Conic quadratic optimization - part 4

Erling D. Andersen
MOSEK ApS,
Fruebjergvej 3, Box 16,
2100 Copenhagen,
Denmark.
Email: e.d.andersen@mosek.com
Personal WWW: http://erling.andersen.name
Company WWW: http://www.mosek.com

April 11, 2008

Introduction

Topics

- Discuss different solution approaches for conic quadratic optimization problems.

Solving conic quadratic optimization problems

The Ben-Tal and Nemirovski linearization

- Can the second order cone be approximated by a polynomial number $(O(n))$ linear inequalities?
- The answer is **yes** as proved by Ben-Tal and Nemirovski [1].
  - Using an appropriate definition of approximated.
- See also the ph.d. thesis of Francois Glineur [2].

What is meant by approximation

Definition 1 *A set $\mathcal{U} \in \mathcal{R}^n$ is said to be an $\varepsilon-$approximation of the second-order cone $\mathcal{K}_q$ if and only if we have*

$$K_q \subseteq \mathcal{U} \subseteq \mathcal{K}_q^\varepsilon = \{x \in R^n : \; x_1(1+\varepsilon) \geq \|x_{2:n}\|\}$$

Main idea:

- Prove that there is $\varepsilon-$ approximation for the 3 dimensional quadratic cone.
- Prove that any quadratic cone can be written using a number of 3 dimensional cones.

Define the set

$$
\begin{array}{rcrcrl}
\multicolumn{6}{c}{\mathcal{F}^k := \{(r, \alpha_0, \ldots, \alpha_k, \beta_0, \ldots \beta_k) \in \mathcal{R}^{2k+3} :} \\
\alpha_{i+1} & = & \alpha_i \cos\left(\frac{\pi}{2^i}\right) & + & \beta_i \sin\left(\frac{\pi}{2^i}\right), & i = 0, \ldots, k-1, \\
\beta_{i+1} & \geq & \beta_i \cos\left(\frac{\pi}{2^i}\right) & - & \alpha_i \sin\left(\frac{\pi}{2^i}\right), & i = 0, \ldots, k-1, \\
-\beta_{i+1} & \geq & \beta_i \cos\left(\frac{\pi}{2^i}\right) & - & \alpha_i \sin\left(\frac{\pi}{2^i}\right), & i = 0, \ldots, k-1, \\
r & = & \alpha_k \cos\left(\frac{\pi}{2^k}\right) & + & \beta_k \sin\left(\frac{\pi}{2^k}\right)\} &
\end{array}
$$

Define

$$
\mathcal{G}^k := \{x \in \mathcal{R}^3 : \ (x_1, x_2, \alpha_1, \ldots, \alpha_k, x_3, \beta_1, \ldots \beta_k) \in \mathcal{F}\}.
$$

Lemma 1 *$\mathcal{G}^k$ is an $\varepsilon-$approximation for K_q^3 where*

$$\varepsilon = \cos\left(\frac{\pi}{2^k}\right)^{-1} - 1$$

For a proof see [2].

Quite good result because

k	$\varepsilon \leq$
2	0.5
4	0.02
8	1.0e-4
16	2.0e-9

The decomposition

Now

$$\mathcal{K}_q = \left\{ x \in \mathcal{R}^n : \; x_1^2 \geq \sum_{j=2}^{n} x_j^2, \; x_1 \geq 0 \right\}$$

Note

$$\begin{array}{rcl} \sum_{j=2}^{\lfloor \frac{n}{2} \rfloor} x_j^2 & \leq & y_l^2, \\ \sum_{j=\lceil \frac{n}{2} \rceil}^{n} x_j^2 & \leq & y_r^2, \\ y_l^2 + y_r^2 & \leq & x_1^2, \\ 0 & \leq & x_1, y_l, y_r. \end{array}$$

is another representation of the quadratic cone.

