ORIE 6300 Mathematical Programming I	September 30, 2008
Lecture 9	
Lecturer: David P. Williamson	*Scribe: Katherine Lai*

1 Recap of last class

Last time, we developed a test for optimality; what we want now is a way of improving a feasible solution. This is the most popular way of solving linear programs. As a recap, we were looking at the following standard primal and dual linear programs:

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

Lemma 1 *A feasible primal solution x is optimal iff there exists a feasible dual solution y such that complementary slackness holds, i.e. $x_j > 0$ implies $\sum_{i=1}^m a_{ij} y_i = c_j$.*

Definition 1 *If A has m linearly independent rows, a set $B \subseteq \{1, \ldots, n\}$ where $|B| = m$ is a* **basis** *if these columns are linearly independent. Let $N = \{1, \ldots, n\} - B$.*

The basic idea is that a basis will correspond to the solution obtained with m equalities $Ax = b$ and $n - m$ equalities $x_j = 0, \forall j \in N$. Let us define A_B as the matrix composed of the columns A_j for $j \in B$ and A_N as the matrix composed of the columns $A_{j'}$ for $j' \in N$. We break apart x and c in the same manner into x_B, x_N, c_B, and c_N.

Observation 1 *Since A has m linearly independent rows, A_B has m linearly independent columns, A_B is an $m \times m$ matrix with full rank. This implies that both A_B^{-1} and $(A_B^T)^{-1}$ exist.*

Lemma 2 *For any basis B, there exists a unique corresponding basic solution (not necessarily feasible), where $x_N = 0$ and some*

$$\begin{aligned} A_B x_B + A_N x_N &= b \\ A_B x_B &= b \\ x_B &= A_B^{-1} b \end{aligned}$$

To see this, note that any such solution has to satisfy

$$\left[\begin{array}{c|c} A_B & A_N \end{array} \right] \begin{bmatrix} x_B \\ - \\ 0 \end{bmatrix} = b$$

Thus, if $Ax = b$ then $A_B x_B = b$, and since A_B is a matrix of full rank, the solution $x_B = A_B^{-1} b$ is uniquely determined.

Lemma 3 *Let x be a basic solution corresponding to a basis B. Then if $A_B^T y = c_B$ has a solution $\bar{y}$ such that $A^T \bar{y} \leq c$ (i.e. $\bar{y}$ is dual feasible), then x is optimal.*

This gives rise to the following optimality condition: Given x and a basis B, compute $\bar{y} = (A_B^T)^{-1} c_B$. If $A^T \bar{y} \leq c$ then x is optimal.

Now, however, suppose the test fails. What do we do then? We will rewrite the optimality condition to make it obvious what you're supposed to do to move to a more optimal solution.

2 Rewriting the Optimality Condition

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

Definition 2 *For any $y \in \mathbb{R}^m$, the* **reduced cost** *$\bar{c}$ with respect to y is $\bar{c} = c - A^T y$.*

Observation 2 *Reduced costs $\bar{c} \geq 0$ with respect to y iff y is dual feasible.*

Lemma 4 *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.

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

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

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

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 need to keep the solution feasible, so we need to maintain the constraint $\bar{A}_{ij} x_j \leq \bar{b}_i$ for all i.

If $\bar{A}_{ij} \leq 0 \ \forall i$, then we can increase x_j as much as we want and still have a feasible solution and a better objective value, which implies that the LP is unbounded in value. Now suppose there exists some i such that $\bar{A}_{ij} > 0$. Then x_j can be no larger than $\frac{\bar{b}_i}{\bar{A}_{ij}}$ for any i where $\bar{A}_{ij} > 0$. Thus we increase x_j to the maximum feasible value:

$$\min_{i : \bar{A}_{ij} > 0} \frac{\bar{b}_i}{\bar{A}_{ij}}$$

Let i^* be the index that achieves the minimum ratio. Recall $x_B = \bar{b} - \bar{A} x_N$. Setting x_j to this new value implies that some variable in x_B will be driven down to 0. In other words, after increasing x_j as much as possible, we will have $x_{i^*} = 0$ for some $i^* \in B$.

Since we now have the same number of variables set to 0 as when we began, this suggests that we have moved to a new "basis" $\hat{B} = B - \{i^*\} \cup \{j\}$. Next lecture, we will show that $\hat{B}$ is indeed a basis. From this new basis, we get a new associated solution $\hat{x}$. We will also show that this $\hat{x}$ is exactly the solution we constructed by increasing x_j as much as possible. The process of switching bases is called "making a pivot". Repeatedly doing this gives us an algorithm for solving LPs, called the simplex method.