On the convexity properties of the barrier parameter in conic programming

Roland Hildebrand

Université Grenoble Alpes / CNRS

RICAM Workshop on Conic optimization, Linz
December 9, 2019

1. Motivation and problem formulation
   - Self-concordant barriers
   - Logarithmically homogeneous barriers

2. Solution
   - Methods
   - Results

Newton method

iterative method minimizing local quadratic approximation of cost function F

$$x_{k+1} = x_k - (F''(x_k))^{-1} F'(x_k) = \arg\min_x q_k(x),$$

where

$$q_k(x) = F(x_k) + \langle F'(x_k), x - x_k \rangle + \frac{1}{2} \langle F''(x_k)(x - x_k), x - x_k \rangle$$

current iterate characterized by Newton decrement

$$\rho_k = ||F'(x_k)||_k = \sqrt{2(q_k(x_k) - q_k(x_{k+1}))} = ||x_{k+1} - x_k||_k$$

here $||\cdot||_k$ is the local metric defined by $F''(x_k)$

Motivation of self-concordance

Newton method will work well if the quadratic approximation q_k is still reasonably good at the new point x_{k+1}

q_{k+1} should not be too far away from q_k:

$$\frac{||q_{k+1} - q_k||}{||x_{k+1} - x_k||} \leq L$$

- $||q_{k+1} - q_k|| \sim$ rate of change F'''
- $||x_{k+1} - x_k||$ measured in local metric $||\cdot||_k \sim F''$

Self-concordant functions

Definition

Let $D \subset \mathbb{R}^n$ be a convex domain and $F : D \to \mathbb{R}$ a convex C^3 function. The function F is called self-concordant if for all $x \in D$ and all $u \in T_x D$ we have

$$|F'''(x)[u, u, u]| \leq 2(F''(x)[u, u])^{3/2},$$

and self-concordant barrier if in addition $\lim_{x \to \partial D} F(x) = +\infty$.

power $\frac{3}{2}$ in order to homogenize inequality with respect to u

Newton method and self-concordance

let $F : D \to \mathbb{R}$ be a self-concordant barrier

apply the Newton method to minimize F on D

self-concordance guarantees that

- if $\rho_k < 1$, then $x_{k+1} \in D$
- $\rho_{k+1} \leq \left(\frac{\rho_k}{1-\rho_k}\right)^2$

quadratic convergence in the vicinity of the minimum

used in interior-point methods for convex programming

Barrier parameter

let $F : D \to \mathbb{R}$ be a self-concordant barrier

the quantity

$$\nu = \sup_{x \in D} \rho^2 = \sup_{x \in D} ||F'(x)||^2_{F''(x)}$$

is called the barrier parameter of F

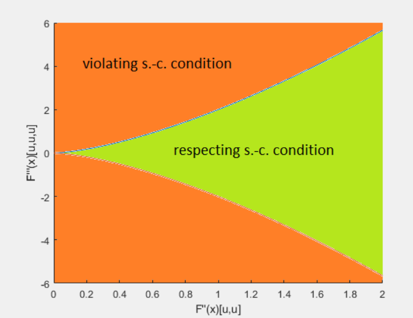
the smaller the parameter ν, the larger the steps of the optimization method and the faster the convergence

number of iterations scales like $\nu^{1/2}$

Problem

For given $D \subset \mathbb{R}$, find the barrier F with the lowest parameter ν on D.

we study the convexity properties of this problem



convex combinations of self-concordant functions are not necessarily self-concordant

convex hull of possible pairs $(F''(x)[u,u], F'''(x)[u,u,u])$ is the whole right half-plane