- The largest cone has about $\frac{1}{2}n$ variables in the new representation.
- Had to introduce 2 cones and 2 variables.
- Recursive application of this idea will produce a problem with about n 3 dimensional cones and n additional variables.

- If each 3 dimensional cone is $\varepsilon-$approximated then the approximation for the big cone is

$$\prod_{l=1}^{q}(1+\varepsilon)-1.$$

where $q \approx \log_2(n)$.

- Hence using $O(1)n$ variables and $O(1)n$ linear constraints it is possible to build a $\varepsilon-$ approximation to the quadratic cone.
- Warning: Approximate an quadratic problem, then no bound can be given on the quality of the objective value. [2].

Computational results

- The results of Glineur suggests
  - The linearized problems are very hard for the simplex algorithm.
  - The linearized problem can be solved using an interior-point reasonably well.
  - The primal-dual conic interior-point algorithm is much better.
- An interesting application in mixed integer conic optimization is reported in [5].

The conic simplex algorithm

A primal simplex like algorithm

Observe that any

$$x \in \mathcal{K}_q = \{x \in \mathcal{R}^n : \ x_1 \geq \|x_{2:n}\|\}$$

can be represented by

$$x = \sum_j \lambda_j \bar{x}^j, \ \lambda \geq 0$$

where

$$\bar{x}^j \in \mathcal{K}_q.$$

- x can be written as a positive sum of points in the cone.
- Extreme rays are sufficient.
- Nevertheless the sum may be infinite.
- Leads to a column generation approach.

- Developed by Goldfarb in [3].
- Not available in any commercial or academic code.
- My experience is that it works poorly in practice.

Smooth convex optimization approach

The conic quadratic constraint

$$x_1 \geq \|x_{2:n}\|$$

is equivalent to

$$x_1^2 \geq \|x_{2:n}\|^2, \; x_1 \geq 0$$

which in turn is equivalent to

$$x_1 \geq \frac{\|x_{2:n}\|^2}{x_1}, \; x_1 > 0.$$

almost. The function

$$f(x) = \frac{\|x_{2:n}\|^2}{x_1} - x_1$$

is a smooth, twice differentiable **convex** function on its domain.

Hence,

$$f(x) \leq 0$$

is a convex inequality.

- The quadratic cone can be handled with a convex inequality.
- Solve the problem using any algorithm/software for smooth convex optimization e.g. conopt, loqo and Minos.
- The $x_1 > 0$ constraint can(read will) cause problems for active set methods.
- A constraint of the form $x_1 > \varepsilon$ can introduce infeasibilities.

Interior point approaches

Interior point methods for linear optimization

The linear optimization problem:

$$\begin{array}{lrcl} \min & c^T x & & \\ \text{st} & Ax & = & b, \\ & x \geq 0. & & \end{array} \tag{1}$$

Assumptions:

- $A \in \mathcal{R}^{m \times n}$ is of full row rank.
- $\exists x^0$ such that $Ax^0 = b$ and $x^0 > 0$.

A barrier function

The function

$$B(x) = -\log(x)$$

is called a *barrier* function for the cone

$$\mathcal{K}_l = \{x \in \mathcal{R} : \; x \geq 0\}.$$

A barrier function is any function such that

$$\lim_{x \rightarrow 0^+} B(x) = +\infty.$$

The barrier problem

$$\begin{array}{lll} \min & c^T x - \mu \sum_{j=1}^{n} \ln(x_j) & \\ \text{st} & Ax & = b. \end{array} \tag{2}$$

- μ is a given **positive** parameter.
- Clearly, any feasible solution to (2) is a feasible solution to (1).
- Claim: As μ goes to 0 the optimal solution to (2) converge to the true optimal solution.

