Conic quadratic optimization - part 1

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 10, 2008

Introduction

Topics

- Who is the speaker.
- What is conic quadratic optimization and why is it interesting.
- Some applications of conic quadratic optimization.

Who is the teacher?

The speaker:

- Holds a ph.d. in Economics (read OR and optimization).
- Main interest is algorithms for linear and convex optimization problems.
- Is CEO and lead developer at MOSEK ApS.

What is MOSEK?

- A software package.
- Solves large-scale sparse optimization problems.
- Handles **linear**, **conic**, and **nonlinear** convex problems.
- Stand-alone as well as embedded.
- Used to solve problems with up to millions of constraints and variables.
- Version 1 release in 1999.
- Version 5 released summer 2007.

For details about interfaces, trials, etc. see

http://www.mosek.com.

Motivation for conic optimization

A detour around linear optimization

$$
\begin{array}{lll}
(P) & \min & c^T x \\
 & \text{st} & Ax = b, \\
 & & x \geq 0.
\end{array}
$$

Pros:

- Huge number applications.
- Powerful theory associated e.g. Farkas' lemma and duality.
- Highly efficient algorithms exists (simplex and interior-point).
- Easy to represent the problem using c, A, and b.
- Several solver software packages are available e.g. MOSEK, CPLEX, Xpress.
- Several modeling software packages are available e.g. GAMS, AMPL, AIMMS.
- In short: Linear optimization is easy to use.

Cons:

- The linearity assumption is restrictive.
- For instance the unit ball

$$x_1^2 + x_2^2 \leq 1$$

can only be approximated.

Question:

- Is it possible to generalize linear optimization while keeping all the good properties?

One answer: nonlinear optimization

$$
\begin{array}{llll}
(NO) & \min & f(x) & \\
 & \text{st} & g(x) & \leq & 0.
\end{array}
$$

Pros:

- Very general.

Cons:

- Duality theory is somewhat fuzzy.
- Lack of good algorithms.
- Local versus global optimums.
- Convexity (how to check).
- Black box model.
- How to compute gradients and Hessians.
- How to handle f and g in software.

Another answer: conic quadratic optimization

The linear optimization problem involves the convex cone

$$\{x \in \mathcal{R} : \; x \geq 0\}$$

denoted the *linear cone*.
Recall a **convex** set $\mathcal{K}$ is a convex cone if

$$x \in \mathcal{K} \Rightarrow \lambda x \in \mathcal{K}, \forall \lambda \geq 0.$$

A generalization of linear optimization is to allow more general cones i.e the *quadratic cone*:

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

The quadratic cone is also known as

- The second order cone.
- The Lorentz cone.
- The ice cream cone.

An example:

$$
\begin{array}{lrcl}
\min & & x_5 & \\
\text{st} & 2x_1 + 3x_2 - 1 & = & x_3, \\
 & 1x_1 + 7x_2 - 2 & = & x_4, \\
 & x_5 \geq \sqrt{x_3^2 + x_4^2}. & &
\end{array}
$$

Equivalent specifications:

- $$\left\{ x \in \mathcal{R}^n : \; x_1 \geq \sqrt{\sum_{j=2}^{n} x_j^2} \right\}$$
- $$\{ x \in \mathcal{R}^n : \; x_1 \geq \| x_{2:n} \| \}$$
- $$\left\{ x \in \mathcal{R}^n : \; x_1^2 \geq \sum_{j=2}^{n} x_j^2, \; x_1 \geq 0 \right\}$$
- $$\{ x \in \mathcal{R}^n : x \succeq_Q 0 \}$$

Notes:

- All norms are 2 norms unless otherwise stated.
- Conic optimization may be seen as a way of generalizing the inequality ($\geq$).
- Many applications will be shown later.

http://www.mosek.com

Conic optimization

The generic problem

The generic conic optimization problem:

$$\begin{array}{lccc} \min & \sum_k (c^k)^T x^k & & \\ \text{st} & \sum_k A^k x^k & = & b, \\ & x^k \in \mathcal{K}^k. & & \end{array}$$

where

- $c^k \in \mathcal{R}^{n^k}$.
- $A^k \in \mathcal{R}^{m \times n^k}$.
- $b \in \mathcal{R}^m$.
- $\mathcal{K}^k$ is a nonempty pointed convex cone i.e.
  - (Convexity) $\mathcal{K}^k$ is a convex set.
  - (Conic) $x \in \mathcal{K}^k \Rightarrow \lambda x \in \mathcal{K}^k, \ \forall \lambda \geq 0$.
  - (Pointed) $x \in \mathcal{K}^k$ and $-x \in \mathcal{K}^k \Rightarrow x = 0$.

