A gentle introduction to Continuous and Mixed-Integer Conic Programming

Università di Bologna
November 23th, 2020

Sven Wiese

www.mosek.com

Overview

1. From Linear to Conic Programming
2. Examples & Applications of cones
   - Symmetric cones
   - Non-symmetric cones
   - Yet more cones
3. Conic duality
4. Numerical solution methods
   - The continuous case
   - The mixed-integer case

1. From Linear to Conic Programming

Linear Programming

Linear Optimization in standard form:

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

Pro:

- Structure is explicit and simple.
- Data is simple: c, A, b.
- Structure implies data-independent convexity.
- Powerful duality theory including Farkas lemma.

Therefore, we have powerful algorithms and software.

Con:

- It is linear only.

Nonlinear Programming

The classical nonlinear optimization problem:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \leq 0. \end{array}$$

Pro

- It is very general.

Con:

- Structure is less explicit.
  - How to specify the problem at all in software?
  - How to compute gradients and Hessians if needed?
  - How to exploit structure?
- Smoothness?
- Verifying convexity is NP-hard!

A fundamental question

Is there a class of nonlinear optimization problems that preserves possibly many of the good properties of Linear Programming?

Good partial orderings

Definition (Ben-Tal & Nemirovski, (2001))

A "good" partial ordering of $\mathbb{R}^n$ is a vector relation that satisfies:

1. *reflexivity*
2. *antisymmetry*
3. *transitivity*
4. *compatibility with linear operations*

The coordinatewise ordering

$$x \geq y \iff x_i \geq y_i \ \forall i = 1, \ldots, n$$

is an example, but not the only one!

Good partial orderings

- If for some good partial ordering "$\succeq$" we define

$$\mathcal{K} := \{a \in \mathbb{R}^n \mid a \succeq 0\},$$

then $\mathcal{K}$ must be a pointed, convex cone:

1. $a, a' \in \mathcal{K} \implies a + a' \in \mathcal{K}$
2. $a \in \mathcal{K}, \lambda \geq 0 \implies \lambda a \in \mathcal{K}$
3. $a \in \mathcal{K}$ and $-a \in \mathcal{K} \implies a = 0$

- Conversely, if $\mathcal{K}$ is a non-empty pointed convex cone, then $x \succeq_{\mathcal{K}} y :\Longleftrightarrow x - y \in \mathcal{K}$ defines a good partial ordering.

The cone $\mathbb{R}^n_+$ is also closed and has a non-empty interior.

(Mixed-Integer) Conic Programming

We thus consider problems of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \cap (\mathbb{Z}^p \times \mathbb{R}^n) \end{array}$$

where $\mathcal{K}$ is a (closed) pointed convex cone (with non-empty interior).

- (MI)LP is a special case!
- Typically, $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2 \times \cdots \times \mathcal{K}_K$ is a product of lower-dimensional cones.
- A conic building block $\mathcal{K}_i$ can be thought of as encoding some type of specific non-linearity.

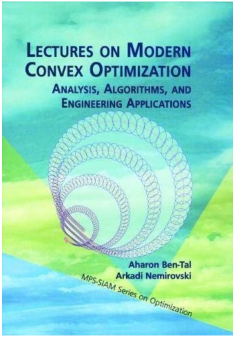
The beauty of conic optimization

- Separation of data and structure:
  - Data: c, A and b.
  - Structure: $\mathcal{K}$.

- No issues with smoothness and differentiability.
- Structural convexity.
- Duality (almost...).

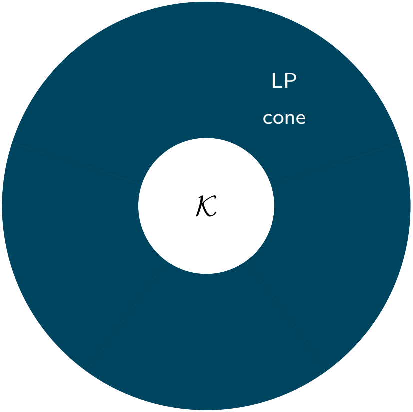
References I

- A. Ben-Tal & A. Nemirovski: *Lectures on Modern Convex Optimization* (2001).



2. Examples & Applications of cones

The conic wheel



Quadratic cones

After the non-negative orthant $\mathbb{R}^n_+$, the quadratic-cone family is arguably most prominent.

- the quadratic cone

$$\mathcal{Q}^n = \{x \in \mathbb{R}^n \mid x_1 \geq \left(x_2^2 + \cdots + x_n^2\right)^{1/2} = \|x_{2:n}\|_2\},$$

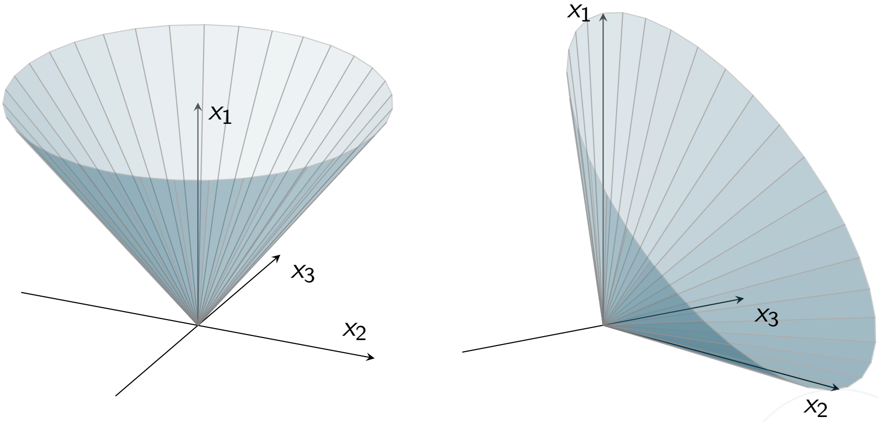
- the rotated quadratic cone

$$\mathcal{Q}_r^n = \{x \in \mathbb{R}^n \mid 2x_1x_2 \geq x_3^2 + \cdots + x_n^2 = \|x_{3:n}\|_2^2,\ x_1, x_2 \geq 0\}.$$

