Solving convex problems involving powers using conic optimization and a new self-concordant barrier

François Glineur

francois.glineur@uclouvain.be

joint work with Robert Chares
UCL/CORE (Center for Operations Research and Econometrics)
UCL/INMA (Department of Mathematical Engineering)

September 19, 2007

CFG 07 Heidelberg University

Overview

1. **Motivation**
    - ◇ Why convex optimization?
    - ◇ Why a conic formulation?
2. **Unified conic formulation**
    - ◇ The power cone
    - ◇ Modelling problems involving powers using the power cone
    - ◇ Modelling problems involving exponentials using the power cone
3. **Concluding remarks**
    - ◇ Future plans

Overview

1. Motivation

◇ Why convex optimization?

◇ Why a conic formulation?

2. Unified conic formulation

◇ The power cone

◇ Modelling problems involving powers using the power cone

◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

◇ Future plans

Overview

1. Motivation

◇ Why convex optimization?

◇ Why a conic formulation?

2. Unified conic formulation

◇ The power cone

◇ Modelling problems involving powers using the power cone

◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

◇ Future plans

Overview

1. Motivation

- ◇ Why convex optimization?
- ◇ Why a conic formulation?

2. Unified conic formulation

- ◇ The power cone
- ◇ Modelling problems involving powers using the power cone
- ◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

- ◇ Future plans

Overview

1. Motivation

- ◇ Why convex optimization?
- ◇ Why a conic formulation?

2. Unified conic formulation

- ◇ The power cone
- ◇ Modelling problems involving powers using the power cone
- ◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

- ◇ Future plans

Overview

1. Motivation

- ◇ Why convex optimization?
- ◇ Why a conic formulation?

2. Unified conic formulation

- ◇ The power cone
- ◇ Modelling problems involving powers using the power cone
- ◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

- ◇ Future plans

Overview

1. **Motivation**
   - ⋄ Why convex optimization?
   - ⋄ Why a conic formulation?
2. **Unified conic formulation**
   - ⋄ The power cone
   - ⋄ Modelling problems involving powers using the power cone
   - ⋄ Modelling problems involving exponentials using the power cone
3. **Concluding remarks**
   - ⋄ Future plans

Overview

1. Motivation

- ◇ Why convex optimization?
- ◇ Why a conic formulation?

2. Unified conic formulation

- ◇ The power cone
- ◇ Modelling problems involving powers using the power cone
- ◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

- ◇ Future plans

Overview

1. **Motivation**
   - ◇ Why convex optimization?
   - ◇ Why a conic formulation?
2. **Unified conic formulation**
   - ◇ The power cone
   - ◇ Modelling problems involving powers using the power cone
   - ◇ Modelling problems involving exponentials using the power cone
3. **Concluding remarks**
   - ◇ Future plans

Convex optimization

Nonlinear optimization

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \le 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

- Variables: finite-dimensional vector $x \in \mathbb{R}^n$
- Constraints: finite number of (in)equalities, indexed by sets I and E

Problem is convex when

- objective function f_0 is convex
- functions f_i defining inequalities $f_i(x) \le 0$ are convex for all $i \in I$
- functions f_i defining equalities $f_i(x) = 0$ are affine for all $i \in E$

Convex optimization

Nonlinear optimization

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

- Variables: finite-dimensional vector $x \in \mathbb{R}^n$
- Constraints: finite number of (in)equalities, indexed by sets I and E

Problem is convex when

- objective function f_0 is convex
- functions f_i defining inequalities $f_i(x) \leq 0$ are convex for all $i \in I$
- functions f_i defining equalities $f_i(x) = 0$ are affine for all $i \in E$

Well-known classes of convex problems

$$\min_{x\in\mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \le 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

- Linear optimization (LO): f_0 and f_i are affine for all $i \in E \cup I$

$$f_i(x) = a_i^{\mathrm{T}} x - b_i$$

- Quadratically constrained quadratic optimization (QCQO): f_0 and f_i are convex quadratic for all $i \in I$

$$f_i(x) = x^{\mathrm{T}} Q_i x + r_i^{\mathrm{T}} x + s_i \text{ with } Q_i \succeq 0$$

(equalities f_i, if present, must still be affine for $i \in E$)

- Convex quadratic can be rewritten using composition of squared Euclidean norm and linear (vector) function:

$$f_i(x) = \|A_i x\|^2 + (r_i^{\mathrm{T}} x + s_i) \text{ with } Q_i = A_i^{\mathrm{T}} A_i$$

Well-known classes of convex problems

$$\min_{x\in\mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

- Linear optimization (LO): f_0 and f_i are affine for all $i \in E \cup I$

$$f_i(x) = a_i^{\mathrm{T}} x - b_i$$

- Quadratically constrained quadratic optimization (QCQO): f_0 and f_i are convex quadratic for all $i \in I$

$$f_i(x) = x^{\mathrm{T}} Q_i x + r_i^{\mathrm{T}} x + s_i \text{ with } Q_i \succeq 0$$

(equalities f_i, if present, must still be affine for $i \in E$)
- Convex quadratic can be rewritten using composition of squared Euclidean norm and linear (vector) function:

$$f_i(x) = \|A_i x\|^2 + (r_i^{\mathrm{T}} x + s_i) \text{ with } Q_i = A_i^{\mathrm{T}} A_i$$

Well-known classes of convex problems

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

- Linear optimization (LO): f_0 and f_i are affine for all $i \in E \cup I$

$$f_i(x) = a_i^{\mathrm{T}} x - b_i$$

- Quadratically constrained quadratic optimization (QCQO): f_0 and f_i are convex quadratic for all $i \in I$

$$f_i(x) = x^{\mathrm{T}} Q_i x + r_i^{\mathrm{T}} x + s_i \text{ with } Q_i \succeq 0$$

(equalities f_i, if present, must still be affine for $i \in E$)

- Convex quadratic can be rewritten using composition of squared Euclidean norm and linear (vector) function:

$$f_i(x) = \|A_i x\|^2 + (r_i^{\mathrm{T}} x + s_i) \text{ with } Q_i = A_i^{\mathrm{T}} A_i$$

Well-known classes of convex problems

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

- Linear optimization (LO): f_0 and f_i are affine for all $i \in E \cup I$

$$f_i(x) = a_i^{\mathrm{T}} x - b_i$$

- Quadratically constrained quadratic optimization (QCQO): f_0 and f_i are convex quadratic for all $i \in I$

$$f_i(x) = x^{\mathrm{T}} Q_i x + r_i^{\mathrm{T}} x + s_i \text{ with } Q_i \succeq 0$$

  (equalities f_i, if present, must still be affine for $i \in E$)
- Convex quadratic can be rewritten using composition of squared Euclidean norm and linear (vector) function:

$$f_i(x) = \|A_i x\|^2 + (r_i^{\mathrm{T}} x + s_i) \text{ with } Q_i = A_i^{\mathrm{T}} A_i$$

More classes of convex problems

- Geometric optimization (GO):
  f_0 and f_i are posynomials (in exponential form) for all $i \in I$

$$f_i(x) = c_i + \sum_{j \in M_i} \exp(a_{ij}x - b_{ij})$$

  Each term in the sum is the composition of exponential and affine scalar function

- l_p-norm optimization (l_pO):
  f_0 linear, f_i are affine plus sum of convex powers with affine scalar arguments for all $i \in I$

$$f_i(x) = a_{i0}x - b_{i0} + \sum_{j \in M_i} |a_{ij}x - b_{ij}|^{p_{ij}} \text{ with } p_{ij} \geq 1$$

More classes of convex problems

- Geometric optimization (GO):
  f_0 and f_i are posynomials (in exponential form) for all $i \in I$

$$f_i(x) = c_i + \sum_{j \in M_i} \exp(a_{ij}x - b_{ij})$$

  Each term in the sum is the composition of exponential and affine scalar function

- l_p-norm optimization (l_pO):
  f_0 linear, f_i are affine plus sum of convex powers with affine scalar arguments for all $i \in I$

$$f_i(x) = a_{i0}x - b_{i0} + \sum_{j \in M_i} |a_{ij}x - b_{ij}|^{p_{ij}} \text{ with } p_{ij} \geq 1$$

◇ Sum-of-norm optimization (SNO):
f_0 (and f_i for all $i \in I$, if any) are convex norms with affine arguments

$$f_i(x) = \sum_{j \in M_i} \|A_{ij}x - b_{ij}\|_{p_{ij}} \quad \text{with } p_{ij} \geq 1$$

with $\|y\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}$

◇ Entropy optimization (EO):
f_0 is a sum of entropy terms, f_i are affine for all $i \in E$

$$f_0(x) = \sum_i x_i \log x_i \quad \text{(implicitly implying } x \geq 0\text{)}$$

