7 Duality in conic programming (continued)

Recall the primal-dual pair of conic programs:

$$\begin{array}{ll} \text{minimise} & \langle c, x\rangle \\ \text{subject to} & \mathcal{A}(x) = b \\ & x \in K \end{array} \tag{1}$$

and

$$\begin{array}{ll} \underset{y\in\mathbb{R}^m, z\in\mathbb{R}^n}{\text{maximise}} & \langle b, y\rangle \\ \text{subject to} & c = z + \mathcal{A}^*(y) \\ & z \in K^*. \end{array} \tag{2}$$

In this lecture we will prove the following theorem.

Theorem 7.1 (Duality for conic programs)**.** *Consider the conic program (1) and let p^* be its optimal value. Also let d^* be the optimal value of the dual program (2). Then the following holds:*

(i) Weak duality: $p^* \geq d^*$

(ii) Strong duality: If the problem (1) is strictly feasible (i.e., there exists $x \in \operatorname{int}(K)$ such that $\mathcal{A}(x) = b$) then $p^* = d^*$.

The condition that there exists $x \in \operatorname{int}(K)$ satisfying $\mathcal{A}(x) = b$ is known as *Slater's condition.* It is a condition that guarantees strong duality.

Proof. Weak duality has been proved in last lecture (see Equation 6 of Lecture 6). We are now going to prove strong duality under the assumption that (1) is strictly feasible. Note that we can assume p^* to be finite: if $p^* = -\infty$ then $d^* = -\infty$ by weak duality, i.e., the dual problem is infeasible.

The following two lemmas are the key part of the proof.

Lemma 1. *Let $K \subseteq \mathbb{R}^n$ be a proper cone, L a linear subspace of $\mathbb{R}^n$ and assume that $K^* + L^\perp$ is closed. Then $(K \cap L)^* = K^* + L^\perp$ (where $K^* + L^\perp \overset{def}{=} \{y + a : y \in K^*, a \in L^\perp\}$).*

Proof. The inclusion $\supseteq$ is easy and simply corresponds to weak duality. We have to prove the inclusion $(K \cap L)^* \subseteq K^* + L^\perp$. To do so we will show instead that $K \cap L \supseteq (K^* + L^\perp)^*$, and the desired inclusion would then follow by the (easy) fact that $A \subseteq B \Rightarrow A^* \supseteq B^*$ and Theorem 2.2, using the assumption that $K^* + L^\perp$ is closed to guarantee that $(K^* + L^\perp)^{**} = K^* + L^\perp$. The inclusion $K \cap L \supseteq (K^* + L^\perp)^*$ is not difficult to show: Assume $z \in (K^* + L^\perp)^*$, i.e., $\langle z, y + a\rangle \geq 0$ for any $y \in K^*$ and $a \in L^\perp$. We want to show that $z \in K \cap L$. Taking $a = 0$ tells us that $\langle z, y\rangle \geq 0$ for any $y \in K^*$, thus $z \in K^{**} = K$ (since K is closed and convex). Similarly if we take $y = 0$ we get that $\langle z, a\rangle \geq 0$ for any $a \in L^\perp$. Since $L^\perp$ is a subspace this also tells us (since $\pm a \in L^\perp$) that $\langle z, a\rangle = 0$ for all $a \in L^\perp$. Thus $z \in (L^\perp)^\perp = L$. We have thus shown that $z \in K \cap L$ as we wanted. This completes the proof of Lemma 1. □

The next lemma connects the Slater condition with the condition that $K^* + L^\perp$ is closed.

Lemma 2. *Assume that $\operatorname{int}(K) \cap L \neq \emptyset$. Then $K^* + L^\perp$ is closed.*

Proof. Let $z_k = y_k + a_k$ be a sequence in $K^* + L^\perp$, with $y_k \in K^*$ and $a_k \in L^\perp$ and assume that $z_k \to z$. We have to show that $z \in K^* + L^\perp$. The main part of the proof is to show that the sequence (y_k) is bounded using the assumption $\operatorname{int}(K) \cap L \neq \emptyset$. After this the proof will be simple.

Let $x_0 \in \operatorname{int}(K) \cap L$. First observe that

$$\langle x_0, y_k \rangle = \langle x_0, z_k - a_k \rangle = \langle x_0, z_k \rangle \to \langle x_0, z \rangle$$

and so the sequence $(\langle x_0, y_k \rangle)$ bounded. Consider $\bar{y}_k = y_k / \|y_k\|$. Since $\bar{y}_k$ is bounded we know it converges (after extracting subsequence) to some $\bar{y}$. Since $\bar{y} \in K^* \setminus \{0\}$ and $x_0 \in \operatorname{int}(K)$ we have $\langle x_0, \bar{y} \rangle > 0$. But then if (y_k) was unbounded we would have $\langle x_0, \bar{y}_k \rangle = \langle x_0, y_k \rangle \frac{1}{\|y_k\|} \to 0$ since $\langle x_0, y_k \rangle$ is bounded, which would be a contradiction. Thus we have shown that (y_k) is bounded.

Since (y_k) is bounded we know that it converges (after extracting subsequence) to some $y \in K$. Thus $a_k = z_k - y_k$ is also bounded and converges to some $a \in L^\perp$. Finally we have $z = y + a \in K^* + L^\perp$ as desired. $\square$

If we combine Lemmas 1 and 2 we get that if $\operatorname{int}(K) \cap L \neq \emptyset$ then $(K \cap L)^* = K^* + L^\perp$. We now see how to use this fact to prove strong duality.