A recap. of linear duality

The linear optimization problem

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

has the dual problem

$$\begin{array}{llll} \max & b^T y & & \\ \text{st} & A^T y + s & = & c, \\ & s \geq 0. & & \end{array} \tag{2}$$

Weak duality

Let (x, y, s) be a primal and dual feasible solution then

$$c^T x \geq b^T y$$

holds.

Comments:

- How to interpretate this fact?
- What can this fact be used to?
- How to prove this fact?

Strong duality

Well-known facts:

- (1) has an optimal solution if and only if a solution (x, y, s) exist such that

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

- (1) is primal infeasible if and only a (y, s) exists such that

$$A^T y + s = 0,\ b^T y > 0,\ s \geq 0.$$

- (1) is dual infeasible (i.e. (2) is infeasible) if and only a x exists such that

$$Ax = 0,\ c^T x < 0,\ x \geq 0.$$

Duality theory is very powerful:

- Makes it easy to verify optimality.
- Makes it easy to certify that a problem is infeasible.
    - Think about how to prove you speak english.
    - And how you prove you do not speak english.
- Employed extensively within algorithms.

Conic duality

Given a convex cone $\mathcal{K}$ then the dual cone $\mathcal{K}^*$ is given by

$$\mathcal{K}^* := \{s : \; s^T x \geq 0, \; \forall x \in \mathcal{K}\}.$$

Given the primal conic optimization

$$\begin{array}{lccc} \min & \sum_k (c^k)^T x^k & & \\ \text{st} & \sum_k A^k x^k & = & b, \\ & x^k \in \mathcal{K}^k. & & \end{array} \tag{3}$$

then the corresponding dual problem is

$$\begin{array}{lccc} \max & b^T y & & \\ \text{st} & (A^k)^T y + s^k & = & c^k, \\ & s^k \in (\mathcal{K}^k)^*. & & \end{array} \tag{4}$$

Conic duality in the linear case

Observe the dual cone corresponding to the linear cone

$$\{x \in \mathcal{R} : \; x \geq 0\}$$

is

$$\{s \in \mathcal{R} : \; s \geq 0\}.$$

- The linear cone is self-dual i.e.

$$\mathcal{K} = \mathcal{K}^*.$$

- In the linear case conic duality is equivalent to the usual linear optimization duality.

Conic duality cont.

Observe

- Weak duality holds:

$$\begin{array}{rcl}\sum_k (c^k)^T x^k - b^T y & = & \sum_k ((A^k)^T y + s^k)^T x^k - b^T y \\ & = & b^T y + \sum_k (x^k)^T s^k - b^T y \\ & = & \sum_k (x^k)^T s^k \\ & \geq & 0.\end{array}$$

- All the usual duality relations holds **ALMOST** in the conic case.
- More details later.

Recap. about conic optimization

- Conic optimization is a way to generalize linear optimization.
- The nonlinearity is placed in the cone.
- Hence, a way of generalizing what is meant by inequality.
- A conic problem has a meaningful dual problem just as in the linear case.
- The problem is completely specified by c, A, and b plus some simple information about the cone structure.
- Hence, conic optimization is almost like linear optimization.

http://www.mosek.com

Conic quadratic optimization

A few facts

Lemma 1 *The set is*

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

is convex cone.

Lemma 2 *The quadratic cone is self dual i.e.*

$$\mathcal{K}_q = \mathcal{K}_q^*.$$

Two special quadratic cones

- One dimensional quadratic cone:

$$\{x \in \mathcal{R}^1 : x_1 \geq 0\}.$$

- Two dimensional quadratic cone:

$$\{x \in \mathcal{R}^2 : x_1 \geq \|x_2\|\} = \{x \in \mathcal{R}^2 : x_1 \geq |x_2|\}.$$

Conclusion:

- Linear optimization is a special case of conic quadratic optimization.

The rotated quadratic cone

Consider the set

$$\frac{1}{2}\|x\|^2 + f^T x \leq g$$

which is equivalent to

$$\begin{array}{rcl} z + f^T x & = & g, \\ y & = & 1, \\ \|x\|^2 & \leq & 2zy, \quad z, y \geq 0. \end{array}$$

Next define

$$z = \frac{u+v}{\sqrt{2}} \text{ and } y = \frac{u-v}{\sqrt{2}}.$$

This implies

$$\begin{array}{rcl} 2zy & = & \frac{u+v}{\sqrt{2}} \frac{u-v}{\sqrt{2}} \\ & = & u^2 - v^2. \end{array}$$

Hence we obtain

$$\begin{array}{rcl} \frac{u+v}{\sqrt{2}} + f^T x & = & g, \\ \frac{u-v}{\sqrt{2}} & = & 1, \\ \|x\|^2 + v^2 & \leq & u^2, \quad u \geq 0. \end{array}$$

Clearly,

$$\|x\|^2 + v^2 \leq u^2, \; u \geq 0.$$

is the standard quadratic cone.

