Conic optimization

Aalborg University, June 26th, 2017

Joachim Dahl

www.mosek.com

Section 1

Linear optimization

Linear optimization

- We minimize a linear function given linear constraints.
- Example: minimize a linear function

$$x_1 + 2x_2 - x_3$$

under the constraints that

$$x_1 + x_2 + x_3 = 1, \quad x_1, x_2, x_3 \geq 0.$$

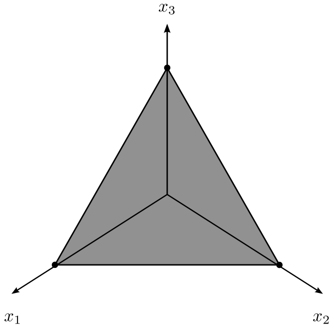
- The function we minimize is called the *objective* function.
- The constraitns are either *equality* or *inequality* constraints.
- **Important:** everything is *linear* in x.

Linear optimization
A simple example

Standard notation:

$$\begin{array}{ll} \text{minimize} & x_1 + 2x_2 - x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0. \end{array}$$

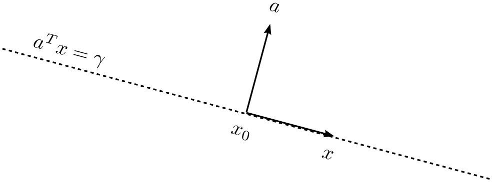
Feasible set:



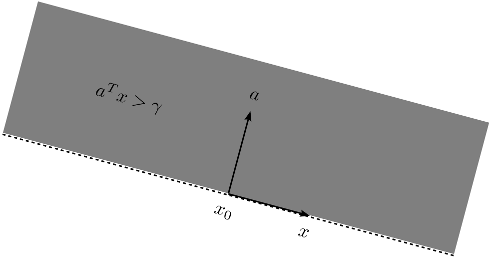
Optimal solution $x^\star = (0, 0, 1)$ with value$=-1$.

Hyperplanes and halfspaces

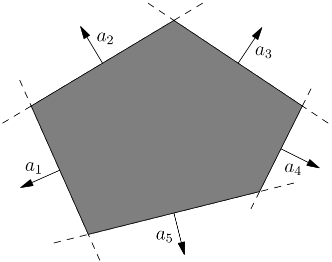
- Hyperplane: $\{x | a^T(x - x_0) = 0\} = \{x \mid a^T x = \gamma\}$



- Halfspace: $\{x | a^T(x - x_0) \geq 0\} = \{x \mid a^T x \geq \gamma\}$

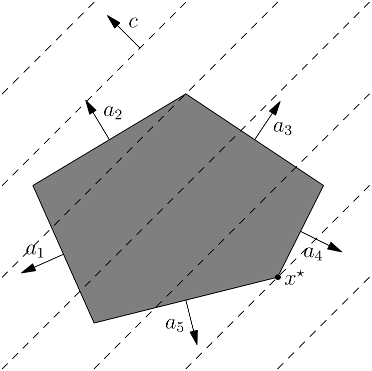


- A polyhedron is an intersection of halfspaces:



- Can be both bounded (as shown) or unbounded.

- The contour lines are shifted hyperplanes

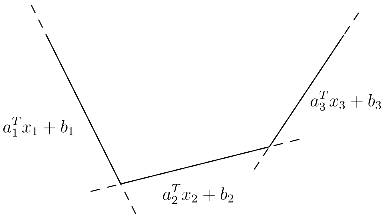


- Optimal solution is a vertex, on a facet, or unbounded.

Convex piecewise-linear functions

Consider f defined as a the maximum of affine functions,

$$f(x) := \max_{i=1,\ldots,m} \{a_i^T x + b_i\}.$$



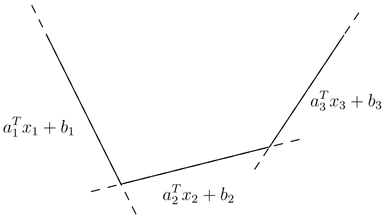
The *epigraph* $f(x) \leq t$ is equivalent to

$$a_i^T x + b_i \leq t, \quad i = 1, \ldots, m.$$

Convex piecewise-linear functions

Consider f defined as a the maximum of affine functions,

$$f(x) := \max_{i=1,\ldots,m} \{a_i^T x + b_i\}.$$



The *epigraph* $f(x) \leq t$ is equivalent to

$$a_i^T x + b_i \leq t, \quad i = 1, \ldots, m.$$

- The absolute value function

$$|\alpha| := \max\{\alpha, -\alpha\}$$

is a convex piecewise-linear function,

$$|\alpha| \leq t \quad \iff \quad -t \leq \alpha \leq t.$$

- The ℓ_∞-norm of a vector $x \in \mathbb{R}^n$ is

$$\|x\|_\infty := \max_{i=1,\dots,n} |x_i|,$$

i.e.,

$$\|x\|_\infty \leq t \quad \iff \quad -t \leq x_i \leq t,\ i = 1, \dots, n.$$

- The absolute value function

$$|\alpha| := \max\{\alpha, -\alpha\}$$

is a convex piecewise-linear function,

$$|\alpha| \leq t \quad \Longleftrightarrow \quad -t \leq \alpha \leq t.$$

- The ℓ_∞-norm of a vector $x \in \mathbb{R}^n$ is

$$\|x\|_\infty := \max_{i=1,\ldots,n} |x_i|,$$

i.e.,

$$\|x\|_\infty \leq t \quad \Longleftrightarrow \quad -t \leq x_i \leq t, \ i = 1, \ldots, n.$$

The ℓ_1-norm of a vector $x \in \mathbb{R}^n$ is

$$\|x\|_1 := |x_1| + |x_2| + \cdots + |x_n|.$$

We can characterize the epigraph

$$\|x\|_1 \leq t$$

as

$$|x_i| \leq z_i, \; i = 1, \ldots, n, \quad \sum_i z_i \leq t.$$

Programming exercises

Given data

```
m=500; n=100;
A=randn(m,n); b=randn(m,1);
```

write a Yalmip program that minimizes $f_i(Ax - b)$ for

1. $f_1(z) = \|z\|_1$
2. $f_2(z) = \|z\|_2$
3. $f_3(z) = \|z\|_\infty$
4. $f_4(z) = \sum_i \max\{0, z_i - 1, -z_i - 1\}$

Plot a histogram comparing $Ax - b$ for the different choices of f.

Duality in linear optimization

We consider a problem in standard form

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

The *Lagrangian* function is a lower bound,

$$L(x, y, s) = c^T x + y^T (b - Ax) - s^T x \leq c^T x$$

where $y \in \mathbb{R}^m$ and $s \in R^n_+$ are *Lagrange multipliers* or *dual variables*.

Note: It's important that $s \geq 0$.

The *dual function* is

$$g(y, s) = \inf_x L(x, y, s) = \inf_x x^T(c - A^T y - s) + b^T y,$$

i.e.,

$$g(y, s) = \begin{cases} b^T y, & c - A^T y - s = 0 \\ -\infty, & \text{otherwise,} \end{cases}$$

which is a global lower bound (valid for all x).

The *dual problem* is the best such lower bound,

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array}$$

Primal problem with optimal value $p^\star$:

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

Dual problem with optimal value $d^\star$:

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array}$$

Weak duality:

$$c^T x - b^T y = x^T (c - A^T y) = x^T s \geq 0,$$

i.e., $p^\star \geq d^\star$.

Convention:

- $p^\star = \infty$ if primal problem is infeasible.
- $d^\star = -\infty$ if dual problem is infeasible.

We then have:

- Primal feasible, dual feasible: $p^\star = d^\star$ and finite.
- Primal infeasible, dual unbounded: $p^\star = \infty, d^\star = \infty$.
- Primal unbounded, dual infeasible: $p^\star = -\infty, d^\star = -\infty$.
- Primal infeasible, dual infeasble: $p^\star = \infty, d^\star = -\infty$.

Only in the last case is $p^\star > d^\star$.

Duality in linear optimization
Example: basis pursuit

Basis pursuit problem:

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & Ax = b. \end{array}$$

Used as heuristic for sparse representation of b.

Equivalent linear problem:

$$\begin{array}{ll} \text{minimize} & e^T z \\ \text{subject to} & Ax = b \\ & -z \leq x \leq z. \end{array}$$

By change of variables

$$u = \frac{1}{2}(z - x), \qquad v = \frac{1}{2}(z + x)$$

we get a standard form linear problem:

$$\begin{array}{ll} \text{minimize} & e^T(v + u) \\ \text{subject to} & A(v - u) = b \\ & u, v \geq 0. \end{array}$$

Dual problem:

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \begin{bmatrix} e \\ e \end{bmatrix} - \begin{bmatrix} A^T \\ -A^T \end{bmatrix} y \geq 0. \end{array}$$

Note that

$$A^T y \leq e, \; -A^T y \leq e \quad \Longleftrightarrow \quad \|A^T y\|_\infty \leq 1.$$

Primal-dual basis pursuit problems:

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & Ax = b. \end{array} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \|A^T y\|_\infty \leq 1. \end{array}$$

Recall the definition of dual norms:

$$\|x\|_{*,p} := \sup\{x^T v \mid \|v\|_p \leq 1\}.$$

Exercise: Derive the dual of the ℓ_∞-norm.
Exercise: Derive the dual of the dual basis pursuit problem.

