MATH464 – Lecture 9 (02/07/2023)

Today: * bfs $\Rightarrow$ extreme point (proof)
* finding all basic solutions

Proof of Theorem 2.3 (contd..) 2. Extreme pt $\Rightarrow$ bfs

We prove, equivalently, that (not bfs) $\Rightarrow$ (not extreme point).

contrapositive argument: $(A \Rightarrow B) \equiv (\neg B \Rightarrow \neg A)$ for logical statements A, B

"equivalent" "negation" or "not"

Let $\bar{x}^* \in P$ be not a bfs. Hence among all the active constraints at $\bar{x}^*$, there do not exist a set of n LI constraints.

Let $\bar{a}_i^T \bar{x}^* = b_i$, $i \in I$ be the set of constraints active at $\bar{x}^*$.

Hence, $\{\bar{a}_i \in \mathbb{R}^n, i \in I\}$ are not LI. index set

Hence, there exists some $\bar{d} \in \mathbb{R}^n$, $\bar{d} \neq \bar{0}$, such that $\bar{a}_i^T \bar{d} = 0 \ \forall i \in I$.

Recall from linear algebra that if the columns of $A = [\bar{a}_1 \cdots \bar{a}_k]$ are not LI, then the homogeneous system $A\bar{x} = \bar{0}$ has a non-trivial solution.

Let $\bar{y} = \bar{x}^* + \epsilon \bar{d}$ and $\bar{z} = \bar{x}^* - \epsilon \bar{d}$ for $\epsilon > 0$, but small.

We want to show that $\bar{y}, \bar{z} \in P$ for ϵ small enough.

We get $\bar{a}_i^T \bar{y} = b_i \ \forall i \in I$ as $\bar{a}_i^T(\bar{x}^* + \epsilon \bar{d}) = b_i + \epsilon 0 = b_i$.

If $i \notin I$, $\bar{a}_i^T \bar{x}^* > b_i$. (We take $P = \{\bar{x} \in \mathbb{R}^n \mid A\bar{x} \geqslant \bar{b}\}$

or $\bar{a}_i^T \bar{x} \geqslant b_i$, $i = 1, \ldots, m$)

constraint i is not active

We also get $\bar{a}_i^T \bar{y} = \bar{a}_i^T(\bar{x}^* + \epsilon \bar{d}) > b_i + \epsilon \underbrace{\bar{a}_i^T \bar{d}}_{\neq 0 \text{ (as } i \notin I)} > b_i$ for ϵ small enough.

In particular, we can choose $\epsilon < \max\limits_{i \in I} \left\{ \dfrac{(\bar{a}_i^T \bar{x}^* - b_i)}{|\bar{a}_i^T \bar{d}|} \right\}$.

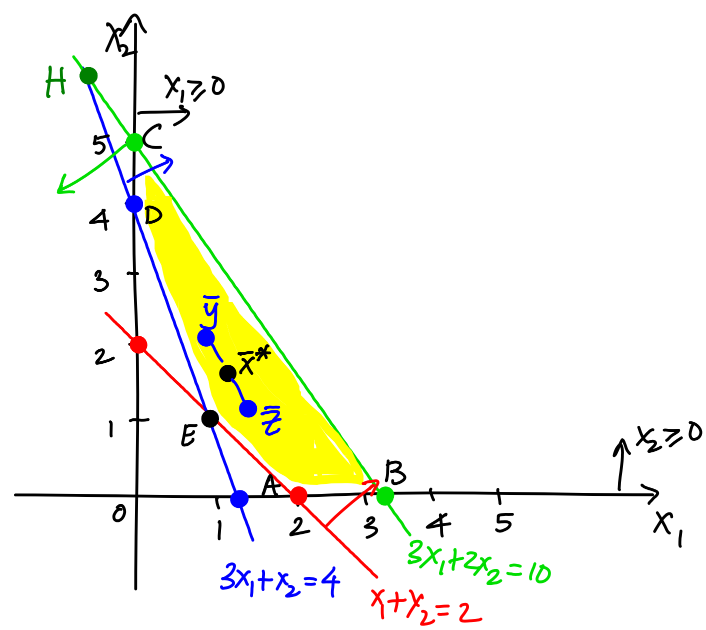
Thus, $\bar{a}_i^T \bar{y} = b_i \ \forall i \in I$, and $\bar{a}_i^T \bar{y} > b_i \ \forall i \notin I$. So $\bar{y} \in P$.

