Towards a simplex-like method for second order cone programming

Kartik Krishnan

Department of Computing and Software
McMaster University
Joint work with Gábor Pataki (UNC), Neha Gupta (IIT Delhi), and Tamás Terlaky (Mac)

ICCOPT, RPI
Aug 4, 2004

Overview

- **Conic optimization**
  - Linear optimization
  - Second Order Cone optimization
- **Contrast simplex-like approaches and IPMs for conic optimization**
- **Notions of extreme points and nondegeneracy**
- **Feasible direction method for conic optimization**
- **Special cases: LP and SOCP**
- **Properties of the algorithm**
- **Preliminary computational results**
- **Conclusions and future work**

Conic optimization

- **Primal**

$$\begin{array}{rcl} \min\ c^T x & & \\ Ax & = & b, \qquad (CP) \\ x & \in & \mathcal{K}. \end{array}$$

- **Dual**

$$\begin{array}{rcl} \max\ b^T y & & \\ A^T y + s & = & c, \qquad (CD) \\ s & \in & \mathcal{K}^*. \end{array}$$

- **Facts:**

  (1) $A \in \mathbb{R}^{m \times n}$ with full row rank.

  (2) $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_r$.

  (3) Each of the cones $\mathcal{K}_i$ is convex, closed, and pointed.

Special cones

- $\mathbb{R}^n_\oplus = \{x \in \mathbb{R}^n : x \geq 0\}$.
  - Linear, self-dual
- $\mathcal{Q}^n = \{x \in \mathbb{R}^n : x_1 \geq \|x_{2:n}\|\}$.
  - Second order, self-dual
- $\mathbb{S}^n_\oplus = \{X \in \mathbb{S}^n : X \succeq 0\}$.
  - Semidefinite, self-dual

Optimality conditions

- (x^*, y^*, s^*) are optimal iff

$$\begin{aligned} Ax^* = b, \quad x^* \in \mathcal{K}, & \qquad (PF) \\ A^T y^* + s^* = c, \quad s^* \in \mathcal{K}^*, & \qquad (DF) \\ x_i^* \circ s_i^* = 0,\ i = 1, \ldots, r. & \qquad (CS) \end{aligned}$$

- For an SOCP cone in $\mathbb{R}^n$

$$x \circ s \;=\; \begin{pmatrix} x^T s \\ x(1)s(2:n) + s(1)x(2:n) \end{pmatrix}.$$

- For any cone $\mathcal{K}$

$$x \in \mathcal{K},\ s \in \mathcal{K}^*,\ x^T s = 0 \;\Rightarrow\; x \circ s = 0.$$

Eigenvalues and Eigenvectors

- Given $x \in \mathbb{R}^n$, we define

$$\text{Arw}(x) = \begin{pmatrix} x(1) & x(2:n)^T \\ x(2:n) & x(1)I_n \end{pmatrix}$$