Are equivalent in the sense that $x \in \mathcal{Q}^n \iff T_n x \in \mathcal{Q}_r^n$ with

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & I_{n-2} \end{pmatrix}.$$

Quadratic cones in dimension 3



Conic quadratic case study: least squares regression

In least squares regression we use the penalty function

$$\phi(r) = \|r\|_2.$$

In its simplest form, given observations $y \in \mathbb{R}^n$ and features $X \in \mathbb{R}^{n \times d}$, it solves

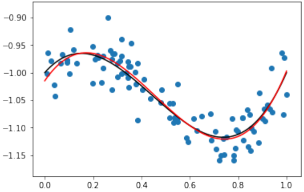
$$\min_{w \in \mathbb{R}^d} \|y - Xw\|_2.$$

Start with a small and simple amount of reformulation:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & t \geq \|y - Xw\|_2 \\ & t \in \mathbb{R}, w \in \mathbb{R}^d \end{array}$$

In the conic framework this would be written as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & s = y - Xw \\ & (t, s) \in \mathcal{Q}^{n+1} \\ & w \in \mathbb{R}^d. \end{array}$$



We usually use the more compact notation

$$(t, y - Xw) \in \mathcal{Q}^{n+1}.$$

More conic quadratic modeling

- Second-order cone inequality:

$$c^T x + d \geq \|Ax + b\|_2 \quad \Longleftrightarrow \quad (c^T x + d, Ax + b) \in \mathcal{Q}^{m+1}.$$

- Squared Euclidean norm:

$$t \geq \|x\|_2^2 \quad \Longleftrightarrow \quad (t, 1/2, x) \in \mathcal{Q}_r^{n+2}.$$

- Convex quadratic inequality:

$$t \geq (1/2)x^T Q x \quad \Longleftrightarrow \quad (t, 1, F^T x) \in \mathcal{Q}_r^{k+2}$$

with $Q = F^T F$, $F \in \mathbb{R}^{n \times k}$.

Any convex (MI)QCQP can be cast in conic form!

More conic quadratic modeling

- Square roots, convex hyperbolic function, some convex negative rational powers...

- Convex positive rational power

$$t \geq x^{3/2},\ x \geq 0 :$$

If we impose

$$(s, t, x), (x, 1/8, s) \in \mathcal{Q}_r^3 \iff 2st \geq x^2, 2x \cdot \frac{1}{8} \geq s^2,$$

it follows that

$$4s^2t^2 \cdot \frac{1}{4}x \geq x^4 s^2 \implies t^2 \geq x^3 \implies t \geq x^{3/2}.$$

The positive semidefinite cone

The positive semidefinite cone can be defined as a subspace of the vector space $\mathbb{R}^{n(n+1)/2}$

$$\mathcal{S}^{n(n+1)/2} := \{x \in \mathbb{R}^{n(n+1)/2} \mid z^T \mathbf{smat}(x) z \geq 0, \ \forall z \in \mathbb{R}^n\},$$

with

$$\mathbf{smat}(x) := \begin{pmatrix} x_1 & x_2/\sqrt{2} & \dots & x_n/\sqrt{2} \\ x_2/\sqrt{2} & x_{n+1} & \dots & x_{2n-1}/\sqrt{2} \\ \vdots & \vdots & & \vdots \\ x_n/\sqrt{2} & x_{2n-1}/\sqrt{2} & \dots & x_{n(n+1)/2} \end{pmatrix}.$$

An equivalent definition via matrix variables:

$$X \in \mathbb{S}^n_+ :\Longleftrightarrow X \in \mathbb{S}^n \text{ and } z^T X z \geq 0 \ \forall z \in \mathbb{R}^n.$$

X is mapped to $\mathcal{S}^{n(n+1)/2}$ via

$$\mathbf{svec}(X) := (X_{11}, \sqrt{2}X_{21}, \dots, \sqrt{2}X_{n1}, X_{22}, \sqrt{2}X_{32}, \dots, X_{nn})^T.$$

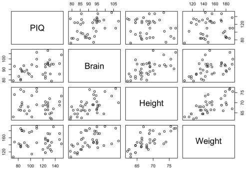
SDP use-case: Nearest correlation matrix

Let $A \in \mathbb{S}^n$ and assume we want to find its nearest correlation matrix

$$X^* \in \mathcal{C} := \{X \in \mathbb{S}^n_+ \mid X_{ii} = 1 \; \forall i = 1, \ldots, n\},$$

i.e.,

$$X^* = \min_{X \in \mathcal{C}} \|A - X\|_F.$$



A conic formulation in vector space is given by

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & x_1 = x_{n+1} = x_{2n} = \ldots = x_{n(n+1)/2} = 1 \\ & (t, \mathbf{svec}(A) - x) \in \mathcal{Q}^{n(n+1)/2+1} \\ & x \in \mathcal{S}^{n(n+1)/2}. \end{array}$$

More semidefinite modeling

- SDP can come in handy in eigenvalue optimization, e.g., if

$$tI - X \succeq_{\mathbb{S}^n_+} 0,$$

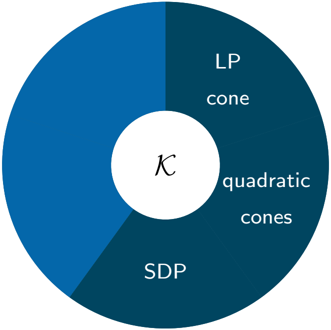
then t is an upper bound on the largest eigenvalue of X.

