A geodesic interior-point method for linear optimization over symmetric cones

Frank Permenter

Toyota Research Institute

Symmetric cones have underlying algebraic structure

A set $\mathcal{K}$ is a *symmetric cone* if $\mathcal{K} = \{x^2 : x \in \mathcal{J}\}$ for a commutative algebra $\mathcal{J}$ over real inner-product space satisfying

$$\langle x \circ y, z \rangle = \langle y, x \circ z \rangle, \qquad (x \circ y) \circ x^2 = x \circ (y \circ x^2)$$

Examples:

- Nonnegative orthant $\{x \in \mathbb{R}^n : x \geq 0\}$

$$[x \circ y]_i := x_i y_i$$

- Second-order-cone $\{(x_0, x) \in \mathbb{R} \times \mathbb{R}^n : \|x\| \leq x_0\}$

$$(x_0, x) \circ (y_0, y) := (x_0 y_0 + x^T y, x_0 y + y_0 x).$$

- Cone of psd matrices $\{VV^T : V \in \mathbb{R}^{n \times n}\}$

$$X \circ Y := \frac{1}{2}(XY + YX)$$

Symmetric cone programs generalize LP/SOCP/SDP

Given symmetric cone $\mathcal{K} = \{z^2 : z \in \mathcal{J}\}$, we consider problem

$$\begin{array}{ll} \text{minimize} & \langle c, x \rangle \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K}. \end{array}$$

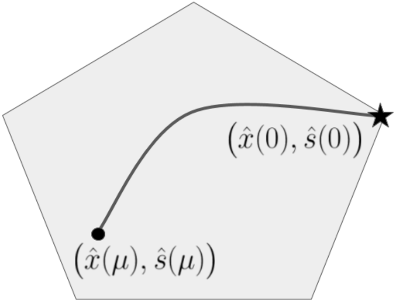
This generalizes linear (LP), second-order-cone (SOCP), and semidefinite programming (SDP).

A well-studied family:

- Algorithms: Faybusovich, Alizadeh/Schmieta, Nesterov/Todd
- Polynomial-time complexity bounds
- Software packages: `SeDuMi`, `SDPT3`, `Mosek`, ...

Symmetric cone programs solved by interior-point methods

IPMs track the *central-path* of $\min_{x \in \mathcal{K}, Ax=b} c^T x$.



$$\begin{aligned} x \circ s &= \mu \mathbf{1}, \\ Ax = b, \;\; s &= c - A^* y \qquad (1) \\ x \in \mathcal{K} \quad & s \in \mathcal{K}. \end{aligned}$$

($\mathbf{1}$ denotes the identity of $\circ$.)

That is, they reduce μ to zero while computing solutions to (1).

Properties of IPMs:

- Move along central path in $\mathcal{O}(\|\mathbf{1}\| \log \frac{\mu_0}{\mu_f})$ iterations
- s and x updated using subspaces:

$$x_{i+1} - x_i \in \operatorname{null} A, \qquad s_{i+1} - s_i \in \operatorname{range} A^*$$

We present a new IPM for symmetric cone optimization.

Key idea: update (s_i, x_i) using *geodesics* of $\mathcal{K}$ instead of *subspaces* of A such that *complementarity* is maintained.

$$\underbrace{Ax_i = b, s_i = c - A^* y_i,}_{\text{existing algs}} \qquad \underbrace{x_i \circ s_i = \mu_i \mathbf{1}}_{\text{this talk}} \qquad \forall \text{ iters. } i$$

Remainder of talk:

Part I: The special-case of *linear* programming

- Log-space transformation of central-path
- A *log-space* IPM and $\mathcal{O}(\sqrt{n})$ complexity.

Part II: The generalization to *symmetric cones*

- From log-space to geodesics
- A *geodesic* IPM and $\mathcal{O}(\|\mathbf{1}\|)$ complexity.

Part I: A log-space interior-point method for linear programming.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0, \quad i.e., x \in \mathbb{R}^n_+ \end{array}$$

We solve log-domain central-path conditions

We rewrite central-path conditions

$$Ax = b \qquad s = c - A^T y, x \geq 0, s \geq 0, \;\; s_i x_i = \mu$$

using a log param. $v \in \mathbb{R}^m$ and elementwise exp. e^v:

$$b = A\sqrt{\mu}e^v, \;\; \sqrt{\mu}e^{-v} = c - A^T y \tag{2}$$

By construction: $x = \sqrt{\mu}e^v$ and $s = \sqrt{\mu}e^{-v}$ satisfy $x_i s_i = \mu$.

Our meta-algorithm:

- Fix μ and apply Newton's method to (2)
- Decrease μ.
- Repeat.