- We have

$$x \in \mathcal{Q}^n \Leftrightarrow \text{Arw}(x) \succeq 0$$

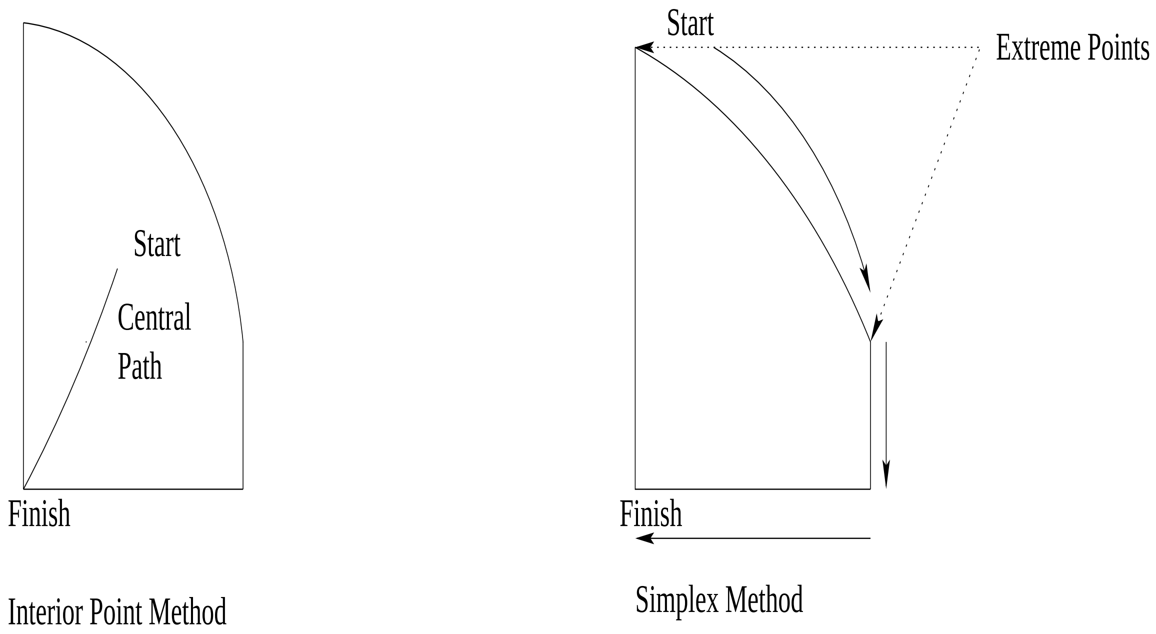
- We define

$$\begin{aligned} \lambda_{min}(x) &= \lambda_{min}(\text{Arw}(x)) \\ &= (x(1) - ||x(2:n)||) \end{aligned}$$

and $v_{min}(x)$ to be the associated eigenvector, where

$$v_{min}(x) = \frac{1}{\sqrt{2}}\left(1; -\frac{x(2{:}n)}{||x(2{:}n)||}\right)$$

Contrast simplex and IPMs for CP I.



Interior Point Methods deal with matrices of full rank.

In the simplex method, the rank of the extreme points satisfy
(1) r <= m (Linear Programming)

(2) r(r+1)/2 <= m (Semidefinite Programming)

Contrast simplex and IPMs for CP II.

Why a simplex method for conic programming?.

1. Listed as an open problem in Alizadeh & Goldfarb (03) and Vandenberghe & Boyd (96).
2. Reoptimization after branching or the addition of cutting planes using the dual simplex method.
3. It is possible in practice in case of SDP to do each iteration more quickly and with less memory.

Subspaces $\mathbb{B}_X$ and $\mathbb{T}_X$ I.

Given a closed convex cone $K \subset \mathbb{R}^n$. Consider $\bar{x} \in K$

- **The subspace $\mathbb{B}_X$:**

$$\mathbb{B}_X \;=\; \{d \in \mathbb{R}^n : \bar{x} \pm \epsilon d \in K, \epsilon > 0\}.$$

- **Tangent space:**

$$\mathbb{T}_X \;=\; \left\{d \in \mathbb{R}^n : \mathrm{dist}(\bar{x} \pm \epsilon d, K) = O(\epsilon^2), \epsilon > 0\right\}.$$

- **Null space of constraint set:**

$$\mathbb{N} \;=\; \{x \in \mathbb{R}^n : Ax = 0\}.$$

Subspaces $\mathbb{B}_X$ and $\mathbb{T}_X$ II.

- Given $\bar{x} = \begin{pmatrix} 1 \\ \bar{x}(2:n) \end{pmatrix} \in \mathrm{bd}(\mathcal{Q}^n)$ with $||\bar{x}(2:n)|| = 1$.
- $\mathbb{B}_X = \mathrm{lin}\,(\bar{x})$.
- $\mathbb{T}_X = \mathrm{lin}\left(\mathbb{B}_X \cup \left\{ \begin{pmatrix} 0 \\ w \end{pmatrix} : ||w|| = 1,\ w^T\bar{x}(2:n) = 0 \right\}\right)$.
- Note $\bar{x} + \epsilon \begin{pmatrix} 0 \\ w \end{pmatrix} \notin \mathcal{Q}^n$.
- However $\bar{x} + \epsilon \begin{pmatrix} 0 \\ w \end{pmatrix} + \epsilon^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathrm{int}(\mathcal{Q}^n)$.

Subspaces $\mathbb{B}_X$ and $\mathbb{T}_X$ III.

- If $\bar{x} = 0$ then $\mathbb{T}_X = \mathbb{B}_X = \{0\}$.
- If $\bar{x} \in \text{int}(\mathcal{Q}^n)$ then $\mathbb{T}_X = \mathbb{B}_X = \mathbb{R}^n$.
- If $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_r$ then

$$\begin{aligned} \mathbb{T}_X &= \mathbb{T}_{X_1} \times \ldots \times \mathbb{T}_{X_r}. \\ \mathbb{B}_X &= \mathbb{B}_{X_1} \times \ldots \times \mathbb{B}_{X_r}. \end{aligned}$$

Subspaces $\mathbb{B}_X$ and $\mathbb{T}_X$ IV.

