotes the ℓ_2-norm. We assume that the B_i's are all different, which is a quite reasonable assumption. This problem is convex, but nonsmooth. It can be expressed as an SDO problem as follows.

Problem (2.11) can be rewritten $\min_{t,v}\{t : \|\mathcal{B}(v)\|_2^2 \leqslant t^2\}$. Using the Schur complement technique, it can be verified that this last problem can also be written as the following SDO problem in dual form [50]:

$$\begin{cases} \max_{(t,v) \in \mathbb{R} \times \mathbb{R}^p} -t \\ \begin{pmatrix} tI^q & \mathcal{B}(v) \\ \mathcal{B}(v)^\mathsf{T} & tI^r \end{pmatrix} \succcurlyeq 0. \end{cases} \tag{2.12}$$

This problem has dimension $n = q + r$ and $m = p + 1$. One can also easily find a strictly feasible primal-dual point (question 2.4). What happens with the SDCO solver when two B_i's are identical? Can you explain its behavior?

Notes

The proximity measure δ defined by (2.6) was introduced by Jiang [18], extending to semidefinite optimization a similar measure introduced by Jansen et al. [17] in linear optimization.

Questions

2.1. {1} *Formula of* $\overline{\mu}(z)$. Show that the solution to problem (2.1) is given by (2.2).

2.2. {2} *Evolution of* $\overline{\mu}(z)$ *along the NT direction.* Show formula (2.10).

2.3. {5} *On test cases* 1*a,* 1*b, and* 1*c.* Using your solver, find a solution to test cases 1*a*, 1*b*, and 1*c*.

2.4. {5} *On test case* 2.

1) {1} Show that, in standard notation, $z_0 = (x_0, y_0, s_0)$, with

$$x_0 = \tfrac{1}{q+r}\, I^{q+r}, \quad y_0 = (t_0, v_0) = (\|B_0\|_2 + 1, 0), \quad \text{and} \quad s_0 = \begin{pmatrix} t_0 I^q & B_0 \\ B_0^{\mathsf{T}} & t_0 I^r \end{pmatrix}$$

is a strictly feasible primal-dual point for problem (2.12).

2) {5} Using your solver, find a solution to problem (2.11), equivalent to problem (2.12), with, say $p = 2$, $q = 2$, $r = 2$, and random data for the matrices $B_i \in \mathbb{R}^{q \times r}$, with $i \in [0\,{:}\,2]$.

3 Finding an appropriate starting point

The predictor-corrector algorithm that was presented in section 2.2.1 assumes that the first iterate is strictly feasible and is in some neighborhood $\mathcal{V}(\theta)$ of the central path. In this section, we consider a method to get such a point, provided a strictly feasible primal-dual point $z_0 = (x_0, y_0, s_0)$ is known. The idea is to minimize a primal-dual merit function that has a central point as unique minimizer, starting the minimization process at z_0. The NT direction can be used to force the decrease of that merit function.

3.1 Getting a feasible point close to the central path

We use on $\mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ the scalar product inherited from those on $\mathbb{E}$ and $\mathbb{R}^m$, which, for $z = (x, y, s)$ and $z' = (x', y', s')$, is denoted by and takes the value

$$\langle z, z' \rangle = \langle x, x' \rangle_{\mathbb{E}} + y^{\mathsf{T}} y' + \langle s, s' \rangle_{\mathbb{E}}.$$

The associated norm is denoted by $\| \cdot \|$.

3.1.1 A primal-dual merit function

We investigate the simple idea that consists in minimizing *approximately* on $\mathcal{F}^s$ a function

$$\varphi_\mu : \mathbb{E} \times \mathbb{R}^m \times \mathbb{E} \to \mathbb{R} \cup \{+\infty\},$$

whose unique minimizer in $\mathcal{F}^s$ is the central point $z(\mu)$, that is the unique solution to the perturbed optimality system (1.13). There are several possibilities for φ_μ, including a function defined only on $\mathbb{E}$ (see the indication given for solving exercise 1.3). We prefer, however, using a function defined on $\mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ that can be minimized using the NT direction, which has already been implemented.

With that objective in mind, we follow [8; § 7.1] by defining the function φ_μ at $z \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ by

$$\varphi_\mu(z) := \langle x, s \rangle_{\mathbb{E}} + \mu \, \psi(x \bullet s), \tag{3.1}$$

where $\mu > 0$ is a fixed parameter and $\psi : \mathbb{F} \to \mathbb{R} \cup \{+\infty\}$ is defined at $v \in \mathbb{F}$ by

$$\psi(v) = \sum_{j=1}^{s} \operatorname{ld} v^{s,j} - \sum_{i=1}^{n_l} \log v_i^l.$$

This expression makes use of the *self-concordant barrier* [34, 41] $\mathrm{ld} : \mathbb{R}^{n\times n} \to \mathbb{R}\cup\{+\infty\}$ of the cone of positive definite matrices that are not necessarily symmetric, defined at $M \in \mathbb{R}^{n\times n}$ by

$$\mathrm{ld}(M) = \begin{cases} -\log\det M & \text{if } \det M > 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{3.2}$$

It can be shown that, when $\mathcal{F}^s \neq \varnothing$ and A is surjective, φ_μ is strictly convex on $\mathcal{F}^s$ and has the central point $z(\mu)$ as unique minimizer on $\mathcal{F}^s$ (see exercise 3.1). This has for consequence that generating iterates in $\mathcal{F}^s$ that minimize φ_μ will eventually provide a point in $\mathcal{V}(\theta)$ for some prescribed $\theta > 0$.

The gradient and Hessian of φ_μ at $z \in \mathbb{E}\times\mathbb{R}^m\times\mathbb{E}$ can be computed by

$$\nabla\varphi_\mu(z) = \begin{pmatrix} s - \mu x^{-1} \\ 0 \\ x - \mu s^{-1} \end{pmatrix} \quad \text{and} \quad \nabla^2\varphi_\mu(z)\,d = \begin{pmatrix} \mu\, x^{-1}\bullet d_x \bullet x^{-1} + d_s \\ 0 \\ d_x + \mu\, s^{-1}\bullet d_x \bullet s^{-1} \end{pmatrix}, \tag{3.3}$$

where $d \in \mathbb{E}\times\mathbb{R}^m\times\mathbb{E}$.

3.1.2 Use of the NT direction

Now, the question arises to see whether φ_μ can be minimized using the NT direction. The next result shows that the NT direction d defined at $z \in \mathcal{F}^s$, with the parameter $\mu > 0$, is a descent direction of φ_μ. It also provides with (3.4) the value of the directional derivative of φ_μ along the NT direction.

Proposition 3.1.1 (decrease of the primal-dual barrier along the NT direction) *Let d be the NT direction at a point $z \in \mathcal{F}^s$ for some $\mu > 0$. Then*

$$\langle\nabla\varphi_\mu(z), d\rangle = -4\,\mu\,\delta_\mu(z)^2, \tag{3.4}$$

where δ_μ is defined by (2.4). Furthermore, the following properties are equivalent:

(i) $z \neq z(\mu)$,
(ii) $d \neq 0$,
(iii) $\langle\nabla\varphi_\mu(z), d\rangle < 0$.

Proof. Using point 1 of exercise 3.1, one gets

$$\langle\nabla\varphi_\mu(z), d\rangle = \langle s - \mu x^{-1}, d_x\rangle_{\mathbb{E}} + \langle x - \mu s^{-1}, d_s\rangle_{\mathbb{E}}.$$

Using the scaling vector w defined by (1.19) and the scaled directions $\tilde{d}_x := w^{-1/2}\bullet d_x \bullet w^{-1/2}$ and $\tilde{d}_s := w^{1/2}\bullet d_s \bullet w^{1/2}$, one gets

$$\langle\nabla\varphi_\mu(z), d\rangle = \langle s - \mu x^{-1}, w^{1/2}\bullet\tilde{d}_x\bullet w^{1/2}\rangle_{\mathbb{E}} + \langle x - \mu s^{-1}, w^{-1/2}\bullet\tilde{d}_s\bullet w^{-1/2}\rangle_{\mathbb{E}}.$$

It is known, from (1.20), that $v := w^{-1/2} \bullet x \bullet w^{-1/2} = w^{1/2} \bullet s \bullet w^{1/2}$ and, from $(1.22)_3$, that $\tilde{d}_x + \tilde{d}_s = \mu v^{-1} - v$, so that

$$\begin{aligned} \langle \nabla \varphi_\mu(z), d \rangle &= \langle v - \mu v^{-1}, \tilde{d}_x + \tilde{d}_s \rangle_{\mathbb{E}} \\ &= -\|v - \mu v^{-1}\|_{\mathbb{E}}^2 \\ &= -\mu \|\mu^{-1/2} v - \mu^{1/2} v^{-1}\|_{\mathbb{E}}^2 \\ &= -4\mu \delta_\mu(z)^2 \qquad [(2.4)], \end{aligned}$$

which is (3.4). The next equivalences follow easily. □

Note that if μ is set to $\overline{\mu}(z)$, if d is the NT direction computed for that μ, and if a stepsize $\alpha > 0$ is taken along d to force the decrease of φ_μ, then $\overline{\mu}(z + \alpha d) = \overline{\mu}(z)$ by (2.10), so that an NT direction at $z + \alpha d$ with $\mu = \overline{\mu}(z + \alpha d)$ will also be a descent direction of the *same* merit function φ_μ. In other words, if one sets $\mu = \overline{\mu}(z)$ at each iterate z, the NT direction is a descent direction for the same merit function φ_μ.

Proposition 3.1.2 (distance to the central path after a NT step) *Let d be the NT direction at a point $z \in \mathcal{F}^s$ for $\mu := \overline{\mu}(z)$. Then the distance δ to the central path satisfies the inequality*

$$\delta(z + d) \leqslant \frac{\delta(z)^2}{\sqrt{2(1 - \delta(z)^2)}}. \tag{3.5}$$

Proof. Adapt the proof of [8; lemma 7.4] to the present framework. □

3.1.3 An algorithm

A consequence of the fact that $z(\mu)$ uniquely minimizes φ_μ (exercise 3.1) and proposition 3.1.1 is that, to get closer to the central path, it makes sense to minimize the function φ_μ, with $\mu = \overline{\mu}(z)$, along the NT direction, using a linesearch technique. Furthermore, from (3.5) in proposition 3.1.2,

$$\delta(z) \leqslant \sqrt{\frac{2\tau^2}{1 + 2\tau^2}} \qquad \Longrightarrow \qquad \delta(z + d) \leqslant \tau \delta(z),$$

so that, choosing $\tau < 1$, the distance decreases linearly without the need to make linesearch, as soon as $\delta(z) \leqslant \sqrt{2\tau^2/(1 + 2\tau^2)}$, which is $< \sqrt{2/3}$. Inequality (3.5) also tells us that the distance $\delta(z)$ converges to zero quadratically.

This discussion leads to the following algorithm, which computes a point in the neighborhood $\mathcal{V}(\theta)$ of the central path, starting from some given point $z \in \mathcal{F}^s$.

Algorithm 3.1.3 (getting a point in $\mathcal{V}(\theta)$) The algorithm uses three parameters: the scalar $\theta > 0$ specifying the neighborhood to reach, a desired linear speed of convergence $\tau < 1$ close to 1, and a linesearch parameter $\omega > 0$ close to zero. Typically $\theta = 1/3$, $\tau \simeq 0.99$, and $\omega \simeq 10^{-4}$. Let $z \in \mathcal{F}^s$.

Repeat until $z \in \mathcal{V}(\theta)$.

1. *NT direction.* Compute the NT direction at z with $\mu := \overline{\mu}(z)$, denote it d^c.
2. *Stepsize* α. If $\delta(z) \leqslant \sqrt{2\tau^2/(1+2\tau^2)}$, take $\alpha = 1$. Otherwise, find the largest α of the form 2^{-i} with $i \in \mathbb{N}$ such that
$$\varphi_\mu(z + \alpha d^c) \leqslant \varphi_\mu(z) - 4\,\omega\,\mu\,\alpha\,\delta(z)^2. \tag{3.6}$$
3. *Next iterate.* Set the next iterate z to $z + \alpha d^c$.

The previous algorithm hides a serious difficulty, which may occur during the linesearch in (3.6), which evaluates the quality of the trial stepsize α. By definition, the function $\varphi_\mu(z) = +\infty$ when $z \notin K^s$. This fact may not be seen by simply evaluating $x^{s,j} s^{s,j}$ (resp. $x_i^l s_i^l$), which may be positive definite (resp. positive) even when $x^{s,j} \not\succ 0$ or $s^{s,j} \not\succ 0$ (resp. when both $x_i^l < 0$ and $s_i^l < 0$). Therefore, the correct evaluation of φ_μ requires to check that, for all $j \in [1:s]$ and all $i \in [1:n_l]$: $x^{s,j} \succ 0$ and $s^{s,j} \succ 0$ (either by the computation of the minimum eigenvalues of $x^{s,j}$ and $s^{s,j}$ or by doing their Cholesky factorizations) and that $x_i^l > 0$ and $s_i^l > 0$.

Notes

Proposition 3.1.1 is taken from [8; p. 117] and proposition 3.1.2 from [8; lemma 7.4].

Questions

3.1. {3} *Computing a central point by primal-dual minimization.* Suppose that $\mathcal{F}^s \neq \varnothing$ and that A is surjective. Let φ_μ be given by (3.1) for some $\mu > 0$ and ld be given by (3.2). Show that

1) {3} (3.3) holds for z and $d \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$,
2) {3} φ_μ is strictly convex on $\mathcal{F}^s$,
3) {3} the problem
$$\inf_{z \in \mathcal{F}^s} \ \varphi_\mu(z) \tag{3.7}$$
has a unique solution, which is the central point $z(\mu)$.

