Introduction to conic optimization: Lecture 1

June 20th 2016 - June 22th 2016

e.d.andersen@mosek.com

www.mosek.com

Section 1

Introduction

The lecturer

- Education: ph.d. in Economics (read OR and optimization).
- Interests: Algorithms and software for linear and convex optimization problems.
- Job: CEO and chief scientist at MOSEK ApS.
- Homepage: `http://erling.andersens.name`
  - Twitter link.
  - Other social media links.

Mosek

- A software package.
- Solves large-scale sparse optimization problems.
- Handles **linear**, **conic**, and **nonlinear** convex problems.
- Stand-alone as well as embedded.
- Version 1 release in 1999.
- Version 8 to be released Fall 2016.

For details about interfaces, trials, academic license etc. see

`https://mosek.com`.

Course objective

Learn:

- Conic optimization.
- Advantages of conic optimization.
- Extreme disciplined modeling.
    - Case studies.
    - Involves programming in Python.

Section 2

Conic optimization

Linear optimization

A detour around linear optimization (LO)

$$
(PLO) \quad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \\ & x \geq 0. \end{array}
$$

Pros:

- Extremely structured.
- Leads to powerful theory e.g. Farkas' lemma and duality.
- Leads to powerful solution algorithms e.g. simplex and interior-point algorithms.
- Easy representation using the data c, A, and b.
- In short: Linear optimization is very powerful when applicable.

Cons:

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

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

  can only be approximated.

Question:

- How to generalize linear optimization to a broader class of problems?

One generalization of LO is nonlinear optimization (NLO):

$$\begin{array}{llll} (PNLO) & \text{minimize} & f(x) & \\ & \text{subject to} & g_i(x) & \leq \;\; 0, \quad \forall i. \end{array}$$

- Very general model.

The cons

- No structure.
- Leads to weak duality theory with local versus global optimum issues.
- Leads to lack of good algorithms.
- Introduce some structure.
  - Assume once or twice differentiability
  - Convexity. Removes global versus local issue.
- Checking convexity is NP hard. (And for users e.g. CVX forum.)
- Differentiability does not say much about structure.
- How to compute gradients and Hessians.
- How to handle f and g in software.

- $(PNLO)$ is very general.
- Too general to obtain good results.
- Let us locate a special class of nonlinear optimization problems with a good structure!

Alternatively replace the linear inequality

$$x \geq 0$$

with

$$x \geq_{\mathcal{K}} 0 \quad \Leftrightarrow \quad x \in \mathcal{K}.$$

where $\mathcal{K}$ is a **convex cone**.
By definition

$$a \geq_{\mathcal{K}} b \quad \Leftrightarrow \quad a - b \geq_{\mathcal{K}} 0 \quad \Leftrightarrow \quad a - b \in \mathcal{K}.$$

Convex cone

$\mathcal{K}^k$ is a nonempty pointed convex cone i.e.

- (Convexity) $\mathcal{K}$ is a convex set.
- (Conic) $x \in \mathcal{K} \Rightarrow \lambda x \in \mathcal{K}, \ \forall \lambda \geq 0.$
- (Pointed) $x \in \mathcal{K}$ and $-x \in \mathcal{K} \Rightarrow x = 0.$

Comments:

- Wikipedia reference: `https://en.wikipedia.org/wiki/Convex_cone`.
- What is the point about convex cones?
  - Using only 3 different cone types a large number of problems can modelled.

Generic conic optimization problem

Primal form

$$
\begin{array}{lll}
\text{minimize} & \sum_k (c^k)^T x^k & \\
\text{subject to} & \sum_k A^k x^k & = \; b, \\
 & x^k \in \mathcal{K}^k &
\end{array}
$$

where

- $c^k \in \mathbb{R}^{n^k}$,
- $A^k \in \mathbb{R}^{m \times n^k}$,
- $b \in \mathbb{R}^m$.

$$
\begin{array}{lrcl}
\text{maximize} & b^T y & & \\
\text{subject to} & c^k - (A^k)^T y & \in & \mathcal{K}^k, \forall k.
\end{array}
$$

- Equally general.
- Problems are convex.
- The objective sense is not important.

