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 lemma is 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 $L\cap\operatorname{int}(K) \neq \emptyset$. Assume $c \in \mathbb{R}^n$ satisfies $\langle c, x\rangle \geq 0$ for all $x \in K \cap L$. Then there exist $c_1 \in K^*$ and $c_2 \in L^\perp$ such that $c = c_1 + c_2$.*

Proof. What we need to prove is that $(K\cap L)^* = K^* + L^\perp$ (actually, just the inclusion $\subseteq$) where the summation in the right-hand side indicates Minkowski sum (i.e., $A + B = \{a + b : a \in A, b \in B\}$). Note that the inclusion $\supseteq$ is easy: if $\alpha \in K^*$ and $\beta \in L^\perp$ then for any $x \in K \cap L$ we have $\langle \alpha + \beta, x\rangle = \langle \alpha, x\rangle + \langle \beta, x\rangle \geq 0$ where $\langle \alpha, x\rangle \geq 0$ since $\alpha \in K^*$ and $x \in K$ and $\langle \beta, x\rangle = 0$ because $\beta \in L^\perp$ and $x \in L$.

It thus remains 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 *assuming $K^* + L^\perp$ is closed* (which we show later – this is where we will use the fact that $L \cap \operatorname{int}(K) \neq \emptyset$, see Lemma 2 to follow). 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.

The remaining part is to show that $K^* + L^\perp$ is indeed closed so we can invoke the fact that $(K^* + L^\perp)^{**} = K^* + L^\perp$. This is where we use the fact that $L \cap \operatorname{int}(K) \neq \emptyset$.

Lemma 2. *Assume that $L \cap \operatorname{int}(K) \neq \emptyset$. 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 $L \cap \operatorname{int}(K) \neq \emptyset$. After this the proof will be simple.

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. □

□

We now show to use Lemma 1 to prove strong duality. We apply Lemma 1 with

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

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 $\mathcal{A}(x_0) = b$. This means that the point $\tilde{x_0} = (x_0, 1)$ lives in $\operatorname{int}(\tilde{K}) \cap \tilde{L}$. The assumptions of Lemma 1 are satisfied, thus we know that there exist $\tilde{c}_1 = (z, \alpha) \in \tilde{K}^* \subseteq \mathbb{R}^n \times \mathbb{R}$ and $\tilde{c}_2 \in \tilde{L}^\perp \subseteq \mathbb{R}^n \times \mathbb{R}$ such that

$$\tilde{c} = \tilde{c}_1 + \tilde{c}_2. \tag{3}$$

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. □

Example where strong duality does not hold Consider the following simple semidefinite program:

$$\begin{array}{ll} \underset{X \in \mathbf{S}^2}{\text{minimise}} & 2X_{12} \\ \text{subject to} & X_{11} = 0, X \succeq 0. \end{array}$$

Here the SDP is specified by $C = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, $\mathcal{A}(X) = X_{11}$ and $b = 0$. The adjoint of $\mathcal{A}$ is $\mathcal{A}^*(y) = \left[\begin{smallmatrix} y & 0 \\ 0 & 0 \end{smallmatrix}\right]$. The dual program is

$$\begin{array}{ll} \underset{y,Z}{\text{maximise}} & 0 \\ \text{subject to} & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = Z + \begin{bmatrix} y & 0 \\ 0 & 0 \end{bmatrix} \\ & Z \succeq 0. \end{array}$$

The value of the primal problem is $p^* = 0$. However the dual problem is infeasible and so $d^* = -\infty$.

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 simply 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).

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 requires saying that the cost vector (for example c, if the objective function is $\langle c, x \rangle$) is a certain sum involving the dual variables.

It is perhaps easier to explain with examples.

1. Let us start with the conic program in standard form (1) and explain how to get the dual (2):

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

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 (4) 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.

2. Consider now another example:

$$\text{minimise} \quad \langle c, x \rangle \quad \text{s.t. } b - \mathcal{A}(x) \in K. \tag{5}$$

This problem has just one constraint "$b - \mathcal{A}(x) \in K$". The linear inequalities we can infer from this constraint are of the form: $\langle z, b - \mathcal{A}(x) \rangle \geq 0$ where $z \in K^*$. Rearranging this inequality to separate the linear term from the constant term we have $\langle -\mathcal{A}^*(z), x \rangle \geq -\langle b, z \rangle$. Since we are interested in the cost function $\langle c, x \rangle$ we want to have $c = -\mathcal{A}^*(z)$. In other words, for any $z \in K^*$ satisfying $c = -\mathcal{A}^*(z)$ we have $p^* \geq -\langle b, z \rangle$ where p^* is the optimal value of (6). Thus the dual problem of (1) is

$$\text{maximise} \quad -\langle b, z \rangle \quad \text{s.t. } -\mathcal{A}^*(z) = c, z \in K^*. \tag{6}$$

Exercise 7.1 (Examples of duals). *Compute the duals of the following problems:*

1. $\underset{x,y\in\mathbb{R}^n}{\text{minimise}} \sum_{i=1}^n y_i$ *s.t.* $y + x \geq 0, y - x \geq 0, Ax = b$

2. $\underset{X,Y\in\mathbf{S}^n}{\text{minimise}} \operatorname{Tr}(Y)$ *s.t.* $Y + X \succeq 0, Y - X \succeq 0, \mathcal{A}(X) = b$

3. $\underset{x,y\in\mathbb{R}^n}{\text{minimise}} 2x + y$ *s.t.* $\begin{bmatrix} 1-x & y \\ y & 1+x \end{bmatrix} \succeq 0$

