Conic optimization: affine geometry of self-concordant barriers

Roland Hildebrand

Laboratoire Jean Kuntzmann / CNRS

Seminar Nonlinear Analysis and Optimization

Laboratoire de Mathématiques d’Avignon
April 12, 2018
Outline

Geometry of self-concordant barriers

- self-concordant barriers
- affine differential geometry
- relationship between barriers and geometry
- canonical barrier
- self-scaled barriers
Regular convex cones

Definition
A regular convex cone $K \subset \mathbb{R}^n$ is a closed convex cone having nonempty interior and containing no lines.

The dual cone

$$K^* = \{ s \in \mathbb{R}^n \mid \langle x, s \rangle \geq 0 \quad \forall \ x \in K \}$$

of a regular convex cone $K$ is also regular.
Conic programs

Definition
A conic program over a regular convex cone $K \subset \mathbb{R}^n$ is an optimization problem of the form

$$\min_{x \in K} \langle c, x \rangle : \quad Ax = b.$$ 

to every conic program we can associate a dual program over the dual cone $K^*$

examples
- linear programs (LP)
- second-order cone programs (SOCP)
- semi-definite programs (SDP)
- geometric programs (GP)
Geometric interpretation

the feasible set is the intersection of $K$ with an affine subspace
History of conic programming

- **LP: Simplex method**
  - [Dantzig 1951], exp. compl.

- **Ellipsoid method**
  - [Yudin, Nemirovski 1976], polynomial-time

- **LP: Interior-point**
  - **projective scaling**
    - [Karmarkar 1984], polynomial-time
  - **affine scaling**
    - [Dikin 1967], rediscovery 1986

- **General cones: IP**
  - [Nesterov, Nemirovski 1988], self-concordant barriers

- **CP: primal, primal-dual IP**
  - [Nesterov, Nemirovski 1994], systematic approach
  - Universal barrier

- **Symmetric cones IP**
  - Euclidean Jordan algebras
    - [Faybusovich 1995]
  - self-scaled barriers
    - [Nesterov, Todd 1994]

- **Classification of self-scaled barriers**
  - [Hauser 1999, 2000]
  - [Hauser, Güler 2002]
  - [Hauser, Lim 2002]
  - [Schmieta 2000]
Logarithmically homogeneous barriers

Definition (Nesterov, Nemirovski 1994)
Let $K \subset \mathbb{R}^n$ be a regular convex cone. A (self-concordant logarithmically homogeneous) barrier on $K$ is a smooth function $F : K^o \rightarrow \mathbb{R}$ on the interior of $K$ such that

- $F(\alpha x) = -\nu \log \alpha + F(x)$ (logarithmic homogeneity)
- $F''(x) \succ 0$ (convexity)
- $\lim_{x \to \partial K} F(x) = +\infty$ (boundary behavior)
- $|F'''(x)[h, h, h]| \leq 2(F''(x)[h, h])^{3/2}$ (self-concordance)

for all tangent vectors $h$ at $x$.

The homogeneity parameter $\nu$ is called the barrier parameter.

Theorem (Nesterov, Nemirovski 1994)
Let $K \subset \mathbb{R}^n$ be a regular convex cone and $F : K^o \rightarrow \mathbb{R}$ a barrier on $K$ with parameter $\nu$. Then the Legendre transform $F^*$ is a barrier on $-K^*$ with parameter $\nu$. 
Barriers as penalty functions

let $K \subset \mathbb{R}^n$ be a regular convex cone
let $F : K^o \rightarrow \mathbb{R}$ be a barrier on $K$
consider the conic program

$$\min_{x \in K} \langle c, x \rangle : \quad Ax = b$$

for $\tau > 0$, solve instead the unconstrained problem

$$\min_{x \in \mathbb{R}^n} \tau \langle c, x \rangle + F(x) : \quad Ax = b$$

- unique minimizer $x^*(\tau) \in K^o$ for every $\tau > 0$
- solution depends continuously on $\tau$ (*central path*)
- $x^*(\tau) \rightarrow x^*$ as $\tau \rightarrow \infty$
alternate Newton steps and increments of $\tau$

the smaller the barrier parameter $\nu$, the faster we can increase $\tau$ safely

