Workshop GeoLMI on the geometry ...

Workshop GeoLMI on the geometry and algebra of linear matrix inequalities

19-20 November 2009

LAAS-CNRS

University of Toulouse, France

The workshop aims at studying connections between real algebraic geometry and semidefinite programming, with the objective of designing algorithms to model convex semi-algebraic sets as linear matrix inequalities (LMI, affine sections and projections of the cone of positive semidefinite matrices).

Quadratic approximation of some convex optimization problems using the arithmetic-geometric mean iteration

François Glineur

francois.glineur@uclouvain.be

UCL/CORE (Center for Operations Research and Econometrics)
UCL/INMA (Département d'ingénierie mathématique)

November 20, 2009

GeoLMI Toulouse

Overview

1. Motivation

- ◇ Convex optimization: problem classes
- ◇ Approximations: direct vs. extended formulations

2. Quadratic approximations of convex optimization problems

- ◇ Arithmetic-geometric mean iteration
- ◇ Application to approximations

3. Generalizations and conclusions

- ◇ Matrix version
- ◇ Conclusions
- ◇ Addendum

Overview

Convex optimization

Nonlinear optimization

$$\min_{x\in\mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

- ◇ Variables: finite-dimensional vector $x \in \mathbb{R}^n$
- ◇ Constraints: finite number of (in)equalities, indexed by sets I and E

Problem is convex when

- ◇ objective function f_0 is convex
- ◇ functions f_i defining inequalities $f_i(x) \leq 0$ are convex for all $i \in I$
- ◇ functions f_i defining equalities $f_i(x) = 0$ are affine for all $i \in E$

Well-known classes of convex problems

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

1. Linear optimization (LO): f_0 and f_i are affine for all $i \in E \cup I$

$$f_i(x) = a_i^{\mathrm{T}} x - b_i$$

2. Quadratically constrained quadratic optimization (QCQO): f_0 and f_i are convex quadratic for all $i \in I$

$$f_i(x) = x^{\mathrm{T}} Q_i x + r_i^{\mathrm{T}} x + s_i \text{ with } Q_i \succeq 0$$

(equalities f_i, if present, must still be affine for $i \in E$)

We call these problems structured convex problems

Other well-known classes of convex problems

Conic optimization (over symmetric cones)

3. Second-order cone optimization (SOCO) involves constraints such as

$$\left\|(c_{i1} - a_{i1}^{\mathrm{T}}x, c_{i2} - a_{i2}^{\mathrm{T}}x, \ldots, c_{in} - a_{in}^{\mathrm{T}}x)\right\| \leq c_{i0} - a_{i0}^{\mathrm{T}}x$$

4. Semidefinite optimization (over symmetric real matrices, SDO)

$$C + \sum_i x_i A_i \succeq 0$$

- ◇ QCQO is a special case of SOCO (i.e. QCQO problems have an equivalent formulation as SOCO problems, although no proof of strict inclusion yet)
- ◇ Both LO and SOCO are special cases of (real) SDO, as well as complex/hermitian SDO

More classes of well-known convex problems

4. Geometric optimization (GO):
f_0 and f_i are posynomials (in exponential form) for all $i \in I$

$$f_i(x) = c_i + \sum_{j \in M_i} \exp(a_j^{\mathrm{T}} x - b_j)$$

Each term in the sum is the composition of exponential and affine scalar function

5. Optimization with powers: l_p-norm optimization (l_pO):
f_0 linear, f_i are affine plus sum of convex powers with affine scalar arguments for all $i \in I$

$$f_i(x) = c_i^{\mathrm{T}} x - d_i + \sum_{j \in M_i} |a_j^{\mathrm{T}} x - b_j|^{p_j} \text{ with } p_j \geq 1$$

Even more classes of well-known convex problems

6. Optimization with norms: sum-of-norm optimization (SNO):
f_0 (and f_i for all $i \in I$, if any) are convex norms with affine arguments

$$f_i(x) = c_i + \sum_{j \in M_i} \left\| A_j^{\mathrm{T}} x - b_j \right\|_{p_j} \text{ with } p_j \geq 1$$

with $\|y\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}$

7. Entropy optimization (EO):
f_0 is a sum of entropy terms, f_i are affine for all $i \in E$

$$f_0(x) = \sum_i x_i \log x_i \quad \text{(implicitly implying } x \geq 0\text{)}$$

8. Analytic centering (AC):
f_0 is a sum of logarithmic terms, f_i are affine for all $i \in I \cup E$

$$f_0(x) = - \sum_{j \in M_0} \log(a_j^{\mathrm{T}} x - b_j)$$

Solving convex problems in practice

- Although each class can be tackled by a general-purpose nonlinear solver, better performance is expected from dedicated solvers
- Such solvers exist for all problems classes described (and others) but typically only handle one (or a few) problem class at a time i.e. there are dedicated solvers for linear optimization, quadratic optimization, geometric optimization, etc.
- Typically some problem classes are (much) easier to solve than others
- Note however that *in theory*, all problems can be solved by unified class of interior-point algorithms, but no single unified efficient solver seems to exist in practice
  Currenly most versatile class of solvers: mixed linear-second-order-semidefinite, such as SeDuMi or SDPT3

Overview

Approximations

We focus on

approximating one class of problems $\mathcal{P}$
with another class of problems $\mathcal{A}$

- ◇ if possible with arbitrarily high accuracy
  (but then typically at the cost of a growing size for the approximating problem)

Examples

- ◇ Linear approximations of quadratic problems
  Previous work:
  very nice construction of Ben-Tal and Nemirovski (2001)
- ◇ Quadratic approximations of nonlinear, nonquadratic convex functions
  This work:
  based on the concept of arithmetic-geometric mean iteration

Why approximate problems $\mathcal{P}$ by problems $\mathcal{A}$?