3.2. {5} *Approaching the central path.* Implement algorithm 3.1.3 and try it on the test case 2 defined in section 2.4.1, with various values of the parameters p, q, r, and

random data for the matrices $B_i \in \mathbb{R}^{q\times r}$, with $i \in [0\,{:}\,p]$. The data of this test case can be recovered by entering

```
[A,b,c,K,x0,y0] = testcase_2 (p,q,r);
```

where `p`, `q`, `r` are the parameters (p, q, r).

4 Starting from an infeasible point

In this chapter, we consider algorithms that do not generate iterates in the feasible set $\mathcal{F}$. It is interesting to have such algorithms since in practice it is usually not obvious to have a strating point in $\mathcal{F}$. Two approaches are considered.

In the big M approach of section 4.1, the starting iterate $z_0 = (x_0, y_0, s_0)$ is supposed to satisfy the affine constraints $A(x_0) = b$ and $A^*(y_0) + s_0 = c$ but x_0 and s_0 are not supposed to belong to the cone K^s. Such an infeasible starting point is easy to find by linear algebra techniques. One then constructs from (P) and (D), other primal-dual pairs of problems, for which z_0 is strictly feasible. These pairs have parameters that must be driven to zero. Hence they are solved repetitively by the robust techniques used in the previous chapters until the parameters vanish.

The technique explored in section 4.2 discards the use of a neighborhood of the central path to control the iterates. The main reason is that this control is not easy to implement. For example, the stepsize in (2.8), which ensures the iterates to stay in the prescribed neighborhood, is conservative, meaning that one could take a larger stepsize (hence converge more rapidly) and stay in the neighborhood. The presented algorithm has a predictor-corrector nature and has the main goal of being practically efficient (while the theoretical convergence issues are ignored). Algorithms implemented in most available codes follow a similar approach. They use heuristics, but are faster than approaches whose polynomial complexity is well documented.

4.1 The big M approach

In all this section, we assume that an initial $x_0 \in \mathbb{E}$ is known, which satisfies the equality constraint

$$A(x_0) = b. \tag{4.1}$$

Having a pair (y_0, s_0) satisfying the affine constraint of the dual problem makes no difficulty, since y_0 can be taken arbitrarily and s_0 can be set to $c - A^*(y_0)$. A way of getting a point $x_0 \in \mathbb{E}$ satisfying (4.1) is described in section 4.1.1.

We then discuss the so-called *big M methods* for solving the SDCO problem from a point $z_0 = (x_0, y_0, s_0)$ satisfying (4.1) and $A^*(y_0) + s_0 = c$ but not the conditions $x_0 \in K^s$ and $s_0 \in K^s$ that would make z_0 strictly feasible and for which the algorithm of chapter 3 could be used. The case when $x_0 \in K^s$ and $s_0 \notin K^s$ is considered in section 4.1.2, the case when $x_0 \notin K^s$ and $s_0 \in K^s$ is considered in section 4.1.3, and the case when $x_0 \notin K^s$ and $s_0 \notin K^s$ is considered in section 4.1.4.

The contents of this section is adapted from [50; § 6].

4.1.1 Getting primal affine feasibility

We assume below that $A : \mathbb{E} \to \mathbb{R}^m$ has the following natural extension from $\mathbb{E}$ to $\mathbb{F}$, denoted

$$A_{\mathbb{F}} : \mathbb{F} \to \mathbb{R}^m.$$

This one is constructed as follows. First, we take the expression (1.8) of A, with vectors $a_i \in \mathbb{E}$, for $i \in [1:m]$. Then, for $x \in \mathbb{F}$, each component $[A_{\mathbb{F}}(x)]_i$ of $A_{\mathbb{F}}(x)$ is defined to be $[A_{\mathbb{F}}(x)]_i = \langle a_i, x\rangle_{\mathbb{F}}$. Obviously, the restriction of $A_{\mathbb{F}}$ to $\mathbb{E}$ is A, which is what we mean by "extension".

Now, if $x_0 \in \mathbb{E}$ satisfying (4.1) is not given by the user of the solver, this one should compute such a point or claim that (4.1) has no solution $x_0 \in \mathbb{E}$. The solver can proceed as follows.

- The first step is to check whether $b \in \mathcal{R}(A_{\mathbb{F}})$. This can be done using linear algebra techniques. If this is not the case, the primal problem is infeasible and there is no reason to go further.
- Otherwise, we claim that (4.1) has a solution (in $\mathbb{E}$). This one can be obtained by first computing $x \in \mathbb{F}$ satisfying the affine constraint

$$A_{\mathbb{F}}(x) = b \tag{4.2a}$$

and then symmetrizing it by taking

$$x_0^l := x^l \qquad \text{and} \qquad x_0^{s,j} = \frac{1}{2}\left(x^{s,j} + (x^{s,j})^{\mathsf{T}}\right), \quad \forall\, j \in [1:s]. \tag{4.2b}$$

It is asked in question **??** to verify that $x_0 \in \mathbb{E}$ obtained in that manner is indeed a solution to (4.1).

4.1.2 Starting from a strictly feasible primal point

We assume in this section that a strictly feasible primal point $x_0 \in \mathcal{F}_P^s$ is known, meaning that

$$A(x_0) = b \qquad \text{and} \qquad x_0 \in K^s, \tag{4.3}$$

but that $s_0 \notin K^s$.

Generally speaking, the *big M method* to solve the SDCO problem when a strictly feasible point is not available, using the technique of chapter 3, consists in introducing a modified SDCO problem that has the same solutions as the original problem, provided a parameter M in the modified SDCO problem is "large enough" (hence the words "big M"), and for which it is easy to find a strictly feasible point. The technique has therefore an effect that is similar to an exact penalty method.

In case (4.3) holds, the modified SDCO problem is obtained from a modification of the dual problem, for which a strictly feasible point is not known. The dual problem becomes the problem in $(y, \eta, s) \in \mathbb{R}^m \times \mathbb{R} \times \mathbb{E}$ (the modified parts are in red):

$$\begin{cases} \sup\ b^{\mathsf{T}}y - M_1\eta \\ A^*(y) - \eta e + s = c \\ s \in K \\ \eta \geqslant 0. \end{cases} \tag{4.4}$$

The new parameter η is used to realize feasibility easily (see below). This nonnegative parameter is then forced to be as small as possible by taking M_1 "sufficiently" large (the rational is that, when $\eta = 0$, the original dual problem (D) is recovered). We claim that

- problem (4.4) can be cast as a standard dual SDCO problem, for which a strictly feasible point is easy to determine,
- its primal-dual solutions are the same as the original problem, provided M_1 is chosen "sufficiently large" and the dual problem is feasible.

Let us see this.

One can cast problem (4.4) as the standard dual SDCO problem with $\tilde{\mathbb{E}} := \mathbb{E} \times \mathbb{R}$, $\tilde{K} = K \times \mathbb{R}_+$, and the dual variable $(\tilde{y}, \tilde{s}) \in \mathbb{R}^{m+1} \times \tilde{\mathbb{E}}$ that solves

$$\begin{cases} \sup\ \tilde{b}^{\mathsf{T}}\tilde{y} \\ \tilde{A}^*(\tilde{y}) + \tilde{s} = \tilde{c} \\ \tilde{s} \in \tilde{K}. \end{cases} \tag{4.5}$$

This problem is identical to problem (4.4) provided the vectors $\tilde{b}$ and $\tilde{y} \in \mathbb{R}^{m+1}$, the adjoint linear map $\tilde{A}^* : \mathbb{R}^{m+1} \to \tilde{\mathbb{E}}$, and the vectors $\tilde{s}$ and $\tilde{c} \in \tilde{\mathbb{E}}$ are defined by

$$\tilde{b} = \begin{pmatrix} b \\ -M_1 \end{pmatrix}, \quad \tilde{y} = \begin{pmatrix} y \\ \eta \end{pmatrix}, \quad \tilde{A}^*(\tilde{y}) = \begin{pmatrix} A^*(y) - \eta e \\ -\eta \end{pmatrix}, \quad \tilde{s} = \begin{pmatrix} s \\ \sigma \end{pmatrix}, \quad \text{and} \quad \tilde{c} = \begin{pmatrix} c \\ 0 \end{pmatrix}.$$

It is easy to get a strictly feasible point for this model by taking

$$\boxed{y_0 \text{ arbitrary}, \quad \eta_0 = \sigma_0 > [\min_K(c - A^*(y_0))]^-, \quad \text{and} \quad s_0 = c - A^*(y_0) + \eta_0 e,}$$

where $\min_K$ has been defined by (1.5) and $\alpha^- := \max(0, -\alpha)$.

We know that (4.5) is the dual of

$$\begin{cases} \inf\ \langle \tilde{c}, \tilde{x} \rangle_{\tilde{\mathbb{E}}} \\ \tilde{A}(\tilde{x}) = \tilde{b} \\ \tilde{x} \in \tilde{K}, \end{cases} \tag{4.6}$$

where $\tilde{x} = (x, \xi_1) \in \mathbb{E} \times \mathbb{R}$. Now, $\tilde{A} : \tilde{\mathbb{E}} \to \mathbb{R}^{m+1}$ is defined at $\tilde{x}$ by

$$\tilde{A}(\tilde{x}) = \begin{pmatrix} A(x) \\ -\langle e, x \rangle_{\mathbb{E}} - \xi_1 \end{pmatrix}.$$

Therefore, problem (4.6) also reads

$$\begin{cases} \inf \langle c, x\rangle_{\mathbb{E}} \\ A(x) = b \\ \langle e, x\rangle_{\mathbb{E}} + \xi_1 = M_1 \\ x \in K \\ \xi_1 \geqslant 0. \end{cases} \tag{4.7}$$

A strictly feasible point $(x_0, \xi_{1,0})$ for (4.6)-(4.7) can be

$$\boxed{\text{the known } x_0 \qquad \text{and} \qquad \xi_{1,0} = M_1 - \langle e, x_0\rangle_{\mathbb{E}},}$$

provided $M_1 > \langle e, x_0\rangle_{\mathbb{E}}$.

The proposed approach consists in solving (4.7)-(4.4), or equivalently (4.6)-(4.5) in standard form, from a strictly feasible primal-dual pair hoping that at the primal-dual solution $(x, \xi_1, y, \eta, s, \sigma)$ the scalar η vanishes, in which case problem (P) is actually solved. If $\eta \neq 0$, one increases M_1 and solve problem (4.7)-(4.4) again. The process is stopped when $\eta = 0$ is found or when M_1 is considered as too large. The latter case may occur if (D) is infeasible.

Algorithm 4.1.1 (getting a solution from a strictly feasible primal point) Let $x \in \mathcal{F}_P^s$ and let $M_1 > \langle e, x\rangle_{\mathbb{E}}$. One iteration of the algorithm is as follows.

1. *Solve* (4.7)-(4.4) from a primal-dual strictly feasible point to get $(x, \xi_1, y, \eta, s, \sigma)$.
2. *Successful stopping test.* If $\eta = 0$, stop and declare that (x, y, s) is a primal-dual solution to (P).
3. *Failure stopping test.* If M_1 is too large, stop and declare that it is likely that *the dual problem* (D) *is infeasible.*
4. *Increase* M_1.

4.1.3 Starting from a strictly feasible dual point

We assume in this section that a strictly feasible dual pair $(y_0, s_0) \in \mathcal{F}_D^s$ is known, meaning that

$$A^*(y_0) + s_0 = c \qquad \text{and} \qquad s_0 \in K^s,$$

but that $x_0 \notin K^s$, and we determine a modified SDCO problem such that it is easy to maintain (y_0, s_0) strictly feasible and to determine a strictly feasible primal point x_0.

Consider the following two equivalent modified primal problems in $(x, \xi_2) \in \mathbb{E} \times \mathbb{R}$:

$$\begin{cases} \inf \langle c, x\rangle_{\mathbb{E}} + M_2\xi_2 \\ A(x) = b \\ x + \xi_2 e \in K \\ \xi_2 \geqslant 0 \end{cases} \quad \text{or} \quad \begin{cases} \inf \langle c, x\rangle_{\mathbb{E}} + (M_2 - \langle c, e\rangle_{\mathbb{E}})\,\xi_2 \\ A(x - \xi_2 e) = b \\ x \in K \\ \xi_2 \geqslant 0. \end{cases} \tag{4.8}$$

The problem in the left-hand side in (4.8) shows that one tries to have an as small as possible perturbation of x in K, namely $x + \xi_e e$, by forcing the nonnegative scalar ξ_2 to be as small as possible thanks to the positive number M_2 in the objective, which is taken "sufficiently" large. This is indeed desirable, since if $\xi_2 = 0$, the original primal problem (P) is recovered. The equivalent problem in the right-hand side, obtained by the redefinition $x + \xi_2 e \curvearrowright x$, is in a form closer to the standard primal SDCO problem. We claim that

- this problem can be cast as the standard primal SDCO problem, for which a strictly feasible point is easy to define,
- its primal-dual solutions are the same as the original problem, provided M_2 is chosen "sufficiently large" and the primal problem is feasible.

Let us see this.

