ENGG 5501: Foundations of Optimization **2016–17 First Term**

Handout 5: Elements of Conic Linear Programming

Instructor: Anthony Man–Cho So October 19, 2016

1 Introduction

As we saw in the previous handout, linear programming is a powerful tool for modelling and analyzing applications that arise in various disciplines. Despite its wide applicability, however, there are situations in reality that are inherently nonlinear and cannot be modelled as LPs. In order to handle such situations, we need to incorporate nonlinear elements in our formulation. In this lecture, we shall focus on a certain nonlinear extension of LP and demonstrate its power through various applications.

To motivate our discussion, let us first recall the standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \\ & x \geq \mathbf{0}, \end{array} \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are given. Now, observe that the inequality $x \geq \mathbf{0}$ is an inequality between two vectors. As such, it requires a definition—we say that $x \geq \mathbf{0}$ iff $x_i \geq 0$ for $i = 1, \ldots, n$. More generally, given two vectors $u, v \in \mathbb{R}^n$, we say that $u \geq v$ iff $u - v \geq \mathbf{0}$, or equivalently, $u_i \geq v_i$ for $i = 1, \ldots, n$. It is easy to verify that the relation $\geq$ defines a *partial order* on vectors in $\mathbb{R}^n$; i.e., it satisfies

(a) *(Reflexivity)* $u \geq u$ for all $u \in \mathbb{R}^n$;

(b) *(Anti–Symmetry)* $u \geq v$ and $v \geq u$ imply $u = v$ for all $u, v \in \mathbb{R}^n$;

(c) *(Transitivity)* $u \geq v$ and $v \geq w$ imply $u \geq w$ for all $u, v, w \in \mathbb{R}^n$.

What is also interesting and important in the development of LP duality theory is that the relation $\geq$ is *compatible with linear operations*; i.e., it satisfies

(d) *(Homogeneity)* for any $u, v \in \mathbb{R}^n$ and $\alpha \geq 0$, if $u \geq v$, then $\alpha u \geq \alpha v$;

(e) *(Additivity)* for any $u, v, w, z \in \mathbb{R}^n$, if $u \geq v$ and $w \geq z$, then $u + w \geq v + z$.

At this point one might ask whether $\geq$ is the only relation that satisfies (a)–(e) above. It turns out that the answer is no. Indeed, as we shall see, one can introduce nonlinearity into (1) by simply replacing the constraint $x \geq \mathbf{0}$ with the constraint $x \succeq \mathbf{0}$, where $\succeq$ is another relation that satisfies (a)–(e) above. In order to characterize such relations, let us begin with some notation. Let E be a finite–dimensional Euclidean space equipped with an inner product $\bullet$ and a relation $\succeq$. We say that the relation $\succeq$ is *good* if it satisfies (a)–(e) above. The key observation is that a good relation $\succeq$ is completely identified by the set $K = \{u \in E : u \succeq \mathbf{0}\}$. In other words, the pairs of vectors $u, v \in E$ for which $u \succeq v$ can be deduced from the set K. Indeed, suppose that $u \succeq v$. By (a), we have $-v \succeq -v$, and hence by (e) we have $u - v \succeq \mathbf{0}$. Conversely, if $u - v \succeq \mathbf{0}$, then we can add $v \succeq v$ to it and obtain $u \succeq v$.

The upshot of the above observation is that it connects the abstract relation $\succeq$ to a concrete geometric object $K \subset E$. To take advantage of this connection, let us dig deeper into the structure of the set K. We claim that K is a **pointed cone**; i.e., it has the following properties:

1. K is *non–empty* and *closed under addition*; i.e., $u + v \in K$ whenever $u, v \in K$.
2. K is a *cone*; i.e., for any $u \in K$ and $\alpha > 0$, we have $\alpha u \in K$.
3. K is *pointed*; i.e., if $u \in K$ and $-u \in K$, then $u = \mathbf{0}$.

Indeed, the first property follows from (a) (which implies that $\mathbf{0} \in K$) and (e); the second follows from (d). To derive the third property, observe that $u \succeq u$ by (a), which together with $-u \succeq \mathbf{0}$ and (e) implies that $\mathbf{0} \succeq u$. Since $u \succeq \mathbf{0}$, it follows from (b) that $u = \mathbf{0}$.

Note that a pointed cone K is automatically convex. To prove this, let $u, v \in K$ and $\alpha \in (0, 1)$. Then, since K is a cone, we have $\alpha u, (1 - \alpha)v \in K$. Since K is closed under addition, we conclude that $\alpha u + (1 - \alpha)v \in K$ as desired.

