LECTURE 4: SIMPLEX METHOD

1. Simplex method
2. Phase one method
3. Big M method

What have we learned so far?

- Consider a standard form LP (primal problem)

$$
\begin{array}{lll}
 & \text{Min} & \mathbf{c}^T\mathbf{x} \\
\text{(LP)} & \text{s. t.} & \mathbf{A}\mathbf{x} = \mathbf{b} \\
 & & \mathbf{x} \geq 0
\end{array}
$$

1. If its feasible domain P is nonempty, it has at least one vertex (extreme point). -- from Resolution Theorem
2. If P is nonempty and the objective value z is not unbounded, then (LP) attains optimal at (at least) one vertex (extreme point). -- from Fundamental Theorem
3. P has finitely many vertices (extreme points). -- C(*n*, *m*)
4. Vertices can be generated algebraically as bsf's.

Implications

- When $C(n, m)$ is small, we can enumerate through all bsf's (vertices) to find the optimal one as our optimal solution. -- Enumeration Method

- When $C(n, m)$ becomes large, we need a systematic and efficient way to do this job. -- Simplex Method

Basic idea of the simplex method

- Conceived by Prof. George B. Dantzig in 1947.
- Basic idea:

Phase I:

Step 1: (Starting)

Find an initial extreme point (ep) or declare P is null.

Phase II:

Step 2: (Checking optimality)

If the current ep is optimal, STOP!

Step 3: (Pivoting)

Move to a better ep.

Return to Step 2.

Observations

- Going back to Step 2 from Step 3 is called an iteration.

- If we don't repeat using the same extreme points, the algorithm will always terminate in a finite number of iterations. -- a finite algorithm

- How to efficiently generate better extreme points?

  -- basic feasible solutions

What else have we learned?

- A point **x** in P is an extreme point if and only if **x** is a basic feasible solution corresponding to some basis B.

- There exists at most *C(n, m)* basic feasible solutions. When rank(A) = $m \leq n$, a bfs is obtained by setting

$$\mathbf{A} = \left[\, \mathbf{B} \mid \mathbf{N} \,\right]$$

$$\mathbf{x} = \left[\begin{array}{c} \mathbf{x}_B \\ \hline \mathbf{x}_N \end{array} \right]$$

and set $\mathbf{x}_N = \mathbf{0}$ to calculate $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}.$

Baseline of the simplex method

Phase I:

Step 1: (Starting)

Find an initial basic feasible solution (bfs), or declare P is null.

Phase II:

Step 2: (Checking optimality)

If the current bfs is optimal, STOP!

Step 3: (Pivoting)

Move to a better bfs.

Return to Step 2.

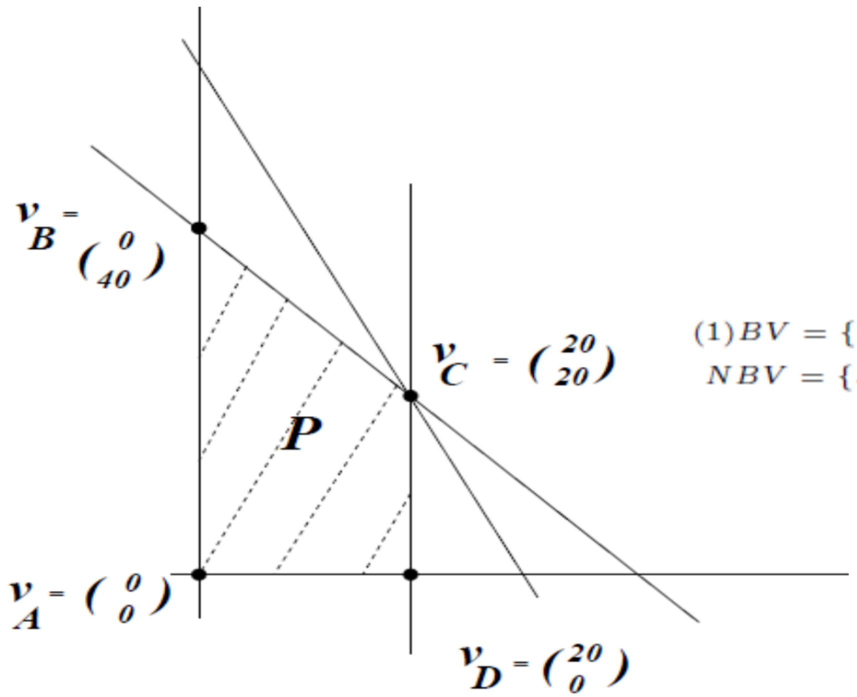
Challenge

- When we move from one bfs to another bfs, do we really move from one extreme point to another extreme point?

