Barriers on Symmetric Cones

Roland Hildebrand
Laboratoire Jean Kuntzmann (Université Grenoble Alpes) / CNRS

ICCOPT 2016, Tokyo, August 9, 2016

Outline

Conic optimization and barriers
- Conic optimization
- Logarithmically homogeneous barriers
- Geometric view on barriers

Symmetric cones and self-scaled barriers
- Symmetric cones
- Parallel extrinsic curvature

Conic programs

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

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

of a regular convex cone K is also regular.

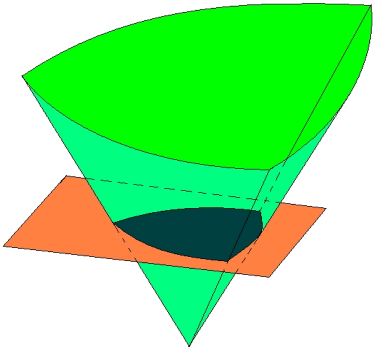
Definition

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

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

every convex optimization problem can be written as a conic program

Geometric interpretation



the feasible set is the intersection of K with an affine subspace

$$\min_x \langle c', x \rangle : \; A'x + b' \in K$$

explicit parametrization

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.

- the map $x \mapsto F'(x)$ takes the level surfaces of F to the level surfaces of F^*
- the map $x \mapsto -F'(x)$ is an isometry between K^o and $(K^*)^o$ with respect to the Hessian metrics defined by F'', $(F^*)''$

Interior-point methods

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

Second fundamental form

let $M \subset \mathcal{M}$ be a submanifold of a (pseudo-)Riemannian space

choose a point $x \in M$ and a tangent vector $h \in T_x M$

consider the geodesics $\gamma_M, \gamma_{\mathcal{M}}$ in M and in $\mathcal{M}$ through x with velocity h

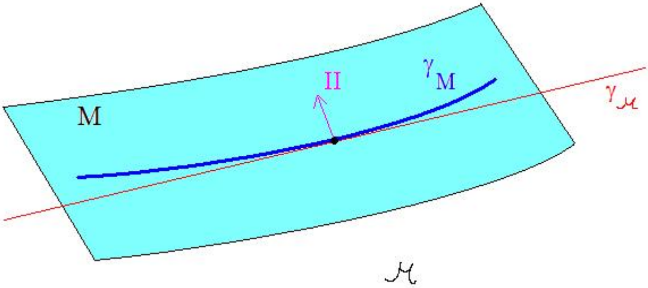
there is a second-order deviation

$$\gamma_M(t) - \gamma_{\mathcal{M}}(t) = \left(\frac{d^2}{dt^2}\bigg|_{t=0} (\gamma_M - \gamma_{\mathcal{M}}) \right) \cdot \frac{t^2}{2} + O(t^3)$$

whose main term depends quadratically on h

the acceleration is called the second fundamental form II of M

$II_x : T_x M \times T_x M \to (T_x M)^\perp$
$T_x M$ tangent subspace, $(T_x M)^\perp$ normal subspace



the second fundamental form measures the deviation of M from a geodesic submanifold

it is also called the extrinsic curvature

Para-Kähler space

consider the product $E_{2n} = \mathbb{R}^n \times \mathbb{R}_n = \{u = (x, p) \,|\, x \in \mathbb{R}^n,\ \ p \in \mathbb{R}_n\}$

for a vector space, we may identify the space with the tangent spaces at its points

E_{2n} carries natural structures:

- $||u||^2 = \langle x, p \rangle$ is a flat pseudo-Riemannian metric G with neutral signature
- $dx \wedge dp$ is a symplectic form ω, $\omega(u_1, u_2) = \frac{1}{2}(\langle x_1, p_2 \rangle - \langle x_2, p_1 \rangle)$
- $(x, p) \mapsto (x, -p)$ is an involution J whose eigenspaces define completely integrable distributions

these structures are compatible:

- $\hat{\nabla}\omega = 0$ ($\hat{\nabla}$ is the parallel transport of G)
- $Jg = \omega$