◇ Analytic centering (AC):
f_0 is a sum of logarithmic terms, f_i are affine for all $i \in I \cup E$

$$f_0(x) = -\sum_{j \in M_j} \log(a_{ij}x - b_{ij})$$

◇ Sum-of-norm optimization (SNO):
f_0 (and f_i for all $i \in I$, if any) are convex norms with affine arguments

$$f_i(x) = \sum_{j \in M_i} \|A_{ij}x - b_{ij}\|_{p_{ij}} \text{ with } p_{ij} \geq 1$$

with $\|y\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}$

◇ Entropy optimization (EO):
f_0 is a sum of entropy terms, f_i are affine for all $i \in E$

$$f_0(x) = \sum_i x_i \log x_i \quad \text{(implicitly implying } x \geq 0)$$

◇ Analytic centering (AC):
f_0 is a sum of logarithmic terms, f_i are affine for all $i \in I \cup E$

$$f_0(x) = -\sum_{j \in M_j} \log(a_{ij}x - b_{ij})$$

◇ Sum-of-norm optimization (SNO):
f_0 (and f_i for all $i \in I$, if any) are convex norms with affine arguments

$$f_i(x) = \sum_{j \in M_i} \left\| A_{ij}x - b_{ij} \right\|_{p_{ij}} \text{ with } p_{ij} \geq 1$$

with $\|y\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}$

◇ Entropy optimization (EO):
f_0 is a sum of entropy terms, f_i are affine for all $i \in E$

$$f_0(x) = \sum_i x_i \log x_i \quad \text{(implicitly implying } x \geq 0\text{)}$$

◇ Analytic centering (AC):
f_0 is a sum of logarithmic terms, f_i are affine for all $i \in I \cup E$

$$f_0(x) = -\sum_{j \in M_j} \log(a_{ij}x - b_{ij})$$

A common formulation

All above-mentioned problems can be described as follows:

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

with functional terms f_0 and f_i defined by

$$f_i(x) = \sum_{j \in M_i} g_{ij}(A_{ij}x - b_{ij})$$

where nonlinearity is confined to functions g_{ij}:

- $x \mapsto x$ (identity): for linear optimization and for all equalities
- $x \mapsto \|x\|^2$: for quadratically constrained quadratic optimization
- $x \mapsto e^x$ for geometric optimization
- $x \mapsto |x|^p$ with $p \geq 1$ for l_p-norm optimization
- $x \mapsto \|x\|_p$ with $p \geq 1$ for sum-of-norm optimization
- $x \mapsto -\log x$ for analytic centering, $x \log x$ for entropy optimization

In summary: separable functions with every term being the composition of a simple (often scalar) convex nonlinear function with an affine function

A common formulation

All above-mentioned problems can be described as follows:

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \le 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

with functional terms f_0 and f_i defined by

$$f_i(x) = \sum_{j \in M_i} g_{ij}(A_{ij}x - b_{ij})$$

where nonlinearity is confined to functions g_{ij}:

- $x \mapsto x$ (identity): for linear optimization and for all equalities
- $x \mapsto \|x\|^2$: for quadratically constrained quadratic optimization
- $x \mapsto e^x$ for geometric optimization
- $x \mapsto |x|^p$ with $p \ge 1$ for l_p-norm optimization
- $x \mapsto \|x\|_p$ with $p \ge 1$ for sum-of-norm optimization
- $x \mapsto -\log x$ for analytic centering, $x \log x$ for entropy optimization

In summary: separable functions with every term being the composition of a simple (often scalar) convex nonlinear function with an affine function

A common formulation

All above-mentioned problems can be described as follows:

$$\min_{x \in \mathbb{R}^n} f_0(x) \text{ such that } f_i(x) \leq 0 \text{ for all } i \in I \text{ and } f_i(x) = 0 \text{ for all } i \in E$$

with functional terms f_0 and f_i defined by

$$f_i(x) = \sum_{j \in M_i} g_{ij}(A_{ij}x - b_{ij})$$

where nonlinearity is confined to functions g_{ij}:

- $x \mapsto x$ (identity): for linear optimization and for all equalities
- $x \mapsto \|x\|^2$: for quadratically constrained quadratic optimization
- $x \mapsto e^x$ for geometric optimization
- $x \mapsto |x|^p$ with $p \geq 1$ for l_p-norm optimization
- $x \mapsto \|x\|_p$ with $p \geq 1$ for sum-of-norm optimization
- $x \mapsto -\log x$ for analytic centering, $x \log x$ for entropy optimization

In summary: separable functions with every term being the composition of a simple (often scalar) convex nonlinear function with an affine function

Properties of convex optimization

Why is it interesting to consider (or restrict yourself to) convex optimization problems?

Passive features:

- every local minimum is a global minimum
- set of optimal solutions is convex
- optimality (KKT) conditions are sufficient (with regularity assumption)

Any algorithm or solver applied to a convex problem will automatically benefit from those features

Properties of convex optimization

Why is it interesting to consider (or restrict yourself to) convex optimization problems?

Passive features:

- every local minimum is a global minimum
- set of optimal solutions is convex
- optimality (KKT) conditions are sufficient (with regularity assumption)

Any algorithm or solver applied to a convex problem will automatically benefit from those features

Properties of convex optimization

Why is it interesting to consider (or restrict yourself to) convex optimization problems?

Passive features:

- every local minimum is a global minimum
- set of optimal solutions is convex
- optimality (KKT) conditions are sufficient (with regularity assumption)

Any algorithm or solver applied to a convex problem will automatically benefit from those features

Properties of convex optimization

Active features:

- possibility of writing down a dual problem strongly related to original problem
  (weak duality and, with regularity assumption, strong duality $\rightarrow$ optimality certificates)
- possibility of designing dedicated algorithm with polynomial algorithmic complexity
  (in most of the cases: an interior-point method based on the theory of self-concordant barriers)

- To use those, additional work is needed for each problem class!
- Need to exploit specific structure of each problem class
- Reward for additional work is better understanding and ability to solve problems more efficiently (including large-scale)

Properties of convex optimization

Active features:

- possibility of writing down a dual problem strongly related to original problem
  (weak duality and, with regularity assumption, strong duality $\rightarrow$ optimality certificates)
- possibility of designing dedicated algorithm with polynomial algorithmic complexity
  (in most of the cases: an interior-point method based on the theory of self-concordant barriers)

- To use those, additional work is needed for each problem class!
- Need to exploit specific structure of each problem class
- Reward for additional work is better understanding and ability to solve problems more efficiently (including large-scale)

Properties of convex optimization

Active features:

- ◇ possibility of writing down a dual problem strongly related to original problem
  (weak duality and, with regularity assumption, strong duality $\rightarrow$ optimality certificates)
- ◇ possibility of designing dedicated algorithm with polynomial algorithmic complexity
  (in most of the cases: an interior-point method based on the theory of self-concordant barriers)

- ◇ To use those, additional work is needed for each problem class!
- ◇ Need to exploit specific structure of each problem class
- ◇ Reward for additional work is better understanding and ability to solve problems more efficiently (including large-scale)

Properties of convex optimization

Active features:

- ◇ possibility of writing down a dual problem strongly related to original problem
  (weak duality and, with regularity assumption, strong duality $\longrightarrow$ optimality certificates)
- ◇ possibility of designing dedicated algorithm with polynomial algorithmic complexity
  (in most of the cases: an interior-point method based on the theory of self-concordant barriers)

- ◇ To use those, additional work is needed for each problem class!
- ◇ Need to exploit specific structure of each problem class
- ◇ Reward for additional work is better understanding and ability to solve problems more efficiently (including large-scale)

Overview

1. Motivation

- Why convex optimization?
- Why a conic formulation?

2. Unified conic formulation

- The power cone
- Modelling problems involving powers using the power cone
- Modelling problems involving exponentials using the power cone

3. Concluding remarks

- Future plans

Conic optimization

Generalization of linear optimization (e.g. dual form)

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \leq c$$

where a new ordering is used instead of $\leq$:

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

- Ordering defined by a set K: $a \preceq_K b \Leftrightarrow 0 \preceq_K b - a \Leftrightarrow b - a \in K$
- Set K has to be a convex cone for useful properties of ordering to hold (and also: closed, solid and pointed for technical reasons)
- Conic optimization problems are clearly convex
- Any convex problem can be cast as a conic optimization problem
- The point of a conic formulation is to make it easier to benefit from active features of convex optimization (duality and algorithms)

Conic optimization

Generalization of linear optimization (e.g. dual form)

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \leq c$$

where a new ordering is used instead of $\leq$:

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

- Ordering defined by a set K: $a \preceq_K b \Leftrightarrow 0 \preceq_K b - a \Leftrightarrow b - a \in K$
- Set K has to be a convex cone for useful properties of ordering to hold (and also: closed, solid and pointed for technical reasons)
- Conic optimization problems are clearly convex
- Any convex problem can be cast as a conic optimization problem
- The point of a conic formulation is to make it easier to benefit from active features of convex optimization (duality and algorithms)

Conic optimization

Generalization of linear optimization (e.g. dual form)

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \leq c$$

where a new ordering is used instead of $\leq$:

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

- Ordering defined by a set K: $a \preceq_K b \Leftrightarrow 0 \preceq_K b - a \Leftrightarrow b - a \in K$
- Set K has to be a convex cone for useful properties of ordering to hold (and also: closed, solid and pointed for technical reasons)
- Conic optimization problems are clearly convex
- Any convex problem can be cast as a conic optimization problem
- The point of a conic formulation is to make it easier to benefit from active features of convex optimization (duality and algorithms)

Conic optimization

Generalization of linear optimization (e.g. dual form)

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \leq c$$

where a new ordering is used instead of $\leq$:

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

- Ordering defined by a set K: $a \preceq_K b \Leftrightarrow 0 \preceq_K b - a \Leftrightarrow b - a \in K$
- Set K has to be a convex cone for useful properties of ordering to hold (and also: closed, solid and pointed for technical reasons)
- Conic optimization problems are clearly convex
- Any convex problem can be cast as a conic optimization problem
- The point of a conic formulation is to make it easier to benefit from active features of convex optimization (duality and algorithms)

Conic optimization

Generalization of linear optimization (e.g. dual form)

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \leq c$$

where a new ordering is used instead of $\leq$:

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

- ◇ Ordering defined by a set K: $a \preceq_K b \Leftrightarrow 0 \preceq_K b - a \Leftrightarrow b - a \in K$
- ◇ Set K has to be a convex cone for useful properties of ordering to hold (and also: closed, solid and pointed for technical reasons)
- ◇ Conic optimization problems are clearly convex
- ◇ Any convex problem can be cast as a conic optimization problem
- ◇ The point of a conic formulation is to make it easier to benefit from active features of convex optimization (duality and algorithms)

Combining several cones

Considering several conic constraints

$$A_1^{\mathrm{T}} y \preceq_{K_1} c_1 \text{ and } A_2^{\mathrm{T}} y \preceq_{K_2} c_2$$

which are equivalent to

$$c_1 - A_1^{\mathrm{T}} y \in K_1 \text{ and } c_2 - A_2^{\mathrm{T}} y \in K_2$$

one introduces the Cartesian product cone $K = K_1 \times K_2$ to write

$$(c_1 - A_1^{\mathrm{T}} y, c_2 - A_2^{\mathrm{T}} y) \in K_1 \times K_2$$

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} - \begin{pmatrix} A_1^{\mathrm{T}} \\ A_2^{\mathrm{T}} \end{pmatrix} \succeq_{K_1 \times K_2} 0 \Leftrightarrow A^{\mathrm{T}} y \preceq_K c$$

$\longrightarrow$ for theory, a single cone can be considered without loss of generality

Combining several cones

Considering several conic constraints

$$A_1^{\mathrm{T}} y \preceq_{K_1} c_1 \text{ and } A_2^{\mathrm{T}} y \preceq_{K_2} c_2$$

which are equivalent to

$$c_1 - A_1^{\mathrm{T}} y \in K_1 \text{ and } c_2 - A_2^{\mathrm{T}} y \in K_2$$

one introduces the Cartesian product cone $K = K_1 \times K_2$ to write

$$(c_1 - A_1^{\mathrm{T}} y, c_2 - A_2^{\mathrm{T}} y) \in K_1 \times K_2$$

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} - \begin{pmatrix} A_1^{\mathrm{T}} \\ A_2^{\mathrm{T}} \end{pmatrix} \succeq_{K_1 \times K_2} 0 \Leftrightarrow A^{\mathrm{T}} y \preceq_K c$$

$\rightarrow$ for theory, a single cone can be considered without loss of generality

Algorithms for conic optimization

- Interior-point methods can easily be applied to conic optimization
- Main ingredient: a good barrier function for every cone K involved
- A good barrier is for example a self-concordant barrier, i.e. $F : \operatorname{int} K \mapsto \mathbb{R}$ satisfying
  - F is convex and three times differentiable
  - $F(x) \to +\infty$ when $x \to \partial K$
  - the following two conditions hold

$$\nabla^3 F(x)[h,h,h] \leq 2 \left(\nabla^2 F(x)[h,h]\right)^{\frac{3}{2}}$$
$$\nabla F(x)^{\mathrm{T}} (\nabla^2 F(x))^{-1} \nabla F(x) \leq \nu$$

  for all $x \in \operatorname{int} C$ and $h \in \mathbb{R}^n$
- Once a good barrier is known, design of a polynomial-time algorithm can be completely straightforward (e.g. using standard short or long step path-following algorithm)

Algorithms for conic optimization

- Interior-point methods can easily be applied to conic optimization
- Main ingredient: a good barrier function for every cone K involved
- A good barrier is for example a self-concordant barrier, i.e. $F : \operatorname{int} K \mapsto \mathbb{R}$ satisfying
  - F is convex and three times differentiable
  - $F(x) \to +\infty$ when $x \to \partial K$
  - the following two conditions hold

$$\nabla^3 F(x)[h,h,h] \leq 2\left(\nabla^2 F(x)[h,h]\right)^{\frac{3}{2}}$$
$$\nabla F(x)^{\mathrm{T}}(\nabla^2 F(x))^{-1}\nabla F(x) \leq \nu$$

  for all $x \in \operatorname{int} C$ and $h \in \mathbb{R}^n$
- Once a good barrier is known, design of a polynomial-time algorithm can be completely straightforward (e.g. using standard short or long step path-following algorithm)

Algorithms for conic optimization

- Interior-point methods can easily be applied to conic optimization
- Main ingredient: a good barrier function for every cone K involved
- A good barrier is for example a self-concordant barrier, i.e. $F : \operatorname{int} K \mapsto \mathbb{R}$ satisfying
  - F is convex and three times differentiable
  - $F(x) \to +\infty$ when $x \to \partial K$
  - the following two conditions hold

$$\nabla^3 F(x)[h,h,h] \leq 2 \left(\nabla^2 F(x)[h,h]\right)^{\frac{3}{2}}$$
$$\nabla F(x)^{\mathrm{T}} (\nabla^2 F(x))^{-1} \nabla F(x) \leq \nu$$

  for all $x \in \operatorname{int} C$ and $h \in \mathbb{R}^n$

- Once a good barrier is known, design of a polynomial-time algorithm can be completely straightforward (e.g. using standard short or long step path-following algorithm)

Algorithms for conic optimization

- Interior-point methods can easily be applied to conic optimization
- Main ingredient: a good barrier function for every cone K involved
- A good barrier is for example a self-concordant barrier, i.e. $F : \operatorname{int} K \mapsto \mathbb{R}$ satisfying
  - F is convex and three times differentiable
  - $F(x) \to +\infty$ when $x \to \partial K$
  - the following two conditions hold

$$\nabla^3 F(x)[h,h,h] \leq 2 \left(\nabla^2 F(x)[h,h]\right)^{\frac{3}{2}}$$
$$\nabla F(x)^{\mathrm{T}} (\nabla^2 F(x))^{-1} \nabla F(x) \leq \nu$$

  for all $x \in \operatorname{int} C$ and $h \in \mathbb{R}^n$
- Once a good barrier is known, design of a polynomial-time algorithm can be completely straightforward (e.g. using standard short or long step path-following algorithm)

Duality for conic optimization

Problem

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

admits a nice symmetrical dual

$$\min c^{\mathrm{T}} x \text{ such that } Ax = b \text{ and } x \succeq_{K^*} 0$$

based on the notion of dual cone

$$K^* = \{z \in \mathbb{R}^n \text{ such that } x^{\mathrm{T}} z \geq 0 \; \forall x \in K\}$$

- Weak duality always holds, strong duality holds with regularity assumption (existence of a strictly interior point)
- Only effort involved in determining a dual problem is computing the dual cone
- Potentially allows design of (symmetrical) primal-dual algorithms

Duality for conic optimization

Problem

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

admits a nice symmetrical dual

$$\min c^{\mathrm{T}} x \text{ such that } Ax = b \text{ and } x \succeq_{K^*} 0$$

based on the notion of dual cone

$$K^* = \{z \in \mathbb{R}^n \text{ such that } x^{\mathrm{T}} z \geq 0 \; \forall x \in K\}$$

- ◇ Weak duality always holds, strong duality holds with regularity assumption (existence of a strictly interior point)
- ◇ Only effort involved in determining a dual problem is computing the dual cone
- ◇ Potentially allows design of (symmetrical) primal-dual algorithms

Examples of conic optimization

Following 3 cones are (by far) most commonly used

1. $K = \mathbb{R}_+$ is the standard ordering, leading to linear optimization
2. $K = \mathbb{L}^n$ leads to second-order cone optimization (including QCQO)

$$\mathbb{L}^n = \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} \mid \sqrt{x_1^2 + \cdots + x_n^2} \leq x_0\}$$

3. $K = \mathbb{S}^n_+$ (positive semidefinite matrices) for semidefinite optimization

- Those cones share additional theoretical properties (symmetric, i.e. homogeneous and self-dual)
- A fourth cone ($K = \{0\}, K^* = \mathbb{R}$) used for modelling convenience
- Many problems from various domains (e.g. mechanical and eletrical engineering, finance) can be modelled using these cones
- Many solvers available for problems involving these cones
- However, no solver seems available (yet) for any other cone!

Examples of conic optimization

Following 3 cones are (by far) most commonly used

1. $K = \mathbb{R}_+$ is the standard ordering, leading to linear optimization
2. $K = \mathbb{L}^n$ leads to second-order cone optimization (including QCQO)

$$\mathbb{L}^n = \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} \mid \sqrt{x_1^2 + \cdots + x_n^2} \leq x_0\}$$

3. $K = \mathbb{S}^n_+$ (positive semidefinite matrices) for semidefinite optimization

- ◇ Those cones share additional theoretical properties (symmetric, i.e. homogeneous and self-dual)
- ◇ A fourth cone $(K = \{0\}, K^* = \mathbb{R})$ used for modelling convenience
- ◇ Many problems from various domains (e.g. mechanical and eletrical engineering, finance) can be modelled using these cones
- ◇ Many solvers available for problems involving these cones
- ◇ However, no solver seems available (yet) for any other cone!