One can cast the problem as the right-hand side of (4.8) in the standard primal SDCO problem in the variable $\tilde{x} \in \tilde{\mathbb{E}} := \mathbb{E} \times \mathbb{R}$:

$$\begin{cases} \inf \langle \tilde{c}, \tilde{x} \rangle_{\tilde{\mathbb{E}}} \\ \tilde{A}(\tilde{x}) = b \\ \tilde{x} \in \tilde{K}. \end{cases} \tag{4.9}$$

This problem is identical to problem (4.8) provided the vectors $\tilde{c}$ and $\tilde{x} \in \tilde{\mathbb{E}}$, the linear map $\tilde{A} : \tilde{\mathbb{E}} \to \mathbb{R}^m$, and the cone $\tilde{K}$ are defined by

$$\tilde{c} = \begin{pmatrix} c \\ M_2 - \langle c, e \rangle_{\mathbb{E}} \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} x \\ \xi_2 \end{pmatrix}, \quad \tilde{A}(\tilde{x}) = A(x - \xi_2 e), \quad \text{and} \quad \tilde{K} := K \times \mathbb{R}_+.$$

A strictly feasible point for (4.9) is given by

$$\tilde{x}_0 := \begin{pmatrix} x_0 + \xi_{2,0} e \\ \xi_{2,0} \end{pmatrix},$$

where

$$\boxed{\begin{array}{c} x_0 \in \mathbb{E} \text{ is an arbitrary solution to } A(x_0) = b, \\ \xi_{2,0} > [\min_K(x_0)]^-. \end{array}}$$

We know that (4.9) has for dual the following problem in $(y, \tilde{s}) \in \mathbb{R}^m \times \tilde{\mathbb{E}}$:

$$\begin{cases} \sup\ b^{\mathsf{T}} y \\ \tilde{A}^*(y) + \tilde{s} = \tilde{c} \\ \tilde{s} \in \tilde{K}. \end{cases} \tag{4.10}$$

Now, $\tilde{A}^* : \mathbb{R}^m \to \tilde{\mathbb{E}}$ is defined at $\tilde{x}$ by

$$\tilde{A}^*(y) = \begin{pmatrix} A^*(y) \\ -A(e)^{\mathsf{T}} y \end{pmatrix}.$$

Writing

$$\tilde{s} := \begin{pmatrix} s \\ \sigma_2 \end{pmatrix},$$

we see that problem (4.10) also reads

$$\begin{cases} \sup\ b^{\mathsf{T}} y \\ A^*(y) + s = c \\ -A(e)^{\mathsf{T}} y + \sigma_2 = M_2 - \langle c, e \rangle_{\mathbb{E}} \\ s \in K \\ \sigma_2 \geqslant 0. \end{cases}$$

Eliminating σ_2, the dual problem (4.10) becomes the following problem in $(y, s) \in \mathbb{R}^m \times \mathbb{E}$:

$$\begin{cases} \sup\ b^{\mathsf{T}} y \\ A^*(y) + s = c \\ s \in K \\ \langle e, s \rangle_{\mathbb{E}} \leqslant M_2. \end{cases} \tag{4.11}$$

It can be shown that the variable ξ_2 in (4.8) is a dual variable associated with the constraint $\langle e, s \rangle_{\mathbb{E}} \leqslant M_2$ in (4.11). The given strictly feasible dual pair for (y_0, s_0) for (D) is still a strictly feasible pair for the modified dual problems (4.10) provided

$$\boxed{M_2 > \langle e, s_0 \rangle_{\mathbb{E}}.}$$

Note that this setting ensures that $M_2 > 0$ (since $s_0 \in K^s$), as desired.

The proposed approach consists in solving (4.8)-(4.11) from a strictly feasible primal-dual pair (instead of the original problem (P)) hoping that at the solution (x, ξ_2, y, s) the scalar ξ_2 vanishes, in which case problem (P) is actually solved. If $\xi_2 \neq 0$, the penalty parameter M_2 is increased and problem (4.8) is solved again. The process is stopped when a solution with $\xi_2 = 0$ is found or when M_2 is considered as too large, which may occur if (P) is infeasible.

Algorithm 4.1.2 (getting a solution from a strictly feasible dual point) Let $(y, s) \in \mathcal{F}_D^s$ and let $M_2 > \langle e, s \rangle_{\mathbb{E}}$. One iteration of the algorithm is as follows.

1. *Solve* (4.8)-(4.11) from a primal-dual strictly feasible point to get (x, ξ_2, y, s).
2. *Successful stopping test.* If $\xi_2 = 0$, then stop and declare that (x, y, s) is a primal-dual solution to (P).
3. *Failure stopping test.* If M_2 is too large, stop and declare that it is likely that *the primal is infeasible.*
4. *Increase* M_2.

4.1.4 Starting without a strictly feasible point

If no strictly feasible primal and/or dual point is known, then one considers the modified primal problem, obtained by combining the methods of sections 4.1.2 and 4.1.3:

$$\left\{\begin{array}{l}\inf \langle c, x\rangle_{\mathbb{E}} + M_2\xi_2 \\ A(x) = b \\ \langle e, x\rangle_{\mathbb{E}} + \xi_1 = M_1 \\ x + \xi_2 e \in K \\ \xi_1 \geqslant 0 \\ \xi_2 \geqslant 0\end{array}\right. \quad \text{or} \quad \left\{\begin{array}{l}\inf \langle c, x\rangle_{\mathbb{E}} + (M_2 - \langle c, e\rangle_{\mathbb{E}})\xi_2 \\ A(x - \xi_2 e) = b \\ \langle e, x\rangle_{\mathbb{E}} + \xi_1 - n_c\xi_2 = M_1 \\ x \in K \\ \xi_1 \geqslant 0 \\ \xi_2 \geqslant 0,\end{array}\right. \tag{4.12}$$

in which both M_1 and M_2 are taken "sufficiently" large (initial values are given below). The problem in the left-hand side in (4.12) satisfies the affine constraint $A(x) = b$ (a constraint that can be easily verified by the technique proposed in section 4.1.1) without $x \in K$, but ensures that $x + \xi_2 e$ is in K, where ξ_2 will be small when M_2 is taken large. Its second affine constraint $\langle e, x\rangle_{\mathbb{E}} + \xi_1 = M_1$ is directly inspired from the one appearing in (4.7) and will allow us to find easily a strictly feasible point for the associated dual problem. The problem in the left-hand side of (4.12) is not in standard primal form because of the constraint $x + \xi_2 e \in K$, which is the reason why it is rewritten as the one in the right-hand side, which is obtained after the change of variable $x + \xi_2 e \curvearrowright x$.

Problem in the right-hand side of (4.12) can be cast as the standard primal SDCO problem in the variable $\tilde{x} \in \tilde{K} := K \times \mathbb{R}^2_+ \subset \tilde{\mathbb{E}} := \mathbb{E} \times \mathbb{R}^2$:

$$\left\{\begin{array}{l}\inf \langle \tilde{c}, \tilde{x}\rangle_{\tilde{\mathbb{E}}} \\ \tilde{A}(\tilde{x}) = \tilde{b} \\ \tilde{x} \in \tilde{K},\end{array}\right. \tag{4.13}$$

where the vectors $\tilde{c}$ and $\tilde{x} \in \tilde{\mathbb{E}}$, the linear map $\tilde{A} : \tilde{\mathbb{E}} \to \mathbb{R}^{m+1}$, and the vector $\tilde{b} \in \mathbb{R}^{m+1}$ are defined by

$$\tilde{c} = \begin{pmatrix} c \\ 0 \\ M_2 - \langle c, e\rangle_{\mathbb{E}} \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} x \\ \xi_1 \\ \xi_2 \end{pmatrix}, \quad \tilde{A}(\tilde{x}) = \begin{pmatrix} A(x - \xi_2 e) \\ -\langle e, x\rangle_{\mathbb{E}} - \xi_1 + n_c\xi_2 \end{pmatrix}, \text{ and } \tilde{b} = \begin{pmatrix} b \\ -M_1 \end{pmatrix}.$$

The opposite of the second constraint in (4.12) adopted in the definitions above is motivated by the desire to have $-M_1$ as the last component of $\tilde{b}$, which will result in requiring $M_1 > 0$, which goes in the same sense as the condition $M_2 > 0$ required above. A strictly feasible point for the modified primal problems (4.12) or (4.13) is

$$\tilde{x}_0 := \begin{pmatrix} x_0 + \xi_{2,0} e \\ \xi_{1,0} \\ \xi_{2,0} \end{pmatrix},$$

where one takes in order:

$$\boxed{\begin{array}{c} x_0 \in \mathbb{E} \text{ is an arbitrary solution to } A(x_0) = b, \\ \xi_{2,0} > [\min_K(x_0)]^-, \\ M_1 > [\langle e, x_0\rangle_{\mathbb{E}}]^+, \\ \xi_{1,0} := M_1 - \langle e, x_0\rangle_{\mathbb{E}}, \end{array}} \tag{4.14}$$

where $\min_K$ has been defined by (1.5), $\alpha^+ = \max(0, \alpha)$, and $\alpha^- = \max(0, -\alpha)$.

We know that the dual of (4.13) reads

$$\left\{ \begin{array}{l} \sup\ \tilde{b}^{\mathsf{T}} \tilde{y} \\ \tilde{A}^*(\tilde{y}) + \tilde{s} = \tilde{c} \\ \tilde{s} \in \tilde{K}. \end{array} \right. \tag{4.15}$$

where $\tilde{A}^* : \mathbb{R}^{m+1} \to \tilde{\mathbb{E}}$ is defined at $\tilde{y} = (y, \eta) \in \mathbb{R}^m \times \mathbb{R}$ by

$$\tilde{A}^*(\tilde{y}) = \begin{pmatrix} A^*(y) - \eta e \\ -\eta \\ -\langle e, A^*(y)\rangle_{\mathbb{E}} + n_c \eta \end{pmatrix}.$$

Therefore, with $\tilde{s} = (s, \sigma_1, \sigma_2)$, problem (4.15) reads

$$\left\{ \begin{array}{l} \sup\ b^{\mathsf{T}} y - M_1 \eta \\ A^*(y) - \eta e + s = c \\ -\eta + \sigma_1 = 0 \\ -\langle e, A^*(y)\rangle_{\mathbb{E}} + n_c \eta + \sigma_2 = M_2 - \langle c, e\rangle_{\mathbb{E}} \\ s \in K \\ \sigma_1 \geqslant 0 \\ \sigma_2 \geqslant 0 \end{array} \right.$$

or

$$\left\{ \begin{array}{l} \sup\ b^{\mathsf{T}} y - M_1 \eta \\ A^*(y) - \eta e + s = c \\ -\eta + \sigma_1 = 0 \\ \langle e, s\rangle_{\mathbb{E}} + \sigma_2 = M_2 \\ s \in K \\ \sigma_1 \geqslant 0 \\ \sigma_2 \geqslant 0. \end{array} \right. \tag{4.16}$$

Recall that the parameter M_1 fixed by (4.14) is positive, which is a desired property, since one would like to have $\eta = 0$ in the above problem. A strictly feasible point for the modified dual problems (4.15) or (4.16) is $\tilde{y}_0 = (y_0, \eta_0)$ and $\tilde{s}_0 = (s_0, \sigma_{1,0}, \sigma_{2,0})$, where one sets successively

$$\boxed{\begin{array}{c} y_0 \text{ arbitrary}, \\ \eta_0 = \sigma_{1,0} > [\min_K(c - A^*(y_0))]^-, \\ s_0 := c - A^*(y_0) + \eta_0 e, \\ M_2 > [\langle e, s_0\rangle_{\mathbb{E}}]^+, \\ \sigma_{2,0} := M_2 - \langle e, s_0\rangle_{\mathbb{E}}, \end{array}} \tag{4.17}$$

where $\min_K$ has been defined by (1.5), $\alpha^- = \max(0, -\alpha)$, and $\alpha^+ = \max(0, \alpha)$. The setting of M_2 above also ensures its positivity, which was a desired property of problems (4.12).

The goal of the following algorithm is to solve a sequence of primal-dual problems (4.13)-(4.15) or (4.12)-(4.16) by increasing M_1 and M_2 at each loop in order to have $\eta = \xi_2 = 0$, which guarantees that the original primal-dual SDCO problem has been solved. The form of the problems (4.12) and (4.16) clearly indicates that one must increase M_1 if $\eta > 0$ and one must increase M_2 when $\xi_2 > 0$.

Algorithm 4.1.3 (getting a solution from a linear feasible point) One loop of the algorithm is as follows.

1. *Solve* (4.13)-(4.15) or (4.12)-(4.16) from a primal-dual strictly feasible point to get $(x, \xi_1, \xi_2, y, \eta, s, \sigma_1, \sigma_2)$.
2. *Successful stopping test.* If $\xi_2 = 0$ and $\eta = 0$, stop and declare that (x, y, s) is a primal-dual solution to (P).
3. *Failure on dual infeasibility.* If M_1 is too large, stop and declare that it is likely that *the dual problem is infeasible.*
4. *Failure on primal infeasibility.* If M_2 is too large, stop and declare that it is likely that *the primal problem is infeasible.*
5. *Increase* M_1 if $\eta > 0$.
6. *Increase* M_2 if $\xi_2 > 0$.

4.1.5 Implementation

Here are a few recommendations.

- Because of the introduction of modified SDCO problems that must be solved by the developed solver, this one has to deal with SDCO problems of various dimensions and data in the same run. It is therefore recommended to have a Matlab function that can find a solution of an SDCO problem when a strictly feasible primal-dual starting point is given (and when the problem has a solution). It is therefore recommended to write this function with the code that has been developed so far in the previous chapter 3. I could have the form

```
[x,y,s] = sdco_solve (A,b,c,K,x0,y0,s0)
```

- To develop the solver, a good idea is to test it on the easy test cases 1 and 2, in which the initial values of x and y has been discarded. Once the solver works correctly on these test cases, you may want to test it on the more special problems of sections 4.3.1, 4.3.2, and 4.3.3.