Duality in linear optimization

Primal infeasibility certificates

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array}$$

- Theorems of strong alternatives (Farkas' lemma): either

$$Ax = b, \quad x \geq 0$$

or

$$A^T y \leq 0, \quad b^T y$$

has a solution.

- The latter is a *certificate* of primal infeasibility.
- If $A^T y < 0$, $b^T y > 0$ then y is an unbounded dual direction.

Duality in linear optimization
Primal infeasibility certificates

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array}$$

- Theorems of strong alternatives (Farkas' lemma): either

$$Ax = b, \quad x \geq 0$$

  or

$$A^T y \leq 0, \quad b^T y$$

  has a solution.
- The latter is a *certificate* of primal infeasibility.
- If $A^T y < 0$, $b^T y > 0$ then y is an unbounded dual direction.

Duality in linear optimization
Primal infeasibility certificates

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array}$$

- Theorems of strong alternatives (Farkas' lemma): either

  $$Ax = b, \quad x \geq 0$$

  or

  $$A^T y \leq 0, \quad b^T y$$

  has a solution.
- The latter is a *certificate* of primal infeasibility.
- If $A^T y < 0$, $b^T y > 0$ then y is an unbounded dual direction.

Duality in linear optimization
Example of primal infeasibility

Consider

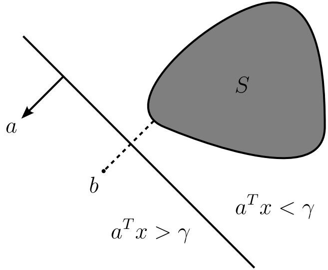
$$\begin{array}{ll} \text{minimize} & -x_1 - x_2 \\ \text{subject to} & x_1 + x_2 = -1 \\ & x_1, x_2 \geq 0 \end{array}$$

with a dual problem

$$\begin{array}{ll} \text{maximize} & -y \\ \text{subject to} & -\left[\begin{array}{c} 1 \\ 1 \end{array}\right] y \geq \left[\begin{array}{c} 1 \\ 1 \end{array}\right]. \end{array}$$

- Primal is trivially infeasible, $p^\star = \infty$.
- Any $y \leq -1$ is a certificate of primal infeasibility, as well as an unbounded dual direction, $d^\star = \infty$.

Separating hyperplane theorem



Theorem: Let S be a closed convex set, and $b \notin S$. Then there exists a separating hyperplane such that

$$a^T b > a^T x, \quad \forall x \in S.$$

Farkas' lemma

Sketch of proof

Either

$$Ax = b, \quad x \geq 0$$

or

$$A^T y \leq 0, \quad b^T y > 0$$

has a solution.

- Both cannot be true, because then $b^T y = x^T A^T y \leq 0$.
- Assume $b \notin S$ where

$$S = \{Ax \mid x \geq 0\}.$$

  Then there exists a separating hyperplane y (for b and S):

$$y^T b > y^T Ax, \quad \forall x \geq 0$$

  implying $b^T y > 0$ and $A^T y \leq 0$.

Farkas' lemma
Sketch of proof

Either

$$Ax = b, \quad x \geq 0$$

or

$$A^T y \leq 0, \quad b^T y > 0$$

has a solution.

- Both cannot be true, because then $b^T y = x^T A^T y \leq 0$.
- Assume $b \notin S$ where

$$S = \{Ax \mid x \geq 0\}.$$

  Then there exists a separating hyperplane y (for b and S):

$$y^T b > y^T Ax, \quad \forall x \geq 0$$

  implying $b^T y > 0$ and $A^T y \leq 0$.

Strong duality

Sketch of proof (using Farkas' lemma)

We assume $d^\star$ is finite. Enough to show that $p^\star \leq d^\star$.

Assume there is no $x \geq 0$ such that $Ax = b$, $c^T x \leq p^\star$, i.e.,

$$\begin{bmatrix} A & 0 \\ c^T & 1 \end{bmatrix} \begin{bmatrix} x \\ \tau \end{bmatrix} = \begin{bmatrix} b \\ d^\star \end{bmatrix}, \quad (x, \tau) \geq 0$$

has no solution. Then (from Farkas' lemma)

$$\begin{bmatrix} A^T & c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y \\ \alpha \end{bmatrix} \leq 0, \quad b^T y + \alpha d^\star > 0, \quad \underbrace{\alpha \neq 0}_{\text{why?}}$$

has a solution. Normalizing $y' := y/\alpha$ gives us

$$c - A^T y' \geq 0, \quad b^T y' > d^\star,$$

contradicting optimality of $d^\star$.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array}$$

- Theorems of strong alternatives (dual variant): either

$$c - A^T y \geq 0$$

or

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

has a solution.

- The latter is a *certificate* of dual infeasibility.

Duality in linear optimization
Dual infeasibility certificates

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \qquad \begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array}$$

- Theorems of strong alternatives (dual variant): either

$$c - A^T y \geq 0$$

or

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

has a solution.

- The latter is a *certificate* of dual infeasibility.

Duality in linear optimization

Example with both primal and dual infeasibility

Consider

$$\begin{array}{ll} \text{minimize} & -x_1 - x_2 \\ \text{subject to} & x_1 = -1 \\ & x_1, x_2 \geq 0 \end{array}$$

with a dual problem

$$\begin{array}{ll} \text{maximize} & -y \\ \text{subject to} & -\begin{bmatrix} 1 \\ 0 \end{bmatrix} y \geq \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \end{array}$$

- $y = -1$ is a certificate of primal infeasibility, $p^\star = \infty$
- $x = (0, 1)$ is a certificate of dual infeasibility, $d^\star = -\infty$.

Section 2

Conic optimization

Proper convex cones

We consider proper convex cones K in $\mathbb{R}^n$:

- Closed.
- Pointed: $K \cap (-K) = \{0\}$.
- Non-empty interior.

Dual-cone:

$$K^* = \{v \in \mathbb{R}^n \mid u^T v \geq 0,\ \forall u \in K\}.$$

If K is a proper cone, then $K^\star$ is also proper.

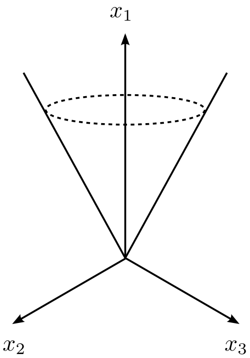
We use the notation:

$$\begin{array}{rcl} x \succeq_K y & \iff & (x - y) \in K \\ x \succ_K y & \iff & (x - y) \in \mathbf{int} K \end{array}$$

Example of cones

Quadratic cone (second-order cone, Lorenz cone)

$$\mathcal{Q}^n = \{x \in \mathbb{R}^n \mid x_1 \geq \sqrt{x_2^2 + x_3^2 + \cdots + x_n^2}\}.$$



$\mathcal{Q}^n$ is self-dual: $(\mathcal{Q}^n)^* = \mathcal{Q}^n$.

Examples of quadratic cones

- Epigraph of absolute value:

$$|x| \leq t \quad \Longleftrightarrow \quad (t, x) \in \mathcal{Q}^2.$$

- Epigraph of Euclidean norm:

$$\|x\|_2 \leq t \quad \Longleftrightarrow \quad (t, x) \in \mathcal{Q}^{n-1},$$

where $x \in \mathbb{R}^n$ and $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$.

- Second-order cone inequality:

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

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

Examples of quadratic cones

- Epigraph of absolute value:

$$|x| \le t \quad \Longleftrightarrow \quad (t, x) \in \mathcal{Q}^2.$$

- Epigraph of Euclidean norm:

$$\|x\|_2 \le t \quad \Longleftrightarrow \quad (t, x) \in \mathcal{Q}^{n-1},$$

where $x \in \mathbb{R}^n$ and $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$.

- Second-order cone inequality:

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

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

Examples of quadratic cones

- Epigraph of absolute value:

$$|x| \le t \quad \iff \quad (t, x) \in \mathcal{Q}^2.$$

- Epigraph of Euclidean norm:

$$\|x\|_2 \le t \quad \iff \quad (t, x) \in \mathcal{Q}^{n-1},$$

where $x \in \mathbb{R}^n$ and $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$.