- Given $\bar{x} = (\bar{x}_R, \bar{x}_I, \bar{x}_0)$ with $|R| = p$.
- The subspace $\mathbb{B}_X$ of dimension $(n_I + p)$ is

$$\mathbb{B}_X = ((\alpha_1 \bar{x}_1) \times \ldots (\alpha_p \bar{x}_p) \times \mathbb{R}^{n_I} \times 0)$$

- The tangent space $\mathbb{T}_X$ of dimension $(n_R + n_I - p)$ is

$$\mathbb{T}_X = \left(\Pi_{i=1}^{p} \mathrm{lin}(\mathbb{B}_{X_i} \cup \{ \begin{pmatrix} 0 \\ w_i \end{pmatrix} : \bar{x}_i(2 : n_i)^T w_i = 0, \forall i\}) \times \mathbb{R}^{n_I} \times 0 \right)$$

Nonredundant and spanning subspaces

Given a subspace $L \subseteq R^n$ we say

- **L is spanning:**

$$\mathbb{N} + L \;=\; \mathbb{R}^n.$$

- **L is nonredundant:**

$$\mathbb{N} \cap L \;=\; \{0\}.$$

Notions of nondegeneracy

A feasible $\bar{x}$ is

- **c-nondegenerate**

$$\mathbb{T}_X + \mathbb{N} \;=\; \mathbb{R}^n.$$

- **f-nondegenerate**

$$\mathbb{B}_X + \mathbb{N} \;=\; \mathbb{R}^n.$$

- **Extreme point**

$$\mathbb{B}_X \cap \mathbb{N} \;=\; \{0\}.$$

Feasible direction method for conic optimization I.

- 1. **Select basis:** Given a feasible $\bar{x}$ in (CP).
  - Select a maximal non-redundant subspace $L_1 \subseteq \mathbb{B}_X$ such that $\bar{x} \in L_1$.
  - If L_1 is spanning then set
    $L = L_1$ and *perturb* $= 0$ and go to step 2.
    Else select a subspace $L_2 \subseteq \mathbb{T}_X \backslash \mathbb{B}_X$ s.t. $L_1 + L_2$ is non-redundant and spanning.
    Set $L = L_1 + L_2$ and *perturb* $= 1$.

Feasible direction method for conic optimization II.

- 2. **Construct complementary dual solution:**
  Solve for (y, s) the system

$$\begin{aligned} A^T y + s &= c, \\ s &\in L^{\perp}. \end{aligned}$$

- 3. **Pricing:**
  If $s \in \mathcal{K}^*$ then $\bar{x}$ and s are optimal. STOP.
  Otherwise, let $s_k = \min_{i=1,\dots,r} \lambda_{min}(s_i)$. Solve for $g \in \mathcal{K}$ where

$$g = \min_{\{h \in \mathcal{K} : ||h|| = 1\}} s_k^T h$$

Feasible direction method for conic optimization III.

- 4. **Find improving direction:**
  - If *perturb* $= 0$, solve for f the system

    $$\begin{array}{rcl} A(f+g) & = & 0, \\ f & \in & L, \end{array}$$

    and let $d = (f+g)$.
  - If *perturb* $= 1$, choose g' such that $f + g' \in \mathrm{dir}(\bar{x}, \mathcal{K})$, $\forall f \in L$ and $s^T(g+g') < 0$. Solve for f the system

    $$\begin{array}{rcl} A(f+g+g') & = & 0, \\ f & \in & L, \end{array}$$

    and let $d = (f+g+g')$.

Feasible direction method for conic optimization IV.

- 5. **Line search:** Find

$$\alpha^* \;=\; \max\{\alpha_i : x_i + \alpha_i d_i \in \mathcal{K}_i, i = 1, \ldots, r\}.$$

If $\alpha^* = \infty$ the primal is unbounded. STOP.

Else, set $\bar{x} = \bar{x} + \alpha^* d$, and go to step 1.

Special case I: Simplex method for LP.

1. Given $\bar{x} \geq 0$ we have

$$\mathbb{B}_X \;=\; \{d : d_i = 0, i \notin \text{support}(\bar{x})\}.$$

2. For a non-degenerate extreme point $\mathbb{B}_X$ is a non-redundant and spanning subspace. So no perturbation is ever needed. The simplex method chooses L as

$$L \;=\; \text{Range}\,(I(:, \text{support}(\bar{x})))\,.$$

   For this choice we also have $x \circ s = 0$ in step 2.

3. The descent direction $d = (f+g)$ where $g = e_k$ (k is the index for which s is the most negative). This descent direction is along an edge of the feasible set.

