

Largest dual ellipsoids inscribed in dual cones

M. J. Todd*

June 23, 2005

Abstract

Suppose \bar{x} and \bar{s} lie in the interiors of a cone K and its dual K^* respectively. We seek dual ellipsoidal norms such that the product of the radii of the largest inscribed balls centered at \bar{x} and \bar{s} and inscribed in K and K^* respectively is maximized. Here the balls are defined using the two dual norms. We provide a solution when the cones are symmetric, that is self-dual and homogeneous. This provides a geometric justification for the Nesterov-Todd primal-dual scaling in symmetric cone programming.

*School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York 14853, USA (miketodd@cs.cornell.edu). This author was supported in part by NSF through grant DMS-0209457 and ONR through grant N00014-02-1-0057.

1 Introduction

Let K be a solid, pointed, closed, convex cone in the finite-dimensional real vector space E . That is, K has a nonempty interior and contains no line. We call such a cone regular. Let E^* be the dual space, with scalar product $\langle \cdot, \cdot \rangle : E^* \times E \rightarrow \mathbb{R}$, and let K^* be the dual cone,

$$K^* := \{s \in E^* : \langle s, x \rangle \geq 0 \text{ for all } x \in K\}.$$

(The reader may think of E and E^* as \mathbb{R}^n and the scalar product as the usual dot inner product without loss, but the more abstract viewpoint is convenient when considering other settings such as the space of symmetric matrices, and has the virtue of clearly distinguishing the primal and dual spaces.) Suppose we are given points $\bar{x} \in \text{int } K$ and $\bar{s} \in \text{int } K^*$. We seek large ellipsoids centered at \bar{x} and \bar{s} and inscribed in K and K^* respectively.

Our motivation comes from convex optimization. Consider the problem

$$\min \langle s_0, x \rangle, \text{ subject to } x \in L + \{x_0\}, x \in K, \quad (1.1)$$

where L is a linear subspace of E . This is a linear programming problem if K is the nonnegative orthant, a second-order cone programming problem if K is a second-order or Lorentz cone, or the product of such cones, and a semidefinite programming problem if K is the cone of symmetric positive semidefinite matrices. See, e.g., Vandenberghe and Boyd [17], Lewis and Overton [8], Lobo et al. [9], Ben Tal and Nemirovski [2], Alizadeh and Goldfarb [1], and [16].

The corresponding dual problem is

$$\min \langle s, x_0 \rangle, \text{ subject to } s \in L^* + \{s_0\}, s \in K^*, \quad (1.2)$$

with $L^* \subseteq E^*$ the orthogonal complement of L . If x and s are feasible solutions to these problems, then by orthogonality of L and L^* and the definition of K^* we have the weak duality inequality

$$\langle s_0, x \rangle + \langle s, x_0 \rangle = \langle s_0, x_0 \rangle + \langle s, x \rangle \geq \langle s_0, x_0 \rangle,$$

so that optimality of both x and s follows if equality holds.

If \bar{x} and \bar{s} are feasible in these problems, we would like to obtain improved feasible solutions. Since the optimization of a linear function over the intersection of an affine subspace and an ellipsoid is straightforward, we can consider the restrictions

$$\min \langle s_0, x \rangle, \text{ subject to } x \in L + \{\bar{x}\}, x \in C, \quad (1.3)$$

and

$$\min \langle s, x_0 \rangle, \text{ subject to } s \in L^* + \{\bar{s}\}, s \in C^*, \quad (1.4)$$

where C and C^* are ellipsoids centered at \bar{x} and \bar{s} that inscribe K and K^* respectively. Many interior-point methods are based on moving towards the solutions of such restrictions, where often the objective functions are replaced by penalized functions including also some multiple of a barrier function for the cone K or K^* .

If we consider these two problems separately, then we seek (unrelated) “large” ellipsoids C and C^* satisfying these conditions. A natural sense for largeness is to maximize the volumes of the ellipsoids, since this measure is invariant (up to a multiplicative constant) under linear transformations. The resulting optimal ellipsoids are then easy to obtain for many standard cones, for instance

$$\{x \in \mathbb{R}^n : \|\bar{X}^{-1}(x - \bar{x})\|_2 \leq 1\}$$

for the nonnegative orthant, where $\bar{X} := \text{Diag}(\bar{x})$, or

$$\{x \in SR^{n \times n} : \|\bar{x}^{-1/2}(x - \bar{x})\bar{x}^{-1/2}\|_F \leq 1\}$$

for the positive semidefinite cone, where $SR^{n \times n}$ denotes the space of symmetric square matrices of order n , $\bar{x}^{-1/2}$ denotes the positive definite square root of the positive definite matrix \bar{x}^{-1} , and subscript “F” denotes the Frobenius norm. In general, if we have a self-concordant barrier F (see Section 2) for K , then

$$\{x \in E : \langle F''(\bar{x})(x - \bar{x}), x - \bar{x} \rangle \leq 1\}$$

will be a suitable ellipsoid; it is optimal for many standard cones if the standard barrier is used, but note that if we choose, say,

$$F(x_1; x_2) := -2 \ln x_1 - \ln x_2$$

as the barrier for the nonnegative orthant in \mathbf{R}^2 , we obtain a suboptimal ellipsoid.

