Design and implementation of a homogeneous interior-point method for conic programming involving exponential cone constraints

Yuan GAO

October 11, 2016

Abstract

We introduce the concept of exponential cone, its basic mathematical properties and designed a homogeneous interior-point method for conic programming involving exponential cone constraints. We tested the implementation on stochastic and robust optimization problems. Techniques of modeling with exponential cone are also discussed.

1 Definitions and preliminaries

In order to begin the discussion on properties of the exponential cone and the homogeneous interior-point method, we summarize the relevant concepts and facts for easy reference. For a more detailed survey on (non-symmetric) conic programming and interior-point methods, see [2] and [4].

Denote $\mathbb{R}_+ = \{x \mid x \in \mathbb{R}, x \geq 0\}$ and $\mathbb{R}_{++} = \{x \mid x \in \mathbb{R}, x > 0\}$. For matrices $M_1, \cdots, M_N$, denote $\text{diag}(M_1, \cdots, M_N)$ the block-diagonal matrix with M_j as diagonal blocks. The relative interior of a set $\mathcal{S} \subset \mathbb{R}^n$ is denoted as $\text{ri}(\mathcal{S})$. 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 relative interior in the usual Euclidean space topology.

For a convex cone $\mathcal{K} \in \mathbb{R}^n$, its *dual cone* is defined as $\mathcal{K}^* = \{y \in \mathbb{R}^n \mid x^T y \geq 0, \forall x \in \mathcal{K}\}$. A proper cone $\mathcal{K}$ is called *self-dual* if $\mathcal{K}^* = \mathcal{K}$. It can be shown that for a proper cone $\mathcal{K} \in \mathbb{R}^n$, its dual cone $\mathcal{K}^*$ is also a proper cone and $(\mathcal{K}^*)^* = \mathcal{K}$ (Section 2.6. in [1]).

For our purpose, we define the Lorentz cone of dimension $n \geq 2$ as

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

It can be easily seen that both the positive orthant and the Lorentz cone are self-dual. 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 all proper and self-dual. They are sometimes referred to as *symmetric cones*.

Next we define the exponential cone and give analytic expressions of it and its dual cone.

Definition 1. 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 $\mathcal{K}_{\exp} = \text{closure}\left(K_{\exp}^0\right)$. 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 1. *One has* $\mathcal{K}_{\exp} = K^0_{\exp} \cup [(-\mathbb{R}_+) \times \mathbb{R}_+ \times \{0\}]$, $\mathcal{P}_{\exp} = \mathcal{P}^0_{\exp} \cup \{0\} \times \mathbb{R}_+ \times \mathbb{R}_+$ *and* $(\mathcal{K}_{\exp})^* = \mathcal{P}_{\exp} \neq \mathcal{K}_{\exp}$.

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 arbitrary number 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 2. Let $\mathcal{K} \subset \mathbb{R}^n$ be a closed convex set with $\text{int}(\mathcal{K}) \neq \emptyset$, and $f : \text{ri}(\mathcal{K}) \to \mathbb{R}$ 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 $\lim x_n \in \mathcal{K} \backslash \text{ri}(\mathcal{K})$, one has $\lim f(x_n) = \infty$,

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

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

Definition 3. Let $\mathcal{K}$ be a proper cone. The function $f : \text{ri}(\mathcal{K}) \to \mathbb{R}$ 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 $x \in \text{ri}(\mathcal{K})$ and $t > 0$, one has

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

Below is an important property of LHSCB (see Theorem 4.3.4 in [4]).

Proposition 2. *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* $\text{ri}(\mathcal{K}^*)$ *by*

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

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 3. (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 4. 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 4. *The function $f_{\exp} : \mathrm{int}\,(\mathcal{K}_{\exp}) \to \mathbb{R}$ 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 $f_{\exp}^* : \mathrm{int}\,(\mathcal{P}_{\exp}) \to \mathbb{R}$, 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 $B\,(\mathrm{int}\,(\mathcal{K}_{\exp})) = \mathrm{int}\,(\mathcal{P}_{\exp})$, 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 LHSCB for $\mathcal{K}_{\exp}$ and $\mathcal{P}_{\exp}$ respectively, for any $\beta \in \mathbb{R}$.