The above discussion shows that every good relation $\succeq$ on E induces a pointed cone $K = \{u \in E : u \succeq \mathbf{0}\}$ with $\mathbf{0} \in K$. It turns out that the converse is also true; i.e., given an arbitrary pointed cone $K \subset E$ with $\mathbf{0} \in K$, we can define a good relation on E. To see this, consider the relation $\succeq_K$ defined by

$$u \succeq_K v \iff u - v \in K. \tag{2}$$

Note that by definition, we have $u \succeq_K v$ iff $u - v \succeq_K \mathbf{0}$. Now, we claim that $\succeq_K$ is good:

(a) (Reflexivity) Since $\mathbf{0} \in K$, we see that for any $u \in E$, we have $u - u \in K$; i.e., $u \succeq_K u$.

(b) (Anti–Symmetry) If $u - v \in K$ and $v - u \in K$, then by the pointedness of K, we have $u - v = \mathbf{0}$; i.e., $u = v$.

(c) (Transitivity) If $u - v \in K$ and $v - w \in K$, then by the addition property, we have $u - w \in K$; i.e., $u \succeq_K w$.

(d) (Homogeneity) Suppose that $u - v \in K$ and $\alpha > 0$. By the conic property, we have $\alpha(u - v) \in K$, which implies that $\alpha u \succeq_K \alpha v$. The case where $\alpha = 0$ trivially follows from reflexivity.

(e) (Additivity) Suppose that $u - v \in K$ and $w - z \in K$. By the addition property, we have $u + w - (v + z) \in K$; i.e., $u + w \succeq_K v + z$.

Thus, the good relations on E are completely characterized by the pointed cones in E that contain the origin. Let us now consider some particularly interesting examples (in the context of optimization) of pointed cones and the good relations they induce.

Example 1 (Representative Pointed Cones)

1. **Non–Negative Orthant.** $\mathbb{R}^n_+ = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0$ *for* $i = 1, \ldots, n\}$. *We consider* $\mathbb{R}^n_+$ *as a pointed cone in* $\mathbb{R}^n$ *equipped with the usual inner product; i.e.,* $u^T v = \sum_{i=1}^n u_i v_i$. *The good relation it induces is the usual coordinate–wise ordering: For* $u, v \in \mathbb{R}^n$*, we have* $u \geq v$ *iff* $u_i \geq v_i$ *for* $i = 1, \ldots, n$.

2. **Lorentz Cone (also known as the Second–Order Cone or the Ice Cream Cone).** $\mathcal{Q}^{n+1} = \{(t,x) \in \mathbb{R} \times \mathbb{R}^n : t \ge \|x\|_2\}$. *We consider* $\mathcal{Q}^{n+1}$ *as a pointed cone in* $\mathbb{R}^{n+1}$ *equipped with the usual inner product. The good relation it induces is the following: For* $(s,u),(t,v) \in \mathbb{R} \times \mathbb{R}^n$*, we have* $(s,u) \succeq_{\mathcal{Q}^{n+1}} (t,v)$ *iff* $s - t \ge \|u - v\|_2$.

3. **Positive Semidefinite Cone.** $\mathcal{S}_+^n = \left\{ X \in \mathcal{S}^n : u^T X u \ge 0 \text{ for all } u \in \mathbb{R}^n \right\}$. *We consider* $\mathcal{S}_+^n$ *as a pointed cone in the space* $\mathcal{S}^n$ *of* $n \times n$ *symmetric matrices equipped with the* Frobenius inner product*; i.e.,*

$$X \bullet Y = \mathrm{tr}(X^T Y) = \mathrm{tr}(XY) = \sum_{i=1}^n \sum_{j=1}^n X_{ij} Y_{ij}$$

(note that $\mathcal{S}^n$ *can be identified with* $\mathbb{R}^{n(n+1)/2}$*). The good relation it induces is the so–called* positive semidefinite ordering*: For* $X, Y \in \mathcal{S}^n$*, we have* $X \succeq Y$ *iff* $X - Y$ *is positive semidefinite (denoted by* $X - Y \succeq \mathbf{0}$*).*

Note that all the cones in Example 1 are closed and have non–empty interiors. These properties have important consequences. First, the closedness of K implies that we can take limits in the relation $\succeq_K$. Specifically, if $\{u^i\}, \{v^i\}$ are sequences in E such that

$$u^i \succeq_K v^i \text{ for } i = 1, 2, \ldots; \quad u^i \to u \in E; \quad v^i \to v \in E,$$

then we may conclude that $u \succeq_K v$. Second, if the pointed cone K has a non–empty interior, then in addition to the *non–strict* relation $\succeq_K$, we may define a *strict* relation $\succ_K$ via $u \succ_K v \iff u - v \in \mathrm{int}(K)$. As we shall see, such a relation is useful in the development of the duality theory for conic optimization problems.

Before we leave this section, let us introduce a useful operation, which allows us to create new pointed cones from old ones.

Proposition 1 *Let* $E_1, \ldots, E_n$ *be finite–dimensional Euclidean spaces and* $K_i \subset E_i$ *be closed pointed cones with non–empty interiors, where* $i = 1, \ldots, n$*. Then, the set*

$$K \equiv K_1 \times \cdots \times K_n = \{(x_1, \ldots, x_n) \in E_1 \times \cdots \times E_n : x_i \in K_i \text{ for } i = 1, \ldots, n\}$$

is a closed pointed cone with non–empty interior.