- If not, we may be trapped into a loop!

Example

$$\begin{cases} x_1 + x_2 & \leq 40 \\ 2x_1 + x_2 & \leq 60 \\ x_1 & \leq 20 \\ x_1, x_2 & \geq 0. \end{cases} \Leftrightarrow \begin{cases} x_1 + x_2 + s_1 & & & = 40 \\ 2x_1 + x_2 & + s_2 & & = 60 \\ x_1 & & + s_3 & = 20 \\ x_1, x_2, s_1, & s_2, & s_3 & \geq 0. \end{cases}$$



(1) $BV = \{x_1, x_2, s_1\}$ $NBV = \{s_2, s_3\}$

(2) $BV = \{x_1, x_2, s_2\}$ $NBV = \{s_1, s_3\}$

(3) $BV = \{x_1, x_2, s_3\}$ $NBV = \{s_1, s_2\}$

Observations

- If an ep is determined by a bfs with exactly *m* positive basic variables and *n* - *m* zero non-basic variables, then the correspondence is one-to-one.

  -- a nondegenerate bfs

- Only when there exists at least one basic variable becoming 0, then the ep may correspond to more than one bfs.

  -- a degenerate bfs

- Terminology:

  An LP is nondegenerate if every bfs is nondegenerate.

Nondegeneracy

- Property 1: If a bfs **x** is nondegenerate, then **x** is uniquely determined by *n* hyperplanes.
- Why? *n* hyperplanes? Where are they?
- Remember that

$$\mathbf{A} = [\,\mathbf{B} \mid \mathbf{N}\,]$$

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} \mathbf{B}^{-1}\mathbf{b} \\ \mathbf{0} \end{bmatrix}$$

- Let

$$\mathrm{M} = \begin{bmatrix} \mathbf{B} & \mathbf{N} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

Then M is nonsingular and

$$\mathrm{M}\mathbf{x} = \begin{bmatrix} \mathbf{B} & \mathbf{N} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}.$$

Hence $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \mathrm{M}^{-1} \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}$ is uniquely determined by n linearly independent hyperplanes.

Fundamental matrix

- Question: $\mathbf{M}^{-1} = ?$

- Answer:

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{B}^{-1} & -\mathbf{B}^{-1}\mathbf{N} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

- Hence, $\mathbf{M}^{-1}$ is known when $\mathbf{B}^{-1}$ is known!

- We call $\mathbf{M}^{-1}$ (or M) the fundamental matrix of LP.

Nondegeneracy

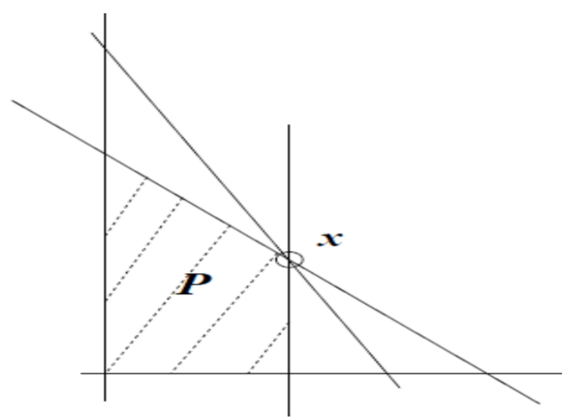
- Property 2: If a bfs **x** is degenerate, then **x** is over-determined by more than n hyperplanes.
- Why? Other than the *n* hyperplanes of

$$\begin{bmatrix} \mathrm{B} & \mathrm{N} \\ 0 & \mathrm{I} \end{bmatrix} \begin{bmatrix} \mathrm{x}_B \\ \mathrm{x}_N \end{bmatrix} = \begin{bmatrix} \mathrm{b} \\ 0 \end{bmatrix}$$

There exists at least one basic variable such that

$$x_i = 0$$

which is another hyperplane.



