ECE285: Semidefinite and sum-of-squares optimization **Winter 2024**

Lecture 7: Duality in conic programming (I)

Lecturer: Yang Zheng *Scribe: Yang Zheng*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. Any typos should be sent to* `zhengy@eng.ucsd.edu`.

Learning goals:

1. Certificate of bounds
2. Duality for general conic programs
3. Dual of LPs, SOCPs, and SDPs
4. The Lagrange dual problem

1 Motivating examples

Given an optimization problem, how do we find a systematic way to bound its optimal value p^*? A trivial answer is to solve the optimization problem and get the exact optimal value. However, there is much more instructive way to answer such as a question, via duality.

Example 7.1. *Let us first look at the following example:*

$$\begin{aligned} p^* = \min_{x_1,\ldots,x_{10}} \quad & x_1 + x_2 + \ldots + x_{10} \\ \text{subject to} \quad & x_1 + 2x_2 + \ldots + 10x_{10} - 1 \geq 0 \\ & 10x_1 + 9x_2 + \ldots + x_{10} - 10 \geq 0 \\ & \ldots \end{aligned} \tag{1}$$

We claim that the optimal value $p^* \geq 1$. This is actually immediate: adding the first two constraints leads to

$$11x_1 + 11x_2 + \ldots + 11x_{10} - 11 \geq 0 \Rightarrow x_1 + x_2 + \ldots + x_{10} \geq 1.$$

We get a valid lower bound with a certificate by certain combinations of the constraints. LP duality is a straightforward generalization of this simple trick. □

Example 7.2. *Consider the following linear program*

$$\begin{aligned} p^* = \min_{x,y} \quad & 2x + y \\ \text{subject to} \quad & x + y + 1 \geq 0 \\ & x + 1 \geq 0 \\ & y + 1 \geq 0 \\ & -x + 1 \geq 0 \\ & -y + 1 \geq 0 \end{aligned} \tag{2}$$

- *Finding an upper bound on p^* is relatively "simple": given any feasible point (x, y), we know by definition that $p^* \leq 2x + y$. (The challenge is that identifying a feasible point is non-trivial sometimes.) In our simple example, we can verify that $(x, y) = (0, 0)$ is feasible. This tells us that $p^* \leq 0$. If we take another feasible point $(x, y) = (-1, 0)$, we get that $p^* \leq -2$. Can we do better?*

- *Finding a lower bound on p^* seems to be more difficult. A trivial lower bound is $p^* \geq -\infty$, which is not informative (but this is the only lower bound in the case that the original problem is unbounded below). Our strategy here is to take linear combinations of the constraints with nonnegative coefficients. For example, if we multiply the second constraint by 2 and add it to the third constraint, we get*

$$2(x+1) + y + 1 \geq 0 \qquad \Rightarrow \qquad 2x + y \geq -3.$$

Thus $p^* \geq -3$, which is a much better bound than $-\infty$. We can try another combination: if we add the first two constraints together, this leads to

$$(x+y+1) + (x+1) \geq 0 \qquad \Rightarrow \qquad 2x + y \geq -2.$$

This means that $p^* \geq -2$.

We have shown that $p^* \leq 2$ by finding a feasible point with objective value -2. Also, by taking appropriate linear combinations with nonnegative coefficients of the constraints, we have shown that $p^* \geq -2$. Therefore, we have established that $p^* = -2$ for this simple LP (3). □

Example 7.3. *Consider the following simple SDP*

$$\begin{aligned} p^* = \min_{x,y} \quad & x + y \\ \textit{subject to} \quad & \begin{bmatrix} 1-x & y \\ y & 1+x \end{bmatrix} \succeq 0. \end{aligned} \tag{3}$$

Consider again the problem of finding a lower bound on p^*. How shall we generalize the idea in the previous example? Let us consider a positive semidefinite matrix

$$\begin{bmatrix} a & c \\ c & b \end{bmatrix} \succeq 0. \tag{4}$$

The inner product of two positive semidefinite matrix is always nonnegative (Problem 1 in HW1). It then follows that any feasible point (x, y) in (3) *must satisfy*

$$\left\langle \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \begin{bmatrix} 1-x & y \\ y & 1+x \end{bmatrix} \right\rangle \geq 0,$$

which is $(c-a)x + 2by \geq -a - c$. Since the objective function is $x + y$, if we choose $c - a = 1$ and $b = 1/2$, we get a lower bound $p^* \geq -a - c$. Note that such a choice of a, b, c should also satisfy (4). *We can verify that*

$$a = \frac{\sqrt{2}-1}{2}, b = \frac{1}{2}, c = \frac{\sqrt{2}+1}{2}$$

satisfy (4). *Using this choice, we get a lower bound $p^* \geq -a - c = -\sqrt{2}$.*

We can actually verify that p^* is indeed $-\sqrt{2}$: take $(x, y) = (-\sqrt{2}/2, -\sqrt{2}/2)$ which is feasible with objective value $-\sqrt{2}$; this proves an upper bound $p^* \leq -\sqrt{2}$. □

2 Duality for general conic programs

In the previous section, we have seen how we can get lower bounds on p^* by taking certain "combination" of the constraints. We will generalize this idea to general conic programs in this section.

