Towards optimal barriers for convex cones

Roland Hildebrand

Université Grenoble Alpes / CNRS

ICCOPT 2019, Berlin
August 7, 2019

1. Motivation and problem formulation
   - Conic programs
   - Barriers
2. Optimization approach
   - Reduction to optimization problem
   - Numerical results

Conic programs

Definition

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

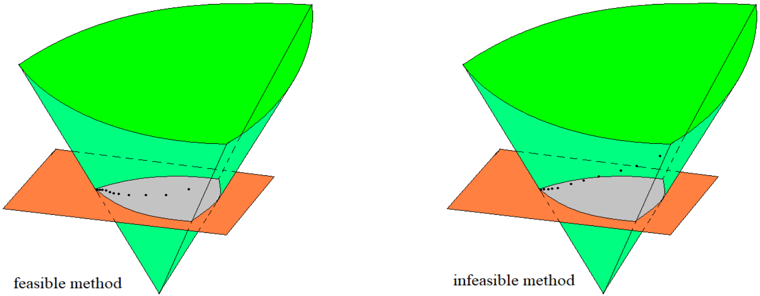
Definition

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 cast in this form

Interior-point methods



iterative methods generating a sequence of interior points

essential ingredient : self-concordant barrier on K

solvability of conic program depends on availability of barrier

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 every $x \in K^o$.
The homogeneity parameter ν is called the barrier parameter.

the smaller ν, the faster the algorithms converge

Examples

common efficiently solvable classes of conic programs

- linear programs (LP)
- second-order cone programs (SOCP)
- semi-definite programs (SDP)

LP: linear inequality constraints, $K = \mathbb{R}^n_+$

SOCP: convex quadratic constraints, $K = \prod_j L_{m_j}$,

$L_m = \{(x_0, \dots, x_{m-1})^T \mid x_0 \geq \sqrt{x_1^2 + \cdots + x_{m-1}^2}\}$

SDP : linear matrix inequalities, $K = \{A \in \mathcal{S}^{n \times n} \mid A \succeq 0\}$

for these cones barriers with the smallest possible parameter are available

Optimal barriers

Problem: for a given cone $K \subset \mathbb{R}^n$, determine the *smallest possible* value of the parameter ν of a barrier on K and construct such a barrier

problem open, e.g., for L_p cones (homogenizations of unit balls of $||\cdot||_p$ norms) and power cones

$$K_p = \{(x,y,z) \,|\, |z| \leq x^{1/p}y^{1/q},\ x,y \geq 0\},\ \frac{1}{p}+\frac{1}{q}=1$$

Bounds on the parameter ν

which parameter values ν can a barrier on $K \subset \mathbb{R}^n$ have?

- if K has a corner, then $\nu \geq k$, where k is the number of half-planes meeting [Nesterov, Nemirovski 1994]
- if K is homogeneous, then $\nu = r$ is optimal, where r is the rank of K [Güler, Tunçel 1998]
- $\nu \leq n$ always possible (canonical barrier [H. 2014], universal barrier [Lee,Yue 2018])
- for $n = 3$, $\nu < 3$ possible if and only if $\mathbb{R}^3_+$ cannot be approximated by linear images of K [Benoist, Hulin 2014]

optimal value of parameter is invariant under linear isomorphisms and duality

Optimizing the barrier parameter

up to now, the analytic description of the cone boundary is used to construct an analytic barrier, either on the cone itself or on a lifting of K

approach: interpret the search for an optimal barrier as an *optimization problem*

- infinite-dimensional (functional space)
- non-convex (self-concordance condition not stable under convex combinations)
- constraints contain derivatives
- boundary constraints

Outline of strategy

reduction to infinite-dimensional, non-convex, but otherwise "classical" optimization problem:

- use logarithmic homogeneity to pass to compact section C of K: values on K^o can be recovered from values on C^o
- for each point $x \in C^o$, introduce a vector variable $p(= \frac{F'(x)}{\nu})$ and a symmetric matrix variable $H(= p'(x))$
- replace boundary condition by reachibility condition
- replace self-concordance and integrability conditions by finite difference relations on p, H
- replace minimization of ν by bisection and feasibility check

Logarithmic homogeneity

homogeneity constraints are *affine-linear* in F and its derivatives

consider compact affine section C of K

- reduces dimension of independent variable
- F on K^o recovered by homogeneity from values on C^o
- boundary condition not affected
- convexity on C^o is necessary, but not sufficient for convexity on K^o
- self-concordance on C^o is necessary, but not sufficient for self-concordance on K^o