- Second-order cone inequality:

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

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

Ellipsoidal set:

$$\begin{aligned}\mathcal{E} &= \{x \in \mathbb{R}^n \mid \|P(x-a)\|_2 \leq 1\} \\ &= \left\{x \in \mathbb{R}^n \mid x = P^{-1}y + a,\ \|y\|_2 \leq 1\right\}.\end{aligned}$$

Worst-case realization of a linear function over $\mathcal{E}$:

$$\sup_{c \in \mathcal{E}} c^T x = a^T x + \sup_{\|y\|_2 \leq 1} y^T P^{-1} x = a^T x + \|P^{-1}x\|_2.$$

Robust LP:

$$\begin{array}{ll} \text{minimize} & \sup_{c\in\mathcal{E}} c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array} \qquad \begin{array}{ll} \text{minimize} & a^T x + t \\ \text{subject to} & Ax = b \\ & (t, P^{-1}x) \in Q^{n+1} \\ & x \geq 0. \end{array}$$

Examples of quadratic cones

Robust optimization with ellipsoidal uncertainty

Ellipsoidal set:

$$\begin{aligned}\mathcal{E} &= \{x \in \mathbb{R}^n \mid \|P(x-a)\|_2 \le 1\} \\ &= \left\{x \in \mathbb{R}^n \mid x = P^{-1}y + a,\ \|y\|_2 \le 1\right\}.\end{aligned}$$

Worst-case realization of a linear function over $\mathcal{E}$:

$$\sup_{c \in \mathcal{E}} c^T x = a^T x + \sup_{\|y\|_2 \le 1} y^T P^{-1} x = a^T x + \|P^{-1}x\|_2.$$

Robust LP:

$$\begin{array}{ll} \text{minimize} & \sup_{c\in\mathcal{E}} c^T x \\ \text{subject to} & Ax = b \\ & x \ge 0, \end{array} \qquad \begin{array}{ll} \text{minimize} & a^T x + t \\ \text{subject to} & Ax = b \\ & (t, P^{-1}x) \in \mathcal{Q}^{n+1} \\ & x \ge 0. \end{array}$$

Ellipsoidal set:

$$\begin{aligned}\mathcal{E} &= \{x \in \mathbb{R}^n \mid \|P(x-a)\|_2 \leq 1\} \\ &= \left\{x \in \mathbb{R}^n \mid x = P^{-1}y + a,\ \|y\|_2 \leq 1\right\}.\end{aligned}$$

Worst-case realization of a linear function over $\mathcal{E}$:

$$\sup_{c \in \mathcal{E}} c^T x = a^T x + \sup_{\|y\|_2 \leq 1} y^T P^{-1} x = a^T x + \|P^{-1}x\|_2.$$

Robust LP:

$$\begin{array}{ll} \text{minimize} & \sup_{c \in \mathcal{E}} c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array} \qquad \begin{array}{ll} \text{minimize} & a^T x + t \\ \text{subject to} & Ax = b \\ & (t, P^{-1}x) \in Q^{n+1} \\ & x \geq 0. \end{array}$$

Rotated quadratic cone:

$$\mathcal{Q}_r^n = \{x \in \mathbb{R}^n \mid 2x_1 x_2 \geq x_3^2 + \ldots x_n^2,\, x_1, x_2 \geq 0\}.$$

Related to standard quadratic cone:

$$x \in \mathcal{Q}_r^n \quad \Longleftrightarrow \quad (T_n x) \in \mathcal{Q}^n$$

for

$$T_n := \left[\begin{array}{ccc} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & I_{n-2} \end{array}\right].$$

$\mathcal{Q}_r^n$ is self-dual: $(\mathcal{Q}_r^n)^* = \mathcal{Q}_r^n$.

Examples of rotated quadratic cones

- Epigraph of squared Euclidean norm:

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

- Convex quadratic inequality:

$$(1/2)x^T Qx \leq c^T x + d \quad \Longleftrightarrow \quad (1/2, c^T x + d, F^T x) \in \mathcal{Q}_r^{k+2}$$

with $Q = F^T F$, $F \in \mathbb{R}^{n \times k}$. So we can write QCQPs as conic problems.

Examples of rotated quadratic cones

- Epigraph of squared Euclidean norm:

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

- Convex quadratic inequality:

$$(1/2)x^T Qx \leq c^T x + d \quad \iff \quad (1/2, c^T x + d, F^T x) \in \mathcal{Q}_r^{k+2}$$

with $Q = F^T F$, $F \in \mathbb{R}^{n \times k}$. So we can write QCQPs as conic problems.

Examples of rotated quadratic cones

- Convex hyperbolic function:

$$\frac{1}{x} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (x, t, \sqrt{2}) \in \mathcal{Q}_r^3.$$

- Square roots:

$$\sqrt{x} \geq t,\ x \geq 0 \quad \Longleftrightarrow \quad (\frac{1}{2}, x, t) \in \mathcal{Q}_r^3.$$

- Convex positive rational power:

$$x^{3/2} \leq t,\ x \geq 0 \quad \Longleftrightarrow \quad (s, t, x), (x, 1/8, s) \in \mathcal{Q}_r^3.$$

- Convex negative rational power:

$$\frac{1}{x^2} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (t, \frac{1}{2}, s), (x, s, \sqrt{2}) \in \mathcal{Q}_r^3.$$

Examples of rotated quadratic cones

- Convex hyperbolic function:

$$\frac{1}{x} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (x, t, \sqrt{2}) \in \mathcal{Q}_r^3.$$

- Square roots:

$$\sqrt{x} \geq t,\ x \geq 0 \quad \Longleftrightarrow \quad (\frac{1}{2}, x, t) \in \mathcal{Q}_r^3.$$

- Convex positive rational power:

$$x^{3/2} \leq t,\ x \geq 0 \quad \Longleftrightarrow \quad (s, t, x), (x, 1/8, s) \in \mathcal{Q}_r^3.$$

- Convex negative rational power:

$$\frac{1}{x^2} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (t, \frac{1}{2}, s), (x, s, \sqrt{2}) \in \mathcal{Q}_r^3.$$

Examples of rotated quadratic cones

- Convex hyperbolic function:

$$\frac{1}{x} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (x, t, \sqrt{2}) \in \mathcal{Q}_r^3.$$

- Square roots:

$$\sqrt{x} \geq t,\ x \geq 0 \quad \Longleftrightarrow \quad (\frac{1}{2}, x, t) \in \mathcal{Q}_r^3.$$

- Convex positive rational power:

$$x^{3/2} \leq t,\ x \geq 0 \quad \Longleftrightarrow \quad (s, t, x), (x, 1/8, s) \in \mathcal{Q}_r^3.$$

- Convex negative rational power:

$$\frac{1}{x^2} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (t, \frac{1}{2}, s), (x, s, \sqrt{2}) \in \mathcal{Q}_r^3.$$

Examples of rotated quadratic cones

- Convex hyperbolic function:

$$\frac{1}{x} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (x, t, \sqrt{2}) \in \mathcal{Q}_r^3.$$

- Square roots:

$$\sqrt{x} \geq t,\ x \geq 0 \quad \Longleftrightarrow \quad (\frac{1}{2}, x, t) \in \mathcal{Q}_r^3.$$

- Convex positive rational power:

$$x^{3/2} \leq t,\ x \geq 0 \quad \Longleftrightarrow \quad (s, t, x), (x, 1/8, s) \in \mathcal{Q}_r^3.$$

- Convex negative rational power:

$$\frac{1}{x^2} \leq t,\ x > 0 \quad \Longleftrightarrow \quad (t, \frac{1}{2}, s), (x, s, \sqrt{2}) \in \mathcal{Q}_r^3.$$

Semidefinite matrices
Basic definitions

- We denote $n \times n$ symmetric matrices by $\mathcal{S}^n$.
- Standard inner product for matrices:

$$\langle V, W \rangle := \mathrm{tr}(V^T W) = \sum_{ij} V_{ij} W_{ij} = \mathbf{vec}(V)^T \mathbf{vec}(W).$$

- X is semidefinite if and only if
  1. $z^T X z \geq 0, \forall z \in \mathbb{R}^n$.
  2. All the eigenvalues of X are nonnegative.
  3. X is a *Grammian* matrix, $X = V^T V$.
- The (semi)definite matrices form a cone $(\mathcal{S}_+)$ $\mathcal{S}_{++}$.

Exercise: Show the three definitions are equivalent.

- We denote $n \times n$ symmetric matrices by $\mathcal{S}^n$.
- Standard inner product for matrices:

$$\langle V, W \rangle := \mathrm{tr}(V^T W) = \sum_{ij} V_{ij} W_{ij} = \mathbf{vec}(V)^T \mathbf{vec}(W).$$

- X is semidefinite if and only if
  1. $z^T X z \geq 0, \forall z \in \mathbb{R}^n$.
  2. All the eigenvalues of X are nonnegative.
  3. X is a *Grammian* matrix, $X = V^T V$.
- The (semi)definite matrices form a cone $(\mathcal{S}_+)$ $\mathcal{S}_{++}$.