However, if we wish to generate iterates in both primal and dual spaces (and computational experience indicates this is usually worthwhile), treating (1.3) and (1.4) separately requires the solution of two ellipsoidal subproblems. We would like to choose C and C^* so that both subproblems can be solved for the price of one. This is possible if C and C^* are balls defined by dual ellipsoidal norms.

Let $H : E \rightarrow E^*$ be a self-adjoint ($\langle Hv, w \rangle = \langle Hw, v \rangle$ for all $v, w \in E$) and positive definite ($\langle Hv, v \rangle > 0$ for all nonzero $v \in E$) linear operator. Then H defines a norm in E by

$$\|v\|_H := \langle Hv, v \rangle^{1/2}$$

and a dual norm in E^* by

$$\|u\|_H^* := \langle u, H^{-1}u \rangle^{1/2}.$$

These are dual norms because it is easy to show that

$$\|u\|_H^* = \max\{\langle u, v \rangle : \|v\|_H \leq 1\} \quad \text{and} \quad \|v\|_H = \max\{\langle u, v \rangle : \|u\|_H^* \leq 1\}.$$

Let

$$C = B_H(\bar{x}; \alpha_{\bar{x}}) := \{x \in E : \|x - \bar{x}\|_H \leq \alpha_{\bar{x}}\} \tag{1.5}$$

and

$$C^* = B_H^*(\bar{s}; \alpha_{\bar{s}}) := \{s \in E^* : \|s - \bar{s}\|_H^* \leq \alpha_{\bar{s}}\}. \tag{1.6}$$

Then, if $L =: \{x : Ax = 0\}$, with A a surjective linear operator from E to some finite-dimensional real vector space D with dual D^* , the solution to (1.3) is given by

$$x = \bar{x} - \alpha_{\bar{x}} \frac{(H^{-1} - H^{-1}A^*(AH^{-1}A^*)^{-1}AH^{-1})s_0}{\langle s_0, (H^{-1} - H^{-1}A^*(AH^{-1}A^*)^{-1}AH^{-1})s_0 \rangle^{1/2}} \tag{1.7}$$

and that to (1.4) by

$$s = \bar{s} - \alpha_{\bar{s}} \frac{A^*(AH^{-1}A^*)^{-1}Ax_0}{\langle A^*(AH^{-1}A^*)^{-1}Ax_0, x_0 \rangle^{1/2}}, \tag{1.8}$$

where $A^* : D^* \rightarrow E^*$ is the adjoint to A , defined by $\langle A^*y, x \rangle := \langle y, Ax \rangle$. These can both be obtained cheaply once (a matrix representation of) $AH^{-1}A^*$ is formed and factored.

Note that, if C and C^* are defined in this way, then

$$\text{vol } C = \alpha_{\bar{x}}^n \cdot (\det \mathbf{H})^{-1/2} \cdot \psi,$$

where n is the dimension of E , \mathbf{H} is the matrix representing H with respect to some dual bases of E and E^* , and ψ is the volume of a certain ball in E , and similarly

$$\text{vol } C^* = \alpha_{\bar{s}}^n \cdot (\det \mathbf{H}^{-1})^{-1/2} \cdot \psi^*,$$

where ψ^* is likewise the volume of a certain ball in E^* . Hence

$$\text{vol } C \cdot \text{vol } C^* = (\alpha_{\bar{x}}\alpha_{\bar{s}})^n \cdot (\psi\psi^*),$$

so that it is equivalent to maximize the product of the volumes of C and C^* and to maximize the product of their radii $\alpha_{\bar{x}}$ and $\alpha_{\bar{s}}$. Also observe that if H is multiplied by a positive scalar τ , then $\alpha_{\bar{x}}$ is multiplied by τ and $\alpha_{\bar{s}}$ by its inverse, and the product stays the same: thus we can equivalently maximize the minimum of $\alpha_{\bar{x}}$ and $\alpha_{\bar{s}}$.

In this paper we consider the problem of choosing H to maximize the product

$$\alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)$$

where $\alpha_{\bar{x}}(H)$ is the largest $\alpha_{\bar{x}}$ such that $B_H(\bar{x}; \alpha_{\bar{x}})$ in (1.5) is contained in K , and similarly $\alpha_{\bar{s}}(H)$ is the largest $\alpha_{\bar{s}}$ such that $B_H^*(\bar{s}; \alpha_{\bar{s}})$ in (1.6) is contained in K^* . The solution to this problem in the case of the nonnegative orthant was basically noted in Section 4 of [15], and used as a motivation for the primal-dual scaling there. Here we obtain an optimal solution to the problem stated above when K is symmetric, that

is, self-dual and homogeneous, and we also obtain a lower bound, tight in many standard cases, for the product in terms of another geometric quantity defined from \bar{x} and \bar{s} in this case. This provides further motivation for the Nesterov-Todd scaling [11, 12] in a primal-dual interior-point algorithm for a symmetric cone programming problem. If K is a hyperbolicity cone, we suggest what may be a good choice for H .