E_{2n} is a (the) flat para-Kähler space form

Barriers as Lagrangian submanifolds

duality $K \subset \mathbb{R}^n \ \leftrightarrow \ K^* \subset \mathbb{R}_n$, $x \ \leftrightarrow \ p = -F'(x)$

to a barrier F on a cone K associate the submanifold

$$M = \{(x, p) \in E_{2n} \,|\, x \in K^o, \ p = -F'(x)\}$$

the structures defined by F on K^o have a natural explanation in terms of the structures defined by E_{2n} on its submanifold M

- the metric $g = F''$ on K^o is ν times the submanifold metric on M, $g = \nu \cdot G|_M$
- M is a non-degenerate definite Lagrangian submanifold, $\omega|_M = 0$
- J is a bijection between the tangent and the normal subspaces to M
- $F''' = \omega \cdot II = Jg \cdot II$

Theorem

The self-concordance condition on F is equivalent to the boundedness of the extrinsic curvature of M. The barrier parameter ν measures the supremum of the norm of the extrinsic curvature.

- ν bounds the deviation of M from a totally geodesic submanifold of E_{2n}
- geodesic submanifolds of E_{2n} correspond to quadratic functions

Symmetric cones

Definition

A self-dual, homogeneous convex cone is called symmetric.

- self-dual: $K = K^*$
- homogeneous: Aut K acts transitively on K^o

conic programs over symmetric cones are efficiently solvable by interior-point methods due to the existence of self-scaled barriers [Nesterov, Nemirovski, 1994]

- linear programs (LP) over $\mathbb{R}^n_+$ $\sim 10^6$ variables
- conic quadratic programs (CQP) over L_n $\sim 10^4$ variables
- semi-definite programs (SDP) over $S_+(n)$ $\sim 10^2$ variables

structure can greatly increase tractable sizes

free (CLP, LiPS, SDPT3, SeDuMi, ...) and commercial (CPLEX, MOSEK, ...) solvers available

Self-scaled barriers on symmetric cones

Theorem (Vinberg, 1960; Koecher, 1962)

Every symmetric cone can be represented as a direct product of a finite number of the following irreducible symmetric cones:

- *Lorentz (or second order) cone* $L_n = \left\{ (x_0, \ldots, x_{n-1}) \mid x_0 \geq \sqrt{x_1^2 + \cdots + x_{n-1}^2} \right\}$
- *matrix cones* $S_+(n)$, $H_+(n)$, $Q_+(n)$ *of real, complex, or quaternionic hermitian positive semi-definite matrices*
- *Albert cone* $O_+(3)$ *of octonionic hermitian positive semi-definite* 3×3 *matrices*

barriers on irreducible symmetric cones

- Lorentz cone L_n: $F(x) = -\log(x_0^2 - x_1^2 - \cdots - x_{n-1}^2)$
- matrix cones: $F(X) = -\log \det X$

barriers on reducible symmetric cones
weighted sums of the barriers on the irreducible components

example: $K = \mathbb{R}_+^n$, $F(x) = -\sum_{k=1}^n \alpha_k \log x_k$, $\alpha_k \geq 1$

Main result

Theorem

Let $K \subset \mathbb{R}^n$ be a regular convex cone, and let $F : K^o \to \mathbb{R}^n$ be a convex, logarithmically homogeneous function such that $\lim_{x \to \partial K} F(x) = +\infty$. Then the following are equivalent:

- *K is a symmetric cone and F a self-scaled barrier,*
- *the product of the inversion J with the extrinsic curvature of the submanifold $M \subset E_{2n}$ is parallel with respect to the geodesic flow on K^o,*
- *the derivative F''' is parallel with respect to the geodesic flow on K^o, $\hat{\nabla} F''' = 0$.*

a barrier is self-scaled if and only if the acceleration of the geodesics on M is invariant with respect to the geodesic flow on M

the barrier F behaves in some precise sense as a primal-dual 3rd order polynomial: it is the mean between the cases when F is cubic and when F^* is cubic

the parallelism condition is local

Explicit equation

we note $\frac{\partial F}{\partial x^\alpha} = F_{,\alpha}$, $\frac{\partial^2 F}{\partial x^\alpha \partial x^\beta} = F_{,\alpha\beta}$ etc.

