Power and Exponential cones
$x^\alpha y^{1-\alpha} \geq |z|$ and $x \geq y e^{z/y}$

June 27, 2018

Ulf Worsøe
MOSEK ApS

www.mosek.com

Who is MOSEK?

- We develop and sell a software package for large scale linear and conic optimization.
- MOSEK v1.0 was released in 1999, v9.0 expected this fall.
- Mainly located in Copenhagen, employing 9 people.

And who am I?

- Employed since 2001
- Work mainly with API ports, MOSEK Fusion (modelling interface) and internal systems.
- Developed most of the Julia/MOSEK interface

Why?

The major changes in the upcoming MOSEK 9.0:

- Remove the API for General Convex optimization
- Add support for the Power cone and the Exponential cone

And a recent major change in JuMP/MathOptInterface:

- Support for constraints on set-form $Ax - b \in C$

I will mainly be talking about the modeling aspects.

MOSEK 8.1 and 9.0

MOSEK 8.1 supports

- Linear inequalities and equalities,
- Second order cone $\mathcal{C}^n = \left\{x \in \mathbb{R}^n | x_1^2 \geq x_2^2 \cdots x_n^2,\ x_1 > 0\right\}$,
- Rotated second order cone $\mathcal{C}_r^n = \left\{x \in \mathbb{R}^n | 2x_1 x_2 \geq x_3^2 \cdots x_n^2,\ x_1, x_2 > 0\right\}$,
- Cone of symmetric positive semidefinite $\mathcal{S}_+^n$ matrixes of dimension $n > 1$.

MOSEK 9.0 will additionally support

- Power cone $\mathcal{P}_\alpha = \left\{(x, y, z) \in \mathbb{R}^3 | x^\alpha y^{1-\alpha} \geq |z|,\ x, y > 0\right\}$ for $0 < \alpha < 1$
- Exponential cone $\mathcal{K}_e = \left\{(x, y, z) \in \mathbb{R}^3 | x \geq y e^{z/y},\ x, y > 0\right\}$

So, what superpowers do these cones give us?

Not a lot, really:

- The power inequality $x > y^a$ and exponential inequality $x > e^y$ have been solvable with general convex methods (e.g. MOSEK)
- The $x > y^a$ for $a = p/q$, $p, q \in \mathbb{N}$ can be modeled with quadratic cones

However,

- The conic framework allows us to mix power, exponential, quadratic and semidefinite cones *and* guarantee convexity.
- There is a stronger theoretical foundation for conic interior-point methods (even if it is weaker for non-self-dual cones)
- The conic methods seem to give increased solver stability

Basic modeling ideas

The conic sets are often not directly useful, but they can be combined to represent complex sets.

We use mainly three constructions:

- Variables fixing, e.g. $(x, 1/2, z) \in \mathcal{Q}_r^3$ meaning

$$(x, y, z) \in \mathcal{Q}_r^3,\ y = 1/2\ \Rightarrow\ x > z^2$$

- Chaining or intersecting cones, e.g.

$$(x, y, z), (z, 1/8, x) \in \mathcal{Q}_r^3\ \Rightarrow\ y > z^{3/2}$$

- Linear transformation of cones, for example the rotated quadratic cone can be written in terms of the quadratic cone:

$$\mathcal{Q}_r^n = \left\{x \in \mathbb{R}^n | (x_1 + x_2, x_1, ..., x_n) \in \mathcal{Q}^{n+1}\right\}$$

Power cone

The Power cone is defined for $0 < \alpha < 1$:

$$\mathcal{P}_\alpha = \left\{ (x, y, z) \in \mathbb{R} | x^\alpha y^{1-\alpha} > |z|,\ x, y > 0 \right\}$$

This is a (scaled) generalization of the rotated quadratic cone:

$$(x, y, z) \in \mathcal{P}_{1/2} \ \Leftrightarrow \ (x, y, \sqrt{2}z) \in \mathcal{Q}_r^3$$

