Geometric view-points on barrier functions as driver of recent and future progress

Roland Hildebrand

MIPT, Skoltech

October 11, 2024
ICOMP 2024, Innopolis

MIPT MOSCOW INSTITUTE OF PHYSICS AND TECHNOLOGY

ICOMP INTERNATIONAL CONFERENCE ON COMPUTATIONAL OPTIMIZATION

Skoltech Skolkovo Institute of Science and Technology

Outline

1. Barrier functions
   - Conic programs
   - Log-homogeneous barriers
   - Symmetric cones
2. Affine differential geometry
   - Log-homogeneity condition and surfaces
   - Affine differential geometry
   - Self-concordance
3. Applications
   - 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.

examples

- orthant $\mathbb{R}^n_+$
- Lorentz cone $L_n = \{(x_0, x) \mid x_0 \geq \|x\|_2\}$
- positive semi-definite matrix cones $\mathcal{S}^n_+, \mathcal{H}^n_+$

Conic programs

Definition

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

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

examples

- linear programs (LP), $K = \mathbb{R}^n_+$
- second-order cone programs (SOCP), $K = \prod_k L_{n_k}$
- semi-definite programs (SDP), $K = \mathcal{S}^n_+, \mathcal{H}^n_+$

solving a conic program necessitates a barrier on K

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 $\alpha > 0$, $x \in K^o$, and tangent vectors h at x.
The homogeneity parameter ν is called the barrier parameter.

allows to build a quantitative theory

Symmetric cones

Definition

A regular convex cone $K \subset \mathbb{R}^n$ is called self-dual if there exists a scalar product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$ such that $K = K^*$.

Definition

A regular convex cone $K \subset \mathbb{R}^n$ is called homogeneous if the automorphism group $\mathtt{Aut}(K)$ acts transitively on K^o.

Definition

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

Jordan algebras

an algebra A is a vector space V equipped with a bilinear operation $\bullet : V \times V \to V$

Definition

An algebra J is a Euclidean Jordan algebra if

- $x \bullet y = y \bullet x$ for all $x, y \in J$ (commutativity)
- $x^2 \bullet (x \bullet y) = x \bullet (x^2 \bullet y)$ for all $x, y \in J$ (Jordan identity)
- $\sum_{k=1}^m x_k^2 = 0$ implies $x_k = 0$ for all k, m (positivity)

where $x^2 = x \bullet x$.

the symmetric cones are *exactly* the cones of squares of Euclidean Jordan algebras

Self-scaled barriers

examples of logarithmically homogeneous barriers

- $K = \mathbb{R}_+^n$: $F(x) = -\sum_k \log x_k$, $\nu = n$
- $K = L_n$: $F(x_0, x) = -\log(x_0^2 - \|x\|_2^2)$, $\nu = 2$
- $K = \mathcal{S}_+^n, \mathcal{H}_+^n$: $F(X) = -\log\det X$, $\nu = n$
- $K = \prod_k K_k$: $F(x) = \sum_k F(x_k)$, $\nu = \sum_k \nu_k$

all these cones are *symmetric*

all these barriers possess an additional property: they are self-scaled
the structure tensor of the Jordan algebra is given by F'''

barriers methods for conic programming work extremely well with self-scaled barriers

Log-homogeneity

What does the log-homogeneity condition mean?

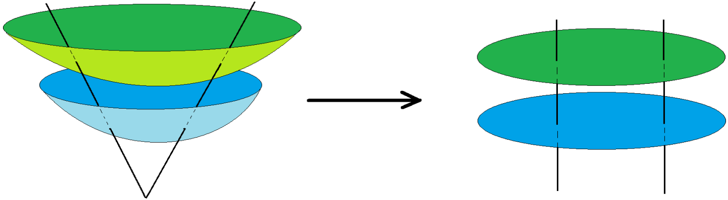
$$F(\alpha x) = -\nu \log \alpha + F(x)$$

$\alpha I : x \mapsto \alpha x$ is an *automorphism* of K
the homogeneity condition implies that F', F'', F''' are invariant under these automorphisms

consequences

- K^o splits into a product of a non-trivial transversal and trivial radial part
- level hypersurfaces of F form a homothetic family
- F can be recovered from ν and $M = \{x \mid F(x) = 0\}$ up to an additive constant

