Long–step Interior Point Methods and Barrier Functions

Osman Güler

`guler@math.umbc.edu`

University of Maryland, Baltimore County
Baltimore, MD, 21250

Programs in conic duality

$$\begin{array}{lll} \min & \langle c, x\rangle \\ s.t. & \mathcal{A}x = b \\ & x \in K \end{array} \qquad \begin{array}{lll} \max & \langle b, y\rangle \\ s.t. & \mathcal{A}^* y + s = c \\ & s \in K^* \end{array}$$

Here $c \in E$, a f.d. Euclidean space, $\mathcal{A} : E \to F$ linear, $K \subseteq E$ a regular convex cone, and $\mathcal{A}^*$ is the conjugate of $\mathcal{A}$.

More geometrically,

$$\begin{array}{lll} \min & \langle s_0, x\rangle \\ s.t. & x \in x_0 + L \\ & x \in K \end{array} \qquad \begin{array}{lll} \min & \langle x_0, s\rangle \\ s.t. & s \in s_0 + L^{\perp} \\ & s \in K^* \end{array}$$

- we assume $\exists$ feasible interior points ($x \in intK$, $s \in intK^*$),
- the dual program above is obtained using standard duality theory of convex programming, cf. Rockafellar.

Outline

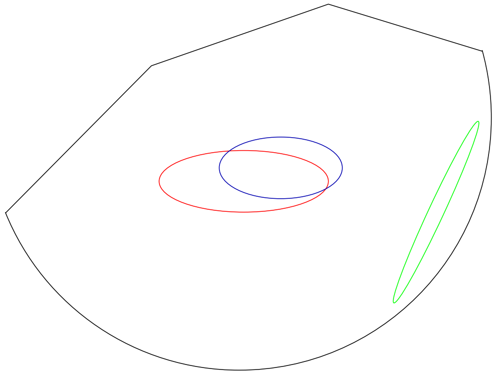
- self–concordant barrier functions (SCB)
- sketch of the theory of interior point methods (IPM)
- universal barrier function (UBF)
- self–scaled barriers (SSB)
- hyperbolic barriers
- homogeneous cones

Self–concordant barriers

A SCB on a regular convex cone K is a convex function F satisfying

- $|D^3F(x)[h,h,h]| \le 2(D^2F(x)[h,h])^{3/2}$ (self–concordance)
- $|DF(x)[h]|^2 \le \vartheta D^2F(x)[h,h]$ (ϑ *barrier param.*)
- $F(x) \to \infty$ as $x \to \partial C$ (barrier property)

- Hessian $D^2F(x)$ defines a Riemannian metric on K, $|u|_x = \sqrt{D^2F(x)[u,u]}$
- SC property is equivalent to $1 - |y-x||_x \le \frac{|u|_y}{||u|_x} \le \frac{1}{1-|y-x|_x}$
  - thus the term *self–concordance*
  - Dikin ellips. $\{y : ||y-x||_x < 1\} \subset K$
- ϑ appears in step–sizes in Newton's method to minimize F



Dikin Ellipsoids

Sketch of IPM

Path following methods:

Follow the *central path* $\{x(t), s(t)) : t > 0\}$ where

$$x(t) = \arg\min\{\langle x, s_0\rangle + tF(x) : x \in x_0 + L\},$$

$$s(t) = \arg\min\{\langle x_0, s\rangle + tF^*(x) : x \in x_0 + L\},$$

and reduce the *duality gap* $\langle x, s\rangle$ by an absolute constant, using *Newton's method*.

Potential reduction methods:

Use Newton–like methods to reduce the Tanabe–Todd–Ye *potential function*

$$\Phi(x, s) = \rho \ln\langle x, s\rangle + F(x) + F^*(s)$$

on $\{(x, s) : x \in x_0 + L, s \in s_0 + L^\perp\}$, by an absolute constant at each step.

