Introduction to conic optimization

June 20th 2016 - June 22th 2016

e.d.andersen@mosek.com

www.mosek.com

Content

- Semidefinite optimization.
- Topological properties
- Duality revisted.
- Complementarity.

Section 1

Semidefinite optimization

The semidefinite cone

Definition

$$\mathcal{K}_s \quad := \quad \{X \in \mathbb{R}^{n\times n} \mid \quad X = X^T, \quad \lambda_{\min}(X) \geq 0\}$$

- The cone of symmetric positive semidefinite matrices.
- Is a cone?

Proof.

$\mathcal{K}_s$ is convex. Since, $X, Y \in \mathcal{K}_s$ implies

$$v^T(\lambda X + (1-\lambda)Y)v \geq 0$$

for all v and $0 \leq \lambda \leq 1$. Moreover, $\alpha X \in \mathcal{K}_s$ for all $\alpha \geq 0$. □

The semidefinite optimization problem

$$
\begin{array}{llll}
\text{minimize} & \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{\bar{n}} \langle \overline{C}_j, \overline{X}_j \rangle & & \\
\text{subject to} & \sum_{j=1}^{n} a_{ij} x_j + \sum_{j=1}^{\bar{n}} \langle \overline{A}_{ij}, \overline{X}_j \rangle & = \quad b_i, & i = 1, \dots, m, \\
& x \in \mathcal{K}, & & \\
& \overline{X}_j \succeq 0, & & j = 1, \dots, \bar{n}.
\end{array}
$$

Explanation:

- x_j is a scalar variable.
- $\overline{X}_j$ is a square matrix variable.

- $\mathcal{K}$ represents Cartesian product of conic quadratic constraints.
- $\overline{C}_j$ and $\overline{A}_j$ are required to be symmetric.
- Inner product between matrices

$$\begin{aligned} \langle A, B \rangle &= \mathrm{tr}(A^T B) \\ &= \sum_i \sum_j A_{ij} B_{ij}. \end{aligned}$$

- Linear constraints + a conic constraint = conic optimization problem.

The nearest correlation matrix

X is a correlation matrix if

$$X \in \mathcal{C} := \{X \in K_s \mid \mathrm{diag}(X) = e\}.$$

Links:

- https://en.wikipedia.org/wiki/Correlation_and_dependence
- https://nickhigham.wordpress.com/2013/02/13/the-nearest-correlation-matrix/

Higham:

> *A correlation matrix is a symmetric matrix with unit diagonal and nonnegative eigenvalues. In 2000 I was approached by a London fund management company who wanted to find the nearest correlation matrix (NCM) in the Frobenius norm to an almost correlation matrix: a symmetric matrix having a significant number of (small) negative eigenvalues. This problem arises when the data from which the correlations are constructed is asynchronous or incomplete, or when models are stress-tested by artificially adjusting individual correlations. Solving the NCM problem (or obtaining a true correlation matrix some other way) is important in order to avoid subsequent calculations breaking down due to negative variances or volatilities, for example.*

The Frobenius norm

$$\|A\|_F := \sqrt{\sum_i \sum_j A_{ij}^2}.$$

Given a symmetric matrix A then the problem is

$$\begin{array}{ll} \text{minimize} & \|A - X\|_F \\ \text{subject to} & X \in \mathcal{K}_s. \end{array}$$

The problem

$$\begin{array}{lccc} \text{minimize} & t & & \\ \text{subject to} & (t; \mathsf{vec}(A - X)) & \in & \mathcal{K}_q, \\ & \mathsf{diag}(X) & = & e, \\ & X & \succeq & 0, \end{array}$$

where

$$\mathsf{vec}(U) = (U_{11}; \sqrt{2}U_{21}; \ldots, \sqrt{2}U_{n1}; U_{22}; \sqrt{2}U_{32}; \ldots, \sqrt{2}U_{n2}; \ldots; U_{nn})^T.$$

Consider a binary problem (Q possible indefinite)

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

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

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

Observe

$$x^T Q x = \langle Q, X \rangle .$$

Still non-convex, since $\text{rank}(X) = 1$.

Clearly,

$$X - xx^T \succeq 0$$

is relaxation of

$$X - xx^T = 0.$$

Also we have

$$X - xx^T \succeq 0 \quad \Leftrightarrow \quad \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0.$$

Combinatorial relaxations

Lifted non-convex problem:

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

Semidefinite relaxation:

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

- Relaxation is exact if the optimal solution satisfies $X = xx^T$.
- Can be strengthened, e.g., by adding $X_{ij} \geq 0$.
- Typical relaxation for combinatorial optimization.

A recap of semidefinite optimization

- The linear and quadratic cones are special cases of the semidefinite cone.
- The semidefinite cone is a very powerful modeling construct.
- The semidefinite cone has a serious drawback.
  - Any suggestions for what that might be?