- SDP-relaxations play a role in Quadratic Programming and in Combinatorial Optimization:

$$X = xx^T \text{ can be relaxed to } X - xx^T \succeq_{\mathbb{S}^n_+} 0.$$

- There are applications for Mixed-Integer SDP, see, e.g., *Gally, Pfetsch and Ulbrich* (2018).

Back to the conic wheel



The three cones we have seen so far are so-called symmetric cones, i.e., they are

1. homogeneous
2. self-dual

The exponential cone

The exponential cone is defined as the closure of the epigraph of the perspective of the exponential function:

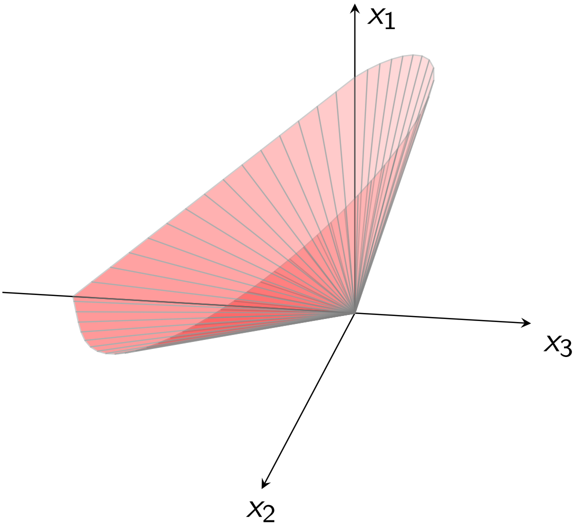
$$\mathcal{K}_{exp} := \mathrm{cl}\{x \in \mathbb{R}^3 \mid x_1 \geq x_2 \exp(x_3/x_2),\ x_2 > 0\},$$

or more explicitly

$$\mathcal{K}_{exp} = \{(x_1, x_2, x_3) \mid x_1 \geq x_2 \exp(x_3/x_2),\ x_2 > 0\} \bigcup \{(x_1, 0, x_3) \mid x_1 \geq 0, x_3 \leq 0\}.$$

The exponential cone is non-symmetric!

The exponential cone



Exponential cone use case: Geometric Programming

Consider the very simple Geometric Program

$$\begin{array}{ll} \text{minimize} & x + y^{0.3} z \\ \text{subject to} & \sqrt{x} + y^{-1} \leq 1 \\ & x, y, z > 0 \end{array}$$

First note that $e^{x_1} + \ldots + e^{x_k} \leq 1$ can be modeled as

$$(u_i, 1, x_i) \in \mathcal{K}_{exp} \ \forall i = 1, \ldots, k \text{ and } \sum_{i=1}^{k} u_i \leq 1,$$

and then substitute $x = e^p, y = e^q, z = e^r$:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & (u_1, 1, p - t), (u_2, 1, 0.3q + w - t) \in \mathcal{K}_{exp}, u_1 + u_2 \leq 1 \\ & (v_1, 1, p/2), (v_2, 1, -q) \in \mathcal{K}_{exp}, v_1 + v_2 \leq 1 \end{array}$$

More exponential cone modeling

- Logarithm:

$$\log x \geq t \quad \Longleftrightarrow \quad (x, 1, t) \in \mathcal{K}_{exp}.$$

- Entropy:

$$-x \log x \geq t \quad \Longleftrightarrow \quad (1, x, t) \in \mathcal{K}_{exp}.$$

- Relative entropy:

$$x \log(x/y) \leq t \quad \Longleftrightarrow \quad (y, x, -t) \in \mathcal{K}_{exp}.$$

- Softplus function:

$$\log(1+e^x) \leq t \quad \Longleftrightarrow \quad (u, 1, x-t), (v, 1, -t) \in \mathcal{K}_{exp},\ u+v \leq 1.$$

The power cone

The power cone is defined as

$$\mathcal{P}_n^\alpha = \{x \in \mathbb{R}^n \mid x_1^\alpha x_2^{(1-\alpha)} \geq \|x_{3:n}\|_2,\ x_1, x_2 \geq 0\},$$

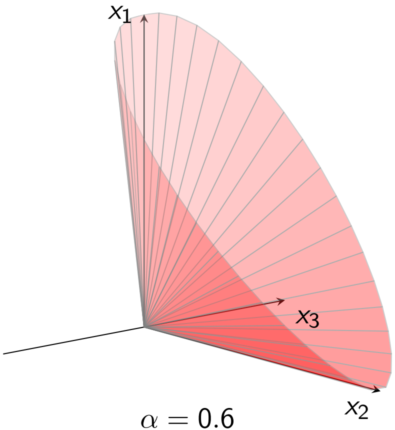
for $0 < \alpha < 1$.

One may also restrict to the three-dimensional power cone without losing any modeling capabilities:

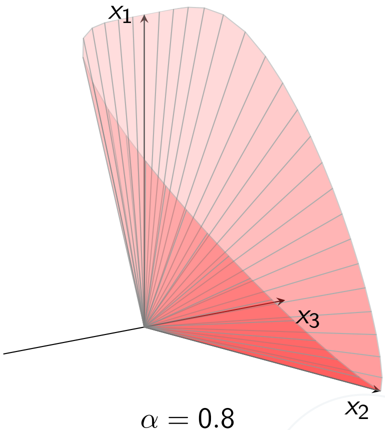
$$(x_1, \dots, x_n) \in \mathcal{P}_n^\alpha \iff (x_1, x_2, z) \in \mathcal{P}_3^\alpha, (z, x_3, \dots, x_n) \in \mathcal{Q}^{n-1}.$$

Also the power cone is non-symmetric!