Examples of conic optimization

Following 3 cones are (by far) most commonly used

1. $K = \mathbb{R}_+$ is the standard ordering, leading to linear optimization
2. $K = \mathbb{L}^n$ leads to second-order cone optimization (including QCQO)

$$\mathbb{L}^n = \{(x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \mid \sqrt{x_1^2 + \cdots + x_n^2} \leq x_0\}$$

3. $K = \mathbb{S}^n_+$ (positive semidefinite matrices) for semidefinite optimization

- Those cones share additional theoretical properties (symmetric, i.e. homogeneous and self-dual)
- A fourth cone $(K = \{0\}, K^* = \mathbb{R})$ used for modelling convenience
- Many problems from various domains (e.g. mechanical and eletrical engineering, finance) can be modelled using these cones
- Many solvers available for problems involving these cones
- However, no solver seems available (yet) for any other cone!

Goal of this talk

How useful are those 3 cones with the convex problems mentioned earlier?

- Linear optimization is modelled using $\mathbb{R}_+$
- QCQO can be modelled using second-order cone $\mathbb{L}^n$
- l_p-norm and sum-of-norm optimization can be modelled directly with $\mathbb{L}^n$ when $p = 2$
- l_p-norm with rational p possible via construction involving several $\mathbb{L}^n$ (size of model increases with "complexity" of p)
- Geometric optimization can only be approximated using several $\mathbb{L}^n$ (size of model increases with accuracy required)
- Missing: sum-of-norm with $p \neq 2$, analytic centering, entropy optimization, irrational p (though some of these can probably be approximated using several $\mathbb{L}^n$)
- Missing: exact model for geometric optimization, simpler model for l_p-norm

Our aim: a single family of convex cones to model all of these exactly

Goal of this talk

How useful are those 3 cones with the convex problems mentioned earlier?

- Linear optimization is modelled using $\mathbb{R}_+$
- QCQO can be modelled using second-order cone $\mathbb{L}^n$
- l_p-norm and sum-of-norm optimization can be modelled directly with $\mathbb{L}^n$ when $p = 2$
- l_p-norm with rational p possible via construction involving several $\mathbb{L}^n$ (size of model increases with "complexity" of p)
- Geometric optimization can only be approximated using several $\mathbb{L}^n$ (size of model increases with accuracy required)
- Missing: sum-of-norm with $p \neq 2$, analytic centering, entropy optimization, irrational p (though some of these can probably be approximated using several $\mathbb{L}^n$)
- Missing: exact model for geometric optimization, simpler model for l_p-norm

Our aim: a single family of convex cones to model all of these exactly

Goal of this talk

How useful are those 3 cones with the convex problems mentioned earlier?

- Linear optimization is modelled using $\mathbb{R}_+$
- QCQO can be modelled using second-order cone $\mathbb{L}^n$
- l_p-norm and sum-of-norm optimization can be modelled directly with $\mathbb{L}^n$ when $p = 2$
- l_p-norm with rational p possible via construction involving several $\mathbb{L}^n$ (size of model increases with "complexity" of p)
- Geometric optimization can only be approximated using several $\mathbb{L}^n$ (size of model increases with accuracy required)
- Missing: sum-of-norm with $p \neq 2$, analytic centering, entropy optimization, irrational p (though some of these can probably be approximated using several $\mathbb{L}^n$)
- Missing: exact model for geometric optimization, simpler model for l_p-norm

Our aim: a single family of convex cones to model all of these exactly

Goal of this talk

How useful are those 3 cones with the convex problems mentioned earlier?

- Linear optimization is modelled using $\mathbb{R}_+$
- QCQO can be modelled using second-order cone $\mathbb{L}^n$
- l_p-norm and sum-of-norm optimization can be modelled directly with $\mathbb{L}^n$ when $p = 2$
- l_p-norm with rational p possible via construction involving several $\mathbb{L}^n$ (size of model increases with "complexity" of p)
- Geometric optimization can only be approximated using several $\mathbb{L}^n$ (size of model increases with accuracy required)
- Missing: sum-of-norm with $p \neq 2$, analytic centering, entropy optimization, irrational p (though some of these can probably be approximated using several $\mathbb{L}^n$)
- Missing: exact model for geometric optimization, simpler model for l_p-norm

Our aim: a single family of convex cones to model all of these exactly

Goal of this talk

How useful are those 3 cones with the convex problems mentioned earlier?

- Linear optimization is modelled using $\mathbb{R}_+$
- QCQO can be modelled using second-order cone $\mathbb{L}^n$
- l_p-norm and sum-of-norm optimization can be modelled directly with $\mathbb{L}^n$ when $p = 2$
- l_p-norm with rational p possible via construction involving several $\mathbb{L}^n$ (size of model increases with "complexity" of p)
- Geometric optimization can only be approximated using several $\mathbb{L}^n$ (size of model increases with accuracy required)
- Missing: sum-of-norm with $p \neq 2$, analytic centering, entropy optimization, irrational p (though some of these can probably be approximated using several $\mathbb{L}^n$)
- Missing: exact model for geometric optimization, simpler model for l_p-norm

Our aim: a single family of convex cones to model all of these exactly

Overview

1. Motivation

- ◇ Why convex optimization?
- ◇ Why a conic formulation?

2. Unified conic formulation

- ◇ The power cone
- ◇ Modelling problems involving powers using the power cone
- ◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

- ◇ Future plans

The power cone $\mathcal{P}_p$

Let us consider the epigraph of the convex power function

$$z \mapsto |z|^p \text{ with } p \geq 1 \quad \rightarrow \quad E_p = \{(x, z) \mid |z|^p \leq x\}$$

and take its conic hull: $(x, y, z) \in K_{E_p} \Leftrightarrow \frac{1}{y}(x, z) \in \mathcal{E}_p$

The resulting 3-dimensional cone will be called the power cone and denoted

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

(with the usual convention $\frac{1}{p} + \frac{1}{q} = 1$)

- $\mathcal{P}_p$ is equivalent to (rotated) $\mathbb{L}^3$ when $p = q = 2$
- $\mathcal{P}_p$ can be obtained from several $\mathbb{L}^n$ for rational values of p (number of cones required increases with "complexity" of p)

The power cone $\mathcal{P}_p$

Let us consider the epigraph of the convex power function

$$z \mapsto |z|^p \text{ with } p \geq 1 \quad \rightarrow \quad E_p = \{(x, z) \mid |z|^p \leq x\}$$

and take its conic hull: $(x, y, z) \in K_{E_p} \Leftrightarrow \frac{1}{y}(x, z) \in \mathcal{E}_p$

The resulting 3-dimensional cone will be called the power cone and denoted

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

(with the usual convention $\frac{1}{p} + \frac{1}{q} = 1$)

- $\mathcal{P}_p$ is equivalent to (rotated) $\mathbb{L}^3$ when $p = q = 2$
- $\mathcal{P}_p$ can be obtained from several $\mathbb{L}^n$ for rational values of p (number of cones required increases with "complexity" of p)

The power cone $\mathcal{P}_p$

Let us consider the epigraph of the convex power function

$$z \mapsto |z|^p \text{ with } p \geq 1 \quad \rightarrow \quad E_p = \{(x, z) \mid |z|^p \leq x\}$$

and take its conic hull: $(x, y, z) \in K_{E_p} \Leftrightarrow \frac{1}{y}(x, z) \in \mathcal{E}_p$

The resulting 3-dimensional cone will be called the power cone and denoted

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

(with the usual convention $\frac{1}{p} + \frac{1}{q} = 1$)

- $\mathcal{P}_p$ is equivalent to (rotated) $\mathbb{L}^3$ when $p = q = 2$
- $\mathcal{P}_p$ can be obtained from several $\mathbb{L}^n$ for rational values of p (number of cones required increases with "complexity" of p)

Duality

The dual of the power cone

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

is . . . the power cone itself !
(up to a simple scaling of the variables/use of a different inner product)
Power cone is therefore self-dual (but not homogeneous), and actual dual cone (for standard inner product) is

$$\mathcal{P}_p^* = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid (px)^{\frac{1}{p}} (qy)^{\frac{1}{q}} \geq |z|\}$$

Therefore the dual for any conic problem based on $\mathcal{P}_p$

- ◇ is a conic problem also based on $\mathcal{P}_p$
- ◇ can be derived in a completely mechanical way

Duality

The dual of the power cone

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

is . . . the power cone itself !