Useful in a few situations, such as

- algorithms for $\mathcal{A}$ are significantly faster than algorithms for $\mathcal{P}$
  $\rightarrow$ hope to obtain an approximate solution to $\mathcal{P}$, possibly with very high accuracy, in less time than required to solve it exactly

- from a more practical point of view:
  need to solve problem $\mathcal{P}$ but only have access to solver for $\mathcal{A}$

- in particular, when dealing with the following type of problems
  1. problem to be solved is discrete, such as (mixed) integer programming
  2. its continuous relaxation belongs to class $\mathcal{P}$
  3. available branch-and-bound type solvers only work with subproblems of type $\mathcal{A}$

  e.g. quadratic integer optimization using a commercial and highly optimized linear (mixed) integer optimization solver

Direct vs. extended formulations

Assume we want linear (polyhedral) approximation of convex set $\mathcal{S} \subset \mathbb{R}^n$

1. Direct approximation
   Look for polytope in $\mathbb{R}^n$ approximating $\mathcal{S}$, i.e.

$$\mathcal{P} = \{y \in \mathbb{R}^n \mid A^{\mathrm{T}} y \leq c\} \text{ such that } \mathcal{P} \approx \mathcal{S}$$

2. Approximation based on extended formulation, i.e. a lifting
   Look for polytope in higher-dimensional space $\mathbb{R}^{n+p}$
   such that its projection on $\mathbb{R}^n$ approximates $\mathcal{S}$

$$\mathcal{E} = \{(y, u) \in \mathbb{R}^n \times \mathbb{R}^p \mid A^{\mathrm{T}} y + B^{\mathrm{T}} u \leq c\}$$

$$\text{such that } \mathcal{E}_y = \mathrm{Proj}_y \mathcal{E} = \{y \mid (y, u) \in \mathcal{E} \text{ for some } u\} \approx \mathcal{S}$$

Optimizing over projection $\mathcal{E}_y$ is not more difficult than on $\mathcal{E}$:

$$\min f(y) \text{ s.t. } y \in \mathcal{S} \approx \min f(y) \text{ s.t. } y \in \mathcal{E}_y \Leftrightarrow \min f(y) \text{ s.t. } (y, u) \in \mathcal{E}$$

Direct vs extended: example

Assume we want polyhedral approximation of disc $\{(x,y) \mid x^2+y^2 \leq 1\}$

1. Direct approximation
   Use m linear inequalities based on tangents (in $\mathbb{R}^2$)
   m-sided approximations $\rightarrow \frac{\pi^2}{2m^2}$ accuracy
   $\Rightarrow$ very expensive ($m > 2000$ for $\varepsilon = 10^{-6}$)

2. Approximation based on extended formulation
   Construction by Ben-Tal and Nemirovsky:
   Explicit description of polytope $\mathcal{E} \subset \mathbb{R}^{2+p}$ with $p+1$ inequalities
   with a projection on $\mathbb{R}^2$ with 2^p sides $\rightarrow \frac{\pi^2}{2^{p+1}}$ accuracy
   $\Rightarrow$ cheap ($p = 24$ for $\varepsilon = 10^{-6}$)

   Has been successfully applied to mechanical engineering problems (limit analysis) and the resolution of integer quadratic problems with linear integer programming solvers

Ben-Tal-Nemirovsky construction (details)

$$(u, v) \in \text{approximated disc in } \mathbb{R}^2 \Leftrightarrow \exists y \in \mathbb{R}^m \mid (u, v, y) \in \mathcal{P}$$

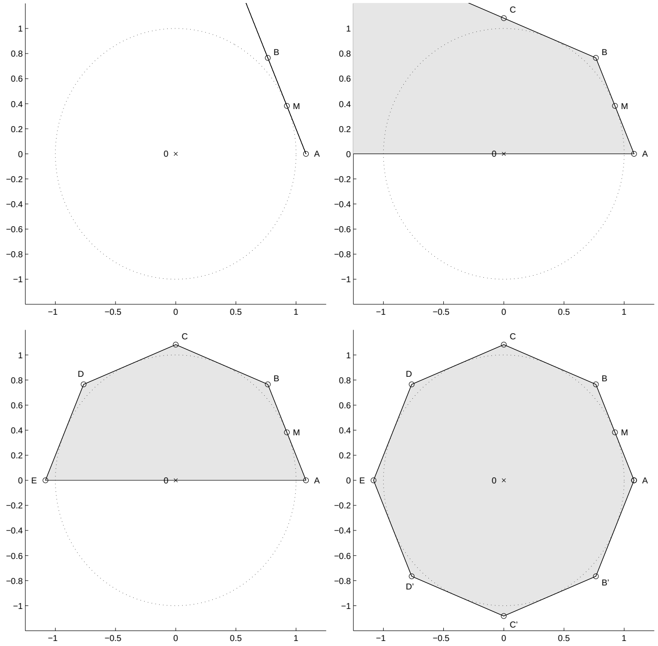
where $\mathcal{P}$ is a polytope in the higher-dimensional space $\mathbb{R}^{2+m}$