The power cone



$\alpha = 0.6$



$\alpha = 0.8$

Power cone modeling

- Simple powers:

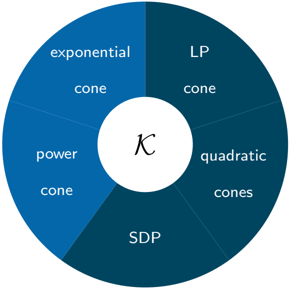
$$|t| \leq x^p, x \geq 0 \text{ with } 0 < p < 1 \iff (x, 1, t) \in \mathcal{P}_3^p.$$

$$t \geq |x|^p \text{ with } p > 1 \iff (t, 1, x) \in \mathcal{P}_3^{1/p}.$$

Example: $t \geq x^{3/2}, x \geq 0 \iff (t, 1, x) \in \mathcal{P}_3^{2/3}$ (instead of $(s, t, x), (x, 1/8, s) \in \mathcal{Q}_r^3 \ldots$)

- p-norms, geometric mean, ...

How general is the conic wheel?



Continuous Optimization Folklore

"Almost all convex constraints which arise in practice are representable using these 5 cones."

More evidence: *Lubin et al.* (2016) show that all convex instances (333) in MINLPLIB2 are conic representable using only 4 of the above cones.

Verifying convexity

Is $\log(1/(1+\exp(-x))) \leq 0$ a convex constraint?

From `ask.cvxr.com`:

B

1 ✎ Oct '18

I encountered a problem, which is was attached.

Is it a convex problem? (how prove?)

If it is not convex, how I convert it to a convex problem?

Verifying convexity can be hard!

Solution: Disciplined Convex Programming (DCP) by *Grant, Boyd and Ye* (2006): only allow for modeling operations that preserve convexity.

We call modeling with the aforementioned 5 cones

Extremely Disciplined Convex Programming.

- More strict than DCP..
- ... but leading to guaranteed convexity and conic-representability.
- Aiming at the development of efficient numerical algorithms.

Are there more cones?

For every convex function $g(x)$ the set

$$\mathcal{K} := \mathrm{cl}\{(y, s, x) \mid y \geq s \cdot g(x/s)\}$$

is a closed pointed convex cone. So

$$y \geq g(x) \iff (y, 1, x) \in \mathcal{K}.$$

But how de we handle $\mathcal{K}$ computationally, and is it tractable?

Do we need more cones?

Tuesday, June 4, 2019

Logarithmic mean temperature difference requires yet another cone?

The logarithmic mean temperature difference

$$RecLMTD^{\beta}(x,y) = \left(\frac{\ln(x/y)}{x-y}\right)^{\beta}, \tag{2}$$

can be extracted as a separately contributing term in the objective function. Capitalizing on the convexity of this term on $(x,y) \in \mathbb{R}^2_{+}$, for all considered $\beta \geq 0$, this leads to better performance when solving the otherwise nonconvex problem as argued in the paper.

A challenge to find the conic reformulation of this function was posed under the Oberwolfach Workshop on Mixed-Integer Nonlinear Optimization (2019) and we accepted. Of course, this is trivial if no restrictions are put on the set of cones as one may just define

$$\mathcal{K} = \mathrm{cl}\{(t,s,x,y) \in \mathbb{R}^4_{++} : t \geq s \cdot RecLMTD^{\beta}(x/s, y/s)\} \tag{3}$$

and call it a day. This cone is nonempty, closed and convex and hence obeys $\mathcal{K} = (\mathcal{K}^*)^*$ as well as all the usual properties of conic duality. Computationally, however, the cone is not particularly desirable and we can do better with a bit of reformulation:

$$y \geq \frac{u}{\exp(u/s)-1}, \qquad u = x - y, \quad s \geq t^{-1/\beta}, \tag{4}$$

where I substitute in the first step, rewrite assuming either $u > 0$ or $u < 0$ (both leads to the same) in the second, and extract a power cone representable subexpression in the third. This means that the representation problem of $RecLMTD^{\beta}$ have been reduced to the representation problem of

$$\mathcal{K} = \mathrm{cl}\left\{(t,s,x) \in \mathbb{R}^2_{+} \times \mathbb{R}_{++} : t \geq \frac{x}{\exp(x/s)-1}\right\}, \tag{5}$$

which, just like the quadratic, power and exponential cones, is defined as the epigraph of the perspective of a univarite convex function; in this case $\frac{x}{\exp(x)-1}$. Whether this cone can be written in terms of the others, or has potential for computationally efficient implementations itself, remains open. We invite anyone interested in barrier functions and interior-point algorithms to take a crack at it.

Posted by HFriberg at 11:21 AM

Do we need more cones?

The Extremely-DCP framework is very general, but does it have limitations?

- A folkloristic saying is not a formal theorem.
- More cones may lead to less reformulation.

Coey, Kapelevich and Vielma (2020) introduce a framework for *Generic Conic Programming*, treating more exotic cones.

In the literature, note the prominent appearance of

- the completely positive,
- the copositive
- and the doubly-non-negative cone.

Exotic cones

- The infinity norm cone

$$\mathcal{K}_{\ell_\infty} = \{x \in \mathbb{R}^n \mid x_1 \geq \|x_{2:n}\|_\infty\}$$

- The relative entropy cone

$$\mathcal{K}_{entr} = \mathrm{cl}\{(x, u, v) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \mid x \geq \sum_{i=1}^{d} u_i \log(u_i/v_i)\}$$

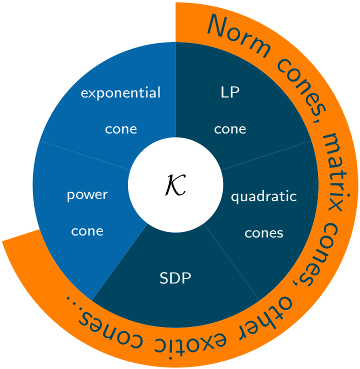
- The spectral norm cone

$$\mathcal{K}_{spec(d_1,d_2)} = \{(x, X) \in \mathbb{R} \times \mathbb{R}^{d_1 \times d_2} \mid x \geq \sigma_1(X)\}$$

- Root-determinant cone, Log-determinant cone, Polynomial weighted sum-of-squares cone, ...