- The linear operator $\tilde{A}$ does not depend on M_1 and M_2 and can be set outside the loop described in algorithm 4.1.3.
- Check whether the linear constraints $\tilde{A}(\tilde{x}_0) = \tilde{b}$ and $\tilde{A}^*(\tilde{y}) + \tilde{s} = \tilde{c}$ are satisfied at the beginning of *each cycle*, after the setting of $\tilde{A}$, $\tilde{x}_0$, and $\tilde{b}$.

4.2 A practical infeasible predictor-corrector algorithm

We present in this section an infeasible interior point algorithm, whose first goal is efficiency, letting aside theoretical issues dealing with convergence and complexity. The algorithm uses a predictor-corrector technique of the Mehrotra-type [28], which provides superquadratic convergence when it is used to solve monotone linear complementarity problems [52]. This is the method implemented in SDPT3 [49].

4.2.1 Second order correction

Direction definition

Here is another approach to get the NT direction. Since the first two equations in the system (1.14) are linear, the difficult part of the derivation of the direction comes from the linearization/symmetrization of its last equation, which reads

$$x \bullet s = \mu e.$$

At the current iterate $z = (x, y, s) \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$, satisfying or not this identity, one would like to compute a displacement $d = (d_x, d_y, d_s) \in \mathbb{E} \times \mathbb{R}^m \times \mathbb{E}$ such that $(x + d_x) \bullet (s + d_s) = \mu e$. This also reads

$$x \bullet d_s + d_x \bullet s = \mu\, e - (x \bullet s + d_x \bullet d_s).$$

Recall the definition of g and g^{T} in (1.24) and (1.3). Multiplying the two sides of the previous identity to the left by g^{-1} and to the right by g, one gets

$$g^{-1} \bullet (x \bullet d_s + d_x \bullet s) \bullet g = \mu\, e - g^{-1} \bullet (x \bullet s + d_x \bullet d_s) \bullet g. \tag{4.18}$$

Given a and $b \in \mathbb{E}$, the symmetrization of $g^{-1} \bullet a \bullet b \bullet g \in \mathbb{F}$ is obtained by replacing it by $\frac{1}{2}(g^{-1} \bullet a \bullet b \bullet g + g^{\mathsf{T}} \bullet b \bullet a \bullet g^{-\mathsf{T}}) \in \mathbb{E}$, where we have set $g^{-\mathsf{T}} := g^{-\mathsf{T}}$ (recall the rule (1.4a) and that the matrix parts of a vector in $\mathbb{E}$ are symmetric). Applying this technique on the previous identity yields

$$\mathcal{E}(d_x) + \mathcal{F}(d_s) = \mu\, e - \mathcal{S}(x \bullet s + d_x \bullet d_s), \tag{4.19}$$

where the linear operators $\mathcal{E} : \mathbb{E} \to \mathbb{E}$ and $\mathcal{F} : \mathbb{E} \to \mathbb{E}$, and the symmetrization linear operator $\mathcal{S} : \mathbb{F} \to \mathbb{E}$ are defined at d_x and $d_s \in \mathbb{E}$, and at $z \in \mathbb{F}$ by

$$\mathcal{E}(d_x) = \frac{1}{2}\left(g^{-1} \bullet d_x \bullet s \bullet g + g^{\mathsf{T}} \bullet s \bullet d_x \bullet g^{-\mathsf{T}}\right),$$
$$\mathcal{F}(d_s) = \frac{1}{2}\left(g^{-1} \bullet x \bullet d_s \bullet g + g^{\mathsf{T}} \bullet d_s \bullet x \bullet g^{-\mathsf{T}}\right),$$
$$\mathcal{S}(z) = \frac{1}{2}\left(g^{-1} \bullet z \bullet g + g^{\mathsf{T}} \bullet \tilde{z} \bullet g^{-\mathsf{T}}\right).$$

Finally, the linearization is obtained by dropping the single nonlinear term $d_x \bullet d_s$ from the right-hand side of (4.19), leading to

$$\mathcal{E}(d_x) + \mathcal{F}(d_s) = \mu\, e - \mathcal{S}(x \bullet s), \tag{4.20}$$

which is identical to the third equation in the NT system (1.22).

The goal of the so-called *second order correction* is to define a better direction than the NT direction (1.22) by reconsidering the linearization phase of the last paragraph. Instead of dropping the term $d_x \bullet d_s$, it is replaced by the one obtained by a prediction phase, say $p = (p_x, p_y, p_s)$ obtained by setting $\mu = 0$ in (1.22). Hence, (4.20) becomes

$$\mathcal{E}(d_x) + \mathcal{F}(d_s) = \mu e - \mathcal{S}(x \bullet s + p_x \bullet p_s). \tag{4.21}$$

Direction computation

Let us see how to compute the solution to (4.21). Using the notation

$$\tilde{d}_x := g^{-1} \bullet d_x \bullet g^{-\mathsf{T}} \qquad \text{and} \qquad \tilde{d}_s := g^{\mathsf{T}} \bullet d_s \bullet g,$$

the identities

$$g^{\mathsf{T}} \bullet s \bullet g = g^{-1} \bullet x \bullet g^{-\mathsf{T}} = \sigma \qquad \text{and} \qquad g^{-1} \bullet x \bullet s \bullet g = \sigma^2 \tag{4.22}$$

recalled from (1.26), and the fact that σ has its matrix part formed of positive definite diagonal matrices, this identity (4.21) becomes

$$\frac{1}{2}\left[\left(\tilde{d}_x + \tilde{d}_s\right) \bullet \sigma + \sigma \bullet \left(\tilde{d}_x + \tilde{d}_s\right)\right] = \mu\, e - \sigma^2 - \frac{1}{2}\left(g^{-1} \bullet p_x \bullet p_s \bullet g + g^{\mathsf{T}} \bullet p_s \bullet p_x \bullet g^{-\mathsf{T}}\right). \tag{4.23}$$

This is a Lyapunov equation in the unknown $\tilde{d}_x + \tilde{d}_s$, which is particularly easy to solve, since σ has diagonal matrix parts. Taking the (i, j) element of a matrix part in both sides gives

$$\frac{1}{2}(\Sigma_{ii} + \Sigma_{jj})(\tilde{d}_x + \tilde{d}_s)_{ij} = (\mu - \Sigma_{ii}^2)\delta_{ij} - \frac{1}{2}\left(g^{-1} \bullet p_x \bullet p_s \bullet g + g^{\mathsf{T}} \bullet p_s \bullet p_x \bullet g^{-\mathsf{T}}\right)_{ij},$$

where δ_{ij} is the Kronecker symbol ($\delta_{ij} = 1$ is $i = j$, $\delta_{ij} = 0$ if $i \neq j$). As a result,

$$\tilde{d}_x + \tilde{d}_s = \mu\, \sigma^{-1} - \sigma + \gamma, \tag{4.24}$$

where the (i, j) element of a matrix part of γ is

$$\boxed{\gamma_{ij}^s := -\left((g^s)^{-1} \bullet p_x^s \bullet p_s^s \bullet g^s + (g^s)^{\mathsf{T}} \bullet p_s^s \bullet p_x^s \bullet (g^s)^{-\mathsf{T}}\right)_{ij} / (\Sigma_{ii} + \Sigma_{jj}).} \tag{4.25a}$$

For the vector part of γ, one immediately get from the Lyapunov equation (4.23):

$$\boxed{\gamma^l := -(\sigma^l)^{-1} \cdot p_x^l \cdot p_s^l.} \tag{4.25b}$$

Left and right-multiplying the two sides of the identity (4.24) by g and g^{T} respectively, using $w = g \bullet g^{\mathsf{T}}$ observed in (1.25), and using (4.22), yield

$$\boxed{d_x + w \bullet d_s \bullet w = \mu s^{-1} - x + g \bullet \gamma \bullet g^{\mathsf{T}}.} \tag{4.26}$$

We see that we have just added the correction $g \bullet \gamma \bullet g^{\mathsf{T}}$ to the right-hand side of $(1.22)_3$.

To summarize, the NT direction with second order correction is the solution to the system

$$\begin{cases} A^*(d_y) + d_s = r_{\mathrm{D}} \\ A(d_x) = r_{\mathrm{P}} \\ d_x + w \bullet d_s \bullet w = r_{\mathrm{S}}, \end{cases} \tag{4.27a}$$

where the residuals r_{D} and r_{P} have been defined by (1.16), and the residual r_{S} is defined by

$$r_{\mathrm{S}} := \mu s^{-1} - x + g \bullet \gamma \bullet g^{\mathsf{T}}. \tag{4.27b}$$

4.2.2 Stepsize computation

Most implemented algorithms do not use a neighborhood of the central path to control the iterates (this is the case for the algorithms implemented in SDPA [12] and SDPT3 [49]). There are at least three reasons for this.

- When the iterates are not feasible, the *feasible* central path cannot be used as a safeguard, so that the neighborhood must take into account infeasibility and are defined with more complexity.
- It is more easy to determine a stepsize to the boundary of the cone of positive semidefinite matrices than to the boundary of the neighborhood of the central path.
- For efficiency reasons, implemented algorithms compute distinct stepsizes for the primal and dual variables. These stepsizes α_{P} and α_{D} are computed from the maximal stepsizes $\bar{\alpha}_{\mathrm{P}}$ and $\bar{\alpha}_{\mathrm{D}}$ maintaining the positive semidefiniteness of matrix parts of x and $s \in \mathbb{E}$.

One proceeds as follows. Consider the case of a matrix variable M, to which corresponds the displacement d_M (M can be a matrix component of the primal or dual variables x and $s \in \mathbb{E}$). The desired stepsize α_M from M along d_M is determined as large as possible in $(0, 1]$, while ensuring a "sufficient" positive definiteness of $M + \alpha_M d_M$. Since $M \succ 0$, it makes sense to take

$$\alpha_M = 1, \qquad \text{if } d_M \succcurlyeq 0. \tag{4.28}$$

If d_M is not positive semidefinite, one looks for the largest $\alpha_M \in (0,1]$, with some safeguard, in order to satisfy the equivalent conditions

$$\lambda_{\min}(M + \alpha_M d_M) \geqslant 0, \tag{4.29a}$$

$$\lambda_{\min}(I + \alpha_M M^{-1} d_M) \geqslant 0, \tag{4.29b}$$

$$\lambda_{\min}(I + \alpha_M L_M^{-1} P_M^{\mathsf{T}} d_M P_M L_M^{-\mathsf{T}}) \geqslant 0. \tag{4.29c}$$

The equivalence between (4.29a) and (4.29b) comes from the fact that $M + \alpha_M d_M$ and $I + \alpha_M M^{-1} d_M$ have the same eigenvalues, which are therefore in $\mathbb{R}$, although the latter matrix is not symmetric [11]. In (4.29c), L_M is a factor of M in the sense that

$$P_M^{\mathsf{T}} M P_M = L_M L_M^{\mathsf{T}}$$

for some possible permutation matrix P_M (it could be the Cholesky factor with permutation matrix P_M). These conditions can be rewritten as follows

$$\alpha_M \lambda_{\min}(M^{-1} d_M) \geqslant -1, \tag{4.30a}$$

$$\alpha_M \lambda_{\min,M}(d_M) \geqslant -1, \tag{4.30b}$$

$$\alpha_M \lambda_{\min}(L_M^{-1} P_M^{\mathsf{T}} d_M P_M L_M^{-\mathsf{T}}) \geqslant -1, \tag{4.30c}$$

where we have written $\lambda_{\min,B}(A)$ the minimal λ such that $Mv = \lambda Bv$ for some nonzero v (the generalized eigenvalues of $Av = \lambda Bv$ are real when both A and B are symmetric and B is positive definite, like here). Since, in the present case, $d_M \not\succcurlyeq 0$, hence the minimal eigenvalues above are negative, the previous inequalities, with some safeguard $\beta \in (0,1)$ becomes respectively

$$\alpha_M = \min\left(1, \frac{-\beta}{\lambda_{\min}(M^{-1} d_M)}\right), \qquad \text{if } d_M \not\succcurlyeq 0, \tag{4.31a}$$

$$\alpha_M = \min\left(1, \frac{-\beta}{\lambda_{\min,M}(d_M)}\right), \qquad \text{if } d_M \not\succcurlyeq 0, \tag{4.31b}$$

$$\alpha_M = \min\left(1, \frac{-\beta}{\lambda_{\min}(L_M^{-1} P_M^{\mathsf{T}} d_M P_M L_M^{-\mathsf{T}})}\right), \qquad \text{if } d_M \not\succcurlyeq 0. \tag{4.31c}$$

In the case of the NT direction, the Cholesky factor L_M of $P_M^{\mathsf{T}} M P_M$ is computed, so that the use of $\lambda_{\min}(L_M^{-1} P_M^{\mathsf{T}} d_M P_M L_M^{-\mathsf{T}})$ is advantageous, in order to have the symmetry of the matrix whose minimum eigenvalue has to be computed, while the computing time for computing the two matrices $L_M^{-1} P_M^{\mathsf{T}} d_M P_M L_M^{-\mathsf{T}}$ and $M^{-1} d_M$ is approximately the same. In Matlab, with `eig` or `eigs`, formula (4.31c) seems to be the one that is the most conservative, in the sense that it gives the most negative minimal eigenvalue; for that reason, it is recommended.

4.2.3 The algorithm

The modifications come from the fact that the centering parameter σ is determined by a prediction step (i.e., with $\mu = 0$ in (1.22)), before the computation of the search

direction d. Furthermore, the computation of the direction is based on a modification of the residual in the third equatio of (1.22), which is now given by r_s in (4.27b).

