Semidefinite and Second Order Cone Programming Seminar
Fall 2012
Lecture 3

Instructor: Farid Alizadeh
Scribe: Marta Cavaleiro

09/24/2012

1 Overview

We first study the moment cone and prove that its dual is the cone of the non-negative polynomials. Next, we introduce the notion of dual problem and prove the Weak Duality Lemma and the Strong Duality Theorem via a Generalized Farkas's Lemma.

2 The Moment Cone and its Dual

2.1 The Moment Cone

Consider $\mathcal{F}_u$ the set of distribution functions, that is functions $F : \mathbb{R} \to \mathbb{R}$ which satisfy the following properties:

1. Nonnegativity over $\mathbb{R}^n$,
2. non-decreasing,
3. right-continuity: if $F \in \mathcal{F}$, then for $x_0 \in \mathbb{R}$ if $x \downarrow x_0$, then $F(x) \downarrow F(x_0)$,
4. bounded from above by $u \geq 0$: $(F(x) \leq u,\ x \in \mathbb{R})$.

Consider the *Moment Space*:

$$M_{n+1} = \left\{ \mathbf{c} \in \mathbb{R}^{n+1} : c = \int \mathbf{u}_x dF,\ \ F \in \mathcal{F}_1 \right\},$$

where $\mathbf{u}_x = \left(1, x, x^2, \ldots, x^n\right)^\top$. In other words, for each vector $\mathbf{c} \in \mathcal{M}_{n+1}$, there is a probability measure F such that c_i is the i^{th} moment of F. Of course, since $\int dF = 1$, $c_0 = 1$ always.

The *Moment Cone* is the cone generated by $\mathcal{M}_{n+1}$:

$$\mathcal{M}_{n+1} = \left\{\alpha \mathbf{c} \in \mathbb{R}^{n+1} : \alpha \geq 0,\ \mathbf{c} \in M_{n+1}\right\}.$$

Theorem 1 *The cone $\mathcal{M}_{n+1}$ is convex, pointed, full-dimensional but not closed (see last class).*

Proof: Clearly $\mathcal{M}_{n+1}$ is a cones, since if $\mathbf{c} = \int \mathbf{u}_x dF$, then $\alpha\mathbf{c} = \int \mathbf{u}_x d(\alpha F)$.

1. **Convexity:** if $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{M}_{n+1}$ then there are functions $F \in \mathcal{F}_u$ and $G \in \mathcal{F}_v$ such that $\mathbf{c}_1 = \int \mathbf{u}_x dF$ and $\mathbf{c}_2 = \int \mathbf{u}_x dG$. And we have $\mathbf{c}_1 + \mathbf{c}_2 = \int \mathbf{u}_x d(F + G)$. And the function $F + G \in \mathcal{F}_{u+v}$.

2. **Pointed:** If $\mathbf{c} = \int \mathbf{u}_x dF$, and $-\mathbf{c} = \int \mathbf{u}_x dG$, then $0 = \int \mathbf{u}_x d(F + G)$. The measure $F + G$ is non-decreasing and nonnegative, and $\int d(F + G) = 0$. This can only happen if $F + G = 0$, and so $\mathbf{c} = 0$.

3. **Full-dimensional**. If we choose $F_x(y) = 0$ if $y < x$, and $F_x(y) = 1$ if $y \geq x$, then clearly $F_x \in \mathcal{F}_1$, and $\int \mathbf{u}_x dF_x = \mathbf{u}_x$. Thus for every $x \in \mathbb{R}$, $\mathbf{x} \in \mathcal{M}_{n+1}$. Now choose $n + 1$ distinct values $x_0 > x_2 > \cdots > x_n$, since the determinant of the Vandermond matrix

$$\text{Det} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_0^n & x_1^n & \cdots & x_n^n \end{pmatrix} = \prod_{i>j} (x_i - x_j) \neq 0$$

   It follows that the $\mathbf{u}_{x_i}, i = 0, \ldots, n$, are linearly independent.