(in short-step methods) the iterates have to stay in a tube around the central path in order for the Newton method to make a controllable iteration

the larger $\nu$, the smaller the diameter of the tube
Outline

**Geometry of self-concordant barriers**
- self-concordant barriers
- affine differential geometry
- relationship between barriers and geometry
- canonical barrier
- self-scaled barriers
Affine connections

an affine connection $\nabla$ on a differentiable manifold defines the parallel transport of tangent vectors $u$ along curves $\sigma(t)$ by

$$\dot{u}^{\gamma} + \nabla_{\alpha\beta}^{\gamma} u^{\alpha} \dot{\sigma}^{\beta} = \left( \frac{\partial u^{\gamma}}{\partial x^{\beta}} + \nabla_{\alpha\beta}^{\gamma} u^{\alpha} \right) \dot{\sigma}^{\beta} = 0$$

the covariant derivative of the vector field $u$ is given by

$$\nabla_{\beta} u^{\gamma} = \frac{\partial u^{\gamma}}{\partial x^{\beta}} + \nabla_{\alpha\beta}^{\gamma} u^{\alpha}$$

we may also define the covariant derivative of general tensors

law of transformation under coordinate changes $x \mapsto y$

$$\nabla_{\alpha\beta}^{\gamma} \mapsto \frac{\partial x^{p}}{\partial y^{\alpha}} \frac{\partial x^{q}}{\partial y^{\beta}} \nabla_{pq}^{r} \frac{\partial y^{\gamma}}{\partial x^{r}} + \frac{\partial y^{\gamma}}{\partial x^{m}} \frac{\partial^{2} x^{m}}{\partial y^{\alpha} \partial y^{\beta}}$$

even can: the flat affine connection on $\mathbb{R}^{n}$ is given by $\nabla_{\alpha\beta}^{\gamma} = 0$ in affine coordinates
Affine differential geometry

Let $M \hookrightarrow \mathbb{R}^{n+1}$ be a hypersurface immersion and $\xi$ a transversal vector field on $M$.

Which objects can be defined on $M$ by the connection on $\mathbb{R}^{n+1}$?
Affine metric, affine connection, cubic form

let \( y^0, \ldots, y^n \) be affine coordinates on \( \mathbb{R}^{n+1} \) and \( x^1, \ldots, x^n \) coordinates on \( M \)
extend these to a neighbourhood of \( M \) and complement with a coordinate \( x^0 \) such that

- \( M \) is a level surface of \( x^0 \)
- \( \xi = \frac{\partial}{\partial x^0} \) on \( M \)

in \( x \) coordinates the flat affine connection of \( \mathbb{R}^{n+1} \) becomes

\[
\nabla^r_{ij} = \frac{\partial x^r}{\partial y^s} \frac{\partial^2 y^s}{\partial x^i \partial x^j}, \quad \nabla^0_{ij} = \frac{\partial x^0}{\partial y^s} \frac{\partial^2 y^s}{\partial x^i \partial x^j}
\]

\( i, j, r = 1, \ldots, n \)

\( \nabla^r_{ij} \) is called the affine connection, \( \nabla^0_{ij} = h_{ij} \) the affine metric, and \( C = \nabla h \) the cubic form on \( M \)
Centro-affine immersions

in centro-affine immersions the transversal vector field $\xi$ equals the position vector field $x$

the cubic form $C = \nabla h$ is totally symmetric
Conormal map

Let $M \hookrightarrow \mathbb{R}^{n+1}$ be a hypersurface immersion.

To each $x \in M$ we associate a vector $p \in \mathbb{R}^{n+1}$ such that:
- $p$ is tangent to $M$ at $x$;
- $\langle p, \xi \rangle = 1$ at $x$.

This hypersurface immersion $M \hookrightarrow \mathbb{R}^{n+1}$ is the conormal map.

The conormal map defines a duality on the class of centro-affine hypersurface immersions.
Outline

Geometry of self-concordant barriers
- self-concordant barriers
- affine differential geometry
- relationship between barriers and geometry
- canonical barrier
- self-scaled barriers
Centro-affine geometry of barriers

let $K \subset \mathbb{R}^n$ a regular convex cone, and $F : K^\circ \to \mathbb{R}$ a logarithmically homogeneous function of degree $-\nu$

Theorem

Let $M$ be a level surface of $F$. Then the centro-affine metric $h$ and the cubic form $C$ of $M$ on a tangent vector $u$ to $M$ are given by

$$h[u, u] = \nu^{-1} F''[u, u],$$
$$C[u, u, u] = \nu^{-1} F'''[u, u, u].$$

The immersion defined by the conormal map is a level surface of the dual barrier $F^*$.

