MATH 464 – Lecture 12 (02/15/2018)

Today:
* simplex method
* feasible and basic direction
* reduced cost

Simplex Method (Chapter 3 in BT-ILO)

0-simplex ≡ vertex
1-simplex ≡ edge
2-simplex ≡ triangle

But in LP, simplex refers to the polyhedron of the problem.

We will describe the method for an LP in standard form $\left\{\begin{array}{ll} \min & \bar{c}^T\bar{x} \\ \text{s.t.} & A\bar{x}=\bar{b} \\ & \bar{x}\geq\bar{0} \end{array}\right\}$.
where $A\in\mathbb{R}^{m\times n}$, $\bar{b}\in\mathbb{R}^m$, $\text{rank}(A)=m$, $m\leq n$.

We want to define **optimality conditions** for $\bar{x}\in P$. If these conditions are satisfied, then $\bar{x}$ is an optimal solution.

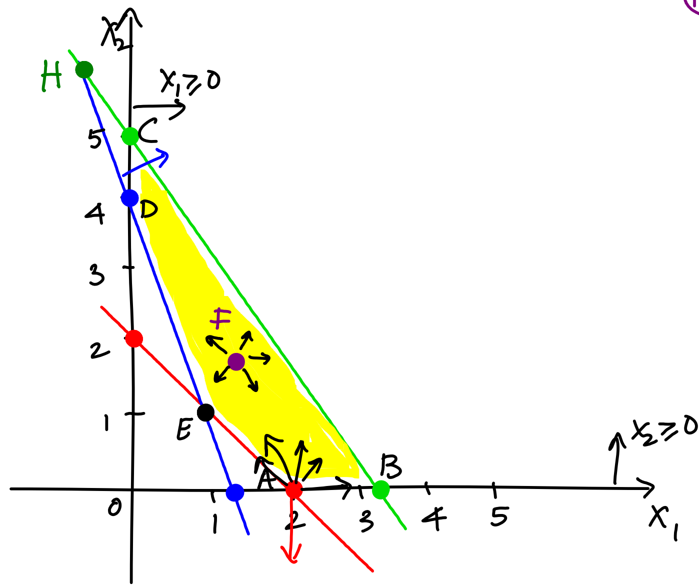
Algorithm: Given some $\bar{x}\in P$, check optimality conditions. If they are not satisfied, we search "nearby" to see if we can improve, i.e., decrease, $\bar{c}^T\bar{x}$.

In 1D calculus, the optimality conditions for $x\in\mathbb{R}$ to be a local minimum of $f(x)$ are $f'(x)=0$ and $f''(x)>0$. If $f(x)$ is a convex function, these conditions also guarantee that x is a global minimum.

For LPs, since $f(\bar{x})=\bar{c}^T\bar{x}$ is linear, and since P is a polyhedron (hence convex), a local optimum is also a global optimum.

When searching "nearby", we want to make sure we always stay feasible, i.e., we do not want to go outside the feasible region.

Suppose we are at $A(2,0)$. We can consider directions to move. If we move straight down, i.e., along $\bar{d}' = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, we will go outside the feasible region. But $d = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a good direction — we can move right all the way up to $B(10/3, 0)$.



In general, from $\bar{x}$, we consider $\bar{x} + \theta \bar{d}$ for $\theta > 0$ to move along $\bar{d}$. We want $\bar{x} + \theta \bar{d} \in P$.

Thus, no $\theta > 0$ exists for $\bar{d}'$, while any $0 < \theta \leq \frac{4}{3}$ works for moving along $\bar{d}$. Similarly, if we move Northwest, i.e., along $\bar{d}' = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, we can go all the way to $E(1,1)$. So any $0 < \theta \leq 1$ works. Further, all directions "in between" $d = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $d'' = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ are good — see figure.

If F is in the interior of P, then any $\bar{d}$ is a good direction (see figure).

We formalize this notion of a "good" direction now.

Def Let $\bar{x} \in P$. A vector $\bar{d} \in \mathbb{R}^n$ is a **feasible direction** at $\bar{x}$ if there exists $\theta > 0$ such that $\bar{x} + \theta \bar{d} \in P$.

