An interior point method for conic programming with exponential cone constraints

July 20, 2016

Abstract

We introduce the concept of exponential cone, expression of its dual cone, barrier functions, and a homogeneous interior-point method for conic programming with exponential cone constraints.

1 Definitions and preliminaries

The following notations, definitions and basic properties are adopted from [2] and [4].

Definition 1. Denote $\mathbb{R}_+ = \{x \mid x \in \mathbb{R}, x \geq 0\}$ and $\mathbb{R}_{++} = \{x \mid x \in \mathbb{R}, x > 0\}$. For a vector $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$, denote the diagonal matrix with entries $x_1, \cdots, x_n$ as $\text{diag}(x)$. A subset $\mathcal{K}$ of an Euclidean space $\mathbb{R}^n$ is a *(pointed) convex cone* if (1) $\mathbf{0} \in \mathcal{K}$, (2) for any $\alpha, \beta \in \mathbb{R}_+$ and $x, y \in \mathcal{K}$, one has $\alpha x + \beta y \in \mathcal{K}$, and (3) the only linear subspace of $\mathbb{R}^n$ contained in $\mathcal{K}$ is $\{0\}$. A *proper cone* is a convex cone that is pointed, closed and has non-empty interior in the usual Euclidean space topology.

Definition 2. For a convex cone $\mathcal{K} \in \mathbb{R}^n$, its *dual cone* is defined as

$$\mathcal{K}^* = \left\{y \in \mathbb{R}^n \mid x^T y \geq 0, \forall x \in \mathcal{K}\right\}.$$

A proper cone $\mathcal{K}$ is called *self-dual* if $\mathcal{K}^* = \mathcal{K}$.

Proposition 1. *For a proper cone $\mathcal{K} \in \mathbb{R}^n$, its dual cone $\mathcal{K}^*$ is also a proper cone and $(\mathcal{K}^*)^* = \mathcal{K}$.*

Definition 3. The Lorentz cone of dimension $n \geq 2$ is defined as

$$\mathcal{Q}_n = \left\{x \in \mathbb{R}^n \mid x_1 \geq \sqrt{x_2 + \cdots + x_n}\right\}.$$

Note that there are slightly different definitions of the Lorentz cone, namely, some authors prefer interchanging the role of x_1 and x_n.

Proposition 2. *Both the positive orthant and the Lorentz cone are self-dual. In other words, $(\mathbb{R}_+^n)^* = \mathbb{R}_+^n$ and $(\mathcal{Q}_m)^* = \mathcal{Q}_m$ for $n \geq 1$, $m \geq 2$.*

The positive orthant, the Lorenz cone and the positive semidefinite cone $\mathcal{S}_+^n = \{X \in \mathcal{S}^n \mid X \succeq 0\}$, (where $\mathcal{S}^n$ denotes the symmetric matrices in $\mathbb{R}^{n \times n}$) are known as *symmetric cones* (there are other less interesting types of symmetric cones).

Definition 4. Let

$$\mathcal{K}_{\exp}^0 = \left\{ (x, y, z) \in \mathbb{R}^3 \mid y \geq 0, z > 0, \exp\left(\frac{x}{z}\right) \leq \frac{y}{z} \right\}$$

and define the *exponential cone* as the closure of $\mathcal{K}_{\exp}^0$ in $\mathbb{R}^3$,

$$\mathcal{K}_{\exp} = \text{closure}\left(K_{\exp}^0\right).$$

Proposition 3. *One has* $\mathcal{K}_{\exp} = K_{\exp}^0 \cup [(-\mathbb{R}_+) \times \mathbb{R}_+ \times \{0\}]$.

Definition 5. Let $\mathcal{P}_{\exp}^0 = \left\{ (u, v, w) \in \mathbb{R}^3 \mid u < 0, v \geq 0, \exp\left(\frac{w}{u}\right) \leq -\frac{e \cdot v}{u} \right\}$ and define $\mathcal{P}_{\exp} = \text{closure}\left(\mathcal{P}_{\exp}^0\right)$, where e denotes the base of natural logarithm.

Proposition 4. *One has* $\mathcal{P}_{\exp} = \mathcal{P}_{\exp}^0 \cup \{0\} \times \mathbb{R}_+ \times \mathbb{R}_+$ *and* $(\mathcal{K}_{\exp})^* = \mathcal{P}_{\exp}$.

We will be primarily working with the above three types of proper cones, namely, the positive orthants $\mathbb{R}_+^n$, the Lorentz cone $\mathcal{Q}^m$ and the exponential cone $\mathcal{K}_{\exp}$. In other words, the interior-point method we present deals with problems with a conic constraint $s \in \mathcal{K}$, where $\mathcal{K}$ is a Cartesian product of these three types of cones.

In theory, all polynomial-time interior-point methods for conic programming problems require a well-defined and computable (up to the Hessian) barrier function for the cone.

Definition 6. Let $\mathcal{K} \subset \mathbb{R}^n$ be a closed convex set with $\text{int}(\mathcal{K}) \neq \emptyset$, and $f : \text{int}(\mathcal{K}) \to \mathbb{R}$ be strictly convex. The function f is called a *self-concordant barrier* for $\mathcal{K}$ if

1. (barrier property) for any sequence of points $\{x_n\}$ such that $\lim x_n \in \mathcal{K} \backslash \text{int}(\mathcal{K})$, one has $\lim f(x_n) = \infty$,

2. (differentiability) f has continuous third-order mixed partial derivatives, and

3. (self-concordance) for any $x \in \text{int}(\mathcal{K})$ and $v \in \mathbb{R}^n$, the function $\phi(t) = f(x + tv)$ satisfies $|\phi'''(t)| \leq M_f \left(\phi''(t)\right)^{2/3}$ for some constant $M_f \geq 0$ and all t such that $tv \in \text{int}(\mathcal{K})$.