Exercise: Show the three definitions are equivalent.

Semidefinite matrices

Basic definitions

- We denote $n \times n$ symmetric matrices by $\mathcal{S}^n$.
- Standard inner product for matrices:

$$\langle V, W \rangle := \mathrm{tr}(V^T W) = \sum_{ij} V_{ij} W_{ij} = \mathbf{vec}(V)^T \mathbf{vec}(W).$$

- X is semidefinite if and only if
  1. $z^T X z \geq 0,\ \forall z \in \mathbb{R}^n$.
  2. All the eigenvalues of X are nonnegative.
  3. X is a *Grammian* matrix, $X = V^T V$.
- The (semi)definite matrices form a cone $(\mathcal{S}_+)$ $\mathcal{S}_{++}$.

Exercise: Show the three definitions are equivalent.

Semidefinite matrices

Basic definitions

- We denote $n \times n$ symmetric matrices by $\mathcal{S}^n$.
- Standard inner product for matrices:

$$\langle V, W \rangle := \mathrm{tr}(V^T W) = \sum_{ij} V_{ij} W_{ij} = \mathbf{vec}(V)^T \mathbf{vec}(W).$$

- X is semidefinite if and only if
  1. $z^T X z \geq 0, \ \forall z \in \mathbb{R}^n$.
  2. All the eigenvalues of X are nonnegative.
  3. X is a *Grammian* matrix, $X = V^T V$.
- The (semi)definite matrices form a cone $(\mathcal{S}_+)$ $\mathcal{S}_{++}$.

Exercise: Show the three definitions are equivalent.

Dual cone:

$$(\mathcal{S}^n_+)^* = \{Z \in \mathbb{R}^{n\times n} \mid \langle X, Z\rangle \geq 0,\ \forall X \in \mathcal{S}^n_+\}.$$

The semidefinite is self-dual: $(\mathcal{S}^n_+)^* = \mathcal{S}^n_+$.

Easy to prove: Assume $Z \succeq 0$ so that $Z = U^T U$ and $X = V^T V$.

$$\langle X, Z\rangle = \langle V^T V, U^T U\rangle = \mathrm{tr}(UV^T)(UV^T)^T = \|UV^T\|_F^2 \geq 0.$$

Conversely assume $Z \not\succeq 0$. Then $\exists w \in \mathbb{R}^n$ such that

$$w^T Z w = \langle ww^T, Z\rangle = \langle X, Z\rangle < 0.$$

Schur's lemma:

$$\left(\begin{array}{cc} B & C^T \\ C & D \end{array} \right) \succ 0 \quad \Longleftrightarrow \quad B - C^T D^{-1} C \succ 0,\ C \succ 0,\ D \succ 0.$$

Example:

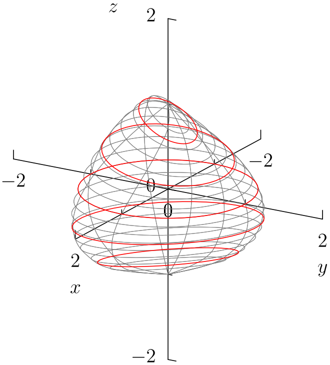
$$\left[\begin{array}{cc} t & x^T \\ x & tI \end{array} \right] \succ 0 \quad \Longleftrightarrow \quad \frac{1}{t} x^T x < t \quad \Longleftrightarrow \quad \|x\| < t,$$

i.e., quadratic cone can be embedded in a semidefinite cone.

The convex set

$$S = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succ 0 \right\},$$

is called a *pillow*.



Exercise: Characterize the restriction $S|_{z=0}$.

Eigenvalue optimization

Symmetric matrices

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathcal{S}_m.$$

- Minimize largest eigenvalue $\lambda_1(F(x))$:

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \gamma I \succeq F(x), \end{array}$$

- Maximize smallest eigenvalue $\lambda_n(F(x))$:

$$\begin{array}{ll} \text{maximize} & \gamma \\ \text{subject to} & F(x) \succeq \gamma I, \end{array}$$

- Minimize eigenvalue spread $\lambda_1(F(x)) - \lambda_n(F(x))$:

$$\begin{array}{ll} \text{minimize} & \gamma - \lambda \\ \text{subject to} & \gamma I \succeq F(x) \succeq \lambda I, \end{array}$$

Eigenvalue optimization

Symmetric matrices

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathcal{S}_m.$$

- Minimize largest eigenvalue $\lambda_1(F(x))$:

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \gamma I \succeq F(x), \end{array}$$

- Maximize smallest eigenvalue $\lambda_n(F(x))$:

$$\begin{array}{ll} \text{maximize} & \gamma \\ \text{subject to} & F(x) \succeq \gamma I, \end{array}$$

- Minimize eigenvalue spread $\lambda_1(F(x)) - \lambda_n(F(x))$:

$$\begin{array}{ll} \text{minimize} & \gamma - \lambda \\ \text{subject to} & \gamma I \succeq F(x) \succeq \lambda I, \end{array}$$

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathcal{S}_m.$$

- Minimize largest eigenvalue $\lambda_1(F(x))$:

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \gamma I \succeq F(x), \end{array}$$

- Maximize smallest eigenvalue $\lambda_n(F(x))$:

$$\begin{array}{ll} \text{maximize} & \gamma \\ \text{subject to} & F(x) \succeq \gamma I, \end{array}$$

- Minimize eigenvalue spread $\lambda_1(F(x)) - \lambda_n(F(x))$:

$$\begin{array}{ll} \text{minimize} & \gamma - \lambda \\ \text{subject to} & \gamma I \succeq F(x) \succeq \lambda I, \end{array}$$

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathbb{R}^{n \times p}.$$

- Frobenius norm: $\|F(x)\|_F := \sqrt{\langle F(x), F(x) \rangle}$,

$$\|F(x)\|_F \leq t \quad \Leftrightarrow \quad (t, \mathbf{vec}(F(x))) \in \mathcal{Q}^{np+1},$$

- Induced ℓ_2 norm: $\|F(x)\|_2 := \max_k \sigma_k(F(x))$,

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} tI & F(x)^T \\ F(x) & tI \end{bmatrix} \succeq 0, \end{array}$$

corresponds to the largest eigenvalue for $F(x) \in \mathcal{S}_+^n$.

Matrix norms

Nonsymmetric matrices

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m, \quad F_i \in \mathbb{R}^{n \times p}.$$

- Frobenius norm: $\|F(x)\|_F := \sqrt{\langle F(x), F(x) \rangle}$,

$$\|F(x)\|_F \leq t \quad \Leftrightarrow \quad (t, \mathbf{vec}(F(x))) \in \mathcal{Q}^{np+1},$$

- Induced ℓ_2 norm: $\|F(x)\|_2 := \max_k \sigma_k(F(x))$,

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} tI & F(x)^T \\ F(x) & tI \end{array}\right] \succeq 0, \end{array}$$

corresponds to the largest eigenvalue for $F(x) \in \mathcal{S}_+^n$.

Nearest correlation matrix

Consider

$$S = \{X \in \mathcal{S}_+^n \mid X_{ii} = 1,\ i = 1, \dots, n\}.$$

For a symmetric $A \in \mathbb{R}^{n \times n}$, the *nearest correlation matrix* is

$$X^\star = \arg\min_{X \in S} \|A - X\|_F,$$

which corresponds to a mixed SOCP/SDP,

$$\begin{array}{lll} \text{minimize} & t & \\ \text{subject to} & \|\mathbf{vec}(A - X)\|_2 & \leq t \\ & \mathbf{diag}(X) & = e \\ & X \succeq 0. & \end{array}$$

MOSEK is limited by the many constraints to, say $n < 200$.

Nearest correlation matrix

Consider

$$S = \{X \in \mathcal{S}_+^n \mid X_{ii} = 1,\ i = 1, \ldots, n\}.$$

For a symmetric $A \in \mathbb{R}^{n\times n}$, the *nearest correlation matrix* is

$$X^\star = \arg\min_{X\in S} \|A - X\|_F,$$

which corresponds to a mixed SOCP/SDP,

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \|\mathbf{vec}(A-X)\|_2 \leq t \\ & \mathbf{diag}(X) = e \\ & X \succeq 0. \end{array}$$

MOSEK is limited by the many constraints to, say $n < 200$.

Combinatorial relaxations

Consider a binary problem

$$\begin{array}{ll} \text{minimize} & x^T Q x + c^T x \\ \text{subject to} & x_i \in \{0,1\}, \quad i = 1,\dots,n. \end{array}$$

where $Q \in \mathcal{S}^n$ can be indefinite.