Previously unanalyzed!

Newton's method uses approx. $e^{v+d} \approx e^v + e^v \circ d$

Newton's method ($\circ :=$ elementwise mult.):

- Solve Newton system for $(y, d) \in \mathbb{R}^m \times \mathbb{R}^n$:

$$\begin{aligned} \sqrt{\mu} A \left(e^v + e^v \circ d \right) &= b \\ \sqrt{\mu} (e^{-v} - e^{-v} \circ d) &= c - A^T y \end{aligned} \tag{3}$$

- Pick step-size α, set $v \leftarrow v + \frac{1}{\alpha} d$ and repeat.

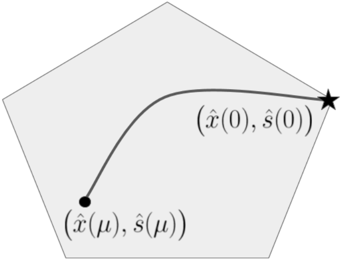
Properties (P., 2020):

- Globally converges if $\alpha = \max(1, \frac{1}{2} \|d\|^2)$.
- Quadratically converges to limit v_* if $\|v - v_*\| \leq \cosh^{-1}(5/4)$.

A log-space IPM for $\min_{x \geq 0, Ax = b} c^T x$

Let $d(\mu)$ denote Newton dir. as function of μ at current $v \in \mathbb{R}^n$.

while $\mu > \mu_f$ **or** $\|d(\mu)\| > \epsilon$ **do**
  Decrease μ
  $\alpha \leftarrow \max(1, \frac{1}{2}\|d(\mu)\|^2)$
  $v \leftarrow v + \frac{1}{\alpha} d(\mu)$
end
$x = \sqrt{\mu} e^v$, $s = \sqrt{\mu} e^{-v}$



Main results (P., 2020):

- Finitely terminates by simply setting $\mu = \mu_f$
- Exists μ-update rule with $\mathcal{O}(\sqrt{n} \log \frac{\mu_0}{\mu_f})$ iteration complexity
- Final log-distance of (x, s) to central-path is $\mathcal{O}(\epsilon)$

Part II: a geodesic-interior point method for symmetric cone optimization.

$$\begin{array}{ll} \text{minimize} & \langle c, x \rangle \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \end{array}$$

Line-segments in log-space are geodesics of $\mathbb{R}^n_+$

For curve $c : [0,1] \to \operatorname{int} \mathbb{R}^n_+$, let

$$L(c) := \int_0^1 \|c(t)^{-1} \circ c'(t)\| dt$$

Let $g(t) := e^{t \log a + (1-t) \log b}$ for $a, b \in \operatorname{int} \mathbb{R}^n_+$.

Properties

- The curve $g(t)$ is a *geodesic*, i.e., it minimizes $L(c)$ over $c(t)$ satisfying $c(0) = a$ and $c(1) = b$.
- $L(g) = \|\log a - \log b\|$.
- $g^{-1}(t)$ is the geodesic between a^{-1} and b^{-1}.

Geodesics of symm. cones have a known parametrization

For curve $c : [0,1] \to \operatorname{int} \mathcal{K}$, define

$$L(c) := \int_0^1 \|Q(c(t))^{-1/2} c'(t)\| dt,$$

where $Q(w) : \mathcal{K} \to \mathcal{K}$ denotes the *quadratic representation* of w.

Properties:

- Geodesics have form $g(t) := Q(w^{1/2}) \exp td$,

$$\exp d := \sum_{m=0}^{\infty} \frac{1}{m!} d^m, \qquad g(0) = w, \qquad L(g) = \|d\|$$

- $g^{-1}(t) = Q(w^{-1/2}) \exp -td$ also a geodesic.

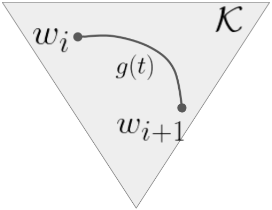
Example ($\mathcal{K} = \mathbb{R}^n_+$, $\mathcal{K} =$ psd matrices)

$$g(t) = w \circ e^{td} = e^{\log w + td}, \qquad g(t) = W^{1/2} e^{tD} W^{1/2}$$

A template geodesic IPM for $\min_{x \in \mathcal{K}, Ax=b} \langle c, x \rangle$

while $\mu > \mu_f$ **do**
  Decrease μ
  Compute search direction d
  Select step-size t.
  $w \leftarrow Q(w^{1/2}) \exp td$
end

$x = \sqrt{\mu} w$, $s = \sqrt{\mu} w^{-1}$



Iterates joined by geodesic curve
$g(t) = Q(w^{1/2}) \exp td$.

