Conic optimization: A refresher

Jean B. Lasserre

LAAS-CNRS and Institute of Mathematics, Toulouse, France

MASTER 2 Optimization: Paris-Orsay, 2016

Conic optimization: A refresher

Jean B. Lasserre

LAAS-CNRS and Institute of Mathematics, Toulouse, France

MASTER 2 Optimization: Paris-Orsay, 2016

- Conic optimization
- LP and Semidefinite Programming

- Conic optimization
- LP and Semidefinite Programming

- Conic optimization
- LP and Semidefinite Programming

Let (V, V^*) and (E, E^*) be two dual pairs of real vector spaces with respective duality brackets $\langle\cdot,\cdot\rangle_{V,V^*}$ and $\langle\cdot,\cdot\rangle_{E,E^*}$ (two bilinear forms):

• $V = \mathbb{R}^n = V^*$ and $\langle x, y\rangle_{V,V^*} = \sum_{i=1}^n x_i y_i$

• $V = \mathcal{S}^n = V^*$, the space of $n \times n$ real symmetric matrices, with $\langle X, Y\rangle_{V,V^*} = \mathrm{trace}(X\,Y)$.

• $V = B(\Omega)$ (the space of bounded measurable functions on a compact set $\Omega \subset \mathbb{R}^n$) and $V^* = M(\Omega)$ (the space of finite signed Borel measures on Ω), with

$$\langle f, \mu\rangle_{V,V^*} = \int_\Omega f(x)\,d\mu(x).$$

(☞ V^* is not the dual of the Banach space $(V, \|\cdot\|)$.)

Let (V, V^*) and (E, E^*) be two dual pairs of real vector spaces with respective duality brackets $\langle \cdot, \cdot \rangle_{V,V^*}$ and $\langle \cdot, \cdot \rangle_{E,E^*}$ (two bilinear forms):

• $V = \mathbb{R}^n = V^*$ and $\langle x, y \rangle_{V,V^*} = \sum_{i=1}^n x_i y_i$

• $V = \mathcal{S}^n = V^*$, the space of $n \times n$ real symmetric matrices, with $\langle X, Y \rangle_{V,V^*} = \mathrm{trace}(X\, Y)$.

• $V = B(\Omega)$ (the space of bounded measurable functions on a compact set $\Omega \subset \mathbb{R}^n$) and $V^* = M(\Omega)$ (the space of finite signed Borel measures on Ω), with

$$\langle f, \mu \rangle_{V,V^*} = \int_\Omega f(x)\, d\mu(x).$$

(☞ V^* is not the dual of the Banach space $(V, \|\cdot\|)$.)

Let (V, V^*) and (E, E^*) be two dual pairs of real vector spaces with respective duality brackets $\langle \cdot, \cdot \rangle_{V,V^*}$ and $\langle \cdot, \cdot \rangle_{E,E^*}$ (two bilinear forms):

• $V = \mathbb{R}^n = V^*$ and $\langle x, y \rangle_{V,V^*} = \sum_{i=1}^n x_i y_i$

• $V = \mathcal{S}^n = V^*$, the space of $n \times n$ real symmetric matrices, with $\langle X, Y \rangle_{V,V^*} = \mathrm{trace}(X\,Y)$.

• $V = B(\Omega)$ (the space of bounded measurable functions on a compact set $\Omega \subset \mathbb{R}^n$) and $V^* = M(\Omega)$ (the space of finite signed Borel measures on Ω), with

$$\langle f, \mu \rangle_{V,V^*} = \int_\Omega f(x)\, d\mu(x).$$

(☞ V^* is not the dual of the Banach space $(V, \|\cdot\|)$.)

Let $K \subset V$ be a closed convex cone, that is, $\lambda x + \theta y \in K$ whenever $\lambda, \theta \geq 0$ and $x, y \in K$.

For instance $\mathbf{K} = \mathbb{R}^n_+$ ($x \in \mathbf{K}$ is just $x \geq 0$), or $\mathbf{K} = \mathcal{S}^n_+$ (real $n \times n$ positive semidefinite symmetric matrices).