Convexity and self-concordance on C

Theorem

The barrier $F = \nu \cdot f$ is self-concordant with parameter ν on K if and only if

$$f''(x) - f'(x) \otimes f'(x) \succ 0$$

for all $x \in C^o$ and

$$f'''(x)[h,h,h] - 6f''(x)[h,h]f'(x)[h] + 4(f'(x)[h])^3 \leq$$

$$\leq 2\gamma(f''(x)[h,h] - (f'(x)[h])^2)^{3/2}$$

for all $x \in C^o$, $h \in T_xC^o$, where $\gamma = \frac{\nu-2}{\sqrt{\nu-1}}$.

smallest possible value $\gamma = 0$ if and only if $\nu = 2$

2-dimensional sections

observation: if a function $F = \nu f : K^o \to \mathbb{R}$ is a self-concordant barrier with parameter ν on every 2-dimensional linear section of K^o, then it is so on K

2-dimensional sections of K become 1-dimensional sections of C — finite intervals I

consider restrictions $\nu^{-1}F|_I = f$ with h parallel to I

- $f' = \langle p, h \rangle, f'' = \langle Hh, h \rangle$ depend linearly on p, H
- boundary condition: $f \to +\infty$ at ends of I
- convexity: $f'' - (f')^2 > 0$
- self-concordance: $|f''' - 6f''f' + 4(f')^3| \leq 2\gamma(f'' - (f')^2)^{3/2}$

Upper and lower bounds

let $I = (a, b)$ be an interval, $f : I \to \mathbb{R}$ satisfying the boundary and self-concordance conditions, and $x_0 \in I$ arbitrary

Which values can $(f'(x_0), f''(x_0))$ take?

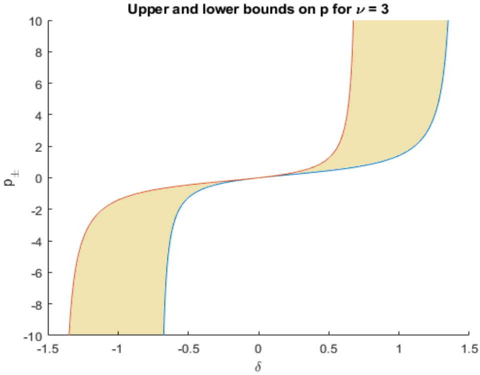
consider the differential inclusion

$$f''' = 6f''f' - 4(f')^3 + 2u\gamma(f'' - (f')^2)^{3/2}, \ u \in [-1, 1]$$

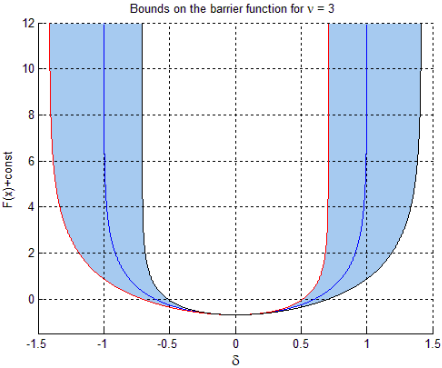
given an initial condition $(f'(x_0), f''(x_0)) = (p_0, g_0^2 + p_0^2)$, the solutions with $u = \pm 1$ yield upper and lower bounds on $f(x)$ and $p = f'(x_0 + \delta)$

$$p_\pm = \frac{p_0 + \delta(g_0^2 - p_0^2 \mp \gamma g_0 p_0)}{(p_0\delta - 1)^2 - g_0^2\delta^2 \pm \gamma g_0 \delta(p_0\delta - 1)}$$

bounds on $p = f'$ for $f'(x_0) = 0$, $f''(x_0) = 1$, $\gamma = \frac{\sqrt{2}}{2}$:



bounds on f for $f'(x_0) = 0$, $f''(x_0) = 1$, $\gamma = \frac{\sqrt{2}}{2}$:



upper bound has to tend to $+\infty$ inside I
lower bound has to tend to $+\infty$ outside I

