In this lecture we continue our discussion of LP duality.

1 Linear Programming Duality Recap

Consider the "primal" maximization LP. Solve for $\mathbf{x}$ to:

$$\begin{aligned} \text{maximize } & \mathbf{c}^T\mathbf{x} \\ \text{subject to } & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned} \tag{1}$$

The constraint $\mathbf{x} \geq \mathbf{0}$ is just short-hand for saying that the $\mathbf{x}$ variables are constrained to be non-negative.[1] Recall that the dual of this is: Solve for $\mathbf{y}$ to:

$$\begin{aligned} \text{minimize } & \mathbf{y}^T\mathbf{b} \\ \text{subject to } & \mathbf{y}^T A \geq \mathbf{c}^T \\ & \mathbf{y} \geq \mathbf{0}, \end{aligned} \tag{2}$$

We defined the dual in order to get the best upper bound we could on the primal. And if you take the dual of (2) to try to get the best lower bound on this LP, you'll get (1). *The dual of the dual is the primal.* The dual and the primal are best upper/lower bounds you can obtain as linear combinations of the inputs.

The natural question is: maybe we can obtain better bounds if we combine the inequalities in more complicated ways, not just using linear combinations. Or do we obtain optimal bounds using just linear combinations? In fact, we get optimal bounds using just linear combinations, as the next theorems show.

1.1 The Theorems

It is easy to show that the dual (2) provides an upper bound on the value of the primal (1):

Theorem 1 (Weak Duality) *If $\mathbf{x}$ is a feasible solution to the primal LP (1) and $\mathbf{y}$ is a feasible solution to the dual LP (2) then*

$$\mathbf{c}^T\mathbf{x} \leq \mathbf{y}^T\mathbf{b}.$$

Proof: This is just a sequence of trivial inequalities that follow from the LPs above:

$$\mathbf{c}^T\mathbf{x} \leq (y^T A)\mathbf{x} = y^T(A\mathbf{x}) \leq y^T b.$$

■

The amazing (and deep) result here is to show that the dual actually gives a perfect upper bound on the primal (assuming some mild conditions).

[1] We use the convention that vectors like $\mathbf{c}$ and $\mathbf{x}$ are column vectors. So $\mathbf{c}^T$ is a row vector, and thus $\mathbf{c}^T\mathbf{x}$ is the same as the inner product $\mathbf{c} \cdot \mathbf{x} = \sum_i c_i x_i$. We often use $\mathbf{c}^T\mathbf{x}$ and $\mathbf{c} \cdot \mathbf{x}$ interchangeably. Also, $\mathbf{a} \leq \mathbf{b}$ means component-wise inequality, i.e., $a_i \leq b_i$ for all i.

Theorem 2 (Strong Duality Theorem) *Suppose the primal LP (1) is feasible (i.e., it has at least one solution) and bounded (i.e., the optimal value is not ∞). Then the dual LP (2) is also feasible and bounded. Moreover, if $\mathbf{x}^*$ is the optimal primal solution, and $\mathbf{y}^*$ is the optimal dual solution, then*

$$\mathbf{c}^T\mathbf{x}^* = (\mathbf{y}^*)^T\mathbf{b}.$$

In other words, the maximum of the primal equals the minimum of the dual.

2 Example #1: Zero-Sum Games

Consider a 2-player zero-sum game defined by an n-by-m payoff matrix R for the row player. That is, if the row player plays row i and the column player plays column j then the row player gets payoff R_{ij} and the column player gets $-R_{ij}$. To make this easier on ourselves (it will allow us to simplify things a bit), let's assume that all entries in R are positive (this is really without loss of generality since as pre-processing one can always translate values by a constant and this will just change the game's value to the row player by that constant). We saw we could write this as an LP:

- Variables: $v, p_1, p_2, \ldots, p_n$.
- Maximize v,
- Subject to:

  $p_i \geq 0$ for all rows i,

  $\sum_i p_i = 1$,

  $\sum_i p_i R_{ij} \geq v$, for all columns j.