Nondegeneracy

- Property 3:

  For a degenerate bfs **x** with p (< m) positive components, we may have up to

$$\begin{pmatrix} n-p \\ n-m \end{pmatrix} = \frac{(n-p)!}{(n-m)!(m-p)!}$$

different bfs corresponding to the same extreme point.

Simplex method under nondegeneracy

- Basic idea:

  Moving from one bfs (ep) to another bfs (ep) with a simple pivoting scheme.

- Instead of considering all bfs (ep) at the same time, just consider some neighboring bfs (ep).

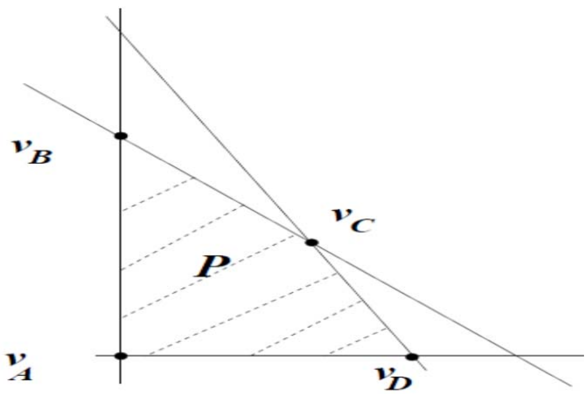
- Definition:

  Two basic feasible solutions are adjacent if they have $m - 1$ basic variables (not their values) in common.

Observations

- Under nondegeneracy, every basic feasible solution (extreme point) has exactly *n* - *m* adjacent neighbors.

- For a bfs, each adjacent bfs can be reached by increasing one nonbaisc variable from zero to positive and decreasing one basic variable from positive to zero. – Pivoting



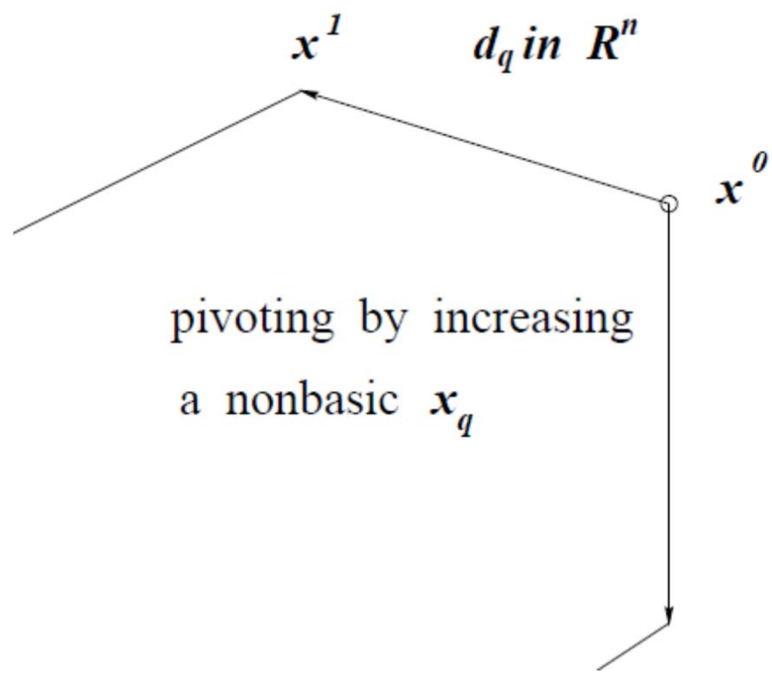
$$v_A = \begin{bmatrix} 0 \\ 0 \\ 40 \\ 60 \end{bmatrix}, v_B = \begin{bmatrix} 0 \\ 40 \\ 0 \\ 20 \end{bmatrix},$$

$$v_C = \begin{bmatrix} 20 \\ 20 \\ 0 \\ 0 \end{bmatrix}, v_D = \begin{bmatrix} 30 \\ 0 \\ 10 \\ 0 \end{bmatrix}.$$

Pivoting

- Concept:

  One nonbasic variable enters (from 0 to positive) the basis and one basic variable leaves the basis (from positive to 0).



$\mathbf{x}^1 = \mathbf{x}^0 + \lambda \mathbf{d}_q$ for $\lambda > 0$.

<u>edge direction</u> <u>step length</u>

Who and where are my neighbors?