Power cone - basic power inequalities

Simple convex power inequalities for the ranges

$$\alpha < -1, \ -1 < \alpha < 0, \ 0 < \alpha < 1, \text{and } 1 < \alpha$$

- $x^\alpha > |z|$ for $0 < \alpha < 1$: $(x, 1, z) \in \mathcal{P}_\alpha$,
- $x > |z|^\alpha$ for $1 < \alpha$: $(x, 1, z) \in \mathcal{P}_{1/\alpha}$,
- $x^\alpha < z$ for $-1 < \alpha < 0$: $(x, 1, u) \in \mathcal{P}_{-\alpha}$, $(u, z, 1/\sqrt{2}) \in \mathcal{Q}_r$,
- $|z|^\alpha < x$ for $\alpha < -1$: $(u, 1, z) \in \mathcal{P}_{-1/\alpha}$, $(u, x, \sqrt{2}) \in \mathcal{Q}_r$

Inequalities for $\alpha \in \{-1, 0, 1\}$ do not requre the power cone.

Power cone - basic power constructions

Example: How we obtained the last power inequality for $\alpha < -1$

$$|z|^\alpha < x$$

$$\begin{aligned}
& (u, 1, z) \in \mathcal{P}_{-1/\alpha},\ (u, x, \sqrt{2}) \in \mathcal{Q}_r \\
\Leftrightarrow \quad & u^{-1/\alpha} \cdot 1^{1+1/\alpha} \geq |z|,\ 2ux \geq (\sqrt{2})^2,\ x, u > 0 \\
\Rightarrow \quad & 1/u \geq |z|^\alpha,\ x \geq 1/u,\ u > 0 \\
\Rightarrow \quad & x \geq |z|^\alpha
\end{aligned}$$

Power cone - Geometric mean inequality

We can model:

$$\left\{(y,x) \in \mathbb{R}^{n+1} | y < \prod_{i=1}^{n} x_i^{1/n},\ y > 0\right\}$$

We can split into two inequalities:

$$\begin{aligned} & \{(y,x) \in \mathbb{R}^{n+1} | y < \prod_{i=1}^{n} x_i^{1/n},\ y > 0\} \\ \Leftrightarrow \quad & (y,x,t) \in \mathbb{R}^{n+2} | y < x_1^{1/n} t^{1-1/n},\ t < \prod_{i=1}^{n-1} x_i^{1/(n-1)},\ y,t > 0 \end{aligned}$$

And by induction we can rewrite the whole inequation into tri-graph power inequalities.

Use case: Portfolio model with market impact

Portfolio optimization with market impact term:

$$\begin{aligned} \text{maximize} \quad & \mu^t x - \sum_{i=1}^{n} \delta_i x_i^{\beta} \\ \text{such that} \quad & x^t Q x \leq \gamma^2 \\ & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0 \end{aligned}$$

Where $\beta > 1$ is the market impact. We can rewrite the objective

$$\mu^t x - \sum_{i=1}^{n} \delta_i z_i, \;\; z_i^{1/\beta} x_i \text{ for } i = 1 \ldots n$$

Exponential cone - basic inequalities

$$\mathcal{K}_e = \left\{(x, y, z) \in \mathbb{R} | x > y e^{z/y},\ y > 0\right\}$$

- Exponential inequality, $x > e^z$: $(x, 1, z) \in \mathcal{K}_e$
- Logarithm inequality, $z < \log(x)$: $(x, 1, z) \in \mathcal{K}_e$
- $t > a_1^{x_1} \cdots a_n^{x_n}$: $(t, 1, \sum_i x_i \log(a_i))$ for positive a_i, arising from

$$t > \exp(\log(a_1^{x_1} \cdots a_n^{x_n})) = \exp\left(\sum_i x_i \log a_i\right)$$

Exponential cone - monomials

We define a monomial for $c > 0, \ a_i \in \mathbb{R}$ as

$$\hat{f}(x) : \mathbb{R}^n_+ \to \mathbb{R} = c x_1^{a_1} \cdots x_n^{a_n}$$

Making a variable substitution with $x_i = e^{y_i}$ we get

$$f(y) : \mathbb{R}^n \to \mathbb{R} = \hat{f}(e^y) = e^{\log c + a^t y}$$

The inequality $f(y) < t$ can be formulated as

$$(t, 1, \log c + a^t y) \in \mathcal{K}_e$$

Note that the original x cannot be mixed with y in the problem, but its solution value can be obtained from the solution value of y.