Proposition 5. *The function $\tilde{f}_{\exp} : \mathrm{ri}\,(\mathcal{K}_{\exp}) \to \mathbb{R}$ 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 $\tilde{f}^*_{\exp} : \mathrm{int}(\mathcal{P}_{\exp}) \to \mathbb{R}$, 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 $(u, v, w)^T \in \mathrm{ri}(\mathcal{P}_{\exp})$ take the following expressions

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

$$\tilde{H}^*\left((u, v, w)^T\right) = \nabla^2 \tilde{f}^*_{\exp}(u, v, w) = \begin{bmatrix} \dfrac{1}{ut} + \dfrac{r^2}{t^2} + \dfrac{1}{u^2} & \dfrac{ur}{vt^2} - \dfrac{1}{vt} & -\dfrac{r}{t^2} \\[3mm] \dfrac{ur}{vt^2} - \dfrac{1}{vt} & \dfrac{1}{v^2} + \dfrac{u}{v^2 t} + \dfrac{u^2}{v^2 t^2} & -\dfrac{u}{vt^2} \\[3mm] -\dfrac{r}{t^2} & -\dfrac{u}{vt^2} & \dfrac{1}{t^2} \end{bmatrix}.$$

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.

Proposition 6. *Suppose f_j are ν_j-LHSCB for proper cones $\mathcal{K}_j$, $j = 1, \cdots, N$ 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 ν-LHSCB for $\mathcal{K}_1 \times \cdots \times \mathcal{K}_N$ with $\nu = \sum_{j=1}^{N} \nu_j$. Furthermore, $\nabla f = [\nabla f_1; \cdots; \nabla f_N]$ and $\nabla^2 f = \mathrm{diag}(\nabla^2 f_1, \nabla^2 f_2, \cdots, \nabla^2 f_N)$.*

2 The standard form, homogeneous self-dual embedding and central path

For different purposes, various standard forms of conic programming have been used. In this paper, we consider the following pair of primal and dual problems

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

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 $m \le n$ and $\mathrm{rank}(A) = m$. By the generalized Slater's condition (see [1]), if there exist $x \in \mathrm{int}(\mathcal{K})$ such that $Ax = b$ and $z \in \mathrm{ri}(\mathcal{K}^*)$ such that $A^T y + z = c$, then strong duality holds

for (PD). In this case, any primal-dual optimal solution (x, y, z) must satisfy the following KKT system

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

We introduce the (full) *homogeneous self-dual embedding* of (PD), which is similar to (25) in [6] and (27) in [7],

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

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 $x^0 \in \mathrm{ri}\,(\mathcal{K})$, $z^0 \in \mathrm{ri}\,(\mathcal{K}^*)$, $\tau^0, \kappa^0, \theta^0 > 0$, namely,

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

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

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

Proposition 7. (Lemma 1 in [3]) *For any given $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ such that $x^0 \in \mathrm{ri}\,(\mathcal{K}_{\exp})$, $z^0 \in \mathrm{ri}\,(\mathcal{P}_{\exp})$ 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 optimal objective is always 0.*

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

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 $b^T y > 0$ 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.

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

Next we introduce the concept of *central path*, which is essential in developing a path-following algorithm for solving (HSD). First we need a simple lemma which is a consequence of the properties of LHSCB (see, e.g. (2.43) in [2]).

5

Lemma 1. *Let $\mathcal{K} \in \mathbb{R}^n$ be a proper cone with ν-LHSCB f, $\mathcal{K}^*$ its dual cone with ν-LHSCB f^*. Denote $g = \nabla f$ and $g^* = \nabla f^*$. For any $x \in \mathrm{ri}\,(\mathcal{K})$, $z \in \mathrm{ri}\,(\mathcal{K}^*)$ and $\mu > 0$, one has $x + \mu g^*(z) = 0$ if and only if $\mu g(x) + z = 0$, in which case $x^T z = \nu \mu$.*

Now we formally define the central path parametrized by $\mu > 0$ in the following proposition. The proof can be found in [4].

Proposition 8. *With the same setting as Lemma 1, for any $\mu > 0$, there exists unique $(x, y, z, \tau, \kappa, \theta)$ with $x \in \mathrm{ri}\,(\mathcal{K})$, $z \in \mathrm{ri}\,(\mathcal{K}^*)$, $\tau > 0$, $\kappa > 0$ that solves the following linear system*

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

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

$$\tau \kappa = \mu.$$

Denote the above unique solution $(x, y, z, \tau, \kappa, \theta)$ as $\bar{x}_\mu$. The set $\{\bar{x}_\mu\}_{\mu > 0}$ is termed *central path*. It can be shown that as $\mu \to 0$, $\bar{x}_\mu$ converges to a KKT point of (PD) (a solution of (1)).