Ultimately, we choose the smaller of the two ϵ values (for $\bar{y}, \bar{z}$).

Similarly, $\bar{z} \in P$. But $\bar{x}^* = \frac{1}{2}(\bar{y} + \bar{z})$, i.e., $\bar{x}^*$ is a convex combination of $\bar{y}, \bar{z} \in P$, $\bar{y} \neq \bar{x}^*$, $\bar{z} \neq \bar{x}^*$.

So $\bar{x}^*$ is not an extreme point.

Intuitively, if $\bar{x}^*$ is feasible but is "in the interior" of P, we can find two other points close enough to $\bar{x}^*$, but still in P, such that $\bar{x}^*$ is the midpoint of the line segment connecting those two points.



3. (bfs $\Rightarrow$ vertex)

see the book (BT-ILO for details). $\square$

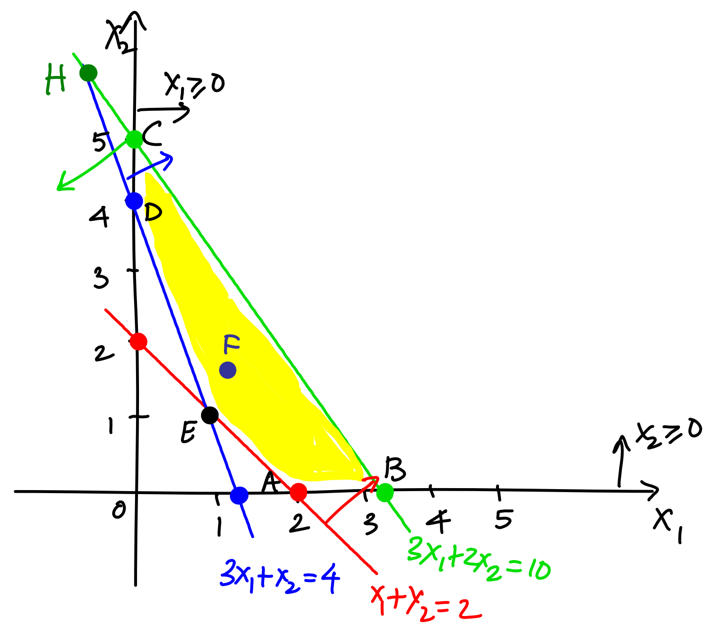
We are almost ready to describe how to go about finding the basic solutions as well as bfs's. But we introduce one more concept first.

Adjacent Basic Solutions

Def Two basic solutions are **adjacent** if there are $(n-1)$ LI constraints that are active at the same $(n-1)$ LI constraints basic solutions are both of them. If two adjacent basic solutions are both feasible, then they are **adjacent bfs's**.

e.g., A and B are adjacent bfs's, as $x_2 \geq 0$ is active at both of them (here, $n-1=1$).

H and C are adjacent basic solutions, as $3x_1+2x_2 \leq 10$ is active at both H and C.



Since H is not feasible, H and C are not adjacent bfs's.

As we will see later, we will move from one bfs to an adjacent bfs that improves the objective function value. We try to repeat this step until we cannot go to an adjacent bfs that improves the objective function value any more. Then we have an optimal bfs.