Algorithm 4.2.1 (IPC) One describes one iteration from $z = (x, y, s)$ to $z^+ = (x^+, y^+, s^+)$. A stepsize safeguard $\beta \in (0, 1)$ is also updated in the loop to become β^+ (take $\beta = 0.9$ at the first iteration).

1. *Stopping criterion.* Stop if the feasibility and $\overline{\mu}(z)$ are small enough.
2. *Prediction step and its stepsizes.*
   - Compute the prediction direction $d' := (d'_x, d'_y, d'_s)$ by (1.22) with $\mu = 0$.
   - Compute the primal and dual stepsizes along that direction d' by (4.31c), for the current stepsize safeguard β, and denote them by α'_P and α'_D (there is one stepsize per matrix component and one for each element of the vector component).
   - Set $z' := (x + \alpha'_P d'_x, y + \alpha'_D d'_y, s + \alpha'_D d'_s)$.
3. *Centering parameter.* For some exponent e discussed below, set the *centering parameter* to
$$\sigma := \min\left(1, \left[\frac{\overline{\mu}(z')}{\overline{\mu}(z)}\right]^e\right). \tag{4.32}$$
4. *Direction.* Compute the search direction $d := (d_x, d_y, d_s)$ by (4.27) with $p = d'$ and μ set to
$$\mu := \sigma\overline{\mu}(z). \tag{4.33}$$
5. *Stepsizes.* Compute the primal and dual stepsizes α_P and α_D by (4.31c), with
$$\beta := 0.9 + 0.09\ \min(\alpha'_P, \alpha'_D).$$
6. *Update of the variables*: $x^+ = x + \alpha_P d_x$, $y^+ = y + \alpha_D d_y$, $s^+ = s + \alpha_D d_s$.
7. *Update of β*:
$$\beta^+ := 0.9 + 0.09\ \min(\alpha_P, \alpha_D).$$

The centering parameter $\sigma \in (0, 1)$ set in step 3 is used in step 4 to determine the degree to which the NT direction must aim a central point close to a solution of the problem. If σ is close to 1 a cautious direction is computed, which favors a move to the central path instead of one to the solution. In contrast, a σ close to zero favors an audacious move towards the solution. Assuming that the exponent $e = 1$, we see that if the prediction step computes a point z' providing an important decrease of the duality gap $\overline{\mu}(z')$, then the centering parameter is taking small in step 3. The exponent e in (4.32) has been introduced to reinforce this behavior of the algorithm. In SDPT3 [48], it has been determined experimentally and set to

$$e := \begin{cases} \max[1, 3\min(\alpha'_P, \alpha'_D)^2] & \text{if } \overline{\mu}(z) > 10^{-6}, \\ 1 & \text{otherwise.} \end{cases}$$

4.3 Test cases

4.3.1 Test case 4a

Consider the problem [50; p. 65]:

$$\begin{cases} \sup & y_1 \\ & \begin{pmatrix} 0 & y_1 & 0 \\ y_1 & y_2 & 0 \\ 0 & 0 & y_1+1 \end{pmatrix} \succcurlyeq 0. \end{cases}$$

4.3.2 Test case 4b

Consider the primal SDO problem written in standard form with the following data:

$$C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{and} \quad b = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

4.3.3 Test case 4c

Consider the primal SDO problem written in standard form with the following data:

$$C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Questions

4.1. {5} *IPC algorithm.* Implement algorithm 4.2.1 and solve/diagnose the test cases of sections 4.3.1, 4.3.2, and 4.3.3.

5 A few applications

This chapter describes a few non trivial applications of self-dual conic optimization, which can be used as test-cases for the developed SDCO solver.

5.1 Global minimization of a univariate polynomial

5.1.1 Unconstrained minimization

We consider the problem of finding the *global minimum of a univariate real polynomial* $p \in \mathbb{R}[x]$ (hence in the variable $x \in \mathbb{R}$), which reads

$$\inf_{x \in \mathbb{R}} \left(p(x) := \sum_{\alpha \in \mathbb{N}} p_\alpha \, x^\alpha \right), \tag{5.1}$$

where only a finite number of coefficients $p_\alpha \in \mathbb{R}$ are nonzero. The degree $d := \deg p$ of the polynomial p is the largest $\alpha \in \mathbb{N}$ such that $p_\alpha \neq 0$. Problem (5.1) can be rewritten as an SDCO problem. Its primal form can be deduced from the SOS approach, while its dual form can be obtained by the moment approach. We start with the SOS approach, which is more intuitive for a reader who is not familiar with measure theory.

The SOS approach

The name of this approach comes from the fact that it is grounded on the representation of a nonnegative univariate polynomial as a sum of squares (SOS) of polynomials (proposition 5.1.1).

The starting point consists in reformulating (5.1) as a problem with an infinite number of constraints and a single variable $s \in \mathbb{R}$:

$$\begin{cases} \sup_{s \in \mathbb{R}} \; s \\ p(x) \geqslant s, \quad \forall \, x \in \mathbb{R}. \end{cases} \tag{5.2}$$

This straightforward reformulation results in an amazing observation: finding the global minimum of a function (not necessarily a polynomial) is a *linear* optimization problem, but with an *infinite* number of *linear* constraints. The question now is to determine whether one can transform this problem in an equivalent one, but with a *finite* number of constraints. It happens that this is possible when p is a polynomial, essentially because

the equality between two polynomials can be expressed by the equality between their coefficients, whose number is finite. Let us introduce the approach progressively.

The infinite number of constraints of the previous problem "disappears" if we trivially express them in terms of membership to the cone $\mathcal{P}$ of univariate nonnegative polynomials:

$$\begin{cases} \sup_{s\in\mathbb{R}} s \\ p - s \in \mathcal{P}. \end{cases} \tag{5.3}$$

It turns out that the membership to $\mathcal{P}$ has an LMI representation (this is no longer true for multivariate polynomials, but there are bypasses [37, 22, 24, 2, 3, 25]). That fact is based on the following two properties: a univariate nonnegative polynomial can be written as a *sum of squares* (SOS) of polynomials (proposition 5.1.1, which is no longer true for multivariate polynomials) and an SOS polynomial can be represented thanks to a positive semidefinite matrix (proposition 5.1.2, which is still true for multivariate polynomials).

Proposition 5.1.1 (nonnegative univariate polynomial on $\mathbb{R}$) *A univariate polynomial $p \in \mathbb{R}[x]$ is nonnegative on $\mathbb{R}$ if and only if it is of even degree, say $2m$, and reads $p = q^2 + r^2$, with q, $r \in \mathbb{R}[x]$, $\deg q = m$, and $\deg r \leqslant m-1$.*

Proof. The given conditions are clearly sufficient. Let us show that they are necessary.

Being nonnegative on $\mathbb{R}$ the polynomial is necessary of even degree, say $2m$, and the leading coefficient is positive. There is therefore no loss of generality in supposing that this leading coefficient is 1. Then, the polynomial can be decomposed in m factors of the form

$$(x-a)^2 + b^2.$$

It is indeed the form of $(x-r)(x-\bar{r})$ when r and $\bar{r}$ are complex conjugate roots $a \pm ib$. On the other hand, any real root has an even multiplicity (otherwise the polynomial would have positive and negative values around the root) and each double root is of the form above with $b = 0$.

The m factors of the form $q^2 + r^2$ are then multiplied successively, by using the following formula

$$(q_j^2 + r_j^2)(q^2 + r^2) = (q_j q + r_j r)^2 + (q_j r - r_j q)^2 =: q_{j+1}^2 + r_{j+1}^2.$$

By induction, we see that $\deg q_j = j$ and $\deg r_j \leqslant j-1$. It is indeed the case for $j = 1$ since $\deg q_1 = 1$ and $\deg r_1 = 0$. Now, be r vanishing or not, $\deg q_{j+1} = \deg q_j + 1 = j+1$. Finally, if $r = 0$, $\deg r_{j+1} = \deg r_j + 1 \leqslant j$ and, if $r \neq 0$, $\deg r_{j+1} \leqslant \max(\deg q_j, \deg r_j + 1) \leqslant j$. □

Proposition 5.1.2 (caracterization of SOS polynomials) *A polynomial $p \in \mathbb{R}[x]$ of degree $\leqslant 2m$ is a sum of r squares of polynomials if and only if there exists a matrix $X \succcurlyeq 0$ of order $m+1$ and of rank $\leqslant r$ such that $p(x) = v_m(x)^\mathsf{T} X v_m(x)$, where $v_m(x)$ is the vector of monomials $(1 \; x \; x^2 \; \dots \; x^m)^\mathsf{T}$.*

Proof. [$\Rightarrow$] If $p \in \mathbb{R}[x]$ is of degree $\leqslant 2m$ and reads $\sum_{i=1}^r \sigma_i^2$, where $\sigma_i \in \mathbb{R}[x]$, the degrees $\deg \sigma_i \leqslant m$. Therefore, on can find vectors $z_i \in \mathbb{R}^{m+1}$ such that

$$p(x) = \sum_{i=1}^r (z_i^\mathsf{T} v_m(x))^2 = \sum_{i=1}^r v_m(x)^\mathsf{T} z_i z_i^\mathsf{T} v_m(x) = v_m(x)^\mathsf{T} X v_m(x),$$

where $X := \sum_{i=1}^r z_i z_i^\mathsf{T} \succcurlyeq 0$ is of rank $\leqslant r$.

[$\Leftarrow$] Conversely, suppose that $p(x) = v_m(x)^\mathsf{T} X v_m(x)$, with $X \succcurlyeq 0$ of rank $\leqslant r$. The spectral decomposition of $X = \sum_{i=1}^r z_i z_i^\mathsf{T}$ allows us to write

$$p(x) = \sum_{i=1}^r v_m(x)^\mathsf{T} z_i z_i^\mathsf{T} v_m(x) = \sum_{i=1}^r (z_i^\mathsf{T} v_m(x))^2,$$

showing that p is the sum of the squares of at most r polynomials. □

Using the above propositions and setting $s = -t$, problem (5.3) can be written

$$\left\{ \begin{array}{l} \inf_{(X,t) \in \mathcal{S}^{m+1} \times \mathbb{R}} t \\ p(x) + t = v_m(x)^\mathsf{T} X v_m(x), \qquad \forall\, x \in \mathbb{R} \\ X \succcurlyeq 0. \end{array} \right. \tag{5.4}$$

Knowing p, the equality constraint in (5.4) is an affine constraint on the unknown $t \in \mathbb{R}$ and the unknown real elements of X, since it is equivalent to the equality between the coefficients of the same monomials in both sides of the identity. To highlight the structure of problem (5.4) and make the determination its dual easier, for $\alpha \in [0\,:\,2m]$, let us introduce the symmetric matrix B_α of order $m+1$, whose $(i,j) \in [1\,:\,m+1]^2$ element is defined by

$$(B_\alpha)_{ij} := \left\{ \begin{array}{ll} 1 & \text{if } i+j-2 = \alpha \\ 0 & \text{otherwise.} \end{array} \right. \tag{5.5}$$

Then,

$$\begin{aligned} v_m(x)^\mathsf{T} X v_m(x) &= \sum_{i,j \in [1\,:\,m+1]} X_{ij} x^{i+j-2} \\ &= \sum_{\alpha \in [0\,:\,2m]} \left(\sum_{i+j-2=\alpha} X_{ij} \right) x^\alpha \\ &= \sum_{\alpha \in [0\,:\,2m]} \langle B_\alpha, X \rangle x^\alpha. \end{aligned}$$

Finally, problem (5.4) reads

$$\begin{cases} \inf_{(X,t)\in\mathcal{S}^{m+1}\times\mathbb{R}} t \\ \langle B_\alpha, X\rangle = p_\alpha + t\delta_{0\alpha}, \qquad \forall\, \alpha \in [0\,{:}\,2m] \\ X \succcurlyeq 0, \end{cases} \tag{5.6}$$

where $\delta_{0\alpha} = 1$ if $\alpha = 0$ and $\delta_{0\alpha} = 0$ otherwise.

Problem (5.6) is almost in the primal form of an SDCO problem. The difference comes from the free variable t, which is only constrained by the affine constraint in (5.6), not by a nontrivial cone. At least two techniques could be considered to solve this problem, which is not in the standard form of the primal SDCO problem.

(F1) Introduce additional *free variables* $x^f \in \mathbb{R}^p$ in the SDCO model (1.7) (by *free variables*, we mean variables that are forced to satisfy the affine constraint but not to belong to an additional cone, hence they are not completely free).

(F2) Write $t = u - v$ and impose $u \geqslant 0$, $v \geqslant 0$. The standard SDCO model (1.7) can then be used with $s = 1$, $n_1 = m + 1$, $n_l = 2$,

$$x = \begin{pmatrix} X \\ u \\ v \end{pmatrix}, \qquad c = \begin{pmatrix} 0_{\mathcal{S}^{m+1}} \\ 1 \\ -1 \end{pmatrix}, \tag{5.7}$$

where $0_{\mathcal{S}^{m+1}}$ is the zero matrix of order $m+1$, and A is the map linking linearly the coefficients of $p(x) + u - v$ and those of $v_m(x)^\mathsf{T} X v_m(x)$.

We suggest using the second approach, which has the advantage of allowing us to use the so far developed code without modification. Actually, some authors [49] decompose each free variable into the difference of two nonnegative variables, as we did it for t above, since otherwise their numerical treatment seems to introduce difficulties.