Definition 7. Let $\mathcal{K}$ be a proper cone. The function $f : \text{int}(\mathcal{K}) \to \mathbb{R}$ is called a logarithmically homogeneous self-concordant barrier (LHSCB) for $\mathcal{K}$ with parameter ν if f is a self-concordant barrier for $\mathcal{K}$ and for all $x \in \text{int}(\mathcal{K})$ and $t > 0$, one has

$$f(tx) = f(x) - \nu \log t.$$

Below is an important property of logarithmically homogeneous self-concordant barriers (see Theorem 4.3.4 in [4]).

Proposition 5. *Let $\mathcal{K}$ be a proper cone and f a LHSCB for $\mathcal{K}$ with parameter ν. The conjugate barrier of f, denoted as f^*, defined on* $\text{int}(\mathcal{K}^*)$ *by*

$$f^*(s) = \sup\left\{ -s^T x - f(x) \mid x \in \text{int}\,(\mathcal{K}) \right\}$$

is a LHSCB for $\mathcal{K}^*$ with parameter ν. Furthermore, one has $(f^*)^* = f$.

Note that the definition of conjugate barrier in this context is slightly different from the definition of convex conjugate adopted in standard references of convex analysis.

We list some useful barriers for the important types of cones.

Proposition 6. *(Section 2 of [7]) The function*

$$f_l(x) = -\sum_{i=1}^{n} \log x_i$$

is a LHCSB for the nonnegative orthant $\mathbb{R}^n_+$ with $\nu = n$. Its gradient and Hessian are

$$\nabla g_l(x) = f_l(x) = \begin{bmatrix} -\dfrac{1}{x_1} \\ \vdots \\ -\dfrac{1}{x_n} \end{bmatrix}, \quad H_l(x) = \nabla^2 f_l(x) = \begin{bmatrix} \dfrac{1}{x_1^2} & & \\ & \ddots & \\ & & \dfrac{1}{x_n^2} \end{bmatrix}.$$

The function

$$f_q(x) = -\frac{1}{2}\log\left(x_1^2 - ||\tilde{x}||^2\right)$$

where $\tilde{x} = (x_2, \cdots, x_n)^T$ and $||\cdot|| = ||\cdot||_2$ is a LHSCB for $\mathcal{Q}_n$ with $\nu = 2$. The gradient and Hessian of f_q are

$$g_q(x) = -\frac{1}{x^T J x} J x, \quad H_q(x) = \frac{1}{(x^T J x)^2}\left(2 J x x^T J - (x^T J x) J\right)$$

where $J = \begin{bmatrix} 1 & \\ & -I_{n-1} \end{bmatrix}.$

In [4], the author listed two different pairs of conjugate barriers (LHSCB) for the exponential cone and its dual cone.

Definition 8. The Wright Omega function $\omega : \mathbb{R} \to \mathbb{R}_{++}$ is defined as the unique solution to the equation

$$\omega(\beta) + \log \omega(\beta) = \beta$$

for $\beta \in \mathbb{R}$.

Proposition 7. *The function $f_{\exp} : \operatorname{int}(\mathcal{K}_{\exp}) \to \mathbb{R}$ defined by*

$$f_{\exp}(x, y, z) = -\log\left(z \log\left(\frac{y}{z}\right) - x\right) - \log y - \log z$$

is a LHSCB for $\mathcal{K}_{\exp}$ with $\nu = 3$. Its conjugate $f_{\exp}^* : \operatorname{int}(\mathcal{P}_{\exp}) \to \mathbb{R}$, which is a LHSCB for $(\mathcal{K}_{\exp})^* = \mathcal{P}_{\exp}$ with $\nu = 3$, takes the following expression

$$f_{\exp}^*(u, v, w) = -2\log(-u) - \log v - \log\left(\frac{(1 - \bar{\omega})^2}{\bar{\omega}}\right) - 3$$

where

$$\bar{\omega} = \omega\left(2 - \frac{w}{u} - \log(-u) + \log v\right).$$

Note that the dual barrier $f_{\exp}^*$ above has a slightly more complicated expression than the primal barrier $f_{\exp}$. It is not too difficult to work out another pair of conjugate barrier functions where the dual barrier takes a slightly simpler expression instead. In fact, there exists an invertible linear transformation $B : \mathbb{R}^3 \to \mathbb{R}^3$ such that $B(\operatorname{int}(\mathcal{K}_{\exp})) = \operatorname{int}(\mathcal{P}_{\exp})$, which is given by the matrix

$$B = \begin{bmatrix} & & -1 \\ & \frac{1}{e} & \\ -1 & & \end{bmatrix}.$$

As such, $\tilde{f}_{\exp}(x, y, z) = f_{\exp}^*(B(x, y, z))$ is a LHSCB for $\mathcal{K}_{\exp}$ and $\tilde{f}_{\exp}^*(u, v, w) = f_{\exp}(B^{-1}(u, v, w))$ is its conjugate barrier, which is a LHSCB for $\mathcal{P}_{\exp}$. Note that by the definition of conjugate barrier and the property of LHSCB, we may also take $\tilde{f}_{\exp} + \beta$ and $\left(\tilde{f}_{\exp} + \beta\right)^* = \tilde{f}_{\exp}^* - \beta$ as the LHSCBs for $\mathcal{K}_{\exp}$ and $\mathcal{P}_{\exp}$, respectively, for any $\beta \in \mathbb{R}$.