For example, $(x, X) \in \mathcal{K}_{spec(d_1,d_2)} \iff \begin{pmatrix} xI_{d_1} & X \\ X^T & xI_{d_2} \end{pmatrix} \in \mathbb{S}_+^{d_1+d_2}$.

The extended conic wheel

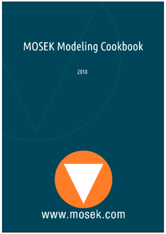


Conic software:

- **MOSEK**: LP, QCP, SDP, Exp, Pow, with MI support
- **SeDuMi**, **CSDP**, **SDPA**, **SDPD**, **SDPT3**: SDP, QCP
- **CPLEX**, **Gurobi**, **XPRESS**: (MI)-LP and -SOCP
- **SCS**: LP, QCP, SDP, Exp, Pow
- **ECOS**: QCP, Exp
- **SCIP-SDP**: MI-SDP
- `Pajarito.jl`: OA-framework for MI, -QCP, -SDP, -Exp
- `Hypatia.jl`: Generic Conic Programming
- Modeling: **CVX**, **Yalmip**, **JuMP**

References II

- The **MOSEK** modeling cookbook (2020).



- T. Gally, M. Pfetsch, S. Ulbrich: *A Framework for Solving Mixed-Integer Semidefinite Programs* (2018).
- M. Lubin and E. Yamangil and R. Bent, J. P. Vielma: *Extended Formulations in Mixed-integer Convex Programming* (2016).
- M. Grant, S. Boyd, Y. Ye: *Disciplined Convex Programming* (2006).
- C. Coey, L. Kapelevich, J. P. Vielma: *Towards Practical Generic Conic Optimization* (2020).

3. Conic duality

Lagrangian duality

Recall the nonlinear optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \leq 0. \end{array}$$

The Lagrangian duality approach defines the Lagrange function

$$L(x, \mu, \lambda) = f(x) + \mu^T h(x) + \lambda^T g(x),$$

and the dual function

$$g(\mu, \lambda) = \inf_x L(x, \mu, \lambda).$$

If $g(\mu, \lambda) > -\infty$, we call (μ, λ) dual feasible.

LP is special...

... because the dual takes on an explicit form:

$$f(x) = c^T x,\ \ h(x) = Ax - b,\ \ g(x) = -x$$

leads to the Lagrange function

$$L(x, \mu, \lambda) = c^T x + \mu^T (Ax - b) - \lambda^T x,$$

and the dual function is finite if (dual feasibility!)

$$A^T \mu + c - \lambda = 0 \text{ and } \lambda \geq 0.$$

Note that $\lambda \geq 0$ guarantees $-\lambda^T x \leq 0$ (for primal feasible x).

Extending duality to Conic Programming

In the conic framework

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax - b = 0 \\ & x \in \mathcal{K}, \end{array}$$

we need dual variables λ that satisfy

$$-\lambda^T x \leq 0 \ \forall x \in \mathcal{K},$$

thus giving rise to the set

$$\mathcal{K}^* = \{y \in \mathbb{R}^n \mid y^T x \geq 0 \ \forall x \in \mathcal{K}\}.$$

For any $\emptyset \neq \mathcal{K}$, $\mathcal{K}^*$ is a closed convex cone, and if $\mathcal{K}$ is a cone, we call $\mathcal{K}^*$ its *dual cone*!

Conic dual problem

The conic dual takes on the (explicit!) form

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & -A^T y + c - \lambda = 0 \\ & \lambda \in \mathcal{K}^*, \end{array}$$

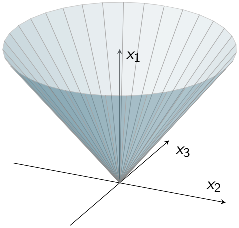
and the feasible set can more compactly be written as

$$c - A^T y \in \mathcal{K}^* \text{ or } c \geq_{\mathcal{K}^*} A^T y.$$

Weak duality comes for free:

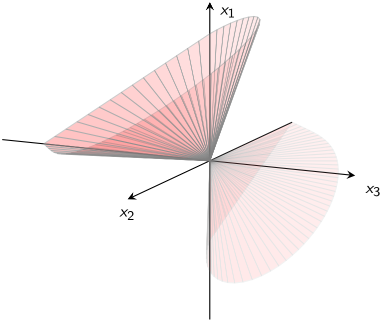
$$b^T y = (Ax)^T y = x^T \cdot A^T y = x^T \cdot (c - \lambda) = c^T x - \lambda^T x \leq c^T x.$$

Dual cones



Self-duality: for $\mathcal{K} \in \{\mathbb{R}^n_+, \mathcal{Q}^n, \mathbb{S}^n_+\}$,