Let $q \geq 1$ a positive parameter and consider the following system

$$\left\{ \begin{array}{rcl} \alpha_{i+1} & = & \alpha_i \cos \frac{\pi}{2^i} + \beta_i \sin \frac{\pi}{2^i} \\ \beta_{i+1} & \geq & \left| -\alpha_i \sin \frac{\pi}{2^i} + \beta_i \cos \frac{\pi}{2^i} \right| \end{array} \right. , \ 0 \leq i < q$$

$$\left\{ \begin{array}{rcl} \beta_q & \leq & 2 \sin \frac{\pi}{2^q} \\ 1 & = & \alpha_q \cos \frac{\pi}{2^q} + \beta_q \sin \frac{\pi}{2^q} \end{array} \right.$$

Its projection on (α_0, β_0) is a regular 2^q-sided polygon, at the cost of $m = 2q + 1$ inequalities and $2q$ additional variables


A
B
M
C
D
E
B'
C'
D'
0
1
0.8
0.6
0.4
0.2
−0.2
−0.4
−0.6
−0.8
−1
−0.5
0.5

Our goal in this talk

Generalize this quadratic by linear approximation result with arbitrary accuracy:

Discover new types of extended formulations

to approximate convex nonlinear (transcendental) functions
and related optimization problems with arbitrary accuracy

using only convex quadratic (second-order cone) inequalities

Main tool: the arithmetic-geometric mean iteration

Overview

1. Motivation
   - ◇ Convex optimization: problem classes
   - ◇ Approximations: direct vs. extended formulations
2. **Quadratic approximations of convex optimization problems**
   - ◇ Arithmetic-geometric mean iteration
   - ◇ Application to approximations
3. Generalizations and conclusions
   - ◇ Matrix version
   - ◇ Conclusions
   - ◇ Addendum

Arithmetic-geometric mean iteration

Let $a_0 = \alpha$, $b_0 = \beta$ with $\alpha > \beta > 0$ and define the iteration

$$a_{n+1} = \frac{a_n + b_n}{2} \text{ and } b_{n+1} = \sqrt{a_n b_n}$$

Inequality between arithmetic and geometric means implies

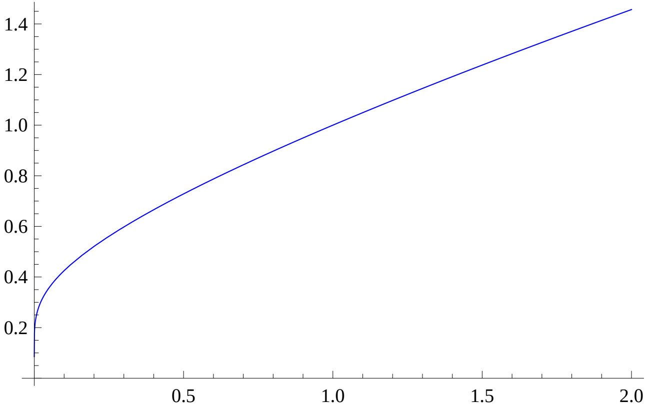
$$\alpha > a_1 > \cdots > a_n > a_{n+1} > \cdots > b_{n+1} > b_n > \cdots > b_1 > \beta$$

so that sequences $\{a_n\}$ and $\{b_n\}$ must admit a joint finite limit $AG(\alpha, \beta)$, called the arithmetic-geometric mean of α and β ; already known from GAUSS, but revisited recently by the BORWEIN brothers (late 80's)
Since $AG(\alpha, \beta)$ is (positively) homogeneous, i.e.

$$AG(\lambda\alpha, \lambda\beta) = \lambda AG(\alpha, \beta) \text{ for any } \lambda > 0 \,,$$

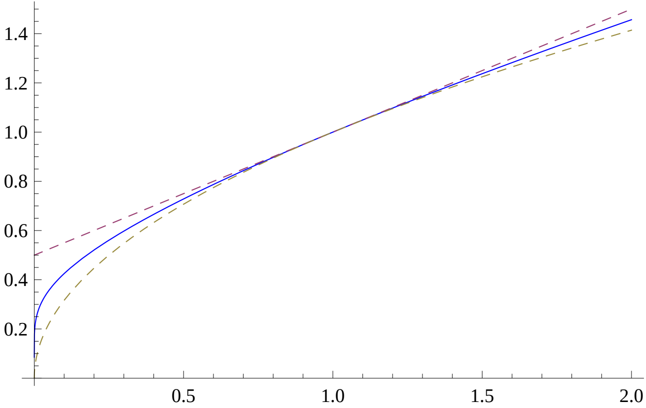
one only has to consider $AG(1, \beta)$, which is a concave function.

