ORIE 6300 Mathematical Programming I	September 22, 2016
Lecture 10	
Lecturer: Damek Davis	*Scribe: Qinru Shi*

1 Last Time

- We shifted to the standard form ($A \in \mathbb{R}^{m\times n}$).

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

  Also, $P(A,b) = \{x \mid A^T x = b, x \ge 0\}$.

- **Assumption:** From now on, we assume that the rows of A are linearly independent ($m \le n$).

- **Proposition 1** *For $P(A,b) = \{x \mid A^T x = b, x \ge 0\}$, $\overline{x} \in P(A,b)$ is a vertex if and only if the columns corresponding to positive components of $\overline{x}$ are linearly independent.*

- **Definition 1** *A set B of m columns of A ($A \in \mathbb{R}^{m\times n}$) is a basis if rank$(A_B) = m$ (A_B is invertible).*

- **Proposition 2** *Every vertex of $P(A,b)$ is a basic feasible solution corresponding to some basis.*

2 Verifying optimality

Given a vertex $\overline{x} \in P(A,b)$, we want to find a "verifying $\overline{y}$" which we use to check whether $\overline{x}$ is optimal. Clearly, any $\overline{y}$ should satisfy complementary slackness, i.e.

$$\overline{x}_i > 0 \iff (A^T \overline{y})_i = c_i.$$

In order to find $\overline{y}$, we first prove the following proposition.

Proposition 3 *Suppose $\overline{x} \in P(A,b)$ has m non-zero components corresponding to basis B. Then, $\overline{x}$ is optimal if and only if there exists $y \in Q(A^T, c)$ such that $A_B^T y = c_B$.*

Proof: By theorem from previous lectures, we know that $\overline{x}$ is optimal if and only if

$$-c \in N_{P(A,b)}(\overline{x}) = \{\hat{c} \mid \exists \hat{y} \in Q(A^T, \hat{c}),\ s.t.\ \ \overline{x}^T(\hat{c} - A^T\hat{y}) = 0\},$$

which is then equivalent to the condition

$$\exists y \in Q(A^T, c)\ \ s.t.\ \ x_i > 0 \Rightarrow (A^T y)_i = c_i.$$

Note that $x_i > 0$ if and only if $i \in B$, so the last condition can be translated into

$$\exists y \in Q(A^T, c) \;\; s.t. \;\; A_B^T y = c_B.$$

□

We also have the following definition regarding non-degeneracy.

Definition 2 *A vertex $\overline{x} \in P(A, b)$ is called non-degenerate if it has exactly m nonzero components and those components correspond to some basis B.*

For each non-degenerate $\overline{x}$ with basis B, we will try to find a $y \in Q(A^T, c)$ such that $A_B^T y = c_B$. Then, $(c - A^T y)_i = 0$ when $i \in B$. For simplicity of notation and computation, we create the following definition.

Definition 3 *For any $y \in \mathbb{R}^m$, the reduced cost $\overline{c}$ with respect to y is $\overline{c} = c - A^T y$.*

Lemma 4 *For all $x \in P(A, b)$, $\text{range}(A^T) + N_{P(A,b)}(x) = N_{P(A,b)}(x)$.*

Proof: Clearly, $N_{P(A,b)}(x) \subseteq \text{range}(A^T) + N_{P(A,b)}(x)$ because $0 \in \text{range}(A^T)$. On the other hand, let $A^T y_0$ be any element in $\text{range}(A^T)$ and $-\hat{c}$ be any element in $N_{P(A,b)}(x)$. Then, there exists $\hat{y}$ such that $A^T \hat{y} \leq \hat{c}$ and $x^T(\hat{c} - A^T \hat{y}) = 0$. Tweaking the equations a bit, we get