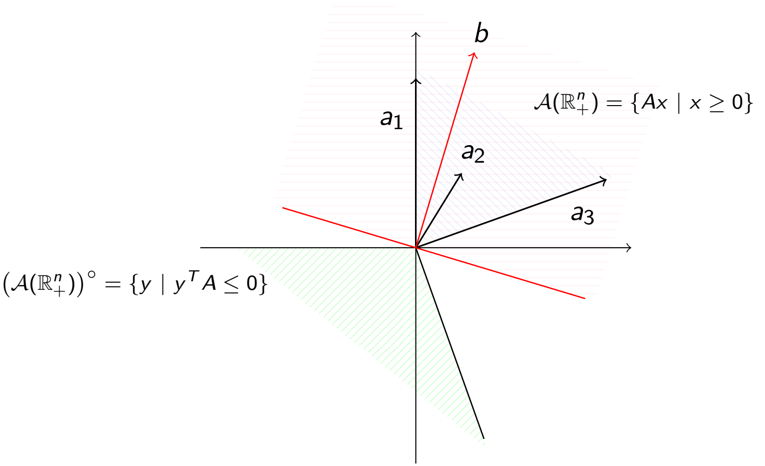
$$\mathcal{K}^* = \mathcal{K}.$$



$\mathcal{K}_{exp}$ is not self-dual:

$$(\mathcal{K}_{exp})^* = \mathrm{cl}\{x \in \mathbb{R}^3 \mid x_1 \geq -x_3 \exp(x_2/x_3), x_3 < 0\}$$

Farkas lemma - the LP case



Either $Ax = b, x \geq 0$ is feasible, or $y^T A \leq 0, y^T b > 0$ is so.

Farkas Lemma - the conic version

Let $\mathcal{A}(\mathcal{K}) = \{Ax \mid x \in \mathcal{K}\}$. The LP case translates almost verbatim to the conic case:

Lemma (Gärtner & Matoušek (2011))

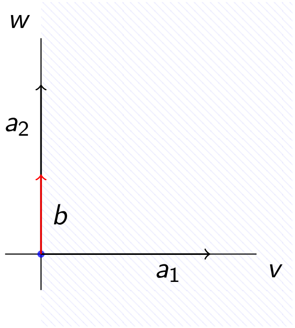
Let $\mathcal{K}$ be a closed convex cone. Exactly of of the following statements is true:

1. *$b \in \mathcal{A}(\mathcal{K})$ (primal system is feasible).*
2. *$-y^T A \in \mathcal{K}^*, b^T y > 0$ is feasible.*
3. *$b \notin \mathcal{A}(\mathcal{K})$ (primal system is infeasible), but $b \in \mathrm{cl}\,(\mathcal{A}(\mathcal{K}))$.*

In the third alternative the primal system is only *limit-feasible*.

An ill-posed example

$$\begin{array}{ll} \text{minimize} & u \\ \text{subject to} & v = 0 \\ & w = \frac{1}{2} \\ & (u, v, w) \in \mathcal{Q}_r^n \end{array}$$

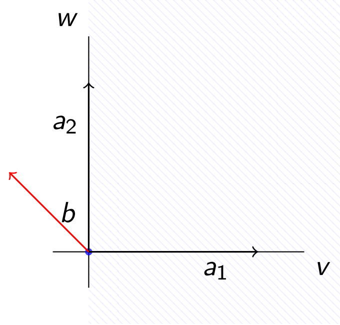


One can show that $\mathcal{A}(\mathcal{Q}_r^n) = (\mathbb{R}_{>0} \times \mathbb{R}) \cup \{(0,0)\}$.

Thus $b \notin \mathcal{A}(\mathcal{Q}_r^n)$ but $b \in \mathrm{cl}\,(\mathcal{A}(\mathcal{Q}_r^n))$!

Well-posed but infeasible - certified!

$$\begin{array}{ll} \text{minimize} & u \\ \text{subject to} & v = -\frac{1}{2} \\ & w = \frac{1}{2} \\ & (u, v, w) \in \mathcal{Q}_r^n \end{array}$$



$y = (-1, 0)^T$ is a certificate of infeasibility:

$$y^T b = \frac{1}{2} > 0 \text{ and } -y^T A = (0, 1, 0) \in (\mathcal{Q}_3^n)^* = \mathcal{Q}_3^n.$$

More generally, duality enables conic solvers to produce certificates of optimality, primal or dual infeasibility.

Strong duality

In the LP case we have:

Theorem (LP strong duality)

If at least one of $c^T x^*$ and $b^T y^*$ is finite, then $c^T x^* = b^T y^*$.

In the conic case we still have strong duality under a regularity assumption:

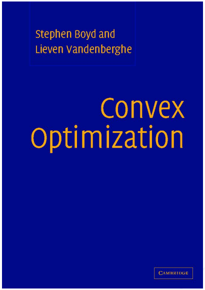
Theorem ((some version of) Conic strong duality)

If there is a strictly feasible point ($\exists x \in \textbf{int}(\mathcal{K}) : Ax = b$) and $c^T x^*$ is finite, then $c^T x^* = b^T y^*$.

In practice, a positive duality gap indicates issues with the problem formulation.

References III

- B. Gärtner, J Matoušek: *Approximation algorithms and semidefinite programming* (2012).
- S. Boyd & L. Vandenberghe: *Convex Optimization* (2013).



4. Numerical solution methods

Reduce a somehow constrained optimization problem

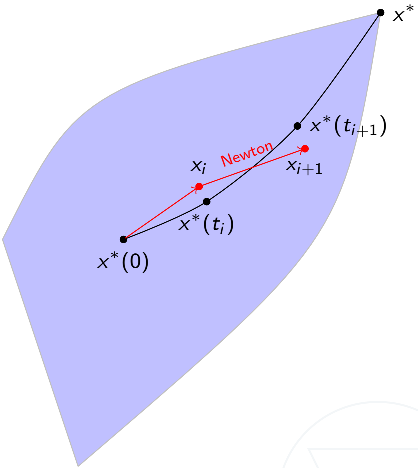
$$\min_{Ax=b, x\in\mathcal{X}} c^t x$$

to a linear equality constrained optimization:

$$\min_{Ax=b} t \cdot c^T x + F(x)$$

where $F(x)$ is such that

$$F(x) \xrightarrow[x\to\delta\mathcal{X}]{} \infty.$$



Conic barriers

A crucial task is to find a barrier function for a given $\mathcal{K}$.