Splitting



interior K^o is diffeomorphic to a direct product of a level surface M and a radial ray

the Hessian F'' defines a Riemannian metric on the interior K^o

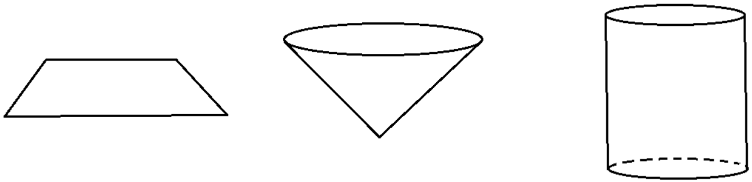
metric also splits into trivial radial factor and non-trivial factor $\tilde{h} = F''|_M$

Recovery of F from h

suppose we are given a $(n-1)$-manifold M with Riemannian metric h

Can we recover the embedding of M into $\mathbb{R}^n$ (and hence the function F) from h?

the answer is NO



does not work (different) case of submanifold metric either: flat metric on surface does not tell how surface is embedded

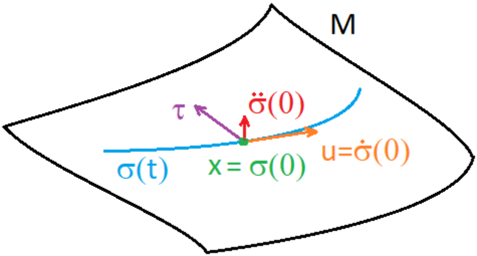
Affine differential geometry

affine differential geometry studies how metrics and other objects arise on hypersurfaces embedded in $\mathbb{R}^n$

only the affine structure of $\mathbb{R}^n$ is used, *without* supposing a Euclidean norm

- classical theory [Titeica 1909, Blaschke 1923, ...]
- motivation: find affine invariants of hypersurface embeddings (or rather immersions)
- actively developing

Affine metric

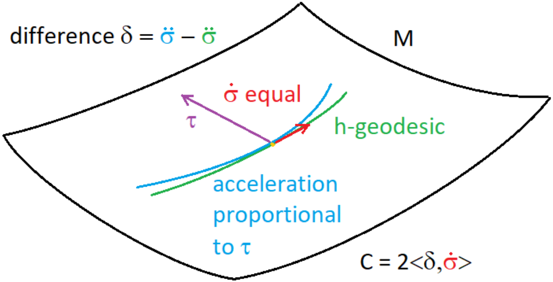


if a *transversal* vector ξ is fixed, then $\ddot{\sigma}(0) = \xi \cdot h(u) + t$ with t a tangent vector

the coefficient h is a quadratic form on $T_x M$ (independent of σ)

if the surface is *convex*, it is a Riemannian metric

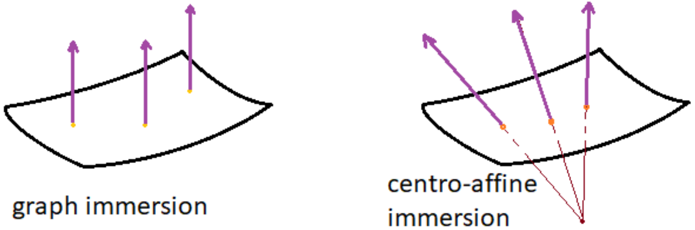
Cubic form



at every point and for every velocity vector we can compare two types of curves

the difference of the accelerations defines a tri-linear form C

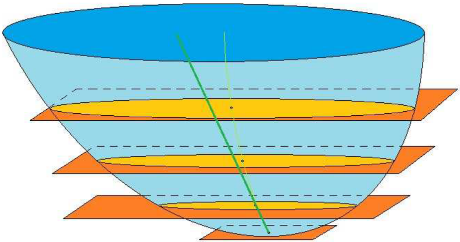
Choice of transversal field



different choices of ξ lead to different theories

- constant ξ: graph immersions
- central ξ: centro-affine immersions
- ξ is affine normal: Blaschke immersions