Let $K \in \mathbb{R}^n$ be a proper cone. A conic program over K is an optimization problem of the form:

$$\begin{aligned} p^* = \min_{x} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Ax = b \\ & x \in K, \end{aligned} \tag{5}$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are problem data. The optimization variable here is $x \in \mathbb{R}^n$. We now use a similar way to establish a lower bound on p^*:

- for the equality constraint $Ax = b$, we have
$$y^\mathsf{T}(Ax - b) = 0, \qquad \forall y \in \mathbb{R}^m.$$
- for the conic constraint $x \in K$, we have
$$z^\mathsf{T} x \geq 0, \qquad \forall z \in K^*.$$

Then, we have
$$y^\mathsf{T}(Ax - b) + z^\mathsf{T} x \geq 0 \Rightarrow (A^\mathsf{T} y + z)^\mathsf{T} x \geq b^\mathsf{T} y.$$
If we can find $z \in K^*, y \in \mathbb{R}^m$ such that $c = z + A^\mathsf{T} y$, then we have a lower bound $p^* \geq b^\mathsf{T} y$. Formally, we have
$$\begin{aligned} c^\mathsf{T} x = (z + A^\mathsf{T} y)^\mathsf{T} x &= z^\mathsf{T} x + y^\mathsf{T} A x \\ &= z^\mathsf{T} x + y^\mathsf{T} b \\ &\geq b^\mathsf{T} y. \end{aligned}$$

Dual problem: A natural thing to do is to ask for the best lower bound on p^* from the procedure above. This amounts to solve the following maximization problem
$$\begin{aligned} d^* = \max_{y,z} \quad & b^\mathsf{T} y \\ \text{subject to} \quad & A^\mathsf{T} y + z = c \\ & z \in K^*. \end{aligned} \tag{6}$$

This problem is called the dual of (5). Note that (6) is a conic program over the dual cone K^*.

In the LP's and SDP's dual problem, the optimal value of the dual is the same as the optimal value of the primal problem. This situation is known as *strong duality*. We have the following theorem.

Theorem 7.1 (Duality for conic programs)**.** *Consider the conic program* (5) *and its dual* (6)*. The following statements hold*

1. *Weak duality: $p^* \geq d^*$;*
2. *Strong duality: $p^* = d^*$ if problem* (5) *is strictly feasible (i.e., there exists $x \in int(K)$ such that $Ax = b$).*

3 Dual of LPs, SOCPs, and SDPs

3.1 Dual of LPs

The standard-form LP is a problem of the form
$$\begin{aligned} \min_{x} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned} \tag{7}$$

Since $\mathbb{R}^n_+$ is self-dual, i.e., $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$, it is clear that the dual problem of (7) is also an LP of the form
$$\begin{aligned} \max_{y,z} \quad & b^\mathsf{T} y \\ \text{subject to} \quad & A^\mathsf{T} y + z = c, \\ & z \geq 0. \end{aligned} \tag{8}$$

Theorem 7.2 (Strong duality)**.** *If the primal LP (7) is feasible and has a finite optimal value, then the dual LP (8) is also feasible. There exist optimal feasible solutions $x^* \geq 0, y^*, z^* \geq 0$ and $c^\mathsf{T} x^* = b^\mathsf{T} y^*$.*

This is a consequence of the separation theorems for convex sets. We will prove this result in Lecture 8.