$$A^T(\hat{y} - y_0) \leq \hat{c} - A^T y_0$$

and

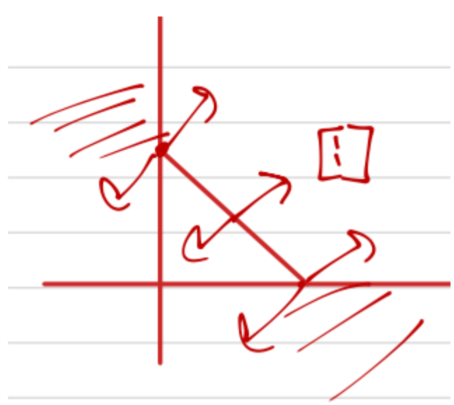
$$x^T(\hat{c} - A^T y_0 - A^T(\hat{y} - y_0)) = 0.$$

Thus,

$$A^T y_0 - \hat{c} \in N_{P(A,b)}(x),$$

which means that $\text{range}(A^T) + N_{P(A,b)}(x) \subseteq N_{P(A,b)}(x)$. □

A simple example to illustrate Lemma 5 is when $A = [1, 1]$, $b = 1$ and $c \in \mathbb{R}^2$. $P(A, b)$ in this case is a line segment connecting $(1, 0)$ and $(0, 1)$. If c is in $N_{P(A,b)}$, then $c + \begin{bmatrix} y \\ y \end{bmatrix}$ is still in $N_{P(A,b)}$.



From Lemma 5, we can easily reach the following conclusion.

Lemma 5 *Consider the two LPs*

$$(LP1) \quad \begin{array}{ll} min & c^T x \\ s.t. & Ax = b \\ & x \geq 0 \end{array} \qquad\qquad (LP2) \quad \begin{array}{ll} min & \overline{c}^T x \\ s.t. & Ax = b \\ & x \geq 0. \end{array}$$

These LPs have the exact same optimal solution.

Proof: We know that $\overline{x}$ solves (1) if and only if $-c \in N_{P(A,b)}(\overline{x})$. By Lemma 5, we have

$$N_{P(A,b)}(\overline{x}) = \text{range}(A^T) + N_{P(A,b)}(\overline{x}),$$

so $c \in N_{P(A,b)}(\overline{x})$ if and only if $\overline{c} = -c + A^T y \in N_{P(A,b)}(\overline{x})$, which equivalent to $\overline{x}$ being an optimal solution of (2). □

We also have the following observation regarding $\overline{c}$.

Observation 1 $\overline{y} \in Q(A^T, c) \iff A^T\overline{y} \leq c \iff \overline{c} = c - A^T\overline{y} \geq 0.$

Using the conclusions above, we can now work with $\overline{c}$ instead of c.

3 Some Simplex Method-type Computations

Consider a basis B and the new linear program with reduced cost $\overline{c}$.

$$(LP2) \quad \begin{array}{ll} \min & \overline{c}^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0. \end{array}$$

We reorganize the vectors and matrices so that $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$, $A = [A_B \ A_N]$ and $\overline{c} = \begin{bmatrix} \overline{c}_B \\ \overline{c}_N \end{bmatrix}$. Note that $x_B > 0$, $x_N = 0$ and

$$[A_B \ A_N] \begin{bmatrix} x_B \\ x_N \end{bmatrix} = b.$$

Multiplying through by A_B^{-1}, we convert LP2 into

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

Now set $y = A_B^{-T} c_B$ and $\overline{c} = c - A^T y$. Then, $\overline{c}_B = c_B - A_B^T A_B^{-T} c_B = 0$. Our problem becomes

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

We can further simplify the problem by discarding x_B. Then, the LP looks like

$$\begin{array}{lll} (LP3) & \min & \bar{c}_N^T x_N \\ & \text{s.t.} & A_B^{-1} A_N x_N \le A_B^{-1} b \\ & & x_N \ge 0. \end{array}$$

If $y \in Q(A^T, c)$, then $\bar{c} = c - A^T y \ge 0$. Because $\bar{c}_N^T x_N \ge 0$, the optimal value of the reduced problem is bounded below by 0. Thus, if $A_B^{-1} b \ge 0$, then $x_N = 0$ is feasible and $\bar{c}^T 0 = 0$, so $x_N = 0$ is an optimal solution. In the original problem, $x = \begin{bmatrix} A_B^{-1} b \\ 0 \end{bmatrix}$ is optimal.

