6 Duality in conic programming

Motivating example in linear programming Consider the following simple linear program:

$$\begin{array}{ll} \text{minimise} & 2x+y \\ \text{subject to} & x+y+1 \geq 0 \\ & x+1 \geq 0 \\ & y+1 \geq 0 \\ & -x+1 \geq 0 \\ & -y+1 \geq 0 \end{array} \tag{1}$$

Let us call p^* the optimal value of (1).

- Finding an *upper bound* on p^* is "simple": if (x, y) is any feasible point then we know, by definition, that $p^* \leq 2x + y$. For example it is easy to verify that the point $(x, y) = (0, 0)$ is feasible. This tells us that $p^* \leq 0$. If we take $(x, y) = (-1, 0)$, which is also feasible, we get that $p^* \leq -2$.

- Consider now the more difficult question of finding a *lower bound* on p^*. How can we do this? One strategy is to take linear combinations of the constraints with nonnegative coefficients. For example if we multiply the second constraint $x+1$ by 2 and add it to the third constraint, we get that any feasible point of (1) must satisfy $2(x + 1) + y + 1 \geq 0$, i.e., $2x + y \geq -3$. In other words this tells us that $p^* \geq -3$. Is this the best possible lower bound we can get on p^* using this strategy? Let's try another combination: if we add the first constraint to the second constraint we get that any feasible (x, y) must satisfy $(x + y + 1) + (x + 1) \geq 0$, i.e., $2x + y \geq -2$. As a consequence we get $p^* \geq -2$.

To summarize: we have shown on the one hand that $p^* \leq -2$ by exhibiting a feasible point of (1) whose objective value is -2. On the other hand, by taking appropriate linear combinations with nonnegative coefficients of the constraints we have shown that $p^* \geq -2$. We have thus shown that $p^* = -2$.

Motivating example in semidefinite programming Consider the now the simple semidefinite programming:

$$\begin{array}{ll} \text{minimise} & 2x+y \\ \text{subject to} & \begin{bmatrix} 1-x & y \\ y & 1+x \end{bmatrix} \succeq 0. \end{array} \tag{2}$$

Let us call p^* the optimal value of (1). Consider again the problem of finding a *lower bound* on p^*. How can we generalise the idea of "taking linear combinations with nonnegative coefficients of the constraints" that we saw in the previous example? Here is how: assume $\left[\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right]$ is a 2×2 symmetric matrix that is positive semidefinite. Since the trace inner product of two positive semidefinite matrices is nonnegative, it follows that any feasible point (x, y) of (2) must satisfy the linear inequality:

$$\operatorname{Tr}\left(\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} 1-x & y \\ y & 1+x \end{bmatrix}\right) \geq 0. \tag{3}$$

Doing the calculation, this gives $a(1-x)+2by+c(1+x) \geq 0$ i.e., $(c-a)x+2by \geq -a-c$. Since our objective function in (2) is $2x+y$, we want a, b, c to satisfy $c-a=2$ and $b=1/2$. Any such choice of (a, b, c) will then tell us that $p^* \geq -a-c$. Consider now the specific choice $a=\alpha-1, c=\alpha+1, b=1/2$ where $\alpha=\sqrt{5}/2$. The matrix $\left[\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right]$ can be easily verified to be positive semidefinite: it trace is $2\alpha \geq 0$ and its determinant is $\alpha^2-1-1/4 \geq 0$. Using this choice of matrix we get that the optimal value p^* satisfies $p^* \geq -a-c=-\sqrt{5}$.

It is not difficult to verify that p^* is indeed equal to $-\sqrt{5}$: take $(x, y)=(-2,-1)/\sqrt{5}$ which is feasible for (2) and note that for this point the objective function evaluates to $-\sqrt{5}$.

Question We saw in both examples how one can get lower bounds on p^* by taking certain "combinations" of the constraints. This method actually allowed us to get lower bounds that are tight, i.e., equal to the optimal value of the optimisation problem. Is it always the case that this method allows us to produce tight lower bounds?