Boundary condition

Lemma

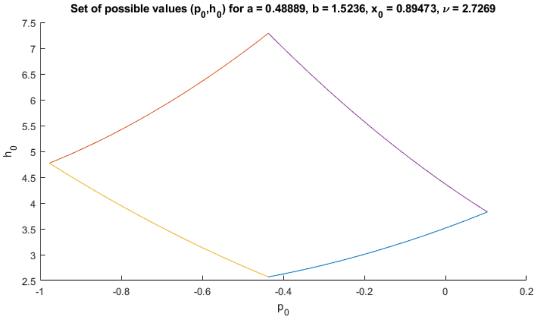
Let $x_0 \in I = (a, b)$ and (p_0, g_0) be given. There exists a self-concordant function $F = \nu f$ with $f'(x_0) = p_0$, $f''(x_0) = g_0^2 + p_0^2$ escaping to infinity when x tends to a or b if and only if

$$\frac{g_0}{\sqrt{\nu - 1}} \leq p_0 + (x_0 - a)^{-1} \leq g_0\sqrt{\nu - 1},$$

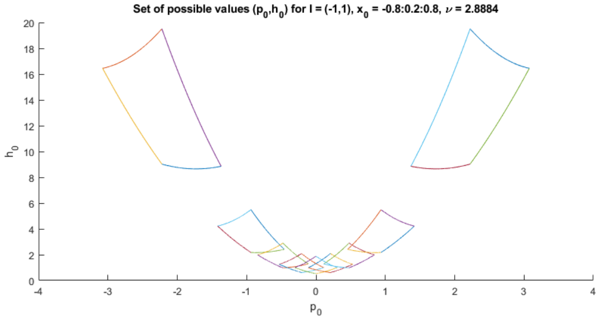
$$-g_0\sqrt{\nu - 1} \leq p_0 - (b - x_0)^{-1} \leq -\frac{g_0}{\sqrt{\nu - 1}}$$

these conditions delineate a compact set of possible values for $(f'(x_0), f''(x_0))$

feasible set for $(f'(x_0), f''(x_0))$



boundary condition is expressed by 2 convex quadratic and 2 concave quadratic constraints on p, H for every $x_0 \in C^o$ and every direction feasible sets of $(f'(x_0), f''(x_0))$ for different values of x_0



Self-concordance condition

Lemma

Let $p, h : I \to \mathbb{R}$ be functions on an interval. There exists a convex function $f : I \to \mathbb{R}$ such that

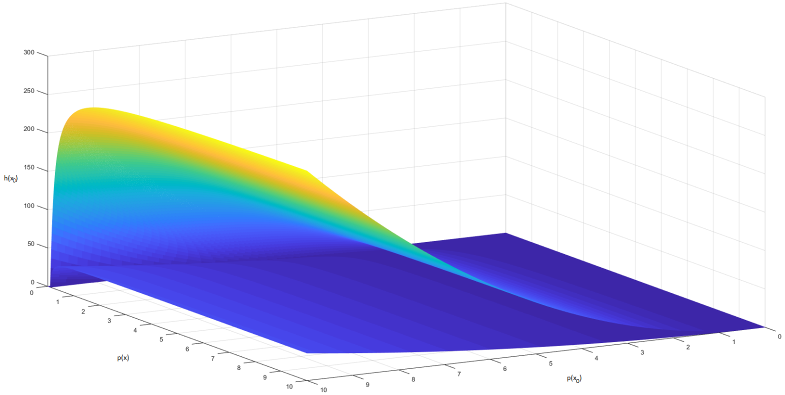
$$|f''' - 6f''f' + 4(f')^3| \leq 2\gamma(f'' - (f')^2)^{3/2}$$

a.e. on I with $p = f'$, $h = f''$ if and only if for every $x_0 \in I$ the condition

$$p_- \leq p \leq p_+$$

holds, where $p_\pm$ are the bounds of p for the initial values $p_0 = p(x_0)$, $g_0 = \sqrt{h(x_0) - p_0^2}$.

set of possible pairs $(p(x_0), h(x_0), p(x))$ allowed by self-concordance



self-concordance is expressed by two algebraic inequalities on p, H for every ordered pair of points $x_0, x \in C^o$

Integrability condition

how to ensure that H is continuous and $H = p'$?

for $x \in C^o$ and non-zero $u \in T_x C^o$ define
$t_-(x,u) = \min t : x + tu \in C$, $t_+ = \max t : x + tu \in C$

let $F = \nu f : C^o \to \mathbb{R}$ be self-concordant, and $x_0, x \in C^o$ such that $\Delta = \frac{\sqrt{t_+(x_0,x-x_0)t_-(x_0,x-x_0)+\frac{\nu}{4}(t_+(x_0,x-x_0)-t_-(x_0,x-x_0))^2}}{t_+(x_0,x-x_0)t_-(x_0,x-x_0)} < \nu^{-1/2}$

then

$$\begin{pmatrix} \left(\frac{\sqrt{\nu}\Delta^2}{1-\sqrt{\nu}\Delta}\right)^2 & \star \\ p(x) - p(x_0) - H(x_0)(x - x_0) & H(x_0) \end{pmatrix} \succeq 0,$$

