ORIE 6300 Mathematical Programming I September 26, 2014

Lecture 10

Lecturer: David P. Williamson *Scribe: Yuhang Ma*

1 From last class

Last time, we introduced the simplex method. In this class we are going to prove that the simplex method indeed works.

Let us first recall what the simplex method does. Consider the standard primal and its dual linear programs:

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

Define a <u>*basis*</u> B as the set of indices of m linearly independent columns of A. Then define

$$A_B = (A_i), \quad for\ i \in B$$

Similarly we have x_B and c_B. Let N denote the set of indices of columns not in B, so that we also have A_N,x_N,c_N.

Suppose we have a basic feasible solution x with associated basis B, so that

$$x_N = 0, \qquad x_B = A_B^{-1} b \geq 0.$$

Now we consider $y = (A_B^T)^{-1} c_B$. We showed last time that if y is dual feasible (i.e. $A^T y \leq c$) then x is an optimal solution to the primal. In the case that y is not dual feasible we introduced the concept of reduced cost:

Definition 1 *For any $y \in \mathbf{R}^m$, the reduced cost $\bar{c}$, with respect to y, is $\bar{c} = c - A^T y$.*

$$\bar{c} \geq 0 \text{ iff y is dual feasible.}$$

We also showed last time that minimizing $\bar{c}^T x$ s.t. $Ax = b, x \geq 0$, has the same optimal solution as original primal problem.

So the primal LP can be rewritten as:

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

Multiplying A_B^{-1} on the both sides of the equality constraints, we get

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

Last time we noticed that $\bar{c}_B = 0$. Set $\bar{A} = A_B^{-1} A_N$,$\bar{b} = A_B^{-1} b$. Then we have

$$\begin{aligned} \min \quad & \bar{c}_N^T x_N \\ s.t. \quad & I x_B + \bar{A} x_N = \bar{b} \\ & x \geq 0. \end{aligned}$$

By setting $x_B = \bar{b} - \bar{A} x_N$, then we turn the primal LP problem into the following form:

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

Note that this program is equivalent to the original primal.

We considered different cases. If $x_N = 0$, $x_B = \bar{b} = A_B^{-1} b$, and $\bar{c} \geq 0$, then this solution is an optimal solution to this new LP and therefore x is an optimal solution to the original primal.

In the case that $\bar{c} \ngeq 0$, since $\bar{c}_B = 0$, it means that there exists a $j \in N$, such that $c_j < 0$. If we increase x_j and keep all other variables in x_N at zero, we can decrease the value of the solution. How much can we increase x_j? We need to keep the solution feasible, that means that we need to maintain the constraints $\bar{A}_{ij} x_j \leq \bar{b}_i, \forall i$.

If $\bar{A}_{ij} \leq 0$ for all i, then as we increase x_j, x remains feasible, so as $x_j \to \infty$, the value of the solution goes to $-\infty$. Therefore the LP is unbounded.

Now suppose there exists some i such that $A_{ij} > 0$. Then x_j can be no larger than $\dfrac{\bar{b}_i}{\bar{A}_{ij}}$ for any i, where $\bar{A}_{ij} > 0$. Thus we increase x_j to the maximum feasible value:

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

Let i^* be the index that achieves the minimum. Recall $x_B = \bar{b} - \bar{A} x_N$. Setting $x_j = \varepsilon$ 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 $i^* \in B$.

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"

$$\hat{B} = B - \{i^*\} \cup j.$$

We will show in the next section 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 "pivoting". Repeatedly doing this gives us an algorithm for solving LPs, called the simplex method, which is due to George Dantzig.

Summarize the above idea as the following steps:

If $\bar{c} \not\geq 0$, $\exists\ j$ *s.t.* $c_j < 0, j \in N$.
(Check for unboundedness)
If $\bar{A}_{ij} \leq 0, \forall i$, *then primal LP unbounded.*
(Ratio Test)
Compute $\varepsilon = \min_{i:\bar{A}_{ij}>0} \frac{\bar{b}_i}{\bar{A}_{ij}}$
Increase x_j *by* ε
Let i^* *be the* i *that attains the minimum. Since* $x_B = \bar{b} - \bar{A}_N(\varepsilon e_j) \Rightarrow x_{i^*} = 0$, *for* $i^* \in B$.
(Update basis)
$\hat{B} \leftarrow B \cup \{j\} - \{i^*\}$
We say j *enters the basis,* i^* *leaves the basis*

2 Some details

Let $\hat{x}$ be the new solution found by the method described above, i.e. $\hat{x}_j = \varepsilon$, $\hat{x}_k = 0$, for all $k \in N, k \neq j$, and $\hat{x}_B = \bar{b} - \bar{A}\hat{x}_N$. Now we want to prove that the simplex method, under some mild conditions, leads to the optimal solution. In order to do this we are going to show 4 claims:

1. $c^T\hat{x} \leq c^T x$ i.e. the new solution is not worse than the old solution;
2. If x is nondegenerate, then $\varepsilon > 0$, i.e. we make progress in our algorithm;
3. The updated basis $\hat{B}$ after a pivot is indeed a basis;
4. $\hat{x}$ is the unique solution corresponding to $\hat{B}$.