To put this into the form of (1), we can replace $\sum_j p_j = 1$ with $\sum_i p_i \leq 1$ since we said that all entries in R are positive, so the maximum will occur with $\sum_i p_i = 1$, and we can also safely add in the constraint $v \geq 0$. We can also rewrite the third set of constraints as $v - \sum_i p_i R_{ij} \leq 0$. This then gives us an LP in the form of (1) with

$$\mathbf{x} = \begin{array}{|c|}\hline v \\ p_1 \\ p_2 \\ \ldots \\ p_n \\ \hline\end{array}, \mathbf{c} = \begin{array}{|c|}\hline 1 \\ 0 \\ 0 \\ \ldots \\ 0 \\ \hline\end{array}, \mathbf{b} = \begin{array}{|c|}\hline 0 \\ 0 \\ \ldots \\ 0 \\ 1 \\ \hline\end{array}, \text{and } A = \begin{array}{|c|ccc|}\hline 1 & & & \\ 1 & & -R^T & \\ \ldots & & & \\ 1 & & & \\ \hline 0 & 1 & \ldots & 1 \\ \hline\end{array}.$$

I.e., maximizing $\mathbf{c}^T\mathbf{x}$ subject to $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$.

We can now write the dual, following (2). Let $\mathbf{y}^T = (y_1, y_2, \ldots, y_{m+1})$. We now are asking to minimize $\mathbf{y}^T\mathbf{b}$ subject to $\mathbf{y}^T A \geq c^T$ and $\mathbf{y} \geq \mathbf{0}$. In other words, we want to:

- Minimize y_{m+1},
- Subject to:

  $y_1 + \ldots + y_m \geq 1$,

  $-y_1 R_{i1} - y_2 R_{i2} - \ldots - y_m R_{im} + y_{m+1} \geq 0$ for all rows i, or equivalently,

$y_1 R_{i1} + y_2 R_{i2} + \ldots + y_m R_{im} \leq y_{m+1}$ for all rows i.

So, we can interpret y_{m+1} as the value to the row player, and $y_1, \ldots, y_m$ as the randomized strategy of the column player, and we want to find a randomized strategy for the column player that minimizes y_{m+1} subject to the constraint that the row player gets *at most* y_{m+1} no matter what row he plays. Now notice that we've only required $y_1 + \ldots + y_m \geq 1$, but since we're minimizing and the R_{ij}'s are positive, the minimum will happen at equality.

Notice that the fact that the maximum value of v in the primal is equal to the minimum value of y_{m+1} in the dual follows from strong duality. Therefore, the minimax theorem is a corollary to the strong duality theorem.

3 Example #1: Shortest Paths

Duality allows us to write problems in multiple ways, which gives us power and flexibility. For instance, let us see two ways of writing the shortest s-t path problem, and why they are equal.

Here is an LP for computing an s-t shortest path with respect to the edge lengths $\ell(u, v) \geq 0$:

$$\begin{aligned} \max \quad & d_t \\ \text{subject to} \quad & d_s = 0 \\ & d_v - d_u \leq \ell(u, v) \qquad \forall (u, v) \in E \end{aligned} \tag{3}$$

The constaints are the natural ones: the shortest distance from s to s is zero. And if the s-u distance is d_u, the s-v distance is at most $d_u + \ell(u, v)$ — i.e., $d_v \leq d_u + \ell(u, v)$. It's like putting strings of length $\ell(u, v)$ between u, v and then trying to send t as far from s as possible—the farthest you can send t from s is when the shortest s-t path becomes tight.