Section 2

Other properties

Topological properties

The interior of the linear cone is given by

$$\text{int}(\mathcal{K}_l) := \{x \in \mathbb{R}^1 : \; x > 0\}.$$

The interior of the quadratic cone is given by

$$\text{int}(\mathcal{K}_q) := \{x \in \mathbb{R}^n : \; x_1 > \|x_{2:n}\|\}.$$

The interior of the semidefinite cone is given by

$$\text{int}(\mathcal{K}_s) := \{X \in \mathbb{R}^{n \times n} : \; X \succ 0\}.$$

Section 3

Duality revisited

A motivating example

Consider

$$\begin{array}{ll} \text{minimize} & x^{-1} \\ \text{subject to} & x \geq 0. \end{array}$$

CQ representation

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & (x; s; \sqrt{2}) \in \mathcal{K}_r. \end{array}$$

Let

$$(x; s; \sqrt{2}) = (\frac{1}{\epsilon^k}; \epsilon^k; \sqrt{2})$$

where $0 < \epsilon < 1$ and k is a positive integer, then it defines a sequence converging towards the optimal solution for $k \to \infty$.

- The optimal solution is not finite.
- Optimal value of 0 is never attained.
    - Replace minimize by $\inf$).

Observations

- Bad things can happen in conic optimization! E.g. nonattainment, infinite values.
- But it did not happen in the (finite dimensional) linear case.
- What is going on?

Duality in conic case

The primal problem

$$\begin{array}{rclrcl} \nu_p & = & \inf & c^T x & & \\ & & \text{subject to} & Ax & = & b, \\ & & & x \in \mathcal{K}. & & \end{array} \tag{1}$$

and the dual problem

$$\begin{array}{rclrcl} \nu_d & = & \sup & b^T y & & \\ & & \text{subject to} & A^T y + s & = & c, \\ & & & s \in \mathcal{K}^*, & & \end{array} \tag{2}$$

where

$$\mathcal{K}^* := \{s : \; s^T x \geq 0, \; \forall x \in \mathcal{K}\}.$$

Dual cones

Lemma

1. *If $\mathcal{K}$ is convex and closed, then $(\mathcal{K}^*)^* = \mathcal{K}$.*
2. *$\mathcal{K}^*$ is closed and convex. (Holds even if $\mathcal{K}$ is not convex but is a cone).*
3. *$\mathcal{K}_1 \subseteq \mathcal{K}_2$ implies $\mathcal{K}_2^* \subseteq \mathcal{K}_1^*$.*

Dual cones for the 3 cones of interest

- Linear case:

$$(\mathcal{K}_l)^* = \mathcal{K}_l.$$

- Conic quadratic case:

$$(\mathcal{K}_q)^* = \mathcal{K}_q.$$

- Semidefinite case:

$$(\mathcal{K}_s)^* = \mathcal{K}_s.$$

All 3 cones are **self-dual**!

Some definitions

- The problem is *primal feasible* if a solution x exists satisfying the constraints of (1).
- The problem is *dual feasible* if a solution (y, s) exists satisfying the constraints of (2).
- If (1) is infeasible, then $\nu_p = \infty$.
- If (2) is infeasible, then $\nu_d = -\infty$.

Primal infeasibility

Lemma

(1) is infeasible if

$$\exists y : \quad b^T y > 0 \quad -A^T y \in \mathcal{K}^*. \tag{3}$$

Proof.

Assume (3) holds and x^* is a feasible solution then

$$\begin{array}{rcl} 0 & < & b^T y \\ & = & (Ax^*)^T y \\ & = & -(-A^T y)^T x^* \\ & \leq & 0 \end{array}$$

which is a contradiction. $\square$

Dual infeasibility

Lemma

(2) is infeasible if

$$\exists x : \quad c^T x < 0, \quad Ax = 0, \quad x \in \mathcal{K}. \tag{4}$$

Assume x^* is a feasible solution and x satisfies (4)

$$\begin{aligned} A(x^* + \alpha x) &= b, \\ x^* + \alpha x &\in \mathcal{K}, \forall \alpha \geq 0. \end{aligned}$$

And

$$\lim_{\alpha \to \infty} c^T (x^* + \alpha x) = -\infty.$$

A conclusion anyone?

Weak duality

Lemma

Given a primal-dual feasible solution (x, y, s) *then*

$$\begin{aligned} \textit{duality gap} \quad &:= \quad \nu_p - \nu_d \\ &\geq \quad 0. \end{aligned}$$

Follows from $x \in \mathcal{K}$ *and* $s \in \mathcal{K}^*$ *implies* $x^T s \geq 0$.

Suggestions/Comments?