We leave the proof as a simple exercise to the reader.

2 Conic Linear Programming

As before, let E be a finite–dimensional Euclidean space equipped with an inner product $\bullet$ and a good relation $\succeq$. Motivated by the development in the previous section, we can define an analog of the LP (1) by defining linear functions using $\bullet$ and replacing the constraint $x \ge \mathbf{0}$ with $x \succeq \mathbf{0}$. More precisely, let $K \subset E$ be the pointed cone induced by $\succeq$ and suppose that K is closed with non–empty interior. Given $c, a_1, \ldots, a_m \in E$ and $b_1, \ldots, b_m \in \mathbb{R}$, we define the **standard form Conic Linear Programming (CLP) problem** as follows:

$$\begin{array}{rccl} v_p^* & = & \inf & c \bullet x \\ & & \text{subject to} & a_i \bullet x = b_i \quad \text{for } i = 1, \ldots, m, \\ & & & x \succeq_K \mathbf{0}. \end{array} \tag{P}$$

The term "conic linear programming" comes from the fact that Problem (P) involves minimizing a linear objective function subject to linear equality constraints and a conic constraint. Note that nonlinearity can be incorporated into (P) through the pointed cone K. As we shall see, this gives Problem (P) tremendous modeling power.

To develop the dual of Problem (P), it is natural to follow the strategy used in the LP case. In the current setting, our goal is to find a vector $y \in \mathbb{R}^m$ such that

$$b^T y = \sum_{i=1}^m (a_i \bullet x) y_i = \left(\sum_{i=1}^m y_i a_i \right) \bullet x \le c \bullet x. \tag{3}$$

For the case of LP, we have $x \ge \mathbf{0}$. Thus, we can ensure that the above inequality holds by requiring $c - \sum_{i=1}^m y_i a_i \ge \mathbf{0}$. For the general case of CLP, however, we only have $x \succeq_K \mathbf{0}$. What then is the condition that can ensure the validity of inequality (3)? One obvious answer is that the vector $c - \sum_{i=1}^m y_i a_i$ should belong to the set

$$K^* = \{ w \in E : x \bullet w \ge 0 \text{ for all } x \in K \}.$$

The set K^* is called the *dual cone* of the cone K and turns out to have a number of important properties. We summarize these properties in the following proposition:

Proposition 2 *Let $K \subset E$ be a non–empty set. Then, the following hold:*

(a) The set K^* is a closed convex cone.

(b) If K is a closed convex cone, then so is K^*. Moreover, we have $(K^*)^* = K$.

(c) If K has a non–empty interior, then K^* is pointed.

(d) If K is a closed pointed cone, then K^* has a non–empty interior.

Proof The proof of (a) is straightforward and is left as an exercise to the reader.

Next, we prove (b). It is clear from the definition that $K \subset (K^*)^*$. To establish the converse, let $v \in (K^*)^*$ be arbitrary. If $v \notin K$, then by the separation theorem (Theorem 7 of Handout 2), there exists a $y \in \mathbb{R}^n$ such that $\inf_{x \in K} y^T x > y^T v$. We claim that $\theta^* \equiv \inf_{x \in K} y^T x = 0$. Clearly, we have $\theta^* \le 0$, since $\mathbf{0} \in K$. Now, if $\theta^* < 0$, then there exists an $x' \in K$ such that $0 > y^T x' > y^T v$. However, since $\alpha x' \in K$ for all $\alpha > 0$, we see that $\alpha y^T x' > y^T v$ for all $\alpha \ge 1$, which is impossible. Thus, the claim is established. In particular, this shows that $y \in K^*$. However, we then have the inequality $0 > y^T v$, which contradicts the fact that $v \in (K^*)^*$. Hence, we conclude that $v \in K$.

To prove (c), suppose that K^* is not pointed. Then, there exists a $w \in K^*$ such that $w \ne \mathbf{0}$ and $x \bullet w = 0$ for all $x \in K$. This implies that K is a subset of the hyperplane $H(w, \mathbf{0}) = \{x \in E : w \bullet x = 0\}$, which shows that $\text{int}(K) = \emptyset$.

Lastly, let us prove (d). Suppose that $\text{int}(K^*) = \emptyset$. Then, there exists a hyperplane $H(s, \mathbf{0}) = \{w \in E : s \bullet w = 0\}$ with $s \ne \mathbf{0}$ such that $K^* \subset H(s, \mathbf{0})$. Since K is a closed convex cone by assumption, using the result in (b), we compute

$$\begin{aligned} K = (K^*)^* &= \{x \in E : x \bullet w \ge 0 \text{ for all } w \in K^*\} \\ &\supset \{x \in E : x \bullet w \ge 0 \text{ for all } w \in H(s, \mathbf{0})\} \\ &= \{\lambda s : \lambda \in \mathbb{R}\}. \end{aligned}$$

