Lecture 11

Lecturer: Damek Davis *Scribe: Lijun Ding*

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. To be clear about notation, let us consider the primal and dual linear program below and the corresponding terminology. Consider the standard primal and its dual linear programs:

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

A <u>basis</u> B is 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. We write the constraints in primal as $x \in P(A, b)$ and the constraints in the dual as $y \in Q(A^T, c)$.

We now give the definition of nondegeneracy and the proposition we proved last time related to it in testing optimality.

Definition 1 *A vertex x of $P(A, b)$ is called nondegenerate if x has exactly m nonzero elements.*

Proposition 1 *Suppose $\bar{x} \in P(A, b)$ has m nonzero components corresponding to a basis B. Then $\bar{x}$ is optimal if and only if, there exists $y \in Q(A^T, c)$ such that $A_B^T y = c_B$.*

In each step of simplex, we consider $y = (A_B^T)^{-1} c_B$, the verifying y. The above proposition shows that if y is dual feasible (i.e. $A^T y \leq c$) then x is an optimal solution. In the case that y is not dual feasible we introduced the concept of reduced cost:

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

The following lemma and linear programs we discussed last time expalin why the reduced cost might be considered.

Lemma 2 *Let $x \in P(A, b)$. Then Range$(A^T) + N_{P(A,b)}(x) = N_{P(A,b)}(x)$.*

Note that for our veryfying $\bar{y} = (A^T)^{-1}c_B$ and reduced cost $\bar{c}$, the following equivalence holds,

$$A^T\bar{y} \le c \iff c - A^T y \ge 0 \iff \bar{c} \ge 0.$$

Starting from Lemma 2, the following program we discussed last time are equivalent to the primal in the sense the solution x can be transoformed back and forth easily. We start with the program with same constraints as our primal but different objective $\bar{c}^T x$.

$$\begin{aligned} \min \quad & \bar{c}^T x \\ s.t. \quad & Ax = b \\ & x \ge 0. \end{aligned}$$

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

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

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

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

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

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

Let us now summarize the simplex-type calculations we did last time. Recall that the goal of those calculations was to find an optimal solution or update our basis B.

- Start with a basis B and its corresponding vertex assumed to be nondegenerate $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}, x_B = A_B^{-1}b, x_N = 0$ (Note: $Ax = A_B x_B + A_N x_N = b$).
- Compute veryfying $y = A_B^{-T}c_B$.
- Case I: Suppose $A^T y \le c$, then x is primal optimal by Proposition 1.
- Case II: Suppose $A^T y \not\le c$. Then $\bar{c} \not\ge 0$ and $\exists j \in \{1, \dots, n\}$ s.t. $\bar{c}_j < 0$.
- In the second case, we need to update our basis B as the "verifying $y = (A_B^T)^{-1}c_B$ " does not work.
- The index j we find above satisfies $j \in N$ because $\bar{c}_B = c_B - A_B^T A_B^{-T} y = c_B - c_B = 0$.

- Let's update x to $\hat{x}$(the new iterates) by increasing $x_{N(j)} = 0$ to some $\epsilon > 0$. Here $N(j)$ means the index of jth entry of x in x_N.
- Can increase the value as long as

$$\forall i \in \{1, 2, \ldots, m\},\ (A_B^{-1} A_N)_{iN(j)} x_{N(j)} \leq (A_B^{-1} b)_i$$

(Note that our x is a vertex of $P(A, b)$ and so $A_B^{-1} b = x_B \geq 0$)

- We Let $\bar{A} = A_B^{-1} A_N$ and $\bar{b} = A_B^{-1} b$.
- If $\forall i$, $\bar{A}_{iN(j)} \leq 0$, then we can take $x_j \to +\infty$. Since $\bar{c}_B = 0$ as shown above, $\bar{c}^T x = \bar{c}_N^T x_N = \bar{c}_{N(j)}^T x_{N(j)} \to -\infty$ as $x_N(j) \to \infty$ and $\bar{A}_{N(j)} \leq 0 \leq \bar{b}$. So the objective of the reduced cost program in Lemma 2 is unbounded below by setting $x = \begin{bmatrix} x_B - t\bar{A}_{N(j)} \\ 0 \end{bmatrix} + te_j$ where e_j is the vector with 1 at jth entry and 0 otherwise.
- if $\exists i$ s.t. $\bar{A}_{iN(j)} > 0$, then we set $i^* = \arg\min_{i:\bar{A}_{iN(j)}>0} \frac{\bar{b}_i}{\bar{A}_{iN(j)}}$ and $\epsilon = \frac{\bar{b}_{i^*}}{\bar{A}_{i^*N(j)}}$.
- For our new iterates $\hat{x} = \begin{bmatrix} \hat{x}_B \\ \hat{x}_N \end{bmatrix}$, we set $\hat{x}_j = \epsilon$, $\hat{x}_k = 0, \forall k \in N/\{j\}$ and $\hat{x}_B = \bar{b} - \bar{A}\hat{x}_N$.
- Observe that $i^* \in B$ because it corresponds to one of the rows of $\bar{A} = A_B^{-1} A_N$, which in turn is one of the columns of A_B. It is the index i^* of $\bar{b} = \bar{x}_B$

Summarize the above idea as the following steps:

If $\bar{c} \ngeq 0$, $\exists\ j$ 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 $\varepsilon = \min_{i:\bar{A}_{ij}>0} \dfrac{\bar{b}_i}{\bar{A}_{ij}}$
  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

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.

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 follwoing 4 claims:

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

Claim 3 $c^T\hat{x} \leq c^T x$

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 $\varepsilon \geq 0$. Now $c^T\hat{x} = (\bar{c} + A^T y)^T\hat{x} = \bar{c}^T\hat{x} + b^T y \geq \bar{c}^T x + b^T y = (\bar{c} + A^T y)^T x = c^T x$ since both $\hat{x}, x$ are feasible. □

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

Claim 4 *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 $\bar{A}_{ij} < 0$. 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 5 *The set $\hat{B}$ 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. $\square$

Claim 6 *The new solution $\hat{x}$ is the basic feasible 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?