- A current ep moves to a neighboring ep by walking on the boundary edge of P.
- There are *n-m* neighbors of the current ep.
- There should be *n-m* edge directions leading to the adjacent extreme points, corresponding to the increase of each nonbasic variable (nbv).
- Let the edge direction $\mathbf{d}_q \in \mathbf{R}^n$ corresponding the increasing of a nonbasic variable $\mathbf{x}_q$.

- Where are these edge directions?

Fundamental matrix and edge direction

- Notice that the fundamental matrix

$$M^{-1} = \begin{bmatrix} B^{-1} & -B^{-1}N \\ 0 & I \end{bmatrix}$$

has *n-m* columns in the part of $\begin{bmatrix} -B^{-1}N \\ I \end{bmatrix}$.

- Could they be the edge directions?

Conjecture

$\mathbf{d}_q$ is in the column in $\mathbf{M}^{-1}$ corresponding to $\mathbf{x}_q$, *i.e.*

$$\mathbf{d}_q = \begin{pmatrix} \underline{-\mathbf{B}^{-1}\mathbf{A}_q} \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

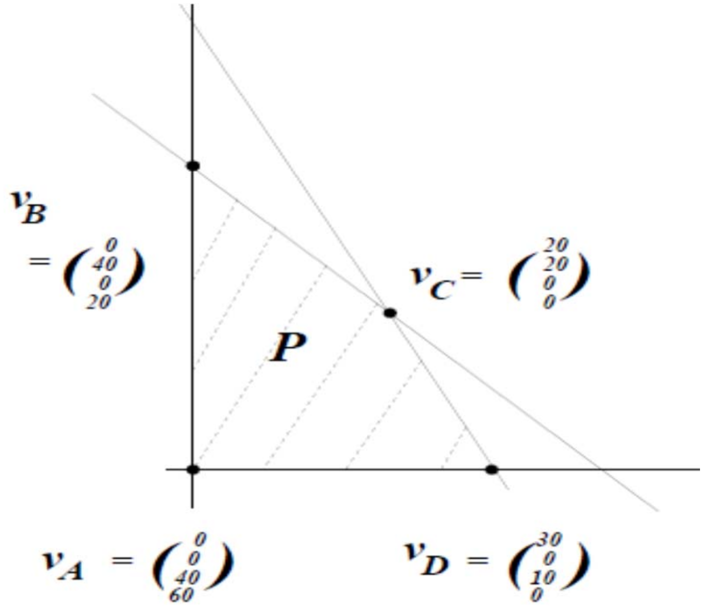
where

$$\mathbf{A} = \left(\mathbf{A}_1 | \mathbf{A}_2 | \cdots | \mathbf{A}_n\right).$$

Example

$$
\begin{cases}
x_1 + x_2 + x_3 & = 40 \\
2x_1 + x_2 \qquad + x_4 & = 60 \\
x_1, x_2, x_3, x_4 & \geq 0.
\end{cases}
$$

$$
\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}.
$$



At v_A, $BV = \{x_3, x_4\}, NBV = \{x_1, x_2\}$

$$
\mathrm{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathrm{N} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}.
$$

$$
\mathrm{M} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathrm{B}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

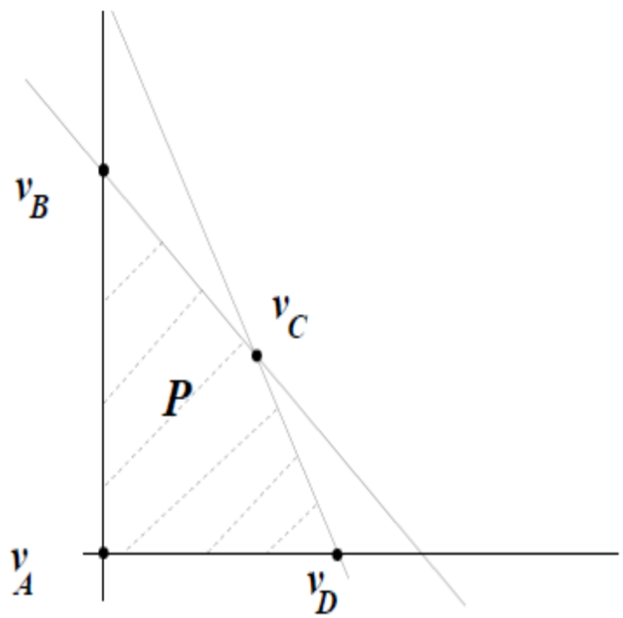
$$
\mathbf{M}^{-1} = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