For $(x, y, z, \tau, \kappa, \theta)$ with $x \in \mathrm{ri}\,(\mathcal{K})$, $z \in \mathrm{ri}\,(\mathcal{K}^*)$, $\tau > 0$, $\kappa > 0$ and $\mu > 0$, define the *predictor search direction* $\Delta \bar{x}^p$ as the solution to the following linear system (where Δv and Δv^p have the same dimension as v)

$$
\begin{bmatrix} 0 & -A & b & -\bar{b} \\ A^T & 0 & -c & \bar{c} \\ -b^T & c^T & 0 & -\bar{g} \\ \bar{b}^T & -\bar{c}^T & \bar{g} & 0 \end{bmatrix} \begin{bmatrix} \Delta y^p \\ \Delta x^p \\ \Delta \tau^p \\ \Delta \theta^p \end{bmatrix} + \begin{bmatrix} 0 \\ \Delta z^p \\ \Delta \kappa^p \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3}
$$

$$\Delta x^p + \mu H^*(z) \Delta z^p = -x$$

$$\tau \Delta \kappa^p + \kappa \Delta \tau^p = -\tau \kappa.$$

and the *combined search direction* $\Delta \bar{x}$ as the solution to the following linear system

$$
\begin{bmatrix} 0 & -A & b & -\bar{b} \\ A^T & 0 & -c & \bar{c} \\ -b^T & c^T & 0 & -\bar{g} \\ \bar{b}^T & -\bar{c}^T & \bar{g} & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta \tau \\ \Delta \theta \end{bmatrix} + \begin{bmatrix} 0 \\ \Delta z \\ \Delta \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4}
$$

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

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

where $\sigma \in [0, 1]$ is a chosen combination parameter. Note that (3) is in fact (4) with $\sigma = 0$.

If we write (2) as $\Phi(\bar{x}, \mu) = 0$, where $\Phi : \mathbb{R}^{2n+m+3} \times \mathbb{R} \to \mathbb{R}^{2n+m+3}$ is a smooth vector-valued map. Viewing $\bar{x} = \bar{x}_\mu$ as a vector-valued function of μ, by the Implicit Function Theorem, it can be shown that when $\bar{x} = \bar{x}_\mu$ for some $\mu > 0$, the direction Δx^p defined by (3) is tangent to the central path. In other words, $\Delta x^p = -\mu \cdot \bar{x}'_\mu$, where $\bar{x}'_\mu$ denotes the derivative of $\bar{x}_\mu$ with respect to μ (differentiated component-wise).

For $\bar{x} = (x, y, z, \tau, \kappa, \theta) \in \bar{K} := \mathrm{ri}\,(\mathcal{K}) \times \mathbb{R}^m \times \mathrm{ri}\,(\mathcal{K}^*) \times \mathbb{R}_+^2 \times \mathbb{R}$ and $\mu > 0$, $\Delta \bar{x}$ defined by (4) with $\sigma = 0$ will be the same as Δx^p defined by (3), while (4) with $\sigma = 1$ gives Δx^{corr} known as the

centering search direction. The term arises since it can be shown that moving along this direction decreases the *central path proximity measure* of the iterate defined as

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

It can be shown that for any $\bar{x} \in \mathrm{ri}\left(\bar{K}\right)$, ones has $\Omega(\bar{x}) \geq 0$. Furthermore, $\bar{x}$ lies on the central path if and only if $\Omega(\bar{x}) = 0$, in which case if (4) with $\sigma = 1$ gives a centering direction $\Delta\bar{x} = 0$. For $\sigma \in (0,1)$, (4) gives $\Delta\bar{x} = (1-\sigma)\Delta x^p + \sigma\Delta x^{corr}$, a combination of the two directions.

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

$$
\begin{aligned}
\text{Primal:} \quad & \min \sum_{i=1}^{N} c_i^T x_i \\
& \text{s.t. } \sum_{i=1}^{N} A_i x_i = b \\
& \qquad x_i \in K_i, \ \forall i
\end{aligned}
\tag{PD$'$}
$$

$$
\begin{aligned}
\text{Dual:} \quad & \max \ b^T y \\
& \text{s.t. } A_i^T y + z_i = c_i, \ \forall i \\
& \qquad y \in \mathbb{R}^m, \ z_i \in K_i^*, \forall i
\end{aligned}
$$

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) $\mathbb{R}^{n_i}$. Note that we define (which is consistent with the definition of dual cone) $\left(\mathbb{R}^k\right)^* = \left\{y \in \mathbb{R}^k \mid y^T x \geq 0, \ \forall x \in \mathbb{R}^k\right\} = \{0\}^k$. When there is no $K_i = \mathbb{R}^{n_j}$, to write (PD$'$) into (PD), one simply let

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

When some $K_i = \mathbb{R}^{n_i}$, we will discuss in the following section how it can be processed so that (PD$'$) can be written into (PD). 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 as $[M_1; \cdots; M_L]$.

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, 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 $\mathcal{K} = \mathbb{R}_+^{n_l} \times \mathcal{Q}_{n_{q_1}} \times \cdots \times \mathcal{Q}_{n_{q_s}} \times (\mathcal{K}_{\exp})^{k_e}$, $n_q = n_{q_1} + ... + n_{q_s}$, $n_e = 3k_e$.