Strongly feasible

A definition

(1) is said to be strongly **feasible** if there $\exists \varepsilon > 0$ such that

$$\{x \in \mathbb{R}^n : \; Ax = \hat{b}, \; x \in \mathcal{K}\} \neq \emptyset$$

for all $\hat{b}$ satisfying

$$\left\| \hat{b} - b \right\| \leq \varepsilon.$$

This is the same as saying that a small perturbation in b does NOT make the problem infeasible.

Srongly infeasible

(2) is said to be strongly **infeasible** if there $\exists \varepsilon > 0$ such that

$$\{x \in \mathbb{R}^n : \; Ax = \hat{b}, \; x \in \mathcal{K}\} = \emptyset$$

for all $\hat{b}$ satisfying

$$\left\| \hat{b} - b \right\| \leq \varepsilon.$$

This is the same as saying that a small perturbation in b does NOT make the problem feasible.

Strong infeasibility

Lemma

(1) is strongly infeasible if and only if

$$b^T y = 1,\ A^T y + s = 0,\ s \in \mathcal{K}^*$$

is strongly feasible.

Lemma

(2) is strongly infeasible if and only if

$$c^T x = -1,\ Ax = 0,\ x \in \mathcal{K}$$

is strongly feasible.

Strong duality

Theorem

(Strong duality) If either (1) or (2) is strong feasible, then $\nu_d = \nu_p$.

Comments on duality

Observe:

- For proofs see [2, p. 73] and [1].
- If A is of full row rank and

$$\text{int}(\{x \in \mathbb{R}^n : \; Ax = b, \; x \in \mathcal{K}\}) \neq \emptyset$$

  then (1) is strongly feasible.
- When does it go wrong?
  - If a small perturbation in the problem data makes the problem status flip from feasible to infeasible or from infeasible to feasible.
- Such problems must be intrinsically hard to solve.
  - Consider that computations are done in finite precision.

Nasty example 1

(1) has an optimal solution but the dual is infeasible

For instance the problem

$$\begin{array}{lll} \text{minimize} & -x_2 & \\ \text{subject to} & x_1 - x_3 & = \quad 0, \\ & \sqrt{x_2^2 + x_3^2} \leq x_1, & \end{array}$$

has the set feasible solutions:

$$\{(x_1, x_2, x_3) : \; x_1 \geq 0, \; x_2 = 0, \; x_3 \geq 0\}.$$

Hence, $x = (0, 0, 0)$ is an optimal solution.

The corresponding dual problem is

$$
\begin{array}{ll}
\text{maximize} & 0 \\
\text{subject to} & y + s_1 \quad = \quad 0, \\
& s_2 \quad = \quad -1, \\
& -y + s_3 \quad = \quad 0, \\
& \sqrt{s_2^2 + s_3^2} \leq s_3.
\end{array}
$$

Hence,

$$\sqrt{s_1^2 + 1} \leq s_1$$

which implies the dual problem is infeasible.

Nasty example 2

Given a primal-dual feasible solution there might be a duality gap

Consider

$$\begin{array}{lrcl} \min & x_2 & & \\ \text{subject to} & \sqrt{x_1^2 + (x_2 - 1)^2} & \leq & x_1, \\ & \sqrt{(-x_1 + x_2)^2} & \leq & x_1. \end{array}$$

From the first constraint it follows

$$x_2 = 1$$

Using this fact and the second constraint then

$$1 \leq 2x_1.$$

The set of primal feasible solutions is

$$\left\{(x_1, x_2) : \; x_1 \geq \frac{1}{2}, \; x_2 = 1\right\}$$

and the optimal objective value is 1.
The corresponding dual problem is

$$\begin{array}{lccc} \max & z_2 & & \\ \text{subject to} & z_1 + w_1 - z_3 + w_2 & = & 0, \\ & z_2 + z_3 & = & 1, \\ & \sqrt{z_1^2 + z_2^2} & \leq & w_1, \\ & \sqrt{z_3^2} & \leq & w_2. \end{array}$$

The two last constraints implies

$$w_1 \geq |z_1| \text{ and } w_2 \geq |z_3|$$

we have

$$w_1 + z_1 \geq 0 \text{ and } w_2 - z_3 \geq 0.$$

Using the first constraint this implies

$$w_1 = -z_1 \text{ and } w_2 = z_3.$$

Now using the second constraint we have that

$$z_2 = 1 - z_3 = 1 - w_2.$$

Therefore, the dual problem is equivalent to

$$\begin{array}{lrcl} \text{maximize} & 1 - w_2 & & \\ \text{subject to} & \sqrt{w_1^2 + (1-w_2)^2} & \leq & w_1, \\ & \sqrt{w_2^2} & \leq & w_2 \end{array}$$

which has the feasible set $\{(w_1, w_2) : \; w_1 \geq 0, \; w_2 = 1\}$ and the optimal objective value is zero. Hence,

$$\begin{aligned} \text{duality gap} &= 1 - 0 \\ &= 1. \end{aligned}$$

It can be verified that if

$$(x_2 - 1)^2$$

with

$$(x_2 - \alpha)^2$$

where $\alpha > 0$ then the dual gap will be α.