The final primal formulation of the problem is therefore

$$\begin{cases} \inf_{(X,u,v)\in\mathcal{S}^{m+1}\times\mathbb{R}\times\mathbb{R}} u - v \\ \langle B_\alpha, X\rangle - (u-v)\delta_{0\alpha} = p_\alpha, \qquad \forall\, \alpha \in [0\,{:}\,2m] \\ X \succcurlyeq 0, \quad u \geqslant 0, \quad v \geqslant 0, \end{cases} \tag{5.8}$$

where $\delta_{0\alpha} = 1$ if $\alpha = 0$ and $\delta_{0\alpha} = 0$ otherwise.

As a result, for finding the global minimum value of a polynomial of degree $d = 2m$ (the degree must be even, otherwise the polynomial is not bounded below), the SDCO approach requires determining a positive semidefinite matrix of order $m+1$ and a nonnegative vector of size 2, subject to $d+1$ affine constraints. For this problem, the linear operator A of the SDCO problem (1.7) has actually a very sparse matrix representation; more specifically, the proportion of nonzero elements is

$$\frac{m^2 + 2m + 1}{2m^3 + 5m^2 + 7m + 3} \quad\sim\quad \frac{1}{2m} = \frac{1}{\deg p} \quad \text{(for large } m\text{)},$$

so that using sparsity techniques allows the approach to compute the global minimal value of polynomials of rather high degree.

The moment approach

Whilst the SOS approach provides the minimal value of a polynomial, the moment approach provides the minimizer of the polynomial. Both approaches are dual to each other in the sense of Lagrange duality, so that an SDCO solver finds a solution to both problems simultaneously.

The dual of problem (5.8) is the following problem

$$\begin{cases} \sup_{(y,S)\in\mathbb{R}^{d+1}\times\mathcal{S}^{m+1}} \sum_{\alpha\in[0\,:\,d]} p_\alpha y_\alpha \\ \sum_{\alpha\in[0\,:\,d]} y_\alpha B_\alpha + S = 0_{\mathcal{S}^{m+1}} \\ y_0 = -1 \\ S \succcurlyeq 0. \end{cases} \tag{5.9}$$

Noting $\bar{y}_\alpha := -y_\alpha$, the problem also reads

$$\begin{cases} \inf_{(\bar{y},S)\in\mathbb{R}^{d+1}\times\mathcal{S}^{m+1}} \sum_{\alpha\in[0\,:\,d]} p_\alpha \bar{y}_\alpha \\ \sum_{\alpha\in[0\,:\,d]} \bar{y}_\alpha B_\alpha \succcurlyeq 0 \\ \bar{y}_0 = 1. \end{cases} \tag{5.10}$$

This problem has an interpretation in terms of the moments $\bar{y}_\alpha$ of a nonnegative measure μ on $\mathbb{R}$, defined by

$$\bar{y}_\alpha = \int_{\mathbb{R}} x^\alpha \,\mathrm{d}\mu.$$

The condition $\bar{y}_0 = 1$ in (5.10) reflects the fact that the sought measure is forced to have a unit mass, i.e., $\int_{\mathbb{R}} \mathrm{d}\mu = 1$. Hence, implicitly, problem (5.10) determines the measure μ through its moments $\bar{y}_\alpha$. When the minimizer of the polynomial p is unique, the optimal measure μ is the Dirac measure at the solution so that the minimizer can be recovered in

$$\bar{y}_1 = \int_{\mathbb{R}} x \,\mathrm{d}\mu.$$

5.1.2 Constrained minimization

Assume now that the problem consists in minimizing the polynomial p in (5.1) on an interval $I \subset \mathbb{R}$, which may take the form $I = \mathbb{R}$ (unconstrained problem), $I = [a, \infty)$, $I = (-\infty, b]$, or $I = [a, b]$, for some a and $b \in \mathbb{R}$:

$$\inf_{x\in I} p(x). \tag{5.11}$$

Like in section 5.1.1, we introduce the methods by the SOS approach. By the same reasoning as the one leading to (5.2), we see that problem (5.11) can obviously be expressed by

$$\begin{cases} \sup_{s\in\mathbb{R}} s \\ p - s \text{ is nonnegative on } I. \end{cases} \tag{5.12}$$

Like with (5.2), this is a linear optimization problem with an infinite number of constraints, one constraint $p(x) \geqslant s$ on s for each x in I. From this point of view, we are

face to the problem of expressing the nonnegativity of a *univariate* polynomial on an interval by a finite number of constraints. Remarkably, it turns out that this property can also be expressed by an LMI.

One can give a representation of a univariate polynomial that is nonnegative on the interval $[0, \infty)$ without sophisticated algebraic concepts, see proposition below 5.1.3 [38, 39]. The case of intervals of the form $[a, \infty)$, $(-\infty, b]$, or $[a, b]$, with a and $b \in \mathbb{R}$, are then dealt with by a change of variable. A common interest of the obtained representations of *univariate* polynomials is to be of the form

$$p = \sum_{j \in J} \sigma_i q_j, \tag{5.13}$$

where $|J| \leqslant 2$, the σ_j's are SOS with a degree that can be determined from the one of p, and the q_j's are polynomials of degree 0, 1, or 2 (the constraints $q_j(x) \geqslant 0$ represent the considered interval). This property is important from a numerical point of view, since the degrees of the σ_j's directly determine the order of the positive semidefinite matrices representing them (proposition 5.1.2).

Nonnegativity on $[a, \infty)$

We start with proposition 5.1.3, which gives the *Pólya and Szegö representation* of a univariate polynomial that is nonnegative on $\mathbb{R}_+$ [38; 1976 ; § VI 45].

Proposition 5.1.3 (nonnegative univariate polynomial on $[0, \infty)$) *A univariate polynomial $p \in \mathbb{R}[x]$ of degree d is nonnegative on $[0, \infty)$ if and only if there exist σ_0 and $\sigma_1 \in \Sigma[x]$ such that $\deg \sigma_0 \leqslant d$, $\deg \sigma_1 \leqslant d - 1$ and, for all $x \in \mathbb{R}$:*

$$p(x) = \sigma_0(x) + \sigma_1(x)x. \tag{5.14}$$

Proof. [⇒] Observe first that if $p_i \in \mathbb{R}[x]$, $i \in \{1, 2\}$, are two polynomials of the required form $\sigma_{0,i} + \sigma_{1,i}x$ with $\sigma_{0,i}$ and $\sigma_{1,i} \in \Sigma[x]$, $\deg \sigma_{0,i} \leqslant \deg p_i$ and $\deg \sigma_{1,i} \leqslant \deg p_i - 1$, then the same property holds for their product, which reads

$$p_1 p_2 = (\sigma_{0,1}\sigma_{0,2} + \sigma_{1,1}\sigma_{1,2}x^2) + (\sigma_{0,1}\sigma_{1,2} + \sigma_{0,2}\sigma_{1,1})x.$$

Indeed, $\sigma_{0,1}\sigma_{0,2} + \sigma_{1,1}\sigma_{1,2}x^2 \in \Sigma[x]$, with a degree $\leqslant \deg p_1 p_2$ and $\sigma_{0,1}\sigma_{1,2} + \sigma_{0,2}\sigma_{1,1} \in \Sigma[x]$ with a degree $\leqslant \deg p_1 p_2 - 1$.

Now, write p like the product of its factors. The product of two factors corresponding to conjugate roots $a \pm ib$ reads $(x - a)^2 + b^2 \in \Sigma[x]$, which is of the required form (with $\sigma_1 = 0$). The positive real roots must have an even multiplicity and the product of the corresponding factors is therefore in $\Sigma[x]$, which is of the required form (with $\sigma_1 = 0$). To each nonopsivite real root corresponds a factor of the form $a + x$, with $a \geqslant 0$, which is of the required form (with $\sigma_0 = a$ and $\sigma_1 = 1$).

[⇐] This is clear since for $x \in [0,\infty)$, $x \geqslant 0$, while σ_0 and σ_1 are nonnegative polynomials. □

Corollary 5.1.4 (nonnegative univariate polynomial on $[a,\infty)$) *Let a be a real number. A univariate polynomial $p \in \mathbb{R}[x]$ of degree d is nonnegative on $[a,\infty)$ if and only if there exist σ_0 and $\sigma_1 \in \Sigma[x]$ such that $\deg \sigma_0 \leqslant d$, $\deg \sigma_1 \leqslant d-1$ and, for all $x \in \mathbb{R}$:*

$$p(x) = \sigma_0(x) + \sigma_1(x)\,(x-a). \tag{5.15}$$

Proof. Let $\tilde{p}$ be the polynomial defined at $x \in \mathbb{R}$ by $\tilde{p}(x) = p(x+a)$. Then, $\tilde{p} \geqslant 0$ on $[0,\infty)$. By proposition 5.1.3, there exist $\tilde{\sigma}_0$ and $\tilde{\sigma}_1 \in \Sigma[x]$, with $\deg \tilde{\sigma}_0 \leqslant d$ and $\deg \tilde{\sigma}_1 \leqslant d-1$, such that $\tilde{p}(x) = \tilde{\sigma}_0(x) + \tilde{\sigma}_1(x)x$ for all $x \in \mathbb{R}$. Since $p(x) = \tilde{p}(x-a)$, it follows that

$$p(x) = \tilde{\sigma}_0(x-a) + \tilde{\sigma}_1(x-a)\,(x-a), \qquad \forall\, x \in \mathbb{R}.$$

Now the polynomials σ_0 and σ_1 defined at x by $\sigma_0(x) = \tilde{\sigma}_0(x-a)$ and $\sigma_1(x) = \tilde{\sigma}_1(x-a)$ are clearly in $\Sigma[x]$ with degrees d and $d-1$ respectively. □

Let us now deduce from the representation (5.19) the SDCO form of the problem of minimizing a polynomial on $(-\infty, b]$. With

$$m_0 := \lceil d/2 \rceil \qquad \text{and} \qquad m_1 := \lceil (d-1)/2 \rceil, \tag{5.16}$$

we have the following table

$d =$	$2m$	$2m+1$
$m_0 =$	m	$m+1$
$m_1 =$	m	m

Note also that the degrees of σ_0 and σ_1 in corollary 5.1.4 are respectively $\leqslant 2m_0$ and $\leqslant 2m_1$. Using proposition 5.1.1, the representation (5.15) of p, and setting $s = -t$, problem (5.12) can be written

$$\begin{cases} \inf_{(X^0,X^1,t)\in \mathcal{S}^{m_0+1}\times\mathcal{S}^{m_1+1}\times\mathbb{R}} t \\ p(x) + t = v_{m_0}(x)^{\mathsf{T}} X^0 v_{m_0}(x) + v_{m_1}(x)^{\mathsf{T}} X^1 v_{m_1}(x)\,(x-a), \qquad \forall\, x \in \mathbb{R} \\ X^0 \succcurlyeq 0 \text{ and } X^1 \succcurlyeq 0. \end{cases} \tag{5.17}$$

Knowing p, the equality constraint in (5.17) is an affine constraint on the unknown $t \in \mathbb{R}$ and the unknown real elements of X^0 and X^1, since it is equivalent to the equality between the coefficients of the same monomials in both sides of the identity. Let us introduce the matrices $B^0_\alpha \in \mathcal{S}^{m_0+1}$ and $B^1_\alpha \in \mathcal{S}^{m_1+1}$ by the formula (5.5). Then, one has

$$\begin{aligned} v_{m_0}(x)^\mathsf{T} X^0 v_{m_0}(x) &= \sum_{\alpha\in[0\,:\,2m_0]} \langle B^0_\alpha, X^0\rangle x^\alpha, \\ v_{m_1}(x)^\mathsf{T} X^1 v_{m_1}(x)\,(x-a) &= \sum_{\alpha\in[1\,:\,2m_1+1]} \langle B^1_{\alpha-1}, X^1\rangle x^\alpha - a \sum_{\alpha\in[0\,:\,2m_1]} \langle B^1_\alpha, X^1\rangle x^\alpha. \end{aligned}$$

Set $t := u - v$ with $u \geqslant$ and $v \geqslant 0$. Then, problem (5.17) becomes

$$\left\{\begin{array}{ll} \inf_{(X^0,X^1,u,v)\in\mathcal{S}^{m_0+1}\times\mathcal{S}^{m_1+1}\times\mathbb{R}\times\mathbb{R}} u - v & \\ \langle B^0_0, X^0\rangle - a\langle B^1_0, X^1\rangle - u + v = p_0 & \\ \langle B^0_\alpha, X^0\rangle + \langle B^1_{\alpha-1} - aB^1_\alpha, X^1\rangle = p_\alpha, & \forall\,\alpha\in[1\,:\,2m_1] \\ \langle B^1_{2m_1}, X^1\rangle = 0, \quad \text{if } d \text{ is even} & \\ \langle B^0_{2m_1+1}, X^0\rangle + \langle B^1_{2m_1}, X^1\rangle = p_{2m_1+1}, & \text{if } d \text{ is odd} \\ \langle B^0_{2m_1+2}, X^0\rangle = 0, \quad \text{if } d \text{ is odd} & \\ X^0 \succcurlyeq 0,\ X^1 \succcurlyeq 0,\ u \geqslant 0, \text{ and } v \geqslant 0. & \end{array}\right. \tag{5.18}$$

To summarize, for finding the global minimum value of a polynomial of degree d, the SOS approach requires determining two positive semidefinite matrices of order $m_0 + 1$ and $m_1 + 1$, as well as a nonnegative vector of size 2, subject to $d + 2$ affine constraints.