In both theory and practice, *self-concordance* of a barrier has proven a desirable property.

- For the quadratic cone $\mathcal{Q}^n$:

$$Q(x) = -\log(x_1^2 - x_2^2 - \ldots - x_n^2)$$

- For the semidefinite cone $\mathbb{S}_+^n$:

$$S(X) = -\log(\det(x))$$

- For the relative entropy cone $\mathcal{K}_{entr}$:

$$E(X) = -\sum_{i=1}^{d}(\log(u_i) + \log(v_i)) - \log(x - \sum_{i=1}^{d} u_i \log(u_i/v_i))$$

Self-dual embedding

Several IPMs for Conic Programming use the homogeneous model (or the self-dual embedding):

$$Ax - b\tau = 0$$
$$c\tau - A^T y - \lambda = 0$$
$$c^T x - b^T y + \kappa = 0$$
$$x \in \mathcal{K},\ \lambda \in \mathcal{K}^*,\ \tau, \kappa \geq 0,$$

encapsulates different duality cases:

- If $\tau > 0$, $\kappa = 0$ then $\frac{1}{\tau}(x, y, \lambda)$ is optimal,
$$Ax = b\tau, \quad c\tau - A^T y = \lambda, \quad c^T x - b^T y = 0.$$
- If $\tau = 0$, $\kappa > 0$ then the problem is infeasible,
$$Ax = 0, \quad -A^T y = \lambda, \quad c^T x - b^T y < 0.$$
- If $\tau = 0$, $\kappa = 0$ then the problem is ill-posed.

1

Symmetric vs. non-symmetric cones

IPMs for symmetric cones are more extensively studied and mature.

- For symmetric cones, the so-called centrality condition is just a perturbed KKT-system.
- For symmetric cones we have the Nesterov-Todd scaling

$$Wx = W^{-1}\lambda = s,$$

- but not for non-symmetric cones:

$$Vx = V^{-T}\lambda = s.$$

Implementing IPMs for non-symmetric cones is an active research area!

- Y. Nesterov and M. J. Todd: *Primal-dual interior-point methods for self-scaled cones* (1998).
- S. H. Schmieta and F. Alizadeh: *Extension of primal-dual interior point methods to symmetric cones* (2003).
- A. Skajaa and Y. Ye: *A homogeneous interior-point algorithm for nonsymmetric convex conic optimization* (2015).
- S. A. Serrano: *Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone* (2015).
- D. Papp and S. Yıldız: *On A homogeneous interior-point algorithm for non-symmetric convex conic optimization* (2017).
- D. Papp and S. Yıldız: *Sum-of-squares optimization without semidefinite programming* (2019).
- J. Dahl and E Andersen: *A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization* (2019).

The Mixed-Integer case

Recall the Mixed-Integer Conic Programming problem:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \cap \left(\mathbb{Z}^p \times \mathbb{R}^{n-p} \right) \end{array}$$

Two (convex) Mixed-Integer Nonlinear Programming approaches have been prominently translated to Mixed-Integer Conic Programming:

- Non-linear Branch-and-Bound $\rightarrow$ Conic Branch-and-Bound.
- Outer approximation: a convex constraint $g(x) \leq 0$ can be approximated by a gradient cut

$$g(\hat{x}) + \nabla g(\hat{x})^T (x - \hat{x}) \leq 0.$$

- In the conic case we have other ways of approximating the feasible set.

Conic outer approximation

Exploit the polar cone $\mathcal{K}^\circ = -\mathcal{K}^*$ (exploit structure!):

- Clearly $\mathcal{K} = \{x \mid a^T x \leq 0 \ \forall a \in \mathcal{K}^\circ\}$, so any point $a \in \mathcal{K}^\circ$ separates $\hat{x} \notin \mathcal{K}$: $a^T \hat{x} > 0$.

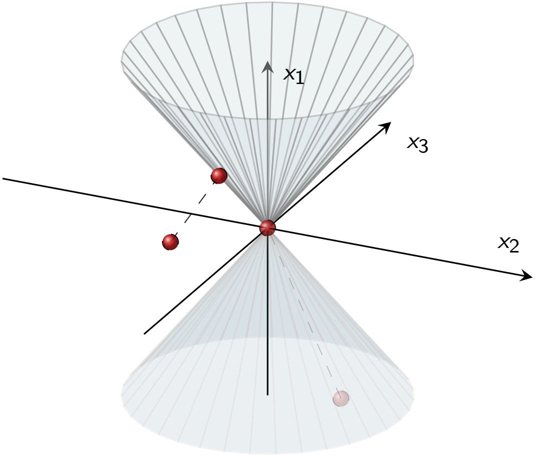
- If $\mathcal{K} = \{x \mid g(x) \leq 0\}$, then $a = \nabla g(\hat{x})$ is a separator, see *Lubin* (2017).

- Otherwise, one can solve the maximal separation problem

$$\max_{a \in \mathcal{K}^\circ, \|a\|_2 \leq 1} a^T \hat{x}.$$

- This is the dual of the projection problem $\min_{x \in \mathcal{K}} \|x - \hat{x}\|_2$.