A number of papers (see Renegar [13, 14] and Freund [4, 5], e.g.) have considered the geometry of the feasible region or the optimal set in conic optimization problems, usually to obtain complexity bounds for algorithms depending on some geometric condition number of the problem. These works address a given problem, rather than the situation at each iteration. Moreover, the measures considered usually depend on the linear constraints and the objective function, whereas we are concerned only with the current primal and dual iterates and the cones.

In Section 2, we define the cones and barriers we will be working with and describe the properties of these cones that we need. Section 3 contains our main results, and in Section 4 we discuss the case of hyperbolicity cones and relate the product of $\alpha_{\bar{x}}(H)$ and $\alpha_{\bar{s}}(H)$ to the improvements in the objective functions when we take steps to the optimal solutions of (1.3) and (1.4).

2 Preliminaries

We call the cone K *self-dual* if there is a linear bijection $J : E \rightarrow E^*$ with

$$J(K) = K^*;$$

we call it *homogeneous* if, for every pair of points x and z in $\text{int } K$, there is a linear bijection $G : E \rightarrow E$ with

$$G(K) = K \quad \text{and} \quad Gx = z.$$

We call G an *automorphism of K* if it satisfies just the first of these conditions. Cones that are self-dual and homogeneous are called *symmetric* and have been studied in depth: see Faraut and Koranyi [3].

A *self-concordant barrier with parameter θ* for a regular cone K (see Nesterov and Nemirovski [10]) is a function $F : \text{int } K \rightarrow \mathbb{R}$ satisfying

- (i) F is C^3 and convex;
- (ii) $F(x) \rightarrow +\infty$ if $x \rightarrow \hat{x} \in \partial K$;
- (iii) $D^2F(x)$ is positive definite for all $x \in \text{int } K$;
- (iv) $|D^3F(x)[h, h, h]| \leq 2(D^2F(x)[h, h])^{3/2}$ for all $x \in \text{int } K$ and all $h \in E$; and
- (v) $|DF(x)[h]| \leq (\theta D^2F(x)[h, h])^{1/2}$ for all $x \in \text{int } K$ and all $h \in E$.

Such a function F is *θ -logarithmically homogeneous* if

- (vi) $F(\tau x) = F(x) - \theta \ln \tau$ for all $x \in \text{int } K$ and all $\tau > 0$.

It can be shown that (i)–(iv) and (vi) imply (v).

The standard barriers

$$F(x) := - \sum_j \ln(x_j)$$

for the nonnegative orthant in \mathbb{R}^n ,

$$F(\xi; x) := - \ln(\xi^2 - \|x\|^2)$$

for the second-order cone $\{(\xi; x) \in \mathbb{R}^{1+n} : \xi \geq \|x\|\}$, and

$$F(x) := - \ln \det x$$

for the cone of symmetric positive semidefinite matrices of order n , can be shown to satisfy these conditions, with parameters θ equal to n , 2, and n respectively.

If F is a θ -logarithmically homogeneous self-concordant barrier for K , then F_* , defined by

$$F_*(s) := \sup\{-\langle s, x \rangle - F(x) : x \in \text{int } K\},$$

is a similar barrier for K^* . We define F' and F'' , and similarly F'_* and F''_* , for $x \in \text{int } K$ and $s \in \text{int } K^*$, by

$$\begin{aligned}\langle F'(x), v \rangle &:= DF(x)[v]; \\ \langle F''(x)v, w \rangle &:= D^2F(x)[v, w]; \\ \langle u, F'_*(s) \rangle &:= DF_*(s)[u]; \text{ and} \\ \langle t, F''_*(s)u \rangle &:= D^2F_*(s)[t, u],\end{aligned}$$

for all $v, w \in E$ and all $t, u \in E^*$. The mapping $-F'$ takes $\text{int } K$ into $\text{int } K^*$, and $-F'_*$ takes $\text{int } K^*$ into $\text{int } K$; these two mappings are inverses of one another. Also, F'' takes $\text{int } K$ into the space of linear maps from E into E^* , and F''_* takes $\text{int } K^*$ into the space of linear maps from E^* into E , and $F''_*(-F'(x)) = F''(x)^{-1}$.