Beauty of conic optimization

- Separation of data and structure:
  - Data: c^k, A^k and b.
  - Structure: $\mathcal{K}$.
- Convexity is **built in**. Given by $\mathcal{K}$.
- Many (but not all) convex models can be formulated with **3** convex cones:
  - The linear.
  - The quadratic.
  - The semidefinite.

The quadratic cone

The most basic nonlinear generalization of the linear cone

$$\{x \in \mathbb{R} : \ x \geq 0\}$$

is the *quadratic cone*:

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

also known as

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

An example

$$\begin{array}{ll} \text{minimize} & x_5 \\ \text{subject to} & \begin{array}{rcl} 2x_1 + 3x_2 - 1 & = & x_3, \\ 1x_1 + 7x_2 - 2 & = & x_4, \end{array} \\ & x_5 \geq \sqrt{x_3^2 + x_4^2}. \end{array}$$

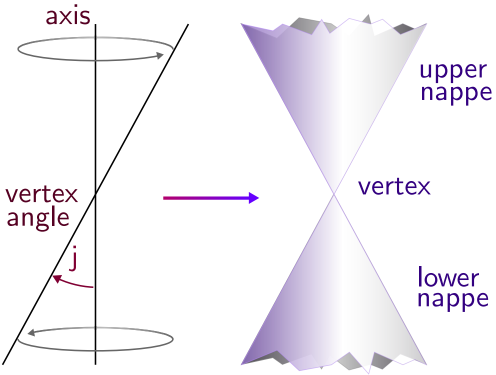
or equivalently

$$\begin{array}{ll} \text{minimize} & x_5 \\ \text{subject to} & \begin{bmatrix} x_5 \\ 2x_1 + 3x_2 - 1 \\ 1x_1 + 7x_2 - 2 \end{bmatrix} \in \mathcal{K}_q \end{array}$$

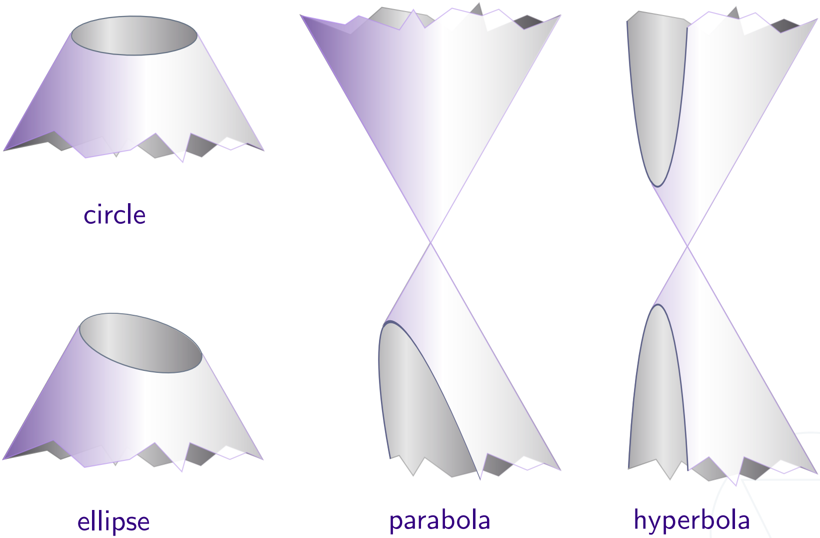
The quadratic cone
Equivalent specifications

$$
\begin{aligned}
&\mathcal{K}_q \\
= \quad & \left\{ x \in \mathbb{R}^n : \ x_1 \geq \sqrt{\sum_{j=2}^{n} x_j^2} \right\} \\
= \quad & \{ x \in \mathbb{R}^n : \ x_1 \geq \|x_{2:n}\| \} \\
= \quad & \left\{ x \in \mathbb{R}^n : \ x_1^2 \geq \sum_{j=2}^{n} x_j^2, \ x_1 \geq 0 \right\} \\
= \quad & \left\{ x \in \mathbb{R}^n : x \succeq_{\mathcal{K}_q} 0 \right\}
\end{aligned}
$$