3.2 Dual of SOCPs

A standard second-order cone problem (SOCP) is a problem of the form

$$\begin{aligned} \min_{x} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & \|A_i x + b_i\| \leq c_i^\mathsf{T} x + d_i, i = 1, \ldots, m \end{aligned} \tag{9}$$

where $A_i \in \mathbb{R}^{k_i \times n}, b_i \in \mathbb{R}^{k_i}, c_i \in \mathbb{R}^n, d_i \in \mathbb{R}$. We put (9) into a conic form

$$\begin{aligned} \min_{x} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & (A_i x + b_i, c_i^\mathsf{T} x + d_i) \in \mathcal{L}^{k_i+1}, i = 1, \ldots, m, \end{aligned} \tag{10}$$

where $\mathcal{L}^{k_i+1} = \{(u, v) \in \mathbb{R}^{k_i+1} \mid \|u\| \leq v\}$ denotes the second-order cone. With a similar argument, for each second-order cone constraint, we have

$$u_i^\mathsf{T}(A_i x + b_i) + v_i(c_i^\mathsf{T} x + d_i) \geq 0, \quad \forall (u_i, v_i) \in \left(\mathcal{L}^{k_i+1}\right)^* = \mathcal{L}^{k_i+1},$$

where we have used the fact that the second-order cone is self-dual (Problem 3 in Homework 2). Now adding these inequalities together leads to

$$\sum_{i=1}^m \left(A_i^\mathsf{T} u_i + c_i v_i\right)^\mathsf{T} x \geq -\sum_{i=1}^m \left(u_i^\mathsf{T} b_i + v_i d_i\right)$$

Therefore, if we can find $(u_i, v_i) \in \mathcal{L}^{k_i+1}, i = 1, \ldots, m$ such that

$$\sum_{i=1}^m (A_i^\mathsf{T} u_i + c_i v_i) = c,$$

we have a valid lower bound on (9)

$$c^\mathsf{T} x \geq -\sum_{i=1}^m (u_i^\mathsf{T} b_i + v_i d_i).$$

The dual of SOCP (9) is

$$\begin{aligned} \max_{u_i, v_i} \quad & -\sum_{i=1}^m (u_i^\mathsf{T} b_i + v_i d_i) \\ \text{subject to} \quad & \sum_{i=1}^m (A_i^\mathsf{T} u_i + c_i v_i) = c \\ & \|u_i\| \leq v_i, i = 1, \ldots, m, \end{aligned} \tag{11}$$

which is also an SOCP.

3.3 Dual of SDPs

The standard form of SDPs is

$$\begin{aligned} \min_{X} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \ldots, m \\ & X \in \mathbb{S}_+^n, \end{aligned} \tag{12}$$

where $C \in \mathbb{S}^n$ and $A_i \in \mathbb{S}^n, i = 1, \ldots, m$, $b_i \in \mathbb{R}, i = 1, \ldots, m$ are problem data. This form is in the standard form of (5). Therefore, it is not difficult to see the dual of (12) is

$$\begin{aligned} \max_{y,Z} \quad & b^{\mathsf{T}} y \\ \text{subject to} \quad & \sum_{i=1}^{m} y_i A_i + Z = C \\ & Z \in \mathbb{S}^n_+, \end{aligned} \tag{13}$$

where we have used the fact that the positive semidefinite cone is self-dual.

Example 7.4 (Strong duality may not hold for SDPs [2, Example 2.14])**.** *Let $\alpha \geq 0$, and consider the primal-dual pair of SDPs*

$$\begin{aligned} \min_{X} \quad & \alpha x_{11} \\ \textit{subject to} \quad & x_{22} = 0, \\ & x_{11} + 2x_{23} = 1, \\ & \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ * & x_{22} & x_{23} \\ * & * & x_{33} \end{bmatrix} \succeq 0, \end{aligned}$$

and

$$\begin{aligned} \max_{y} \quad & y_2 \\ \textit{subject to} \quad & \begin{bmatrix} y_2 & 0 & 0 \\ 0 & y_1 & y_2 \\ 0 & y_2 & 0 \end{bmatrix} \preceq \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

For a primal feasible point, X being positive semidefinite and $x_{22} = 0$ imply $x_{23} = 0$, and thus $x_{11} = 1$. The primal optimal cost p^* is then equal to α (and is achieved). On the dual side, y_2 must be zero, and thus $d^* = 0$. There is a nonzero duality gap $p^* - d^* = \alpha$. □

4 Lagrange dual formulation

We have seen a simple "trick" to derive a useful lower bound for the conic programming. In fact, there is a more systematic framework to derive the dual formulation of any (possibly non-convex) optimization problem (with nonlinear cost function). We give a brief introduction here; see [3, Chapter 4] for more details.

Let us consider a possibly non-convex optimization problem

$$\begin{aligned} p^* = \min \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, i = 1, \ldots, m \\ & h_j(x) = 0, j = 1, \ldots, p, \end{aligned} \tag{14}$$

with variable $x \in \mathbb{R}^n$. We denote the domain of the problem ($f_i, i = 0, 1, \ldots, m$ and $h_j, j = 1, \ldots, p$) as $\mathcal{D}$, and the feasible region as $\mathcal{X}$. We refer to the problem above as the primal problem and the decision variable x as the primal variable. One purpose of Lagrange duality is to find a lower bound on the minimization problem (14).