We describe the method to identify basic (feasible) solutions for polyhedra in standard form. Recall that the LP $\left\{\begin{array}{l}\min \bar{c}^T\bar{x} \\ \text{s.t. } A\bar{x}=\bar{b} \\ \quad\;\; \bar{x}\geqslant\bar{0}\end{array}\right\}$ is in standard form.

Polyhedra in standard form

Def $P=\{\bar{x}\in\mathbb{R}^n \mid A\bar{x}=\bar{b}, \bar{x}\geqslant\bar{0}\}$, $A\in\mathbb{R}^{m\times n}$, $\bar{b}\in\mathbb{R}^m$, and $\underbrace{\text{rank}(A)=m}_{\text{"full row-rank"}}$, $m\leq n$, is a polyhedron in **standard form**.

How do we find basic solutions of polyhedra in standard form?

$A\bar{x}=\bar{b}$ gives m LI constraints, as $\text{rank}(A)=m$.
Hence we need $n-m$ more binding constraints from $\bar{x}\geqslant\bar{0}$ (i.e., $x_j\geqslant 0$). Which $x_j\geqslant 0$ should we choose?

We cannot pick just any $(n-m)$ x_j's! The following theorem describes how to pick them.

Theorem 2.4 (BT-ILO)

$\bar{x}^*\in\mathbb{R}^n$ is a **basic solution** of P in standard form iff $A\bar{x}^*=\bar{b}$, and there exist indices $\underbrace{B(1),\ldots,B(m)}_{\text{numbers in } \{1,\ldots,n\}}$ such that

(a) the columns $A_{B(1)},\ldots,A_{B(m)}$ are LI, and

(b) if $i\neq B(1),\ldots,B(m)$, then $x_i^*=0$.

Based on this theorem, we could describe a procedure for constructing basic solutions of a polyhedron in standard form.

Procedure for constructing basic solutions

1. Choose m LI columns $A_{B(1)}, \dots, A_{B(m)}$.
2. Set $x_i = 0$ for $i \neq B(1), \dots, B(m)$,
3. Solve $A\bar{x} = \bar{b}$ for unknown $x_{B(1)}, \dots, x_{B(m)}$.

   If $\bar{x} \geq \bar{0}$, then it is a bfs.

We make the following observations.

* $x_{B(1)}, \dots, x_{B(m)}$ are basic variables.

* $A_{B(1)}, \dots, A_{B(m)}$ (the corresponding columns of A) are **basic columns**, they are LI and span $\mathbb{R}^m$.

* Let $B = \left[A_{B(1)} \; A_{B(2)} \; \cdots \; A_{B(m)} \right]$, $\bar{x}_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}$ $\quad \bar{x}_N \to$ remaining x_i's

  $\hookrightarrow$ basis matrix, is invertible. $\qquad$ we set $\bar{x}_N = \bar{0}$.

* We can "split" the original system $A\bar{x} = \bar{b}$ as follows:

  We write $A = [B \; N]$, where N are the (remaining) nonbasic columns,

  and $\bar{x} = \begin{bmatrix} \bar{x}_B \\ \bar{x}_N \end{bmatrix}$.

We will finish the discussion in the next lecture...

Procedure for constructing basic solutions (continued)

* With $A = [B\ N]$, $\bar{x} = \begin{bmatrix} \bar{x}_B \\ \bar{x}_N \end{bmatrix}$, $A\bar{x} = \bar{b}$ is equivalent to

$$[B\ N]\begin{bmatrix} \bar{x}_B \\ \bar{x}_N \end{bmatrix} = \bar{b}$$

$$\Rightarrow \quad B\bar{x}_B + N\bar{x}_N = \bar{b}$$

We then set $\bar{x}_N = \bar{0}$.