$h, C$ are the projective counterparts of the derivatives $F'', F'''$

indeed, Karmarkar used a metric proportional to $h$ on the simplex in his algorithm
Theorem
Let $K \subset \mathbb{R}^n$, $n \geq 2$, be a regular convex cone and $F : K^o \rightarrow \mathbb{R}$ a logarithmically homogeneous locally strongly convex function with homogeneity parameter $\nu$. Let $M$ be a level surface of $F$. Then $F$ is self-concordant if and only if

$$|C[u, u, u]| \leq 2\gamma (h[u, u])^{3/2}$$

for all vectors $u$ which are tangent to $M$. Here $\gamma = \frac{\nu-2}{\sqrt{\nu-1}}$.

Corollary
On cones $K \subset \mathbb{R}^n$, $n \geq 2$, there exist no barriers with parameter $\nu < 2$. 
Dependence between $\gamma$ and $\nu$
Extreme case $\nu = 2$

**Corollary**

Let $K \subset \mathbb{R}^n$ be a regular convex cone, and $n \geq 2$. Let $F : K^o \to \mathbb{R}$ be a self-concordant barrier on $K$. Then $F$ has parameter $\nu \geq 2$, with equality if and only if $K$ is isomorphic to the Lorentz cone and $F$ to the hyperbolic barrier on $K$.

The Lorentz cone $L_n \subset \mathbb{R}^n$ is the cone

$$\left\{ x = (x_0, x_1, \ldots, x_{n-1})^T \mid x_0 \geq \sqrt{x_1^2 + \cdots + x_{n-1}^2} \right\}$$

its hyperbolic barrier is given by

$$F(x) = -\frac{1}{2} \log \left( x_0^2 - x_1^2 - \cdots - x_{n-1}^2 \right)$$

the level surfaces are isometric to hyperbolic space
Outline

**Geometry of self-concordant barriers**
- self-concordant barriers
- affine differential geometry
- relationship between barriers and geometry
- canonical barrier
- self-scaled barriers
Affine normal

non-degenerate convex hypersurface in $\mathbb{R}^n$

the **affine normal** is the tangent to the curve made of the gravity centers of the sections

a hypersurface immersion with the affine normal as transversal vector field is called a **Blaschke immersion**
Affine spheres

a hyperbolic proper affine sphere is a convex surface such that all affine normals meet at a point outside of the convex hull.

a centro-affine immersion is a proper affine sphere if and only if

▶ the affine normal is proportional to the position vector
▶ the cubic form is traceless, $C_{\alpha\beta\gamma} h^{\beta\gamma} = 0$

Theorem (Calabi conjecture; Fefferman 76, Cheng-Yau 86, Li 90, and others)

Let $K \subset \mathbb{R}^n$ be a regular convex cone. Then there exists a unique foliation of $K^\circ$ by a homothetic family of affine complete and Euclidean complete hyperbolic affine hyperspheres which are asymptotic to $\partial K$.

Every affine complete, Euclidean complete hyperbolic affine hypersphere is asymptotic to the boundary of a regular convex cone.
the foliating hyperspheres are asymptotic to the boundary of $K$
Monge-Ampère equation

characterisation of the log-homogeneous functions \( F : K^o \rightarrow \mathbb{R} \) of degree \( n \) whose level surfaces are affine spheres up to an additive constant, \( F \) is the convex solution of the Monge-Ampère equation

\[
\log \det F'' = 2F
\]

with boundary condition

\[
\lim_{x \to \partial K} F(x) = +\infty
\]

properties

- exists and is unique
- real analytic
- invariant w.r.t. unimodular linear maps
- respects Legendre duality
Canonical barrier

Theorem (H., 2014; independently D. Fox, 2015)
Let $K \subset \mathbb{R}^n$ be a regular convex cone. Then the convex solution of the Monge-Ampère equation $\log \det F'' = 2F$ with boundary condition $F|_{\partial K} = +\infty$ is a logarithmically homogeneous self-concordant barrier (the canonical barrier) on $K$ with parameter $\nu = n$.

main idea of proof: use non-positivity of the Ricci curvature [Calabi 1972]

already conjectured by O. Güler

- invariant under the action of $SL(\mathbb{R}, n)$
- fixed under unimodular automorphisms of $K$
- additive under the operation of taking products
- respects Legendre duality
Universal constructions: comparison