Universal barrier function I

Nesterov & Nemirovskii (1988): $u(x) = c\ln(v^{ol_n}(Q^*(x))$, where $Q^*(x) = \{y \in \mathbb{R}^n : \langle z - x, y\rangle \leq 1, \quad \forall z \in Q\}$.

Equivalently, G (1994): $F(x) = c\ln(\underbrace{\int_{C^*} e^{-\langle x,v\rangle} dv}_{characteristic\ fu^{nctio^n}})$

- $F(x)$ is SCB on any cone C (Nesterov & Nemirovskii)
  - F satisfies other nice properties (discussed below)
- F is hard to compute in general
- F is *invariant*: $F(Ax) = F(x) - \ln\det A$ if A is linear, $A(C) = C$.
  - If C is a *homogeneous cone*, $F(x) = const - c\ln detA_x$.
  - C homog. $\Rightarrow$ F is easy to compute

Universal barrier function II

- Santaló (1949) defines the function $vol_n(Q^*(x)$ for use in convex geometry. Its minimizer, *analytic center for us* is also called the *Santaló point*.
- Bochner (1944) defines the function

$$\int_{T(K^*)} e^{i\langle x+iy,u\rangle} du$$

  calls it the *Cauchy kernel*. He shows that this is the Cauchy-Szegö reproducing kernel for the Hardy space $H^2(T(K^*))$, $T(K^*) := \mathbb{R}^n + iK^*$, the *tube domain* on K^*. Cauchy kernel is also related to *Bergman kernel* over $T(K^*)$, see Faraut–Korányi, Chapter 9.
- Koecher (1957) defines *characteristic function* $\varphi(x) = \int_{C^*} e^{-\langle x,v\rangle} dv$ to study symmetric cones. Note that it is just the real part of the Cauchy kernel.

Universal barrier function III

The UBF also satisfies:

- $|D^m F(x)[h,h,\ldots,h]| \le c(m)(D^2F(x)[h,h])^{m/2}, \quad \forall m \ge 3.$

  Here $D^m F(x)[h,h,\ldots,h] = \frac{d^m}{dt^m} F(x+th)|_{t=0}$ is the kth directional derivative at x along h. (If $m=3$, this is the self–concordance inequality.)

- $|D^4 F(x)[h,h,h,h]| \le c(4)(1+3\vartheta)^2 D^2F(x)[h,h]\,||h||^2_{K,x}.$

  This can be used to get some long–step IPM (NN).

These are obtained by writing

$\varphi(x+th) = c^{1-n}\int_0^\infty s^{n-1}e^{-s}\int_{\{y\in K^*:\langle x,y\rangle=c\}} e^{-st\langle h,y\rangle}\,dy\,ds$, and expanding the exponential in power series. A result of Bourgain (also Kannan-Lovász-Simonovits)

$$\left(\int_Q |p(x)|^r dx\right)^{1/r} \le c(d,r)\int_Q |p(x)|dx, \quad \forall r>0,$$

where $|Q|_n = 1$, a convex body, and $deg(p) = d$.

Universal barrier function IV

We see that UBF has a quite history. It has given rise to

- Blaschke–Santaló inequality in convex geometry: estimating the size of $\min\{|Q_x| : x \in Q\}$
- the study of Santaló regions (level sets of UBF) and its relation to *floating bodies* and *affine surface area* (Meyer & Werner)
- the study of boundary behavior of UBF using similar ideas in Bergman kernel theory (Sasaki)
- Fourier analysis can also be used to study UFB (after all, the characteristic function is Laplace–Fourier transform)
- a lot of work has been done in IPM on long–step methods, but still further work needs to be done, perhaps using these diverse approaches
- is it possible to devise other integrals to define SCB? (answer seems to be yes; this is true for hyperbolic barriers)

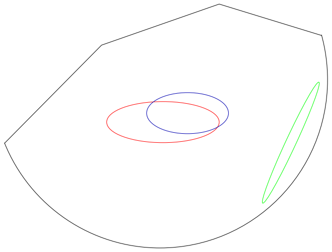
Self–scaled barriers (SSB)