Suppose F further satisfies

- (a) $F''(w)(K) \subseteq K^*$ for all $w \in \text{int } K$; and
- (b) $F_*(F''(w)x) = F(x) - 2F(w) - \theta$ for all $w, x \in \text{int } K$;

then we call F a *self-scaled* barrier (Nesterov and Todd [11]). It follows that F_* is then necessarily a self-scaled barrier for K^* . The barriers above are all self-scaled.

A cone that admits a self-scaled barrier is called *self-scaled*. It turns out from the work of Güler [6] that self-scaled cones are precisely symmetric cones. Moreover, the latter have been characterized: they are exactly cones that are isomorphic to direct sums of second-order cones, cones of symmetric positive semidefinite real matrices, cones of Hermitian positive semidefinite complex matrices, cones of Hermitian positive semidefinite quaternion matrices, and an exceptional 27-dimensional cone (see Faraut and Koranyi [3]). While such cones are restrictive, they arise in a wide variety of optimization applications: see, e.g., Vandenberghe and Boyd [17], Lewis and Overton [8], Lobo et al. [9], Ben Tal and Nemirovski [2], Alizadeh and Goldfarb [1], and [16].

The key property we use here of self-scaled cones is the following: for every $x \in \text{int } K$ and $s \in \text{int } K^*$, there is a unique $w \in \text{int } K$ such that

$$F''(w)(K) = K^* \quad \text{and} \quad F''(w)x = s; \tag{2.1}$$

see Nesterov and Todd [11], Theorems 3.1 and 3.2. Thus $F''(w)$ takes the cone K into K^* and the point x into s , so that the geometry of K^* around s is exactly that of K around x . The existence of such a linear bijection follows directly from K being symmetric: if J is a linear bijection from K onto K^* , and G an automorphism of K taking x into $J^{-1}s$, then JG takes K into K^* and x into s . However, self-scaled cones have a precise form for such a mapping, using their associated self-scaled barriers. We call w the *scaling point* for x and s . Algorithms whose search directions are obtained by solving restricted subproblems of the form (1.3) and (1.4), where C and C^* are balls defined using $H = F''(w)$ with w the scaling point for \bar{x} and \bar{s} , are said to be using the Nesterov-Todd scaling.

We now briefly discuss hyperbolicity cones. Let p be a polynomial function defined on E , homogeneous of degree m . We say that p is *hyperbolic* in direction $d \in E$ if, for every $x \in E$, the polynomial of τ defined by $p(x - \tau d)$ has only real roots. Given such a polynomial, the *hyperbolicity cone* of p in the direction d is the closure of the connected component of the set $\{x \in E : p(x) \neq 0\}$ containing d (see Güler [7], e.g.). The nonnegative orthant, second-order cone, and semidefinite cone are all hyperbolicity cones; the corresponding polynomials are

$$p(x) := \prod_j x_j, \quad p(\xi; x) := \xi^2 - \|x\|_2^2, \quad \text{and} \quad p(x) = \det x$$

respectively. However, hyperbolicity cones are much more general: for example, all homogeneous cones are hyperbolicity cones, whether or not they are self-dual; see Güler [7]. It turns out that $F(x) = -\ln p(x)$ is always an m -logarithmically homogeneous self-concordant barrier function. Moreover, under the weak condition that p is not constant on any lines, Güler shows in Theorem 6.2 of [7] that, for every $x \in \text{int } K$, $s \in \text{int } K^*$, there is a unique scaling point $w \in \text{int } K$ satisfying

$$F''(w)x = s \quad \text{and} \quad F''(w)(K) \subseteq K^*. \tag{2.2}$$

Thus the strong property (2.1) of self-scaled cones almost holds.

3 Results

We can now answer the question raised in the Introduction in the case that K is self-scaled, with self-scaled barrier F . We are given $\bar{x} \in \text{int } K$ and $\bar{s} \in \text{int } K^*$, and we wish to find a self-adjoint positive definite $H : E \rightarrow E^*$ to maximize $\alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)$. We show that $H = F''(w)$ solves this problem, where w is the scaling point of \bar{x} and \bar{s} , so that

$$F''(w)(K) = K^*, \quad F''(w)\bar{x} = \bar{s}. \quad (3.1)$$

Recall that

$$\alpha_{\bar{x}}(H) := \max\{\alpha_{\bar{x}} : B_H(\bar{x}; \alpha_{\bar{x}}) \subseteq K\}, \quad \alpha_{\bar{s}}(H) := \max\{\alpha_{\bar{s}} : B_H^*(\bar{s}; \alpha_{\bar{s}}) \subseteq K^*\}.$$