This shows that K is not pointed. □

In view of Proposition 2, the following is immediate:

Corollary 1 *Let $K \subset E$ be a closed pointed cone with non–empty interior. Then, the dual cone $K^* \subset E$ is also a closed pointed cone with non–empty interior.*

We also have the following proposition, which shows that the dual of a Cartesian product of cones is simply the Cartesian product of the corresponding dual cones:

Proposition 3 *Let $E_1, \dots, E_n$ be finite–dimensional Euclidean spaces equipped with the inner products $\bullet_1, \dots, \bullet_n$, respectively. Let $E = E_1 \times \cdots \times E_n$ and define the inner product $\bullet$ on E by*

$$u \bullet v = \sum_{i=1}^{n} u_i \bullet_i v_i \quad \text{where } u_i, v_i \in E_i;\ i = 1, \dots, n.$$

Suppose that $K_i \subset E_i$ (where $i = 1, \dots, n$) are closed pointed cones with non–empty interiors and let $K = K_1 \times \cdots \times K_n$. Then, the dual cone K^* is given by $K^* = K_1^* \times \cdots \times K_n^*$ and is a closed pointed cone with non–empty interior.

Proof Let $(y_1, \dots, y_n) \in K^*$. Then, we have $\sum_{i=1}^n x_i \bullet_i y_i \geq 0$ for all $(x_1, \dots, x_n) \in K$. We claim that $x_i \bullet_i y_i \geq 0$ for all $x_i \in K_i$, where $i = 1, \dots, n$. Indeed, suppose that this is not the case. Then, there exists a $j \in \{1, \dots, n\}$ and an $\bar{x}_j \in K_j$ such that $\bar{x}_j \bullet_j y_j < 0$. Since K_j is a cone, we have $\alpha \bar{x}_j \in K$ for any $\alpha > 0$. Now, let $x_i' \in K_i$ be arbitrary, where $i \in \{1, \dots, j-1, j+1, \dots, n\}$. Then, we have

$$\sum_{i \neq j} x_i' \bullet_i y_i + (\alpha \bar{x}_j) \bullet_j y_j < 0$$

for sufficiently large α, which is a contradiction. Hence, we have $x_i \bullet_i y_i$ for all $x_i \in K_i$, where $i = 1, \dots, n$. In particular, we have $(y_1, \dots, y_n) \in K_1^* \times \cdots \times K_n^*$.

Conversely, suppose that $(y_1, \dots, y_n) \in K_1^* \times \cdots \times K_n^*$. Then, we have $x_i \bullet_i y_i \geq 0$ for all $x_i \in K_i$, where $i = 1, \dots, n$. It follows that $\sum_{i=1}^n x_i \bullet_i y_i \geq 0$ for all $(x_1, \dots, x_n) \in K$; i.e., $(y_1, \dots, y_n) \in K^*$.

The claim that K^* is a closed pointed cone with non–empty interior follows from the assumption on K and Corollary 1. □

The above discussion leads us to the following dual of Problem (P), whose objective value gives the largest lower bound on v_p^*:

$$\begin{array}{rccl} v_d^* & = & \sup & b^T y \\ & & \text{subject to} & \displaystyle\sum_{i=1}^{m} y_i a_i + s = c, \\ & & & y \in \mathbb{R}^m,\ s \succeq_{K^*} \mathbf{0}. \end{array} \qquad (D)$$

By Corollary 1, if K is a closed pointed cone with non–empty interior, then so is K^*. In this case, both (P) and (D) are of the same nature; i.e., they both involve optimizing a linear function over a set defined by linear equality constraints and a conic constraint that is associated with a closed pointed cone with non–empty interior.

Before we study the relationship between the primal–dual pair of problems (P) and (D), let us consider some concrete instances of them.

Example 2 (Representative CLP Problems)

1. **Linear Programming (LP).** *By taking* $E = \mathbb{R}^n$, $K = \mathbb{R}^n_+$, *and* $u \bullet v = u^T v$ *for* $u, v \in E$, *Problem* (P) *becomes*

$$\begin{array}{lll} \inf & c^T x & \\ \text{subject to} & a_i^T x = b_i & \text{for } i = 1, \ldots, m, \\ & x \in \mathbb{R}^n_+, & \end{array}$$

which is nothing but an LP in standard primal form. To obtain an explicit description of its dual problem (D)*, we need to compute* $K^* = (\mathbb{R}^n_+)^*$*. We claim that* $\mathbb{R}^n_+$ *is* self–dual*; i.e.,* $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$*. Indeed, on one hand, we have* $\mathbb{R}^n_+ \subset (\mathbb{R}^n_+)^*$ *because* $y^T x \geq 0$ *if* $x, y \geq \mathbf{0}$*. On the other hand, suppose that* $y \in (\mathbb{R}^n_+)^*$*. Then, we have* $x^T y \geq 0$ *for all* $x \in \mathbb{R}^n_+$*. In particular, we have* $e_i^T y = y_i \geq 0$ *for* $i = 1, \ldots, n$*, where* $e_i \in \mathbb{R}^n$ *is the* i*–th standard basis vector. This shows that* $y \in \mathbb{R}^n_+$*, as desired.*

Based on the above discussion, we see that Problem (D) *becomes*

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & \displaystyle\sum_{i=1}^{m} y_i a_i + s = c, \\ & y \in \mathbb{R}^m,\ s \in \mathbb{R}^n_+, \end{array}$$

which is an LP in standard dual form.