A SCB F is *self-scaled* if $F(tx) = F(x) - \vartheta \ln t$ and

$$F''(w)x \in K^*, \qquad F_*\big(F''(w)x\big) = F(x) - 2F(w) - \vartheta, \quad \forall x, w \in K.$$

If K allows such a barrier, then K is a *self-scaled cone* (Nesterov and Todd, 1994)

- Self–scaled barriers allow *long–step* primal–dual IPM
- as Hessian norms can be compared at any two points in K
- Applications include SDP and QCP



Dikin Ellipsoids

A cone is *symmetric* if it is homogeneous and $K^* = K$.

- Self–scaled cones = symmetric cones (G)
- symmetric cones = *Euclidean Jordan algebras* (Koecher)
  - basic symmetric cones: psd cones over $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$; quadratic cone; and 3×3 psd cone $\mathbb{O}$ (Jordan, von Neumann, and Wigner)
  - Jordan algebra formalism simplifies treatment of IPM
- Self–scaled barriers have been completely classified

Properties of SSB

- $g(x) = -\langle DF(x), y\rangle$ is convex on $intK$ when $y \in intK$.
- This implies (Nesterov–Todd)

  $$\frac{1}{(1+\alpha\sigma_x(p))^2} D^2F(x) \preceq D^2F(x-\alpha p) \preceq \frac{1}{(1+\alpha\sigma_x(p))^2} D^2F(x),$$

  where $\sigma_x(p) = 1/sup\{\alpha : x - \alpha p \in intK\}$.
- $$f(x-\alpha p) \leq F(x) - \alpha DF(x)[p] + \frac{||p||_x^2}{\sigma_x(p)^2}(-\alpha\sigma_x(p) - \ln(1-\alpha\sigma_x(p))).$$

... and these inequalities lead to powerful, long–step IPM.

Hyperbolic barriers

Hyperbolic polynomials:

$$\{p : \mathbb{R}^n \to \mathbb{R} : p \text{ homog. }, t \mapsto p(x+td) \text{ has all real roots}\}.$$

Hyperbolic polynomials give rise to IPM for which there exists a nice barrier and

long–step IPM

- Hyperbolic cone:
  $K := \{x \in \mathbb{R}^n : t \mapsto p(x+td)$ has all negative roots$\}$
- $F(x) = -\ln p(x)$ is self–concordant barrier on K
- $F(x)$ satisfies $g(x) = -\langle DF(x), y\rangle$ is convex, as in the self–scaled case; thus leading to long–step IPM
- however, F^* may not satisfy similar properties, so primal–dual methods may be harder to come by

Homogeneous cones

A *homogeneous cone* K in a f.d. Euclidean space $(E, \langle,\rangle)$ is a regular convex cone which has a *transitive automorhism group*, that is, if $x, y \in int(K)$, $\exists A : E \to E$ linear, $A(K) = K$ and $Ax = y$.

Because of this rich symmetry, the universal barrier function can be calculated explicitly, in fact $F(x) = const - \ln|\det A_x|$, where $A_x e = x$.

- homogeneous cones have been classified (Vinberg, 1960-1962) in terms of t–algebras invented for this purpose. However, this is a rough classification
    - the characteristic function φ is heavily used in the classification
    - this work leads to a recursive, *Siegel domain* construction of homogeneous cones (Vinberg, Gindikin) (discussed below)
    - Rothaus (1966) describes a dual, recursive construction of homog. cones (discussed below)
- Dorfmeister (1977-79, Ph.D. thesis) uses Jordan algebras to give a tighter classification of homog. cones
- Koszul (1962) gives a Lie group construction of homog. cones

Homogeneous cones II

Homogeneous cones form a much larger class than symmetric cones. Here are two examples:

Epigraph of Matrix Norms: $\{(x,u,t) \in S_m^\times \mathbb{R}^{n\times m} \times \mathbb{R}_+ : tx - u^T u \succeq 0\}$ is a Siegel cone over S_m^+, and is not a symmetric cone.