(up to a simple scaling of the variables/use of a different inner product)

Power cone is therefore self-dual (but not homogeneous), and actual dual cone (for standard inner product) is

$$\mathcal{P}_p^* = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid (px)^{\frac{1}{p}} (qy)^{\frac{1}{q}} \geq |z|\}$$

Therefore the dual for any conic problem based on $\mathcal{P}_p$

- ◇ is a conic problem also based on $\mathcal{P}_p$
- ◇ can be derived in a completely mechanical way

Duality

The dual of the power cone

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

is . . . the power cone itself !
(up to a simple scaling of the variables/use of a different inner product)
Power cone is therefore self-dual (but not homogeneous), and actual dual cone (for standard inner product) is

$$\mathcal{P}_p^* = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid (px)^{\frac{1}{p}} (qy)^{\frac{1}{q}} \geq |z|\}$$

Therefore the dual for any conic problem based on $\mathcal{P}_p$

- is a conic problem also based on $\mathcal{P}_p$
- can be derived in a completely mechanical way

Duality

The dual of the power cone

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

is ... the power cone itself !
(up to a simple scaling of the variables/use of a different inner product)
Power cone is therefore self-dual (but not homogeneous), and actual dual cone (for standard inner product) is

$$\mathcal{P}_p^* = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid (px)^{\frac{1}{p}} (qy)^{\frac{1}{q}} \geq |z|\}$$

Therefore the dual for any conic problem based on $\mathcal{P}_p$

- ◇ is a conic problem also based on $\mathcal{P}_p$
- ◇ can be derived in a completely mechanical way

Barrier function for $\mathcal{P}_p$

- A self-concordant barrier for

$$\mathcal{P}_p = \{(x,y,z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

with parameter $\nu = 4$ was proposed by NESTEROV

$$F_p(x,y,z) = -\log\left(x^{\frac{2}{p}} y^{\frac{2}{q}} - z^2\right) - \log x - \log y$$

- This can be improved: the following barrier

$$F_p(x,y,z) = -\log\left(x^{\frac{2}{p}} y^{\frac{2}{q}} - z^2\right) - \frac{1}{q}\log x - \frac{1}{p}\log y$$

is self-concordant with lower parameter $\nu = 3$

- This implies complexity of solving conic problems involving $\mathcal{P}_p$ depends only the number of cones (not on parameter p)

Barrier function for $\mathcal{P}_p$

- A self-concordant barrier for

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

with parameter $\nu = 4$ was proposed by Nesterov

$$F_p(x, y, z) = -\log\left(x^{\frac{2}{p}} y^{\frac{2}{q}} - z^2\right) - \log x - \log y$$

- This can be improved: the following barrier

$$F_p(x, y, z) = -\log\left(x^{\frac{2}{p}} y^{\frac{2}{q}} - z^2\right) - \frac{1}{q} \log x - \frac{1}{p} \log y$$

is self-concordant with lower parameter $\nu = 3$

- This implies complexity of solving conic problems involving $\mathcal{P}_p$ depends only the number of cones (not on parameter p)

Barrier function for $\mathcal{P}_p$

- A self-concordant barrier for

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

with parameter $\nu = 4$ was proposed by NESTEROV

$$F_p(x, y, z) = -\log\left(x^{\frac{2}{p}} y^{\frac{2}{q}} - z^2\right) - \log x - \log y$$

- This can be improved: the following barrier

$$F_p(x, y, z) = -\log\left(x^{\frac{2}{p}} y^{\frac{2}{q}} - z^2\right) - \frac{1}{q}\log x - \frac{1}{p}\log y$$

is self-concordant with lower parameter $\nu = 3$

- This implies complexity of solving conic problems involving $\mathcal{P}_p$ depends only the number of cones (not on parameter p)

Overview

Modelling with the power cone

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

- Linear optimization is OK, since $\mathcal{P}_p$ contains 2 copies of $\mathbb{R}_+$
- Epigraph of convex power $\{(x, z) \mid |z|^p \leq x\}$ can be obtained by fixing $y = 1$: in a dual conic formulation, this allows any l_p-norm optimization problem to be modelled
- For example, modelling the constraint

$$|u_1 + u_2|^3 + |u_1 - u_2|^{4.5} \leq 2u_2 + 1$$

will be done with

$$\begin{pmatrix}0\\1\\0\end{pmatrix} - \begin{pmatrix}0 & 0 & -1\\0 & 0 & 0\\-1 & -1 & 0\end{pmatrix}\begin{pmatrix}u_1\\u_2\\v\end{pmatrix} \in \mathcal{P}_3, \begin{pmatrix}1\\1\\0\end{pmatrix} - \begin{pmatrix}0 & -2 & 1\\0 & 0 & 0\\-1 & 1 & 0\end{pmatrix}\begin{pmatrix}u_1\\u_2\\v\end{pmatrix} \in \mathcal{P}_{4.5}$$

Modelling with the power cone

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

- Linear optimization is OK, since $\mathcal{P}_p$ contains 2 copies of $\mathbb{R}_+$
- Epigraph of convex power $\{(x, z) \mid |z|^p \leq x\}$ can be obtained by fixing $y = 1$: in a dual conic formulation, this allows any l_p-norm optimization problem to be modelled
- For example, modelling the constraint

$$|u_1 + u_2|^3 + |u_1 - u_2|^{4.5} \leq 2u_2 + 1$$

will be done with

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ v \end{pmatrix} \in \mathcal{P}_3, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ v \end{pmatrix} \in \mathcal{P}_{4.5}$$

Modelling with the power cone

$$\mathcal{P}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid x^{\frac{1}{p}} y^{\frac{1}{q}} \geq |z|\}$$

- Linear optimization is OK, since $\mathcal{P}_p$ contains 2 copies of $\mathbb{R}_+$
- Epigraph of convex power $\{(x, z) \mid |z|^p \leq x\}$ can be obtained by fixing $y = 1$: in a dual conic formulation, this allows any l_p-norm optimization problem to be modelled
- For example, modelling the constraint

$$|u_1 + u_2|^3 + |u_1 - u_2|^{4.5} \leq 2u_2 + 1$$

will be done with

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ v \end{pmatrix} \in \mathcal{P}_3, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ v \end{pmatrix} \in \mathcal{P}_{4.5}$$

$$\begin{pmatrix}0\\1\\0\end{pmatrix} - \begin{pmatrix}0 & 0 & -1\\0 & 0 & 0\\-1 & -1 & 0\end{pmatrix}\begin{pmatrix}u_1\\u_2\\v\end{pmatrix} \in \mathcal{P}_3, \begin{pmatrix}1\\1\\0\end{pmatrix} - \begin{pmatrix}0 & -2 & 1\\0 & 0 & 0\\-1 & 1 & 0\end{pmatrix}\begin{pmatrix}u_1\\u_2\\v\end{pmatrix} \in \mathcal{P}_{4.5}$$

$$\Leftrightarrow (v, 1, u_1 + u_2) \in \mathcal{P}_3 \text{ and } (1 + 2u_2 - v, 1, u_1 - u_2) \in \mathcal{P}_{4.5}$$

$$\Leftrightarrow |u_1 + u_2|^3 \leq v \text{ and } |u_1 - u_2|^{4.5} \leq 1 + 2u_2 - v$$

$$\Leftrightarrow |u_1 + u_2|^3 + |u_1 - u_2|^{4.5} \leq v + (1 + 2u_2 - v) = 2u_2 + 1$$

◇ Those are examples of constraints where nonlinearity is additively separable into scalar convex components

$$\begin{pmatrix}0\\1\\0\end{pmatrix} - \begin{pmatrix}0 & 0 & -1\\0 & 0 & 0\\-1 & -1 & 0\end{pmatrix}\begin{pmatrix}u_1\\u_2\\v\end{pmatrix} \in \mathcal{P}_3, \begin{pmatrix}1\\1\\0\end{pmatrix} - \begin{pmatrix}0 & -2 & 1\\0 & 0 & 0\\-1 & 1 & 0\end{pmatrix}\begin{pmatrix}u_1\\u_2\\v\end{pmatrix} \in \mathcal{P}_{4.5}$$

$$\Leftrightarrow (v, 1, u_1 + u_2) \in \mathcal{P}_3 \text{ and } (1 + 2u_2 - v, 1, u_1 - u_2) \in \mathcal{P}_{4.5}$$

$$\Leftrightarrow |u_1 + u_2|^3 \leq v \text{ and } |u_1 - u_2|^{4.5} \leq 1 + 2u_2 - v$$

$$\Leftrightarrow |u_1 + u_2|^3 + |u_1 - u_2|^{4.5} \leq v + (1 + 2u_2 - v) = 2u_2 + 1$$

◇ Those are examples of constraints where nonlinearity is additively separable into scalar convex components

Modelling non-separable nonlinearity

- Norm constraint

$$\|z\|_p \leq t \quad \Leftrightarrow \quad |z_1|^p + |z_2|^p \cdots |z_n|^p \leq t^p \text{ and } t \geq 0$$

(with $p \geq 1$) is not separable and cannot be formulated using epigraphs of convex powers $|z_i|^p \leq x$