For any $v \in \text{int } K$, we write $\|\cdot\|_v$, $B_v(\bar{x}; \alpha_{\bar{x}})$, and $\alpha_{\bar{x}}(v)$ for $\|\cdot\|_{F''(v)}$, $B_{F''(v)}(\bar{x}; \alpha_{\bar{x}})$, and $\alpha_{\bar{x}}(F''(v))$, and define $\|\cdot\|_v^*$, $B_v^*(\bar{s}; \alpha_{\bar{s}})$, and $\alpha_{\bar{s}}(v)$ similarly. Then there is some $\bar{z} \in \partial K$ (the boundary of K) with $\|\bar{z} - \bar{x}\|_w = \alpha_{\bar{x}}(w)$, so by the separation theorem, for some nonzero $\bar{u} \in K^*$, $\langle \bar{u}, \bar{z} \rangle = 0$, and since $B_w(\bar{x}; \alpha_{\bar{x}}(w)) \subseteq K$,

$$0 = \langle \bar{u}, \bar{z} \rangle = \min\{\langle \bar{u}, z \rangle : z \in B_w(\bar{x}; \alpha_{\bar{x}}(w))\}. \quad (3.2)$$

Similarly, there is some $\bar{t} \in \partial K^*$ with $\|\bar{t} - \bar{s}\|_w^* = \alpha_{\bar{s}}(w)$, and for some nonzero $\bar{v} \in K$,

$$0 = \langle \bar{t}, \bar{v} \rangle = \min\{\langle t, \bar{v} \rangle : t \in B_w^*(\bar{s}; \alpha_{\bar{s}}(w))\}. \quad (3.3)$$

In fact, by applying $F''(w)$ to K , \bar{x} , $B_w(\bar{x}; \alpha_{\bar{x}}(w))$, and \bar{z} , we see that we can take

$$\bar{t} = F''(w)\bar{z}, \quad \bar{v} = F''(w)^{-1}\bar{u}, \quad \text{and that} \quad \alpha_{\bar{x}}(w) = \alpha_{\bar{s}}(w). \quad (3.4)$$

To get an explicit formula for the minima in (3.2) and (3.3), we use

Lemma 3.1 *We have*

$$\min\{\langle \bar{u}, z \rangle : z \in B_H(\bar{x}; \alpha_{\bar{x}})\} = \langle \bar{u}, \bar{x} \rangle - \alpha_{\bar{x}}\|\bar{u}\|_H^*,$$

attained by $z = \bar{x} - (\alpha_{\bar{x}}/\|\bar{u}\|_H^*)H^{-1}\bar{u}$. *Correspondingly, we have*

$$\min\{\langle t, \bar{v} \rangle : t \in B_H^*(\bar{s}; \alpha_{\bar{s}})\} = \langle \bar{s}, \bar{v} \rangle - \alpha_{\bar{s}}\|\bar{v}\|_H,$$

attained by $t = \bar{s} - (\alpha_{\bar{s}}/\|\bar{v}\|_H)H\bar{v}$.

Proof: These follow immediately from the Karush-Kuhn-Tucker conditions. \square

Note that the optimizing z and t are special cases of (1.7) and (1.8), with A the null operator and the identity respectively.

Without loss of generality, we suppose that $\|\bar{u}\|_w^* = 1$, so that, with $\bar{v} = F''(w)^{-1}\bar{u}$, $\|\bar{v}\|_w = 1$. Then Lemma 3.1 together with (3.2) and (3.3) give

$$\langle \bar{u}, \bar{x} \rangle = \alpha_{\bar{x}}(w)\|\bar{u}\|_w^* = \alpha_{\bar{x}}(w), \quad \langle \bar{s}, \bar{v} \rangle = \alpha_{\bar{s}}(w)\|\bar{v}\|_w = \alpha_{\bar{s}}(w). \quad (3.5)$$

We observe that $\langle \bar{u}, \bar{x} \rangle = \langle F''(w)\bar{v}, \bar{x} \rangle = \langle F''(w)\bar{x}, \bar{v} \rangle = \langle \bar{s}, \bar{v} \rangle$, so again we see that $\alpha_{\bar{x}}(w) = \alpha_{\bar{s}}(w)$.

Now let $H : E \rightarrow E^*$ be an arbitrary self-adjoint positive definite operator. Since $B_H(\bar{x}; \alpha_{\bar{x}}(H)) \subseteq K$, $\min\{\langle \bar{u}, z \rangle : z \in B_H(\bar{x}; \alpha_{\bar{x}}(H))\} \geq 0$, so from Lemma 3.1,

$$\alpha_{\bar{x}}(H)\|\bar{u}\|_H^* \leq \langle \bar{u}, \bar{x} \rangle.$$

Similarly, $B_H^*(\bar{s}; \alpha_{\bar{s}}(H)) \subseteq K^*$ yields

$$\alpha_{\bar{s}}(H)\|\bar{v}\|_H \leq \langle \bar{s}, \bar{v} \rangle.$$

Since all the quantities involved are positive, we can multiply these inequalities to obtain

$$\alpha_{\bar{x}}(w)\alpha_{\bar{s}}(w) = \langle \bar{u}, \bar{x} \rangle \langle \bar{s}, \bar{v} \rangle \quad (3.6)$$

$$\geq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)\|\bar{u}\|_H^*\|\bar{v}\|_H \quad (3.7)$$

$$\geq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)\langle \bar{u}, \bar{v} \rangle \quad (3.8)$$

$$= \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H), \quad (3.9)$$

since $\|\cdot\|_H^*$ and $\|\cdot\|_H$ are dual norms and $\langle \bar{u}, \bar{v} \rangle = \langle \bar{u}, F''(w)^{-1}\bar{u} \rangle = (\|\bar{u}\|_w^*)^2 = 1$. We have therefore proved