Vinberg cone: Let E be the space of 3×3 symmetric matrices $\begin{bmatrix} a & b_1 & b_2 \\ b_1 & c_1 & 0 \\ b_2 & 0 & c_2 \end{bmatrix}$ with inner product $\langle x, y = tr(xy)$.

- Vinberg's cone is $K = \{x \in E : a \geq 0, ac_1 - b_1^2, ac_2 - b_2^2 \geq 0\}$.
- K^* the psd cone in E; there is no linear map $A : E \to E$ satisfying $A(K) = K^*$.

Homogeneous cones III

Siegel domain construction of homog. cones:

Consider the commutative diagram

$$\begin{tikzcd}
\mathbb{R}^p \times \mathbb{R}^p \arrow[r, "\overline{g}\times\overline{g}"] \arrow[d, "B"'] & \mathbb{R}^p \times \mathbb{R}^p \arrow[d, "B"] \\
\mathbb{R}^k \arrow[r, "g"] & \mathbb{R}^k
\end{tikzcd}$$

where $g \in Aut(K)$, K homog cone in $\mathbb{R}^k$, B symmetric, bilinear, $B(u,u) \in K$, and $B(u,u) = 0$ iff $u = 0$ (B non–singular), and $\overline{g}$ linear is assumed to exist and make the diagram commutative.

The Siegel cone of K:

$$SC(K,B) = cl\{(x,u,t) \in \mathbb{R}^k \times \mathbb{R}^p \times \mathbb{R} : t > 0, tx - B(u,u) \in K\}.$$

- every homogeneous cones can obtained by a Siegel construction (Vinberg, Gindikin)
- UBF of SC(K,B): $F(x,u,t) = F_K(x - B(u/t,u)) + \ln(|\det \overline{g}|) - ((p+2)/2)\ln t$,
- *optimal barrier* G-Tunçel: $F(x,u,t) = F_K(x - B(u/t,u)) - \ln t$

Homogeneous cones IV

Dual Siegel domain construction of homog. cones (Rothaus):

Consider the commutative diagram

$$\begin{tikzcd}
\mathbb{R}^k \arrow[r, "g^*"] \arrow[d, "\mathrm{U}"'] & \mathbb{R}^k \arrow[d, "\mathrm{U}"] \\
S_p \arrow[r, "T_{\overline{g}}"] & S_p
\end{tikzcd}$$

Here $g^* \in Aut(K^*)$, $U(y)$ is psd if $y \in K^*$, and pd if $y \in int K^*$, $T_{\overline{g}}$ is assumed to exist and make the diagram commute, and S_p is the space of $p \times p$ symmetric matrices.

- the connection to the Siegel construction is: $\langle U(y)u, v\rangle = \langle B(u,v), y\rangle$,
- $int\ SC(K,B)^* = \{(y,v,s) \in int\ K^* \times \mathbb{R}^p \times \mathbb{R} : s > \langle U(y)^{-1}v, v\rangle\}$.
- $s > \langle U(y)^{-1}v, v\rangle \iff \begin{bmatrix} s & v^T \\ v & U(y) \end{bmatrix}$ pd.

Homogeneous cones V

- homog. cones admit hyperbolic barriers, hence long–step IPM
  - however their barrier parameter may be too high, and they may be hard to construct (Gindikin's idea is too theoretical, based on generalized eigenvalues)
  - primal–dual methods may not work well
- currently working on getting optimal/near optimal barriers, with nice duality properties, and long–step IPM properties
  - recursive constructions, t-algebras, and Koszul approach, and Dorfmeister's Jordan algebra approaches are explored
- we also need good application areas. May get better approximation algorithms using homogeneous cones, not SDP alone?