Example - continue

- From v_A to v_B ,

$$\begin{bmatrix} 0 \\ 40 \\ 0 \\ 20 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 40 \\ 60 \end{bmatrix} = 40 \begin{bmatrix} 0 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

- From v_A to v_D ,
- $$\begin{bmatrix} 30 \\ 0 \\ 10 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 40 \\ 60 \end{bmatrix} = 30 \begin{bmatrix} 1 \\ 0 \\ -1 \\ -2 \end{bmatrix}$$



$$\mathrm{M}^{-1} = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$v_A = \begin{bmatrix} 0 \\ 0 \\ 40 \\ 60 \end{bmatrix}, v_B = \begin{bmatrix} 0 \\ 40 \\ 0 \\ 20 \end{bmatrix}, v_c = \begin{bmatrix} 20 \\ 20 \\ 0 \\ 0 \end{bmatrix}, v_D = \begin{bmatrix} 30 \\ 0 \\ 10 \\ 0 \end{bmatrix}.$$

General case

In general, for $\lambda \geq 0$

$$\mathbf{x}(\lambda) = \mathbf{x} + \lambda \mathbf{d}_q = \begin{pmatrix} \mathbf{x}_B \\ \hline \mathbf{x}_N \end{pmatrix} + \lambda \begin{pmatrix} -\mathbf{B}^{-1}\mathbf{A}_q \\ \hline 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

(1) For nonbasic variables, all are kept at zero, except $\mathbf{x}_q$ increases by λ. *i.e.*

$$\mathbf{x}_N(\lambda) = \mathbf{x}_N + \lambda \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

(2) For basic variables, since $\mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$, thus $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N$,
when x_q increases by λ and the rest n.b.v are kept at 0, then $\mathbf{x}_B(\lambda) = \mathbf{B}^{-1}\mathbf{b} - \lambda\mathbf{B}^{-1}\mathbf{A}_q$,

Hence

$$\mathbf{d}_q = \begin{pmatrix} -\mathbf{B}^{-1}\mathbf{A}_q \\ \hline e_q \end{pmatrix}$$

Question

- Is an edge direction $\mathbf{d}_q$ always a feasible direction?
- That means for a small enough step length $\lambda > 0$, we need

$$\mathbf{x}(\lambda) = \mathbf{x} + \lambda \mathbf{d}_q \in P.$$

- Must show that $\mathbf{A}\mathbf{x}(\lambda) = \mathbf{b}$ and $\mathbf{x}(\lambda) \geq 0$.
- Equivalently, we need to show that $\mathbf{A}\mathbf{d}_q = \mathbf{0}$ and $\mathbf{x}(\lambda) \geq 0$.

Answer - I

- Yes, every edge direction is a feasible direction when the problem is nondegenerate.

- Proof:

(1) $\mathbf{A}\mathbf{d}_q = \mathbf{0}$ can be derived from $\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$.

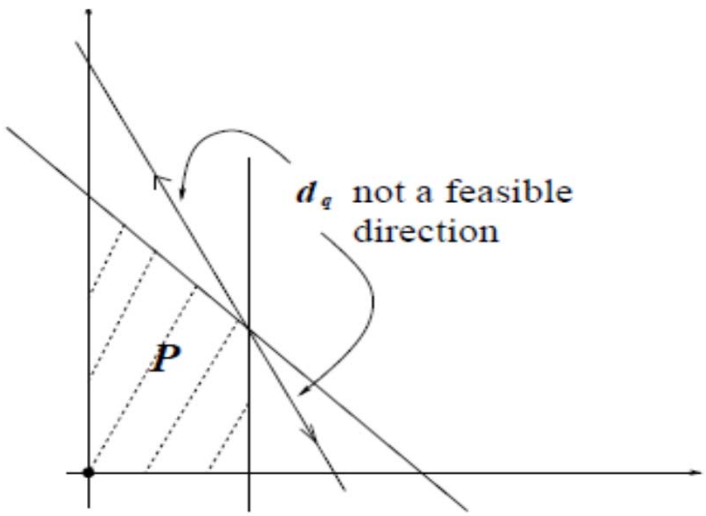
(2) For nondegenerate case,

$$\mathbf{x}(\lambda) = \mathbf{x} + \lambda \begin{pmatrix} -\mathbf{B}^{-1}\mathbf{A}_q \\ \hline e_q \end{pmatrix}$$

Hence $\mathbf{x}(\lambda) \geq 0$ when λ is small enough.
i.e., under nondegeneracy,
an edge direction $\mathbf{d}_q$ is a feasible direction!