Theorem 3.1 *The product of the inscribed radii $\alpha_{\bar{x}}(H)$ and $\alpha_{\bar{s}}(H)$ is maximized when $H = F''(w)$, where w is the scaling point of \bar{x} and \bar{s} . \square*

This result provides the desired further justification of the Nesterov-Todd scaling in interior-point methods.

We now relate $\alpha_{\bar{x}}(v)$ for any $v \in \text{int } K$ to another geometric quantity. Recall that [11] defines

$$\sigma_{\bar{x}}(z) := (\sup\{\alpha : \bar{x} - \alpha z \in K\})^{-1} = \min\{\beta \geq 0 : \beta\bar{x} - z \in K\}.$$

Then $|z|_{\bar{x}} := \max\{\sigma_{\bar{x}}(z), \sigma_{\bar{x}}(-z)\}$ behaves like a scaled ℓ_∞ -norm in the same way that $\|z\|_{\bar{x}}$ behaves like a scaled ℓ_2 -norm.

Proposition 3.1 *For any $v \in \text{int } K$, we have*

$$\alpha_{\bar{x}}(v) \geq \frac{1}{\sigma_{\bar{x}}(v)} \quad \text{and} \quad \alpha_{\bar{s}}(v) \geq \frac{1}{\sigma_{\bar{s}}(-F'(v))}.$$

Proof: Since $\alpha_{\bar{x}}(v)$ scales inversely and $\sigma_{\bar{x}}(v)$ directly with v , we can assume that v is scaled up if necessary so that $\bar{x} - v \notin K$. From Theorem 4.1 of [11], with $\alpha = 1$, $p = \bar{x} - v$, $x = \bar{x}$, we get

$$\frac{1}{(1 + \sigma_{\bar{x}}(v - \bar{x}))^2} F''(\bar{x}) \leq F''(v),$$

so that

$$B_v(\bar{x}; \frac{1}{1 + \sigma_{\bar{x}}(v - \bar{x})}) \subseteq B_{\bar{x}}(\bar{x}; 1) \subseteq K. \quad (3.10)$$

Now

$$\begin{aligned} 1 + \sigma_{\bar{x}}(v - \bar{x}) &= 1 + \min\{\beta \geq 0 : \beta\bar{x} - (v - \bar{x}) \in K\} \\ &= 1 + \min\{\beta \geq 0 : (\beta + 1)\bar{x} - v \in K\} \\ &= \min\{\gamma \geq 1 : \gamma\bar{x} - v \in K\}. \end{aligned}$$

But since $\bar{x} - v \notin K$, $\gamma\bar{x} - v \in K$ implies $\gamma > 1$, so $\min\{\gamma \geq 1 : \gamma\bar{x} - v \in K\} = \min\{\gamma \geq 0 : \gamma\bar{x} - v \in K\} = \sigma_{\bar{x}}(v)$. Combining this with (3.10) yields the first result, and the second follows similarly. \square

In fact, the bounds given in the proposition are tight for several standard cones. For the nonnegative orthant, $\sigma_{\bar{x}}(v) = \max\{v_j/\bar{x}_j\}$, so $1/\sigma_{\bar{x}}(v) = \min\{\bar{x}_j/v_j\}$, and it is easy to see that a step of this length in the v -norm in the negative direction of the corresponding coordinate vector goes to the boundary of K . For the semidefinite cone, we find that $\sigma_{\bar{x}}(v)$ is the maximum eigenvalue of $\bar{x}^{-1/2}v\bar{x}^{-1/2}$, or equivalently of $v^{1/2}\bar{x}^{-1}v^{1/2}$. So its reciprocal is the minimum eigenvalue of $v^{-1/2}\bar{x}v^{-1/2}$. Let q be a corresponding unit eigenvector. Then $v^{-1/2}\bar{x}v^{-1/2} - (1/\sigma_{\bar{x}}(v))qq^T$ lies in the boundary of K , and hence so does $\bar{x} - (1/\sigma_{\bar{x}}(v))v^{1/2}qq^Tv^{1/2}$. It is easy to check that this displacement from \bar{x} has the appropriate length in the v -norm. Finally consider the second-order cone. The norms associated with its self-scaled barrier are somewhat hard to compute. Nevertheless, given any \bar{x} and v in $\text{int } K$, we can perform a transformation of the form $F''(w')^{-1}F''(w)$ to bring v to $v' := (1; 0)$ and \bar{x} to $\bar{x}' := (\xi; x)$. Moreover, using Theorem 3.2 of [11], the norms transform in the appropriate way. We find the reciprocal of $\sigma_{\bar{x}'}(v')$ is $\xi - \|x\|_2$. The norm associated with v' is just $\sqrt{2}$ times the Euclidean norm, and so a step from \bar{x}' to

$$\left(\xi - \frac{\xi - \|x\|_2}{2}; x + \frac{\xi - \|x\|_2}{2\|x\|_2}x \right)$$

shows the equality.