- We use the following trick:

$$\begin{aligned} & |z_1|^p + |z_2|^p \cdots |z_n|^p \leq t^p \\ \Leftrightarrow \quad & |\frac{z_1}{t}|^p + |\frac{z_2}{t}|^p \cdots |\frac{z_n}{t}|^p \leq 1 \\ \Leftrightarrow \quad & |\frac{z_1}{t}|^p \leq \frac{x_1}{t}, \ \cdots, \ |\frac{z_n}{t}|^p \leq \frac{x_n}{t} \text{ and } \frac{x_1}{t} + \cdots + \frac{x_n}{t} = 1 \\ \Leftrightarrow \quad & (x_1, t, y_1) \in \mathcal{P}_p, \ \cdots, \ (x_n, t, y_n) \in \mathcal{P}_p \text{ and } x_1 + \cdots + x_n = t \end{aligned}$$

which can be modelled using conic optimization

Modelling non-separable nonlinearity

- Norm constraint

$$\|z\|_p \le t \quad \Leftrightarrow \quad |z_1|^p + |z_2|^p \cdots |z_n|^p \le t^p \text{ and } t \ge 0$$

(with $p \ge 1$) is not separable and cannot be formulated using epigraphs of convex powers $|z_i|^p \le x$

- We use the following trick:

$$\begin{aligned} & |z_1|^p + |z_2|^p \cdots |z_n|^p \le t^p \\ \Leftrightarrow \quad & |\frac{z_1}{t}|^p + |\frac{z_2}{t}|^p \cdots |\frac{z_n}{t}|^p \le 1 \\ \Leftrightarrow \quad & |\frac{z_1}{t}|^p \le \frac{x_1}{t}, \ \cdots, \ |\frac{z_n}{t}|^p \le \frac{x_n}{t} \text{ and } \frac{x_1}{t} + \cdots + \frac{x_n}{t} = 1 \\ \Leftrightarrow \quad & (x_1, t, y_1) \in \mathcal{P}_p, \ \cdots, \ (x_n, t, y_n) \in \mathcal{P}_p \text{ and } x_1 + \cdots + x_n = t \end{aligned}$$

which can be modelled using conic optimization

Norm constraints (continued)

- Using the previous construction, we can model any constraint $\|z\|_p \leq t$ for $p \geq 1$
- We can therefore model sum-of-norm optimization problems
- When $p = 2$, this is the second-order cone: we can therefore also model all second-order cone optimization problems, including QCQO problems
- We can also model more complicated non-separable expressions, such as

$$z_1^{2.5} z_2^{-4.5} z_3^4 + 2z_1 z_2 + z_3^4 \leq t^2$$

(crucial condition for convexity is that degree of every term on l.h.s. should be greater that degree of r.h.s.)

- This looks similar to posynomials involved in geometric optimization but is completely different (no exponential transformation)

Norm constraints (continued)

- Using the previous construction, we can model any constraint $\|z\|_p \leq t$ for $p \geq 1$
- We can therefore model sum-of-norm optimization problems
- When $p = 2$, this is the second-order cone: we can therefore also model all second-order cone optimization problems, including QCQO problems
- We can also model more complicated non-separable expressions, such as

$$z_1^{2.5} z_2^{-4.5} z_3^4 + 2 z_1 z_2 + z_3^4 \leq t^2$$

(crucial condition for convexity is that degree of every term on l.h.s. should be greater that degree of r.h.s.)
- This looks similar to posynomials involved in geometric optimization but is completely different (no exponential transformation)

Norm constraints (continued)

- Using the previous construction, we can model any constraint $\|z\|_p \leq t$ for $p \geq 1$
- We can therefore model sum-of-norm optimization problems
- When $p = 2$, this is the second-order cone: we can therefore also model all second-order cone optimization problems, including QCQO problems
- We can also model more complicated non-separable expressions, such as

$$z_1^{2.5} z_2^{-4.5} z_3^4 + 2 z_1 z_2 + z_3^4 \leq t^2$$

  (crucial condition for convexity is that degree of every term on l.h.s. should be greater that degree of r.h.s.)
- This looks similar to posynomials involved in geometric optimization but is completely different (no exponential transformation)

Overview

What about exponentials?

Some convex problem classes described earlier are still missing, among which the exponential function seems to play a central role

We use the following well-known limit

$$\lim_{p \to +\infty} (1 + \frac{x}{p})^p = e^x ,$$

valid for any real x, to obtain the exponential function. Letting $z = y + \frac{z'}{p}$ (a linear transformation), the definition of the power cone becomes:

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \geq \left| y + \frac{z'}{p} \right| \;\Leftrightarrow\; x^{\frac{1}{p}} y^{(\frac{1}{q}-1)} \geq \left| 1 + \frac{z'/y}{p} \right| \Leftrightarrow xy^{-1} \geq \left| 1 + \frac{1}{p}\frac{z'}{y} \right|^p$$

$$\Leftrightarrow \; \left| 1 + \frac{1}{p}\left(\frac{z'}{y}\right) \right|^p \leq \frac{x}{y}$$

What about exponentials?

Some convex problem classes described earlier are still missing, among which the exponential function seems to play a central role

We use the following well-known limit

$$\lim_{p\to+\infty}(1+\frac{x}{p})^p = e^x \ ,$$

valid for any real x, to obtain the exponential function. Letting $z = y + \frac{z'}{p}$ (a linear transformation), the definition of the power cone becomes:

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \geq \left| y + \frac{z'}{p} \right| \quad \Leftrightarrow \quad x^{\frac{1}{p}} y^{(\frac{1}{q}-1)} \geq \left| 1 + \frac{z'/y}{p} \right| \Leftrightarrow xy^{-1} \geq \left| 1 + \frac{1}{p}\frac{z'}{y} \right|^p$$

$$\Leftrightarrow \quad \left| 1 + \frac{1}{p}\left(\frac{z'}{y}\right) \right|^p \leq \frac{x}{y}$$

What about exponentials?

Some convex problem classes described earlier are still missing, among which the exponential function seems to play a central role

We use the following well-known limit

$$\lim_{p \to +\infty} (1 + \frac{x}{p})^p = e^x \ ,$$

valid for any real x, to obtain the exponential function. Letting $z = y + \frac{z'}{p}$ (a linear transformation), the definition of the power cone becomes:

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \geq \left| y + \frac{z'}{p} \right| \ \Leftrightarrow \ x^{\frac{1}{p}} y^{(\frac{1}{q}-1)} \geq \left| 1 + \frac{z'/y}{p} \right| \Leftrightarrow xy^{-1} \geq \left| 1 + \frac{1}{p} \frac{z'}{y} \right|^p$$

$$\Leftrightarrow \ \left| 1 + \frac{1}{p} \left(\frac{z'}{y} \right) \right|^p \leq \frac{x}{y}$$

What about exponentials?

Some convex problem classes described earlier are still missing, among which the exponential function seems to play a central role

We use the following well-known limit

$$\lim_{p \to +\infty} (1 + \frac{x}{p})^p = e^x \ ,$$

valid for any real x, to obtain the exponential function. Letting $z = y + \frac{z'}{p}$ (a linear transformation), the definition of the power cone becomes:

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \geq \left| y + \frac{z'}{p} \right| \ \Leftrightarrow \ x^{\frac{1}{p}} y^{(\frac{1}{q}-1)} \geq \left| 1 + \frac{z'/y}{p} \right| \Leftrightarrow xy^{-1} \geq \left| 1 + \frac{1}{p}\frac{z'}{y} \right|^p$$

$$\Leftrightarrow \ \left| 1 + \frac{1}{p}\left(\frac{z'}{y}\right) \right|^p \leq \frac{x}{y}$$

Letting now p tend to $+\infty$ we obtain

$$\left|1 + \frac{1}{p}\Big(\frac{z'}{y}\Big)\right|^p \leq \frac{x}{y} \quad \longrightarrow_{p\to\infty} \quad \exp\Big(\frac{z'}{y}\Big) \leq \frac{x}{y}$$

which defines the exponential cone:

$$\mathcal{E}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid \exp\Big(\frac{z}{y}\Big) \leq \frac{x}{y}\}$$

- We can now model the epigraph of exponential function (take $y = 1$) and therefore geometric optimization
- We can also model the epigraph of minus logarithm: $(x, 1, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z) \leq x \Leftrightarrow -\log x \leq z$ and therefore analytic centering
- We can also model the epigraph of entropy: $(1, y, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z/y) \leq 1/y \Leftrightarrow -z/y \leq -\log y \Leftrightarrow y \log y \leq z$ and therefore entropy optimization

Letting now p tend to $+\infty$ we obtain

$$\left|1 + \frac{1}{p}\left(\frac{z'}{y}\right)\right|^p \leq \frac{x}{y} \quad \longrightarrow_{p\to\infty} \quad \exp\left(\frac{z'}{y}\right) \leq \frac{x}{y}$$

which defines the exponential cone:

$$\mathcal{E}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid \exp\left(\frac{z}{y}\right) \leq \frac{x}{y}\}$$

- We can now model the epigraph of exponential function (take $y = 1$) and therefore geometric optimization
- We can also model the epigraph of minus logarithm: $(x, 1, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z) \leq x \Leftrightarrow -\log x \leq z$ and therefore analytic centering
- We can also model the epigraph of entropy: $(1, y, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z/y) \leq 1/y \Leftrightarrow -z/y \leq -\log y \Leftrightarrow y \log y \leq z$ and therefore entropy optimization

Letting now p tend to $+\infty$ we obtain

$$\left|1 + \frac{1}{p}\left(\frac{z'}{y}\right)\right|^p \leq \frac{x}{y} \quad \longrightarrow_{p\to\infty} \quad \exp\left(\frac{z'}{y}\right) \leq \frac{x}{y}$$

which defines the exponential cone:

$$\mathcal{E}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid \exp\left(\frac{z}{y}\right) \leq \frac{x}{y}\}$$