Define:

$$\tilde{K} = K \times \mathbb{R}_+, \quad \tilde{L} = \{(x,t) : \mathcal{A}(x) = tb\} \subseteq \mathbb{R}^{n+1}, \text{ and } \langle \tilde{c}, \begin{bmatrix} x \\ t \end{bmatrix} \rangle = \langle c, x \rangle - p^* t.$$

Our assumption that $\langle c, x \rangle \geq p^*$ for all $x \in K$ such that $\mathcal{A}(x) = b$ means that $\langle \tilde{c}, \tilde{x} \rangle \geq 0$ for all $\tilde{x} \in \tilde{K} \cap \tilde{L}$. Furthermore we know that there exists $x_0 \in \operatorname{int}(K)$ such that $\mathcal{A}(x_0) = b$. Thus we know that there exist $\tilde{z} = (z, \alpha) \in \tilde{K}^* \subseteq \mathbb{R}^n \times \mathbb{R}$ and $\tilde{y} \in \tilde{L}^\perp \subseteq \mathbb{R}^n \times \mathbb{R}$ such that

$$\tilde{c} = \tilde{z} + \tilde{y}. \tag{3}$$

It is not difficult to verify that $\tilde{K}^* = K^* \times \mathbb{R}_+$ and $\tilde{L}^\perp = \{(\mathcal{A}^*(y), -\langle b, y \rangle) : y \in \mathbb{R}^m\}$. Thus we get from (3) that: $c = z + \mathcal{A}^*(y)$ where $z \in K^*$ and $-p^* = \alpha - \langle b, y \rangle$ where $\alpha \geq 0$. The last equality implies that $p^* \leq \langle b, y \rangle$. Since we know by weak duality that $p^* \geq \langle b, y \rangle$ we have $p^* = \langle b, y \rangle$. This completes the proof. $\square$

How to compute duals To compute the dual of a given conic program, one way is to put it in the standard form (1), identity $\mathcal{A}, b, c$ and then use (2). This can be a bit tedious. Here we explain a simple way of computing duals of conic programs without having to put them in standard form.

1. We assume the problem is a **minimisation** problem (if it is a maximisation problem you can, for example, negate the objective to get a minimisation).

2. Process the constraints of the problem one by one. For each constraint, identify the *linear inequalities* that you can infer from it that will be valid for any point satisfying the constraint. For example if your constraint is "$x \in K$" then you can infer the linear inequality $\langle \lambda, x \rangle \geq 0$ where $\lambda \in K^*$. If your constraint is "$Ax \leq b$", (componentwise inequality) you can infer the inequality $\langle \lambda, b - Ax \rangle \geq 0$ where $\lambda \geq 0$. Finally if your constraint is $Ax = b$ you can infer the (in)equality $\langle \lambda, Ax - b \rangle = 0$ where λ is arbitrary. In general any constraint will give rise to a certain *dual variable* (here λ). Let's say for example your problem is:

$$\underset{x}{\text{minimise}} \quad \langle c, x \rangle \quad \text{s.t. } b - \mathcal{A}(x) \in K. \tag{4}$$

Then from the single constraint we can infer $\langle \lambda, b - \mathcal{A}(x) \rangle \geq 0$ assuming $\lambda \in K^*$.

3. Having processed all the constraints, we now have identified linear inequalities that are valid on our feasible set. What we want is to have these linear inequalities say something about our objective function (say $\langle c, x \rangle$). This imposes certain linear equalities on our dual variables. For the example (4) the inequality $\langle \lambda, b - \mathcal{A}(x) \rangle \geq 0$ can be rewritten as $\langle -\mathcal{A}^*(\lambda), x \rangle \geq -\langle b, \lambda \rangle$. Since we are interested in the linear function $\langle c, x \rangle$ we want: $-\mathcal{A}^*(\lambda) = c$. The dual problem consists in finding the best lower bound on $\langle c, x \rangle$ one obtains this way. Thus the dual of (4) is:
$$\text{maximise} \quad -\langle b, \lambda \rangle \quad \text{s.t.} \; -\mathcal{A}^*(\lambda) = c, \lambda \in K^*. \tag{5}$$

Here is another example: let us start with the conic program in standard form (1) and explain how to get the dual (2):
$$\text{minimise} \quad \langle c, x \rangle \quad \text{s.t.} \; \mathcal{A}(x) = b, x \in K. \tag{6}$$
There are two constraints and so we will have two dual variables. For the first constraint "$\mathcal{A}(x) = b$", the only valid (in)equalities that we can write are $\langle y, \mathcal{A}(x) - b \rangle = 0$ where y can be arbitrary. The second constraint is "$x \in K$" and we know that the valid inequalities we can infer from this constraint are of the form $\langle z, x \rangle \geq 0$ where $z \in K^*$. Thus we know that for any x feasible of (6) and any y and $z \in K^*$ we have $\langle y, \mathcal{A}(x) - b \rangle + \langle z, x \rangle \geq 0$. Rearranging this inequality to group together the linear term gives $\langle \mathcal{A}^*(y) + z, x \rangle \geq \langle b, y \rangle$. Since we are interested in the cost function $\langle c, x \rangle$ we want to have $c = \mathcal{A}^*(y) + z$. In other words, for any y and $z \in K^*$ satisfying $c = \mathcal{A}^*(y) + z$ we have $p^* \geq \langle b, y \rangle$. Thus the dual problem is
$$\text{maximise} \quad \langle b, y \rangle \quad \text{s.t.} \; c = z + \mathcal{A}^*(y), z \in K^*$$
as we saw before.