Linear versus conic duality

- Almost identical.
  - Dual problems look a like.
  - Weak duality is identical
  - Infeasibility certificates exists.
- Linear optimization:
  - No duality gap occur.
  - The optimal value is always attained.
- Conic optimization:
  - Any bad situation imaginable can occur.

In linear optimization complementarity means something like

$$x_i s_i = 0.$$

What does the complementarity conditions look like for conic quadratic optimization?
First define the arrow head matrix

$$V := \mathrm{mat}(v) = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \\ v_2 & v_1 & & \\ \vdots & & \ddots & \\ v_n & & & v_1 \end{bmatrix}.$$

Observe

$$
\begin{aligned}
\mathsf{mat}(x)s &= \begin{bmatrix} x_1 & x_{2:n} & \cdots & x_n \\ x_2 & x_1 & & \\ \vdots & & \ddots & \\ x_n & & & x_1 \end{bmatrix} s \\
&= \begin{bmatrix} x^T s \\ x_1 s_2 + s_1 x_2 \\ \vdots \\ x_1 s_n + s_1 x_n \end{bmatrix}
\end{aligned}
$$

Complementarity for the quadratic cone

Lemma

Assume $\mathcal{K} = \mathcal{K}^1 \times \cdots \times \mathcal{K}^r$ *and each* $\mathcal{K}^k$ *is a quadratic cone. If* $x, s \in \mathcal{K}$, *then* x *and* s *are complementary, i.e.* $x^T s = 0$, *if and only if*

$$X^k S^k e^k = S^k X^k e^k = 0, \quad k = 1, \ldots, r,$$

where $X^k := \mathit{mat}(x^k)$, $S^k := \mathit{mat}(s^k)$ *and* $e^k = (0, 0, \ldots, 1, \ldots, 0)^T \in \mathbb{R}^{n^k}$.

Proof:
Clearly

$$X^k S^k e^k = 0 \Rightarrow (x^k)^T s^k = 0$$

because

$$\begin{aligned} 0 &= \sum_{i=1}^{n} (e^k)^T X^k S^k e^k \\ &= \sum_{i=1}^{k} (x^k)^T s^k \\ &= x^T s. \end{aligned}$$

Next we prove if

$$(x^k)^T s^k = 0 \Rightarrow X^k S^k e^k = 0$$

This is clearly true if $x_1^k = 0$ or $s_1^k = 0$. Therefore, we can assume that $x_1^k > 0$ and $s_1^k > 0$.

Now

$$\begin{aligned} 0 &= x^T s \\ &= \sum_{k=1}^{r} (x^k)^T (s^k) \\ &= \sum_{k=1}^{r} \left(x_1^k s_1^k + (x_{2:n^k}^k)^T s_{2:n^k}^k \right) \\ &\geq \sum_{k=1}^{r} \left(x_1^k s_1^k - \left\| (x_{2:n^k}^k) \right\| \left\| s_{2:n^k}^k \right\| \right) \\ &\geq 0. \end{aligned}$$

We can conclude

$$\begin{aligned} x_1^k s_1^k &= \left\| x_{2:n^k}^k \right\| \left\| s_{2:n^k}^k \right\|, \\ x_1^k &= \left\| x_{2:n^k}^k \right\| \\ s_1^k &= \left\| s_{2:n^k}^k \right\|. \end{aligned}$$

(Why?).

Now

$$|(x^k_{2:n^k})^T s^k_{2:n^k}| = \left\| x^k_{2:n^k} \right\| \left\| s^k_{2:n^k} \right\|$$

can only be the case if

$$\exists \alpha : \ x^k_{2:n^k} = \alpha s^k_{2:n^k}.$$

Therefore,

$$\begin{aligned} 0 &= (x^k)^T s^k \\ &= x^k_1 s^k_1 + \alpha \left\| s^k_{2:n^k} \right\|^2 \\ &= x^k_1 s^k_1 + \alpha (s^k_1)^2 \end{aligned}$$

and

$$\alpha = -\frac{x^k_1}{s^k_1}$$

implying that the complementarity conditions $X^k s^k = 0$ are satisfied.

Lemma

Let $X, S \in K_s$ then they are complementarity if

$$XS = 0.$$

Proof.

See ☐

Section 4

Summary

Recap.

- Have introduced semidefinite optimization.
- Reviewed conic duality.
  - Shown that robust feasibility is important.
- Discussed complementarity conditions.

References I

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

[2] J. Renegar.
A mathematical view of interior-point methods in convex optimization.
MPS/SIAM Series on Optimization. SIAM, 2001.