The quadratic cone illustrated



Slices of the quadratic cone



History of the quadratic cone

- The old Greeks studied conics.
- Graphics and info. from:

  http://platonicrealms.com/encyclopedia/conics

Section 3

Conic modelling

Modeling with the quadratic cone

- What can be modelled using linear and quadratic cones only?
- What are the typical tricks used in conic modeling?

Consider the linear least squares problem

$$\begin{array}{ll} \text{minimize} & \|Ax - b\| \\ \text{subject to} & x \geq 0. \end{array}$$

Is that conic quadratic representable?
Let us linearize the objective

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \|Ax - b\| \leq t, \\ & x \geq 0. \end{array}$$

Are the constraints conic representable?

Yes, using

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} t \\ Ax - b \end{bmatrix} \in \mathcal{K}_q, \\ & x \geq 0. \end{array}$$

Observer

- Arbitrary linear constraints can easily be added to the problem!

ℓ_1 norm minimization

Consider the ℓ_1 problem

$$\text{minimize} \quad \|Ax - b\|_1 .$$

where

$$\|x\|_1 = \sum_j |x_j|.$$

Observe if $x \in \mathbb{R}^1$ then

$$\begin{bmatrix} t \\ x \end{bmatrix} \in \mathcal{K}_q$$

implies

$$t \geq |x|.$$

Therefore,

$$\begin{array}{ll} \text{minimize} & e^T t \\ \text{subject to} & \begin{bmatrix} t_i \\ A_{i:}x - b_i \end{bmatrix} \in \mathcal{K}_q, \\ & x \geq 0. \end{array}$$

where e is the vector of all ones. I.e.

$$e^T t = \sum_j t_j.$$

The rotated quadratic cone

Consider the set

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

which is equivalent to

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

Next define

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

This implies

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

Hence we obtain

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

The last constraint is equivalent to

$$\begin{bmatrix} u \\ v \\ x \end{bmatrix} \in \mathcal{K}_q.$$

Comments:

- The set

$$\mathcal{K}_r := \left\{ x \in \mathbb{R}^n : \; 2x_1x_2 \geq \sum_{j=3}^{n} x_j^2, \; x_1, x_2 \geq 0 \right\}$$

  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 cone whenever we like.

Fractional quadratical function

Consider the set

$$(s, t, x) \in \mathcal{K}_r$$

this implies

$$2st \geq \|x\|^2 \text{ and } s, t \geq 0$$

or

$$2s \geq \frac{\|x\|^2}{t} \text{ and } s, t \geq 0.$$

A branch of a hyperbola

Consider the set

$$(s, t, \sqrt{2}) \in \mathcal{K}_r$$

this implies

$$2st \geq 2 \text{ and } s, t \geq 0$$

and hence (sloppy!)

$$s \geq \frac{1}{t}, \quad t \geq 0.$$

The function

$$f(x) = x^T H x$$

is convex if and only if

$$H$$

is positive semidefinite.

Lemma

The following statements are equivalent.

- H *is positive semidefinite.*
- $\exists L : \quad H = LL^T$ *e.g.* L *is a Cholesky factor.*
- $\lambda_{\min}(H) \geq 0.$

The quadratic optimization problem

$$\begin{array}{lrcl} \text{minimize} & 0.5x^T H^0 x + c^T x & & \\ \text{subject to} & 0.5x^T H^i x + a_{i:} x & \leq & b_i, \forall i = 1, 2, \ldots \end{array}$$

is convex if and only if

$$\exists Q^i : \quad H^i = Q^i (Q^i)^T$$

An equivalent reformulation:

$$\begin{array}{lrcl} \text{minimize} & 0.5 \left\| (Q^0)^T x \right\|^2 + c^T x & & \\ \text{subject to} & 0.5 \left\| (Q^i)^T x \right\|^2 + a_{i:} x & \leq & b_i, \forall i = 1, 2, \ldots. \end{array}$$