note $F^{,\alpha\beta}$ for the inverse of the Hessian

we adopt the Einstein summation convention over repeating indices, e.g.,

$$F^{,\alpha\beta} F_{,\beta\gamma} := \sum_{\beta=1}^{n} F^{,\alpha\beta} F_{,\beta\gamma} = \delta_\gamma^\alpha$$

then $\hat{\nabla} F''' = 0$ is equivalent to the 4-th order quasi-linear PDE

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

F is self-scaled if and only if it is a solution to this PDE

a solution can be recovered from the values of F, F', F'', F''' at a single point

Idea of proof

differentiating with respect to x^η and substituting the fourth order derivatives by the right-hand side, we get

$$
\begin{aligned}
F_{,\alpha\beta\gamma\delta\eta} = \frac{1}{4}F^{,\rho\sigma}F^{,\mu\nu}\,\big(&F_{,\beta\eta\nu}F_{,\alpha\rho\mu}F_{,\gamma\delta\sigma} + F_{,\alpha\eta\mu}F_{,\rho\beta\nu}F_{,\gamma\delta\sigma} \\
+ \quad & F_{,\gamma\eta\nu}F_{,\alpha\rho\mu}F_{,\beta\delta\sigma} + F_{,\alpha\eta\mu}F_{,\rho\gamma\nu}F_{,\beta\delta\sigma} + F_{,\beta\eta\nu}F_{,\gamma\rho\mu}F_{,\alpha\delta\sigma} \\
+ \quad & F_{,\gamma\eta\mu}F_{,\rho\beta\nu}F_{,\alpha\delta\sigma} + F_{,\beta\eta\nu}F_{,\delta\rho\mu}F_{,\alpha\gamma\sigma} + F_{,\delta\eta\mu}F_{,\rho\beta\nu}F_{,\alpha\gamma\sigma} \\
+ \quad & F_{,\delta\eta\nu}F_{,\alpha\rho\mu}F_{,\beta\gamma\sigma} + F_{,\alpha\eta\mu}F_{,\rho\delta\nu}F_{,\beta\gamma\sigma} + F_{,\delta\eta\nu}F_{,\gamma\rho\mu}F_{,\alpha\beta\sigma} \\
+ \quad & F_{,\gamma\eta\mu}F_{,\rho\delta\nu}F_{,\alpha\beta\sigma}\big)
\end{aligned}
$$

anti-commuting δ, η gives the integrability condition

$$
\begin{aligned}
F^{,\rho\sigma}F^{,\mu\nu}\,\big(&F_{,\beta\eta\nu}F_{,\delta\rho\mu}F_{,\alpha\gamma\sigma} + F_{,\alpha\eta\mu}F_{,\rho\delta\nu}F_{,\beta\gamma\sigma} + F_{,\gamma\eta\mu}F_{,\rho\delta\nu}F_{,\alpha\beta\sigma} \\
&-F_{,\beta\delta\nu}F_{,\eta\rho\mu}F_{,\alpha\gamma\sigma} - F_{,\alpha\delta\mu}F_{,\rho\eta\nu}F_{,\beta\gamma\sigma} - F_{,\gamma\delta\mu}F_{,\rho\eta\nu}F_{,\alpha\beta\sigma}\big) = 0.
\end{aligned}
$$

define a multiplication on the tangent space by

$$(u \bullet v)^\alpha = \frac{1}{2}F^{,\alpha\delta}F_{,\delta\beta\gamma}u^\beta v^\gamma$$

this defines a commutative algebra satisfying the Jordan identity

$$(u^2 \bullet v) \bullet u = (u \bullet v) \bullet u^2$$

connection between Jordan algebras and symmetric cones is long known

References

- Hessian potentials with parallel derivatives. *Results in Mathematics* **65**(3-4):399–413, 2014
- Centro-affine hypersurface immersions with parallel cubic form. *Contributions to Algebra and Geometry* **56**(2):593-640, 2015

Thank you