2. **Second–Order Cone Programming (SOCP).** *Let* $E = \mathbb{R}^{n+1}$, $K = \mathcal{Q}^{n+1}$, *and* $u \bullet v = u^T v$ *for* $u, v \in E$*. Then, Problem* (P) *becomes*

$$\begin{array}{lll} \inf & c^T x & \\ \text{subject to} & a_i^T x = b_i & \text{for } i = 1, \ldots, m, \\ & x \in \mathcal{Q}^{n+1}, & \end{array}$$

which is an SOCP in standard primal form. To derive its dual, we need to compute $(\mathcal{Q}^{n+1})^*$*. It is an easy exercise to show that* $\mathcal{Q}^{n+1}$ *is self–dual; i.e.,* $(\mathcal{Q}^{n+1})^* = \mathcal{Q}^{n+1}$*. Thus, Problem* (D) *becomes*

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & \displaystyle\sum_{i=1}^{m} y_i a_i + s = c, \\ & y \in \mathbb{R}^m,\ s \in \mathcal{Q}^{n+1}, \end{array} \tag{4}$$

which is an SOCP in standard dual form. To get a better understanding of Problem (4)*, let us write it out explicitly. Towards that end, let* $a_i = (u_i, a_{i,1}, \ldots, a_{i,n}) \in \mathbb{R}^{n+1}$ *for* $i = 1, \ldots, m$ *and* $c = (v, d) \in \mathbb{R}^{n+1}$ *with* $d \in \mathbb{R}^n$*. Then, we have*

$$\sum_{i=1}^{m} y_i a_i = (u^T y, A^T y),$$

where $A \in \mathbb{R}^{m \times n}$ *is the matrix whose* i*–th row contains the entries* $a_{i,1}, \ldots, a_{i,n}$*. It follows that the constraint* $s = c - \sum_{i=1}^{m} y_i a_i \in \mathcal{Q}^{n+1}$ *is equivalent to* $(v - u^T y, d - A^T y) \in \mathcal{Q}^{n+1}$, *which implies that Problem* (4) *takes the form*

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & (v - u^T y, d - A^T y) \in \mathcal{Q}^{n+1}. \end{array} \tag{5}$$

In other words, the problem of optimizing a linear function subject to the constraint that the image of an affine map belongs to a second–order cone is an SOCP.

A natural extension of Problem (5) *is to allow multiple second–order cone constraints; i.e.,*

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & (v_j - (u^j)^T y, d^j - (A^j)^T y) \in \mathcal{Q}^{n_j+1} \qquad \textit{for } j = 1, \ldots, p, \end{array} \tag{6}$$

where $u^j \in \mathbb{R}^m$*,* $v_j \in \mathbb{R}$*,* $A^j \in \mathbb{R}^{m \times n_j}$*, and* $d^j \in \mathbb{R}^{n_j}$ *for* $i = 1, \ldots, p$*. This can also be viewed as an SOCP, as it can be put into the form*

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & \displaystyle\sum_{i=1}^{m} y_i \bar{a}_i + \bar{s} = \bar{c}, \\ & y \in \mathbb{R}^m, \; s \in \mathcal{Q}^{n_1+1} \times \cdots \times \mathcal{Q}^{n_p+1}. \end{array}$$

Here, $\bar{c} = (c^1, \ldots, c^p)$*,* $\bar{s} = (s^1, \ldots, s^p)$*, and* $\bar{a}_i = (a_i^1, \ldots, a_i^p)$ *are vectors in* $\mathbb{R}^{n_1+1} \times \cdots \times \mathbb{R}^{n_p+1}$*, with*

$$c^j = (v_j, d^j) \in \mathbb{R}^{n_j+1} \quad \textit{and} \quad a_i^j = (u_i^j, a_{i,1}^j, \ldots, a_{i,n_j}^j) \in \mathbb{R}^{n_j+1}.$$

Using (6)*, it is immediate that the class of SOCPs includes the class of LPs as a special case. Indeed, the standard form LP (1) can be formulated as the following SOCP:*

$$\begin{array}{ll} \inf & c^T x \\ \text{subject to} & \|Ax - b\|_2 \le 0, \\ & \|\mathbf{0}x - \mathbf{0}\|_2 \le e_i^T x \quad \textit{for } i = 1, \ldots, n. \end{array}$$

