ORIE 6300 Mathematical Programming I	September 23, 2014

Lecture 9

Lecturer: David P. Williamson — *Scribe: Yi Xu*

Consider a primal and dual LP in the generic form in which we have been studying LPs so far in the course, in the case when both are feasible. We know the optimal values of the LPs are equal, but is there a good procedure to tell whether a given x is optimal?

1 Verifying optimality

Let's look at the following LP primal and dual pair:

$$\begin{array}{llcll} \min & c^T x & \qquad & \max & y^T b \\ \text{s.t.} & Ax = b & & \text{s.t.} & A^T y = c \\ & x \geq 0 & & & \end{array} \tag{1}$$

Answer 1 *We know x is optimal if there exists a dual feasible y such that $c^T x = b^T y$, by strong duality.*

This is true as far as it goes, but it doesn't seem that useful. Let's think about other ways in which we can show the optimality of x.

Let x and y be feasible for the primal and dual, respectively. Recall our proof of strong duality:

$$c^T x = \sum_{j=1}^{n} c_j x_j \geq \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} y_i \right) x_j = \sum_{i=1}^{m} y_i \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{m} y_i b_i = b^T y.$$

where the inequality follows from $A^T y \leq c$. From the strong duality theorem, we know if x and y are optimal, then $c^T x = b^T y$. For this to be true, in the inequalities above, we need that if $\sum_{i=1}^{m} a_{ij} y_i < c_j$ then $x_j > 0$. Call these conditions $(*)$.

Definition 1 *We say that a primal feasible solution x, and a dual feasible solution y obey the* complementary slackness conditions *if $(*)$ holds.*

So we see from the above that if x and y are optimal solutions, then complementary slackness holds. But actually we can say something stronger than this.

Lemma 1 *Given a primal feasible solution x, and a dual feasible solution y, x and y are optimal if and only if the complementary slackness conditions hold.*

Hence we have another answer to our question.

Answer 2 *x is optimal if there exists a dual feasible y such that the complementary slackness conditions hold.*

This still doesn't seem like such a useful way of verifying optimality, but it will prove to be a step in the right direction.

So far we haven't been taking advantage of something that we know about optimal solutions. We've said that there exists an optimal solution that is a vertex, and have shown this on a problem set for bounded polyhedra, and in a recitation for pointed polyhedra. We've also shown in a problem set that if x is not a vertex, we can find a vertex $\tilde{x}$ such that $c^T\tilde{x} \leq c^T x$. So we can assume that x is a vertex.

Recall that x is a vertex if and only if $rank(A_=) = n$. Note that $a_j x = b_j$ for $j = 1, \ldots, m$. The remaining $n - m$ inequalities met with equality (modulo linear dependence) must be of the form $x_i = 0$. Assume that the variables are numbered such that $x_1, ..., x_k > 0$ and $x_{k+1}, ..., x_n = 0$. Then

$$\left[\begin{array}{c|c} \multicolumn{2}{c}{A} \\ \hline 0 & I \end{array}\right] x = \left[\begin{array}{c} b \\ \hline 0 \end{array}\right].$$

This matrix, $A^=$, has rank n, so all its columns are linearly independent. So the columns of A corresponding to positive x_i variables are linearly independent. This gives us the following lemma.

Lemma 2 *A feasible solution x is a vertex iff the columns corresponding to its positive coordinates are linearly independent.*

This gives us an easy way to check if a feasible solution is a vertex or not. It's worth encoding this into a definition. First, we need an assumption though. We assume without loss of generality that the m rows of A are linearly independent. It's without loss of generality since otherwise a constraint is redundant (if a constraint can be expressed as a linear combination of other constraints) or the system $Ax = b$ is infeasible (if the right-hand side of the

Definition 2 *A set B of m columns of A is a* basis *if these columns are linearly independent.*

We will focus on a subset of columns of A which correspond to a basis B.

$$\begin{array}{ccc} A\text{: } m \text{ lin. ind. rows} & \left[\begin{array}{ccc} A & \left(A_i\right) & \quad \end{array}\right] & \rightarrow \quad A_B \\ & \uparrow \;\; \uparrow \;\; \uparrow & \\ & m \text{ columns } B & \end{array}$$

We will denote by x_B the coordinates of x corresponding to basis B. We do the same for the *nonbasic* variables N, which correspond to all the columns of A not in B, and define A_N and x_N similarly. In the basic solution corresponding to basis B, we set the nonbasic variables to zero, so that $A_N = 0$.

Lemma 3 *For any basis B, there is a unique corresponding basic solution to $Ax = b$.*

Proof: To see this, notice that any such solution has to satisfy

$$\left[\begin{array}{c|c} A_B & A_N \end{array}\right] \begin{bmatrix} x_B \\ - \\ x_N \end{bmatrix} = b$$

Notice that $A_N x_N = 0$, $Ax = b \Rightarrow A_B x_B + A_N x_N = b \Rightarrow A_B x_B = b$. Since A_B is an $m \times m$ matrix of rank m, the solution $x_B = A_B^{-1} b$ is uniquely determined. $\square$

What if $x_B = A_B^{-1} b$ has some $i \in B$ such that $x_i = 0$?

Definition 3 *x is a* degenerate *basic solution if $x_i = 0$ for $i \in B$.*

We can finally give another optimality criterion.