■

But $\mathcal{M}_{n+1}$ is not closed. For instance, for $n = 3$, consider the vector $(1, N, 2N^2)^\top$. This vector is the moment of the normal distribution with mean N and variance N^2. On the other hand the set of vectors $\frac{1}{2N^2}(1, N, 2N^2)^\top$ converges to $(0, 0, 1)^\top$ as $N \to \infty$, which is not a nonnegative multiple of any moment vector. As a result, $\mathcal{M}_{n+1}$ is not closed. However the cone

$$\overline{\mathcal{M}}_{n+1} = \mathcal{M}_{n+1}^u \cup \left\{\alpha(0, \ldots, 0, 1)^\mathsf{T},\ \alpha \geq 0\right\}$$

can be proved to be closed.

Theorem 2 *The extreme rays of the Moment Cone constitute the following set:*

$$\mathsf{EXT}(\mathcal{M}_{n+1}) = \left\{(1, x, x^2, \ldots, x^n)^\mathsf{T},\ x \in \mathbb{R})\right\}.$$

Proof: We observe that functions $F_x(y)$ defined by $F_x(y) = 0$ for $y < x$ and $F_x(y) = 1$ for $y \geq x$ are extreme rays of the (infinite-dimensional) convex set $\mathcal{F}_1$. This is easily seen by noting that if $F + G = F_x$, $F(y) \leq F_x(y)$ and $G(y) < F_x(y)$ for y. In particular, for $y < x$ we must have $F(y) = G(y) = 0$. For $y \geq 0$, F is nondecreasing and $G = 1 - F$ can be nondecreasing only if $-F$ is decreasing. Thus F must be constant, and since it belongs to $\mathcal{F}_1$, it must equal to 1. By symmetry, $G(y) = 1$ for all $y \geq x$. Thus $F = G = F_x$, and F_x is an extreme point. On the other hand since $\mathbf{u}_x = \int \mathbf{u}_x dF_x$, then if $\mathbf{u}_x = \mathbf{c}_1 + \mathbf{c}_2$ we must have $\int \mathbf{u}_x dF_x = \int \mathbf{u}_x dF + \int \mathbf{u}_x dG$, which implies that $F_x = F + G$, and by the preceding argument $F = G = F_x$. ■

It follows that

$$\mathsf{EXT}(\overline{\mathcal{M}}_{n+1}) = \left\{(1, x, x^2, \ldots, x^n)^T,\ x \in \mathbb{R})\right\} \cup \left\{(0, \ldots, 0, 1)^T\right\}.$$

From now on $\mathcal{M}_{n+1}$ will represent the closure $\overline{M}_{n+1}$.

2.2 The dual of the Moment Cone: the Cone of Non-Negative Polynomials

Consider the cone of non-negative polynomials of degree n:

$$\mathcal{P}_{n+1} = \left\{(p_0, \ldots, p_n)^T : p_0 + p_1 x + \cdots + p_n x^n \geq 0,\ \forall x \in \mathbb{R}\right\}.$$

Note that if n is odd, then $p_n = 0$, that is a nonnegative polynomial must have even degree. Otherwise, as $x \to -\infty$, then $p(x) \to -\infty$ as well, and cannot be a positive polynomial.

Theorem 3 $\mathcal{M}^*_{n+1} = \mathcal{P}_{n+1}$.

Proof: Firstly we prove that $\mathcal{M}^*_{n+1} \subseteq \mathcal{P}_{n+1}$. Let $\mathbf{p} \in \mathcal{M}^*_{n+1}$ which particularly implies that

$$\langle \mathbf{p}, \mathbf{x} \rangle \geq 0,\ \forall \mathbf{x} \in \mathsf{EXT}(\mathcal{M}_{n+1}),$$

that is,

$$\forall x \in \mathbb{R},\ \langle \mathbf{p}, (1, x, x^2, \ldots, x^n)^T \rangle = p_0 + p_1 x + p_2 x^2 \cdots + p_n x^n \geq 0,$$

which means that $\mathbf{p} \in \mathcal{P}_{n+1}$.

Now we show that $\mathcal{P}_{n+1} \subseteq \mathcal{M}^*_{n+1}$. Let $p \in \mathcal{P}_{n+1}$. By Caratheodory's Theorem (see last lecture) to show that $\mathbf{p} \in \mathcal{M}^*_{n+1}$ it is enought to show that

$$\langle \mathbf{p}, \mathbf{x} \rangle \geq 0,\ \forall \mathbf{x} \in \mathsf{EXT}(\mathcal{M}_{n+1}),$$

which, as we saw previously, is equivalent to prove that

$$\forall x \in \mathbb{R},\ p_0 + p_1 x + p_2 x^2 \cdots + p_n x^n \geq 0.$$

But this is trivially true since $\mathbf{p}$ is a non-negative polynomial.