Next we set the iterate $(x^0, y^0, z^0, \tau^0, \kappa^0, \theta^0)$ such that $x^0 \in \mathrm{ri}\,(\mathcal{K})$, $z^0 \in \mathrm{ri}\,(\mathcal{K}^*)$, $y^0 \in \mathbb{R}^m$, $\tau^0 > 0$, $\kappa^0 > 0$, $\theta^0 > 0$. 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 simple fact. 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 found a LHSCB f for the product cone $\mathcal{K}$. For the initial iterate, take $y = 0$, $\tau^0 = \kappa^0 = \theta^0 = 1$, $x \in \mathrm{ri}\,(\mathcal{K})$, $z \in \mathrm{ri}\,(\mathcal{K}^*)$. Specifically, let $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_N$ where each $\mathcal{K}_j$ is either a nonnegative orthant, a Lorentz cone or an exponential cone. For each j, if $\mathcal{K}_j = \mathbb{R}_+^{n_j}$, set the corresponding sub-vectors $\left(x^0\right)_j = \left(z^0\right)_j = (1, \cdots, 1)^T \in \mathbb{R}_+^{n_j}$; if $\mathcal{K}_j = \mathcal{Q}^{n_j}$, set $\left(x^0\right)_j = \left(z^0\right)_j = (1, 0, \cdots, 0)^T \in \mathcal{Q}^{n_j}$; if $\mathcal{K} = \mathcal{K}_{\exp}$, set $\left(x^0\right)_j = \left(z^0\right)_j = \iota^0$, where $\iota^0 = (-1.0151, 1.2590, 0.5560)^T$ is the approximate solution to $\iota + \tilde{g}^*(\iota) = 0$. Finally, compute the auxillary parameters $\bar{b}$, $\bar{c}$, $\bar{g}$, $\bar{\alpha}$ and set the initial central-path parameter $\mu^0 = \theta^0$.

3.2   A typical iteration

Henceforth the current and next iterates are denoted as $\bar{x} = (x, y, \tau, z, \kappa, \theta)$ and $\bar{x}^+$ respectively.

Given the current iterate $\bar{x}$, compute $\mu = \dfrac{x^T z + \tau\kappa}{\nu}$, where $\nu = \sum_{j=1}^N \nu_j$ is the parameter for the LHSCB for $\mathcal{K}$. Note that with the above initialization, it always holds that $\mu = \theta$.

Solve for the predictor search direction $\Delta\bar{x}^p$ from (3) and approximately find

$$\alpha^p = \sup\left\{\alpha > 0 \mid \bar{x} + \alpha\Delta\bar{x}^p \in \mathrm{ri}\left(\bar{K}\right)\right\}.$$

Note that we are not maintaining a central path neighborhood in addition to the underlying conic constraints.

Set $\sigma = \max\left\{(1-\alpha^p)^3, 0\right\}$ and solve for the combined search direction $\Delta\bar{x}$ from (4).

$$\alpha^c = \left\{\alpha > 0 \mid \bar{x} + \alpha\Delta\bar{x}^p \in \mathrm{ri}\left(\bar{K}\right)\right\}.$$

Set $\alpha' = 0.98\alpha^c$ and update the iterate

$$\bar{x}^+ = \bar{x} + \alpha'\Delta\bar{x}.$$

For approximately finding α^p (and similarly α^c), we adopt a bisection strategy and find $\alpha^p = \beta \cdot \sup\left\{\alpha > 0 \mid \bar{x} + \alpha\Delta\bar{x}^p \in \mathrm{ri}\left(\bar{K}\right)\right\}$ for some $\beta \in (0.95, 1]$. To solve the linear systems (3) and (4), we adopt the strategy in [6] of first solving a Schur complement equation.

3.3   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

$$\|Ax - \tau b\|_\infty \le \epsilon \cdot \max\{1, \|[A, b]\|_\infty\} \tag{5}$$

$$\|A^T y + z - c\tau\|_\infty \le \epsilon \cdot \max\{1, \|A^T, I, -c\|_\infty\} \tag{6}$$

$$|c^T x/\tau - b^T y/\tau| \le \epsilon \cdot (1 + |b^T y/\tau|). \tag{7}$$

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

2. We declare primal and/or dual infeasibility if (5), (6) and

$$|-c^T x + b^T y - \kappa| \le \epsilon \cdot \max\{1, \|[-c^T, b^T, 1]\|_\infty\} \tag{8}$$

$$\tau \le \epsilon \cdot 10^{-2} \cdot \{1, \kappa\} \tag{9}$$

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

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

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

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.