4. $\underset{X\in\mathbf{S}^n}{\text{minimise}} \operatorname{Tr}(CX)$ *s.t.* $X_{ii} = 1 \ \forall i = 1, \ldots, n, \quad X \succeq 0$

5. *Compute the dual of* (2) *and show that you get* (1)

Exercise 7.2 (Another proof of strong duality theorem). *In this exercise we look at a different proof of item (ii) in Theorem 7.1 (strong duality).*

1. *Before we start, we need a certain extension of the separating hyperplane theorem: show that if* $C, D \subseteq \mathbb{R}^n$ *are two disjoint convex sets, then there exists a hyperplane that separates* C *from* D*, i.e., there exists* $a \in \mathbb{R}^n \setminus \{0\}$ *and* $b \in \mathbb{R}$ *such that* $\langle a, x\rangle \geq b$ *for all* $x \in C$ *and* $\langle a, x\rangle \leq b$ *for all* $x \in D$*.*

2. *Consider now problems* (1) *and its dual* (2)*. We want to show that* $p^* = d^*$ *under the assumption that there exists* $x \in \operatorname{int}(K)$ *such that* $\mathcal{A}(x) = b$*. Let*

$$C = \Big\{(\langle c, x\rangle - p^*, x) : \mathcal{A}(x) = b\Big\} \subseteq \mathbb{R} \times \mathbb{R}^n.$$

   Using the assumption that $\langle c, x\rangle \geq p^*$ *for all* $x \in K$ *satisfying* $\mathcal{A}(x) = b$*, show that* C *is disjoint from another convex set* $D \subset \mathbb{R} \times \mathbb{R}^n$*.*

3. *Use the separating hyperplane theorem between* C *and* D*, and then the assumption that there exists* $x \in \operatorname{int}(K)$ *satisfying* $\mathcal{A}(x) = b$ *to show that there exists* $z \in K^*$ *such that*

$$\langle c, x\rangle - p^* - \langle z, x\rangle \geq 0 \quad \forall x : \mathcal{A}(x) = b. \tag{7}$$

4. *Show that inequality* (7) *implies that* $c - z \in \ker(\mathcal{A})^\perp = \operatorname{im}(\mathcal{A}^*)$*. Deduce that* $c = z + \mathcal{A}^*(y)$ *with* $\langle b, y\rangle = p^*$*.*

Exercise 7.3. *(Closure of image of a cone).*

1. *Let* K *be an arbitrary set in* $\mathbb{R}^n$ *and* $M : \mathbb{R}^n \to \mathbb{R}^m$ *a linear map. Show that* $M(K)$ *is closed if and only if* $K + \ker M$ *is closed.*

2. *Give an example of a proper cone* K *and linear map* M *such that* $M(K)$ *is not closed (hint: use* $K = \mathbf{S}^2_+$*).*

Exercise 7.4 (Farkas' lemma). *Let $\ell_1, \ldots, \ell_k$ be k linear forms on $\mathbb{R}^n$, and assume that ℓ is a linear form satisfying, for any $x \in \mathbb{R}^n$:*

$$\ell_1(x) \geq 0, \ldots, \ell_k(x) \geq 0 \Rightarrow \ell(x) \geq 0.$$

Show that there exist $\lambda_1, \ldots, \lambda_k \geq 0$ such that $\ell = \sum_{i=1}^k \lambda_i \ell_i$.

Exercise 7.5 (Polyhedral and finitely generated cones (1)). *A* polyhedral *cone K is a convex cone defined by a finite number of linear inequalities, i.e.,:*

$$K = \{x \in \mathbb{R}^n : Ax \geq 0\} \tag{8}$$

where $A \in \mathbb{R}^{m \times n}$. A cone K is called finitely generated *if is the conic hull of a finite number of vectors*

$$K = \operatorname{cone}(v_1, \ldots, v_k) = \left\{ \sum_{i=1}^k a_i v_i : a_1, \ldots, a_k \geq 0 \right\}. \tag{9}$$

The purpose of this exercise is to show that finitely generated and polyhedral cones are the same thing (i.e., K is finitely generated if and only if it is polyhedral).

1. *Show that polyhedral and finitely generated cones are closed.*
2. *Show that if K is polyhedral then K^* is finitely generated. Conversely show that if K is finitely generated then K^* is polyhedral (hint: see Exercise 7.4).*
3. *Show that if K is polyhedral then it is finitely generated (hint[a]: show that if x_0 is an extreme ray of a polyhedral cone K then there exists $I \subseteq \{1, \ldots, m\}$ such that $\ker(A_I) = \mathbb{R}x_0$, where A_I is the matrix obtained from A by keeping only the rows in I). How many extreme rays can K have at most?*
4. *Conversely show that if K is finitely generated, then it is polyhedral.*

[a] The hint was updated from a previous version of the exercise where the converse was also present, namely that if $Ax_0 \geq 0$ and $\ker(A_I) = \mathbb{R}x_0$ for some I then x_0 is an extreme ray. This direction however requires additional assumptions on K, for example that K is pointed. Also this direction is actually not needed for the question.

Exercise 7.6 (Polyhedral and finitely generated cones (2)). *In Exercise 7.5 we saw that a polyhedral cone is a cone having the form either (8) or (9).*

1. *Show that if K is polyhedral and L is a subspace then $K^* + L^\perp$ is closed.*
2. *Deduce that if K is polyhedral then strong duality holds in Theorem 7.1(ii) without the need for Slater condition (provided either the primal problem or the dual problem is feasible).*