Properties of w-update:

- Equivalent to $w^{-1} \leftarrow Q(w^{-1/2}) \exp -td$.
- Formulae for LP and SDP:

$$w \leftarrow e^{\log w + td}, \qquad W \leftarrow W^{1/2} e^{tD} W^{1/2}$$

Linearizing w-update yields a geodesic Newton method

Geodesic Newton method:

- Solve Newton system for $(y, d) \in \mathbb{R}^m \times \mathbb{R}^n$:

$$\sqrt{\mu} A Q(w^{1/2})(\mathbf{1} + d) = b,$$
$$\sqrt{\mu} Q(w^{-1/2})(\mathbf{1} - d) = c - A^T y$$

- Set $w \leftarrow Q(w^{1/2}) \exp \frac{1}{\alpha} d$ using step-size α and repeat.

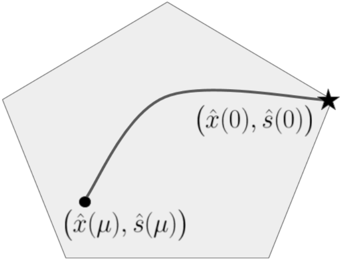
Properties (P., 2020):

- Based on approx. $Q(w^{1/2}) \exp d \approx Q(w^{1/2})(\mathbf{1} + d)$
- Globally converges to limit w_* if $\alpha = \max(1, \frac{1}{2}\|d\|^2)$.
- Quad. converges if geodesic distance $\delta(w, w_*) \leq \cosh^{-1}(5/4)$.

A geodesic IPM for $\min_{x \in \mathcal{K}, Ax=b} \langle c, x \rangle$

Let $d(\mu)$ denote Newton dir. as function of μ at current $w \in \mathcal{K}$.

while $\mu > \mu_f$ **or** $\|d(\mu)\| > \epsilon$ **do**
  Decrease μ
  $\alpha \leftarrow \max(1, \frac{1}{2}\|d(\mu)\|^2)$
  $w \leftarrow Q(w^{1/2}) \exp \frac{1}{\alpha} d(\mu)$
end
$x = \sqrt{\mu} w$, $s = \sqrt{\mu} w^{-1}$



Main results (P. 2020):

- Finitely terminates by simply setting $\mu = \mu_f$.
- Exists μ-update with $\mathcal{O}(\|\mathbf{1}\| \log \frac{\mu_0}{\mu_f})$ iteration complexity
- Final geodesic distance of (x, s) to central-path is $\mathcal{O}(\epsilon)$

Geodesic IPM implemented in software package conex

Currently developing `conex`, a software package for:

$$\begin{array}{ll} \text{minimize} & \langle c, x \rangle \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \end{array}$$

Features:

- Supports all symmetric cones $\mathcal{K}$
  - LP, SDP, SOCP
  - Hermitian psd matrices with complex and quaternion entries
  - The exceptional one (3x3 octonions)
- Sparse (supernodal) linear algebra.
- Approximation methods for matrix exponential.
- Lanczos methods for generalized eigenvalues.

Comparison with SDPT3 solver

Parameters	Solver Time (sec)		$\|\|Ax - b\|\|$		Duality Gap	
(n, m)	spdt3	conex	spdt3	conex	sdpt3	conex
(20, 20)	1.1e-01	4.1e-03	1.4e-12	3.9e-12	1.4e-09	8.9e-10
(50, 50)	7.0e-01	1.1e-01	1.0e-12	1.5e-12	1.1e-09	1.9e-09
(100, 100)	3.1e+00	9.8e-01	2.0e-12	3.9e-12	9.7e-10	2.4e-09
(20, 40)	1.4e-01	1.6e-02	6.9e-11	7.7e-13	4.6e-10	7.2e-10
(50, 250)	1.8e+00	5.6e-01	1.5e-11	9.8e-12	5.3e-09	6.6e-10
(100, 1000)	1.9e+01	1.4e+01	3.4e-11	3.1e-11	6.5e-10	6.9e-10

Table: SDPs of order n with m equality constraints.

Remarks:

- Our solver `conex` faster and just as accurate.
- Speed-up diminishes with $m > n$ since computation of Newton step dominates both solvers.

Thanks very much!

In summary,

- Presented new IPM for symmetric cone programming
- For LP, reduces to central-path tracking in log domain
- $\mathcal{O}(\|\mathbf{1}\|)$ complexity bounds match state-of-the-art
- Software package `conex` in development (demo on Thursday).

Paper and software:

`www.mit.edu/~fperment/`
`www.github.com/FrankPermenter/`