Separable quadratic reformulation:

$$\begin{array}{lrcl} \text{minimize} & 0.5 \left\| y^0 \right\|^2 + c^T x & & \\ \text{subject to} & 0.5 \left\| y^i \right\|^2 + a_{i:} x & \leq & b_i, \forall i = 1, 2, \ldots \\ & (Q^i)^T x - y^i & = & 0, \forall i = 0, 2, \ldots \end{array}$$

CQ reformulation

$$
\begin{array}{llcll}
\text{minimize} & c^T x + t_0 & & & \\
\text{subject to} & t_i + a_{i:} x & = & b_i, & \forall i = 1, 2, \ldots, \\
& (Q^i)^T x - y^i & = & 0, & \forall i = 0, 1, \ldots, \\
& z_i & = & 1, & \forall i = 0, 1, \ldots, \\
& \left\| y^i \right\|^2 & \leq & 2 t_i z_i, & \forall i = 0, 1, \ldots.
\end{array}
$$

because

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

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.
- See discussion in [2].

Ellipsoidal sets

The set

$$\mathcal{E} = \{x \in \mathbb{R}^n \mid \|P(x-c)\|_2 \leq 1\}$$

describes an ellipsoid centered at c.
Has a natural conic quadratic representation, i.e., $x \in \mathcal{E}$ if and only if

$$y = P(x-c), \quad (t, y) \in \mathcal{K}_q^{n+1}, \quad t = 1.$$

Suppose P is nonsingular then

$$\mathcal{E} = \{x \in \mathbb{R}^n \mid x = P^{-1}y + c, \ \|y\|_2 \leq 1\}.$$

is an alternatively characterization

$$\begin{array}{ll} \text{minimize} & \sum_k \left\| x^k \right\| \\ \text{subject to} & \sum_k A^k x^k \;=\; b, \end{array}$$

CQ reformulation

$$\begin{array}{lll} \text{minimize} & \sum_k t_k & \\ \text{subject to} & \sum_k A^k x^k & = \; b, \\ & \begin{bmatrix} t_k \\ x^k \end{bmatrix} \in \mathcal{K}_q. & \end{array}$$

Applications:

- Image denoising.
- Location problems.

The Fermat-Weber problem

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

$$\text{minimize} \sum_k w_k \left\| x - d^k \right\|$$

i.e. the total weighted distance to the costumers are minimized.

- Wikpedia: `https://en.wikipedia.org/wiki/Weber_problem`.

Image restoration

- A $n \times n$ image is represented by $n \times n$ matrix.
- An observed image:

$$F \in \mathbb{R}^{n \times n}.$$

- Original unknown image:

$$U \in \mathbb{R}^{n \times n}$$

- Noise in the image:

$$V \in \mathbb{R}^{n \times n}$$

We have

$$U + V = F.$$

Problem:

- How to estimate V?

The total variation model:

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

where

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

and σ is a user specified constant. Usually chosen related to amount of expected amount of noise.
See [5] for more details.

The problem:

$$\text{minimize} \quad \text{maximize}_i \left\| A^i x^i + b^i \right\|$$

CQ reformulation

$$\begin{array}{ll} \text{minimize} & v \\ \text{subject to} & \begin{bmatrix} v \\ A^i x^i + b^i \end{bmatrix} \in \mathcal{K}_q, \quad \forall i \end{array}$$

because that implies

$$v \geq \left\| A^i x^i + b^i \right\|, \forall i.$$

Lemma

The following five propositions are true.

i) $\left\{(t,x) \mid t \geq \frac{1}{x},\ x \geq 0\right\} = \left\{(t,x) \mid (x,t,\sqrt{2}) \in \mathcal{K}_r^3\right\}$.