Answer - II

- No, an edge direction is not necessarily a feasible direction when the problem is degenerate.
- Proof:



$$\mathbf{x}(\lambda) = \mathbf{x} + \lambda \begin{pmatrix} -\mathbf{B}^{-1}\mathbf{A}_q \\ \hline e_q \end{pmatrix}$$

say $x_i = 0$, no mater how small λ is, $\mathbf{x}_i(\lambda) < 0$!!

Which neighbor is a good one?

- If current bsf is not optimal, which neighboring bsf is a better one?
- That means, along which edge direction to move?

  or, which nonbasic variable is a good candidate to pivot in?
- Observation:

$$
\begin{aligned}
\mathbf{z}(\mathbf{x}(\lambda)) &= \mathbf{c}^T\mathbf{x}(\lambda) \\
&= \mathbf{c}^T(\mathbf{x} + \lambda \mathbf{d}_q) \\
&= \mathbf{z}(\mathbf{x}) + \lambda[\mathbf{c}_B^T | \mathbf{c}_N^T] \left(\frac{-\mathbf{B}^{-1}\mathbf{A}_q}{e_q} \right) \\
&= \mathbf{z}(\mathbf{x}) + \lambda[c_q - \mathbf{c}_B^T \mathbf{B}^{-1}\mathbf{A}_q] \\
&= \mathbf{z}(\mathbf{x}) + \lambda r_q
\end{aligned}
$$

If $r_q = \mathbf{c}^T\mathbf{d}_q = c_q - \mathbf{c}_B^T B^{-1}\mathbf{A}_q < 0$, then d_q is a good direction!

Reduced cost

- Definition: The quantity of

$$r_q = \mathbf{c}^T \mathbf{d}_q = c_q - \mathbf{c}_B^T B^{-1} \mathbf{A}_q$$

is called a reduced cost with respect to the variable $\mathbf{x}_q$.

Theorem:

If $\mathbf{x} = \left(\frac{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}} \right)$ is a bfs with $\mathbf{B}$ and $r_q < 0$ for some n.b.v. x_q, then $\mathbf{d}_q = \left(\frac{-\mathbf{B}^{-1}\mathbf{A}_q}{e_q} \right) \in \mathbf{R}^n$ leads to an improved objective value.

Observations

- Observation 1:

For a basic variable $x_q \in \mathbf{B}$,

$$\begin{aligned} r_q &= c_q - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{A}_q \\ &= c_q - c_q \\ &= 0. \end{aligned}$$

- Observation 2:

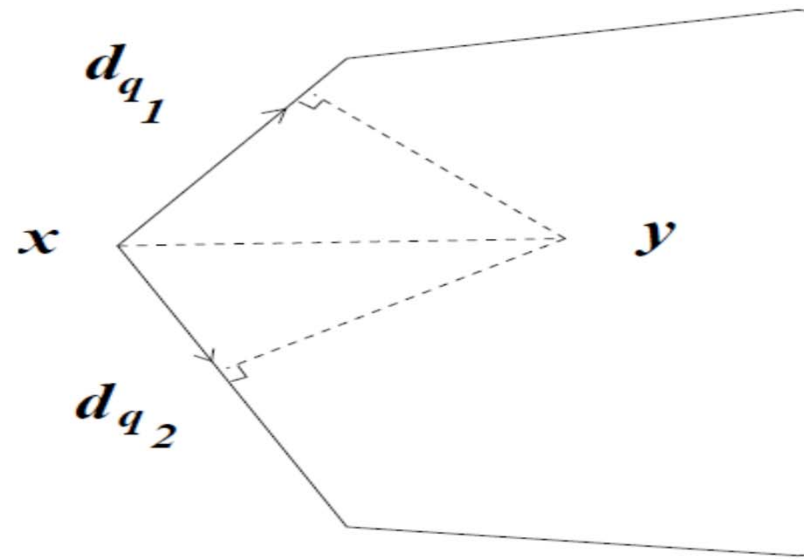
Any $\mathbf{d}_q$ (x_q n.b.v.) with $r_q < 0$ will do for the simplex method. The one with most reduced cost can be found by

$$\min_{j:\text{nonbasic}} \left\{ \frac{\mathbf{c}^T \mathbf{d}_j}{\|\mathbf{d}_j\|} \right\}.$$