Here is another LP that also computes the s-t shortest path:

$$\begin{aligned} & \min \textstyle\sum_e \ell(e)\, y_e \\ \text{subject to} \quad & \textstyle\sum_{w:(s,w) \in E} y_{sw} = 1 \\ & \textstyle\sum_{v:(v,t) \in E} y_{vt} = 1 \\ & \textstyle\sum_{v:(u,v) \in E} y_{uv} = \sum_{v:(v,w) \in E} y_{vw} \qquad \forall w \in V \setminus \{s, t\} \\ & y_e \geq 0. \end{aligned} \tag{4}$$

In this one we're sending one unit of flow from s to t, where the cost of sending a unit of flow on an edge equals its length ℓ_e. Naturally the cheapest way to send this flow is along a shortest s-t path length. So both the LPs should compute the same value. Let's see how this follows from duality.

3.1 Duals of Each Other

Take the first LP. Since we're setting d_s to zero, we could hard-wire this fact into the LP. So we could rewrite (3) as

$$\begin{aligned} & \max \quad d_t \\ \text{subject to} \quad & d_v - d_u \le \ell(u,v) \qquad \forall (u,v) \in E, s \notin \{u,v\} \\ & d_v \le \ell(s,v) \qquad \forall (s,v) \in E \\ & -d_u \le \ell(u,s) \qquad \forall (u,s) \in E \end{aligned} \tag{5}$$

Moreover, the distances are never negative for $\ell(u,v) \ge 0$, so we can add in the constraint $d_v \ge 0$ for all $v \in V$.

How to find an upper bound on the value of this LP? The LP is in the standard form, so we can do this mechanically. But let us do this from starting from the definition of the dual as the "best upper bound".

Let us define $E_s^{out} := \{(s,v) \in E\}$, $E_s^{in} := \{(u,s) \in E\}$, and $E^{rest} := E \setminus (E_s^{out} \cup \mathrm{E}_s^{in})$. For every arc $e = (u,v)$ we will have a variable $y_e \ge 0$. We want to get the best upper bound on d_t by linear combinations of the the constraints, so we should find a solution to

$$\sum_{e \in E^{rest}} y_{uv}\,(d_v - d_u) + \sum_{e \in E_s^{out}} y_{sv}\,d_v - \sum_{e \in E_s^{in}} y_{us}\,d_u \ge d_t \tag{6}$$

(this is like $\mathbf{y}^T A \ge c$) and the objective function is to

$$\text{minimize} \sum_{(u,v) \in E} y_{uv}\,\ell(u,v). \tag{7}$$

(This is like $\min \mathbf{y}^T \mathbf{b}$.) Great, the objective function (7) is exactly what we want, but what about the craziness in (6)? Just collect all copies of each of the variables d_v, and it now says

$$\sum_{v \neq s} d_v \left(\sum_{u:(u,v) \in E} y_{uv} - \sum_{w:(v,w) \in E} y_{vw} \right) \ge d_t.$$

First, this must be an equality at optimality (since otherwise we could reduce the y values). Moreover, these equalities must hold regardless of the d_v values, so this is really the same as

$$\begin{aligned} \sum_{u:(u,v) \in E} y_{uv} - \sum_{w:(v,w) \in E} y_{vw} &= 0 \qquad \forall v \notin \{s,t\}. \\ \sum_{u:(u,t) \in E} y_{ut} - \sum_{w:(t,w) \in E} y_{tw} &= 1. \end{aligned} \tag{8}$$

Summing all these inequalities for all nodes $v \in V \setminus \{s\}$ gives us the missing equality:

$$\sum_{w:(s,w) \in E} y_{sw} - \sum_{u:(u,s) \in E} y_{us} = 1.$$

Finally, observe that since there's flow conservation at all nodes, and the net unit flow leaving s and reaching t, this means we must have a possibly-empty circulation (i.e., flow going around in circles) plus one unit of s-t flow. Removing the circulation can only lower the objective function, so at optimality we're left with one unit of flow from s to t. This is precisely the LP (4), showing that the dual of LP (3) is LP (4), after a small amount of algebra.