ORIE 6300 Mathematical Programming I August 25, 2016

Lecture 2

Lecturer: Damek Davis *Scribe: Johan Bjorck*

Last time, we considered the dual of linear programs in our "basic" form: $\max(c^T x : Ax \le b)$. We also considered an algebraic method for testing optimality of solution: To help find the optimal solution, we generate upper bounds on the value of feasible solutions by taking linear combinations of rows of constraints equal to c. That is, we find y such that $A^T y = c$, with $y \ge 0$. This gives an upper bound $y^T b$ on $c^T x$. Then, the best upper bound is given by the following LP: $\min(y^T b : A^T y = c,\ y \ge 0)$. We call this the *dual LP*.

We also mentioned two theorems regarding the dual LP: The weak duality theorem, which states that the value of the primal LP is at most the value of the dual, and the strong duality theorem, which states that if there exists a feasible solution to either the primal or the dual LP, then the two LPs have the same value.

Today we will consider linear programs in other standard forms. We will derive rules for how to take the duals of these other linear programs. See the Chapter 5 of Chvátal for a good summary of these rules.

We will consider the "standard form" of linear programs, which has as a precise, technical definition . Most algorithms used for solving linear programs assume that the input is given in the standard form, and the formulation is not restrictive in any representational sense. The standard form is defined as: $\min(\bar{c}^T x : x \ge 0, \bar{A}x = \bar{b})$. This formulation is the dual of our "basic" form, and we have introduced the bars on each of A, b, and c to distinguish this from the dual that we derived last class: $\min(y^T b : y \ge 0, A^T y = c)$; Our new primal is equivalent to that dual in the following way:

$$\boxed{\text{Translation: } \bar{b} = c,\ \bar{c} = b,\ \bar{A} = A^T}$$

We saw last time in the bakery example that duality was helpful for finding solutions to linear programs, and there are two ways to think about taking the dual of the standard form of the LP.

1. By reducing to the form considered in last lecture: $\max(c^T x : Ax \le b)$

2. By finding best bound on primal

First, we will rewrite the linear program $\min(\bar{c}^T x : x \ge 0, \bar{A}x = \bar{b})$ in our "basic" form (namely, as a maximization problem with $\le$ constraints), and take its dual using the rules we derived last time. This leads to a messy dual, and we will work on simplifying what we obtain. To understand better what the dual is, we will also derive the dual using the idea that the dual is a way to derive sharp bounds on the primal.

We start by converting our LP into our "basic" maximization form that uses inequalities only. First we need to change from min to max, so instead of minimizing $\bar{c}^T x$, we can consider maximizing $-\bar{c}^T x$. This only switches the sign of the result (but not the absolute value, nor the optimal solution). We will switch the sign back later. Furthermore, our "basic" form consisted only inequalities, so we consider:

$$\max -\bar{c}^T x$$

$$\begin{aligned} \bar{A}x &\leq \bar{b} \\ -\bar{A}x &\leq -\bar{b} \\ -Ix &\leq 0, \end{aligned}$$

where I is the identity matrix. To get the dual, we introduce one dual variable for each row (i.e. for each inequality constraint) of the primal, and so we have one dual variable corresponding to each row of the resulting matrix above. We have three sets of rows, so we will use three different names for the three sets of variables: $y = (s, t, w)$ where the variables s, t and w correspond to the blocks of rows in $\bar{A}$, $-\bar{A}$ and $-I$, respectively. So, using the dual formulation from last lecture, we have the following dual LP

$$\min y^T \begin{bmatrix} \bar{b} \\ -\bar{b} \\ 0 \end{bmatrix}$$

$$\begin{aligned} \begin{bmatrix} \bar{A} \\ -\bar{A} \\ -I \end{bmatrix}^T y &= -\bar{c} \\ y &\geq 0. \end{aligned}$$

Since $y = (s, t, w)$, we obtain

$$\min \begin{bmatrix} s \\ t \\ w \end{bmatrix}^T \begin{bmatrix} \bar{b} \\ -\bar{b} \\ 0 \end{bmatrix}$$

$$\begin{aligned} \begin{bmatrix} \bar{A} \\ -\bar{A} \\ -I \end{bmatrix}^T \begin{bmatrix} s \\ t \\ w \end{bmatrix} &= -\bar{c} \\ s,\ t,\ w &\geq 0, \end{aligned}$$

which can be rewritten as

$$\min s^T\bar{b} - t^T\bar{b}$$

$$\begin{aligned} \bar{A}^T s - \bar{A}^T t - Iw &= -\bar{c} \\ s,\ t,\ w &\geq 0. \end{aligned}$$

We will simplify this. First recall that when creating the "basic" form equivalent, we switched the sign of the objective, so let's switch this back now: instead of minimizing $s^T\bar{b} - t^T\bar{b}$ we will maximize $-s^T\bar{b} + t^T\bar{b}$. Further notice that s and t never occur by themselves any place, only $t - s$, so we can use a new variable $z = t - s$. Both s and t had to be nonnegative, but $t - s$ can be of any sign and any value. The objective function then reads as $\max z^T\bar{b}$, and the constraints will be $-\bar{A}^T z - Iw = -\bar{c}$. Note that this change makes an equivalent LP in the sense that given any feasible solution using t and s, we get a feasible solution of the same value by setting $z = t - s$, and similarly if we have a feasible solution for the LP using the variable z, we can get a feasible solution to the previous LP of the same value by setting $t_i = z_i$ and $s_i = 0$ if $z_i \geq 0$, and $s_i = |z_i|$ and $t_i = 0$ if $z_i < 0$.