Arithmetic-geometric mean $\beta \mapsto AG(1, \beta)$



Arithmetic-geometric mean $\beta \mapsto AG(1, \beta)$

(between arithmetic and geometric means)



Properties of AGM iterates

Denote $AG_k(1,\beta) = a_k$, i.e. the k^{th} AGM iterate starting from $(1,\beta)$:

$$AG_k(\alpha,\beta) \to AG(1,\beta) \text{ where convergence is quadratic}$$

i.e. the error is squared after each iteration, which implies

$$0 \leq AG_k(\alpha,\beta) - AG(1,\beta) \leq C^{-2^k}$$

Function $AG(1,\beta)$ can be defined in terms of a complete elliptic integral of the first kind

$$AG(1,\beta) = \frac{\pi}{2I(1,\beta)} \text{ with } I(\alpha,\beta) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{\alpha^2\cos^2\theta + \beta^2\sin^2\theta}}$$

for which there exists no closed form in terms of elementary functions

Properties of elliptic integral $I(\alpha, \beta)$

◇

$$I(\alpha,\beta) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{\alpha^2\cos^2\theta + \beta^2\sin^2\theta}} = \int_0^{\infty} \frac{dx}{\sqrt{(x^2+a^2)(x^2+b^2)}}$$

◇ Some careful but straightforward computations show that

$$I(\frac{a+b}{2}, \sqrt{ab}) = I(a,b)$$

which implies (since $I(a_n, b_n)$ must be independent from n)

$$I(a_0,b_0) = I(a_n,b_n) = \lim_{n\to\infty} I(a_n,b_n) = I(L,L) = \frac{\pi}{2L}$$

and therefore

$$AG(1,\beta) = \frac{\pi}{2I(1,\beta)}$$

From an algorithm to an extended formulation

Computing the AGM only requires linear and quadratic operations

But our goal is not to compute the AGM for a given value of β, but for any value of β (β is potentially a variable in an optimization model)

The computation of $AG_k(1, \beta)$ must therefore be embedded into the optimization model, which leads to an extended formulation

How to convert an algorithm into an extended formulation ?

- ◇ operations become equalities
- ◇ intermediate results become additional variables
- ◇ every operation must preserve convexity, i.e.
  - equalities can only be linear
  - unless you can prove they can be relaxed into inequalities
  - and those inequalities can only be in the form "convex $\leq$ concave"

The end result is an extended formulation for the epigraph/hypograph of the (convex/concave) function you want to compute (approximately)

A quadratic extended formulation for the AGM

For each value of k, function $\beta \to AG_k(1,\beta)$ is concave and its hypograph

$$H_k = \{(\beta,t) \mid AG_k(1,\beta) \geq t\}\ ,$$

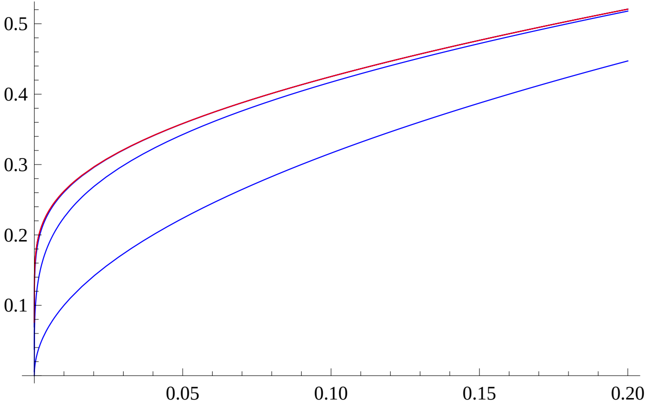
a convex set, admits the following quadratic extended formulation

$$\begin{aligned} a_0 &= 1, b_0 = \beta \\ a_n &= \frac{a_{n-1}+b_{n-1}}{2}, b_n \leq \sqrt{a_{n-1}b_{n-1}}\ \forall 1 \leq n \leq k \\ a_k &= t \end{aligned}$$

Each quadratic inequality $X \leq \sqrt{YZ}$ is convex and corresponds to a second-order cone: $X^2 \leq YZ \Leftrightarrow X^2 + (\frac{Y-Z}{2})^2 \leq (\frac{Y+Z}{2})^2$ (slightly more general than convex quadratic inequalities, but still convex)

Hypograph of $AG(1,\beta)$ can be approximated with arbitrary accuracy using only quadratic inequalities ; similarly for convex epigraph of $I(1,\beta)$ (using one additional quadratic inequality $\frac{\pi}{2a_k} \leq t$)

Arithmetic-geometric mean $\beta \mapsto AG(1,\beta)$ and approximations $\beta \mapsto AG_k(1,\beta)$ for $k = 1, 2, 3, 4$



Overview

1. Motivation
   - ◇ Convex optimization: problem classes
   - ◇ Approximations: direct vs. extended formulations
2. **Quadratic approximations of convex optimization problems**
   - ◇ Arithmetic-geometric mean iteration
   - ◇ Application to approximations
3. Generalizations and conclusions
   - ◇ Matrix version
   - ◇ Conclusions
   - ◇ Addendum