Nonnegativity on $(-\infty, b]$

Corollary 5.1.5 (nonnegative univariate polynomial on $(-\infty, b]$) *Let b be a real number. A univariate polynomial $p \in \mathbb{R}[x]$ of degree d is nonnegative on $(-\infty, b]$ if and only if there exist σ_0 and $\sigma_1 \in \Sigma[x]$ such that $\deg\sigma_0 \leqslant d$, $\deg\sigma_1 \leqslant d-1$ and, for all $x \in \mathbb{R}$:*

$$p(x) = \sigma_0(x) + \sigma_1(x)\,(b - x). \tag{5.19}$$

Proof. Let $\tilde{p}$ be the polynomial defined at $x \in \mathbb{R}$ by $\tilde{p}(x) = p(b - x)$. Then, $\tilde{p} \geqslant 0$ on $[0, \infty)$. By proposition 5.1.3, there exist $\tilde{\sigma}_0$ and $\tilde{\sigma}_1 \in \Sigma[x]$, with $\deg\tilde{\sigma}_0 \leqslant d$ and $\deg\tilde{\sigma}_1 \leqslant d - 1$, such that $\tilde{p}(x) = \tilde{\sigma}_0(x) + \tilde{\sigma}_1(x)x$ for all $x \in \mathbb{R}$. Since $p(x) = \tilde{p}(b - x)$, it follows that

$$p(x) = \tilde{\sigma}_0(b - x) + \tilde{\sigma}_1(b - x)\,(b - x), \qquad \forall\, x \in \mathbb{R}.$$

Now the polynomials σ_0 and σ_1 defined at x by $\sigma_0(x) = \tilde{\sigma}_0(b-x)$ and $\sigma_1(x) = \tilde{\sigma}_1(b-x)$ are clearly in $\Sigma[x]$ with degrees d and $d - 1$ respectively. □

Let us now deduce from the representation (5.19) the SDCO form of the problem of minimizing a polynomial on $(\infty, b]$. Define again m_0 and m_1 by (5.16) and use the table that follows. Note that the degrees of σ_0 and σ_1 in corollary 5.1.5 are, like in the previous section, $\leqslant 2m_0$ and $\leqslant 2m_1$ respectively. Using proposition 5.1.1, the representation (5.19) of p, and setting $s = -t$, problem (5.12) can be written

$$\begin{cases} \inf_{(X^0,X^1,t)\in\mathcal{S}^{m_0+1}\times\mathcal{S}^{m_1+1}\times\mathbb{R}} t \\ p(x)+t = v_{m_0}(x)^{\mathsf{T}} X^0 v_{m_0}(x) + v_{m_1}(x)^{\mathsf{T}} X^1 v_{m_1}(x)\,(b-x), \qquad \forall\, x\in\mathbb{R} \\ X^0 \succcurlyeq 0 \text{ and } X^1 \succcurlyeq 0. \end{cases} \tag{5.20}$$

Knowing p, the equality constraint in (5.20) is an affine constraint on the unknown $t \in \mathbb{R}$ and the unknown real elements of X^0 and X^1, since it is equivalent to the equality between the coefficients of the same monomials in both sides of the identity. Let us introduce the matrices $B^0_\alpha \in \mathcal{S}^{m_0+1}$ and $B^1_\alpha \in \mathcal{S}^{m_1+1}$ by the formula (5.5). Then, one has

$$\begin{aligned} v_{m_0}(x)^{\mathsf{T}} X^0 v_{m_0}(x) &= \sum_{\alpha\in[0\,:\,2m_0]} \langle B^0_\alpha, X^0\rangle x^\alpha, \\ v_{m_1}(x)^{\mathsf{T}} X^1 v_{m_1}(x)\,(b-x) &= b \sum_{\alpha\in[0\,:\,2m_1]} \langle B^1_\alpha, X^1\rangle x^\alpha - \sum_{\alpha\in[1\,:\,2m_1+1]} \langle B^1_{\alpha-1}, X^1\rangle x^\alpha. \end{aligned}$$

Set $t := u - v$ with $u \geqslant$ and $v \geqslant 0$. Then, problem (5.20) becomes

$$\begin{cases} \inf_{(X^0,X^1,u,v)\in\mathcal{S}^{m_0+1}\times\mathcal{S}^{m_1+1}\times\mathbb{R}\times\mathbb{R}} u - v \\ \langle B^0_0, X^0\rangle + b\langle B^1_0, X^1\rangle - u + v = p_0 \\ \langle B^0_\alpha, X^0\rangle + \langle bB^1_\alpha - B^1_{\alpha-1}, X^1\rangle = p_\alpha, \qquad \forall\,\alpha\in[1\,:\,2m_1] \\ \langle B^1_{2m_1}, X^1\rangle = 0, \qquad \text{if } d \text{ is even} \\ \langle B^0_{2m_1+1}, X^0\rangle - \langle B^1_{2m_1}, X^1\rangle = p_{2m_1+1}, \qquad \text{if } d \text{ is odd} \\ \langle B^0_{2m_1+2}, X^0\rangle = 0, \qquad \text{if } d \text{ is odd} \\ X^0 \succcurlyeq 0,\ X^1 \succcurlyeq 0,\ u \geqslant 0, \text{ and } v \geqslant 0. \end{cases} \tag{5.21}$$

To summarize, for finding the global minimum value of a polynomial of degree d, the SOS approach requires determining two positive semidefinite matrices of order m_0+1 and m_1+1, as well as a nonnegative vector of size 2, subject to $d+2$ affine constraints.

Nonnegativity on $[a,b]$ ▲

Corollary 5.1.6 (nonnegative univariate polynomial on $[a,b]$) *Let a and b be real numbers and $p \in \mathbb{R}[x]$ be a univariate polynomial of degree d. Then, the following properties are equivalent:*

(i) p is nonnegative on $[a,b]$,

(ii) there exist σ_0 and $\sigma_1 \in \Sigma[x]$ such that, for all $x \in \mathbb{R}$, one has

$$p(x) = \sigma_0(x) + \sigma_1(x)(x-a)(b-x), \tag{5.22}$$

where each term has a degree $\leqslant 2\lceil d/2\rceil$,

(iii) if d is odd, there exist σ_1 and $\sigma_2 \in \Sigma[x]$ such that, for all $x \in \mathbb{R}$, one has

$$p(x) = \sigma_1(x)(x-a) + \sigma_2(x)(b-x), \tag{5.23}$$

where each term has a degree $\leqslant d$.

PROOF. □

5.1.3 Support tool

The Matlab function

```
[A,b,c,K] = polmin2sdco (p,a,b);
```

transforms in the SDCO format (A, b, c, K) the problem of finding the global minimum of the following 1D polynomial of degree d

$$\sum_{i=1}^{d+1} p_i \, x^{d+1-i},$$

on the inteval $[a, b] \subset \mathbb{R}$, along the lines explained in sections 5.1.1 and 5.1.2. In this expression, $p_i \in \mathbb{R}$ is the ith element of the Matlab vector `p`, which has length $d + 1$. The function purges p from its first zero elements and makes sure that the minimzation problem is not unbounded, by examining the sign of p_1, the parity of d, and the possible infinite value of a and b.

5.2 Rank relaxation of a QCQO problem

A quadratically constrained quadratic optimization (QCQO) problem has the following structure

$$\begin{cases} \inf_x \; x^{\mathsf{T}} A_0 x \\ x^{\mathsf{T}} A_k x = b_k, \quad \forall\, k \in E, \\ x^{\mathsf{T}} A_k x \leqslant b_k, \quad \forall\, k \in I, \end{cases}$$

where the matrices $A_k \in \mathcal{S}^n$ (for $k \in [0\,{:}\,m]$), the scalars $b_k \in \mathbb{R}$ (for $k \in [1\,{:}\,m]$), and the index sets E and I form a partition of $[1\,{:}\,m]$. Since the matrices are not assumed to be positive semidefinite, the problem is nonconvex. It is actually NP-hard.

This problem covers a large family of problems, including algebraic optimization, whose problems have polynomial objective and constraints. It also includes problems with binary constraints since $x_j \in \{0, 1\}$ can be expressed by the polynomial constraint $x_j(x_j - 1) = 0$. Whilst a QCQO problem is NP-hard, its optimal value can be approached from below by a semidefinite optimization problem.

5.2.1 QCQO formulation of an OPF problem

One of the emblematic instances of the so-called Optimal Power Flow (OPF) problem, in alternating current, consists in minimizing the Joule heating losses in an electricity transmission network, while satisfying the electricity demand, by determining appropriately the powers of the generators installed at given buses (or nodes of the network) [7, 6, 5]. This optimization problem is NP hard, implying that there is (presently) no algorithm to find the (global) solution in polynomial time (i.e., in a reasonable time when the network size is like the one in most countries or union of countries), while there are several good reasons for being only interested in the global solutions.

Structurally, one can view (or reduce) the problem to a nonconvex quadratic optimization problem with nonconvex quadratic constraints (sometimes abbreviated by the acronym QCQO for Quadratically Constrained Quadratic Programming). Recalling that any polynomial optimization problem can be written that way [45, 10, 27], one understands the potential difficulty of such a problem.

A QCQO representation of an OPF instance can be obtained by the Matlab function `qcqp_opf` [19], available in Matpower [53] (a Matlab package for simulating power systems). This function writes a specified OPF instance as a QCQO problem *in complex numbers*:

$$\begin{cases} \inf_z \ z^{\mathsf{H}} C z \\ z^{\mathsf{H}} A_k z = a_k, \quad \text{for } k \in [1\,{:}\,m] \\ z^{\mathsf{H}} B_k z \leqslant b_k, \quad \text{for } k \in [1\,{:}\,p], \end{cases} \tag{5.24}$$

where $z \in \mathbb{C}^n$ is a complex vector, C, A_k, and $B_k \in \mathbb{C}^{n\times n}$ are Hermitian matrices[1], a and b are real vectors, and the exponent $\cdot^{\mathsf{H}}$ is used to denote the conjugate transpose. In this representation, n is the number of buses (or nodes of the network).

To solve (5.24) (more precisely a relaxation of it) by an SDCO solver in real numbers, the first task is to rewrite the problem in real numbers (see the discussion in [15] on the interest in having an all-complex approach). For $M \in \mathbb{C}^{n\times n}$, we write $M = \Re(M) + \mathrm{i}\Im(M)$, where $\Re(M)$ and $\Im(M) \in \mathbb{R}^{n\times n}$, and $\mathrm{i} \in \mathbb{C}$ the pure imaginary number ($\mathrm{i}^2 = -1$). Similarly, a vector $z \in \mathbb{C}^n$ is decomposed in $z = \Re(z) + \mathrm{i}\Im(z)$, with $\Re(z)$ and $\Im(z) \in \mathbb{R}^n$. It is easy to see that

$$M \text{ is Hermitian} \quad \Longleftrightarrow \quad \begin{cases} \Re(M) \text{ is symmetric,} \\ \Im(M) \text{ is skew symmetric.} \end{cases}$$

With this notation and a Hermitian matrix M, there holds

$$z^{\mathsf{H}} M z = \tilde{z}^{\mathsf{T}} \tilde{M} \tilde{z},$$

where

$$\tilde{z} := \begin{pmatrix} \Re(z) \\ \Im(z) \end{pmatrix} \in \mathbb{R}^{2n} \qquad \text{and} \qquad \tilde{M} := \begin{pmatrix} \Re(M) & -\Im(M) \\ \Im(M) & \Re(M) \end{pmatrix} \in \mathcal{S}^{2n}.$$

Note indeed that $\tilde{M}$ is symmetric. Then, problem (5.24) becomes the QCQO in real numbers

[1] Recall that $M \in \mathbb{C}^{n\times n}$ is Hermitian if $M^{\mathsf{H}} = M$.

$$\left\{\begin{array}{ll}\inf_{\tilde{z}} & \tilde{z}^{\mathsf{T}}\tilde{C}\tilde{z}\\ & \tilde{z}^{\mathsf{T}}\tilde{A}_k\tilde{z} = a_k, \quad \text{for } k \in [1:m]\\ & \tilde{z}^{\mathsf{T}}\tilde{B}_k\tilde{z} \leqslant b_k, \quad \text{for } k \in [1:p].\end{array}\right. \tag{5.25}$$

The dimension of this problem is twice the one of (5.24).