Define the Lagrange function

$$L(x, y) := c^T x - \mu \sum_{j=1}^{n} \ln(x_j) - y^T (Ax - b)$$

then the optimality conditions to (2) are

$$\begin{array}{rcccc} \nabla_x L(x,y) & = & c - \mu X^{-1} e - A^T y & = & 0, \\ \nabla_y L(x,y) & = & -Ax + b & = & 0, \\ & & x > 0 & & \end{array}$$

where

$$X^{-1} = diag(x_1^{-1}, \ldots, x_n^{-1}) = \begin{bmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & & \\ \vdots & & \ddots & 0 \\ 0 & & & x_n \end{bmatrix}^{-1}$$

and

$$e = (1, \ldots, 1)^T.$$

Now define

$$s := \mu X^{-1} e.$$

then the optimality conditions can be written as

$$\begin{array}{rcl} c - A^T y - s & = & 0, \\ -Ax + b & = & 0, \\ s - \mu X^{-1} & = & 0, \end{array}$$

or equivalently

$$\begin{array}{rcl} c - A^T y - s & = & 0, \\ -Ax + b & = & 0, \\ Xs & = & \mu e. \end{array} \tag{3}$$

Observe

$$Xs = \mu e$$

is equivalent to

$$x_j s_j = \mu.$$

The optimality conditions (3) says:

- Dual feasibility.
- Primal feasibility.
- Perturbed complementarity.

Observe that

$$
\begin{aligned}
c^T x - b^T y &= c^T x - (Ax)^T y \\
&= (c - A^T y)^T x \\
&= s^T x \\
&= e^T X s \\
&= \mu e^T e \\
&= \mu n.
\end{aligned}
$$

- Conclusion: Find a solution to the barrier problem (2) for μ sufficiently small using Newton's method.
- The barrier term gets rid of the inequalities!

http://www.mosek.com

The central path

Define $(x(\mu), y(\mu), s(\mu))$ as the solution to

$$\begin{array}{rcl} c - A^T y - s & = & 0, \\ -Ax + b & = & 0, \\ Xs & = & \mu e. \end{array}$$

- $(x(\mu), y(\mu), s(\mu))$ defines a continuous curve parameterized by μ,
- Called the *central path*.
- Algorithmic idea.
  - Follow the central path to the optimum.
  - Leads to polynomial complexity.

A primal algorithm

Define

$$F^p_\mu(x,y,s) = \begin{bmatrix} c - A^T y - s \\ -Ax + b \\ s - \mu X^{-1} e \end{bmatrix}.$$

Assume we are given a (x^0, y^0, s^0) such that

$$Ax^0 = b \text{ and } x^0 > 0.$$

Also let

$$\mu^0 > 0$$

be given.
Solve

$$F^p_{\gamma\mu^0}(x,y,s) = 0 \tag{4}$$

approximately where

$$0 < \gamma < 1.$$

How to solve (4) approximately? Use Newtons method i.e.

$$\nabla F^p_{\gamma\mu^0}(x^0, y^0, s^0) \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix} = F^p_{\gamma\mu^0}(x^0, y^0, s^0)$$

and

$$\begin{bmatrix} x^+ \\ y^+ \\ s^+ \end{bmatrix} = \begin{bmatrix} x^0 \\ y^0 \\ s^0 \end{bmatrix} + \alpha \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix}$$

for suitable chosen step size $\alpha \in (0, 1]$ i.e.

$$x^+ > 0.$$

The explicit Newton equation system

$$
\begin{array}{rcl}
-A^T d_y - d_s & = & -(c - A^T y^0 - s^0), \\
-A d_x & = & -(-A x^0 + b), \\
d_s - \gamma \mu^0 (X^0)^{-2} d_x & = & -(s^0 - \gamma \mu^0 (X^0)^{-1})
\end{array}
$$

- Easy fact:

$$Ax^+ = b.$$

- For suitable chosen starting point and γ:
  - $\alpha = 1$ is allowed.
  - Leads to convergence of order $\frac{1}{\sqrt{n}}$.
  - Polynomial complexity ($O(\sqrt{n})$ iterations).