$\mathbf{K} = B(\Omega)_+$ (nonnegative measurable functions on Ω)

The DUAL cone

$K^* \subset V^*$ of K is also a convex cone and is defined by:

$$K^* = \{y \in V^* : \langle x, y \rangle_{V,V^*} \geq 0 \quad \forall x \in K.\}$$

For instance if $\mathbf{K} = \mathbb{R}^n_+$ or $\mathcal{S}^n_+$ then $\mathbf{K}^* = \mathbf{K}$ ($\mathbf{K}$ is self-dual).

If $\mathbf{K} = B(\Omega)_+$ then $\mathbf{K}^*$ is the cone of nonnegative Borel measures on Ω)

Let $K \subset V$ be a closed convex cone, that is, $\lambda x + \theta y \in K$ whenever $\lambda, \theta \geq 0$ and $x, y \in K$.

For instance $\mathbf{K} = \mathbb{R}^n_+$ ($x \in \mathbf{K}$ is just $x \geq 0$), or $\mathbf{K} = \mathcal{S}^n_+$ (real $n \times n$ positive semidefinite symmetric matrices).

$\mathbf{K} = B(\Omega)_+$ (nonnegative measurable functions on Ω)

The DUAL cone

$K^* \subset V^*$ of K is also a convex cone and is defined by:

$$K^* \;=\; \{y \in V^* : \langle x, y \rangle_{V,V^*} \;\geq\; 0 \quad \forall x \;\in\; K.\}$$

For instance if $\mathbf{K} = \mathbb{R}^n_+$ or $\mathcal{S}^n_+$ then $\mathbf{K}^* = \mathbf{K}$ ($\mathbf{K}$ is self-dual).

If $\mathbf{K} = B(\Omega)_+$ then $\mathbf{K}^*$ is the cone of nonnegative Borel measures on Ω)

If $A : V \to E$ is a linear map,

Its ADJOINT

$A^* : E^* \to V^*$ is defined by:

$$\langle A\,x, u\rangle_{E,E^*} = \langle x, A^*\,u\rangle_{V,V^*}, \quad \forall x, u \in V \times E^*.$$

For instance if $V = \mathbb{R}^n$ and $E = \mathbb{R}^m$ then A can be identified with a $m \times n$ real matrix, and so $A^* = A^T$ (the transpose of A).

Conic programs

A CONIC program

is a linear optimization problem **P** of the form:

$$\mathbf{P}: \quad \rho = \inf_x \{\langle x, c\rangle_{V,V^*} : A\,x = b; \quad x \in K\}$$

where $c \in V^*$, $b \in E$ and $\mathbf{K} \subset V$ is a closed convex cone.

The DUAL $\mathbf{P}^*$ of the conic program P

is also a conic program, which reads:

$$\mathbf{P}^*: \quad \rho^* = \sup_y \{\langle b, y\rangle_{E,E^*} : c - A^* y \in \mathbf{K}^*\}.$$

Conic programs

A CONIC program

is a linear optimization problem **P** of the form:

$$\mathbf{P}: \quad \rho = \inf_x \{\langle x, c\rangle_{V,V^*} : A\,x = b; \quad x \in K\}$$

where $c \in V^*$, $b \in E$ and $\mathbf{K} \subset V$ is a closed convex cone.

The DUAL $\mathbf{P}^*$ of the conic program P

is also a conic program, which reads:

$$\mathbf{P}^*: \quad \rho^* = \sup_y \{\langle b, y\rangle_{E,E^*} : c - A^*\,y \in \mathbf{K}^*\}.$$

Conic programs

A CONIC program

is a linear optimization problem **P** of the form:

$$\mathbf{P}: \quad \rho = \inf_x \{\langle x, c\rangle_{V,V^*} : A\,x \;=\; b; \quad x \in K\,\}$$

where $c \in V^*$, $b \in E$ and $\mathbf{K} \subset V$ is a closed convex cone.