$$\Rightarrow \quad B\bar{x}_B = \bar{b} \quad \Rightarrow \quad \bar{x}_B = B^{-1}\bar{b} \quad \text{(recall, } B \text{ is invertible)}$$

$$\left.\begin{aligned} x_1 + x_2 - x_3 \qquad\quad &= 2 \\ 3x_1 + x_2 \qquad - x_4 \quad &= 4 \\ 3x_1 + 2x_2 \qquad\quad + x_5 &= 10 \end{aligned}\right\} \quad A = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 3 & 1 & 0 & -1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{bmatrix} \quad \bar{b} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}$$

$$x_j \geq 0 \ \forall j$$

Correspondence between corner pts & bfs's

We can use the above procedure to identify the bfs's in the standard form that correspond to the corner points marked out in the 2D graphical representation (of the original polyhedron). The maximum number of basic solutions possible in this case is $\binom{n}{m} = \binom{5}{3} = 10$. In this example, it indeed is true that each choice of $m=3$ columns of A produce a basic solution. In a general case, not all of the $\binom{n}{m}$ choices of m columns may produce a basis (i.e., an invertible matrix B). Hence, the number of basic solutions may be smaller than $\binom{n}{m}$.

Also, note that not all of these $\binom{n}{m}$ bases may lead to bfs's.
We can use MATLAB to demonstrate the correspondences between the bfs's (and basic solutions) in the standard form LP and the corner points in the 2D picture. See the course web page for the MATLAB files.

e.g., $B(1)=1, B(2)=2, B(3)=3$ gives $B = A(:,[1\ 2\ 3])$ — all rows, columns 1,2,3

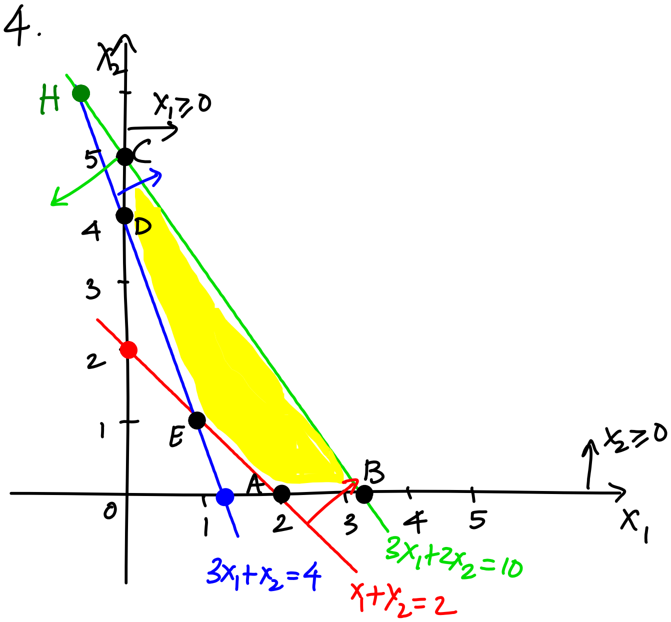
$$B = \begin{bmatrix} 1 & 1 & -1 \\ 3 & 1 & 0 \\ 3 & 2 & 0 \end{bmatrix}, \quad \det(B) = -3 \text{ (hence } B \text{ is invertible).}$$

We set $x_4 = x_5 = 0$ $(\bar{x}_N = \bar{0})$, and solve for $\bar{x}_B = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = B^{-1}\bar{b}$, to get

$$\bar{x}_B = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 6 \\ 10/3 \end{bmatrix}.$$

This point is not feasible (as all entries are not ≥ 0); hence it is not a bfs.

This vector correponds to the vertex $H(-2/3, 6)$ in the picture.



As another example, consider $B(1)=2, B(2)=3, B(3)=4$.

$$B = A(:,[2\ 3\ 4]) = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 2 & 0 & 0 \end{bmatrix}. \quad \det(B) = 2,$$

hence B is invertible. We get

$$\bar{x}_B = \begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix} = B^{-1}\bar{b} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}.$$

Thus,

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 3 \\ 1 \\ 0 \end{bmatrix} \text{ is a bfs.}$$

It Corresponds to the vertex $C(0,5)$.