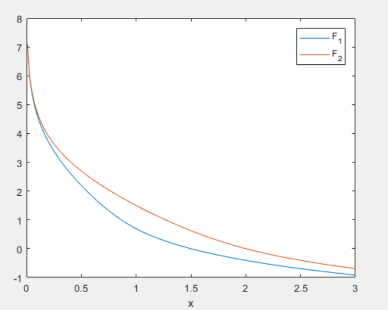
Example

consider the domain $D = \mathbb{R}_{++}$ and the self-concordant functions

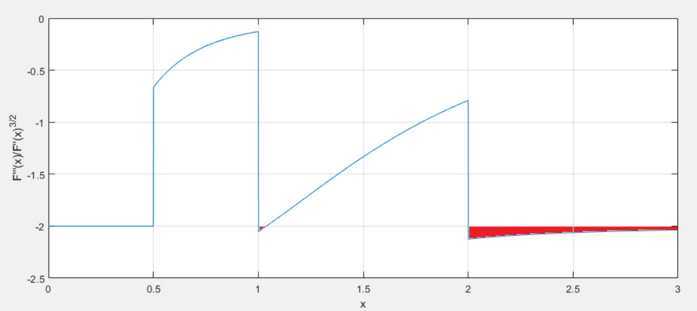
$$F_1(x) = \begin{cases} -\log x + \frac{5}{2} - 2x, & x \leq \frac{1}{2}, \\ \log 2 + 2(1-x) + 2(1-x)^2, & \frac{1}{2} < x \leq 1, \\ -\log(x - \frac{1}{2}), & x > 1, \end{cases}$$

$$F_2(x) = \begin{cases} -\log x + \frac{5}{2} - x, & x \leq 1, \\ (2-x) + \frac{(2-x)^2}{2}, & 1 < x \leq 2, \\ -\log(x-1), & x > 2, \end{cases}$$

set $F = \frac{2}{3}F_1 + \frac{1}{3}F_2$



F_1, F_2 are self-concordant



$F = \frac{2}{3}F_1 + \frac{1}{3}F_2$ is not self-concordant

Cones and conic programs

Definition

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

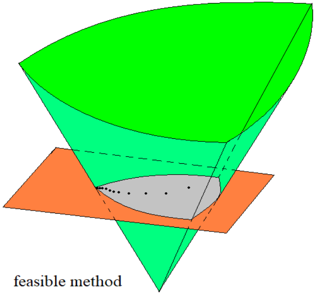
Definition

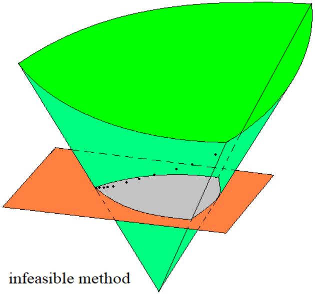
A conic program over a regular convex cone $K \subset \mathbb{R}^n$ is an optimization problem of the form

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

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

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, $K = \prod_j L_{m_j}$,

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

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

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)
- $|F'''(x)[u,u,u]| \leq 2(F''(x)[u,u])^{3/2}$ (self-concordance)

for all tangent vectors u at every $x \in K^o$.

ν is the parameter: $F''(x)x = -F'(x)$, $\langle F'(x), x \rangle = -\nu$
$\Rightarrow \langle (F''(x))^{-1} F'(x), F'(x) \rangle = \nu$

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 $x \in K^o$.

Compatibility with convexity

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

cost function linear but feasible set not convex

but:

- the feasible set is mapped into itself by multiplication by constants ≥ 1
- the cost function increases under multiplication by constants ≥ 1

Scaling

by multiplying the function F by a constant α we achieve the transformation

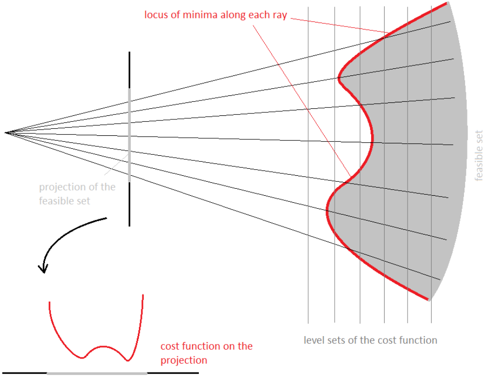
$$(\nu, F'', F''') \mapsto (\alpha\nu, \alpha F'', \alpha F''')$$