The DUAL $\mathbf{P}^*$ of the conic program P

is also a conic program, which reads:

$$\mathbf{P}^*: \quad \rho^* = \sup_y \{\,\langle b, y\rangle_{E,E^*} : c - A^*\,y \in \mathbf{K}^*\,\}.$$

Example: Linear programming

- $V = V^* = \mathbb{R}^n$, $E = E^* = \mathbb{R}^m$;
- $\langle x, c\rangle_{V,V^*} = \sum_{i=1}^n c_i\, x_i$,
- $\langle b, y\rangle_{E,E^*} = \sum_{i=1}^m b_i\, y_i$,
- $\mathbf{K} = \mathbb{R}^n_+ = \mathbf{K}^*$ and $A : \mathbb{R}^n \to \mathbb{R}^m$ (represented by a $m \times n$ matrix A).
- Hence the primal **P** reads:

$$\mathbf{P}: \quad \rho = \inf_x \{x^T c : \; A\,x = b; \quad x \geq 0\},$$

- and its dual $\mathbf{P}^*$ reads:

$$\mathbf{P}^*: \quad \rho^* = \sup_y \{b^T y : \; A^T y \leq c\}$$

Example: Generalized Moment Problem

- $V = M(\Omega)$ and $V^* = B(\Omega)$ where $\Omega \subset \mathbb{R}^n$ is compact.
- $\langle \mu, f \rangle_{V,V^*} = \int_\Omega f \, d\mu$
- $E = E^* = \mathbb{R}^m$; $\langle b, y \rangle_{E,E^*} = \sum_{i=1}^m b_i \, y_i$,
- $\mathbf{K} = M(\Omega)_+$, $\mathbf{K}^* = B(\Omega)_+$
- $A : V \to E$ is given by:

$$\mu \mapsto A\mu = \left[\begin{array}{c} \int_\Omega f_1 \, d\mu \\ \cdots \\ \int_\Omega f_m \, d\mu \end{array} \right] \in \mathbb{R}^m,$$

for some functions $f_j \in B(\Omega)$, $j = 1, \ldots, m$.

Hence the primal **P** reads:

$$\mathbf{P}: \quad \rho = \inf_{\mu} \{ \int_{\Omega} f_0 \, d\mu : \int_{\Omega} f_j \, d\mu = b_j, j = 1, \ldots, m; \mu \in M(K)_+ \}$$

and its dual **P*** reads:

$$\mathbf{P}^*: \quad \rho^* = \sup_{y} \{ b^T y : f_0(x) - \sum_{j=1}^{m} y_j f_j(x) \geq 0 \; \forall x \in \Omega \}$$

- Both **P** and $\mathbf{P}^*$ are CONVEX optimization problems.
- For certain classes of finite-dimensional convex cones $\mathbf{K}$, there exist extremely efficient algorithms (based on interior point methods) which can solve **P** and $\mathbf{P}^*$ in time polynomial in the input size of the problem.

- Both **P** and **P*** are CONVEX optimization problems.
- For certain classes of finite-dimensional convex cones **K**, there exist extremely efficient algorithms (based on interior point methods) which can solve **P** and **P*** in time polynomial in the input size of the problem.

Theorem (WEAK DUALITY)

*For every feasible solution x of **P** and y of $\mathbf{P}^*$,*
$\langle x, c\rangle \geq \langle b, y\rangle$*, and so* $\rho \geq \rho^*$.

Proof.