■

We are now interested in studying the extreme rays of $\mathcal{P}_{n+1}$. For that we will introduce the following Lemma:

Lemma 4 *Every non-negative polynomial* p *can be written as a sum of the squares of two polynomials, that is, there exists polynomials* p_1, p_2 *such that*

$$p(x) = p_1(x)^2 + p_2(x)^2, \ x \in \mathbb{R}.$$

Proof: Since p is a non-negative polynomial we know that all its real roots have even multiplicity so, by the Fundamental Theorem of Algebra, we can write p in the form

$$p(x) = p_n \prod_{j=1}^{m} (x - c_j)^2 \prod_{j=1}^{n-2m} (x - (a_j + ib_j))(x - (a_j - ib_j))$$

where c_j, $j = 1, \ldots, m$, are its real roots (not necessarily all different) and $a_j \pm ib_j$, $j = 1, \ldots, n - 2m$, are its complex roots.

Note that,

$$\prod_{j=1}^{n-2m} (x - (a_j + ib_j))(x - (a_j - ib_j)) = \prod_{j=1}^{n-2m} (x - a_j)^2 + b_j^2. \tag{1}$$

Now we apply the following identity successively many times to the right side of (1):

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2,$$

obtaining in the end that

$$\prod_{j=1}^{n-2m} (x - (a_j + ib_j))(x - (a_j - ib_j)) = q_1(x)^2 + q_2(x)^2,$$

for some polynomials q_1 and q_2, which allows us to conclude that

$$p(x) = p_n \prod_{j=1}^{m} (x - c_j)^2 q_1(x)^2 + p_n \prod_{j=1}^{m} (x - c_j)^2 q_2(x)^2.$$

■

Remark 5 *Note that Lemma 4 is not true in general for multivariate polynomials. In fact, there are multivariate polynomials that cannot be sums of squares of any number of polynomials. Hilbert conjectured that each nonnegative multivariate polynomial is sum of squares of* rational *functions. This is the 17th of his* 23 Problems *and was proved by E. Artin to be true.*

From Lemma 4 we can conclude that the extreme rays must be a square of some polynomial. However this is not sufficient. A counterexample is the following polynomial

$$\begin{aligned} p(x) &= ((x - (a + ib))(x - (a - ib)))^2 \\ &= ((x - a)^2 + b^2)^2 \\ &= (x - a)^4 + 2(x - a)^2 b^2 + b^4. \end{aligned}$$

We see that although p is a perfect square it can be written as a sum of non-negative polynomials. So, it could never be an extreme ray.

Theorem 6 (Extreme rays of $\mathcal{P}_{n+1}$) *The extreme rays of $\mathcal{P}_{n+1}$ are the polynomials r such that:*

1. *there exists a polynomial q such that $r(x) = (q(x))^2$, $x \in \mathbb{R}$,*
2. *q has only real roots.*

Proof: Consider r a non-negative polynomial such that $r(x) = (q(x))^2$ where q has only real roots. Suppose that there exists polynomials p_1, p_2 such that

$$r(x) = p_1^2(x) + p_2^2(x).$$

We will prove that p_1^2 is in fact a multiple of r. Observe the following:

- $\deg p_1^2 \leq \deg r$ (otherwise, for a sufficiently large x_0 we would have that $p_1^2(x) > r(x)$ for $x > x_0$);
- all roots of r are also roots of p_1^2 (consequence of r being a sum of two non-negative polynomials);
- if x_i is a common root of r, with multiplicity m_i, and of p_1^2, with mutiplicity n_i, then $n_i \geq m_i$ (note that on a sufficiently small neighborhood of x_i we have $(x - x_i)^{n_i} \leq (x - x_i)^{m_i}$ which implies that $n_i \geq m_i$). This observation implies that $\deg p_1^2 \geq \deg r$.

We conclude that r and p_1^2 have the same degree, and by the second observation they both have the same roots. So, we conclude that they are the same polynomial modulo a constant. Similarly, we conclude the same about p_2^2, proving that r cannot be written as a sum of two other non-negative polinomials (different from r).