- Rewrite binary constraints $x_i \in \{0,1\}$:

$$x_i^2 = x_i \quad \Longleftrightarrow \quad X = xx^T, \quad \textbf{diag}(X) = x.$$

- Semidefinite relaxation:

$$X \succeq xx^T, \quad \textbf{diag}(X) = x.$$

Combinatorial relaxations

Consider a binary problem

$$\begin{array}{ll} \text{minimize} & x^T Q x + c^T x \\ \text{subject to} & x_i \in \{0,1\}, \quad i = 1,\ldots,n. \end{array}$$

where $Q \in \mathcal{S}^n$ can be indefinite.

- Rewrite binary constraints $x_i \in \{0,1\}$:

$$x_i^2 = x_i \quad \Longleftrightarrow \quad X = xx^T, \quad \mathbf{diag}(X) = x.$$

- Semidefinite relaxation:

$$X \succeq xx^T, \quad \mathbf{diag}(X) = x.$$

Combinatorial relaxations

Consider a binary problem

$$\begin{array}{ll} \text{minimize} & x^T Q x + c^T x \\ \text{subject to} & x_i \in \{0,1\}, \quad i = 1, \ldots, n. \end{array}$$

where $Q \in \mathcal{S}^n$ can be indefinite.

- Rewrite binary constraints $x_i \in \{0,1\}$:

$$x_i^2 = x_i \quad \Longleftrightarrow \quad X = xx^T, \quad \mathbf{diag}(X) = x.$$

- Semidefinite relaxation:

$$X \succeq xx^T, \quad \mathbf{diag}(X) = x.$$

Combinatorial relaxations

Lifted non-convex problem:

$$\begin{array}{ll} \text{minimize} & \langle Q, X\rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & X = xx^T. \end{array}$$

Semidefinite relaxation:

$$\begin{array}{ll} \text{minimize} & \langle Q, X\rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0. \end{array}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.

Combinatorial relaxations

Lifted non-convex problem:

$$\begin{array}{ll} \text{minimize} & \langle Q, X \rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & X = xx^T. \end{array}$$

Semidefinite relaxation:

$$\begin{array}{ll} \text{minimize} & \langle Q, X \rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0. \end{array}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.

Combinatorial relaxations

Lifted non-convex problem:

$$\begin{array}{ll} \text{minimize} & \langle Q, X\rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & X = xx^T. \end{array}$$

Semidefinite relaxation:

$$\begin{array}{ll} \text{minimize} & \langle Q, X\rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0. \end{array}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.

Combinatorial relaxations

Lifted non-convex problem:

$$\begin{array}{ll} \text{minimize} & \langle Q, X\rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & X = xx^T. \end{array}$$

Semidefinite relaxation:

$$\begin{array}{ll} \text{minimize} & \langle Q, X\rangle + c^T x \\ \text{subject to} & \mathbf{diag}(X) = x \\ & \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0. \end{array}$$

- Relaxation is exact if $X = xx^T$.
- Otherwise can be strengthened, e.g., by adding $X_{ij} \geq 0$.

Relaxations for boolean optimization

Same approach used for boolean constraints $x_i \in \{-1, +1\}$.

Lifting of boolean constraints

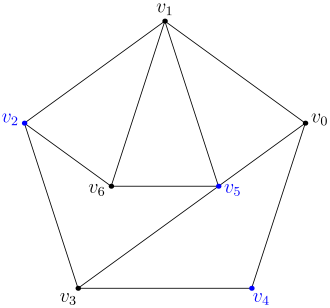
Rewrite boolean constraints $x_i \in \{-1, 1\}$:

$$x_i^2 = 1 \quad \Longleftrightarrow \quad X = xx^T, \quad \mathbf{diag}(X) = e.$$

Semidefinite relaxation of boolean constraints

$$X \succeq xx^T, \quad \mathbf{diag}(X) = e.$$

Undirected graph G with vertices V and edges E.



A cut partitions V into disjoint sets S and T with cut-set

$$I = \{(u, v) \in E \mid u \in S,\, v \in T\}.$$

The capacity of a cut is $|I|$. The cut $\{v_2, v_4, v_5\}$ has capacity 9.

Let

$$x_i = \begin{cases} +1, & v_i \in S \\ -1, & v_i \notin S \end{cases}$$

and assume $x_i \in S$. Then

$$1 - x_i x_j = \begin{cases} 2, & v_j \in S \\ 0, & v_j \notin S \end{cases}.$$

If A is the adjancency matrix for G, then the capacity is

$$\text{cap}(x) = \frac{1}{2} \sum_{(i,j) \in E} (1 - x_i x_j) = \frac{1}{4} \sum_{i,j} (1 - x_i x_j) A_{ij},$$

i.e, the MAXCUT problem is

$$\begin{array}{ll} \text{maximize} & \frac{1}{4} e^T A e - \frac{1}{4} x^T A x \\ \text{subject to} & x \in \{-1, +1\}^n. \end{array}$$

Exercise: Implement a SDP relaxation for G on the previous slide.

Sums-of-squares relaxations

- f: multivariate polynomial of degree $2d$.
- $v_d = (1, x_1, x_2, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_n^2, \ldots, x_n^d)$.
  Vector of monomials of degree d or less.

Sums-of-squares representation

f is a sums-of-squares (SOS) iff

$$f(x_1, \ldots, x_n) = v_d^T Q v_d, \quad Q \succeq 0.$$

If $Q = LL^T$ then

$$f(x_1, \ldots, x_n) = v_d^T L L^T v_d = \sum_{i=1}^{m} (l_i^T v_d)^2.$$

Sufficient condition for $f(x_1, \ldots, x_n) \geq 0$.

Sums-of-squares relaxations

- f: multivariate polynomial of degree $2d$.
- $v_d = (1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^2, \ldots, x_n^d)$.
  Vector of monomials of degree d or less.

Sums-of-squares representation

f is a sums-of-squares (SOS) iff

$$f(x_1, \ldots, x_n) = v_d^T Q v_d, \quad Q \succeq 0.$$

If $Q = LL^T$ then

$$f(x_1, \ldots, x_n) = v_d^T LL^T v_d = \sum_{i=1}^{m} (l_i^T v_d)^2.$$

Sufficient condition for $f(x_1, \ldots, x_n) \geq 0$.

A simple example

Consider

$$f(x,z) = 2x^4 + 2x^3z - x^2z^2 + 5z^4,$$

homogeneous of degree 4, so we only need

$$v = \begin{pmatrix} x^2 & xz & z^2 \end{pmatrix}.$$

Comparing cofficients of $f(x,z)$ and $v^TQv = \langle Q, vv^T \rangle$,

$$\langle Q, vv^T \rangle = \left\langle \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix}, \begin{pmatrix} x^4 & x^3z & x^2z^2 \\ x^3z & x^2z^2 & xz^3 \\ x^2z^2 & xz^3 & z^4 \end{pmatrix} \right\rangle$$

we see that $f(x,z)$ is SOS iff $Q \succeq 0$ and

$$q_{00} = 2, \quad 2q_{10} = 2, \quad 2q_{20} + q_{11} = -1, \quad 2q_{21} = 0, \quad q_{22} = 5.$$

A simple example

Consider

$$f(x,z) = 2x^4 + 2x^3 z - x^2 z^2 + 5z^4,$$

homogeneous of degree 4, so we only need

$$v = \begin{pmatrix} x^2 & xz & z^2 \end{pmatrix}.$$

Comparing cofficients of $f(x,z)$ and $v^T Q v = \langle Q, vv^T \rangle$,

$$\langle Q, vv^T \rangle = \langle \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix}, \begin{pmatrix} x^4 & x^3 z & x^2 z^2 \\ x^3 z & x^2 z^2 & xz^3 \\ x^2 z^2 & xz^3 & z^4 \end{pmatrix} \rangle$$

we see that $f(x,z)$ is SOS iff $Q \succeq 0$ and

$$q_{00} = 2, \quad 2q_{10} = 2, \quad 2q_{20} + q_{11} = -1, \quad 2q_{21} = 0, \quad q_{22} = 5.$$

Applications in polynomial optimization

$$f(x,z) = 4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xz - 4z^2 + 4z^4$$

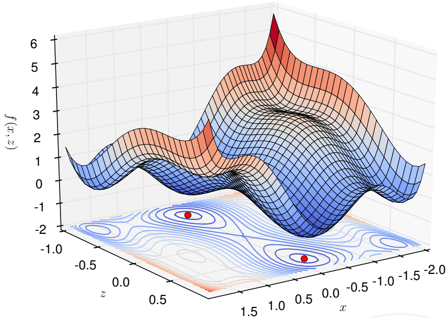
Global lower bound

Replace non-tractable problem,

$$\text{minimize} \quad f(x,z)$$

by a tractable lower bound

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & f(x,z) - t \text{ is SOS.} \end{array}$$



Relaxation finds the global optimum $t = -1.031$.