Let x and y be feasible for **P** and $\mathbf{P}^*$ respectively. Then:

$$\begin{aligned} 0 \leq \langle \underbrace{x}_{\in \mathbf{K}}, \underbrace{c - A^* y}_{\in \mathbf{K}^*} \rangle_{V,V^*} &= \langle x, c\rangle_{V,V^*} - \langle x, A^* y\rangle_{V,V^*} \\ &= \langle x, c\rangle_{V,V^*} - \langle A x, y\rangle_{E,E^*} \\ &= \langle x, c\rangle_{V,V^*} - \langle b, y\rangle_{E,E^*} \end{aligned}$$

□

Theorem (STRONG DUALITY (in finite dimensional-case))

(a) If there exists a strict feasible solution $x \in \operatorname{int} \mathbf{K}$ of $\mathbf{P}$ then $\rho = \rho^*$ and if $\rho > -\infty$ then $\mathbf{P}^*$ has an optimal solution y^*.

(b) If there exists a strict feasible solution $c - A^* y \in \operatorname{int} \mathbf{K}^*$ of $\mathbf{P}^*$ then $\rho = \rho^*$ and if $\rho^* < \infty$ then $\mathbf{P}$ has an optimal solution x^*.

(c) If there exists a strict feasible solution $x \in \operatorname{int} \mathbf{K}$ of $\mathbf{P}$ and a strict feasible solution $c - A^* y \in \operatorname{int} \mathbf{K}^*$ of $\mathbf{P}^*$ then $\rho = \rho^*$ and both $\mathbf{P}$ and $\mathbf{P}^*$ have an optimal solution.

Theorem (STRONG DUALITY (in finite dimensional-case))

(a) If there exists a strict feasible solution $x \in \mathrm{int}\,\mathbf{K}$ *of* $\mathbf{P}$ *then* $\rho = \rho^*$ *and if* $\rho > -\infty$ *then* $\mathbf{P}^*$ *has an optimal solution* y^*.

(b) If there exists a strict feasible solution $c - A^*y \in \mathrm{int}\,\mathbf{K}^*$ *of* $\mathbf{P}^*$ *then* $\rho = \rho^*$ *and if* $\rho^* < \infty$ *then* $\mathbf{P}$ *has an optimal solution* x^*.

(c) If there exists a strict feasible solution $x \in \mathrm{int}\,\mathbf{K}$ *of* $\mathbf{P}$ *and a strict feasible solution* $c - A^*y \in \mathrm{int}\,\mathbf{K}^*$ *of* $\mathbf{P}^*$ *then* $\rho = \rho^*$ *and both* $\mathbf{P}$ *and* $\mathbf{P}^*$ *have an optimal solution.*

Theorem (STRONG DUALITY (in finite dimensional-case))

(a) If there exists a strict feasible solution $x \in \mathrm{int}\,\mathbf{K}$ of $\mathbf{P}$ then $\rho = \rho^*$ and if $\rho > -\infty$ then $\mathbf{P}^*$ has an optimal solution y^*.

(b) If there exists a strict feasible solution $c - A^* y \in \mathrm{int}\,\mathbf{K}^*$ of $\mathbf{P}^*$ then $\rho = \rho^*$ and if $\rho^* < \infty$ then $\mathbf{P}$ has an optimal solution x^*.

(c) If there exists a strict feasible solution $x \in \mathrm{int}\,\mathbf{K}$ of $\mathbf{P}$ and a strict feasible solution $c - A^* y \in \mathrm{int}\,\mathbf{K}^*$ of $\mathbf{P}^*$ then $\rho = \rho^*$ and both $\mathbf{P}$ and $\mathbf{P}^*$ have an optimal solution.

Semidefinite programming

- $V = V^* = \mathcal{S}^n$, $E = E^* = \mathbb{R}^m$;
- $\langle X, C\rangle_{V,V^*}, = \mathrm{trace}(X\,C)$
- $\langle b, y\rangle_{E,E^*} = \sum_{i=1}^m b_i\, y_i$,
- $\mathbf{K} = \mathcal{S}^n_+ = \mathbf{K}^*$ and $A : \mathcal{S}^n \to \mathbb{R}^m$
- Hence the primal **P** reads:

$$\mathbf{P}: \quad \rho = \inf_X \{\mathrm{trace}(X\,C) : \; A\,X = b; \quad X \succeq 0\}$$

- the dual $\mathbf{P}^*$ reads:

$$\mathbf{P}^*: \quad \rho^* = \sup_y \{b^T y : \; C - A^* y \succeq 0\}$$

where $X \in \mathcal{S}^n_+$ is denoted $X \succeq 0$.