ii) $\left\{(t,x) \mid t \geq x^{3/2},\ x \geq 0\right\} = \left\{(t,x) \mid (s,t,x), (x,1/8,s) \in \mathcal{K}_r^3\right\}$.

iii) $\left\{(t,x) \mid t \geq x^{5/3},\ x \geq 0\right\} = \left\{(t,x) \mid (s,t,x), (1/8,z,s), (s,x,z) \in \mathcal{K}_r^3\right\}$.

iv) $\left\{(t,x) \mid t \geq \frac{1}{x^2},\ x \geq 0\right\} = \left\{(t,x) \mid (t,1/2,s), (x,s,\sqrt{2}) \in \mathcal{K}_r^3\right\}$.

v) $\left\{(t,x,y) \mid t \geq \frac{|x|^3}{y^2}, y \geq 0\right\} = \left\{(t,x,y) \mid (z,x) \in \mathcal{K}_q^2, (\frac{y}{2},s,z), (\frac{t}{2},z,s) \in \mathcal{K}_r^3\right\}$.

Proof of i)

$$
\begin{aligned}
& (x, t, \sqrt{2}) \in \mathcal{K}_r^3 \\
\Leftrightarrow \quad & 2xt \geq 2,\ x, t \geq 0 \\
\Leftrightarrow \quad & t \geq \frac{1}{x},\ x \geq 0.
\end{aligned}
$$

Proof of ii)

$$
\begin{aligned}
& (s,t,x), (x, \frac{1}{8}, s) \in \mathcal{K}_r^3 \\
\Leftrightarrow \quad & 2st \geq x^2,\ \frac{1}{4}x \geq s^2,\ s,t,x \geq 0 \\
\Leftrightarrow \quad & \sqrt{x}t \geq x^2,\ t,x \geq 0 \\
\Leftrightarrow \quad & t \geq x^{3/2},\ x \geq 0.
\end{aligned}
$$

Proof of iii)

$$
\begin{aligned}
& (s,t,x),(1/8,z,s),(s,x,z) \in \mathcal{K}_r^3 \\
\Leftrightarrow\quad & \frac{1}{4}z \geq s^2,\ 2sx \geq z^2,\ 2st \geq x^2,\ s,t,x \geq 0 \\
\Leftrightarrow\quad & 2sx \geq (4s^2)^2,\ 2st \geq x^2,\ s,t,x \geq 0 \\
\Leftrightarrow\quad & x \geq 8s^3,\ 2st \geq x^2,\ x,s \geq 0 \\
\Leftrightarrow\quad & x^{1/3}t \geq x^2,\ x \geq 0 \\
\Leftrightarrow\quad & t \geq x^{5/3},\ x \geq 0.
\end{aligned}
$$

Proof of iv)

$$
\begin{aligned}
& (t, \frac{1}{2}, s), (x, s, \sqrt{2}) \in \mathcal{K}_r^3 \\
\Leftrightarrow \quad & t \geq s^2,\ 2xs \geq 2,\ t, x \geq 0 \\
\Leftrightarrow \quad & t \geq \frac{1}{x^2}.
\end{aligned}
$$

Proof of v)

$$
\begin{aligned}
& (z,x) \in \mathcal{K}_q^2, (\frac{y}{2}, s, z), (\frac{t}{2}, z, s) \in \mathcal{K}_r^3 \\
\Leftrightarrow \quad & z \geq |x|,\ ys \geq z^2,\ zt \geq s^2,\ z, y, s, t \geq 0 \\
\Leftrightarrow \quad & z \geq |x|,\ zt \geq \frac{z^4}{y^2},\ z, y, t \geq 0 \\
\Leftrightarrow \quad & t \geq \frac{|x|^3}{y^2},\ y \geq 0.
\end{aligned}
$$

The geometric mean

The geometric mean is defined by

$$\sqrt[n]{\prod_{j=1}^{n} x_j}$$

Is the hypograph of the geometric mean i.e.

$$\mathcal{G}^n = \{(t,x) \mid \quad (\prod_{j=1}^{n} x_j)^{\frac{1}{n}} \geq t, \quad x \geq 0\}$$

conic quadratic representable?