Proposition 8. *The function $\tilde{f}_{\exp} : \operatorname{int}(\mathcal{K}_{\exp}) \to \mathbb{R}$ defined by*

$$\tilde{f}_{\exp}(x, y, z) = -2\log z - \log y - \log\left(\frac{(1 - \bar{\omega})^2}{\bar{\omega}}\right) - 3$$

where

$$\bar{\omega} = \omega\left(1 - \frac{x}{z} - \log z - \log y\right)$$

is a LHSCB for $\mathcal{K}_{\exp}$ with $\nu = 3$. Its conjugate $\tilde{f}_{\exp}^* : \operatorname{int}(\mathcal{P}_{\exp}) \to \mathbb{R}$, which is a LHSCB for $(\mathcal{K}_{\exp})^* = \mathcal{P}_{\exp}$ with $\nu = 3$, takes the following expression

$$\tilde{f}_{\exp}^*(u, v, w) = -\log(-t) - \log(-u) - \log v$$

where $r = \log(-v/u)$ and $t = u - w + ur$. Furthermore, the gradient and Hessian of the dual barrier at $(u, v, w)^T \in \operatorname{int}(\mathcal{P}_{\exp})$ are

4

$$\tilde{g}^* \left((u,v,w)^T \right) = \nabla \tilde{f}^*_{\exp}(u,v,w) = \begin{bmatrix} -\dfrac{r}{t} - \dfrac{1}{u} \\[2ex] -\dfrac{1}{v} - \dfrac{u}{vt} \\[2ex] \dfrac{1}{t} \end{bmatrix},$$

$$\tilde{H}^* \left((u,v,w)^T \right) = \nabla^2 \tilde{f}^*_{\exp}(u,v,w) = \begin{bmatrix} \dfrac{1}{ut} + \dfrac{r^2}{t^2} + \dfrac{1}{u^2} & \dfrac{ur}{vt^2} - \dfrac{1}{vt} & -\dfrac{r}{t^2} \\[2ex] \dfrac{ur}{vt^2} - \dfrac{1}{vt} & \dfrac{1}{v^2} + \dfrac{u}{v^2 t} + \dfrac{u^2}{v^2 t^2} & -\dfrac{u}{vt^2} \\[2ex] -\dfrac{r}{t^2} & -\dfrac{u}{vt^2} & \dfrac{1}{t^2} \end{bmatrix}.$$

Note that once we obtain the expressions of the LHSCB, their gradients and Hessians for the positive orthant, the Lorentz cone and the exponential cone, we can easily construct LHSCB, their gradients and Hessians for any Cartesian product of these cones through vertical and block-diagonal concatenation.

2 Standard forms of conic programming

For different purposes, there have been different standard forms in conic programming. The simplest formulation is the following pair of primal and dual problems denoted as (PD),

$$\begin{aligned}
\text{Primal:} \quad & \min \ c^T x \\
& \text{s.t. } Ax = b, \ x \in \mathcal{K} \\
\text{Dual:} \quad & \max \ b^T y \\
& \text{s.t. } A^T y + z = c, \ s \in \mathcal{K}^*, \ y \in \mathbb{R}^m
\end{aligned} \tag{PD}$$

where $c, x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $\mathcal{K} \subset \mathbb{R}^n$ is a proper cone. Without loss of generality, we further assume $m \leq n$ and $\text{rank}(A) = m$. By the generalized Slater's condition (see [1]), if there exist $x \in \text{int}(\mathcal{K})$ such that $Ax = b$ and $z \in \text{int}(\mathcal{K}^*)$ such that $A^T y + z = c$, then strong duality holds for (PD). In this case, any primal-dual optimal solution (x, y, s) must satisfy the following KKT system

$$\begin{aligned}
Ax - b &= 0 \\
A^T y + z - c &= 0 \\
x^T z &= 0 \\
x \in \mathcal{K}_{\exp}, \ z \in \mathcal{P}_{\exp}, & \ y \in \mathbb{R}^m
\end{aligned} \tag{1}$$

We introduce the (full) *homogeneous self-dual embedding* of (PD), which is similar to