Notice that $A^* y$ being a real square $n \times n$ symmetric matrix, linear in y, it can always be written

$$A^* y \;=\; \sum_{i=1}^{m} y_i \, \Theta_i,$$

for some real $n \times n$ matrices Θ_i. Hence,

$$\begin{aligned} A^* y \;=\; \sum_{i=1}^{m} y_i \, \Theta_i, \;\; &= \;\; \sum_{i=1}^{m} y_i \underbrace{\left(\Theta_i + \Theta_i^T\right)/2}_{B_i} \\ &= \;\; \sum_{i=1}^{m} y_i \, B_i \end{aligned}$$

for some real $n \times n$ symmetric matrices B_i.

Therefore:

$$\begin{aligned}
\langle A\,X, y\rangle_{E,E^*} &= \langle X, A^* y\rangle_{V,V^*} \\
&= \langle X, \sum_i B_i y_i\rangle_{V,V^*} = \sum_i y_i \langle X, B_i\rangle_{V,V^*} \\
&= \left\langle \begin{bmatrix} \langle X, B_1\rangle_{V,V^*} \\ \cdots \\ \langle X, B_m\rangle_{V,V^*} \end{bmatrix}, y \right\rangle_{E,E^*}
\end{aligned}$$

and so the linear mapping $A : \mathcal{S}^n \to \mathbb{R}^m$ can be written

$$A\,X = \begin{bmatrix} \operatorname{trace}(X\,B_1) \\ \cdots \\ \operatorname{trace}(X\,B_m) \end{bmatrix} \in \mathbb{R}^m$$

Hence the primal $\mathbf{P}$ and dual $\mathbf{P}^*$ can be written:

$$\mathbf{P}: \quad \rho = \inf_{X \succeq 0} \{\mathrm{trace}(X\,C) : \mathrm{trace}\,(X\,B_i) = b_i,\; i = 1,\ldots,m\}$$

$$\mathbf{P}^*: \quad \rho^* = \sup_{y} \{b^T y : C - \sum_{i=1}^{m} y_i\,B_i \succeq 0\}$$

which is the canonical form of a semidefinite program and its dual.

Example

Consider the semidefinite program:

$$\mathbf{P}: \quad \inf_{X \succeq 0} \left\{ \langle X, -\begin{bmatrix} 3 & -5 \\ -5 & 0 \end{bmatrix} \rangle : \langle X, \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \rangle = 1; \right.$$
$$\left. \langle X, \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \rangle = 1 \right\},$$

or, equivalently,

$$\mathbf{P}: \quad \inf_{X} \left\{ -3x_1 + 10x_2 \ : \ 2x_1 + 2x_2 + x_3 = 1; x_1 - 2x_3 = 1; \right.$$
$$\left. \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \succeq 0 \right\}$$

Its DUAL $\mathbf{P}^*$ reads:

$$\mathbf{P}^* : \sup_{y} \left\{ y_1 + y_2 : \text{s.t.} \begin{bmatrix} 3 & -5 \\ -5 & 0 \end{bmatrix} - y_1 \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} - y_2 \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \succeq 0 \right\},$$

or, equivalently

$$\mathbf{P}^* : \sup_{y} \left\{ y_1 + y_2 : \text{s.t.} \begin{bmatrix} 3 - 2y_1 - y_2 & -5 - y_1 \\ -5 - y_1 & -y_1 + 2y_2 \end{bmatrix} \succeq 0 \right\}$$

Several academic SDP software packages exist, (e.g. MATLAB "LMI toolbox", SeduMi, SDPT3, ...). However, so far, size limitation is more severe than for LP software packages.

Pioneer contributions by A. Nemirovsky, Y. Nesterov, N.Z. Shor, B.D. Yudin,...