hence we may replace the objective ν by the homogeneous function

$$\sup_{x \in D, u \in T_x D} \frac{\nu(F'''(x)[u,u,u])^2}{4(F''(x)[u,u])^3}$$

and consider only the slice $\nu = 1$

Transfer of non-convexity



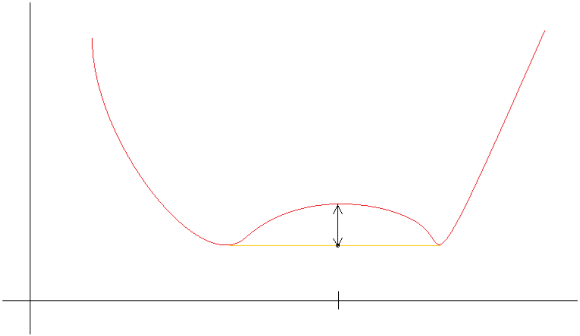
the feasible set becomes convex, the cost function non-convex

Problem formulation

How much do we lose by convexification of the cost function?

in other words:

Given a convex combination F of barriers on K with parameter ν, what is the minimal value $c > 1$ such that $c \cdot F$ is guaranteed to be self-concordant?



How large ist the performance loss when using a minimum of the convexified cost function instead of a minimum of the original non-convex cost?

Level set

let $K \subset \mathbb{R}^n$ be a regular convex cone
let $\nu > 0$ be fixed

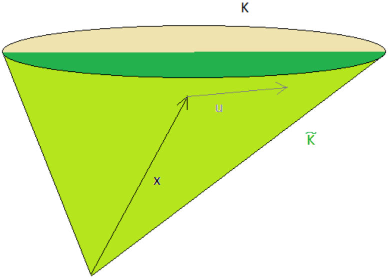
let $F : K^o \to \mathbb{R}$ in the sub-level set corresponding to ν

- $F(\alpha x) = -\log \alpha + F(x)$, $\alpha > 0$, $x \in K^o$
- $F''(x) \succ 0$, $x \in K^o$
- $\lim_{x \to \partial K} F(x) = +\infty$
- $|F'''(x)[u,u,u]| \leq 2\sqrt{\nu}(F''(x)[u,u])^{3/2}$

setting $u = x$ we get $\nu \geq 1$

Reduction to dimension 2

let $x \in K^o$ be an interior point
let $x \not\parallel u \in T_x K^o$ be a tangent vector
let L be the span of x, u



then $F''(x)[u,u]$, $F'''(x)[u,u,u]$ depend only on restriction of F to $\tilde{K}^o = (K \cap L)^o$

Normalization

normalize coordinate system such that
$K = L_2 = \{(x_0, x_1) \mid x_0 \geq |x_1|\}$, $x = (1, 0)$, $u = (\tau, 1)$, then

$$F'''(x)[\cdot, \cdot, e_0] = -2 \begin{pmatrix} 1 & a \\ a & b \end{pmatrix}, \quad F'''(x)[\cdot, \cdot, e_1] = -2 \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

$$F''(x) = \begin{pmatrix} 1 & a \\ a & b \end{pmatrix}, \qquad F'(x) = -\begin{pmatrix} 1 \\ a \end{pmatrix}$$

$$|\tau^3 + 3a\tau^2 + 3b\tau + c| \leq \sqrt{\nu}(\tau^2 + 2a\tau + b)^{3/2} \qquad \forall\, \tau \in \mathbb{R}$$

Positivity condition

replace τ by $\tau - a$, then we get

$$(\tau^3 + 3(b - a^2)\tau + c - 3ab + 2a^3)^2 \leq \nu(\tau^2 + b - a^2)^3 \quad \forall\, \tau$$

this can be rewritten as

$$|c - 3ab + 2a^3| \leq \frac{\nu - 2}{\sqrt{\nu - 1}}(b - a^2)^{3/2} \qquad (\Rightarrow \;\; \nu \geq 2)$$

the expressions $b - a^2$, $c - 3ab + 2a^3$ encode the *affine metric* and the *cubic form* of the level curve of F

