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

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

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

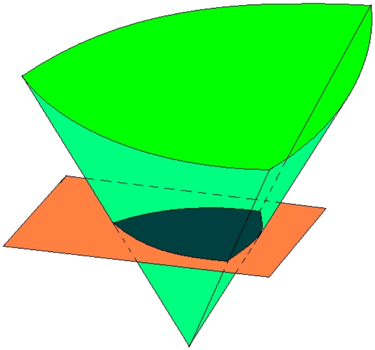
$$\min_{x \in K} \langle c, x \rangle : \quad Ax = b.$$

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

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]

LP: Interior-point
affine scaling
[Dikin 1967]
rediscovery 1986

LP: Primal-dual IP
[Kojima, Mizuno, Yoshise 1989]
[Monteiro, Adler 1989]
[Todd, Ye 1990]

Symmetric cones IP
[Nesterov, Todd 1994]
self-scaled barriers

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 $F : K^o \to \mathbb{R}$ 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 behaviour)
- $|F'''(x)[h, h, h]| \leq 2(F''(x)[h, h])^{3/2}$ (self-concordance)

for all tangent vectors h at x.
The homogeneity parameter ν is called the barrier parameter.

Theorem (Nesterov, Nemirovski 1994)

Let $K \subset \mathbb{R}^n$ be a regular convex cone and $F : K^o \to \mathbb{R}$ a barrier on K with parameter ν. Then the Legendre transform F^* is a barrier on $-K^*$ with parameter ν.

Barriers as penalty functions

let $K \subset \mathbb{R}^n$ be a regular convex cone
let $F : K^o \to \mathbb{R}$ 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 τ (*central path*)
- $x^*(\tau) \to x^*$ as $\tau \to \infty$

Path-following methods

alternate Newton steps and increments of τ

the smaller the barrier parameter ν, the faster we can increase τ 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 ν, 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 ∇ on a differentiable manifold defines the parallel transport of tangent vectors u along curves $\sigma(t)$ by

$$\dot{u}^\gamma + \nabla^\gamma_{\alpha\beta} u^\alpha \dot{\sigma}^\beta = (\frac{\partial u^\gamma}{\partial x^\beta} + \nabla^\gamma_{\alpha\beta} u^\alpha)\dot{\sigma}^\beta = 0$$

the covariant derivative of the vector field u is given by

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

we may also define the covariant derivative of general tensors

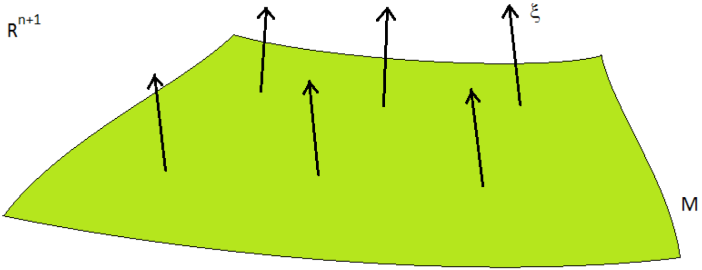
law of transformation under coordinate changes $x \mapsto y$

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

example: the flat affine connection on $\mathbb{R}^n$ is given by $\nabla^\gamma_{\alpha\beta} = 0$ in affine coordinates

Affine differential geometry

let $M \hookrightarrow \mathbb{R}^{n+1}$ be a hypersurface immersion and ξ 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 $y^0, \dots, y^n$ be affine coordinates on $\mathbb{R}^{n+1}$ and $x^1, \dots, x^n$ 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

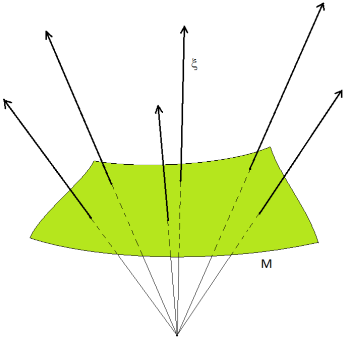
$$\nabla^r_{ij} = \frac{\partial x^r}{\partial y^s} \frac{\partial^2 y^s}{\partial x^i \partial x^j}, \qquad \nabla^0_{ij} = \frac{\partial x^0}{\partial y^s} \frac{\partial^2 y^s}{\partial x^i \partial x^j}$$