Comments:

- The set
$$\left\{ x \in \mathcal{R}^n : \; 2x_1x_2 \geq \sum_{j=3}^{n} x_j^2, \; x_1, x_2 \geq 0 \right\}$$
is called the rotated quadratic cone.
- The rotated quadratic cone is identical to the quadratic cone under a linear transformation.
- Implies we can use the rotated quadratic whenever we like.

Applications

Convex quadratic optimization

Quadratic optimization

$$
\begin{array}{lllll}
\min & 0.5||Q^0 x||^2 + c^T x & & \\
\text{st} & 0.5||Q^i x||^2 + a_{i:}x & \leq & b_i, \forall i = 1, 2, \ldots .
\end{array}
$$

Conic quadratic equivalent:

$$
\begin{array}{llll}
\min & c^T x + t_0 & & \\
\text{st} & t_i + a_{i:}x & = & b_i, & \forall i = 1, 2, \ldots, \\
 & Q^i x - y^i & = & 0, & \forall i = 0, 1, \ldots, \\
 & z_i & = & 1, & \forall i = 0, 1, \ldots, \\
 & ||y^i||^2 & \leq & 2 t_i z_i, & \forall i = 0, 1, \ldots .
\end{array}
$$

Because

$$
\frac{1}{2}||Q^i x||^2 \leq t_i, \ \forall i = 0, 1, \ldots .
$$

Applications:

- Finance.
- Approximation of more general nonlinear problems.
- Constrained linear least squares.

Notes:

- The model contains fixed variables naturally.
  - Eliminating the fixed variables destroys the duality.
  - Fixed variables can be exploited computationally.
- A problem size expansion may occur when stating the problem on conic form.

Minimum sum of norms

Consider the minimum sum of norms problem

$$\begin{array}{lllll} \min & \sum_k \left\| x^k \right\| & & \\ \text{st} & \sum_k A^k x^k & = & b, \end{array}$$

Conic quadratic reformulation

$$\begin{array}{lllll} \min & \sum_k t_k & & \\ \text{st} & \sum_k A^k x^k & = & b, \\ & t_k \geq \left\| x^k \right\| & & \end{array}$$

Applications:

- Image denoising.
- Location problems.

The dual problem is

$$
\begin{array}{lll}
\max & b^T y & \\
\text{st} & (A^k)^T y + s^k & = & c^k, \\
 & u_k & = & 1, \\
 & u_k \geq \left\| s^k \right\| & &
\end{array}
$$

where u_k and s^k are dual variables corresponding to t_k and x^k respectively.

Equivalent dual problem after eliminating u is

$$
\begin{array}{llll}
\max & b^T y & & \\
\text{st} & (A^k)^T y + s^k & = & c^k, \\
 & \left\| s^k \right\|^2 & \leq & 1.
\end{array}
$$

Observe how easily the last dual problem was obtained using conic duality.

The Fermat-Weber problem

- Assume k customers are given each located at position d^k.
- Assume we want to place a new facility at position x such that

$$\min \sum_k \left\| x - d^k \right\|$$

i.e. the total distance to the costumers are minimized.
- A special case of the minimizing a sum of norms problem.

Image restoration

- A $n \times n$ image is represented by $n \times n$ matrix.
- Let
$$F \in \mathbf{R}^{n \times n}$$
be the observed image.
- Let
$$U \in \mathbf{R}^{n \times n}$$
be the original image.
- Let
$$V \in \mathbf{R}^{n \times n}$$
be some noise in the image.

We have
$$U + V = F.$$

Problem:

- U and V are unknown.
- How to estimate V?

The total variation model

The model:

$$\begin{array}{lccc} \sum_{ij} & t_{i,j} & & \\ & U+V & = & F, \\ & \left\| \begin{array}{c} u_{i,j} - u_{(i+1),j} \\ u_{i,j} - u_{i,(j+1)} \end{array} \right\| & \leq & t_{i,j}, \\ & \|V\|_F & \leq & \sigma \end{array}$$

where

$$\|V\|_F := \sqrt{\sum_{i,j} v_{i,j}^2}$$

and σ is user specified constant. Usually chosen related to amount of expected amount of noise.
See [3] for more details.