3. **Semidefinite Programming (SDP).** *By taking* $E = \mathcal{S}^n$*,* $K = \mathcal{S}_+^n$*,* $X \bullet Y = \mathrm{tr}(X^T Y)$ *for* $X, Y \in E$*, Problem* (P) *becomes*

$$\begin{array}{ll} \inf & C \bullet X \\ \text{subject to} & A_i \bullet X = b_i \quad \textit{for } i = 1, \ldots, m, \\ & X \in \mathcal{S}_+^n, \end{array}$$

which is an SDP in standard primal form. We leave it as an exercise to the reader to show that $\mathcal{S}_+^n$ *is also self–dual; i.e.,* $(\mathcal{S}_+^n)^* = \mathcal{S}_+^n$*. Based on this result, Problem* (D) *becomes*

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & \displaystyle\sum_{i=1}^{m} y_i A_i + S = C, \\ & y \in \mathbb{R}^m, \; S \in \mathcal{S}_+^n, \end{array}$$

which is an SDP in standard dual form.

Note that the class of SDPs includes the class of SOCPs as a special case. To see this, we first observe that the SOCP (5) *is equivalent to*

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & \begin{bmatrix} (v - u^T y) I & d - A^T y \\ (d - A^T y)^T & v - u^T y \end{bmatrix} \succeq \mathbf{0}. \end{array} \tag{7}$$

Indeed, let $\bar{y} \in \mathbb{R}^m$ *be a feasible solution to Problem (5). Then, we have* $v - u^T \bar{y} \geq 0$*. If* $v - u^T \bar{y} = 0$*, then* $d - A^T \bar{y} = \mathbf{0}$*, which shows that* $\bar{y} \in \mathbb{R}^m$ *is feasible for (7). On the other hand, if* $v - u^T \bar{y} > 0$*, then we can write the constraint in (5) as*

$$v - u^T \bar{y} - \frac{\|d - A^T \bar{y}\|_2^2}{v - u^T \bar{y}} \geq 0, \tag{8}$$

which, by the Schur complement, is equivalent to

$$\begin{bmatrix} (v - u^T \bar{y}) I & d - A^T \bar{y} \\ (d - A^T \bar{y})^T & v - u^T \bar{y} \end{bmatrix} \succeq \mathbf{0}; \tag{9}$$

i.e., $\bar{y} \in \mathbb{R}^m$ *is feasible for (7).*

Conversely, let $\bar{y} \in \mathbb{R}^m$ *be a feasible solution to Problem (7). Then, we have* $(v - u^T \bar{y}) I \succeq \mathbf{0}$*, which implies that* $v - u^T \bar{y} \geq 0$*. Now, if* $v - u^T \bar{y} = 0$*, then* $(v - u^T \bar{y}) I = \mathbf{0}$*, which, by (9), implies that* $d - A^T \bar{y} = \mathbf{0}$*. On the other hand, if* $v - u^T \bar{y} > 0$*, then constraint (9) is equivalent to constraint (8). In either case, we see that* $\bar{y} \in \mathbb{R}^m$ *is feasible for (4). Hence, we conclude that* (5) *and* (7) *are equivalent.*

We now leave it as an exercise to the reader to show that Problem (7) can be expressed as an SDP in standard dual form.

Let us now return to study the relationship between the primal–dual pair of CLPs (P) and (D). Our first result is the CLP Weak Duality Theorem, which essentially follows from our construction of (D).

Theorem 1 (CLP Weak Duality) *Let* $\bar{x} \in K$ *be feasible for* (P) *and* $(\bar{y}, \bar{s}) \in \mathbb{R}^m \times K^*$ *be feasible for* (D)*. Then,* $b^T \bar{y} \leq c \bullet \bar{x}$*.*

Proof We compute

$$c \bullet \bar{x} = \left(\sum_{i=1}^m \bar{y}_i a_i + \bar{s} \right) \bullet x = \sum_{i=1}^m \bar{y}_i (a_i \bullet x) + \bar{s} \bullet \bar{x} = b^T \bar{y} + \bar{s} \bullet \bar{x} \geq b^T \bar{y},$$

where the last inequality follows from the fact that $\bar{x} \in K$ and $\bar{s} \in K^*$. This completes the proof. □

To establish strong duality between (P) and (D), one natural approach is to develop a conic version of the Farkas lemma. Towards that end, recall that the classic Farkas lemma is concerned with the following alternative *linear* systems:

$$(A) \quad a_i^T x = b_i \quad \text{for } i = 1, \ldots, m, \;\; x \in \mathbb{R}_+^n.$$

$$(B) \quad -\sum_{i=1}^m y_i a_i \in \mathbb{R}_+^n, \;\; b^T y > 0,$$

where $a_1, \dots, a_m \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ are given. This suggests that a Farkas–type lemma should hold for the following *conic linear* systems:

$$\text{(I)} \quad a_i \bullet x = b_i \quad \text{for } i = 1, \dots, m, \;\; x \in K.$$

$$\text{(II)} \quad -\sum_{i=1}^{m} y_i a_i \in K^*, \;\; b^T y > 0.$$

Here, as before, E is a finite–dimensional Euclidean space equipped with an inner product $\bullet$, $K \subset E$ is a closed pointed cone with non–empty interior, and $a_1, \dots, a_m \in E$ and $b \in \mathbb{R}^m$ are given vectors. It is straightforward to show that (I) and (II) cannot both have solutions. Indeed, suppose that $\bar{x} \in E$ is a solution to (I) and $\bar{y} \in \mathbb{R}^m$ is a solution to (II). Then, we have

$$0 < b^T y = \sum_{i=1}^{m} y_i (a_i \bullet x) = -\left[\left(-\sum_{i=1}^{m} y_i a_i \right) \bullet x \right] \leq 0,$$

which is a contradiction. Unfortunately, as the following example shows, it is possible that neither (I) nor (II) has a solution.