they are projectively invariant and hence the above inequality is valid for any

$(a, b, c) = (-F'(x)[u], F''(x)[u, u], -\frac{1}{2}F'''(x)[u, u, u])$

Controlled system

set $f(t) = F(1,t)$, $t \in (-1,1)$, then

$$f'' > (f')^2, \quad \lim_{t \to \pm 1} f(t) = +\infty$$

$$\frac{1}{2} f''' - 3f'f'' + 2(f')^3 = \frac{\nu - 2}{\sqrt{\nu - 1}} u(f'' - (f')^2)^{3/2}, \quad u \in [-1,1]$$

reachability problem: for which initial conditions

$$f'(0) = -a, \ f''(0) = b, \ f'''(0) = -2c$$

the problem has a solution?

Feasible set for a, b, c

Theorem

The preceding problem has a solution if and only if

$$\frac{\sqrt{b-a^2}}{\sqrt{\nu-1}} \leq 1-a \leq \sqrt{\nu-1}\sqrt{b-a^2},$$

$$\frac{\sqrt{b-a^2}}{\sqrt{\nu-1}} \leq 1+a \leq \sqrt{\nu-1}\sqrt{b-a^2},$$

$$|c-3ab+2a^3| \leq \frac{\nu-2}{\sqrt{\nu-1}}(b-a^2)^{3/2}.$$

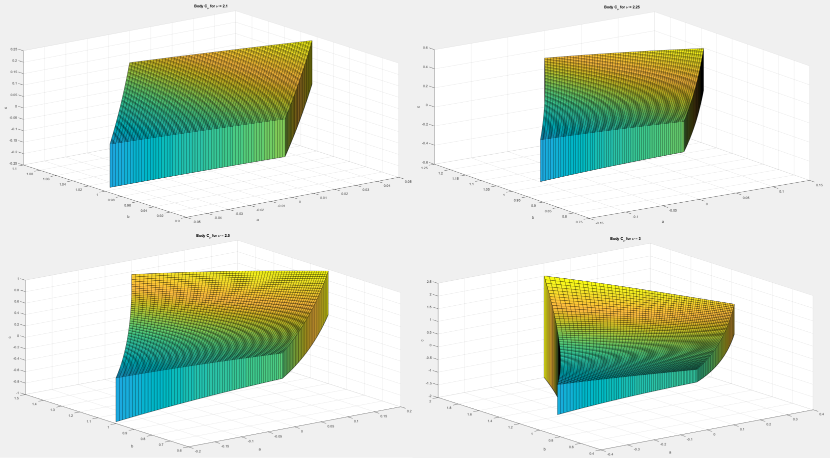
denote the corresponding non-convex body in $\mathbb{R}^3$ by C_ν

Bodies C_ν

properties of C_ν

- C_ν compact
- $C_{\nu'} \subset C_\nu$ for $\nu' \leq \nu$
- $C_2 = \{(0, 1, 0)\}$
- $\bigcup_{\nu \geq 2} C_\nu = \{(a, b, c) \mid b > 0\}$
- for every ν there exists $\tilde{\nu} \geq \nu$ such that $C_{\tilde{\nu}}$ contains the convex hull of C_ν
- there exists a minimal such $\tilde{\nu}$

Bodies C_ν



Main result

Theorem

Let $K \subset \mathbb{R}^n$ be a regular convex cone, and let $\nu \geq 2$.

- *Let $\tilde{\nu} \geq 2$ be such that $C_{\tilde{\nu}}$ contains the convex hull of C_ν. Then every convex combination of barriers with parameter ν on K yields a barrier in the generalized sense with parameter $\tilde{\nu}$ on K when multiplied by $\frac{\tilde{\nu}}{\nu}$.*
- *Let $\tilde{\nu}$ be such that $C_{\tilde{\nu}}$ does not contain the convex hull of C_ν. Then there exists a convex combination of barriers with parameter ν on K which cannot be scaled into a barrier with parameter $\tilde{\nu}$ on K.*

Convex hull

for fixed a the expressions $\pm(b - a^2)^{3/2}$ are convex (concave) in b

hence the convex hull of C_ν equals the convex hull of the edges, which are rational curve segments