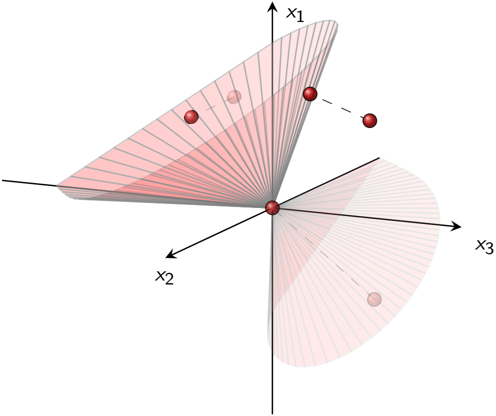
Cone projections



For the symmetric cones, the projection problem can be solved algebraically!

Cone projections

For the exponential and power cones, the projection problem is at most a univariate root-finding problem, shown by *Hien* (2015) and *Friberg* (2018).

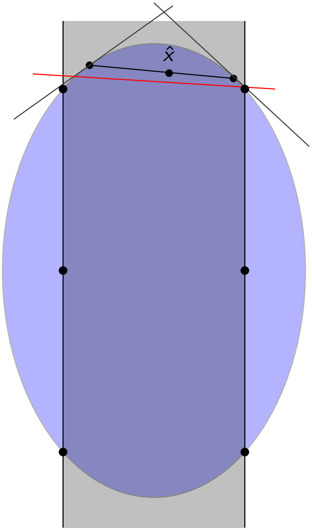


Mixed-Integer Conic Programming as a research area

When developing (components of) Mixed-Integer Conic Programming solvers, we may:

- exploit the fact of dealing with a cone $\mathcal{K}$:
  - Deriving disjunctive cuts: *Lodi, Tanneau, Vielma* (2020).
  - Conic outer approximation: *Coey, Lubin, Vielma* (2018).
- exploit the fact of dealing with a specific cone (limited structure!):
  - Cutting planes for $\mathcal{Q}^n$: *Andersen, Jensen* (2013) and others.
  - Primal heuristics for $\mathcal{Q}^n$: *Çay, Pólik, Terlaky* (2018).
  - Disjunctive Programming techniques: *Bernal* (2019).

Disjunctive cuts



- In the general convex case, *Bonami* (2011) proposed to
  1. solve NLP,
  2. build OA,
  3. solve Cut Generating LP
- In the conic case, *Lodi, Tanneau, Vielma* (2020)
  1. solve Cut Generating Conic Program

An application of conic duality!

Disjunctive programming

When dealing with nonlinear disjunctive, or indicator constraints

$$z = 1 \implies g(x) \leq 0,$$

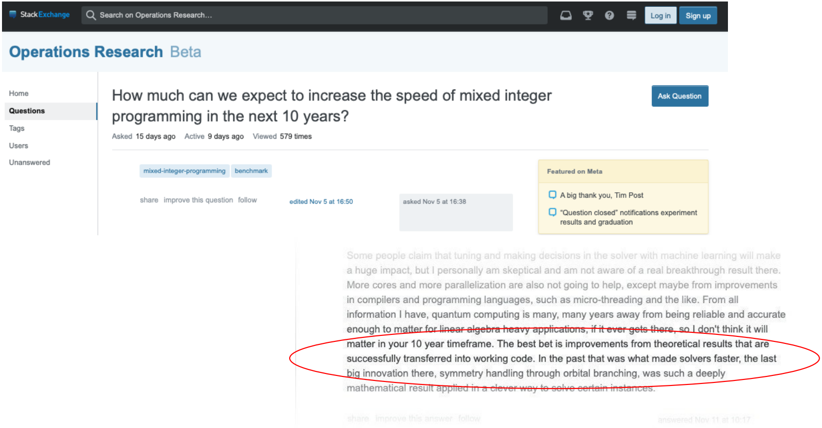
Ceria and Soares (1999) show that the perspective function $z \cdot g(x/z)$ can be used for building strong continuous relaxations.

This is just another cone, maybe a well-known one:

- $z = 1 \implies (y, 1, x) \in \mathbb{Q}_r^3$ leads to $(y, z, x) \in \mathbb{Q}_r^3$.
- $z = 1 \implies (y, 1, x) \in \mathcal{K}_{exp}$ leads to $(y, z, x) \in \mathcal{K}_{exp}$.

There are no differentiability issues!

Mixed-Integer Conic Programming as a research area



"Mixed-Integer Conic Programming is very immature yet, so good improvements can be expected as theory and practice develop."

References IV b

- S. Ceria and J. Soares: *Convex programming for disjunctive convex optimization* (1999).
- P. Bonami: *Lift-and-Project Cuts for Mixed Integer Convex Programs* (2011).
- K. Andersen and A. Jensen: *Intersection Cuts for Mixed Integer Conic Quadratic Sets* (2013).
- P. Belotti, J. Góez, I. Pólik, T. Ralphs and T. Terlaky: *On Families of Quadratic Surfaces Having Fixed Intersections with Two Hyperplanes* (2013).
- F. Kılınç-Karzan and S. Yıldız: *Two-Term Disjunctions on the Second-Order Cone* (2015).
- A. Lodi, M. Tanneau and J. P. Vielma: *Disjunctive cuts in Mixed-Integer Conic Optimization* (2020).

- L. T. K. Hien: *Differential properties of Euclidean projection onto power cone* (2015).
- H. Friberg: *Projection onto the exponential cone: a univariate root-finding problem* (2018).
- M. Lubin: *Mixed-integer convex optimization: outer approximation algorithms and modeling power* (2017).
- C. Coey, M Lubin and J.P. Vielma: *Outer Approximation With Conic Certificates For Mixed-Integer Convex Problems* (2018).
- S. Çay, I. Pólik and T. Terlaky: *The first heuristic specifically for mixed-integer second-order cone optimization* (2018).
- D. Bernal: *Easily Solvable Convex Mixed-Integer Nonlinear Programs Derived from Generalized Disjunctive Programming using Cones* (AIChE Annual Meeting, 2019)