Largest dual ellipsoids inscribed in dual cones

Michael J. Todd

School of Operations Research and Industrial Engineering,
Cornell University
www.orie.cornell.edu/ ˜miketodd/todd.html

FoCM Conference, Santander
July 4, 2005

Problem

E: finite dimensional real vector space with dual E^*;

$K \subseteq E$: regular ($=$ solid and pointed) closed convex cone;

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

$\bar{x} \in$ int K, $\qquad \bar{s} \in$ int K^*.

(Can think of E and E^* as $\Re^n$ with $\langle \cdot, \cdot \rangle$ the usual dot product.)

Problem:

Find “large” ellipsoids centered at $\bar{x}$ and $\bar{s}$ and contained in K and K^*.

Motivation

Consider the conic programming problem

$$\min \langle s_0, x\rangle : \quad x \in L + \{x_0\}, \quad x \in K,$$

and its dual

$$\min \langle s, x_0\rangle : \quad s \in L^* + \{s_0\}, \quad s \in K^*,$$

where L is a linear subspace of E.

Cases:

K nonnegative orthant → linear programming;

K second-order cone → second-order cone programming;

K positive semidefinite cone → semidefinite programming.

Suppose $\bar{x}$ and $\bar{s}$ are feasible. We wish to find improved iterates.

Restrictions

Let us choose ellipsoids $C \subseteq K$ and $C^* \subseteq K^*$ centered at $\bar{x}$ and $\bar{s}$ and solve the simpler subproblems

$$\min \langle s_0, x \rangle : \quad x \in L + \{\bar{x}\}, \quad x \in C,$$

and

$$\min \langle s, x_0 \rangle : \quad s \in L^* + \{\bar{s}\}, \quad s \in C^*.$$

Good approximation: want C and C^* "large" — choose maximum volume.

There are known solutions for the cases above.

But ... C and C^* are unrelated, so ...

Need to solve two subproblems.

Dual Ellipsoids

Let $H : E \to E^*$ be self-adjoint and positive definite.

H defines the norms on E and E^*:

$$\|v\|_H := \langle Hv, v\rangle^{1/2}, \qquad \|u\|_H^* := \left\langle u, H^{-1}u\right\rangle^{1/2},$$

dual to each other.

Let

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

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

Then both restrictions can be solved for the price of one.

So want the “largest” inscribed dual balls.

Goal

If $\dim E =: n$, we have

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

so that

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

Let

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

Goal: Maximize the product $\alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)$ of the radii of the balls.

Solution: If K is symmetric (homogeneous and self-dual), then we know the form of the answer. If we have a self-scaled barrier for K, then we get an explicit solution. This gives a geometric motivation for the Nesterov-Todd scaling.

Symmetric Cones

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

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

homogeneous if, for every x, z in int K, there is a linear bijection $G : E \longrightarrow E$ with

$$G(K) = K, \qquad Gx = z;$$

and symmetric if it satisfies both properties. Then for every $x \in$ int K, $s \in$ int K^*, there exist such J and G with

$$(J \circ G)(K) = K^*, \qquad (J \circ G)x = s.$$

Barriers

Nesterov and Nemirovskii defined θ-logarithmically homogeneous self-concordant barriers F : int $K \to \Re$ (think of $F(x) = -\sum \ln x_j$ for the nonnegative orthant or $F(x) = -\ln \det x$ for the semidefinite cone). These exist for all regular closed convex cones.

Nesterov and Todd defined self-scaled barriers.

Güler observed that a cone admitted a self-scaled barrier (i.e., is self-scaled) iff it is symmetric.

NT: if $x \in$ int K, $s \in$ int K^*, then there is a unique $w \in$ int K with

$$F''(w)(K) = K^*, \qquad F''(w)x = s;$$

w is called the scaling point for x and s.

Güler: this almost holds (instead $F''(w)(K) \subseteq K^*$) for the wider class of hyperbolicity cones.

Main Result

Theorem: If K is self-scaled, then $\alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)$ is maximized by $H = F''(w)$, where w is the scaling point for $\bar{x}$ and $\bar{s}$.

Sketch of proof: There is some $\bar{z} \in \partial K$ with $\|\bar{z} - \bar{x}\|_w = \alpha_{\bar{x}}(w)$. So there is some nonzero $\bar{u} \in K^*$ with

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

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

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

Without loss of generality, we can take $\|\bar{u}\|_w^* = \|\bar{v}\|_w = 1$.

In fact, we can take $\bar{t} = F''(w)\bar{z}$, $\bar{v} = F''(w)^{-1}\bar{u}$.

Proof, ctd.

This implies that

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

Now, for any H, $B_H(\bar{x}, \alpha_{\bar{x}}(H)) \subseteq K$ and $B^*_H(\bar{s}, \alpha_{\bar{s}}(H)) \subseteq K^*$, so

$$\min\{\langle \bar{u}, z \rangle : z \in B_H(\bar{x}, \alpha_{\bar{x}}(H))\} \geq 0,$$

$$\min\{\langle t, \bar{v} \rangle : t \in B^*_H(\bar{s}, \alpha_{\bar{s}}(H))\} \geq 0.$$

Hence we obtain

$$\begin{aligned} \alpha_{\bar{x}}(w)\alpha_{\bar{s}}(w) &= \langle \bar{u}, \bar{x} \rangle \langle \bar{s}, \bar{v} \rangle \geq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)\|\bar{u}\|^*_H \|\bar{v}\|_H \\ &\geq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H) \langle \bar{u}, \bar{v} \rangle = \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)\|\bar{v}\|^2_w = \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H). \end{aligned}$$

□

Final Remarks

If K is not self-scaled, but is a hyperbolicity cone, we can again find the “scaling point” w for $\bar{x}$ and $\bar{s}$, but because we only have $F''(w)(K) \subseteq K^*$, the argument above does not go through. We can still define $\bar{z}$, $\bar{u}$, $\bar{t}$, and $\bar{v}$ as before, but we no longer have $\bar{v} = F''(w)^{-1}\bar{u}$. So we obtain:

$$\alpha_{\bar{x}}(w)\alpha_{\bar{s}}(w) = \langle \bar{u}, \bar{x} \rangle \langle \bar{s}, \bar{v} \rangle \geq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H)\|\bar{u}\|_H^*\|\bar{v}\|_H \geq \alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H) \langle \bar{u}, \bar{v} \rangle .$$

If we can find $\bar{u}$ and $\bar{v}$, $H = F''(w)$ is optimal to within the factor $\langle \bar{u}, \bar{v} \rangle \leq \|\bar{u}\|_w^*\|\bar{v}\|_w = 1$.

Also, $\alpha_{\bar{x}}(H)\alpha_{\bar{s}}(H) \langle \bar{s}, \bar{x} \rangle$ can be shown to be an upper bound on the product of the improvement in the primal objective function in moving from $\bar{x}$ to the optimal solution of the primal ellipsoidal restriction, and the corresponding improvement in the dual objective function. Hence we are equivalently maximizing this upper bound.