■

2.3 Optimization over the Moment Cone - an SDP case

Let $\mathbf{c} \in \mathcal{M}_{2n+1}$. By Caratheodory's Theorem and since $\mathcal{M}_{2n+1}$ is full-dimensional we know that

$$\mathbf{c} = \sum_{i=1}^{2n+1} \alpha_i \mathbf{u}_{a_i} \tag{2}$$

for some extreme rays $\mathbf{u}_{a_i} = (1, a_i, a_i^2, \ldots, a_i^{2n})^T$, $i = 1, \ldots, 2n+1$, and $\alpha_i \geq 0$, $i = 1, \ldots, 2n+1$.

Observation 7 *Observe that:*

$$\begin{pmatrix} 1 \\ a \\ a^2 \\ \vdots \\ a^n \end{pmatrix} . \begin{pmatrix} 1 & a & a^2 & \cdots & a^n \end{pmatrix} = \begin{bmatrix} 1 & a & a^2 & \cdots & a^n \\ a & a^2 & \cdot & \iddots & a^{n+1} \\ a^2 & \cdot & \iddots & \iddots & \vdots \\ \vdots & \cdot & \iddots & \iddots & \vdots \\ \vdots & a^n & \iddots & & \vdots \\ a^n & a^{n+1} & \cdots & \cdots & a^{2n} \end{bmatrix} . \tag{3}$$

Consider now the matrix

$$H(\mathbf{c}) = \sum_{i=1}^{2n+1} \alpha_i \bar{\mathbf{u}}_{a_i} \bar{\mathbf{u}}_{a_i}^T \tag{4}$$

for $\bar{\mathbf{u}}_{a_i} = (1, a_i, a_i^2, \ldots, a_i^n)^T$, $i = 1, \ldots, 2n+1$. From the observation and considering $\mathbf{c} = (1, c_1, c_2, \ldots, c_{2n})$, we can conclude that

$$H(\mathbf{c}) = \begin{bmatrix} 1 & c & c^2 & \cdots & c^n \\ c & c^2 & \cdot & \iddots & c^{n+1} \\ c^2 & \cdot & \iddots & \iddots & \vdots \\ \vdots & \cdot & \iddots & \iddots & \vdots \\ \vdots & c^n & \iddots & & \vdots \\ c^n & c^{n+1} & \cdots & \cdots & c^{2n} \end{bmatrix} .$$

which tells us that each element of $\mathcal{M}_{n+1}$ can be represented by a positive semidefinite Hankel matrix. (Note that if $\boldsymbol{e}_n = (0, 0, \ldots, 0, 1)^\top$, then $\boldsymbol{e}_n \boldsymbol{e}_n^\top$ is also a positive semidefinite Hankel matrix.)

Conversely, suppose $H(\mathbf{c}) \succcurlyeq 0$, but $\mathbf{c} \notin \mathcal{M}_{n+1}$. Then there is a polynomial $p(x) = q^2(x)$ with only real roots such that $\mathbf{p}^\top \mathbf{c} < 0$, where $\mathbf{p}$ is the vector of coefficients of $p(x)$. But, $\mathbf{p}^\top \mathbf{c} = \mathbf{q}^\top H(\mathbf{c}) \mathbf{q}$ (check!), which means $H(\mathbf{c})$ is not positive semidefinite.

We have shown that there is a one-to-one correspondence between vectors $\mathbf{c} \in \mathcal{M}_{n+1}$ written as in (2) and the matrices $H(\mathbf{c}) \succcurlyeq 0$ in (4) (note that the first row and last column of $H(\mathbf{c})$ are constituted by the elements of $\mathbf{c}$ and, because of its structure, these are the only elements that define the matrix).

Remark 8 *By construction, the matrix in (3) has rank 1. A matrix with this structure (the elements of the opposite diagonals are all equal) is called a Hankel matrix.*

We saw that $\mathbf{c}$ is in $\mathcal{M}_{n+1}$ iff the corresponding Hankel matrix $H(\mathbf{c})$ is positive semi-definite. This is a very important observation since it allows us to conclude that if we want to optimize over the moment cone we can reformulate the problem to a SPD one where we would replace the constraint $\mathbf{x} \in \mathcal{M}_{n+1}$ by $H(\mathbf{x})$ being positive semidefinite, where $H(\mathbf{x})$ is the corresponding Hankel matrix.