We can now establish our lower bound (tight in the cases above) for $\alpha_{\bar{x}}(w)\alpha_{\bar{s}}(w)$:

Proposition 3.2 *If we choose $v = w$, the scaling point for \bar{x} and \bar{s} , then*

$$\alpha_{\bar{x}}(w) = \alpha_{\bar{s}}(w) \geq \frac{1}{\sigma_{\bar{x}}(w)} = \frac{1}{\sigma_{\bar{s}}(-F'(w))} \quad (3.11)$$

and

$$\alpha_{\bar{x}}(w)\alpha_{\bar{s}}(w) \geq \frac{1}{\sigma_{\bar{x}}(w)^2} = \frac{1}{\sigma_{\bar{x}}(-F'_*(s))} = \frac{1}{\sigma_{\bar{s}}(-F'(\bar{x}))}. \quad (3.12)$$

Proof: The previous result gives $\alpha_{\bar{x}}(w) \geq 1/\sigma_{\bar{x}}(w)$ and $\alpha_{\bar{s}}(w) \geq 1/\sigma_{\bar{s}}(-F'(w))$. We have already noted (due to the action of $F''(w)$) that the two left-hand sides are equal, and Lemma 3.4 of [12] shows that the right-hand sides are also equal.

Then the second part follows since $\sigma_{\bar{x}}(w)^2 = \sigma_{\bar{x}}(-F'_*(s)) = \sigma_{\bar{s}}(-F'(\bar{x}))$, also by Lemma 3.4 of [12]. \square

It would be nice to have a more direct geometric proof of these results.

4 Discussion

The question we have raised can be posed with respect to any regular cone, but we do not know how to solve it in general. Let us consider the case of a hyperbolicity cone. By (2.2), given $\bar{x} \in \text{int } K$ and $\bar{s} \in \text{int } K^*$, we can find a unique $w \in \text{int } K$ with

$$F''(w)\bar{x} = \bar{s} \quad \text{and} \quad F''(w)(K) \subseteq K^*.$$

It therefore seems natural to choose $H = F''(w)$ as in the case of self-scaled cones. Let us examine how the arguments of the previous section need to be modified.

Equations (3.2) and (3.3) still hold, for the same reasons, but now it is no longer the case that $\bar{v} = F''(w)^{-1}\bar{u}$ as in (3.4). Indeed, $F''(w)^{-1}\bar{u}$ may not lie in K , because \bar{u} may lie in $K^* \setminus F''(w)(K)$.

Lemma 3.1 remains true. We can again scale \bar{u} so that $\|\bar{u}\|_w^* = 1$, but this no longer implies that $\|\bar{v}\|_w = 1$. Nevertheless, we can scale \bar{v} separately so that this last relation holds. Then, using Lemma 3.1, we see that (3.5) still holds. However, since $\langle \bar{u}, \bar{x} \rangle$ may not equal $\langle \bar{s}, \bar{v} \rangle$, we cannot conclude that $\alpha_{\bar{x}}(w)$ and $\alpha_{\bar{s}}(w)$ are equal. Instead, since the images of \bar{x} , $B_w(\bar{x}; \alpha_{\bar{x}}(w))$, and K by $F''(w)$ are \bar{s} , $B_w^*(\bar{s}; \alpha_{\bar{x}}(w))$, and $F''(w)(K) \subseteq K^*$, we can conclude that

$$\alpha_{\bar{s}}(w) \geq \alpha_{\bar{x}}(w).$$

Finally, we still have (3.6)–(3.8), but (3.9) may not hold since \bar{u} and \bar{v} are no longer related: indeed, we have $\langle \bar{u}, \bar{v} \rangle \leq \|\bar{u}\|_w^* \|\bar{v}\|_w = 1$, and this inequality is in the reverse direction of what we need. If we can compute the largest inscribed ellipsoids corresponding to the dual w -norms, and obtain the corresponding vectors \bar{u} and \bar{v} , we know that these norms are optimal to within at least the factor $\langle \bar{u}, \bar{v} \rangle$.