(25) in [6] and (27) in [7],

$$\min \ \bar{\alpha}\theta$$

$$\text{s.t.} \begin{bmatrix} 0 & -A & b & -\bar{b} \\ A^T & 0 & -c & \bar{c} \\ -b^T & c^T & 0 & -\bar{g} \\ \bar{b}^T & -\bar{c}^T & \bar{g} & 0 \end{bmatrix} \begin{bmatrix} y \\ x \\ \tau \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ z \\ \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \bar{\alpha} \end{bmatrix} \qquad \text{(HSD)}$$

$$x \in \mathcal{K}, \ z \in \mathcal{K}^*, \ \tau \geq 0, \ \kappa \geq 0, \ y \in \mathbb{R}^m, \ \theta \in \mathbb{R},$$

where the auxillary parameters $\bar{b}, \bar{c}, \bar{g}$ and $\bar{\alpha}$ are determined by a given $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ such that $x^0 \in \text{int}(\mathcal{K})$, $z^0 \in \text{int}(\mathcal{K}^*)$ and $\tau^0, \kappa^0, \theta^0 > 0$, namely,

$$\bar{b} = \frac{1}{\theta^0}\left(b\tau^0 - Ax^0\right), \qquad \bar{c} = \frac{1}{\theta^0}\left(c\tau^0 - A^Ty^0 - z^0\right),$$

$$\bar{g} = \frac{1}{\theta^0}\left(c^Tx^0 - b^Ty + \kappa\right), \qquad \bar{\alpha} = \frac{1}{\theta^0}\left((x^0)^Tz^0 + \tau^0\kappa^0\right).$$

The properties of (HSD) are summarized in the following proposition.

Proposition 9. (Lemma 1 in [3]) *For any given $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ such that $x^0 \in \text{int}(\mathcal{K}_{\exp})$, $z^0 \in \text{int}(\mathcal{P}_{\exp})$ and $\tau^0, \kappa^0, \theta^0 > 0$, the auxiliary parameters $\bar{b}$, $\bar{c}$, $\bar{g}$, $\bar{\alpha}$ and hence the problem (HSD) are well-defined. One has $\bar{\alpha} > 0$. Further more, the following facts hold.*

1. *The problem (HSD) is self-dual. In other words, the dual of the problem is equivalent to itself.*

2. *$(x, y, z, \tau, \kappa, \theta) = (x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ is a strictly feasible (primal and dual) solution.*

3. *The (attained) optimal objective is 0.*

4. *Assume the solution $(x, y, \tau, z, \kappa, \theta)$ is feasible. One has $\theta \geq 0$ and $x^Tz + \tau\kappa = \bar{\alpha}\theta$. Furthermore, the solution is optimal if and only if $\theta = 0$, in which case one has $x^Tz = \tau\kappa = 0$.*

5. *Assume $(x, y, \tau, z, \kappa, 0)$ is an optimal solution. If $\tau > 0$ then $(x, y, z)/\tau$ is an optimal solution to (PD). If $\kappa > 0$ then either $b^Ty > 0$ or $c^Tx < 0$ or both hold.*

   If $b^Ty > 0$ then (PD) is primal-infeasible.

   If $c^Tx < 0$ then (PD) is dual infeasible.

6. *(Proposition 3 in [3]) For any $\epsilon \geq 0$, there exists a feasible solution of (HSD) with objective value equal to ϵ.*

Note that if $(y, x, \tau, z, \kappa, 0)$ is an optimal solution to (HSD) with $\tau = \kappa = 0$, no conclusion can be made regarding (PD).

For implementation purpose, we will adopt the standard form in [6] which can be written in the form of (PD). Specifically, we consider the following pair of primal and dual problem

$$\text{Primal:} \quad \min \sum_{i=1}^{N} c_i^T x_i$$

$$\text{s.t.} \sum_{i=1}^{N} A_i x_i = b$$

$$x_i \in K_i, \ \forall i \tag{PD'}$$

$$\text{Dual:} \quad \max \ b^T y$$

$$\text{s.t.} \ A_i^T y + z_i = c_i, \ \forall i$$

$$y \in \mathbb{R}^m, \ z_i \in K_i^*, \forall i$$

where $x_i, c_i \in \mathbb{R}^{n_i}$, $A_i \in \mathbb{R}^{m \times n_i}$, $b \in \mathbb{R}^m$ and $n = \sum_{i=1}^{N} n_i \geq m$. Each K_i is one of the following: (i) the nonnegative orthant $\mathbb{R}_+^{n_i}$, (ii) product of Lorentz cones $\mathcal{Q}^{q_1} \times \cdots \times \mathcal{Q}^{q_{k_i}}$, $\sum_{j=1}^{k_i} q_j = n_i$, (iii) product of the exponential cone $(\mathcal{K}_{\exp})^{k_i}$, $3k_i = n_i$ or (iv) $\mathbb{R}_i^n$. Note that we define (which is consistent with the definition of dual cone) $\left(\mathbb{R}^k\right)^* = \left\{ y \in \mathbb{R}^k \mid y^T x \geq 0, \ \forall x \in \mathbb{R}^k \right\} = \{0\}^k$. To write (PD') into (PD), one simply let

$$A = \begin{bmatrix} A_1, \ \ldots, \ A_N \end{bmatrix}, \ x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}, \ c = \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix}, \ z = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}, \ \mathcal{K} = K_1 \times \cdots \times K_N, \ \mathcal{K}^* = K_1^* \times \cdots \times K_N^*.$$

For notational convenience, vertical concatenation of matrices (or vectors) of appropriate dimensions $(M_1^T, \ M_2^T, \ \cdots, \ M_L^T)^T$ will occasionally be denoted $[M_1; \cdots; M_L]$.

3 A homogeneous interior-point algorithm

Next we describe the algorithm for solving (HSD) (and hence find either a solution to (PD) or certificates of infeasibility) in detail. Henceforth the gradients of the barriers for $\mathcal{K}$ and $\mathcal{K}^*$ will be denoted as g and g^* (in particular, $\mathcal{K}_{\exp}$ and $\mathcal{P}_{\exp}$) respectively, and the Hessians H and H^* respectively.