$$f(x,z) - t = 4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xz - 4z^2 + 4z^4 - t$$

$$vv^T = \begin{pmatrix} 1 & x & z & x^2 & xz & z^2 & x^3 & x^2z & xz^2 & z^3 \\ x & x^2 & xz & x^3 & x^2z & xz^2 & x^4 & x^3z & x^2z^2 & xz^3 \\ z & xz & z^2 & x^2z & xz^2 & z^3 & x^3z & x^2z^2 & xz^3 & z^4 \\ x^2 & x^3 & x^2z & x^4 & x^3z & x^2z^2 & x^5 & x^4z & x^3z^2 & x^2z^3 \\ xz & x^2z & xz^2 & x^3z & x^2z^2 & xz^3 & x^4z & x^3z^2 & x^2z^3 & xz^4 \\ z^2 & xz^2 & z^3 & x^2z^2 & xz^3 & z^4 & x^3z^2 & x^2z^3 & xz^4 & y^5 \\ x^3 & x^4 & x^3z & x^5 & x^4z & x^3z^2 & x^6 & x^5z & x^4z^2 & x^3z^3 \\ x^2z & x^3z & x^2z^2 & x^4z & x^3z^2 & x^2z^3 & x^5z & x^4z^2 & x^3z^3 & x^2z^4 \\ xz^2 & x^2z^2 & xz^3 & x^3z^2 & x^2z^3 & xz^4 & x^4z^2 & x^3z^3 & x^2z^4 & xz^5 \\ z^3 & xz^3 & z^4 & x^2z^3 & xz^4 & z^5 & x^3z^3 & x^2z^4 & xz^5 & z^6 \end{pmatrix}$$

By comparing cofficients of v^TQv and $f(x,z) - t$:

$$q_{00} = -t, \quad (2q_{30} + q_{11}) = 4, \quad (2q_{72} + q_{44}) = -\frac{21}{10}, \quad q_{77} = \frac{1}{3}$$

$$2(q_{51} + q_{32}) = 1, \quad (2q_{61} + q_{33}) = -4, \quad (2q_{10,3} + q_{66}) = 4$$

$$2q_{10} = 0, \quad 2q_{20} = 0, \quad 2(q_{71} + q_{42}) = 0, \quad \ldots$$

A standard SDP with a 10×10 variable and 28 constraints.

Nonnegative polynomials

- Univariate polynomial of degree $2n$:

$$f(x) = c_0 + c_1 x + \cdots + c_{2n} x^{2n}.$$

- Nonnegativity is **equivalent** to SOS, i.e.,

$$f(x) \geq 0 \qquad \Longleftrightarrow \qquad f(x) = v^T Q v, \quad Q \succeq 0$$

  with $v = (1, x, \ldots, x^n)$.
- Simple extensions for nonnegativity on a subinterval $I \subset \mathbb{R}$.

Nonnegative polynomials

- Univariate polynomial of degree $2n$:

$$f(x) = c_0 + c_1 x + \cdots + c_{2n} x^{2n}.$$

- Nonnegativity is **equivalent** to SOS, i.e.,

$$f(x) \geq 0 \qquad \Longleftrightarrow \qquad f(x) = v^T Q v, \quad Q \succeq 0$$

with $v = (1, x, \ldots, x^n)$.

- Simple extensions for nonnegativity on a subinterval $I \subset \mathbb{R}$.

Nonnegative polynomials

- Univariate polynomial of degree $2n$:

$$f(x) = c_0 + c_1 x + \cdots + c_{2n} x^{2n}.$$

- Nonnegativity is **equivalent** to SOS, i.e.,

$$f(x) \geq 0 \qquad \Longleftrightarrow \qquad f(x) = v^T Q v, \quad Q \succeq 0$$

with $v = (1, x, \ldots, x^n)$.

- Simple extensions for nonnegativity on a subinterval $I \subset \mathbb{R}$.

Polynomial interpolation

Fit a polynomial of degree n to a set of points (x_j, y_j),

$$f(x_j) = y_j, \quad j = 1, \ldots, m,$$

i.e., linear equality constraints in c,

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \ldots & x_m^n \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Semidefinite shape constraints:

- Nonnegativity $f(x) \geq 0$.
- Monotonicity $f'(x) \geq 0$.
- Convexity $f''(x) \geq 0$.

Polynomial interpolation

Fit a polynomial of degree n to a set of points (x_j, y_j),

$$f(x_j) = y_j, \quad j = 1, \ldots, m,$$

i.e., linear equality constraints in c,

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \ldots & x_m^n \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Semidefinite shape constraints:

- Nonnegativity $f(x) \geq 0$.
- Monotonicity $f'(x) \geq 0$.
- Convexity $f''(x) \geq 0$.

Polynomial interpolation

A specific example

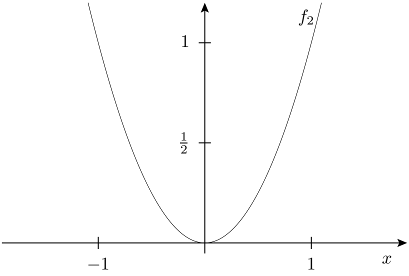
Smooth interpolation

Minimize largest derivative,

$$\begin{array}{ll} \text{minimize} & \max_{x\in[-1,1]} |f'(x)| \\ \text{subject to} & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1 \end{array}$$

or equivalently

$$\begin{array}{ll} \text{minimize} & z \\ \text{subject to} & -z \leq f'(x) \leq z \\ & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1. \end{array}$$



$f_2(x) = x^2 \qquad f_2'(1) = 2$

Polynomial interpolation

A specific example

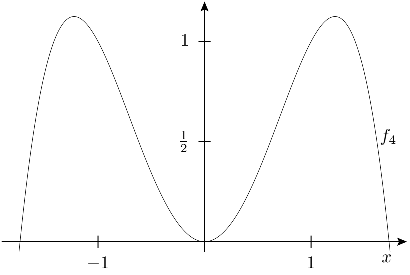
Smooth interpolation

Minimize largest derivative,

$$\begin{aligned} \text{minimize} \quad & \max_{x\in[-1,1]} |f'(x)| \\ \text{subject to} \quad & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1 \end{aligned}$$

or equivalently

$$\begin{aligned} \text{minimize} \quad & z \\ \text{subject to} \quad & -z \le f'(x) \le z \\ & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1. \end{aligned}$$



$$f_2(x) = x^2 \qquad f_2'(1) = 2$$

$$f_4(x) = \frac{3}{2}x^2 - \frac{1}{2}x^4 \qquad f_4'(\frac{1}{\sqrt{2}}) = \sqrt{2}$$

Polynomial interpolation

A specific example

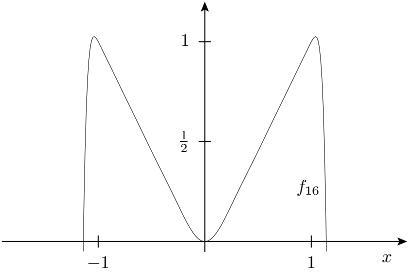
Smooth interpolation

Minimize largest derivative,

$$\begin{array}{ll} \text{minimize} & \max_{x\in[-1,1]} |f'(x)| \\ \text{subject to} & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1 \end{array}$$

or equivalently

$$\begin{array}{ll} \text{minimize} & z \\ \text{subject to} & -z \le f'(x) \le z \\ & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1. \end{array}$$



$f_2(x) = x^2 \qquad f_2'(1) = 2$

$f_4(x) = \frac{3}{2}x^2 - \frac{1}{2}x^4 \qquad f_4'(\frac{1}{\sqrt{2}}) = \sqrt{2}$

Polynomial interpolation

A specific example

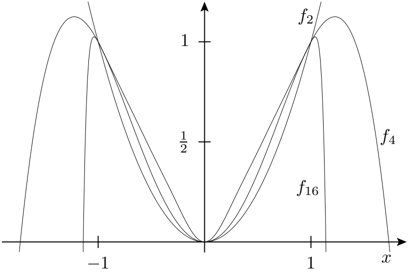
Smooth interpolation

Minimize largest derivative,

$$\begin{aligned} \text{minimize} \quad & \max_{x\in[-1,1]} |f'(x)| \\ \text{subject to} \quad & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1 \end{aligned}$$

or equivalently

$$\begin{aligned} \text{minimize} \quad & z \\ \text{subject to} \quad & -z \leq f'(x) \leq z \\ & f(-1) = 1 \\ & f(0) = 0 \\ & f(1) = 1. \end{aligned}$$



$$f_2(x) = x^2 \qquad f_2'(1) = 2$$

$$f_4(x) = \frac{3}{2}x^2 - \frac{1}{2}x^4 \qquad f_4'\left(\frac{1}{\sqrt{2}}\right) = \sqrt{2}$$

Optimizing over Hermitian semidefinite matrices

Let $X \in \mathcal{H}_+^n$ be a Hermitian semidefinite matrix of order n with inner product

$$\langle V, W \rangle := \mathrm{tr}(V^H W) = \sum_{ij} V_{ij}^* W_{ij} = \mathbf{vec}(V)^H \mathbf{vec}(W).$$

Then

$$\begin{aligned} z^H X z &= (\Re z - i \Im z)^T (\Re X + i \Im X)(\Re z + i \Im z) \\ &= \begin{bmatrix} \Re z \\ \Im z \end{bmatrix}^T \begin{bmatrix} \Re X & -\Im X \\ \Im X & \Re X \end{bmatrix} \begin{bmatrix} \Re z \\ \Im z \end{bmatrix} \geq 0, \quad \forall z \in \mathbb{C}^n. \end{aligned}$$

In other words,

$$X \in \mathcal{H}_+^n \quad \Longleftrightarrow \quad \begin{bmatrix} \Re X & -\Im X \\ \Im X & \Re X \end{bmatrix} \in \mathcal{S}_+^{2n}.$$

Note skew-symmetry $\Im X = -\Im X^T$.