First example: computation of π

One can show that when $(\alpha, \beta) = (1, \frac{1}{\sqrt{2}})$, AGM iterates satisfy

$$\frac{2a_k^2}{1 - \sum_{i=0}^{k} 2^i c_i^2} \to \pi \text{ when } k \to \infty \qquad (\text{where } c_i = \frac{a_i - b_i}{2})$$

Therefore arbitrary accuracy approximations of π can be computed with a rational (quadratic) second-order cone optimization problem

k		approximation	correct digits
1	$\to$	3.1876726427	2
2	$\to$	3.1416802933	4
3	$\to$	3.1415926539	9
$\vdots$		$\vdots$	
∞	$\pi =$	3.14159265358979323 ...	

Three inequalities suffice for near double-precision floating-point accuracy

Second application: transcendental functions

- One can show that $I(1, x)$ gives a good approximation of $\log(\frac{4}{x})$ near the origin

$$\lim_{x \to 0^+} \left[\log\left(\frac{4}{x}\right) - I(1, x) \right] = 0$$

- Moreover, for each $m \geq 3$, one has for all $0 < x < 1$

$$\left| \log x - \left(I(1, 10^{-m}) - I(1, x10^{-m}) \right) \right| < m10^{-2(m-1)}$$

  i.e. $\log x$ can be approximated very accurately using function $I(1, \beta)$, which can itself be cheaply approximated by a few quadratic inequalities.

- cheap and accurate quadratic approximation of the logarithm

Quadratic extended formulation for the logarithm

Main result: the hypograph of the logarithm

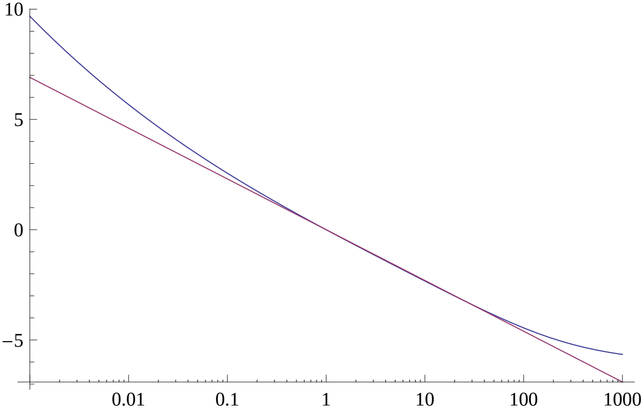
$$\{(x,t) \in \mathbb{R}_{++} \times \mathbb{R} \text{ s.t. } \log x \geq t\}$$

can be approximated with arbitrary accuracy using the following k-step quadratic extended formulation for sufficiently large values of m and k

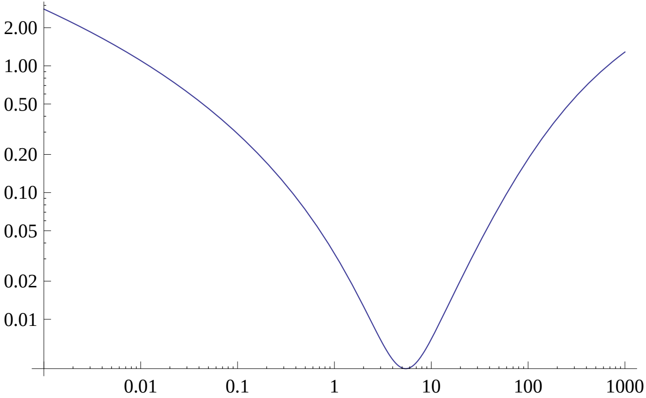
$$\left\{ \begin{array}{rcl} \multicolumn{3}{l}{(x, a_0, b_0, a_1, b_1, \ldots, a_k, b_k, u, t) \in \mathbb{R}_{++} \times \mathbb{R}^{2k+4} \text{ s.t.}} \\ a_0 & = & 1 \text{ and } b_0 = x10^{-m} \\ a_n & = & \dfrac{a_{n-1} + b_{n-1}}{2} \; \forall 1 \leq n \leq k \\ b_n & \leq & \sqrt{a_{n-1} b_{n-1}} \; \forall 1 \leq n \leq k \\ \pi & \leq & (a_k + b_k) u \\ I(1, 10^{-m}) - u & = & t \end{array} \right\}$$

where the $k+1$ inequalities can be formulated with second-order cones

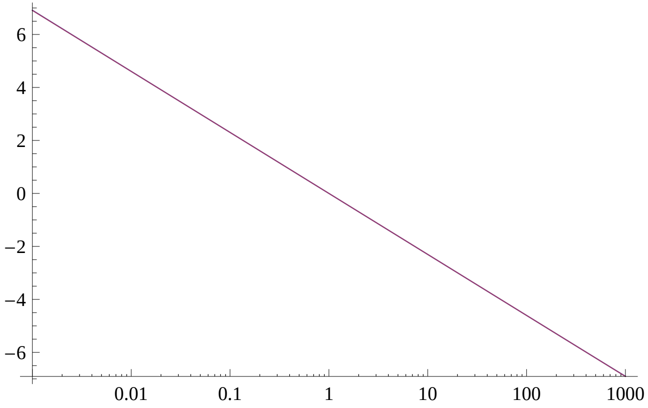
log vs. approximation for $k = 3$ iterations and $m = 2$ (loglinear plot)