$$(1 - \sqrt{\nu}\Delta)^2 H(x_0) \preceq H(x) \preceq (1 - \sqrt{\nu}\Delta)^{-2} H(x_0)$$

we have $\Delta = O(x - x_0)$ for x close to x_0

Main result

Let $p : C^o \to \mathbb{R}^n$, $H : C^o \to \mathcal{S}^n_+$. There exists a self-concordant barrier $F = \nu f : C^o \to \mathbb{R}$ with $p = f'$, $H = f''$ if and only if

$$\frac{g_0}{\sqrt{\nu - 1}} \leq p_0 + (x_0 - a)^{-1} \leq g_0\sqrt{\nu - 1},$$

$$-g_0\sqrt{\nu - 1} \leq p_0 - (b - x_0)^{-1} \leq -\frac{g_0}{\sqrt{\nu - 1}}$$

hold for every $x_0 \in C$ and every interval I containing x_0, and

$$p_- \leq p \leq p_+,$$

$$\begin{pmatrix} \left(\frac{\sqrt{\nu}\Delta^2}{1-\sqrt{\nu}\Delta}\right)^2 & \star \\ p(x) - p(x_0) - H(x_0)(x - x_0) & H(x_0) \end{pmatrix} \succeq 0,$$

$$(1 - \sqrt{\nu}\Delta)^2 H(x_0) \preceq H(x) \preceq (1 - \sqrt{\nu}\Delta)^{-2} H(x_0)$$

hold for every ordered pair $x_0, x \in C^o$.

Convexification

the resulting optimization problem is still non-convex and infinite-dimensional

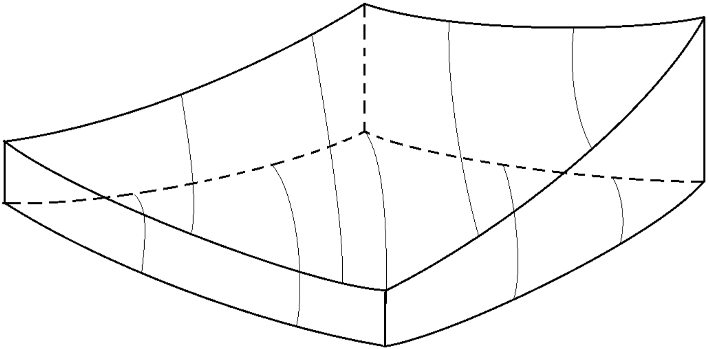
for $x_0, x \in C^o$, $\delta = x - x_0$, convexify the set of

$$(p_0, h_0, p) = (\langle p(x_0), \delta\rangle, \langle H(x_0)\delta, \delta\rangle, \langle p(x), \delta\rangle)$$

satisfying the scalar inequality constraints

for x close enough to x_0 this set is compact and its convex hull equals the convex hull of 8 rational curve segments of order 4

$\Rightarrow$ convex hull described by LMIs of order ≤ 3



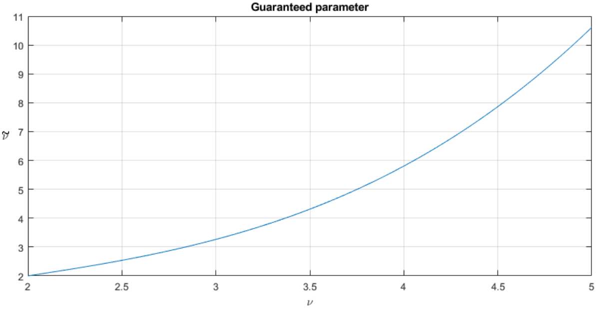
for x close enough to x_0 the upper bound p_+ of p is convex with respect to h_0 and the lower bound p_- is concave

Performance guarantee

Lemma

Let $p : C^o \to \mathbb{R}^n$, $H : C^o \to \mathcal{S}_+^n$ be functions satisfying the convexified conditions for some value $\nu > 2$.
Then there exists $f : C^o \to \mathbb{R}$ with $f' = p$, $f'' = H$ such that for every $\epsilon > 0$ the function $F = (\tilde{\nu} + \epsilon)f$ is self-concordant with parameter $\tilde{\nu} + \epsilon$ on C^o. Here $\tilde{\nu}$ is a piece-wise algebraic function of ν satisfying

$$\tilde{\nu} = \nu + \frac{3}{8}(\nu - 2)^2 + O((\nu - 2)^3).$$


Guaranteed parameter
11
10
9
8
7
6
5
4
3
2
2
2.5
3
3.5
4
4.5
5
$\tilde{\nu}$
ν

Generalized self-concordance

a more natural definition of self-concordance

Definition

Let K be a regular convex cone. We call a C^2 function $F : K^o \to \mathbb{R}$ a logarithmically homogeneous self-concordant barrier in the generalized sense on K with parameter ν if

- $F(\alpha x) = -\nu \log \alpha + F(x)$
- $F''(x) \succ 0$
- $\lim_{x \to \partial K} F(x) = +\infty$
- $\limsup_{\epsilon \to 0} \frac{|F''(x+\epsilon h)[h,h] - F''(x)[h,h]|}{\epsilon} \leq 2(F''(x)[h,h])^{3/2}$

for all tangent vectors h and $x \in K^o$.