Duality for general conic programs Let $K \subseteq \mathbb{R}^n$ be a proper cone and consider the conic program

$$\begin{array}{ll} \text{minimise} & \langle c, x\rangle \\ \text{subject to} & \mathcal{A}(x)=b \\ & x \in K \end{array} \tag{4}$$

Here $\mathcal{A}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear map and $b \in \mathbb{R}^m$. Let us call p^* the optimal value of (4). We now describe a way to find a lower bound on p^* in the same way we did for the examples considered above. Assume we can find $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$ that satisfy the following:

$$c=z+\mathcal{A}^*(y) \quad \text{ and } \quad z \in K^* \tag{5}$$

where $\mathcal{A}^*$ denotes the *adjoint* of $\mathcal{A}$ (in the matrix representation of $\mathcal{A}$ in the canonical basis then $\mathcal{A}^*$ is simply the transpose of $\mathcal{A}$). It is now an easy calculation to show that if (y, z) satisfy (5) then we have the following lower bound on p^*: $p^* \geq\langle b, y\rangle$. Indeed if x is any feasible point to (4), then:

$$\begin{aligned} \langle c, x\rangle=\langle z+\mathcal{A}^*(y), x\rangle & \stackrel{(a)}{=}\langle z, x\rangle+\langle y, \mathcal{A}(x)\rangle \\ & \stackrel{(b)}{=}\langle z, x\rangle+\langle y, b\rangle \\ & \stackrel{(c)}{\geq}\langle y, b\rangle \end{aligned} \tag{6}$$

where in (a) we used the definition of adjoint, namely that $\langle\mathcal{A}^*(u), v\rangle=\langle u, \mathcal{A}(v)\rangle$; in (b) we used the fact that $\mathcal{A}(x)=b$; and in (c) we used the fact that $z \in K^*$ and $x \in K$ to conclude that $\langle z, x\rangle \geq 0$.

A natural thing to do is to look at the best lower bound on p^* one can obtain in this way. This amounts to the following maximisation problem:

$$\begin{array}{ll} \underset{y \in \mathbb{R}^m, z \in \mathbb{R}^n}{\operatorname{maximise}} & \langle b, y\rangle \\ \text{subject to} & c=z+\mathcal{A}^*(y) \\ & z \in K^* . \end{array} \tag{7}$$

The optimisation problem (7) is called the *dual* of (4). Observe that (7) is a conic program over K^*. In the two simple examples we considered in the beginning we saw that the optimal value of the dual was equal to the optimal value of our (primal) problem. This phenomenon is known as strong duality. We will state and prove a theorem next time giving us (fairly mild) conditions under which strong duality holds for conic programs.

Exercise 6.1 (Closest positive semidefinite matrix in spectral norm). *Recall the spectral norm of a (symmetric) matrix $M \in \mathbf{S}^n$ is defined as*

$$\|M\| = \max_{\|x\|_2=1} \|Mx\| = \max(|\lambda_1(M)|, \ldots, |\lambda_n(M)|)$$

where $\lambda_1(M), \ldots, \lambda_n(M)$ are the eigenvalues of M. Let $A \in \mathbf{S}^n$ and consider the optimisation problem:

$$\underset{X \in \mathbf{S}^n}{\textit{minimise}} \quad \|A - X\| \quad \textit{s.t.} \quad X \succeq 0. \tag{8}$$

In other words, we want to find the closest positive semidefinite matrix to A in the spectral norm.

1. *Show that the optimisation problem (8) is equivalent to the semidefinite program:*

   $$\underset{X \in \mathbf{S}^n, t \in \mathbb{R}}{\textit{minimise}} \quad t \quad \textit{s.t.} \quad A - X \preceq tI, \quad A - X \succeq -tI, \quad X \succeq 0 \tag{9}$$

   where I denotes the identity matrix (hint: prove that $\|M\| \leq t$ if and only if $-tI \preceq M \preceq tI$).

2. *Let p^* be the optimal value of (9). Use a duality argument (similar to the example (2) we discussed in the lecture) to show that $p^* \geq 0$.*

3. *Let $\lambda_{\min}$ be the smallest eigenvalue of A. Use another duality argument to show that $p^* \geq -\lambda_{\min}$.*

4. *Show that $p^* = \max(0, -\lambda_{\min})$ by exhibiting a feasible point (X, t) of (9) with that value.*