Example 3 ("Failure" of the Conic Farkas Lemma) *Let $E = \mathcal{S}^2$ and $K = \mathcal{S}^2_+$. Define*

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

Consider the following systems:

$$\text{(I)} \quad A_1 \bullet X = b_1, \;\; A_2 \bullet X = b_2, \;\; X = \begin{bmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{bmatrix} \in \mathcal{S}^2_+.$$

$$\text{(II)} \quad -(y_1 A_1 + y_2 A_2) \in \mathcal{S}^2_+, \;\; b^T y > 0.$$

Observe that (I) *is equivalent to*

$$X_{11} = 0, \;\; X_{12} = 1, \;\; X \in \mathcal{S}^2_+$$

and (II) *is equivalent to*

$$-\begin{bmatrix} y_1 & y_2 \\ y_2 & 0 \end{bmatrix} \in \mathcal{S}^2_+, \;\; y_2 > 0,$$

both of which are insolvable.

To see why a Farkas–type lemma need not hold for the systems (I) and (II), recall that in the proof of Farkas' lemma for linear systems, we show that the set $S = \{Ax \in \mathbb{R}^m : x \in \mathbb{R}^n_+\}$ is closed and convex, so that we can apply the separation theorem to $\{b\}$ and S if system (A) is insolvable. However, for a general closed pointed cone K, the set $S' = \{(a_1 \bullet x, \dots, a_m \bullet x) \in \mathbb{R}^m : x \in K\}$ need not be closed. Indeed, in the setting of Example 3, we have

$$S' = \left\{ (A_1 \bullet X, A_2 \bullet X) : X \in \mathcal{S}^2_+ \right\} = \left\{ (X_{11}, 2X_{12}) : X \in \mathcal{S}^2_+ \right\}.$$

Since $X \in \mathcal{S}_+^2$ if and only if all of its principal minors are non–negative, it can be verified that

$$S' = \{(0,0)\} \cup \{(x,y) : x > 0, y \in \mathbb{R}\} \subset \mathbb{R}^2,$$

which is clearly not closed. Such observation raises a natural question: When is the linear image of a closed convex cone closed? In [2], Pataki addressed this question and gave a fairly complete answer. We shall not discuss Pataki's results here and refer the interested reader to his paper.

The above discussion shows that the systems (I) and (II) may not be alternatives to each other. However, not all is lost. In fact, by imposing extra conditions, one can establish a conic version of the Farkas lemma for systems (I) and (II).

Theorem 2 (Conic Farkas' Lemma) *Let E be a finite–dimensional Euclidean space equipped with an inner product $\bullet$, $K \subset E$ be a closed pointed cone with non–empty interior, and $a_1, \ldots, a_m \in E$ and $b \in \mathbb{R}^m$ be given vectors. Suppose that the* Slater condition *holds; i.e., there exists a $\bar{y} \in \mathbb{R}^m$ such that $-\sum_{i=1}^m \bar{y}_i a_i \in \text{int}(K^*)$. Then, exactly one of the systems* (I) *and* (II) *has a solution.*

Using Theorem 2, we can establish the following CLP Strong Duality Theorem:

Theorem 3 (CLP Strong Duality)

(a) *Suppose that (P) is bounded below and* strictly feasible*; i.e., there exists a feasible solution $\bar{x}$ to (P) such that $\bar{x} \in \text{int}(K)$. Then, we have $v_p^* = v_d^*$. Moreover, there exists a feasible solution $(\bar{y}, \bar{s})$ to (D) such that $b^T\bar{y} = v_d^* = v_p^*$; i.e., the common optimal value is attained by some dual feasible solution.*

(b) *Suppose that (D) is bounded above and* strictly feasible*; i.e., there exists a feasible solution $(\bar{y}, \bar{s})$ to (D) such that $\bar{s} \in \text{int}(K^*)$. Then, we have $v_p^* = v_d^*$. Moreover, there exists a feasible solution $\bar{x}$ to (P) such that $c \bullet \bar{x} = v_p^* = v_d^*$; i.e., the common optimal value is attained by some primal feasible solution.*

(c) *Suppose that either (P) or (D) is bounded and strictly feasible. Then, given a feasible solution $\bar{x}$ to (P) and a feasible solution $(\bar{y}, \bar{s})$ to (D), the following are equivalent:*