3 Duality Theory in Linear Conic Programming

Consider the cone $\mathcal{K}$ and the following conic linear program:

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \succcurlyeq_{\mathcal{K}} 0. \end{array} \tag{5}$$

We call problem (5) the *primal* and define its *dual* as being the problem:

$$\begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ \text{s.t.} & A^T\mathbf{y} + \mathbf{s} = \mathbf{c} \\ & \mathbf{s} \succcurlyeq_{\mathcal{K}^*} 0, \end{array} \tag{6}$$

which is equivalent to

$$\begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ \text{s.t.} & A^T\mathbf{y} \preccurlyeq_{\mathcal{K}^*} \mathbf{c}. \end{array}$$

In practice an optimization problem may not be in the primal or dual format above. However, by adding extra variables and constraints, it is possible to transform any conic optimization problem into the above format.

3.1 Weak Duality

Lemma 9 (Weak Duality) *If $\mathbf{x}$ is feasible for (5) and $(\mathbf{y}, \mathbf{s})$ is feasible for (6) then*

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\mathbf{y}.$$

Proof: We have that

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

■

since $\mathbf{s} \in \mathcal{K}^*$. To the quantity $\mathbf{s}^T\mathbf{x}$ we call *duality gap*.

Observation 10 *From the Weak Duality Lemma we can conclude the following:*

1. *If primal is unbounded then the dual is infeasible (any feasible dual solution would constitute a lower bound for the primal);*
2. *If the dual is unbounded then the primal is infeasible (analogous reason to the one above);*

3. *If $\mathbf{x}$ is feasible for the primal and $(\mathbf{y}, \mathbf{s})$ is feasible for the dual and $\mathbf{c}^\top \mathbf{x} = \mathbf{b}^\top \mathbf{y}$, then $\mathbf{x}$ must be optimal for the primal and $(\mathbf{y}, \mathbf{s})$ must be optimal for the dual.*

3.2 The Generalized Farkas' Lemma and Strong Duality

Before presenting a generalization of the well-known Farkas' Lemma we recall the Hyperplane Separation Theorem for convex sets.

Theorem 11 (Hyperplane Separation Theorem) *Consider $C \subset \mathbb{R}^n$ a closed and convex set and $\bar{\mathbf{x}} \notin C$. Then, there exists $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $b \in \mathbb{R}$ such that*

$$\langle \mathbf{a}, \bar{\mathbf{x}} \rangle < b \quad \textit{and} \quad \langle \mathbf{a}, \mathbf{z} \rangle \geq b, \ \forall \mathbf{z} \in C.$$

That is, the hyperplane $\langle \mathbf{a}, \mathbf{x} \rangle = b$ separates C from $\bar{\mathbf{x}}$.

It is easy to see that the following is a consequence of this result:

Corollary 12 (Hyperplane Separation Theorem for Cones) *Consider $\mathcal{K} \subset \mathbb{R}^n$ a closed and convex cone and $\bar{\mathbf{x}} \notin \mathcal{K}$. Then, there exists $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that*

$$\langle \mathbf{a}, \bar{\mathbf{x}} \rangle < 0 \quad \textit{and} \quad \langle \mathbf{a}, \mathbf{z} \rangle \geq 0, \ \forall \mathbf{z} \in \mathcal{K}.$$

Now, recall Farkas' Lemma:

Theorem 13 (Farkas' Lemma) *Given $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$:*

- *Either $\exists \mathbf{x} \geq \mathbf{0}$ s.t. $A\mathbf{x} = \mathbf{b}$,*
- *Or $\exists \mathbf{y}$ s.t. $A^\top \mathbf{y} \geq \mathbf{0}$ and $\mathbf{b}^\top \mathbf{y} < 0$.*

Proof: Consider $A = [\mathbf{a}_1 \mathbf{a}_2 \cdots \mathbf{a}_n]$. If $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$ and each $x_i \geq 0$, then

$$\mathbf{b} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n,$$

that is, $\mathbf{b}$ is a non-negative combination of columns of A which means that $\mathbf{b}$ is in the (polyhedral) cone generated by the columns of A. Since polyhedral cones are always closed, So either $\mathbf{b}$ is in that cone or it is not. If it is not, by the Corollary 12, since a polyhedral cone is convex and closed:

$$\exists \mathbf{y} : \ \mathbf{a}_i^\top \mathbf{y} \geq 0, \ \forall i = 1, \ldots, n, \ \text{and} \ \mathbf{b}^\top \mathbf{y} < 0.$$