Claim 1 $\bar{c}^T x \leq c^T x$

Proof: Since we already know $\bar{c}_B^T = 0$, $x_N = 0$, $\hat{x}_j = \varepsilon, \hat{x}_k = 0$, for all $k \in N, k \neq j$,

$$\bar{c}^T\hat{x} = \bar{c}_B^T\hat{x}_B + \bar{c}_N^T\hat{x}_N = \bar{c}_j\varepsilon \leq 0 = \bar{c}_B^T x_B + \bar{c}_N^T x_N = \bar{c}^T x$$

where the inequality holds because $\bar{c}_j < 0$ and $\varepsilon \geq 0$. □

This means that at least our solution value does not decrease by applying the simplex method, but do we in fact make progress? In fact, if $\varepsilon > 0$, then the inequality holds strict: $\bar{c}^T x < c^T x$.

Claim 2 *If* x *is nondegenerate, then* $\varepsilon > 0$ *and thus* $\bar{c}^T x < c^T x$.

Proof: Since x is a nondegenerate basic solution, we know that $x_j > 0$ for all $j \in B$, i.e. $x_B > 0$.
Recall

$$x_B = A_B^{-1} b = \bar{b} > 0,$$

then

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

for all i with $\bar{A}_{ij}$, therefore $\varepsilon > 0$. □

Let us for now assume that all basic feasible solutions are nondegenerate. We will treat the case of degenerate solutions in a later lecture. So far, we know that we make progress with the simplex method, but what do we get after making a pivot?

Claim 3 *The set $\hat{B}$ after a pivot is a basis.*

Proof: By definition of a basis, $\hat{B}$ is a basis if and only if $A_{\hat{B}}$ has full rank. To get $A_{\hat{B}}$ we substituted the jth column of A for the i^*th column into A_B.

$$\begin{aligned}
A_{\hat{B}} &= [\text{ old columns}|A_j|\text{ old columns }] \\
&= A_B \left[\begin{array}{ccc|c|ccc}
1 & 0 & \dots & & \dots & 0 & 0 \\
0 & 1 & & & & 0 & 0 \\
\vdots & \vdots & \ddots & A_B^{-1}A_j & \ddots & \vdots & \vdots \\
0 & 0 & & & & 1 & 0 \\
0 & 0 & \dots & & \dots & 0 & 1
\end{array}\right] \\
&= A_B \left[\begin{array}{ccc|c|ccc}
1 & 0 & \dots & \bar{A}_{1j} & \dots & 0 & 0 \\
0 & 1 & & \bar{A}_{2j} & & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & & \bar{A}_{(n-1)j} & & 1 & 0 \\
0 & 0 & \dots & \bar{A}_{nj} & \dots & 0 & 1
\end{array}\right],
\end{aligned}$$

recalling that $\bar{A} = A_B^{-1}A_N$. B is a basis so A_B is non-singular. In order to show that $A_{\hat{B}}$ is non-singular we need to show that the big matrix on the righthand-side is non-singular. But we chose $i*$ in the ε ratio such that $\bar{A}_{i^*j} > 0$ and therefore the matrix is non-singular and $\hat{B}$ is a basis. $\square$

Lemma 4 *The new solution $\hat{x}$ is the solution corresponding to $\hat{B}$.*

Proof: We want to show that $\hat{x}$ is the solution we get by setting $x_{\hat{N}} = 0$, $x_{\hat{B}} = A_{\hat{B}}^{-1}b$.

Note $\hat{x}_k = 0$ for all $k \notin \hat{B}$ (i.e. for all k $\in N - \{j\} \cup \{i^*\}$). We have

$$\bar{A}\hat{x}_N + I\hat{x}_B = \bar{A}(\varepsilon e_j) + I(\bar{b} - \bar{A}(\varepsilon e_j)) = \bar{b}$$

Recall that $\bar{A} = A_B^{-1}A_N$, $\bar{b} = A_B^{-1}b$, and therefore

$$A_N\hat{x}_N + A_B\hat{x}_B = b$$

and we get $A\hat{x} = b$, $\hat{x} \geq 0$, so $\hat{x}$ is a feasible solution with corresponding basis $\hat{B}$. $\square$

3 Some Issues to Deal with

There are some issues we have to address when using the simplex method. We will go over these issues in the coming lectures.

1. About running time:
   (a) How much work is involved in every pivot step?
   (b) How many pivots do we need to reach the optimal solution?
   (If all solutions encountered are nongenerate, then from Claim 2, we know that each basis encountered is unique $\Rightarrow$ # of pivots $\leq$ # of bases$=\binom{n}{m}$)

2. Starting point: We assume that we have a feasible solution to begin our algorithm, but how do we find such a initial feasible solution?

3. How can we guarantee progress towards optimality, if x is degenerate?

4. Assume we are in the case where $\bar{c} \ngeq 0$, i.e. there exists j such that $c_j < 0$. Which one of these c_j's do we choose?

   (a) j that gives the most improvement?

   (b) First j such that $\bar{c}_j < 0$?

   (c) j such that $\bar{c}_j$ most negative?