The Lagrangian The basic idea is to move the constraints (14) into its cost function. We define the Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ as

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \nu_i h_i(x).$$

The variables $\lambda \in \mathbb{R}^m$ are called Lagrange variables associated with the inequality constraints $f_i(x) \leq 0, i = 1, \ldots, m$, and we refer to $\nu \in \mathbb{R}^p$ as the Lagrange variable associated with the equality constraint $h_j(x) = 0, j = 1, \ldots, p$.

One simple but important observation is that for any feasible $x \in \mathcal{X}$, and any $\lambda \in \mathbb{R}^m_+$, $\nu \in \mathbb{R}^p$, the cost value $f_0(x)$ is bounded below by $L(x, \lambda, \nu)$,

$$f_0(x) \geq f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{j=1}^p \nu_i h_i(x) = L(x, \lambda, \nu), \quad \forall x \in \mathcal{X}, \lambda \in \mathbb{R}^m_+, \nu \in \mathbb{R}^p. \tag{15}$$

The Lagrangian can be used to write the primal problem (14) as an unconstrained form, as

$$p^* = \min_x \max_{\lambda \in \mathbb{R}^m_+, \nu \in \mathbb{R}^p} \ L(x, \lambda, \nu),$$

where we have used the fact that

$$\max_{\lambda \in \mathbb{R}^m_+} \lambda^\mathsf{T} f = \begin{cases} 0 & \text{if } f \leq 0 \\ +\infty & \text{otherwise,} \end{cases} \qquad \max_{\nu \in \mathbb{R}^p} \nu^\mathsf{T} h = \begin{cases} 0 & \text{if } h = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Lagrange dual function. We now define the Lagrange dual function as

$$g(\lambda, \nu) := \min_{x \in \mathcal{D}} \ L(x, \lambda, \nu).$$

Note that for some λ, ν, the Lagrangian may be unbounded below in x, then the dual function $g(\lambda, \nu)$ takes on the value $-\infty$. Since the dual function is pointwise minimum of affine functions in λ, ν, it is always concave. Considering the fact in (15), we obtain

$$f_0(x) \geq \min_{x \in \mathcal{D}} \ L(x, \lambda, \nu) = g(\lambda, \nu), \qquad \forall x \in \mathcal{X}, \lambda \in \mathbb{R}^m_+, \nu \in \mathbb{R}^p.$$

Therefore, the dual function always provide a lower bound on p^*, i.e.,

$$p^* \geq g(\lambda, \nu), \quad \forall \lambda \in \mathbb{R}^m_+, \nu \in \mathbb{R}^p.$$

The Lagrange dual problem The best lower bound that we can obtain is through the optimization problem

$$\begin{aligned} d^* = \max_{\lambda, \nu} \quad & g(\lambda, \nu) \\ \text{subject to} \quad & \lambda \geq 0. \end{aligned} \tag{16}$$

We refer to the problem (16) as the dual problem, and the vector $\lambda \in \mathbb{R}^m, \nu \in \mathbb{R}^p$ as the dual variables. The dual problem is always convex.

Theorem 7.3 (Weak duality)**.** *For the (possibly non-convex) problem* (14)*, weak duality holds:* $p^* \geq d^*$*.*

We can also derive the dual LP (8) by defining a Lagrange function

$$L(x, \lambda, \nu) = c^\mathsf{T} x + \nu^\mathsf{T}(b - Ax) + \lambda^\mathsf{T}(-x).$$

Then the Lagrange dual function is

$$\begin{aligned} g(\lambda, \nu) &= \min_x \ L(x, \lambda, \nu) \\ &= \min_x \ (c - A^\mathsf{T}\nu - \lambda)^\mathsf{T} x + b^\mathsf{T}\nu \\ &= \begin{cases} b^\mathsf{T}\nu & \text{if } c - A^\mathsf{T}\nu - \lambda = 0 \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

Thus, the Lagrange dual problem is

$$\begin{aligned} \max_{\lambda,\nu} \quad & b^\top \nu \\ \text{subject to} \quad & A^\top \nu + \lambda = c \\ & \lambda \geq 0, \end{aligned}$$

which is the same as (8).