$i, j, r = 1, \dots, n$

∇^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 ξ equals the position vector field x



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

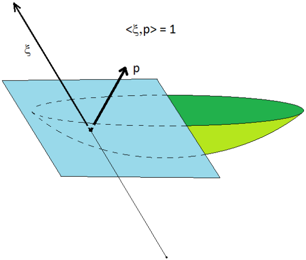
Conormal map

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 $F : K^o \to \mathbb{R}$ 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

Self-concordance and boundedness of cubic form

Theorem

Let $K \subset \mathbb{R}^n$, $n \geq 2$, be a regular convex cone and $F : K^o \to \mathbb{R}$ a logarithmically homogeneous locally strongly convex function with homogeneity parameter ν. Let M be a level surface of F.
Then F is self-concordant if and only if

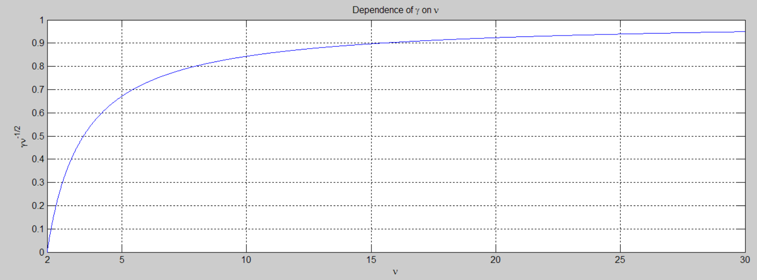
$$|C[u,u,u]| \leq 2\gamma \left(h[u,u]\right)^{3/2}$$

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 γ and ν



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

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

its hyperbolic barrier is given by

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

the level surfaces are isometric to hyperbolic space

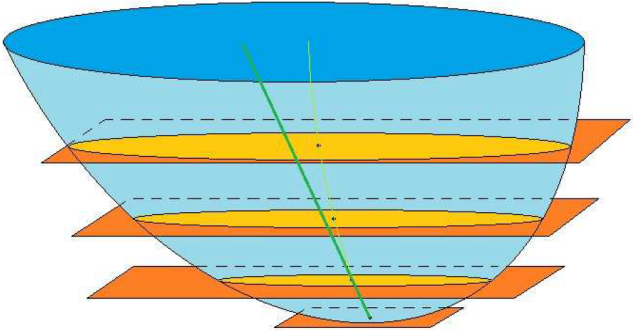
Outline

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

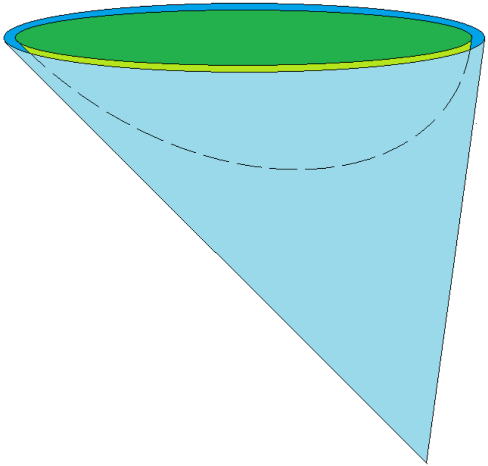
a centro-affine immersion is a proper affine sphere if and only if

- 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^o by a homothetic family of affine complete and Euclidean complete hyperbolic affine hyperspheres which are asymptotic to ∂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 $F : K^o \to \mathbb{R}$ 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

Property	Universal barrier	Canonical barrier
$SL(\mathbb{R}, n)$-invariance	Yes	Yes
$\mathrm{Aut}(K)$-invariance	Yes	Yes
product additivity	Yes	Yes
parameter	$O(n)$	$\leq n$
duality	No	Yes
computability	No	No