3.1 Initialization

Suppose one has a problem in the form (PD'). First, for all i with $K_i = \mathbb{R}^{n_i}$, define $x_i^{new} = [x_i^+; x_i^-]$, $c_i^{new} = [c_i; -c_i]$, $A_i^{new} = [A_i, -A_i]$, $K_i^{new} = \mathbb{R}_+^{n_i^{new}}$, $n_i^{new} = 2n_i$ and replace c_i^{new}, A_i, K_i, n_i by c_i^{new}, A_i^{new}, K_i^{new} and n_i^{new} respectively. After the above adjustment and necessary rearrangement (grouping together the same types of K_i), one has $\mathcal{K} = \mathbb{R}_+^{n_l} \times \mathcal{Q}_{n_{q_1}} \times \cdots \times \mathcal{Q}_{n_{q_s}} \times (\mathcal{K}_{\exp})^{k_e}$, $n_q = n_{q_1} + \ldots + n_{q_s}$, $n_e = 3k_e$.

Next we set the iterate $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ such that $x^0 \in \text{int}(\mathcal{K})$, $z^0 \in \text{int}(\mathcal{K}^*)$, $y^0 \in \mathbb{R}^m$, $\tau^0 > 0$, $\kappa^0 > 0$, $\theta^0 > 0$. One strategy is to set $y^0 = 0$, $\tau^0 = \kappa^0 = \theta^0 = 1$. For

x^0 and z^0, we will take the analytical centers of the respective barrier function of $\mathcal{K}$ (and equivalently that of $\mathcal{K}^*$). In other words, we first find a LHSCB for the product cone $\mathcal{K}$ as the sum of its parts using the following fact.

Proposition 10. *Suppose $f_1, \cdots, f_N$ are LHSCB for proper cones $\mathcal{K}_1, \cdots, \mathcal{K}_N$ respectively, then $f : \mathcal{K}_1 \times \cdots \times \mathcal{K}_N \to \mathbb{R}$ defined by $f([x_1; \cdots; x_N]) = f_1(x_1) + \cdots + f_N(x_N)$ is a LHSCB for $\mathcal{K}_1 \times \cdots \times \mathcal{K}_N$.*

Since we have found analytic expressions for the LHSCB for the nonnegative orthant, the Lorentz cone, the exponential cone and its dual cone, we have obtained a LHSCB f for the product cone $\mathcal{K}$ and a LHSCB f^* for $\mathcal{K}^*$. Note that we choose the second pair of barrier functions $\tilde{f}_{\exp}$ for $\mathcal{K}_{\exp}$ and $\tilde{f}^*_{\exp}$ for $\mathcal{K}_{\exp}$. The reason is that the barrier function for the dual exponential cone $\tilde{f}^*_{\exp}$, whose gradient and Hessian will be frequently evaluated in the implementation, takes a slightly simpler expression than $f^*_{\exp}$. Denote the gradient and Hessian of f as g and H, and those of f^* and g^* and H^* respectively.

Similar to the initialization strategy in [4], we find $x^0 \in \mathcal{K}$ such that $x^0 = -g(x^0)$. By the property of LHSCB (see 10.2.2 in [2]), the vector $z^0 = x^0$ will also satisfy $z^0 = -g^*(z^0)$. For $\mathcal{K}$ defined above, let $\iota_l = (1, \cdots, 1)^T \in \mathbb{R}^{n_l}$, $\iota_q = [\iota_{q_1}; \cdots; \iota_{q_s}]$ where $\iota_{q_j} = (1, 0, \cdots, 0)^T \in \mathbb{R}^{n_{q_j}}$, $\iota_e = [\iota_e^0; \cdots; \iota_e^0] \in \mathbb{R}^{n_e}$ where $\iota_e^0 = (-1.0151, 1.2590, 0.5560)^T$ and set

$$x^0 = z^0 = [\iota_l; \iota_q; \iota_e].$$

Note that $\iota^0 \in \text{int}(\mathcal{K}_{\exp})$ and $||\iota^0 + \tilde{g}(\iota^0)|| \approx 0$. Also compute the auxillary parameters $\bar{b}, \bar{c}, \bar{g}, \bar{\alpha}$ to fully construct (HSD).

3.2 A typical iteration

Henceforth the current and next iterates are denoted as $(x, y, \tau, z, \kappa, \theta)$ and $(x^+, y^+, \tau^+, z^+, \kappa^+, \theta^+)$ respectively. Similarly we also use $(\cdot)^+$ to denote the updated values of the parameters for the next iteration. Similar to Algorithm 6 in Chapter 10 of [4], we set the *central-path parameter* $\mu = \theta$, the *central-path proximity measure*

$$\Omega(x, z, \tau, \kappa) = (\nu + 1) \log \left(\frac{x^T z + \tau \kappa}{\nu + 1} \right) + (\nu + 1) + f(x) + f^*(z) - \log(\tau) - \log(\kappa)$$

and the *neighborhood threshold* $\bar{\Omega} = \Omega(x^0, z^0, \tau^0, \kappa^0) + 1$. Solve the following system for the *predictor search direction* $(\Delta x, \Delta y, \Delta z, \Delta \tau, \Delta \kappa, \Delta \theta)$.