- *$\bar{x}$ and $(\bar{y}, \bar{s})$ are optimal for (P) and (D), respectively.*
- *The* duality gap *is zero; i.e., $c \bullet \bar{x} = b^T\bar{y}$.*
- *We have* complementary slackness*; i.e., $\bar{x} \bullet \bar{s} = 0$.*

Upon closer inspection, we see that the CLP Strong Duality Theorem is weaker than the LP Strong Duality Theorem. Indeed, recall from Handout 3 that in the case of LP, whenever one of (P) or (D) is *bounded* and *feasible*, then (i) the optimal value v_p^* of (P) and the optimal value v_d^* of (D) are equal, and (ii) there exists a primal feasible solution $\bar{x}$ and a dual feasible solution $(\bar{y}, \bar{s})$ such that $c^T\bar{x} = v_p^* = v_d^* = b^T\bar{y}$. In other words, primal and dual attainment of the common optimal value is implied by the boundedness and feasibility of either the primal or the dual LP problem. However, such a claim is absent from the statement of Theorem 3. It is thus natural to ask whether one can strengthen the conclusion of Theorem 3 so that it directly generalizes the LP Strong Duality Theorem. Curiously, as the following examples show, several things can go wrong in the case of CLP:

Example 4 (Pathologies in Conic Duality)

1. Both the primal problem (P) and the dual problem (D) are bounded and feasible, but the duality gap is non–zero.

Consider the SDP

$$\begin{array}{ll} \inf & X_{12} \\ \text{subject to} & X = \begin{bmatrix} 0 & X_{12} & 0 \\ X_{12} & X_{22} & 0 \\ 0 & 0 & 1+X_{12} \end{bmatrix} \in \mathcal{S}^3_+. \end{array} \tag{10}$$

It is a routine exercise to show that the dual of (10) *is given by*

$$\begin{array}{ll} \sup & y_4 \\ \text{subject to} & S = \begin{bmatrix} -y_1 & (1+y_4)/2 & -y_2/2 \\ (1+y_4)/2 & 0 & -y_3/2 \\ -y_2/2 & -y_3/2 & -y_4 \end{bmatrix} \in \mathcal{S}^3_+. \end{array} \tag{11}$$

Since $X \in \mathcal{S}^3_+$, *we must have* $X_{12} = 0$, *which implies that the optimal value of (10) is* 0. *Similarly, since* $S \in \mathcal{S}^3_+$, *we must have* $(1+y_4)/2 = 0$, *or equivalently,* $y_4 = -1$. *Hence, the optimal value of (11) is* -1.

2. The primal problem (P) is bounded below and strictly feasible, but the optimal value is not attained by any primal feasible solution.

Consider the SOCP

$$\begin{array}{ll} \inf & x_1 \\ \text{subject to} & (x_1+x_2, 1, x_1-x_2) \in \mathcal{Q}^3. \end{array} \tag{12}$$

Note that the constraint in (12) is equivalent to

$$x_1 + x_2 \ge \sqrt{1 + (x_1 - x_2)^2},$$

which in turn is equivalent to

$$4x_1x_2 \ge 1, \quad x_1 + x_2 > 0. \tag{13}$$

By (13), we see that the optimal value of (12) is bounded below by 0. *Moreover, for* $x_1 = x_2 = 1$, *we have* $(2,1,0) \in \text{int}(\mathcal{Q}^3)$, *which implies that (12) is strictly feasible. Now, by setting* $x_1 = 1/(4x_2)$ *and letting* $x_2 \to \infty$, *we see that the optimal value of (12) is* 0. *However, such an optimal value is not attained by any feasible solution to (12).*

It is instructive to inspect the dual of (12), *which is given by*

$$\begin{array}{ll} \sup & -y_2 \\ \text{subject to} & y_1 + y_3 = 1, \\ & y_1 - y_3 = 0, \\ & (y_1, y_2, y_3) \in \mathcal{Q}^3. \end{array} \tag{14}$$

The feasible set of (14) *is* $\{(1/2, 0, 1/2)\}$, *which shows that the optimal value of* (14) *is* 0. *However, it is clear that* (14) *is not strictly feasible.*

For further results on CLP duality, we refer the reader to [1, 3, 4].

References

[1] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Duality and Self–Duality for Conic Convex Programming. Technical Report No. EI 9620/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1996.

[2] G. Pataki. On the Closedness of the Linear Image of a Closed Convex Cone. *Mathematics of Operations Research*, 32(2):395–412, 2007.

[3] I. Pólik and T. Terlaky. Exact Duality for Optimization over Symmetric Cones. Technical Report No. 2007/10, Advanced Optimization Laboratory, McMaster University, Hamilton, ON, Canada, 2007.

[4] S. P. Schurr, A. L. Tits, and D. P. O'Leary. Universal Duality in Conic Convex Optimization. *Mathematical Programming, Series A*, 109(1):69–88, 2007.