Approximation error for $k = 3$ iterations and $m = 2$ (loglog plot)

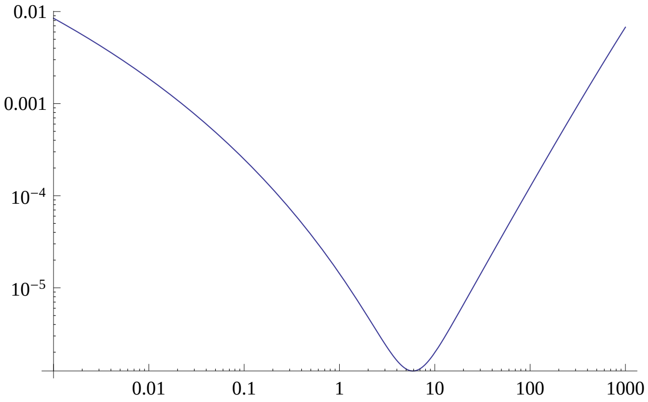


Approximation for $k = 5$ iterations and $m = 4$
(loglinear plot)

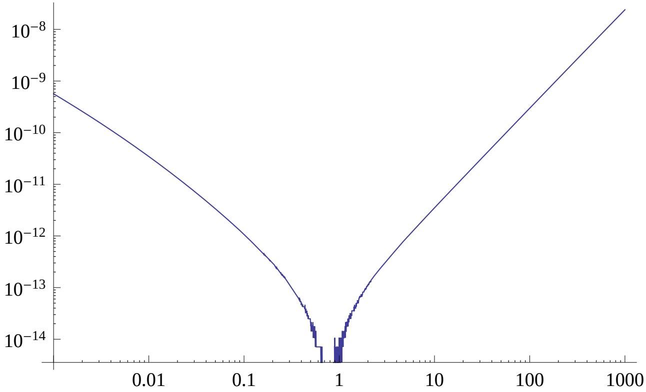


Approximation error for $k = 5$ iterations and $m = 4$

(loglog plot)



Approximation error for $k = 7$ iterations and $m = 7$ (loglog plot)



Logarithm brings many other transcendental functions

- Since constraint $\log x \geq t$ can be approximated, so can exponential $e^t \leq x$ (inverse has same graph), its conic hull $\log(\frac{x}{u}) \geq \frac{t}{u}$ and entropy $u \log u \leq -t$ (using conic hull with $x = 1$)
- Convex powers can also obtained: $x^p \leq t$ (with $p \geq 1$) is equivalent to $p \log x \leq \log t$, itself equivalent to the pair of constraints $x \log x \leq \frac{u}{p-1}$ and $u/x \leq \log(t/x)$ (another lifting)
- Hyperbolic and inverse functions, such as $\cosh x = \frac{1}{2}(e^x + e^{-x})$ and $\cosh^{-1} z = \log(z + \sqrt{z^2 - 1})$ can also be similarly approximated (e.g. $e^x \leq t_1$, $e^{-x} \leq t_2$ and $t_1 + t_2 = 2t$) ; another useful example is the Lambert W function (inverse of xe^x, with no closed-form)
- A whole class of convex optimization problems involving powers and exponentials can be approximated (including geometric optimization, ℓ_p norm-optimization, entropy optimization, analytic centering, etc.)

Overview

Possible generalizations

Arithmetic-geometric mean iteration can be generalized to different settings:

- ◇ A complex variant of the iteration can be used to approximate trigonometric functions (via exponential with imaginary argument): can this approach also be translated into convex quadratic inequalities ? I don't know

- ◇ A matrix variant of the iteration can be used to approximate the matrix logarithm/exponential function (on positive definite matrices): can this approach also be translated into convex quadratic inequalities ? linear matrix inequalities ? quadratic matrix inequalities ? I know how to do it with (lifted) LMIs

Third application: matrix logarithm

The arithmetic-geometric mean iteration is well defined for

(commuting) symmetric positive semidefinite matrices !