$$\begin{bmatrix} 0 & -A & b & -\bar{b} \\ A^T & 0 & -c & \bar{c} \\ -b^T & c^T & 0 & -\bar{g} \\ \bar{b}^T & -\bar{c}^T & \bar{g} & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta \tau \\ \Delta \theta \end{bmatrix} + \begin{bmatrix} 0 \\ \Delta z \\ \Delta \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2}$$

$$\tau \Delta \kappa + \kappa \Delta \tau = -\tau \kappa \tag{3}$$

$$\theta H^*(z) \Delta z + \Delta x = -x. \tag{4}$$

Note that the above system (2) - (4) can be compactly written as $\bar{G}\Delta\bar{x}_p = \bar{R}_p$, where

$$
\bar{G} = \begin{bmatrix}
-A & 0 & 0 & b & 0 & -\bar{b} \\
0 & A^T & I & -c & 0 & \bar{c} \\
c^T & -b^T & 0 & 0 & 1 & -\bar{g} \\
-\bar{c}^T & \bar{b}^T & 0 & \bar{g} & 0 & 0 \\
0 & 0 & 0 & \kappa & \tau & 0 \\
I & 0 & \theta H^*(z) & 0 & 0 & 0
\end{bmatrix}, \quad
\Delta\bar{x}_p = \begin{bmatrix}
\Delta x_p \\
\Delta y_p \\
\Delta z_p \\
\Delta\tau_p \\
\Delta\kappa_p \\
\Delta\theta_p
\end{bmatrix}, \quad
\bar{R}_p = \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
-\tau\kappa \\
-x
\end{bmatrix}. \tag{5}
$$

Set the *maximum step-length* based on the predictor search direction and the neighborhood characterization

$$
\alpha_p = \max\left\{\alpha > 0 \mid \Omega(x + \alpha\Delta x_p, z + \alpha\Delta z_p, \tau + \alpha\Delta\tau_p, \kappa + \alpha\Delta\kappa_p) \le \bar{\Omega}\right\}.
$$

Alternatively, one may find sufficiently large $\alpha_p > 0$ such that $x + \alpha_p\Delta x_p \in \text{int}\,(\mathcal{K}_{\exp})$, $z + \alpha_p\Delta z_p \in \text{int}\,(\mathcal{P}_{\exp})$, $\tau + \alpha_p\Delta\tau_p > 0$, $\kappa + \alpha_p\Delta\kappa_p > 0$. Set $\sigma = \max\left\{(1 - \alpha_p)^3, 0\right\}$.

Solve the following system for the (combined) search direction $(\Delta x, \Delta y, \Delta z, \Delta\tau, \Delta\kappa, \Delta\theta)$.

$$
\begin{bmatrix}
0 & -A & b & -\bar{b} \\
A^T & 0 & -c & \bar{c} \\
-b^T & c^T & 0 & -\bar{g} \\
\bar{b}^T & -\bar{c}^T & \bar{g} & 0
\end{bmatrix}
\begin{bmatrix}
\Delta y \\
\Delta x \\
\Delta\tau \\
\Delta\theta
\end{bmatrix}
+
\begin{bmatrix}
0 \\
\Delta z \\
\Delta\kappa \\
0
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
0 \\
0
\end{bmatrix}
\tag{6}
$$

$$
\tau\Delta\kappa + \kappa\Delta\tau = -\tau\kappa + \sigma\theta - \Delta\tau_p\Delta\kappa_p \tag{7}
$$

$$
\mu H^*(z)\Delta z + \Delta x = -x - \sigma\theta g^*(z). \tag{8}
$$

Note that similarly (6) - (8) can be written as $\bar{G}\Delta\bar{x} = \bar{R}$, where $\bar{G}$ is defined in (5) and

$$
\Delta\bar{x} = \begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta z \\
\Delta\tau \\
\Delta\kappa \\
\Delta\theta
\end{bmatrix}, \quad
\bar{R} = \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
-\tau\kappa + \sigma\theta - \Delta\tau_p\kappa_p \\
-x - \sigma\theta g^*(z)
\end{bmatrix}.
$$

Set the *step-length* based on the combined search direction and the neighborhood characterization

$$
\alpha_c = \max\left\{\alpha > 0 \mid \Omega(x + \alpha\Delta x, z + \alpha\Delta z, \tau + \alpha\Delta\tau, \kappa + \alpha\Delta\kappa) \le \bar{\Omega}\right\}.
$$

Alternatively, one may find $\alpha_c > 0$ such that $x + \alpha_c\Delta x \in \text{int}\,(\mathcal{K}_{\exp})$, $z + \alpha_c\Delta z \in \text{int}\,(\mathcal{P}_{\exp})$, $\tau + \alpha_c\Delta\tau > 0$, $\kappa + \alpha_c\Delta\kappa > 0$. Set $\alpha' = 0.98\alpha_c$ and the next iterate

$$
(x^+, y^+\tau^+, z^+, \kappa^+, \theta^+) = (x, y, \tau, z, \kappa, \theta) + \alpha'(\Delta x, \Delta y, \Delta z, \Delta\tau, \Delta\kappa, \Delta\theta).
$$

3.3 Termination conditions

A   A simplified homogeneous self-dual model

Below we describe a *simplified* homogeneous self-dual embedding (SHSD) for (PD) and a path-following algorithm that solves (SHSD) and consequently solves or detects infeasibility of (PD). Consider the following feasibility problem

$$\min\ 0$$
$$\text{s.t.}\ \begin{bmatrix} 0 & A & -b \\ -A^T & 0 & c \\ b^T & -c^T & 0 \end{bmatrix} \begin{bmatrix} y \\ x \\ \tau \end{bmatrix} - \begin{bmatrix} 0 \\ z \\ \kappa \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{(SHSD)}$$
$$x \in \mathcal{K},\ z \in \mathcal{K}^*,\ \tau \geq 0,\ \kappa \geq 0,\ y \in \mathbb{R}^m.$$

It can be shown that the above problem is self-dual (Lemma 1 in [5]). Similar to Proposition 9, the following facts hold (Lemma 2 in [5]).