Optimizing over Hermitian semidefinite matrices

Let $X \in \mathcal{H}_+^n$ be a Hermitian semidefinite matrix of order n with inner product

$$\langle V, W \rangle := \mathrm{tr}(V^H W) = \sum_{ij} V_{ij}^* W_{ij} = \mathbf{vec}(V)^H \mathbf{vec}(W).$$

Then

$$\begin{aligned} z^H X z &= (\Re z - i\Im z)^T (\Re X + i\Im X)(\Re z + i\Im z) \\ &= \begin{bmatrix} \Re z \\ \Im z \end{bmatrix}^T \begin{bmatrix} \Re X & -\Im X \\ \Im X & \Re X \end{bmatrix} \begin{bmatrix} \Re z \\ \Im z \end{bmatrix} \geq 0, \quad \forall z \in \mathbb{C}^n. \end{aligned}$$

In other words,

$$X \in \mathcal{H}_+^n \qquad \Longleftrightarrow \qquad \begin{bmatrix} \Re X & -\Im X \\ \Im X & \Re X \end{bmatrix} \in \mathcal{S}_+^{2n}.$$

Note skew-symmetry $\Im X = -\Im X^T$.

Nonnegative trigonometric polynomials

Consider a trigonometric polynomial:

$$f(z) = x_0 + 2\Re(\sum_{i=1}^{n} x_i z^{-i}), \qquad |z| = 1$$

parametrized by $x \in \mathbb{R} \times \mathbb{C}^n$. Let T_i be Toeplitz matrices with

$$[T_i]_{kl} = \begin{cases} 1, & k - l = i \\ 0, & \text{otherwise} \end{cases} \qquad i = 0, \dots, n.$$

Then $f(z) \geq 0$ on the unit-circle iff

$$X \in \mathcal{H}_+^{n+1}, \quad x_i = \langle X, T_i \rangle, \quad i = 0, \dots, n.$$

Proved by Nesterov. Simple extensions for nonnegativity on subintervals.

Consider a *transfer function*:

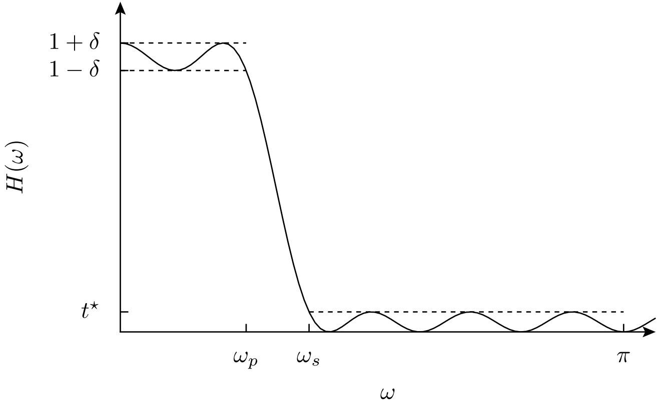
$$H(\omega) = x_0 + 2\Re(\sum_{k=1}^{n} x_k e^{-j\omega k}).$$

We can design a *lowpass filter* by solving

$$\begin{array}{lrclcl} \text{minimize} & & & t & & \\ \text{subject to} & 0 & \leq & H(\omega) & & \forall\omega \in [0,\pi] \\ & 1-\delta & \leq & H(\omega) & \leq \quad 1+\delta & \forall\omega \in [0,\omega_p] \\ & & & H(\omega) & \leq \quad t & \forall\omega \in [\omega_s,\pi], \end{array}$$

where ω_s and ω_s are design parameters.

The constraints all have simple semidefinite characterizations.



Transfer function for $n = 10, \delta = 0.05, \omega_p = \pi/4, \omega_s = \omega_p + \pi/8$.

Power cone

The $(n+1)$-dimensional power-cone is

$$K_\alpha = \left\{ x \in \mathbb{R}^{n+1} \mid x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \geq |x_{n+1}|,\ x_1, \ldots, x_n \geq 0 \right\}$$

for $\alpha > 0,\ e^T \alpha = 1$. Dual cone:

$$K_\alpha^* = \left\{ s \in \mathbb{R}^{n+1} \mid (s_1/\alpha_1)^{\alpha_1} \cdots (s_n/\alpha_n)^{\alpha_n} \geq |s_{n+1}|,\ s_1, \ldots, s_n \geq 0 \right\}$$

The power cone is self-dual:

$$T_\alpha K_\alpha^* = K_\alpha$$

where $T_\alpha := \mathbf{Diag}(\alpha_1, \ldots, \alpha_n, 1) \succ 0$.

Power cone

Simple examples

Three dimensional power cone:

$$\mathcal{Q}_\alpha = \{x \in \mathbb{R}^3 \mid x_1^\alpha x_2^{1-\alpha} \geq |x_3|,\ x_1, x_2 \geq 0\}.$$

- Epigraph of convex power $p \geq 1$:

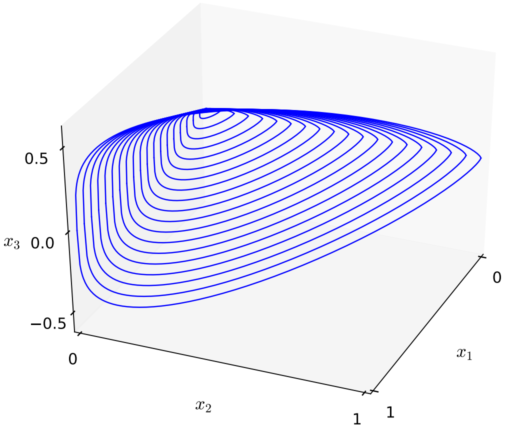
$$|x|^p \leq t \quad \Longleftrightarrow \quad (t, 1, x) \in \mathcal{Q}_{1/p}.$$

- Epigraph of p-norm:

$$\|x\|_p \leq t \quad \Longleftrightarrow \quad (z_i, t, x_i) \in \mathcal{Q}_{1/p}, \quad e^T z = t.$$

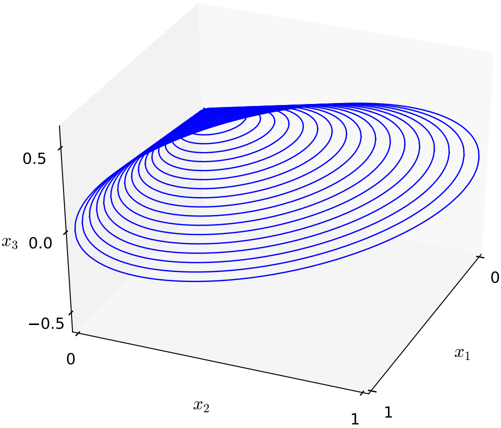
where $\|x\|_p := \left(\sum_i |x_i|^p\right)^{1/p}$.