Let $A_0 = A$ and $B_0 = B$ symmetric positive definite commuting matrices ; define the iteration

$$A_{n+1} = \frac{A_n + B_n}{2} \text{ and } B_{n+1} = \sqrt{A_n B_n}$$

(note square-root is uniquely defined over symmetric positive definite matrices, and commutativity is preserved at each iteration)

The nonlinear B_{n+1} iteration can be relaxed as a semidefinite constraint

$$B_{n+1} \preceq \sqrt{A_n B_n} \quad \Leftrightarrow \quad \begin{pmatrix} A_n & B_{n+1} \\ B_{n+1} & B_n \end{pmatrix} \succeq 0$$

Integral representation and matrix logarithm

Sequences $\{A_n\}$ and $\{B_n\}$ admit a joint finite limit $AG^M(A, B)$, called the matrix arithmetic-geometric mean of A and B

As in the scalar case, one has

$$AG^M(A, B) = \frac{\pi}{2} I^M(A, B)^{-1}$$

with (using matrix square root and inverse)

$$I^M(A, B) = \int_0^\infty \frac{dx}{\sqrt{(x^2 \cdot \mathbb{I} + A^2)(x^2 2 \cdot \mathbb{I} + B^2)}}$$

and I^M can now be used to approximate the matrix logarithm

$$\log^M X \approx \left(\; I(1, 10^{-m}) \cdot \mathbb{I} \; - \; I^M(\mathbb{I}, 10^{-m} \cdot X) \; \right)$$

(where $\mathbb{I}$ is the identity matrix)

Example

Simply computing the matrix logarithm of a given matrix M can be done using

$$\max \operatorname{tr} T \text{ such that } \log^M X \geq T$$

where the nonlinear constraint can be approximated by

$$\begin{aligned}
A_0 &= \mathbb{I} \text{ and } B_0 = 10^{-m} \cdot X \\
A_n &= \frac{A_{n-1} + B_{n-1}}{2} \quad \forall 1 \leq n \leq k \\
\begin{pmatrix} A_n & B_{n+1} \\ B_{n+1} & B_n \end{pmatrix} &\succeq 0 \quad \forall 1 \leq n \leq k \\
\begin{pmatrix} A_k + B_k & \mathbb{I} \\ \mathbb{I} & \frac{1}{\pi} U \end{pmatrix} &\succeq 0 \\
I(1, 10^{-m}) \cdot \mathbb{I} - U &= T
\end{aligned}$$

A (recorded) MATLAB demo

Output of a 10-line MATLAB script using the YALMIP toolbox (Löfberg) and SeDuMi solver (AdvOL-McMaster)

```
Approximation of 5 x 5 matrix logarithm with 6-step matrix AGM (param m=5):
M =
    5.4548   -0.3147   -0.4327    1.9221   -0.0614
   -0.3147    0.7643    1.0283   -1.1224    0.0496
   -0.4327    1.0283    3.5340    0.0114   -3.0764
    1.9221   -1.1224    0.0114    7.0892   -1.9941
   -0.0614    0.0496   -3.0764   -1.9941    5.4030
trueLog =
    1.6362    0.0013   -0.1025    0.3255    0.0139
    0.0013   -1.6375    1.3469   -0.4630    0.5525
   -0.1025    1.3469    0.3247   -0.0292   -1.1642
    0.3255   -0.4630   -0.0292    1.8033   -0.3606
    0.0139    0.5525   -1.1642   -0.3606    1.1707
aproxLog =
    1.6361    0.0013   -0.1024    0.3255    0.0139
    0.0013   -1.6372    1.3467   -0.4630    0.5524
   -0.1024    1.3467    0.3247   -0.0292   -1.1642
    0.3255   -0.4630   -0.0292    1.8032   -0.3606
    0.0139    0.5524   -1.1642   -0.3606    1.1707
err =
  3.7750e-004
```

```
>> M=randn(5,5)*2;log_demo(M'*M,6);
SeDuMi 1.21 by AdvOL, 2005-2008 and Jos F. Sturm, 1998-2003.
Alg = 2: xz-corrector, theta = 0.250, beta = 0.500
Put 150 free variables in a quadratic cone
eqs m = 180, order n = 63, dim = 752, blocks = 8
nnz(A) = 700 + 0, nnz(ADA) = 28350, nnz(L) = 14265
 it :     b*y       gap    delta  rate   t/tP*  t/tD*   feas cg cg  prec
  0 :            1.41E+000 0.000
  1 :  4.93E+000 7.19E-001 0.000 0.5111 0.9000 0.9000   0.23  1  1  3.3E+000
  2 :  4.80E+000 2.06E-001 0.000 0.2860 0.9000 0.9000   1.50  1  1  7.4E-001
  3 :  3.51E+000 5.97E-002 0.000 0.2904 0.9000 0.9000   1.37  1  1  2.0E-001
  4 :  2.80E+000 1.87E-002 0.000 0.3136 0.9000 0.9000   0.99  1  1  7.5E-002
  5 :  2.42E+000 6.99E-003 0.000 0.3732 0.9000 0.9000   0.79  1  1  3.7E-002
  6 :  2.19E+000 2.72E-003 0.000 0.3894 0.9000 0.9000   0.70  1  1  2.0E-002
  7 :  2.05E+000 1.40E-003 0.000 0.5160 0.9000 0.9000   0.55  1  1  1.4E-002
...
 19 :  1.55E+000 1.94E-007 0.000 0.5190 0.9000 0.9000   0.82  1  1  8.4E-006
 20 :  1.54E+000 4.70E-008 0.000 0.2426 0.9000 0.9000   0.91  1  1  2.1E-006
 21 :  1.54E+000 1.06E-008 0.000 0.2246 0.9000 0.9000   0.96  1  1  4.9E-007
 22 :  1.54E+000 5.54E-010 0.000 0.0524 0.9900 0.9900   0.99  1  1  2.6E-008
 23 :  1.54E+000 1.52E-011 0.000 0.0274 0.9900 0.9900   1.00  1  1  7.1E-010

iter seconds digits       c*x               b*y
 23      0.6   Inf  1.5435050372e+000  1.5435050430e+000
|Ax-b| =  2.1e-009, [Ay-c]_+ =  2.8E-010, |x|= 9.7e+002, |y|= 1.6e+000
```

```
function log_demo(M, k)
n=size(M,2);



A0=M/10^5;B0=eye(n);
for i=1:k
    A{i}=sdpvar(n,n); B{i}=sdpvar(n,n);
end
cons = set(A{1}==(A0+B0)/2) + set([A0 B{1};B{1} B0]>0);
for i=2:k
    cons = cons + set(A{i}==(A{i-1}+B{i-1})/2) + set([A{i-1} B{i};B{i} B{i-1}]>0);
end
L = sdpvar(n);
cons = cons + [(A{k}+B{k})/pi eye(n) ; eye(n) 12.8992*eye(n)-L]
solvesdp(cons, -trace(L));


disp(sprintf('\nApproximation of %d x %d matrix logarithm with %d-step matrix AGM:'
M
trueLog=logm(M)
aproxLog=double(L)
err=norm(trueLog-aproxLog)
```