- We can now model the epigraph of exponential function (take $y = 1$) and therefore geometric optimization
- We can also model the epigraph of minus logarithm: $(x, 1, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z) \leq x \Leftrightarrow -\log x \leq z$ and therefore analytic centering
- We can also model the epigraph of entropy: $(1, y, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z/y) \leq 1/y \Leftrightarrow -z/y \leq -\log y \Leftrightarrow y \log y \leq z$ and therefore entropy optimization

Letting now p tend to $+\infty$ we obtain

$$\left|1 + \frac{1}{p}\Big(\frac{z'}{y}\Big)\right|^p \leq \frac{x}{y} \quad \longrightarrow_{p\to\infty} \quad \exp\Big(\frac{z'}{y}\Big) \leq \frac{x}{y}$$

which defines the exponential cone:

$$\mathcal{E}_p = \{(x, y, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid \exp\Big(\frac{z}{y}\Big) \leq \frac{x}{y}\}$$

- We can now model the epigraph of exponential function (take $y = 1$) and therefore geometric optimization
- We can also model the epigraph of minus logarithm:
  $(x, 1, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z) \leq x \Leftrightarrow -\log x \leq z$
  and therefore analytic centering
- We can also model the epigraph of entropy: $(1, y, -z) \in \mathcal{E}_p \Leftrightarrow \exp(-z/y) \leq 1/y \Leftrightarrow -z/y \leq -\log y \Leftrightarrow y \log y \leq z$ and therefore entropy optimization

Barrier for exponential cone

Exponential cone is the limit of a suitably linearly transformed power cone
Since self-concordancy is preserved by linear transformations and limits, we should be able to easily compute a self-concordant barrier for $\mathcal{E}_p$

$$F_{\exp}(x,y,z) = \lim_{p \to +\infty} F_p(x,y,y+z/p) = \cdots = +\infty$$

Is something wrong? Can be corrected by adding a missing constant term:

$$\begin{aligned} F_{\exp}(x,y,z) &= \lim_{p \to +\infty} \left(F_p(x,y,y+z/p) - \log \frac{p}{2} \right) \\ &= \cdots = -\log\big(y \log(x/y) - z\big) - \log(x) - \log(y) \end{aligned}$$

Using this unified barrier, one can solve any conic problem involving exponential cones $\mathcal{E}_p$
(and recompute the standard barriers for exponential, minus logarithm and entropy epigraphs)

Barrier for exponential cone

Exponential cone is the limit of a suitably linearly transformed power cone
Since self-concordancy is preserved by linear transformations and limits, we should be able to easily compute a self-concordant barrier for $\mathcal{E}_p$

$$F_{\exp}(x,y,z) = \lim_{p\to+\infty} F_p(x,y,y+z/p) = \cdots = +\infty$$

Is something wrong? Can be corrected by adding a missing constant term:

$$\begin{aligned} F_{\exp}(x,y,z) &= \lim_{p\to+\infty}\Big(F_p(x,y,y+z/p) - \log\frac{p}{2}\Big) \\ &= \cdots = -\log\big(y\log(x/y) - z\big) - \log(x) - \log(y) \end{aligned}$$

Using this unified barrier, one can solve any conic problem involving exponential cones $\mathcal{E}_p$
(and recompute the standard barriers for exponential, minus logarithm and entropy epigraphs)

Barrier for exponential cone

Exponential cone is the limit of a suitably linearly transformed power cone
Since self-concordancy is preserved by linear transformations and limits, we should be able to easily compute a self-concordant barrier for $\mathcal{E}_p$

$$F_{\exp}(x, y, z) = \lim_{p \to +\infty} F_p(x, y, y + z/p) = \cdots = +\infty$$

Is something wrong? Can be corrected by adding a missing constant term:

$$\begin{aligned} F_{\exp}(x, y, z) &= \lim_{p \to +\infty} \Big(F_p(x, y, y + z/p) - \log \frac{p}{2} \Big) \\ &= \cdots = -\log\big(y \log(x/y) - z\big) - \log(x) - \log(y) \end{aligned}$$

Using this unified barrier, one can solve any conic problem involving exponential cones $\mathcal{E}_p$
(and recompute the standard barriers for exponential, minus logarithm and entropy epigraphs)

Combining different types of constraints

Part of the usefulness of this framework is that it allows combinations of different types of constraints in a completely seamless way

An example: the Lambert W function, defined by $W(x) \exp W(x) = x$

From *MathWorld*: Banwell and Jayakumar (2000) showed that a W-function describes the relation between voltage, current and resistance in a diode, and Packel and Yuen (2004) applied the W-function to a ballistic projectile in the presence of air resistance. Other applications have been discovered in statistical mechanics, quantum chemistry, combinatorics, enzyme kinetics, the physiology of vision, the engineering of thin films, hydrology, and the analysis of algorithms (Hayes 2005).

Combining different types of constraints

Part of the usefulness of this framework is that it allows combinations of different types of constraints in a completely seamless way

An example: the Lambert W function, defined by $W(x) \exp W(x) = x$

From *MathWorld*: Banwell and Jayakumar (2000) showed that a W-function describes the relation between voltage, current and resistance in a diode, and Packel and Yuen (2004) applied the W-function to a ballistic projectile in the presence of air resistance. Other applications have been discovered in statistical mechanics, quantum chemistry, combinatorics, enzyme kinetics, the physiology of vision, the engineering of thin films, hydrology, and the analysis of algorithms (Hayes 2005).

An example: the Lambert W function

$W(x)$ is real for $x \geq 0$, and concave on that domain ; therefore, we can try to model the convex set defined by $0 \leq y \leq W(x)$ (intersection of its hypograph with nonnegative orthant)

$$0 \leq y \leq W(x) \Leftrightarrow 0 \leq y \exp y \leq W(x) \exp W(x) \Leftrightarrow 0 \leq y \exp y \leq x$$

which can be obtained using

- a exponential constraint $\exp\left(\frac{z}{y}\right) \leq \frac{x}{y}$ and
- a quadratic constraint $z \geq y^2$

Indeed, we can check that

$$0 \leq y \exp y = y \exp(y^2/y) \leq y \exp(z/y) \leq x$$

In summary, combining a quadratic and an exponential constraint, we have shown that

$$0 \leq y \leq W(x) \Leftrightarrow (x, y, z) \in \mathcal{E}_p \text{ and } (z, 1, x) \in \mathcal{P}_2$$

An example: the Lambert W function

$W(x)$ is real for $x \geq 0$, and concave on that domain ; therefore, we can try to model the convex set defined by $0 \leq y \leq W(x)$ (intersection of its hypograph with nonnegative orthant)

$$0 \leq y \leq W(x) \Leftrightarrow 0 \leq y \exp y \leq W(x) \exp W(x) \Leftrightarrow 0 \leq y \exp y \leq x$$

which can be obtained using

- a exponential constraint $\exp\left(\frac{z}{y}\right) \leq \frac{x}{y}$ and
- a quadratic constraint $z \geq y^2$

Indeed, we can check that

$$0 \leq y \exp y = y \exp(y^2/y) \leq y \exp(z/y) \leq x$$

In summary, combining a quadratic and an exponential constraint, we have shown that

$$0 \leq y \leq W(x) \Leftrightarrow (x, y, z) \in \mathcal{E}_p \text{ and } (z, 1, x) \in \mathcal{P}_2$$

An example: the Lambert W function

$W(x)$ is real for $x \geq 0$, and concave on that domain ; therefore, we can try to model the convex set defined by $0 \leq y \leq W(x)$ (intersection of its hypograph with nonnegative orthant)

$$0 \leq y \leq W(x) \Leftrightarrow 0 \leq y \exp y \leq W(x) \exp W(x) \Leftrightarrow 0 \leq y \exp y \leq x$$

which can be obtained using

- a exponential constraint $\exp\left(\frac{z}{y}\right) \leq \frac{x}{y}$ and
- a quadratic constraint $z \geq y^2$

Indeed, we can check that

$$0 \leq y \exp y = y \exp(y^2/y) \leq y \exp(z/y) \leq x$$

In summary, combining a quadratic and an exponential constraint, we have shown that

$$0 \leq y \leq W(x) \Leftrightarrow (x, y, z) \in \mathcal{E}_p \text{ and } (z, 1, x) \in \mathcal{P}_2$$

An example: the Lambert W function

$W(x)$ is real for $x \geq 0$, and concave on that domain ; therefore, we can try to model the convex set defined by $0 \leq y \leq W(x)$ (intersection of its hypograph with nonnegative orthant)

$$0 \leq y \leq W(x) \Leftrightarrow 0 \leq y \exp y \leq W(x) \exp W(x) \Leftrightarrow 0 \leq y \exp y \leq x$$

which can be obtained using

- a exponential constraint $\exp\left(\frac{z}{y}\right) \leq \frac{x}{y}$ and
- a quadratic constraint $z \geq y^2$