■

Remark 14 *It is essential in both the Hyperplane Separation Theorem and its conic version, that the convex set (cone) be closed. If it is not, the theorem in this format is not true, since for example $\bar{x} \notin \mathcal{K}$ but could be on its boundary. In this case the strong separation mentioned is not true. This point is critical for the following discussion of duality theory.*

We now introduce a generalized version Farkas Lemma for cones:

Theorem 15 (Generalized Farkas' Lemma) *Let $\mathcal{K}$ be a proper cone and $\mathcal{K}^*$ its dual. Consider $\mathbf{b} \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$ such that $A(\mathcal{K}) = \{A\mathbf{x} : \mathbf{x} \in \mathcal{K}\}$ is closed. Then:*

- *Either $\exists \mathbf{x} \succcurlyeq_{\mathcal{K}} \mathbf{0}$ s.t. $A\mathbf{x} = \mathbf{b}$,*
- *Or $\exists \mathbf{y}$ s.t. $A^T\mathbf{y} \succcurlyeq_{\mathcal{K}^*} \mathbf{0}$ and $\mathbf{b}^T\mathbf{y} < 0$.*

Proof: Firstly note that $A(\mathcal{K})$ is a convex cone.

Now, either $b \in A(\mathcal{K})$ (the first alternative holds) or not. If not, then by the Hyperplane Separation Theorem for Cones:

$$\exists \mathbf{y} : \ \mathbf{y}^T\mathbf{z} \geq 0, \ \forall \mathbf{z} \in A(\mathcal{K}), \text{ and } \mathbf{b}^T\mathbf{y} < 0.$$

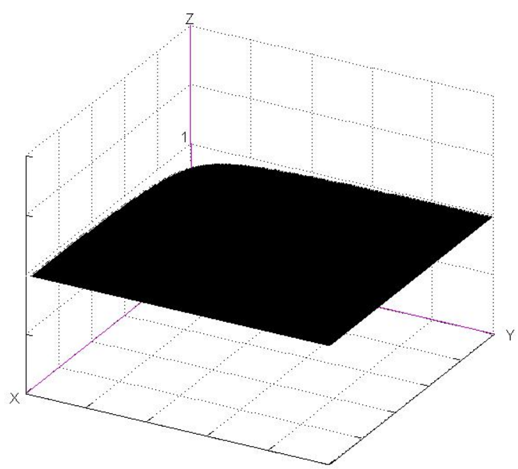
Now note that:

$$\begin{aligned} \mathbf{y}^T\mathbf{z} \geq 0, \ \forall \mathbf{z} \in A(\mathcal{K}) \quad &\Leftrightarrow \quad \mathbf{y}^T A\mathbf{x} \geq 0, \ \forall \mathbf{x} \in \mathcal{K} \\ &\Leftrightarrow \quad (A^T\mathbf{y})^T\mathbf{x} \geq 0, \ \forall \mathbf{x} \in \mathcal{K} \\ &\Leftrightarrow \quad A^T\mathbf{y} \succcurlyeq_{\mathcal{K}^*} \mathbf{0}, \end{aligned}$$

that is, the second alternative holds. ■

Observation 16 *The Generalized Farkas' Lemma for $\mathcal{K}$ equal the positive orthant of $\mathbb{R}^n$ is exactly the usual version of the Farkas' Lemma (Theorem 13).*

Observation 17 *In spite of $\mathcal{K}$ being proper, the cone $A(\mathcal{K})$, though convex, may not be pointed or closed. In fact, closedness may not be preserved by linear transformations. Such an example is the following cone:*



It is clear that when we project the cone in the figure onto the $x - z$ *plane we get* $(0, +\infty)$*, an open set.*

Clearly, we did not need this assumption in the ordinary Farkas' Lemma since the image of a polyhedral cone through a linear map is closed (a polyhedral cone does not have asymptotes).

The following example shows that the Generalized Farkas' Lemma fails when $A(\mathcal{K})$ is not closed.