5.2.2 Rank relaxation

The second task is to define an SDCO relaxation of (5.25). Recalling that $\langle\cdot,\cdot\rangle_{\mathcal{S}^{2n}}$ denotes the scalar product on the space of symmetric matrices (see section 1.1.1), problem (5.25) reads

$$\left\{\begin{array}{ll}\inf_{\tilde{z}} & \langle\tilde{C}, \tilde{z}\tilde{z}^{\mathsf{T}}\rangle_{\mathcal{S}^{2n}}\\ & \langle\tilde{A}_k, \tilde{z}\tilde{z}^{\mathsf{T}}\rangle_{\mathcal{S}^{2n}} = a_k, \quad \text{for } k \in [1:m]\\ & \langle\tilde{B}_k, \tilde{z}\tilde{z}^{\mathsf{T}}\rangle_{\mathcal{S}^{2n}} \leqslant b_k, \quad \text{for } k \in [1:p]\end{array}\right.$$

or

$$\left\{\begin{array}{ll}\inf_{\tilde{X}} & \langle\tilde{C}, \tilde{X}\rangle\\ & \langle\tilde{A}_k, \tilde{X}\rangle = a_k, \quad \text{for } k \in [1:m]\\ & \langle\tilde{B}_k, \tilde{X}\rangle \leqslant b_k, \quad \text{for } k \in [1:p]\\ & \tilde{X} \succcurlyeq 0\\ & \operatorname{rank}(\tilde{X}) \leqslant 1.\end{array}\right.$$

The rank constraint of this formulation is very annoying (it takes integer values, hence is discontinuous). The rank relaxation [50, 33, 31, 32, 23, 4] of the problem drops that constraint and therefore reads

$$\left\{\begin{array}{ll}\inf_{\tilde{X}} & \langle\tilde{C}, \tilde{X}\rangle\\ & \langle\tilde{A}_k, \tilde{X}\rangle = a_k, \quad \text{for } k \in [1:m]\\ & \langle\tilde{B}_k, \tilde{X}\rangle \leqslant b_k, \quad \text{for } k \in [1:p]\\ & \tilde{X} \succcurlyeq 0.\end{array}\right.$$

We still have to write the relaxed problem in the standard primal SDCO form in (1.7), which looks like the closest to the previous one. This form can be obtained by introducing p *slack variables* $v_k \in \mathbb{R}$, for $k \in [1:p]$, and by writing the problem as follows

$$\left\{\begin{array}{ll}\inf_{(\tilde{X},\underline{D})} & \langle\tilde{C}, \tilde{X}\rangle\\ & \langle\tilde{A}_k, \tilde{X}\rangle = a_k, \quad \text{for } k \in [1:m]\\ & \langle\tilde{B}_k, \tilde{X}\rangle + v_k = b_k, \quad \text{for } k \in [1:p]\\ & \tilde{X} \succcurlyeq 0\\ & v \geqslant 0,\end{array}\right. \tag{5.26}$$

which is in the primal form (1.7), in the unknown $(\tilde{X}, v)$.

5.2.3 Test cases

The real SDCO relaxation of the OPF problem, namely the standard primal SDCO form deduced in (5.26), is available by

```
[A,b,c,K,x0,y0] = testcase_5 (testcase);
```

where `testcase` is a string giving the name of a Matpower test case (see the table below for a short list), `A`, `b`, and `c` stand for the data A, b, and c of the standard primal SDCO problem in (1.7), while `x0` and `y0` are empty variables (no initial primal-dual point is provided).

The rank relaxation of a QCQO can be viewed as the Lasserre [22] (or moment-sos) relaxation of degree one [23]. For the OPF problem, when the degree one relaxation of its QCQO formulation is inexact (i.e., it does not provide the solution to the original QCQO), it is often the case that the degree 2 moment-sos relaxation is exact [20].

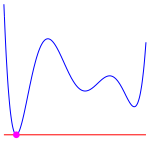
Notes

The proof of proposition 5.1.1 is taken from [38; § VI 44]. The proof of proposition 5.1.3 is taken from [39; 2000 ; proposition 2] and is a rewriting of the proof of Pólya and Szegö [38; 1976 ; § VI 45]. For more advanced technique to minimize polynomials subject to polynomial constraints (the *algebraic optimization problem*), see [37, 22, 24, 2, 3, 25].

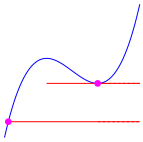
Questions

5.1. *Global minimization of a univariate polynomial.* Find the global minimum of a few univariate polynomials p on intervals $I \subset \mathbb{R}$, using the moment-SOS technique presented in sections 5.1.1 and 5.1.2, with the help of the Matlab function `polmin2sdco` described in section 5.1.3:

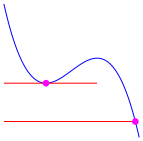
1) $p(x) = x^6 - 7x^5 + 7x^4 + 35x^3 - 56x^2 - 28x + 48$, $I = \mathbb{R}$ [43],



2) $p(x) = x^3 + 3x^2 - 9x$, $I = [-6, \infty)$ and $I = [-3, \infty)$.



3) $p(x) = -x^3 + 3x^2 + 9x$, $I = (-\infty, 6]$ and $I = (-\infty, 3]$.



References

[1] M.F. Anjos, J.B. Lasserre (2012). *Handbook on Semidefinite, Conic and Polynomial Optimization*. Springer. [doi]. 1

[2] M.F. Anjos, J.B. Lasserre (2012). *Introduction to semidefinite, conic and polynomial optimization*, volume 166 of *International Series in Operations Research & Management Science*, chapter 1. Springer. 50, 61

[3] G. Blekherman, P.A. Parrilo, R.R. Thomas (2013). *Semidefinite Optimization and Convex Algebraic Geometry*. MOS-SIAM Series on Optimization. SIAM and MPS, Philadelphia. [doi]. 50, 61

[4] I.M. Bomze, E. de Klerk (2002). Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24, 163–185. [doi]. 60

[5] F. Capitanescu (2016). Critical review of recent advances and further developments needed in AC optimal power flow. *Electric Power Systems Research*, 136, 57–68. [doi]. 59

[6] F. Capitanescu, J.L. Martinez Ramos, P. Panciatici, D. Kirschen, A. Marano Marcolini, L. Platbrood, L. Wehenkel (2011). State-of-the-art, challenges, and future trends in security-constrained optimal power flow. *Electric Power Systems Research*, 81(8), 1731–1741. [doi]. 59

[7] M.J. Carpentier (1962). Contribution à l'étude du dispatching économique. *Bulletin de la Société Française des Électriciens*, 8(3), 431–447. 59

[8] E. de Klerk (2002). *Aspects of Semidefinite Programming - Interior Point Algorithms and Selected Applications*. Kluwer Academic Publishers, Dordrecht. [doi]. 11, 22, 24, 27, 29, 30

[9] A. Deza, E. Nematollahi, R. Peyghami, T. Terlaky (2006). The central path visits all the vertices of the Klee-Minty cube. *Optimization Methods and Software*, 21, 849–863. [doi]. 22

[10] C. Ferrier (1998). Hilbert's 17th problem and best dual bounds in quadratic minimization. *Cybernetics and Systems Analysis*, 34, 696–709. [doi]. 59

[11] K. Fujisawa, M. Fukuda, M. Kojima, K. Nakata (2000). Numerical evaluation of SDPA (semidefinite programming algorithm). In H. Frenk, K. Roos, T. Terlaky, S. Zhang (editors), *High Peformance Optimization*, pages 267–301. Springer Science+Business Media, Dordrecht. 45

[12] M. Fukuda, M. Kojima, K. Murota, K. Nakata (2012). *Latest developments in the SDPA family for solving large-scale SDPs*, volume 166 of *International Series in Operations Research & Management Science*, chapter 24, pages 687–713. Springer. 44

[13] J.Ch. Gilbert (2016). *Fragments d'Optimisation Différentiable – Théorie et Algorithmes*. Syllabus de cours à l'ENSTA, Paris. [internet]. 9

[14] J.Ch. Gilbert (2019). *Step by step design of an interior-point solver in self-dual conic optimization with applications*. Lecture notes of the Master-2 "Optimization" at Paris-Saclay University, Paris, France. [hal]. 1

[15] J.Ch. Gilbert, C. Josz (2017). Plea for a semidefinite optimization solver in complex numbers. *Mathematical Programming Computation* (submitted). [hal]. 59

[16] C. Helmberg, F. Rendl, R.J. Vanderbei, H. Wolkowicz (1996). An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6(2), 342–361. [doi]. 14

[17] B. Janssen (1994). Primal-dual algorithms for linear programming based on the logarithmic barrier method. *Journal of Optimization Theory and Applications*, 83, 1–26. 25

[18] J. Jiang (1998). A long step primal-dual path following method for semidefinite programming. *Operations Research Letters*, 23(1-2), 53–62. [doi]. 22, 25

[19] C. Josz, S. Fliscounakis, J. Maeght, P. Panciatici (2015). The `Matlab` function `qcqp_opf`. Piece of software, Réseau de Transport d'Electricité, France. Personal communication. 59

[20] C. Josz, J. Maeght, P. Panciatici, J.Ch. Gilbert (2015). Application of the moment-SOS approach to global optimization of the OPF problem. *IEEE Transactions on Power Systems*, 30(1), 463–470. [doi]. 61

[21] M. Kojima, S. Shindoh, S. Hara (1997). Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM Journal on Optimization*, 7, 86–125. [doi]. 14

[22] J.B. Lasserre (2001). Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11, 796–817. [doi]. 50, 61

[23] J.B. Lasserre (2001). Convergent LMI relaxations for nonconvex quadratic programs. Technical report. Graduate seminar at MIT, fall 2001. http://www.mit.edu/~6.454/www_fall_2001/cmcaram/lasserre_1.pdf. 60, 61

[24] J.B. Lasserre (2010). *Moments Positive Polynomials and Their Applications*. Imperial College Press Optimization Series 1. Imperial College Press. 50, 61

[25] J.B. Lasserre (2015). *An Introduction to Polynomial and Semi-Algebraic Optimization*. Cambridge Texts in Applied Mathematics. Cambridge University Press. 50, 61

[26] C.-J. Lin, R. Saigal (1995). A predictor-corrector method for semi-definite linear programming. Technical Report 95-20, Department of Industrial & Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109-2117, USA. 11

[27] R. Madani, G. Fazelnia, J. Lavaei (2014). Rank-2 matrix solution for semidefinite relaxations of arbitrary polynomial optimization problems. Technical report. 59

[28] S. Mehrotra (1992). On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, 2(4), 575–601. [doi]. 42

[29] H.D. Mittelmann (2012). *The state-of-the-art in conic optimization software*, volume 166 of *International Series in Operations Research & Management Science*, chapter 23. Springer. 1

[30] S. Mizuno, M.J. Todd, Y. Ye (1993). On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18(4), 964–981. [doi]. 22

[31] Y. Nesterov (2000). Squared functional systems and optimization problems. In J.B.G. Frenk, C. Roos, T. Terlaky, S. Zhang (editors), *High Performance Optimization*, pages 405–440. Kluwer Academic Publishers. 60

[32] Y. Nesterov, H. Wolkowicz, Y. Ye (2000). Semidefinite programming relaxations of nonconvex quadratic optimization. In H. Wolkowicz, R. Saigal, L. Vandenberghe (editors), *Handbook of Semidefinite Programming – Theory, Algorithms, and Applications*, volume 27 of *International Series in Operations Research & Management Science*, chapter 13, pages 361–419. Kluwer Academic Publishers. [doi]. 60

[33] Y.E. Nesterov (1998). Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software*, 9, 141–160. [doi]. 60

[34] Y.E. Nesterov, A.S. Nemirovskii (1994). *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics 13. SIAM, Philadelphia, PA, USA. 28

[35] Y.E. Nesterov, M.J. Todd (1997). Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22(1), 1–42. [doi]. 11

[36] Y.E. Nesterov, M.J. Todd (1998). Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8(2), 324–364. [doi]. 11

[37] P.A. Parrilo (2000). On a decomposition for multivariable forms via LMI methods. In *Proceedings of the American Control Conference*. 50, 61

[38] G. Pólya, G. Szegö (1976). *Problems and Theorems in Analysis II*. Springer-Verlag, Berlin. 54, 61

[39] V. Powers, B. Reznick (2000). Polynomials that are positive on an interval. *Translations of the American Mathematical Society*, 352, 4677–4692. 54, 61

[40] F. Rendl, R. Sotirov, H. Wolkowicz (2002). A simplified/improved HKM direction for certain classes of semidefinite programming. Technical report. 14

[41] J. Renegar (2001). *A Mathematical View of Interior-Point Methods in Convex Optimization*. MPS-SIAM Series on Optimization 3. SIAM. 28

[42] S.M. Robinson (1976). Stability theory for systems of inequalities, part II: differentiable nonlinear systems. *SIAM Journal on Numerical Analysis*, 13, 497–513. [doi]. 9

[43] C. Roos (2004). Introduction to conic optimizatiion and SeDuMi – in memory of Jos Sturm (1971-2003). 61

[44] SeDuMi. [internet]. 18

[45] N.Z. Shor (1987). Class of global minimum bounds of polynomial functions. *Kibernetica*, 6, 9–11. [English translation: *Cybernetics*, 23 (1987) 731-734]. 59

[46] J.F. Sturm (1999). Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11, 625–653. 18

[47] M.J. Todd, K.-C. Toh, R.H. Tütüncü (1998). On the Nesterov-Todd direction in semidefinite programming. *SIAM Journal on Optimization*, 8, 769–796. [doi]. 11, 12

[48] M.J. Todd, K.-C. Toh, R.H. Tütüncü (1999). SDPT3 - A MATLAB software package for semidefinite programming, version 1.3. *Optimization Methods and Software*, 11/12(1-4), 545–581. [doi]. 46

[49] K.-C. Toh, M.J. Todd, R.H. Tütüncü (2012). On the implementation and usage of SDPT3 - a Matlab software package for semidefinite-quadratic-linear programming, version 4.0. In M.F. Anjos, J.B. Lasserre (editors), *Handbook on Semidefinite, Conic and Polynomial Optimization*, International Series in Operations Research and Management Science 166, pages 715–754. Springer. [doi]. 1, 18, 42, 44, 52

[50] L. Vandenberghe, S. Boyd (1996). Semidefinite programming. *SIAM Review*, 38, 49–95. [doi]. 25, 34, 47, 60

[51] H. Wolkowicz, R. Saigal, L. Vandenberghe (editors) (2000). *Handbook of Semidefinite Programming – Theory, Algorithms, and Applications*, volume 27 of *International Series in Operations Research & Management Science*. Kluwer Academic Publishers. 1

[52] S.J. Wright, Y. Zhang (1996). A superquadratic infeasible-interior-point method for linear complementarity problems. *Mathematical Programming*, 73, 269–289. 42

[53] R.D. Zimmerman, C.E. Murillo-Sánchez, R.J. Thomas (2011). MATPOWER steady-state operations, planning and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1), 12–19. [doi]. 59

Index