Proposition 11. *Assume* (x, y, z, τ, κ) *solves (SHSD).*

1. *One has* $x^T z + \tau \kappa = 0$ *(complimentarity).*

2. *If* $\tau > 0$ *then* $(x, y, z)/\tau$ *is an optimal solution to (PD).*

3. *If* $\kappa > 0$ *then either* $b^T y > 0$ *or* $c^T x < 0$ *or both hold.*
    If $b^T y > 0$ *then (PD) is primal-infeasible.*
    If $c^T x < 0$ *then (PD) is dual infeasible.*

Next we describe the algorithm in moderate detail. For notational convenience, denote

$$G = \begin{bmatrix} 0 & A & -b \\ -A^T & 0 & c \\ b^T & -c^T & 0 \end{bmatrix}.$$

A.1   Initialization

Given an initial point $(x^0, y^0, z^0, \tau^0, \kappa^0)$ such that $x^0 \in \text{int}(\mathcal{K})$, $z^0 \in \text{int}(\mathcal{K}^*)$, $\tau > 0$ and $\kappa > 0$, calculate the residual of the linear constraints

$$R^0 = G \begin{bmatrix} y^0 \\ x^0 \\ \tau^0 \end{bmatrix} - \begin{bmatrix} 0 \\ z^0 \\ \kappa^0 \end{bmatrix}.$$

Define $\mu^0 = \left((x^0)^T z^0 + \tau^0 \kappa^0\right) / (\nu + 1)$. When $\mathcal{K} = (\mathcal{K}_{\exp})^N$ (only exponential cone constraints are present), one has $\nu = 3N$. Take a pair of barriers f for $\mathcal{K}$ and f^* for $\mathcal{K}^*$. Denote their gradients and Hessians as g, g^* and H, H^* respectively. For this algorithm, we take $f = \tilde{f}_{\exp}$ and $f^* = \tilde{f}^*_{\exp}$ (stacked N times if necessary when necessary). It can be shown

that for any $\mu \in (0, \mu^0]$, there exists a unique vector $\bar{x}(\mu) := (x_\mu, y_\mu, z_\mu, \tau_\mu, \kappa_\mu)$ that satisfies

$$G \begin{bmatrix} y \\ x \\ \tau \end{bmatrix} - \begin{bmatrix} 0 \\ z \\ \kappa \end{bmatrix} = \mu R^0$$

$$\tau \kappa = \mu$$

$$x - \mu g^*(z) = 0$$

$$x \in \mathcal{K}, \ z \in \mathcal{K}^*, \ \tau \geq 0, \ \kappa \geq 0, \ y \in \mathbb{R}^m.$$

Define the *central-path* parametrized by μ as $\Gamma = \left\{ \bar{x}(\mu) \mid \mu \in (0, \mu^0] \right\}$. It can be shown that as $\mu \to 0$, $\bar{x}(\mu)$ converges to a solution to (SHSD).

A.2 A typical iteration

We use $(\cdot)$ and $(\cdot)^+$ to denote the current and updated values of a varaible or parameter. First, set $\mu = \left(x^T z + \tau \kappa \right) / (\nu + 1)$, calculate the current residual of the linear constraints

$$R = G \begin{bmatrix} y \\ x \\ \tau \end{bmatrix} - \begin{bmatrix} 0 \\ z \\ \kappa \end{bmatrix}$$

and solve the following system for the predictor search direction $(\Delta x_p, \Delta y_p, \Delta z_p, \Delta \tau_p, \Delta \kappa_p)$.

$$G \begin{bmatrix} \Delta y_p \\ \Delta x_p \\ \Delta \tau_p \end{bmatrix} - \begin{bmatrix} 0 \\ \Delta z_p \\ \Delta \kappa_p \end{bmatrix} = -R$$

$$\tau \Delta \kappa_p + \kappa \Delta \tau_p = -\tau \kappa \tag{9}$$

$$\mu H^*(z) \Delta z_p + \Delta x_p = -x.$$

Note that we can write (9) compactly as $\bar{G} \Delta \bar{x}_p = \bar{R}$ where

$$\bar{G} = \begin{bmatrix} A & 0 & 0 & -b & 0 \\ 0 & -A^T & -I & c & 0 \\ -c^T & b^T & 0 & 0 & -1 \\ 0 & 0 & 0 & \kappa & \tau \\ I & 0 & \mu H^*(z) & 0 & 0 \end{bmatrix}, \quad \Delta \bar{x}_p = \begin{bmatrix} \Delta x_p \\ \Delta y_p \\ \Delta z_p \\ \Delta \tau_p \\ \Delta \kappa_p \end{bmatrix}, \quad \bar{R}_p = \begin{bmatrix} -R \\ -\tau \kappa \\ -x \end{bmatrix}. \tag{10}$$