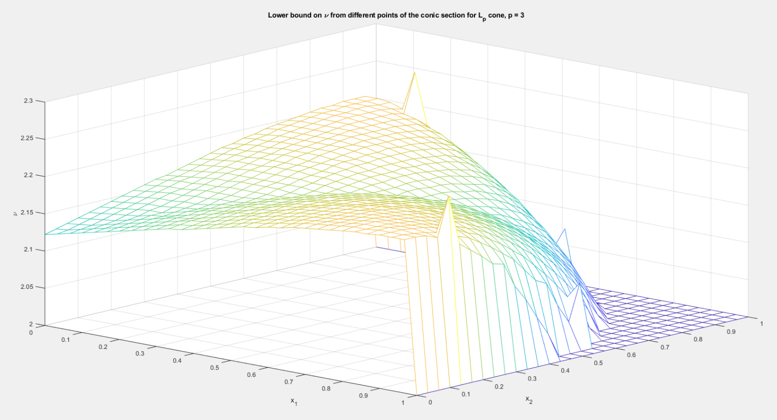
Sampling of C^o

to obtain an ordinary SDP we consider only a sample of points from C^o and take into account convexified constraints for each pair

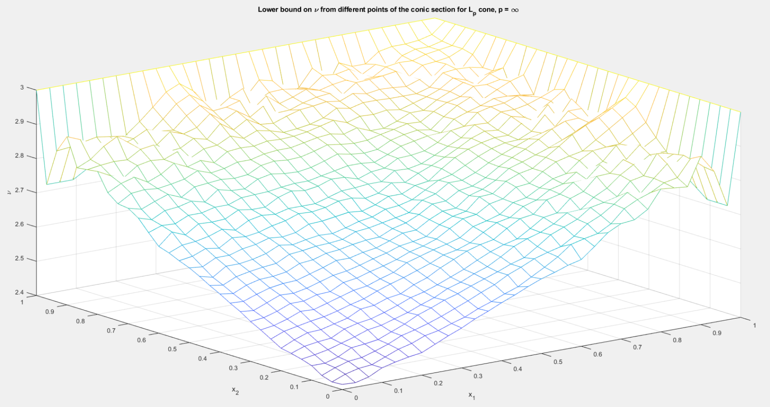
- SDP consists of many low order LMIs
- suitable only for cones of small dimension
- sample must be comparatively large in order to activate self-concordance and integrability conditions
- no performance guarantee

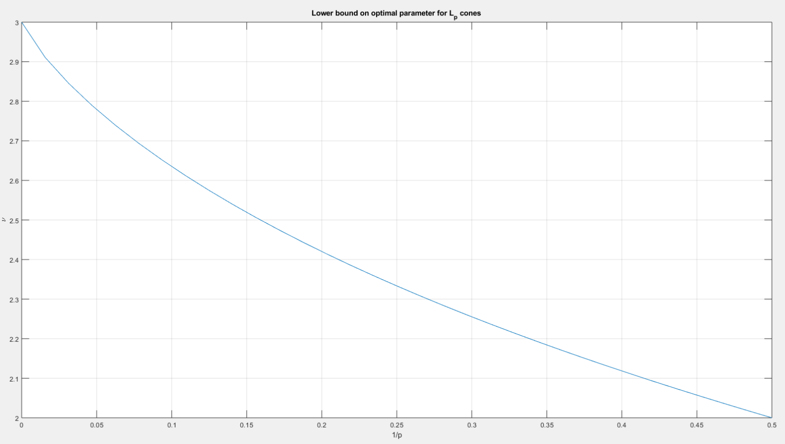
future work:

- tailored samples
- obstacles for decreasing of ν from dual variables



only convexified constraints from boundary conditions (separate SOCP for each point in C^o)


Lower bound on ν from different points of the conic section for L_p cone, p = ∞
ν
3
2.9
2.8
2.7
2.6
2.5
2.4
x_2
x_1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Lower bound on optimal parameter for L_p cones
3
2.9
2.8
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
1/p

Thank you