Notice that θ can be arbitrarily small, as long as it is > 0.

For instance, at $\bar{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\bar{d} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ is not a feasible direction, as

$$\bar{x} + \theta \bar{d} = \begin{bmatrix} 2 + 0\theta \\ 0 + (-1)\theta \end{bmatrix} \overset{A}{=} \begin{bmatrix} 2 \\ -\theta \end{bmatrix} \notin P \text{ for any } \theta > 0.$$

We now describe how to move from one bfs to an adjacent bfs using (linear) algebra — how do we actually implement the "move"?

Let $\bar{x}$ be a bfs for $\underbrace{B(1), \ldots, B(m)}_{\text{basis's}}$. With $\bar{x}_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}$, and $\bar{x}_N$ collecting the remaining nonbasic variables, we set $B = [A_{B(1)} \cdots A_{B(m)}]$ as the basis's matrix (made of the basic columns of A). We get that B^{-1} exists, and after setting $\bar{x}_N = \bar{0}$, we can find $\bar{x}_B = B^{-1}\bar{b}$.

Consider the direction $\bar{d}$ at $\bar{x}$ such that $d_j = 1$ for a non-basic x_j, and $d_i = 0$ for $i \neq B(1), \ldots, B(m), j$. all other non basic d_i's are set to zero.

Let $\bar{d}_B = \begin{bmatrix} d_{B(1)} \\ \vdots \\ d_{B(m)} \end{bmatrix}$. $\bar{d}_N = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ corresponding to x_j

We want to move to $\bar{x} + \theta\bar{d}$, $\theta > 0$, and stay feasible. So we need $A(\bar{x} + \theta\bar{d}) = \bar{b}$, and $\bar{x} + \theta\bar{d} \geq 0$.

But $A\bar{x} = \bar{b}$, so we get $\theta(A\bar{d}) = \bar{0}$, i.e., $A\bar{d} = \bar{0}$, which can be written as

$$B\bar{d}_B + \sum_{i \notin B(1), \ldots, B(m)}' A_i d_i = \bar{0}, \text{ i.e., } B\bar{d}_B + A_j = \bar{0}$$

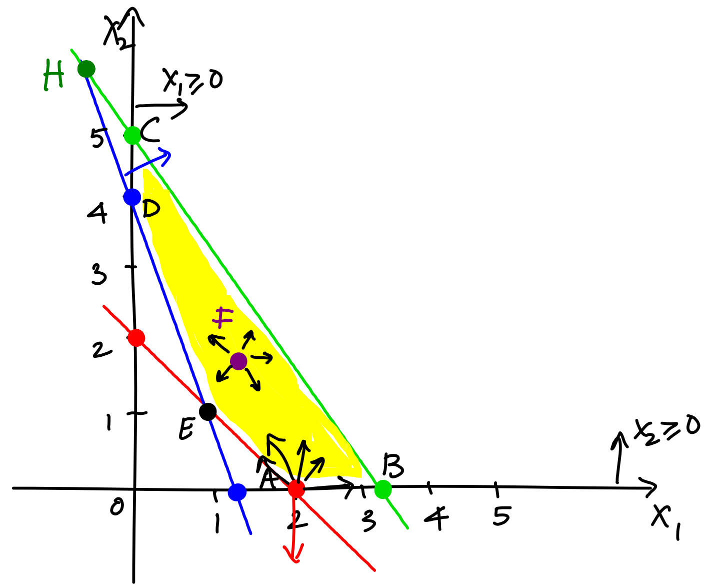
A_j: j^{th} column of A

so $\boxed{\bar{d}_B = -B^{-1}A_j}$

This $\bar{d}_B$ defines the **j^{th} basic direction at bfs $\bar{x}$.**

Notice that there are $(n-m)$ basic directions at a bfs $\bar{x}$.

Back to our example:



$$A = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 3 & 1 & 0 & -1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{bmatrix} \qquad \bar{b} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}$$

BFS corresponding to $A(2,0)$.