A primal-dual algorithm

Assume we are given a (x^0, y^0, s^0) such that

$$\begin{array}{rcll} Ax^0 & = & b, & x^0 > 0, \\ A^T y^0 + s^0 & = & c, & s^0 > 0. \end{array}$$

i.e. a primal-dual feasible interior solution.
Define

$$F_\mu^{pd}(x, y, s) = \begin{bmatrix} c - A^T y - s \\ -Ax + b \\ Xs - \mu e \end{bmatrix}.$$

The primal-dual search direction

$$\nabla F_{\gamma\mu^0}^{pd}(x^0, y^0, s^0) \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix} = F_{\gamma\mu^0}^{pd}(x^0, y^0, s^0)$$

Or explicitly

$$
\begin{array}{rcl}
-A^T d_y - d_s & = & -(c - A^T y^0 - s^0), \\
-A d_x & = & -(-A x^0 + b), \\
X^0 d_s + S^0 d_x & = & -(X^0 s^0 - \gamma \mu^0 e)
\end{array}
$$

- (x^+, y^+, s^+) is computed the same way as in the primal algorithm
  - Except $s^+ > 0$ is required too.
- Primal and dual feasibility are preserved.
- Is the basis for all commercially available interior-point based software for linear optimization.
- Target: Generalize this algorithm to conic quadratic optimization.

A primal interior-point approach to conic quadratic optimization

First algorithm

Primal problem:

$$
\begin{array}{lccc}
\min & c^T x & & \\
\text{st} & Ax & = & b, \\
& x_1 - \frac{\|x_{2:n}\|^2}{x_1} & \geq & 0, \\
& x_1 & \geq & 0.
\end{array}
\tag{5}
$$

- We assume ONE cone for simplicity.
- The barrier function:

$$
B(x) = -\frac{1}{2}\left(\ln\left(x_1 - \frac{\|x_{2:n}\|^2}{x_1}\right) + \ln(x_1)\right).
$$

Primal barrier problem:

$$\begin{array}{lrcl} \min & c^T x - \frac{\mu}{2}\left(\ln\left(x_1 - \frac{\|x_{2:n}\|^2}{x_1}\right) + \ln(x_1)\right) & & \\ & Ax & = & b, \end{array}$$

- μ is positive parameter.
- The optimum is in the interior for $\mu > 0$.
- Claim: For $\mu \to 0$ the optimum converges to the true optimum.

http://www.mosek.com

Optimality conditions:

$$
\begin{aligned}
Ax &= b, \\
c_1 - \frac{\mu}{2(x_1 - \frac{\|x_{2:n}\|^2}{x_1})}\left(1 + \frac{\|x_{2:n}\|^2}{x_1^2}\right) - \frac{\mu}{2} x_1^{-1} - a_{:1}^T y &= 0, \\
c_{2:n} + \frac{\mu}{x_1 - \frac{\|x_{2:n}\|^2}{x_1}} \frac{x_{2:n}}{x_1} - A_{:(2:n)}^T y &= 0.
\end{aligned}
$$

Comments on the primal algorithm

- Solve optimality conditions using Newton's method.
- While decreasing μ.
- Given the right assumptions it leads to polynomial complexity.
- Does not exploit primal-dual structure.
- A primal-dual algorithm should work better.

The reformulated optimality conditions

Define:

$$
\begin{aligned}
\bar{y} &:= \frac{\mu}{2(x_1 - \frac{\|x_{2:n}\|^2}{x_1})}, \\
s_1 &:= \frac{\mu}{2x_1}.
\end{aligned}
$$

Reformulated system:

$$
\begin{aligned}
Ax &= b, \\
c_1 - \bar{y}(1 + \frac{\|x_{2:n}\|^2}{x_1^2}) - a_{:1}^T y - s_1 &= 0, \\
c_{2:n} + \bar{y}\frac{x_{2:n}}{x_1} - A_{:(2:n)}^T y &= 0, \\
\bar{y}(x_1 - \frac{\|x_{2:n}\|^2}{x_1}) &= \frac{\mu}{2}, \\
x_1 s_1 &= \frac{\mu}{2}.
\end{aligned}
$$

Observations:

- Is perturbed KKT system to (5).
- $\bar{y}$ and s_1 are dual multipliers.