We conclude by relating the criterion we have been maximizing, $\alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)$, to another natural criterion. Suppose we move from \bar{x} to the solution to (1.3) when $C = B_H(\bar{x}; \alpha_{\bar{x}}(H))$. Then the improvement in the objective function of (1.1) is easily seen to be

$$\begin{aligned} \delta_{\bar{x}} &:= \alpha_{\bar{x}}(H) \langle s_0, (H^{-1} - H^{-1}A^*(AH^{-1}A^*)^{-1}AH^{-1})s_0 \rangle^{1/2} \\ &= \alpha_{\bar{x}}(H) \langle \bar{s}, (H^{-1} - H^{-1}A^*(AH^{-1}A^*)^{-1}AH^{-1})\bar{s} \rangle^{1/2}, \end{aligned}$$

using (1.7) and the fact that $s_0 - \bar{s}$ lies in the range of A^* since \bar{s} is feasible in (1.2). Similarly, the improvement in the objective function of (1.2), when moving from \bar{s} to the solution to (1.4) when $C^* = B_H^*(\bar{s}; \alpha_{\bar{s}}(H))$, is seen to be

$$\begin{aligned} \delta_{\bar{s}} &:= \alpha_{\bar{s}}(H) \langle A^*(AH^{-1}A^*)^{-1}Ax_0, x_0 \rangle^{1/2} \\ &= \alpha_{\bar{s}}(H) \langle A^*(AH^{-1}A^*)^{-1}A\bar{x}, \bar{x} \rangle^{1/2}. \end{aligned}$$

It is at least plausible to choose H to maximize the product of these improvements, although the sum would perhaps be preferable.

Now since $H^{-1}A^*(AH^{-1}A^*)^{-1}AH^{-1}$ is positive semidefinite, we obtain

$$\langle \bar{s}, (H^{-1} - H^{-1}A^*(AH^{-1}A^*)^{-1}AH^{-1})\bar{s} \rangle^{1/2} \leq \langle \bar{s}, H^{-1}\bar{s} \rangle^{1/2} = \|\bar{s}\|_H^*,$$

and since $H - A^*(AH^{-1}A^*)^{-1}A$ is positive semidefinite, we get

$$\langle A^*(AH^{-1}A^*)^{-1}A\bar{x}, \bar{x} \rangle^{1/2} \leq \langle H\bar{x}, \bar{x} \rangle^{1/2} = \|\bar{x}\|_H.$$

Hence

$$\delta_{\bar{x}}\delta_{\bar{s}} \leq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)\|\bar{s}\|_H^*\|\bar{x}\|_H \leq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)\langle \bar{s}, \bar{x} \rangle.$$

Maximizing $\alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)$ is therefore equivalent to maximizing this upper bound on the product of the improvements possible by taking steps to the boundaries of the dual inscribed ellipsoids.

References

- [1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming **95** (2003), no. 1, 3–51.
- [2] A. Ben-Tal and A. Nemirovskii, *Lectures on modern convex optimization: Analysis, algorithms, and engineering applications*, MPS/SIAM Series on Optimization, vol. 2, SIAM, Philadelphia, 2001.

- [3] J. Faraut and A. Koranyi, *Analysis on symmetric cone*, Oxford University Press, Oxford, 1994.
- [4] R. M. Freund, *On the primal-dual geometry of level sets in linear and conic optimization*, SIAM Journal on Optimization **13** (2003), no. 4, 1004–1013.
- [5] ———, *Complexity of convex optimization using geometry-based measures and a reference point*, Mathematical Programming **99** (2004), no. 2, 197–221.
- [6] O. Güler, *Barrier functions in interior point methods*, Mathematics of Operations Research **21** (1996), 860–885.
- [7] ———, *Hyperbolic polynomials and interior point methods for convex programming*, Mathematics of Operations Research **22** (1997), 350–377.
- [8] A. S. Lewis and M. L. Overton, *Eigenvalue optimization*, Acta Numerica, Cambridge University Press, 1996, pp. 149–190.
- [9] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebert, *Applications of second-order cone programming*, Linear Algebra and its Applications **284** (1998), no. 1–3, 193–228.
- [10] Y. E. Nesterov and A. S. Nemirovskii, *Interior point polynomial methods in convex programming: Theory and algorithms*, SIAM Publications, SIAM, Philadelphia, USA, 1993.
- [11] Y. E. Nesterov and M. J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research **22** (1997), 1–42.
- [12] ———, *Primal-dual interior-point methods for self-scaled cones*, SIAM Journal on Optimization **8** (1998), 324–364.
- [13] J. Renegar, *Some perturbation theory for linear programming*, Mathematical Programming **65** (1994), 73–91.
- [14] ———, *Linear programming, complexity theory and elementary functional analysis*, Mathematical Programming **70** (1995), 279–351.
- [15] M. J. Todd, *Scaling, shifting and weighting in interior-point methods*, Computational Optimization and Applications **3** (1994), 305–315.
- [16] ———, *Semidefinite optimization*, Acta Numerica **10** (2001), 515–560.
- [17] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Review **38** (1996), 49–95.