4.1 Dual of QPs

A convex quadratic program is an optimization problem of the form

$$\begin{aligned} \min_{x} \quad & \frac{1}{2} x^\top Q x + q^\top x + c \\ \text{subject to} \quad & Ax \leq b \end{aligned} \tag{17}$$

where $Q \in \mathbb{S}^n_+, q \in \mathbb{R}^n, c \in \mathbb{R}, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Here, we consider $Q \in \mathbb{S}^n_{++}$. The Lagrangian is

$$L(x, \lambda) = \frac{1}{2} x^\top Q x + q^\top x + c + \lambda^\top (Ax - b).$$

Since $L(x, \lambda)$ is a convex quadratic function in x, we can find the minimizing x from the optimality condition

$$\nabla_x L(x, \lambda) = Qx + q + A^\top \lambda$$

which yields $x = -Q^{-1}(q + A^\top \lambda)$. Therefore, the dual function is

$$\begin{aligned} g(\lambda) &= \frac{1}{2}(q + A^\top \lambda)^\top Q^{-1}(q + A^\top \lambda) - q^\top Q^{-1}(q + A^\top \lambda) + c + \lambda^\top(-AQ^{-1}(q + A^\top \lambda) - b) \\ &= -\frac{1}{2}(q^\top Q^{-1} q + \lambda^\top A Q^{-1} A^\top \lambda) - (AQ^{-1} q + b)^\top \lambda + c. \end{aligned}$$

The Lagrange dual problem of (17) is

$$\begin{aligned} \max_{\lambda} \quad & -\frac{1}{2} \lambda^\top A Q^{-1} A^\top \lambda - (AQ^{-1} q + b)^\top \lambda + \frac{1}{2} q^\top Q^{-1} q + c \\ \text{subject to} \quad & \lambda \geq 0. \end{aligned}$$

4.2 Dual of QCQPs

A quadratically constrained quadratic programming (QCQP) corresponds to a problem of the form

$$\begin{aligned} \min_{x} \quad & x^\top Q_0 x + q_0^\top x + c_0 \\ \text{subject to} \quad & x^\top Q_i x + q_i^\top x + c_i \leq 0, i = 1, \ldots, m \end{aligned} \tag{18}$$

where $Q_0, Q_i \in \mathbb{S}^n, q_0, q_i \in \mathbb{R}^n, c_0, c_i \in \mathbb{R}$.

The Lagrange function is

$$\begin{aligned} L(x, \lambda) &= x^\top Q_0 x + q_0^\top x + c_0 + \sum_{i=1}^m \lambda_i (x^\top Q_i x + q_i^\top x + c_i) \\ &= x^\top \left(Q_0 + \sum_{i=1}^m \lambda_i Q_i \right) x + \left(q_0 + \sum_{i=1}^m \lambda_i q_i \right)^\top x + c_0 + \sum_{i=1}^m \lambda_i c_i. \end{aligned}$$

The dual function is

$$g(\lambda) = \min_x L(x, \lambda).$$

If

$$L(x, \lambda) - \xi \geq 0, \forall x \in \mathbb{R}^n, \tag{19}$$

then $g(\lambda) \geq \xi$. The condition (19) is equivalent to

$$\begin{bmatrix} c_0 + \sum_{i=1}^m \lambda_i c_i - \xi & \frac{1}{2}(q_0 + \sum_{i=1}^m \lambda_i q_i)^\mathsf{T} \\ \frac{1}{2}(q_0 + \sum_{i=1}^m \lambda_i q_i) & Q_0 + \sum_{i=1}^m \lambda_i Q_i \end{bmatrix} \succeq 0.$$

Therefore, we get a dual problem for (18) as follows

$$\begin{aligned} \max_{\xi, \lambda} \quad & \xi \\ \text{subject to} \quad & \begin{bmatrix} c_0 + \sum_{i=1}^m \lambda_i c_i - \xi & \frac{1}{2}(q_0 + \sum_{i=1}^m \lambda_i q_i)^\mathsf{T} \\ \frac{1}{2}(q_0 + \sum_{i=1}^m \lambda_i q_i) & Q_0 + \sum_{i=1}^m \lambda_i Q_i \end{bmatrix} \succeq 0, \\ & \lambda \geq 0, \end{aligned} \tag{20}$$

which is a semidefinite program.

Notes

The preparation of this lecture was based on [4, Lectures 5 & 6]. Further reading for this lecture can refer to [1, Chapter 1] and [3, Chapter 5].

References

[1] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[2] Grigoriy Blekherman, Pablo A Parrilo, and Rekha R Thomas. *Semidefinite optimization and convex algebraic geometry*. SIAM, 2012.

[3] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[4] Hamza Fawzi. Topics in Convex Optimisation, Michaelmas 2018.