Now suppose B is an arbitrary basis. Let $x_N = 0$ and $x_B = A_B^{-1} b$. Based on the analysis above, we devise the following simplex scheme for finding the optimal solution:

- If $A_B^{-1} b \ge 0$, then $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$ is feasible and a vertex.
- If, in addition, x has m nonzero components, then solve $y = A_B^{-1} c_B$ and compute reduced cost $\bar{c}$ with respect to y.
- If $\bar{c} \ge 0$, then x and y are optimal by Proposition 3.
- **Failure Case**: If there exists i such that $\bar{c}_j < 0$, then $j \in N$ since $c_B = 0$. Increase x_j and we get an improvement in the objective value of LP3. Then, all entries of x_N are zero except for $x_{N(j)}$ which corresponds to x_j. Note that we can increase $x_{N(j)}$ as long as

  $$(A_B^{-1} A_N)_{iN(j)} x_j \le (A_B^{-1} b)_i$$

  for all i.

  We have assumed that $x = A_B^{-1} b$ is feasible so $A_B^{-1} b \ge 0$.

  If $(A_B^{-1} A_N)_{iN(j)} \le 0$ for all i, then the constraint is never violated as $x_j \to \infty$, and $c_j x_j \to -\infty$ as $x_j \to \infty$. Therefore, in this case, LP is unbounded.

  If there exists i such that $(A_B^{-1} A_N)_{iN(j)} > 0$, then we cannot improve $x_{N(j)}$ over $\frac{(A_B^{-1} b)_i}{(A_B^{-1} A_N)_{iN(j)}}$. Define

  $$i^* = \text{argmin}_{i:(A_B^{-1} A_N)_{iN(j)} > 0} \frac{(A_B^{-1} b)_i}{(A_B^{-1} A_N)_{iN(j)}}$$

  and

  $$\epsilon = \frac{(A_B^{-1} b)_{i^*}}{(A_B^{-1} A_N)_{i^* N(i)}}.$$

  Hence, we can increase x_j to ϵ and get a new point $\hat{x}_N$. The new point has a strictly smaller cost $\bar{c}_N^T \hat{x}_N$ than $\bar{c}_N^T x_N = 0$.

  Now we update x_B through the equation

  $$\hat{x}_B = A_B^{-1} b - A_B^{-1} A_N \hat{x}_N.$$

$\hat{x}_B \geq 0$ by construction. Clearly, $\hat{x} = \begin{bmatrix} \hat{x}_B \\ \hat{x}_N \end{bmatrix}$ satisfies $A\hat{x} = b$.

Notice that $\hat{x}_{B(i^*)} = 0$ by construction. Thus, we can update our basis to

$$\hat{B} = (B \setminus \{i^*\}) \cup \{j\}.$$

To summarize, the main update step works as follows:

- If $\overline{c} \ngeq 0$, find j, s.t. $\overline{c}_j < 0$. ($j \in N$).
- Check for unboundedness: If $(A_B^{-1}A_N)_{iN(j)} \leq 0$ for all i, then LP is unbounded.
- Ratio test: Compute

  $$i^* = \arg \min_{i:(A_B^{-1}A_N)_{iN(j)}>0} \frac{(A_B^{-1}b)_i}{(A_B^{-1}A_N)_{iN(j)}}$$

  and

  $$\epsilon = \frac{(A_B^{-1}b)_{i^*}}{(A_B^{-1}A_N)_{i^*N(j)}}.$$

- Update x: $x_i \leftarrow \epsilon$ and $x_B \leftarrow A_B^{-1}b - A_B^{-1}A_N x_N$.
- Update basis: $B \leftarrow (B \setminus \{i^*\}) \cup \{j\}$. We say that j enters the basis and i^* leaves the basis.