Noisy picture

Estimated picture

Other norm minimization problems

Consider the problem of minimizing the maximum of some norms

$$\min \quad \max_k \left\| A^k x^k + b^k \right\|$$

Conic quadratic reformulation

$$\begin{array}{lll} \min & v & \\ \text{st} & A^k x^k + b^k & = & -y^k, \\ & t_k & \leq & v, \\ & t_k \geq \left\| y^k \right\|. & & \end{array}$$

mosek http://www.mosek.com

Hyperbolic optimization

Consider the problem

$$\begin{array}{lllll} \min & \sum_j \frac{c_j}{x_j} & & \\ \text{st} & Ax & = & b, \\ & x \geq 0, & & \end{array}$$

where $c_j > 0$.
Conic quadratic reformulation:

$$\begin{array}{llll} \min & \sum_j c_j t_j & & \\ \text{st} & Ax & = & b, \\ & z_j & = & \sqrt{2}, \\ & z_j^2 & \leq & 2x_j t_j, \\ & x \geq 0. & & \end{array}$$

http://www.mosek.com

Applications:

- Equilibrium in TCP networks.
- Stratified sampling.
- Stock optimization models.

The square root

Consider

$$\begin{array}{rcl} \sqrt{x} & \geq & |t|, \\ x & \geq & 0, \end{array}$$

where both t and x are variables. CQ reformulation

$$\begin{array}{rcl} t^2 & \leq & 2xz, \\ z & = & 0.5, \\ x, z & \geq & 0. \end{array}$$

Given $c_j > 0$ then

$$\begin{array}{rcl} \sum_j c_j \sqrt{x_j} & \geq & v, \\ x \geq 0 & & \end{array}$$

is equivalent to

$$\begin{array}{rcl} \sum_j c_j t_j & \geq & v, \\ t_j^2 & \leq & 2 x_j z_j, \\ z_j & = & 0.5, \\ x_j, z_j, & \geq & 0. \end{array}$$

($t_j < 0.0$ may occur but does it matter?).

The 3/2 power function

Consider

$$\begin{array}{rcl} x^{1.5} & \leq & t, \\ 0 & \leq & x. \end{array}$$

Note the simple fact

$$x^{1.5} = \frac{x^2}{\sqrt{x}}.$$

First define the set

$$\begin{array}{rcl} x^2 & \leq & 2st, \\ s, t & \geq & 0. \end{array}$$

Now if we can make sure that

$$2s \leq \sqrt{x},$$

then we have the desired result because this implies

$$x^{1.5} = \frac{x^2}{\sqrt{x}} \leq \frac{x^2}{2s} \leq t.$$

Observe s can be chosen freely and $\sqrt{x} = 2s$ is a valid choice. Let

$$\begin{array}{rcl} x^2 & \leq & 2st, \\ w^2 & \leq & 2vr, \\ x & = & v, \\ s & = & w, \\ r & = & \frac{1}{8}, \\ s,t,v,r & \geq & 0, \end{array} \tag{5}$$

then

$$
\begin{aligned}
s^2 &= w^2 \\
&\le 2vr \\
&= \frac{v}{4} \\
&= \frac{x}{4}
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
x^2 &\le 2st, \\
&\le 2\sqrt{\frac{x}{4}}t
\end{aligned}
$$

leading to the conclusion

$$x^{1.5} \le t.$$

A note

Consider the problem

$$\begin{array}{lrcl} \max & x & & \\ \text{st} & t & = & 1, \\ & z & = & 1, \\ & x^2 & \leq & 2tz, \end{array}$$

- The optimal solution is $x = \sqrt{2}$.
- All data are rational but the solution is irrational.
- Has important implications for complexity.
- For linear optimization rational data implies a rational solution.

Summary

- The conic optimization model has been introduced.
  - Is a generalization of linear optimization.
  - The $\geq$ is generalized.
- The special case of conic quadratic optimization is introduced.
  - Some applications has be shown.
- A lot material for this lecture is available in [1, 2, **?**, 4].

http://www.mosek.com

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Programming*, 95(1):3–51, 2003.

[2] A. Ben-Tal and A Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS/SIAM Series on Optimization. SIAM, 2001.

[3] D. Goldfarb and W. Win. Second-order cone programming methods for total variation-based image restoration. *SIAM Journal Scientific Computing*, 27(2):622–645, 2005.

[4] M. S. Lobo, L. Vanderberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra Appl.*, 284:193–228, November 1998.

[5] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004. http://www.stanford.edu/~boyd/cvxbook/.

Exercises

1: Convexity

1. Assume $f(x)$ is a twice differentiable convex function. What holds about the Hessian?
2. Given the function

$$f(x) = 0.5x^T Hx + c^T x + b$$

then specify under which conditions on H, v and b

- The set $\{x : \ f(x) = 0\}$ is convex.
- The set $\{x : \ f(x) \leq 0\}$ is convex.
- The set $\{x : \ f(x) \geq 0\}$ is convex.