Lemma 4 *Let x be a basic feasible solution and let B be the associated basis. Then:*

1. *If there is a solution y to the system $A^T y = c_B$ such that $A^T y \leq c$, then x is optimal.*
2. *If x is nondegenerate and optimal, then there is a y such that $A_B^T y = c_B$ and $A^T y \leq c$.*

Proof: Suppose we have a y such that $A_B^T y = c_B$ and $A^T y \leq c$. Then for all $i \in B$, $\sum_{i=1}^m a_{ij} y_i = c_i$. Note that $x_j = 0$ for all $j \in N$. Thus for all i such that $x_i > 0$, we have $\sum_{i=1}^m a_{ij} y_i = c_i$. Therefore since x is primal feasible and y is dual feasible and the complementary slackness conditions are obeyed, then x and y must be optimal.

If x is optimal, then there is a dual feasible solution y such that complementary slackness conditions are obeyed. Thus $x_i > 0$ implies that $\sum_{i=1}^m a_{ij} y_i = c_i$. Because x is nondegenerate, $x_i > 0$ for all $i \in B$. Thus $\sum_{i=1}^m a_{ij} y_i = c_i$ for all $i \in B$, and $A^T y = c_B$. Since y is dual feasible, it is also the case that $A^T y \leq c$. $\square$

This brings us to our final answer on how to determine if x is optimal. Since A_B is an $m \times m$ matrix of rank m, $(A_B^T)^{-1}$ exists. So we can solve $A_B^T y = c_B$ for y by setting $y = (A_B^T)^{-1} c_B$. If y is dual feasible, then the lemma above tells us that x must be optimal.

Answer 3 *Given a basic feasible solution x and associated basis B, if $y = A_B^T)^{-1} c_B$ is dual feasible ($A^T y \leq c$), then x must be optimal.*

Call such an y a "verifying y".

Finally, this seems like an answer such that we can actually carry out a reasonably short computation and determine if x is optimal. The real question then is what do we do if x is not optimal.

2 Rewriting the Optimality Condition

Here we introduce the idea of reduced costs, which will be very useful later.

Definition 4 *For any $y \in \mathbb{R}^m$, the* **reduced cost** *$\bar{c}$ with respect to y is $\bar{c} = c - A^T y$.*

Observation 1 *Reduced costs $\bar{c} \geq 0$ with respect to y iff y is dual feasible.*

Lemma 5 *Consider the LP [min $c^T x$ s.t. $Ax = b$, $x \geq 0$] and the alternative LP [min $\bar{c}^T x$ s.t. $Ax = b$, $x \geq 0$] for some $y \in \mathbb{R}^m$. Then $\hat{x}$ is an optimal solution for one iff it is optimal for the other.*

Proof: We have

$$\bar{c}^T x \;=\; (c - A^T y)^T x \;=\; c^T x - y^T A x \;=\; c^T x - y^T b$$

since x satisfies $Ax = b$. So $c^T x - \bar{c}^T x = y^T b$, which is constant since both y and b are given. Since the objective function is just shifted from the other by some constant, an optimal solution for one must be optimal for the other. $\square$

Note 1 *What we have just proved is rather remarkable, for any feasible x the objective function values $c^T x$ and $\bar{c}^T x$ move in tandem, hence minimizing one of them also minimizes the other.*

3 Finding a Better Solution

The main idea now is that given some feasible solution x associated with some basis B, set $y = (A_B^T)^{-1} c_B$. We are now interested in solving the linear problem with the new cost $\bar{c}$:

$$\begin{array}{rc} \min & \bar{c}^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$$

Let us rewrite in terms of the basis B:

$$\begin{array}{rc} \min & \bar{c}_B^T x_B + \bar{c}_N^T x_N \\ \text{s.t.} & A_B x_B + A_N x_N = b \\ & x \geq 0 \end{array}$$

We can multiply the first set of constraints with A_B^{-1} to yield

$$\begin{array}{rc} \min & \bar{c}_B^T x_B + \bar{c}_N^T x_N \\ \text{s.t.} & I x_B + A_B^{-1} A_N x_N = A_B^{-1} b \\ & x \geq 0. \end{array}$$

Since by definition $\bar{c}_B = c_B - A_B^T y$ and because we set $y = (A_B^T)^{-1} c_B$, we have that $\bar{c}_B = c_B - A_B^T (A_B^T)^{-1} c_B = 0$. We can again rewrite the linear program as the following.

$$\begin{array}{rc} \min & 0 x_B + \bar{c}_N^T x_N \\ \text{s.t.} & I x_B + A_B^{-1} A_N x_N = A_B^{-1} b \\ & x \geq 0. \end{array} \tag{2}$$

Next, we simplify the LP by letting $\bar{A} = A_B^{-1} A_N$ and $\bar{b} = A_B^{-1} b$. Now we can do the opposite of what we usually do when starting with inequality constraints. View the variables x_B as slack variables that are constrained to be non-negative and transform the equality constraints into inequality constraints to yield the equivalent minimization problem

$$\begin{array}{rc} \min & \bar{c}_N^T x_N \\ \text{s.t.} & \bar{A} x_N \leq \bar{b} \\ & x_N \geq 0. \end{array}$$

To get a solution of the same value for the previous LP, we set $x_B = \bar{b} - \bar{A}x_N$, which implies the constraint $x_B \geq 0$. We now have a couple of cases.

First, if $\bar{b} \geq 0$ and $\bar{c} \geq 0$, then $x_N = 0$ is optimal because it is feasible and minimizes $\bar{c}_N^T x_N \geq 0$. As a result, $x_B = \bar{b} = A_B^{-1}b$. x_B is feasible since by assumption $\bar{b} \geq 0$. Thus the solution x associated with B is feasible and x is optimal since $\bar{c} \geq 0$.

For the second case, suppose $\bar{c} \ngeq 0$, then there must exist some $j \in N$ such that $\bar{c}_j < 0$. This means that if we increase x_j and keep all other variables in x_N set to zero, we can decrease the value of the solution. How much can we increase x_j? We will look at this problem next time.