A nonsymmetric primal-dual interior-point algorithm

- Solve system using Newton's method.
- While decreasing perturbation.
- Authors suggesting this idea:
  - McCormick,
  - Vial, Vial and Anstreicher,
  - Andersen and Ye,
  - Vanderbei.

A small example

Study

$$\begin{array}{lrcl} \min & x_1 & & \\ \text{s.t.} & \frac{x_2^2}{x_1} & \leq & x_1, \\ & x_1 & \geq & 0. \end{array}$$

Perturbed primal-dual system:

$$\begin{array}{rcl} 1 - \bar{y}(1 + \frac{x_2^2}{x_1^2}) - s_1 & = & 0, \\ \bar{y}\frac{x_2}{x_1} & = & 0, \\ \bar{y}(x_1 - \frac{x_2^2}{x_1}) & = & \frac{\mu}{2}, \\ x_1 s_1 & = & \frac{\mu}{2}. \end{array}$$

Observations:

- One solution

$$(x_1, x_2, \bar{y}, s_1) = (\epsilon, \beta\epsilon, 0, 1)$$

  where $\beta \in (0, 1)$. Is not strictly complementarity.
- The nonsymmetric primal-dual algorithm converge to

$$(x_1, x_2, \bar{y}, s_1) = (\epsilon, \beta\epsilon, 0.5, 0.5)$$

  for "bad" starting points. Point is strictly complementarity.
- See also the discussion of this example in [4].

- Implies dual feasibility is not achieved.
- Computational results indicates:
  - It works for nice problems.
  - Computes primal optimal solution.
  - Has severe problems with dual feasibility.

Summary

- Shown that conic quadratic problems can be approximated by a linear problem.
  - Is inefficient in most applications.
  - Generates hard problems for the simplex methods.
- Outline the conic simplex algorithm.
  - But has not proven to be efficient in practice.
- Discussed a couple of interior-point approaches.
  - Primal interior-point methods works well at least in theory.

References

[1] A. Ben-Tal and A. Nemirovski. On polyhedral approximations of the second-order cone. *Math. Oper. Res.*, 26(2):193–205, 2001.

[2] F. Glineur. *Topics in Convex Optimization: Interior-Point Methods, Conic Duality and Approximations*. PhD thesis, Service de Mathematiqué et de Recherche Opérationelle, Faculté Polytechnique de Mons, Rue de Houdain, 9, B-7000 Mons, Belgium, 2001.

[3] D. Goldfarb. The simplex method for conic programming. Technical Report "TR-2002-05", IOER, Columbia University, New York, 2002.

[4] Robert J. Vanderbei and Hande Yurttan. Using LOQO to solve second-order cone programming problems . Technical Report SOR-98-9, Department of Civil engineering and Operations Research, Princeton University, 1998.

[5] Juan Pablo Vielma, Shabbir Ahmed, and George L. Nemhauser. A lifted linear programming branch-and-bound algorithm for mixed integer conic quadratic programs. Technical report, Georgia Institute of Technology, 2007.

Exercises

1. Prove that the function

$$f(x) = \frac{\|x_{2:n}\|^2}{x_1} - x_1$$

is convex on its domain.

2. Consider the problem

$$\begin{array}{lll} \min & x_1 & \\ \text{st} & x_1^2 + x_2^2 & \leq & 1. \end{array}$$

- Rewrite the problem as conic quadratic optimization problem.
- Specify an linear relaxation that is $\varepsilon = 1.0e - 3$ accurate.
- Solve the linear relaxation using you favorite software.

3. Implement the primal-dual interior-point algorithm for linear problems in MATLAB.