Example 1 *Consider* $\mathcal{K}$ *the cone of the symmetric positive semi-definite* 2×2 *matrices:*

$$\mathcal{K} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \succcurlyeq 0 \right\},$$

and consider

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Note that $\mathcal{K} = \mathcal{K}^*$ *(see last class).*

Let us see if the Generalized Farkas' Lemma holds for this cone:

- $\exists \mathbf{x} \in \mathcal{K}$ *s.t.* $A\mathbf{x} = \mathbf{b}$*? If yes that implies that* $x_1 = 0, x_2 = 1$*. However the following matrix is not postive definite:*

$$\begin{bmatrix} 0 & 1 \\ 1 & x_3 \end{bmatrix}$$

  since it always has a negative eigenvalue. So, the first alternative is not true, so let us check the second.

- $\exists \mathbf{y}$ *s.t.* $A^T\mathbf{y} \in \mathcal{K}^*$ *and* $\mathbf{b}^T\mathbf{y} < 0$*? If it is true then*

$$b^T\mathbf{y} = y_2 < 0 \quad \text{and} \quad A^T\mathbf{y} \in \mathcal{K}^* \Leftrightarrow A^T\mathbf{y} \in \mathcal{K} \Leftrightarrow \begin{bmatrix} y_1 & y_2 \\ y_2 & 0 \end{bmatrix} \succcurlyeq 0.$$

  This is not true either because this matrix is again not positive-definite.

The reason why Generalized Farkas' Lemma is failing is because $A(\mathcal{K})$ *is not closed. In fact* $A(\mathcal{K})$ *is the following set, which clearly is not closed:*

$$A(\mathcal{K}) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 > 0, x_2 \in \mathbb{R} \right\}. \tag{7}$$

Let us prove (7)*. We have that by definition:*

$$A(\mathcal{K}) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \exists x_3 \geq 0 \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \succcurlyeq 0 \right\}. \tag{8}$$

The eigenvalues of the matrices in (8) *are the following:*

$$\lambda = \frac{(x_1 + x_3) \pm \sqrt{(x_1 - x_3)^2 + 4x_2^2}}{2}.$$

- *If* $x_1 = 0$ *then in order to have* $\lambda \geq 0$ *we need to have* $x_2 = 0$*;*
- *If* $x_1 > 0$ *then* $\lambda \geq 0$ *for any* $x_2 \in \mathbb{R}$*;*
- *If* $x_1 < 0$ *then* $\lambda < 0$ *because* $x_3 \geq 0$*.*

Now, we have the tool to prove the Strong Duality Theorem.

Theorem 18 (Strong Duality Theorem) *Let the primal problem (5) be feasible and bounded with optimal value* z_P*, and the dual problem (6) be feasible and bounded with optimal value* z_D*. If* $\overline{A}(\mathcal{K})$ *is closed, with*

$$\overline{A} = \begin{bmatrix} A \\ \mathbf{c}^T \end{bmatrix},$$

then $z_P = z_D$*.*

Proof: Firstly observe that if $\overline{A}(\mathcal{K})$ is closed then $A(\mathcal{K})$ is also closed.

Suppose, by contradiction, that $z_D < z_P$. Then by Weak Duality the following system of equations is infeasible:

$$\begin{array}{c} \mathbf{c}^T\mathbf{x} = z_D \\ A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \succcurlyeq_{\mathcal{K}} 0 \end{array} \Leftrightarrow \bar{\mathbf{b}} := \begin{pmatrix} \mathbf{b} \\ z_D \end{pmatrix} \notin \overline{A}(\mathcal{K}).$$

So by the Generalized Farkas' Lemma, since $\overline{A}(\mathcal{K})$ is closed, the following must hold:

$$\exists \bar{\mathbf{y}} : \ \overline{A}^T \bar{\mathbf{y}} \succcurlyeq_{\mathcal{K}^*} 0 \quad \text{and} \quad \bar{\mathbf{b}}^T \bar{\mathbf{y}} < 0,$$

that is,

$$\exists \bar{\mathbf{y}} := \begin{pmatrix} \mathbf{y} \\ y_0 \end{pmatrix} : \ A^T\mathbf{y} + y_0\mathbf{c} \succcurlyeq_{\mathcal{K}^*} 0 \quad \text{and} \quad \mathbf{b}^T\mathbf{y} + y_0 z_D < 0. \tag{9}$$

- If $y_0 = 0$ then $A^T\mathbf{y} \succcurlyeq_{\mathcal{K}^*} 0$ and $\mathbf{b}^T\mathbf{y} < 0$ applying again the Generalized Farkas' Lemma ($A(\mathcal{K})$ is closed) we conclude that

  $$\begin{array}{c} A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \succcurlyeq_{\mathcal{K}} 0 \end{array}$$

  is infeasible, contradicting the assumption.