<table>
<thead>
<tr>
<th>Property</th>
<th>Universal barrier</th>
<th>Canonical barrier</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SL(\mathbb{R}, n)$-invariance</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Aut($K$)-invariance</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Product additivity</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Parameter</td>
<td>$O(n)$</td>
<td>$\leq n$</td>
</tr>
<tr>
<td>Duality</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Computability</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

for $K \subset \mathbb{R}^3$ with non-trivial automorphism group, the canonical barrier is given generically by elliptic integrals

for homogeneous cones the two constructions coincide

for compact sets there exists also the entropic barrier with parameter $n + O(\log n\sqrt{n})$
Outline

Geometry of self-concordant barriers
- self-concordant barriers
- affine differential geometry
- relationship between barriers and geometry
- canonical barrier
- self-scaled barriers
Self-scaled barriers

Definition
Let $K \subseteq \mathbb{R}^n$ be a regular convex cone, let $K^*$ be its dual cone, let $F$ be a self-concordant barrier on $K$ with parameter $\nu$, and let $F^*_*$ be the dual barrier on $K^*$. Then $F$ is called self-scaled if for every $x, w \in K^\circ$ we have

$$F''(w)x \in \text{int } K^*, \quad F^*_*(F''(w)x) = F(x) - 2F(w) - \nu.$$ 

A cone $K$ admitting a self-scaled barrier is called self-scaled cone.

Hauser, Güler, Lim, Schmieta 1998 – 2002:

- self-scaled cone $\iff$ symmetric cone
- self-scaled barriers on products are sums of self-scaled barriers on irreducible components
- self-scaled barriers on irreducible cones are log-determinants
Parallelism conditions

the affine connection $\nabla$ is generated by the primal immersion
the dual immersion generates the dual connection $\bar{\nabla}$
the primal-dual symmetric connection $\hat{\nabla} = \frac{1}{2}(\nabla + \bar{\nabla})$ is the
Levi-Civita connection of the affine metric

the most simple class of barriers are the hyperbolic barriers, on
whose level surfaces $C = 0$

the next class, ordered by complexity, are the barriers whose level
surfaces have constant cubic form
constant means preserved by the geodesic flow of the affine metric

$\hat{\nabla} C = 0$
Equivalence between self-scaledness and parallelism

**Theorem**

Let $K \subset \mathbb{R}^n$ be a regular convex cone and $F$ a self-concordant barrier on it. Then the following are equivalent:

- $F$ is a self-scaled barrier (and $K$ a self-scaled cone)
- on the level surfaces of $F$ the condition $\hat{\nabla} C = 0$ holds.

Every convex hyperbolic centro-affine hypersurface immersion satisfying $\hat{\nabla} C = 0$ can be completed to the level surface of a self-scaled barrier on some symmetric cone.

This yields a **local** characterization of self-scaled barriers.
Sketch of proof

\[ \nabla C = 0 \] can be rewritten as the 4-th order quasi-linear PDE

\[ F_{,\alpha\beta\gamma\delta} = \frac{1}{2} F^{,\rho\sigma}(F_{,\alpha\beta\rho} F_{,\gamma\delta\sigma} + F_{,\alpha\gamma\rho} F_{,\beta\delta\sigma} + F_{,\alpha\delta\rho} F_{,\beta\gamma\sigma}) \]

here \( F^{,\rho\sigma} \) is the inverse Hessian and \( F_{,\gamma\delta\sigma} \) etc. the partial derivatives

the integrability condition of this PDE is the Jordan identity for the algebra defined by the structure tensor \( (u \cdot v = K^{\gamma}_{\alpha\beta} u^\alpha v^\beta) \)

\[ K^{\gamma}_{\alpha\beta} = -\frac{1}{2} F_{,\gamma\delta} F_{,\alpha\beta\delta} \]

the barrier can be recovered from a metrised Euclidean Jordan algebra by

\[ F(x) = \sum_{k=2}^{\infty} \frac{(-1)^k}{k} g[x, x^{k-1}] \]
most of the proof remains valid if the convexity assumption is dropped
the appropriate framework is the theory of Koechers $\omega$-domains

<table>
<thead>
<tr>
<th>convex case</th>
<th>general case</th>
</tr>
</thead>
<tbody>
<tr>
<td>symmetric cone</td>
<td>$\omega$-domain</td>
</tr>
<tr>
<td>Euclidean Jordan algebra</td>
<td>semi-simple Jordan algebra</td>
</tr>
<tr>
<td>irreducible Euclidean Jordan algebra</td>
<td>simple Jordan algebra</td>
</tr>
<tr>
<td>canonical barrier</td>
<td>logarithmic potential $\Phi$</td>
</tr>
<tr>
<td>determinant of Jordan algebra</td>
<td>$\omega$-function</td>
</tr>
</tbody>
</table>
Affine spheres with $\hat{\nabla} C = 0$

the classification of affine spheres with parallel cubic form reduces
to the classification of semi-simple Jordan algebras
irreducible spheres / simple factors:

<table>
<thead>
<tr>
<th>vector space</th>
<th>real dimension</th>
<th>range</th>
<th>$\Phi$</th>
<th>$\omega$</th>
<th>affine sphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{C}$</td>
<td>2</td>
<td></td>
<td>$Re(\log x)$</td>
<td>$</td>
<td>x</td>
</tr>
<tr>
<td>$\mathbb{C}^m$</td>
<td>$2m$</td>
<td>$m \geq 3$</td>
<td>$Re(\log x^T x)$</td>
<td>$</td>
<td>x^T x</td>
</tr>
<tr>
<td>$S_m(\mathbb{C})$</td>
<td>$m(m+1)$</td>
<td>$m \geq 3$</td>
<td>$Re(\log \det A)$</td>
<td>$\det A^{m+1}$</td>
<td>$\det A = const$</td>
</tr>
<tr>
<td>$M_m(\mathbb{C})$</td>
<td>$2m^2$</td>
<td>$m \geq 3$</td>
<td>$Re(\log \det A)$</td>
<td>$\det A^{2m}$</td>
<td>$\det A = const$</td>
</tr>
<tr>
<td>$A_{2m}(\mathbb{C})$</td>
<td>$2m(2m-1)$</td>
<td>$m \geq 3$</td>
<td>$Re(\log \text{pf } A)$</td>
<td>$\text{pf } A^{2(2m-1)}$</td>
<td>$\text{pf } A = const$</td>
</tr>
<tr>
<td>$H_3(O, \mathbb{C})$</td>
<td>54</td>
<td></td>
<td>$Re(\log \det A)$</td>
<td>$\det A^{18}$</td>
<td>$\det A = const$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>vector space</th>
<th>real dimension</th>
<th>range</th>
<th>$\Phi$</th>
<th>$\omega$</th>
<th>affine sphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{R}$</td>
<td>1</td>
<td></td>
<td>$\log</td>
<td>x</td>
<td>$</td>
</tr>
<tr>
<td>$\mathbb{R}^m$</td>
<td>$m$</td>
<td>$m \geq 3$</td>
<td>$\log</td>
<td>x^T Q x</td>
<td>$</td>
</tr>
<tr>
<td>$M_m(\mathbb{R})$</td>
<td>$m^2$</td>
<td>$m \geq 3$</td>
<td>$\log</td>
<td>\det A</td>
<td>$</td>
</tr>
<tr>
<td>$M_m(\mathbb{H})$</td>
<td>$4m^2$</td>
<td>$m \geq 2$</td>
<td>$\log</td>
<td>\det S</td>
<td>$</td>
</tr>
<tr>
<td>$S_m(\mathbb{R})$</td>
<td>$\frac{m(m+1)}{2}$</td>
<td>$m \geq 3$</td>
<td>$\log</td>
<td>\det A</td>
<td>$</td>
</tr>
<tr>
<td>$H_m(\mathbb{C})$</td>
<td>$m^2$</td>
<td>$m \geq 3$</td>
<td>$\log</td>
<td>\det A</td>
<td>$</td>
</tr>
<tr>
<td>$H_m(\mathbb{H})$</td>
<td>$m(2m-1)$</td>
<td>$m \geq 3$</td>
<td>$\log</td>
<td>\det S</td>
<td>$</td>
</tr>
<tr>
<td>$A_{2m}(\mathbb{R})$</td>
<td>$m(2m-1)$</td>
<td>$m \geq 3$</td>
<td>$\log</td>
<td>\text{pf } A</td>
<td>$</td>
</tr>
<tr>
<td>$SH_m(\mathbb{H})$</td>
<td>$m(2m+1)$</td>
<td>$m \geq 2$</td>
<td>$\log</td>
<td>\det S</td>
<td>$</td>
</tr>
<tr>
<td>$H_3(O)$</td>
<td>27</td>
<td></td>
<td>$\log</td>
<td>\det A</td>
<td>$</td>
</tr>
<tr>
<td>$H_3(O, \mathbb{R})$</td>
<td>27</td>
<td></td>
<td>$\log</td>
<td>\det A</td>
<td>$</td>
</tr>
</tbody>
</table>
Thank you