Lemma

For $l = 1, 2, \ldots$ *and* $n = 2^l$ *and* $g \in \mathbb{R}_+^{2n-1}$. *Given*

$$(g_{2i}, g_{2i+1}, g_i) \in \mathcal{K}_r, \text{ for } i = 1, \ldots, n-1,$$

then

$$\sqrt{n}^n \prod_{i=n}^{2n-1} g_i \geq g_1^n$$

A fact:

$$\sum_{i=0}^{l} 2^i = 2n - 1.$$

Proof of lemma

We will prove the lemma using induction on l.
For $l = 1$ we have

$$2g_2g_3 \geq g_1^2$$

which is correct. Now assume the lemma is true for l i.e.

$$\sqrt{2^l}^{2^l} \prod_{i=2^l}^{2(2^l)-1} g_i \geq g_1^{2^l}.$$

For $l + 1$ is holds

$$(g_{2i}, g_{2i+1}, g_i) \in \mathcal{K}_r, \text{ for } i = 1, \ldots, 2^{l+1} - 1.$$

This implies

$$\prod_{i=2^l}^{2^{l+1}-1} \sqrt{2g_{2i}g_{2i+1}} \geq \prod_{i=2^l}^{2^{l+1}-1} g_i$$

or

$$\sqrt{2}^{2^l} \prod_{i=2^{l+1}}^{2(2^{l+1})-1} \sqrt{g_i} \geq \prod_{i=2^l}^{2^{l+1}-1} g_i.$$

Therefore,

$$\sqrt{2^l}^{2^l} \sqrt{2}^{2^l} \prod_{i=2^{l+1}}^{2(2^{l+1})-1} \sqrt{g_i} \geq g_1^{2^l}$$

and taken squares on both sides leads to the conclusion

$$\sqrt{2^{l+1}}^{l+1} \prod_{i=2^{l+1}}^{2(2^{l+1})-1} g_i \geq g_1^{2^{l+1}}$$

because

$$\begin{aligned}\left(\sqrt{2^l}^{2^l}\sqrt{2}^{2^l}\right)^2 &= 2^{l2^l+2^l} \\ &= 2^{(l+1)(0.5)2^{l+1}} \\ &= \sqrt{2^{l+1}}^{2^{l+1}}.\end{aligned}$$

Therefore

$$
\begin{aligned}
& \mathcal{G}^{2^l} \\
= \; & \{(x,t) \in \mathbb{R}_+^{2^l+1} | \quad (\prod_{j=1}^{n} x_j)^{\frac{1}{n}} \geq t\} \\
= \; & \{(x,t) \in R_+^{2^l+1} | \quad (g_{2i}, g_{2i+1}, g_i) \in \mathcal{K}_r, \text{ for } i = 1, \ldots, 2^l - 1, \\
& \qquad\qquad\qquad\quad g_{2^l-1+i} = x_i, \text{ for } i = 1, \ldots, 2^l, \\
& \qquad\qquad\qquad\quad g_1 = \sqrt{2^l} t\}
\end{aligned}
$$

- $n = 2^l$ quadratic cones is needed represent the geometric mean.
- What if n is uneven.

What if n is not a power of 2

Let us assume $n = 6$. We then wish to characterize the set

$$t \leq (\prod_{j=1}^{6} x_j)^{1/6}$$

which is equivalent to

$$t \leq (\prod_{j=1}^{8} x_j)^{1/8}, \quad x_7 = x_8 = t, \quad x \geq 0.$$

Now use the result for $\mathcal{G}^8$.
Thus, if n is not a power of two, we take $l = \lceil \log_2 n \rceil$ and build the conic quadratic quadratic representation for that set, and we add $2^l - n$ simple equality constraints.