Affine normal



- intersect hypersurface with planes parallel to tangent plane
- assemble gravity centers into curve
- affine normal is tangent to this curve
- length determined by volume compatibility of metric and embedding space

Fundamental theorem

let $M \subset \mathbb{R}^n$ be a hypersurface with given transversal field ξ

knowing the metric h on M is not sufficient to recover how M is embedded

Theorem (Fundamental theorem of affine differential geometry)

If the metric h and the cubic form C are known on M, then the immersion of M into $\mathbb{R}^n$ can be recovered up to an affine transformation of $\mathbb{R}^n$.

for the mentioned special cases C is symmetric

Graph and centro-affine immersions

graph immersions:

- let ξ be a vertical unit basis vector
- consider the surface M as a graph $\{(z, f(z)) \mid z \in D \subset \mathbb{R}^{n-1}\}$
- then $h = f''$, $C = f'''$

centro-affine immersions:

- move centre to the origin of $\mathbb{R}^n$
- set $f(x) = \log \alpha$ such that $\alpha x \in M$
- then $h = f''|_M$, $C = f'''|_M$

Reformulation of self-concordance

let F be a self-concordant barrier and set $f = \nu^{-1} F$

then log-homogeneity becomes

$$f(\alpha x) = -\log \alpha + f(x)$$

and self-concordance becomes

$$|f'''(x)[h,h,h]| \leq 2\sqrt{\nu}(f''(x)[h,h])^{3/2}$$

hence the level surface of the barrier F can be seen as *centro-affine hypersurface immersion*

Reformulation of self-concordance

Theorem (H., 2022)

Let F be a log-homogeneous locally strongly convex C^3 function on the interior of a regular convex cone K. Then the following are equivalent:

- *F is self-concordant with parameter ν*
- *the level surfaces of F are centro-affine hypersurfaces with cubic form C bounded by $\frac{2(\nu-2)}{\sqrt{\nu-1}}$*

self-concordance is equivalent to the bounded-ness of the cubic form when measured in the centro-affine metric

Applications

affine differential geometry benefits optimization

- canonical barrier
- PDE for self-scaled cones

optimization benefits geometry

- algebraic characterization of parallelism conditions
- classification of graph immersions and centro-affine immersions with parallel C

Affine spheres

a classical question in affine differential geometry:

For which hypersurfaces $M \subset \mathbb{R}^n$ the affine normal is constant or centro-affine?

these are the definitions of improper affine spheres and proper affine spheres

Theorem (Calabi theorem)

Let $K \subset R^n$ be a regular convex cone. Then there exists a unique complete proper affine sphere with centre at the origin which is asymptotic to the boundary of K.

Canonical barrier

Theorem (H. 2014; Fox 2015)

Let $K \subset R^n$ be a regular convex cone, set $\nu = n$. Let $M \subset K^o$ be the proper affine sphere which is asymptotic to ∂K. Then the log-homogeneous function F defined from M as level surface $M = \{x \mid F(x) = 0\}$ is a self-concordant barrier, the canonical barrier.

the Calabi theorem helped to develop a *universal construction* of self-concordant barriers

Parallel cubic form

another question in affine differential geometry:

On which hypersurfaces $M \subset \mathbb{R}^n$ the cubic form C has a vanishing covariant derivative (is parallel)?

for centro-affine immersions the condition translates to a 4-th order PDE on the corresponding log-homogeneous function F

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

Connection with Jordan algebras

key observation:

if F''' is interpreted as the structure tensor of an algebra, then the integrability condition of the PDE is the Jordan identity

consequence:

- self-scaled barriers are exactly those which generate a parallel C
- self-scaled barriers can be *locally* characterized by the 4-th order PDE
- extends beyond convex case: classification of surfaces with parallel C reduced to classification of Jordan algebras

Prospective research directions

knowing a computable barrier on a cone K allows to use the cone in the formulation of efficiently solvable conic programs

the canonical barrier is not efficiently computable in general, but behind is the rich theory of affine spheres which allows to find more computable cases

- case $n = 3$ connected to the theory of Riemann surfaces and holomorphic functions
- symmetries allow to simplify the equations for the barrier

for some special non-symmetric 3D cones (power cone, exponential cone) the canonical barrier has already been computed explicitly

Thank you