Power cone



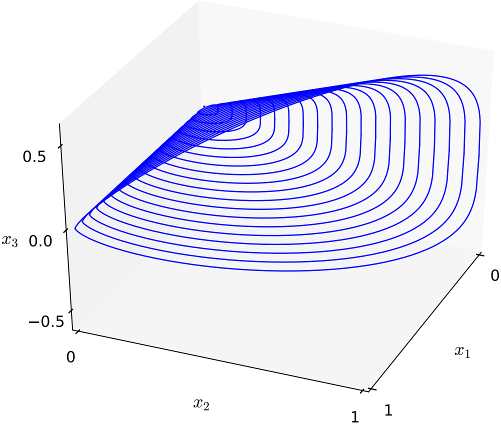
$\mathcal{Q}_{3/4}$

Power cone



$\mathcal{Q}_{1/2}$

Power cone



$\mathcal{Q}_{1/4}$

Exponential cone

Exponential cone:

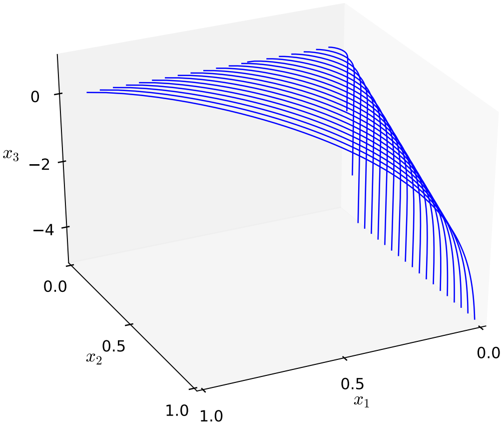
$$K_{\exp} = \mathbf{cl}\left\{x \in \mathbb{R}^3 \mid x_1 \geq x_2 e^{x_3/x_2},\ x_2 > 0\right\}$$
$$= \left\{x \in \mathbb{R}^3 \mid x_1 \geq x_2 e^{x_3/x_2},\ x_2 > 0\right\} \cup \left(\mathbb{R}_+ \times \{0\} \times \mathbb{R}_-\right)$$

Dual cone:

$$K_{\exp}^* = \mathbf{cl}\left\{s \in \mathbb{R}^3 \mid s_1 \geq (-s_3)\exp\left(\frac{s_3 - s_2}{-s_3}\right),\ s_3 < 0\right\}$$
$$= \left\{s \in \mathbb{R}^3 \mid s_1 \geq (-s_3)\exp\left(\frac{s_3 - s_2}{-s_3}\right),\ s_3 < 0\right\} \cup \left(\mathbb{R}_+^2 \times \{0\}\right).$$

Not a self-dual cone.

Exponential cone



$K_{\exp}$

Exponential cone

Simple examples

- Epigraph of negative logarithm:

$$-\log(x) \le t \quad \iff \quad (x, 1, -t) \in K_{\exp}.$$

- Epigraph of negative entropy:

$$x \log x \le t \quad \iff \quad (1, x, -t) \in K_{\exp}.$$

- Epigraph of Kullback-Leibler divergence (with variable p):

$$D(p \parallel q) = \sum_i p_i \log \frac{p_i}{q_i} \le t \quad \iff$$

$$p_i \log p_i \le p_i \log q_i, \quad \sum_i p_i \log q_i \le t$$

Exponential cone
Simple examples

- Epigraph of exponential:

$$e^x \leq t \quad \iff \quad (t, 1, x) \in K_{\exp}.$$

- Epigraph of log of sum of exponentials:

$$\log \sum_i e^{a_i^T x + b_i} \leq t \quad \iff \quad (z_i, 1, a_i^T x + b_i - t) \in K_{\exp}, \quad e^T z = 1.$$

Section 3

Primal-dual methods for conic optimization

The homogeneous model for conic problems

The homogenous model:

$$\begin{bmatrix} s \\ 0 \\ \kappa \end{bmatrix} + \begin{bmatrix} 0 & A^T & -c \\ A & 0 & -b \\ c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \tau \end{bmatrix} = 0, \quad x, s \in K, \ \tau, \kappa \geq 0.$$

Encapsulates different duality cases:

- If $\tau > 0$, $\kappa = 0$ then $\frac{1}{\tau}(x, y, s)$ is optimal,

$$Ax = b\tau, \quad c\tau - A^T y = s, \quad c^T x - b^T y = x^T s = 0.$$

- If $\tau = 0$, $\kappa > 0$ then the problem is infeasible,

$$Ax = 0, \quad -A^T y = s, \quad c^T x - b^T y < 0.$$

- If $\tau = 0$, $\kappa = 0$ then the problem is ill-posed.

The homogeneous model for conic problems

The homogenous model:

$$\begin{bmatrix} s \\ 0 \\ \kappa \end{bmatrix} + \begin{bmatrix} 0 & A^T & -c \\ A & 0 & -b \\ c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \tau \end{bmatrix} = 0, \quad x, s \in K, \ \tau, \kappa \geq 0.$$

Encapsulates different duality cases:

- If $\tau > 0$, $\kappa = 0$ then $\frac{1}{\tau}(x, y, s)$ is optimal,

$$Ax = b\tau, \quad c\tau - A^T y = s, \quad c^T x - b^T y = x^T s = 0.$$

- If $\tau = 0$, $\kappa > 0$ then the problem is infeasible,

$$Ax = 0, \quad -A^T y = s, \quad c^T x - b^T y < 0.$$

- If $\tau = 0$, $\kappa = 0$ then the problem is ill-posed.

The homogeneous model for conic problems

The homogenous model:

$$\begin{bmatrix} s \\ 0 \\ \kappa \end{bmatrix} + \begin{bmatrix} 0 & A^T & -c \\ A & 0 & -b \\ c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \tau \end{bmatrix} = 0, \quad x, s \in K, \ \tau, \kappa \geq 0.$$

Encapsulates different duality cases:

- If $\tau > 0$, $\kappa = 0$ then $\frac{1}{\tau}(x, y, s)$ is optimal,

$$Ax = b\tau, \quad c\tau - A^T y = s, \quad c^T x - b^T y = x^T s = 0.$$

- If $\tau = 0$, $\kappa > 0$ then the problem is infeasible,

$$Ax = 0, \quad -A^T y = s, \quad c^T x - b^T y < 0.$$

- If $\tau = 0$, $\kappa = 0$ then the problem is ill-posed.

The homogeneous model for conic problems

The homogenous model:

$$\begin{bmatrix} s \\ 0 \\ \kappa \end{bmatrix} + \begin{bmatrix} 0 & A^T & -c \\ A & 0 & -b \\ c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \tau \end{bmatrix} = 0, \quad x, s \in K, \; \tau, \kappa \geq 0.$$

Encapsulates different duality cases:

- If $\tau > 0$, $\kappa = 0$ then $\frac{1}{\tau}(x, y, s)$ is optimal,

$$Ax = b\tau, \quad c\tau - A^T y = s, \quad c^T x - b^T y = x^T s = 0.$$

- If $\tau = 0$, $\kappa > 0$ then the problem is infeasible,

$$Ax = 0, \quad -A^T y = s, \quad c^T x - b^T y < 0.$$

- If $\tau = 0$, $\kappa = 0$ then the problem is ill-posed.

Symmetric cones

Symmetric cones can be written as squares

$$x^2 = x \circ x$$

for appropriate product $x \circ y$.

Products for three symmetric cones:

- Nonnegative orthant: $x \circ y = \mathbf{diag}(X)y$.
- Second-order cone with $x = (x_1, x_2)$ and $y = (y_1, y_2)$:

$$x \circ y = \begin{bmatrix} x^T y \\ x_1 y_2 + y_1 x_2 \end{bmatrix}.$$

- Semidefinite cone with $X = \mathbf{mat}(x)$ and $Y = \mathbf{mat}(y)$:

$$x \circ y = (1/2)\mathbf{vec}(XY + YX).$$

Central path for homogeneous model

Given initial point $z^0 := (x^0, y^0, s^0, \tau^0, \kappa^0)$.

Central path:

$$\begin{bmatrix} s \\ 0 \\ \kappa \end{bmatrix} + \begin{bmatrix} 0 & A^T & -c \\ A & 0 & -b \\ c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \tau \end{bmatrix} = \gamma \begin{bmatrix} A^T y^0 + s^0 - c\tau^0 \\ Ax^0 - b\tau^0 \\ c^T x^0 - b^T y^0 + \kappa^0 \end{bmatrix}$$

$$x \circ s = \gamma \mu^0 e, \quad \tau^0 \kappa^0 = \gamma \mu^0$$

where e is the unit-element and $\mu^0 := \dfrac{(x^0)^T s^0 + \tau^0 \kappa^0}{n+1}$.

Continuously connects z^0 to $z^\star$ as γ goes from 1 to 0.

Nesterov-Todd scaling for symmetric cones

Properties of symmetric Nesterov-Todd scaling W:

- Maps x and s to the same *scaling point* λ.

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

- Leaves the cone invariant.

$$x, s \succeq 0 \quad \iff \quad \lambda \succeq 0$$

- Preserves the central path.

$$x \circ s = (Wx) \circ (W^{-1}s) = \lambda \circ \lambda = \lambda^2$$

Computing a search-direction

Essential part of a primal-dual method

Linearizing the scaled central path:

$$\begin{bmatrix} \Delta s \\ 0 \\ \Delta\kappa \end{bmatrix} + \begin{bmatrix} 0 & A^T & -c \\ A & 0 & -b \\ c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\tau \end{bmatrix} = \begin{bmatrix} r_x \\ r_y \\ r_\tau \end{bmatrix}$$

$$\lambda \circ (W\Delta x + W^{-1}\Delta s) = \gamma\mu e - \lambda^2, \quad \tau\Delta\kappa + \kappa\Delta\tau = \gamma\mu - \tau\kappa,$$

where r_x, r_y and r_z depend on previous iteration.

Most expensive step (after block-elimination):

$$AW^{-2}A^T\Delta y = \tilde{r}_y.$$

Solved using a Cholesky factorization.