We will make use of the same central-path proximity measure Ω defined in Section 3.2 and take the threshold $\bar{\Omega} = \Omega \left(x^0, z^0, \tau^0, \kappa^0 \right) + 1$. Note that for any $x \in \text{int}(\mathcal{K})$, $z \in \text{int}(\mathcal{K}^*)$, $\tau, \kappa > 0$, one has $\Omega(x, z, \tau, \kappa) \geq 0$. Furthermore, $\Omega(x, z, \tau, \kappa) = 0$ if and only if (x, z, τ, κ) is on the central-path. Find

$$\alpha_{\max} = \max \left\{ \alpha > 0 \mid \Omega(x + \alpha \Delta x_p, z + \alpha \Delta z_p, \tau + \alpha \Delta \tau_p, \kappa + \alpha \Delta \kappa_p) \leq \bar{\Omega} \right\}.$$

Note that the maximum is attained due to the compactness of the set. Set $\sigma = \max\left\{(1 - \alpha_{\max})^3, 0\right\}$. Solve the following system for the combined search direction $(\Delta x, \Delta y, \Delta z, \Delta \tau, \Delta \kappa)$.

$$G\begin{bmatrix} \Delta y \\ \Delta x \\ \Delta \tau \end{bmatrix} - \begin{bmatrix} 0 \\ \Delta z \\ \Delta \kappa \end{bmatrix} = -(1-\sigma)\,R$$

$$\tau\Delta\kappa + \kappa\Delta\tau = -\tau\kappa + \sigma\mu - \Delta\tau_p\Delta\kappa_p$$

$$\mu H^*(z)\Delta z + \Delta x = -x - \sigma\mu g^*(z).$$

(11)

Similarly, (11) can be written compactly as $\bar{G}\Delta\bar{x} = \bar{R}$, where $\bar{G}$ is definedd This in (10) and

$$\Delta\bar{x} = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta \tau \\ \Delta \kappa \end{bmatrix}, \quad \bar{R} = \begin{bmatrix} -(1-\sigma)\,R \\ -\tau\kappa + \sigma\mu - \Delta\tau_p\Delta\kappa_p \\ -x - \sigma\mu g^*(z) \end{bmatrix}.$$

(12)

Find $\alpha'_{\max} = \max\left\{\alpha > 0 \mid \Omega(x + \alpha\Delta x, z + \alpha\Delta z, \tau + \alpha\Delta\tau, \kappa + \alpha\Delta\kappa) \leq \bar{\Omega}\right\}$. Set the step-length $\alpha' = 0.98\alpha'_{\max}$. Update the iterate

$$(x^+, y^+, z^+, \tau^+, \kappa^+) = (x, y, z, \tau, \kappa) + \alpha'(\Delta x, \Delta y, \Delta z, \Delta\tau, \Delta\kappa).$$

Repeat until a termination condition is met.

A.3 Termination conditions

We are interested in either a solution or a certificate of infeasibility of (PD). Similar to Section 5.4 in [5], there are 3 possibilities.

1. We declare optimality (primal and dual feasibility with an approximate optimal solution found) if

$$\|Ax - \tau b\|_\infty \leq \epsilon \cdot \max\left\{1, \|[A, b]\|_\infty\right\}$$

(13)

$$\|A^T y + z - c\tau\|_\infty \leq \epsilon \cdot \max\left\{1, \|A^T, I, -c\|_\infty\right\}$$

(14)

$$\left|c^T x/\tau - b^T y/\tau\right| \leq \epsilon \cdot \left(1 + \left|b^T y/\tau\right|\right).$$

(15)

In this case, the approximate optimal solution $(x, y, z)/\tau$ is returned.

2. We declare primal and/or dual infeasibility if (13), (14) and

$$\left|-c^T x + b^T y - \kappa\right| \leq \epsilon \cdot \max\left\{1, \left\|\left[-c^T, b^T, 1\right]\right\|_\infty\right\}$$

(16)

$$\tau \leq \epsilon \cdot 10^{-2} \cdot \{1, \kappa\}$$

(17)

hold. If $b^T y > 0$ ($c^T x < 0$, respectively), the problem is primal infeasible (dual infeasible, respectively).

3. We declare that the problem is ill-posed if

$$\kappa \leq \epsilon \cdot 10^{-2} \cdot \min\{1, \tau\}$$

$$\mu \leq \epsilon \cdot 10^{-2} \cdot \mu^0.$$

References

[1] S. Boyd and L. Vandenberghe. *Convex Optimization*, chapter 5. Cambridge University Press, 2009.

[2] P. R. Charles. *Cones and interior-point algorithms for structured convex optimization involving powers and exponentials.* PhD thesis, Universite catholique de Louvain and Ecole Polytechnique de Louvain, 2007.

[3] R.M. Freund. On the behavior of the homogeneous self-dual model for conic convex optimization. 106:527–545, 2005.

[4] S. A. Serrano. *Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone.* PhD thesis, Stanford University, 2015.

[5] A. Skarjaa and Y. Ye. A homogeneous interior-point algorithm for nonsymmetric convex conic optimization. *Mathematical Programming, Series A*, 150:391–422, 2015.

[6] K.C. Toh, M.J. Todd, and R.H. Tütüncü. On the implementation and usage of sdpt3 – a matlab software package for semidefinite-quadratic-linear programming, version 4.0.

[7] L. Vandenberghe. The cvxopt linear and quadratic cone program solvers.