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

$$F''(w)x \in \operatorname{int} K^*, \qquad F_*(F''(w)x) = F(x) - 2F(w) - \nu.$$

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

Hauser, Güler, Lim, Schmieta 1998 – 2002:

- self-scaled cone $\Leftrightarrow$ 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 ∇ 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

$\hat{\nabla} C = 0$ 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 ($u \bullet v = K^{\gamma}_{\alpha\beta} u^{\alpha} v^{\beta}$)

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

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

$$F(x) = \sum_{k=2}^{\infty} \frac{(-1)^k}{k} g[x, x^{k-1}]$$

Non-convex case

most of the proof remains valid if the convexity assumption is dropped
the appropriate framework is the theory of Koechers ω-domains

convex case	general case
symmetric cone	ω-domain
Euclidean Jordan algebra	semi-simple Jordan algebra
irreducible Euclidean Jordan algebra	simple Jordan algebra
canonical barrier	logarithmic potential Φ
determinant of Jordan algebra	ω-function

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:

vector space	real dimension	range	Φ	ω	affine sphere
$\mathbb{C}$	2		$Re(\log x)$	$\|x\|^2$	$\|x\| = const$
$\mathbb{C}^m$	$2m$	$m \geq 3$	$Re(\log x^T x)$	$\|x^T x\|^m$	$\|x^T x\| = const$
$S_m(\mathbb{C})$	$m(m+1)$	$m \geq 3$	$Re(\log \det A)$	$\|\det A\|^{m+1}$	$\|\det A\| = const$
$M_m(\mathbb{C})$	$2m^2$	$m \geq 3$	$Re(\log \det A)$	$\|\det A\|^{2m}$	$\|\det A\| = const$
$A_{2m}(\mathbb{C})$	$2m(2m-1)$	$m \geq 3$	$Re(\log \operatorname{pf} A)$	$\|\operatorname{pf} A\|^{2(2m-1)}$	$\|\operatorname{pf} A\| = const$
$H_3(O, \mathbb{C})$	54		$Re(\log \det A)$	$\|\det A\|^{18}$	$\|\det A\| = const$
$\mathbb{R}$	1		$\log \|x\|$	$\|x\|$	point
$\mathbb{R}^m$	m	$m \geq 3$	$\log \|x^T Q x\|$	$\|x^T Q x\|^{m/2}$	quadric
$M_m(\mathbb{R})$	m^2	$m \geq 3$	$\log \|\det A\|$	$\|\det A\|^m$	$\det A = const$
$M_m(\mathbb{H})$	$4m^2$	$m \geq 2$	$\log \det S$	$(\det S)^{2m}$	$\det S = const$
$S_m(\mathbb{R})$	$\frac{m(m+1)}{2}$	$m \geq 3$	$\log \|\det A\|$	$\|\det A\|^{(m+1)/2}$	$\det A = const$
$H_m(\mathbb{C})$	m^2	$m \geq 3$	$\log \|\det A\|$	$\|\det A\|^m$	$\det A = const$
$H_m(\mathbb{H})$	$m(2m-1)$	$m \geq 3$	$\log \det S$	$(\det S)^{m-1/2}$	$\det S = const$
$A_{2m}(\mathbb{R})$	$m(2m-1)$	$m \geq 3$	$\log \|\operatorname{pf} A\|$	$\|\operatorname{pf} A\|^{2m-1}$	$\operatorname{pf} A = const$
$SH_m(\mathbb{H})$	$m(2m+1)$	$m \geq 2$	$\log \det S$	$(\det S)^{m+1/2}$	$\det S = const$
$H_3(\mathbb{O})$	27		$\log \|\det A\|$	$\|\det A\|^9$	$\det A = const$
$H_3(O, \mathbb{R})$	27		$\log \|\det A\|$	$\|\det A\|^9$	$\det A = const$

Thank you