$\{B(1), B(2), B(3)\} = \{1, 4, 5\}$

$$B = A(:, [1\,4\,5]) = \begin{bmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

$j=2$: $\quad A_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \qquad \bar{d}_B = -B^{-1}A_2 = \begin{array}{c} \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \end{array}\!\!\begin{array}{c} 1 \\ 4 \\ 5 \end{array} \implies \bar{d} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}$ move toward E along $\overline{AE}$

$d_2 = 1, \; d_3 = 0$

$j=3$: $\quad d_3 = 1, \; d_2 = 0 \qquad \bar{d}_B = -B^{-1}A_3 = \begin{array}{c} \begin{bmatrix} 1 \\ 3 \\ -3 \end{bmatrix} \end{array}\!\!\begin{array}{c} 1 \\ 4 \\ 5 \end{array} \implies \bar{d} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 3 \\ -3 \end{bmatrix}$ move toward B along $\overline{AB}$

See Octave session on the course web page for details.

When we move along $\bar{d}$ from $\bar{x}$ to $\bar{x} + \theta\bar{d}$, $A\bar{x} = \bar{b}$ is satisfied.

How about $\bar{x} \geq \bar{0}$? For the nonbasic variables, $x_j > 0$, and $x_i = 0$ for $i \neq j$, i nonbasic. What about $\bar{x}_B$?

(a) If $\bar{x}$ is non-degenerate, i.e., $\bar{x}_B > 0$, $\bar{x}_B + \theta \bar{d}_B \geq \bar{0}$ as long as θ is sufficiently small.

(b) If $\bar{x}$ is degenerate, a good $\theta > 0$ may not exist, as there might be some $B(i)$ such that $x_{B(i)} = 0$ and $d_{B(i)} = -1$, which will make $x_{B(i)} = -\theta < 0$!

We will describe the non-degenerate case first.

Now, let's incorporate the objective function $\min \bar{c}^T \bar{x}$. Let $\bar{d}$ be the j^{th} basic direction. We move from $\bar{x}$ to $\bar{x} + \theta \bar{d}$, and observe how $\bar{c}^T \bar{x}$ changes, i.e., from $\bar{c}^T \bar{x}$ to $\bar{c}^T(\bar{x} + \theta \bar{d}) = \bar{c}^T \bar{x} + \theta (\bar{c}^T \bar{d})$.

$\bar{c}^T \bar{d}$ is the rate of change of the cost function when moving along $\bar{d}$.

$$\bar{c}^T \bar{d} = \bar{c}_B^T \bar{d}_B + \bar{c}_N^T \bar{d}_N = \bar{c}_B^T \bar{d}_B + c_j \quad \text{as } \bar{d}_N = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j^{th} \text{ entry}$$

$$= \bar{c}_B^T(-B^{-1}A_j) + c_j$$

$$= c_j - \bar{c}_B^T B^{-1} A_j$$

$c'_j = \underbrace{c_j}_{\text{cost per unit increase of } x_j} - \underbrace{\bar{c}_B^T B^{-1} A_j}$ is the **reduced cost** of x_j.

book uses $\bar{c}$

compensating change so that $A\bar{x} = \bar{b}$ is still satisfied — reduce the cost by this amount.

In vector form $\bar{c}'^T = \bar{c}^T - \bar{c}_B^T B^{-1} A$.

In particular, for the basic columns, we get

$$\bar{c}_B'^T = \bar{c}_B^T - \bar{c}_B^T \underbrace{B^{-1}B}_{I} = \bar{c}_B^T - \bar{c}_B^T = \bar{0}.$$

Thus, the reduced cost of every basic variable is zero.

e.g., At $A(2,0)$, the bfs in standard form is $\bar{x} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \\ 4 \end{bmatrix}$,

and the reduced cost is $\bar{c}' = \begin{bmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 0 \end{bmatrix}$.

Since $c_2' = -1$, if we move along the 2^{nd} basic direction, the cost will decrease.

If $\bar{c}' \geq \bar{0}$, then none of the basic directions could give an improvement in $\bar{c}^T\bar{x}$. In other words, the solution is optimal!