Overview

Concluding remarks

Further research needed:

- Compare with other proposed approximation techniques for elementary functions (e.g. CORDIC, Padé, Brent, etc.)
- Perform computational experiments with off-the-shelf quadratic solvers
- Test applicability to geometric optimization and (mixed) integer geometric optimization
- Is it possible to guarantee accuracy of the solution (e.g. error on the objective function) ?
- Implication on the difficulty to compute exact solutions to systems of convex quadratic inequalities ?
- Adapt Ben-Tal-Nemirovsky construction to obtain other efficient linear approximations

Overview

1. Motivation
   - ◇ Convex optimization: problem classes
   - ◇ Approximations: direct vs. extended formulations
2. Quadratic approximations of convex optimization problems
   - ◇ Arithmetic-geometric mean iteration
   - ◇ Application to approximations
3. Generalizations and conclusions
   - ◇ Matrix version
   - ◇ Conclusions
   - ◇ Addendum

A small addendum

Huge gap in class of convex sets efficiently representable or approximable by quadratic inequalities:

LMI-representable sets

(i.e. solve (approximately) semidefinite optimization with second-order cone optimization) What we can do exactly with quadratic inequalities ?

- ◇ 2 by 2 matrices: $\begin{pmatrix} a & b \\ b & c \end{pmatrix} \succeq 0 \Leftrightarrow a + c \geq \|(a - c, 2b)\|$
- ◇ arrow matrices: $\begin{pmatrix} x_0 & x_1 & \cdots & x_n \\ x_1 & x_0 & & \\ \vdots & & \ddots & \\ x_n & & & x_0 \end{pmatrix} \succeq 0 \Leftrightarrow x_0 \geq \|(x_1, \cdots, x_n)\|$
- ◇ and ... ? What about the Cayley cubic: $\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succeq 0$?

Cayley cubic and generalizations

After trying for a while without success, I started to think it was impossible ... until Y. Nesterov showed me how to do it:

$$\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succeq 0 \Leftrightarrow a^2 + b^2 + c^2 - 2abc \leq 1 \text{ and } -1 \leq a, b, c \leq 1$$

$$\Leftrightarrow \frac{(a+b)^2}{1+c} + \frac{(a-b)^2}{1-c} \leq 2 \text{ and } -1 \leq c \leq 1$$

$$\Leftrightarrow (a+b)^2 \leq (1+c)u, (a-b)^2 \leq (1-c)v, -1 \leq c \leq 1 \text{ and } u + v = 2$$

$$\Leftrightarrow \begin{pmatrix} 1+c & a+b \\ a+b & u \end{pmatrix} \succeq 0, \begin{pmatrix} 1-c & a-b \\ a-b & u \end{pmatrix} \succeq 0 \text{ and } u + v = 2$$

(this is an extended formulation - the two others were direct)

Can the Cayley cubic representation be generalized ?

- Non constant diagonal elements:

  I can handle $\begin{pmatrix} \alpha & a & b \\ a & \beta & c \\ b & c & \beta \end{pmatrix} \succeq 0$ but not $\begin{pmatrix} \alpha & a & b \\ a & \beta & c \\ b & c & \gamma \end{pmatrix} \succeq 0$

  (but maybe somebody can show me how to do the latter ...)

- Higher dimensions:

  Question: Why is there a trick for Cayley and (apparently) not for a general 3×3 symmetric matrix ?

  My (current) answer: Cayley and the first example above works because they admit a 2×2 minor with constant eigenvectors

  $\begin{pmatrix} \beta & c \\ c & \beta \end{pmatrix}$ always admits $(1, 1)$ and $(-1, -1)$ as eigenvectors,

  and hence is diagonalizable in a constant basis

  $\rightarrow$ can be generalized to handle larger matrices with that constant-basis structure augmented by one row/column

Thank you for your attention!