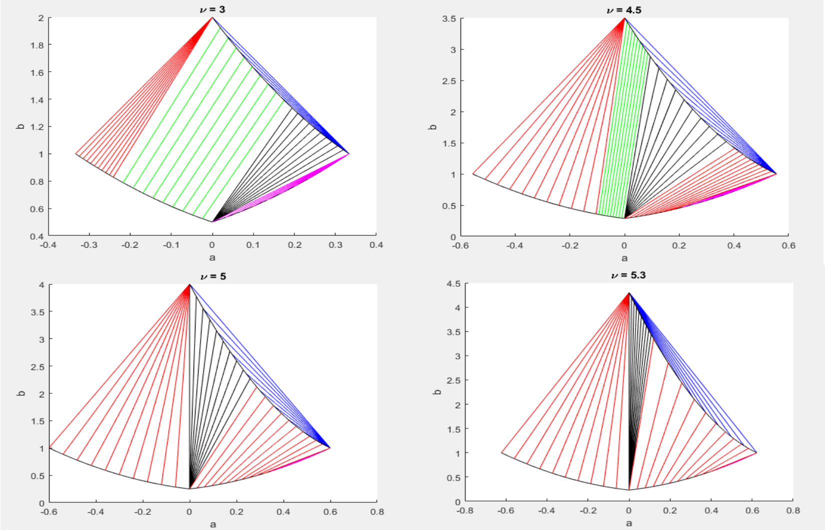
$\Rightarrow$ convex hull is semi-definite representable

convex hull can be analytically computed

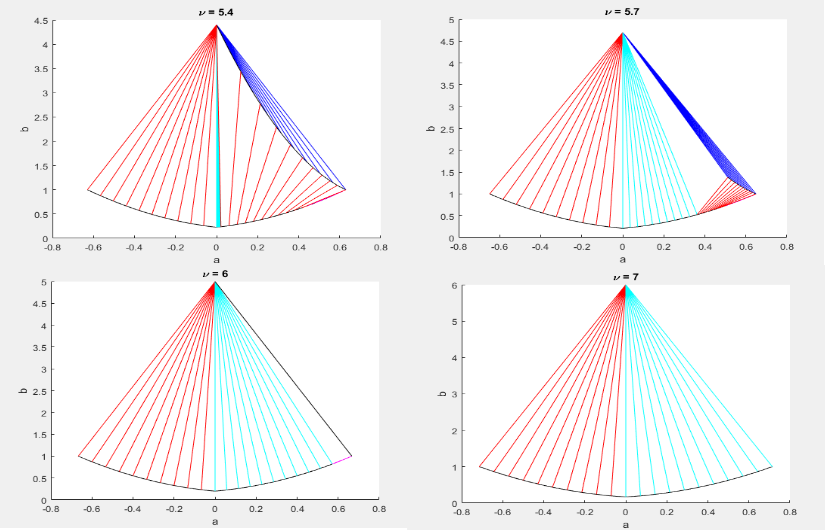
qualitatively different for different ν

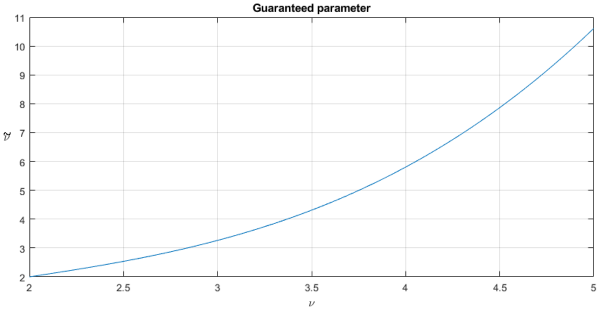
structural changes at $\nu =$ $3.9718553726, 4.8473221018, 5.2360679774, 5.3770889307, 5.4716822838, 5.8892812008, 6.2802453362$

Convex hull of C_ν



Convex hull of C_ν





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

piece-wise algebraic function of ν
for large ν we get $\tilde{\nu} \sim \nu^3$

Thank you