Lemma

The set

$$t \geq x^{\frac{p}{q}}, \quad x \geq 0$$

is conic quadratic representable where p and q are integers such $p \geq q \geq 1$.

Proof.

Let

$$\begin{array}{rcl} 0 \leq & x & \leq \left(\prod_{j=1}^{p} y_j\right)^{\frac{1}{p}} \\ & t & = \; y_j \qquad \text{for } j = 1, \ldots, q, \\ & 1 & = \; y_j \qquad \text{for } j = q+1, \ldots, p. \end{array}$$

and it follows

$$0 \leq x^p \leq t^q.$$

However, the set

$$\begin{array}{rcll} (x, y) & \in & \mathcal{G}^n & \\ t & = & y_j & \text{for } j = 1, \ldots, q, \\ 1 & = & y_j & \text{for } j = q+1, \ldots, p \end{array}$$

is conic quadratic representable. $\square$

Lemma

$$\|x\|_p \leq t$$

is CQ representable for a integer $p \geq 1$*.*

Lemma

The set

$$t \geq x^{\frac{-p}{q}}, \quad x \geq 0$$

where p *and* q *are nonnegative integers is CQ representable.*

For more examples CQ representable sets see [7].

Section 4

Duality

The linear optimization problem

$$\begin{array}{lrcl} \text{minimize} & c^T x & & \\ \text{subject to} & Ax & = & b, \\ & x \geq 0 & & \end{array} \tag{1}$$

has the dual problem

$$\begin{array}{lrcl} \text{maximize} & b^T y & & \\ \text{subject to} & A^T y + s & = & c, \\ & s \geq 0. & & \end{array} \tag{2}$$

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

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

- 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.

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

$$\begin{array}{lrcl} \text{minimize} & \sum_k (c^k)^T x^k & & \\ \text{subject to} & \sum_k A^k x^k & = & b, \\ & x^k \in \mathcal{K}^k & & \end{array} \tag{3}$$

then the corresponding dual problem is

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

Observe the dual cone corresponding to the linear cone

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

is

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

- 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.

- Weak duality holds:

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

- The (rotated) quadratic cone is self-dual.
- Strong duality and the other relations holds **ALMOST** in the conic case.
- To be continued.

Section 5

Summary

Extremely disciplined modeling

- A conic optimization model.
- Restricted to a limited set of cone types.
- Advantages:
  - Convex by construction.
  - Explicit structure.
  - Much more general than linear only.
  - Behaves in most aspects as the linear case.

A recap

- Introduced conic optimization.
- The quadratic cone has been introduced.
- Leads to **extremely disciplined modeling**.
- Some applications has be shown.
- Amazing how general the quadratic cone is!
- Background material:
  - Primary: [8, 7].
  - Secondary: [1, 3, 4, 6].

References I

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

[2] Erling D. Andersen.
On formulating quadratic functions in optimization models.
Technical Report TR-1-2013, MOSEK ApS, 2013.
Last revised 23-feb-2016.

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

[4] S. Boyd and L. Vandenberghe.
Convex optimization.
Cambridge University Press, 2004.
http://www.stanford.edu/~boyd/cvxbook/.

[5] 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.

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

[7] MOSEK ApS.
The MOSEK Modeling Cookbook.
MOSEK ApS, Fruebjergvej 3, Boks 16, 2100 Copenhagen O, 2012.
Last revised September 2015.

[8] A. Nemirovski.
Advances in convex optimization: Conic programming.
In Marta Sanz-Sol, Javier Soria, Juan L. Varona, and Joan Verdera, editors, *Proceedings of International Congress of Mathematicians, Madrid, August 22-30, 2006, Volume 1*, pages 413–444. EMS - European Mathematical Society Publishing House, April 2007.

Section 6

Exercises

1: Quadratic over linear

1. Prove that the function

$$f(x,t) = \frac{\|x\|^2}{t}$$

is convex on its domain $(t > 0)$.