Indeed, we can check that

$$0 \leq y \exp y = y \exp(y^2/y) \leq y \exp(z/y) \leq x$$

In summary, combining a quadratic and an exponential constraint, we have shown that

$$0 \leq y \leq W(x) \Leftrightarrow (x, y, z) \in \mathcal{E}_p \text{ and } (z, 1, x) \in \mathcal{P}_2$$

Overview

Concluding remarks

In conclusion, the single family of 3-dimensional cones $\mathcal{P}_p$ can

- model a very large class of structured convex problems (with the notable exception of semidefinite optimization)
- enable their resolution with powerful interior-point methods
- allow the easy computation of their dual problems

Convex problems covered include linear, quadratic, second-order cone, quadratically constrained, geometric, l_p-norm, sum-of-norm, entropy optimization and others, as well as any combinations of these

Potential drawback: conic modelling sometimes require the introduction of a large number of additional variables (e.g. $\|x\|_p \leq t$ constraint)

Concluding remarks

In conclusion, the single family of 3-dimensional cones $\mathcal{P}_p$ can

- model a very large class of structured convex problems (with the notable exception of semidefinite optimization)
- enable their resolution with powerful interior-point methods
- allow the easy computation of their dual problems

Convex problems covered include linear, quadratic, second-order cone, quadratically constrained, geometric, l_p-norm, sum-of-norm, entropy optimization and others, as well as any combinations of these

Potential drawback: conic modelling sometimes require the introduction of a large number of additional variables (e.g. $\|x\|_p \leq t$ constraint)

Concluding remarks

In conclusion, the single family of 3-dimensional cones $\mathcal{P}_p$ can

- model a very large class of structured convex problems (with the notable exception of semidefinite optimization)
- enable their resolution with powerful interior-point methods
- allow the easy computation of their dual problems

Convex problems covered include linear, quadratic, second-order cone, quadratically constrained, geometric, l_p-norm, sum-of-norm, entropy optimization and others, as well as any combinations of these

Potential drawback: conic modelling sometimes require the introduction of a large number of additional variables (e.g. $\|x\|_p \leq t$ constraint)

Concluding remarks

In conclusion, the single family of 3-dimensional cones $\mathcal{P}_p$ can

- model a very large class of structured convex problems (with the notable exception of semidefinite optimization)
- enable their resolution with powerful interior-point methods
- allow the easy computation of their dual problems

Convex problems covered include linear, quadratic, second-order cone, quadratically constrained, geometric, l_p-norm, sum-of-norm, entropy optimization and others, as well as any combinations of these

Potential drawback: conic modelling sometimes require the introduction of a large number of additional variables (e.g. $\|x\|_p \leq t$ constraint)

Concluding remarks

In conclusion, the single family of 3-dimensional cones $\mathcal{P}_p$ can

- model a very large class of structured convex problems (with the notable exception of semidefinite optimization)
- enable their resolution with powerful interior-point methods
- allow the easy computation of their dual problems

Convex problems covered include linear, quadratic, second-order cone, quadratically constrained, geometric, l_p-norm, sum-of-norm, entropy optimization and others, as well as any combinations of these

Potential drawback: conic modelling sometimes require the introduction of a large number of additional variables (e.g. $\|x\|_p \leq t$ constraint)

Concluding remarks

In conclusion, the single family of 3-dimensional cones $\mathcal{P}_p$ can

- model a very large class of structured convex problems (with the notable exception of semidefinite optimization)
- enable their resolution with powerful interior-point methods
- allow the easy computation of their dual problems

Convex problems covered include linear, quadratic, second-order cone, quadratically constrained, geometric, l_p-norm, sum-of-norm, entropy optimization and others, as well as any combinations of these

Potential drawback: conic modelling sometimes require the introduction of a large number of additional variables (e.g. $\|x\|_p \leq t$ constraint)

Conic solver

joint work with Robert Chares, CORE

- Our current (working) implementation relies on a dual long-step path-following interior-point algorithm
- It currently handles the nonnegative cone $\mathbb{R}^n_+$, the power cone $\mathcal{P}_p$, the exponential cone $\mathcal{E}_p$ and the $\mathbb{R}^n/\{0\}$ cones for better handling of primal free variable/dual equality constraints
- In the future: take advantage of self-duality and implement a primal-dual interior-point method

Conic solver

joint work with Robert Chares, CORE

- Our current (working) implementation relies on a dual long-step path-following interior-point algorithm
- It currently handles the nonnegative cone $\mathbb{R}^n_+$, the power cone $\mathcal{P}_p$, the exponential cone $\mathcal{E}_p$ and the $\mathbb{R}^n/\{0\}$ cones for better handling of primal free variable/dual equality constraints
- In the future: take advantage of self-duality and implement a primal-dual interior-point method

Conic solver

joint work with Robert Chares, CORE

- Our current (working) implementation relies on a dual long-step path-following interior-point algorithm
- It currently handles the nonnegative cone $\mathbb{R}^n_+$, the power cone $\mathcal{P}_p$, the exponential cone $\mathcal{E}_p$ and the $\mathbb{R}^n/\{0\}$ cones for better handling of primal free variable/dual equality constraints
- In the future: take advantage of self-duality and implement a primal-dual interior-point method

Conic solver

joint work with Robert Chares, CORE

- Our current (working) implementation relies on a dual long-step path-following interior-point algorithm
- It currently handles the nonnegative cone $\mathbb{R}^n_+$, the power cone $\mathcal{P}_p$, the exponential cone $\mathcal{E}_p$ and the $\mathbb{R}^n/\{0\}$ cones for better handling of primal free variable/dual equality constraints
- In the future: take advantage of self-duality and implement a primal-dual interior-point method

Overview

1. Motivation

◇ Why convex optimization?

◇ Why a conic formulation?

2. Unified conic formulation

◇ The power cone

◇ Modelling problems involving powers using the power cone

◇ Modelling problems involving exponentials using the power cone

3. Concluding remarks

◇ Future plans

Modelling language

- Even if a large class of problems can be modelled with our cones, writing down the equivalent formulation can be long and error-prone
- However, the procedure of converting a given problem problem into its equivalent conic formulation is completely understood and can be carried out completely automatically
- Our medium-term goal is to add an additional software layer around our solver to handle automatically all known modelling tricks for convex problems involving our cones (similar to current environments such as YALMIP, CVX or CVXOPT, which currently work only with standard cones $\mathbb{R}^n_+$, $\mathbb{L}^n$, $\mathbb{S}^n_+$ and $\mathbb{R}^n$)
- Stay tuned !
- If you have any problem that you think could be modelled in our framework, please do not hesitate to contact us!

Modelling language

- Even if a large class of problems can be modelled with our cones, writing down the equivalent formulation can be long and error-prone
- However, the procedure of converting a given problem problem into its equivalent conic formulation is completely understood and can be carried out completely automatically
- Our medium-term goal is to add an additional software layer around our solver to handle automatically all known modelling tricks for convex problems involving our cones (similar to current environments such as YALMIP, CVX or CVXOPT, which currently work only with standard cones $\mathbb{R}^n_+$, $\mathbb{L}^n$, $\mathbb{S}^n_+$ and $\mathbb{R}^n$)
- Stay tuned !
- If you have any problem that you think could be modelled in our framework, please do not hesitate to contact us!

Modelling language

- Even if a large class of problems can be modelled with our cones, writing down the equivalent formulation can be long and error-prone
- However, the procedure of converting a given problem problem into its equivalent conic formulation is completely understood and can be carried out completely automatically
- Our medium-term goal is to add an additional software layer around our solver to handle automatically all known modelling tricks for convex problems involving our cones (similar to current environments such as YALMIP, CVX or CVXOPT, which currently work only with standard cones $\mathbb{R}^n_+$, $\mathbb{L}^n$, $\mathbb{S}^n_+$ and $\mathbb{R}^n$)
- Stay tuned !
- If you have any problem that you think could be modelled in our framework, please do not hesitate to contact us!

Modelling language

- Even if a large class of problems can be modelled with our cones, writing down the equivalent formulation can be long and error-prone
- However, the procedure of converting a given problem problem into its equivalent conic formulation is completely understood and can be carried out completely automatically
- Our medium-term goal is to add an additional software layer around our solver to handle automatically all known modelling tricks for convex problems involving our cones (similar to current environments such as YALMIP, CVX or CVXOPT, which currently work only with standard cones $\mathbb{R}^n_+$, $\mathbb{L}^n$, $\mathbb{S}^n_+$ and $\mathbb{R}^n$)
- Stay tuned !
- If you have any problem that you think could be modelled in our framework, please do not hesitate to contact us!

Modelling language

- ◇ Even if a large class of problems can be modelled with our cones, writing down the equivalent formulation can be long and error-prone
- ◇ However, the procedure of converting a given problem problem into its equivalent conic formulation is completely understood and can be carried out completely automatically
- ◇ Our medium-term goal is to add an additional software layer around our solver to handle automatically all known modelling tricks for convex problems involving our cones (similar to current environments such as YALMIP, CVX or CVXOPT, which currently work only with standard cones $\mathbb{R}^n_+$, $\mathbb{L}^n$, $\mathbb{S}^n_+$ and $\mathbb{R}^n$)
- ◇ Stay tuned !
- ◇ If you have any problem that you think could be modelled in our framework, please do not hesitate to contact us!

Thank you for your attention!