The LP looks better if we multiply the constraints by -1. What we have so far is

$$\max z^T \bar{b}$$

$$\begin{aligned} w &\geq 0 \\ \bar{A}^T z + Iw &= \bar{c}. \end{aligned}$$

We will further simplify what we get by eliminating the variables w. Our equations can be rewritten to say that $Iw = \bar{c} - \bar{A}^T z$. The only constraint we have about w is that it is nonnegative, and there is no w in the objective function, so we can write the constraint as $\bar{c} - \bar{A}^T z \geq 0$, and not have the variables w at all. The final form that we obtained for the dual is

$$\max z^T \bar{b}$$

$$\bar{A}^T z \leq \bar{c}.$$

Recall that we use $\bar{A}$, $\bar{b}$, and $\bar{c}$ to denote A^T, c, and b, respectively. Substituting these back into the the dual LP above, we note that we obtain an LP that is exactly the same as the primal LP that we started from last time:

$$\max z^T c$$

$$Az \leq b.$$

We have proved the following theorem.

Theorem 1 *The dual of the dual LP is the same as the original primal LP.*

Next we will work on understanding the dual, by creating it as a bound for the primal. Here is the bounding idea via an example:

$$\begin{aligned} \min x_1 + x_2 & \\ 2x_2 + x_3 &= 7 \\ x_1 - x_3 &= 2 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

If we add these two equations, we get that $x_1 + 2x_2 = 9$ for all feasible solutions (and that this is the only way to eliminate x_3 from the result). The objective function is to minimize $x_1 + x_2$, and not $x_1 + 2x_2$. An equivalent form of what we have is that $(1/2)x_1 + x_2 = 9/2$ for any primal feasible x vector. From this form, we see that $x_1 + x_2$ must be at least $9/2$ since x_1 is non-negative.

Thus we get a lower bound on the LP by taking linear combinations of the equality constraints

$$\begin{aligned} & y_1 \times (\bar{a}_1 x = \bar{b}_1) \\ & y_2 \times (\bar{a}_2 x = \bar{b}_2) \\ & \ldots \\ + \quad & y_n \times (\bar{a}_n x = \bar{b}_n) \end{aligned}$$

such that the number of copies of each x_j is at most c_j. This implies that:

$$y_1 \bar{a}_{1j} + y_2 \bar{a}_{2j} + \ldots + y_n \bar{a}_{nj} \leq c_j.$$

The bound we obtain is $y^T \bar{b}$. Thus for any feasible x, we have that $\bar{c}^T x \geq y^T \bar{b}$ if $\bar{A}^T y \leq \bar{c}$. Hence, the best bound of this form is the following linear program:

$$\max y^T \bar{b}$$

$$\bar{A}^T y \leq \bar{c}.$$

This is once again the dual of the original primal LP.

So far, we saw how to derive the dual for LPs in two forms: $\max(c^T x : Ax \leq b)$, and $\min(c^T x : x \geq 0, Ax = b)$. The idea of understanding the dual as a bound for the primal leads to very simple rules for creating dual linear programs without having to convert to any normal form. Here is a set of rules for maximization type problems. For simplicity, we will only have $=$ and $\leq$ type constraints, i.e., I suggest that you take all $\geq$ type constraints for a maximization problem, and multiply them by -1 to get $\leq$ type constraints before taking the dual. This form has some nonnegative variables, and some unconstrained ones, some equations and some inequalities as constraints.

$$\begin{aligned} \max c^T x + \bar{c}^T \bar{x} & & \\ Ax + \bar{A}\bar{x} &\leq b & \\ \hat{A}x + \tilde{A}\bar{x} &= \bar{b} & \qquad \text{(primal LP)} \\ x &\geq 0 & \end{aligned}$$

The dual has nonnegative variables y corresponding to the inequality constraints, variables $\bar{y}$ unconstrained in sign corresponding to the equations, and the nonnegative variables x give rise to inequalities in the dual, whereas the unconstrained variables $\bar{x}$ give rise to equations in the dual. The dual LP is:

$$\begin{aligned} \min y^T b + \bar{y}^T \bar{b} & & \\ A^T y + \hat{A}^T \bar{y} &\geq c & \\ \bar{A}^T y + \tilde{A}^T \bar{y} &= \bar{c} & \qquad \text{(dual LP)} \\ y &\geq 0 & \end{aligned}$$

In summary,

Primal		Dual
Max	$\Leftrightarrow$	Min
$\leq$ constraint	$\Leftrightarrow$	variable ≥ 0
$=$ constraint	$\Leftrightarrow$	variable unconstrained
variable ≥ 0	$\Leftrightarrow$	$\geq$ constraint
variable unconstrained	$\Leftrightarrow$	$=$ constraint

Recall that the dual of the dual is the primal. So for an LP in the minimization form we can use the above forms to take the dual.

As an aside it can be mentioned that solving linear inequalitites is essentially as difficult as solving linear programs. Earlier discussions also suggests an algorithm for solving $\max(c^T x : Ax \leq b)$ knowing only that $\max(c^T x : Ax \leq b) \leq B_U^0$. Assuming that we have a lower bound $B_L^0 = c^T x^0$ for some x^0 in the feasable region, we can then follow the scheme below

for $k = 0, 1, 2...K$ **do**
  if $\exists x^* \ s.t. \ c^T x^* \geq (B_L^k + B_U^k)/2 \ and \ Ax^* \leq b$ **then**
    $B_L^{k+1} = c^T x^*$;
    $B_U^{k+1} = B_U^k$;
    $x^{k+1} = x^*$;
  else
    $B_U^{k+1} = (B_L^k + B_U^k)/2$;
    $B_K^{k+1} = B_L^k$;
    $x^{k+1} = x^k$;
  end
end
return (x^K)

As an exercise the reader is welcome to prove that $0 \leq c^T x^* - c^T x^K \leq (B_U^0 - B_L^0)/2^K$.