Special case II: Simplex method for SOCP I.

- 1. Given $\bar{x} = (\bar{x}_R, \bar{x}_I, \bar{x}_O)$, we choose $L = L_1 + L_2$ as follows:
  - L_1 contains vectors l_i, $i \in R \cup I$, whose ith block is the normalized $\bar{x}_i$ with zeros elsewhere (this ensures $\bar{x} \in L$), and elements that are the unit vectors in the blocks $i \in I$ and zeros elsewhere.
  - If $\bar{x}$ is f-degenerate, we choose L_2 to be the elements of tangent space for the blocks $i \in R$ and zeros elsewhere.

Special case II: Simplex method for SOCP II.

- 2. The improving direction $d = (f + g + g')$, where
  - $g = (0, \ldots, g_k, \ldots, 0)$ with $g_k = \frac{1}{\sqrt{2}}\left(1, -\frac{s_k(2:n_k)}{||s_k(2:n_k)||}\right)$ (k is the index for which $\lambda_{min}(s_i) = (s_i(1) - ||s_i(2:n_i)||)$, $i = 1, \ldots, r$ is the most negative).
  - If *perturb* $= 0$, then $g' = 0$. Else, g' is a vector which is some multiple of e_1 for all the blocks $i \in R$ which contributed to the subspace L_2, and zeros elsewhere.

- 3. The maximal step length computed in step 5 of the algorithm has a closed analytic expression.

Properties of the algorithm

Theorem. *Let $\{(x^k, y^k, s^k)\}$ be a sequence generated by the algorithm. Then for all k, x^k is primal feasible and $(x^k)^T s^k = 0$. At the k-th iteration, one of the following alternative cases arises:*

1. *If it stops in Step 2 then x^k, (y^k, s^k) are primal and dual optimal solutions to (CP) and (CD) respectively.*
2. *If it stops in Step 5, then (CP) is unbounded and (CD) is infeasible.*
3. *Otherwise, if x^k is also c-nondegenerate we have $c^T x^{k+1} < c^T x^k$.*

Preliminary computational results

Prob	m	n	lp	r	Opt	Obj(0)	Obj(iter)	Iter
hs51	14	22	10	1(12)	-5.99	-1.78	-5.99(*)	10
slp1	30	50	20	10(3)	52.83	86.01	52.83	200
slp2	50	130	100	10(3)	14.61	1669.43	14.61	200
slp3	50	160	100	[50 10]	38.51	1773.46	38.51(*)	183
slp4	100	180	150	10(3)	426.64	681.72	427.7	200
slp5	50	80	50	10(3)	45.67	103.66	45.67(*)	153
slp6	50	75	25	10(5)	24.05	85.19	24.17	500
slp7	50	150	100	10(5)	21.27	109.21	21.41	1000
slp8	50	110	100	[10]	82.10	300.68	82.10(*)	157
slp9	100	160	150	[10]	16.92	565.37	16.92(*)	128
slp10	100	300	150	[50 50 50]	705.79	2511.93	715.99	5000

Conclusions and future work

1. A primal simplex approach for conic optimization of which the primal simplex method for LP is a special case.
2. The simplex approach exploits the well known facial structure of conic optimization problems.
3. We have the framework for solving conic optimization problems over LP, SOCP and SDP cones.
4. We also have a dual simplex variant which mimics the dual simplex method for LP.
5. Currently investigating fast basis inverse (LU) updates to speed up the algorithm.
6. Future use in warm start after branching or the addition of cutting planes.

References

- F. Alizadeh and D. Goldfarb. Second-order cone programming, Mathematical Programming, 95(2003), pp. 3-51.

- F. Alizadeh, J.P.A. Haeberly, and M.L. Overton. Complementarity and Nondegeneracy in SDP, Mathematical Programming, 77(1997), pp. 111-128.

- M.S. Bazaraa, H.D. Sherali and C.M. Shetty. Nonlinear Programming: Theory and Algorithms, 2nd edition, John Wiley 1993.

- K. Krishnan, G. Pataki, N. Gupta and T. Terlaky. A feasible direction method for SOCP, forthcoming.

- I. Maros. Computational techniques of the simplex method. Kluwer International Series, 2003.

- G. Pataki. Cone-LP's and Semidefinite Programs: Geometry and a Simplex-type Method, Proceedings of the 5th IPCO conference, Springer Verlag 1996.

- G. Pataki. The Geometry of Semidefinite Programming, Handbook of SDP edited by H. Wolkowicz et al., Kluwer Academic Publishers, 2000, pp. 29-65.