Exponential cone - Geometric Problem

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1\ldots p_0} \hat{f}_{0,k}(x) \\ \text{such that} \quad & \sum_{k=1\ldots p_i} \hat{f}_{i,k}(x) \leq 1, \text{ for } i = 1\ldots m \\ & x_i > 0 \end{aligned}$$

Substitution $x_j = e^{y_j}$ and skipping forward a few steps we end up with a conic formulation

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1\ldots p_0} u_{0,k} \\ \text{such that} \quad & \sum_{k=1\ldots p_i} u_{i,k} \leq 1, \text{ for } i = 1\ldots m \\ & (u_{i,k}, 1, a_{i,k}^t y + \log c_{i,k}) \in \mathcal{K}_e, \text{ for } i = 0\ldots m, \ k = 1\ldots p_i \end{aligned}$$

Exponential cone - Geometric Problem

$$\begin{array}{ll} \text{minimize} & \sum_{k=1\ldots p_0} \hat{f}_{0,k}(x) \\ \text{such that} & \sum_{k=1\ldots p_i} \hat{f}_{i,k}(x) \leq 1, \text{ for } i = 1 \ldots m \\ & x_i > 0 \end{array}$$

Substitution $x_j = e^{y_j}$ and skipping forward a few steps we end up with a conic formulation

$$\begin{array}{ll} \text{minimize} & \sum_{k=1\ldots p_0} u_{0,k} \\ \text{such that} & \sum_{k=1\ldots p_i} u_{i,k} \leq 1, \text{ for } i = 1 \ldots m \\ & (u_{i,k}, 1, a_{i,k}^t y + \log c_{i,k}) \in \mathcal{K}_e, \text{ for } i = 0 \ldots m, \; k = 1 \ldots p_i \end{array}$$

Exponential cone - Use case: Balanced portfolio

Risk-minimizing Markowitz portfolio model

$$\begin{aligned} \text{minimize} \quad & \sqrt{x^y Q x} \\ \text{such that} \quad & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0,\ i = 1 \ldots n \end{aligned}$$

The coviariance matrix Q estimates the risk of assets. Problem: The portfolio may end up being very unbalanced — we wish to add a penalty for having very small positions:

$$\begin{aligned} \text{minimize} \quad & \sqrt{x^y Q x} + c \sum_{i=1}^{n} \log x_i \\ \text{such that} \quad & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0,\ i = 1 \ldots n \end{aligned}$$

Exponential cone - A few other use cases

- Geometric Programming allows a long range of problems in engineering and electronics.
- Entropy function maximization $H(x) = -x \log x$ as max $t : \ (1, t, x) \in \mathcal{K}_e$
- Logistic regression
- Many, many more — that we don't know yet!

Conclusions - being a bit insubstantial here

- Power cone - currently:
  - Can be approximated and solved using SOCP, but it is *complex*
  - We can not currenly conclude whether the Power Cone is more efficient
  - Simpler infeasibility certificates and dual solutions
- Exponential cone
  - This replaces the General Convex formulation
  - Allows mixing of SOCP and SDP with exponential terms
  - Simpler infeasibility certificates and dual solutions
  - Possibly yields more stable solve times

See also:

- MOSEK Cookbook:
  https://docs.mosek.com/modeling-cookbook/index.html
- General MOSEK documenation at
  https://www.mosek.com/documentation/
- Tutorials at Github:
  https://github.com/MOSEK/Tutorials
- “A Tutorial on Geometric Programming”, S. Boyd et al., 2007.