- If $y_0 < 0$ then by dividing both inequalities in (9) by y_0 we get:

  $$A^T\left(\frac{\mathbf{y}}{y_0}\right) + \mathbf{c} \preccurlyeq_{\mathcal{K}^*} 0 \quad \text{and} \quad \mathbf{b}^T\left(\frac{\mathbf{y}}{y_0}\right) + z_D > 0, \tag{10}$$

  Consider now $\mathbf{y}^*$ the optimal solution of the dual, that is,

  $$A^T\mathbf{y}^* - \mathbf{c} \preccurlyeq_{\mathcal{K}^*} 0 \quad \text{and} \quad \mathbf{b}^T\mathbf{y}^* - z_D = 0. \tag{11}$$

Summing the inequalities in (10) and (11) we get:

$$A^T\left(\mathbf{y}^* + \frac{\mathbf{y}}{y_0}\right) \preccurlyeq_{\mathcal{K}^*} \mathbf{0} \quad \text{and} \quad \mathbf{b}^T\left(\mathbf{y}^* + \frac{\mathbf{y}}{y_0}\right) > 0,$$

which is equivalent to

$$A^T\left(-\mathbf{y}^* - \frac{\mathbf{y}}{y_0}\right) \succcurlyeq_{\mathcal{K}^*} \mathbf{0} \quad \text{and} \quad \mathbf{b}^T\left(-\mathbf{y}^* - \frac{\mathbf{y}}{y_0}\right) < 0.$$

Now by the Generalized Farkas' Lemma since the second alternative is true, the first one must not, so we conclude that

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\succcurlyeq_{\mathcal{K}} \mathbf{0} \end{aligned}$$

is infeasible, contradicting the assumption.

- If $y_0 > 0$ then again by dividing both inequalities in (9) by $-y_0$ we get:

$$A^T\left(-\frac{\mathbf{y}}{y_0}\right) - \mathbf{c} \preccurlyeq_{\mathcal{K}^*} \mathbf{0} \quad \text{and} \quad \mathbf{b}^T\left(-\frac{\mathbf{y}}{y_0}\right) - z_D > 0,$$

that is and $\left(-\frac{\mathbf{y}}{y_0}\right)$ is feasible for the dual and $\mathbf{b}^T\left(-\frac{\mathbf{y}}{y_0}\right) > z_D$, which contradicts the fact that z_D is the optimal value.

■

If a problem is presented in a form other than the standard form above, we can still find its dual. Here is atable (adapted from Bazaara, Sherali and Jarvis,) that finds the dual of any arbitrary conic optimization problem.

In the next table we present the correspondences between constraints of the primal and variables of the dual, and vice-versa:

min		max	
Constraints	$\succcurlyeq_{\mathcal{K}}$	$\succcurlyeq_{\mathcal{K}^*} \mathbf{0}$	Variables
	$\preccurlyeq_{\mathcal{K}}$	$\preccurlyeq_{\mathcal{K}^*} \mathbf{0}$	
	$=$	unconstrained	
Variables	$\succcurlyeq_{\mathcal{K}} \mathbf{0}$	$\preccurlyeq_{\mathcal{K}^*}$	Constraints
	$\preccurlyeq_{\mathcal{K}} \mathbf{0}$	$\succcurlyeq_{\mathcal{K}^*}$	
	unconstrained	$=$	

Table 1: Relationship between the primal and dual problems.

Example 2 *Consider the following problem:*

$$\begin{aligned} \min \quad & x \\ s.t. \quad & \begin{bmatrix} 0 & x & 0 \\ x & 0 & 0 \\ 0 & 0 & x+1 \end{bmatrix} \succcurlyeq 0 \end{aligned}$$

It is possible to prove that its dual is:

$$\begin{array}{ll} \max & -z_2 \\ s.t. & \begin{bmatrix} z_1 & \frac{1-z_2}{2} & z_3 \\ \frac{1-z_2}{2} & 0 & z_4 \\ z_3 & z_4 & z_2 \end{bmatrix} \succcurlyeq 0 \end{array}$$

We will take this up in the next lecture.