Optimality check by reduced cost

- Question:

  If $r_q \geq 0, \ \forall$ n.b.v. x_q, is the current bfs optimal?



- Guess:

$\forall \mathbf{y} \in P,$

$\mathbf{y} = \mathbf{x} + y_{q_1}\mathbf{d}_{q_1} + y_{q_2}\mathbf{d}_{q_2}, \quad y_{q_1}, y_{q_2} \geq 0$

Hence

$\mathbf{c}^T\mathbf{y} = \mathbf{c}^T\mathbf{x} + y_{q_1}\mathbf{c}^T\mathbf{d}_{q_1} + y_{q_2}\mathbf{c}^T\mathbf{d}_{q_2} \geq \mathbf{c}^T\mathbf{x} + 0 = \mathbf{c}^T\mathbf{x}$

Optimality condition

- Theorem: Given a bfs $\mathbf{x}^0 = \begin{pmatrix} \mathbf{B}^{-1}\mathbf{b} \\ \hline \mathbf{0} \end{pmatrix}$ with basis $\mathbf{B}$, if $r_q \geq 0, \ \forall$ n.b.v x_q, then $\mathbf{x}$ is optimal.
- Proof:

$$\forall \ \mathbf{y} \in P, \ \mathbf{y} = \begin{pmatrix} \mathbf{y}_B \\ \hline \mathbf{y}_N \end{pmatrix} \geq 0, \ \mathbf{A}\mathbf{y} = \mathbf{b}$$

Note $\mathbf{x}_N^0 = 0$ and $\mathbf{A}\mathbf{x}^0 = \mathbf{b}$

Thus

$$\begin{aligned} \mathbf{M}(\mathbf{y} - \mathbf{x}^0) &= \begin{bmatrix} \mathbf{B} & \mathbf{N} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{y}_B - \mathbf{x}_B^0 \\ \hline \mathbf{y}_N \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{b} - \mathbf{b} \\ \hline \mathbf{y}_N \end{bmatrix} \\ &= \begin{bmatrix} \underline{0} \\ \mathbf{y}_N \end{bmatrix}. \end{aligned}$$

$$\begin{aligned} \mathbf{y} - \mathbf{x}^0 &= \mathbf{M}^{-1} \begin{bmatrix} 0 \\ \mathbf{y}_N \end{bmatrix} \quad \text{with } \mathbf{y}_N = \begin{bmatrix} \vdots \\ y_q \\ \vdots \end{bmatrix} \geq 0 \\ &= \begin{bmatrix} \mathbf{B}^{-1} & -\mathbf{B}^{-1}\mathbf{N} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{y}_N \end{bmatrix} \\ &= \begin{bmatrix} -\mathbf{B}^{-1}\mathbf{N} \\ \hline \mathbf{I} \end{bmatrix} \mathbf{y}_N \\ &= \textstyle\sum_{q \in N} y_q \mathbf{d}_q \end{aligned}$$

$$i.e., \ \mathbf{y} = \mathbf{x}^0 + \textstyle\sum_{q \in N} y_q d_q$$

$$\text{Hence } \mathbf{c}^T\mathbf{y} \geq \mathbf{c}^T\mathbf{x}^0, \ \forall \ \mathbf{y} \in P.$$

Uniqueness of optimal solution

- Corollary 1: If the reduced cost $r_q > 0$ for every nbv x_q , then the bfs **x** is the unique optimal solution.

- Corollary 2: If **x** is an optimal bfs with some

$$r_{q_1}, r_{q_2}, \ldots, r_{q_k} = 0,$$

then any point $\mathbf{y} \in P$ such that $\mathbf{y} = \mathbf{x} + \sum_{i=1}^{k} y_{q_i} d_{q_i}$ is also optimal.

Question

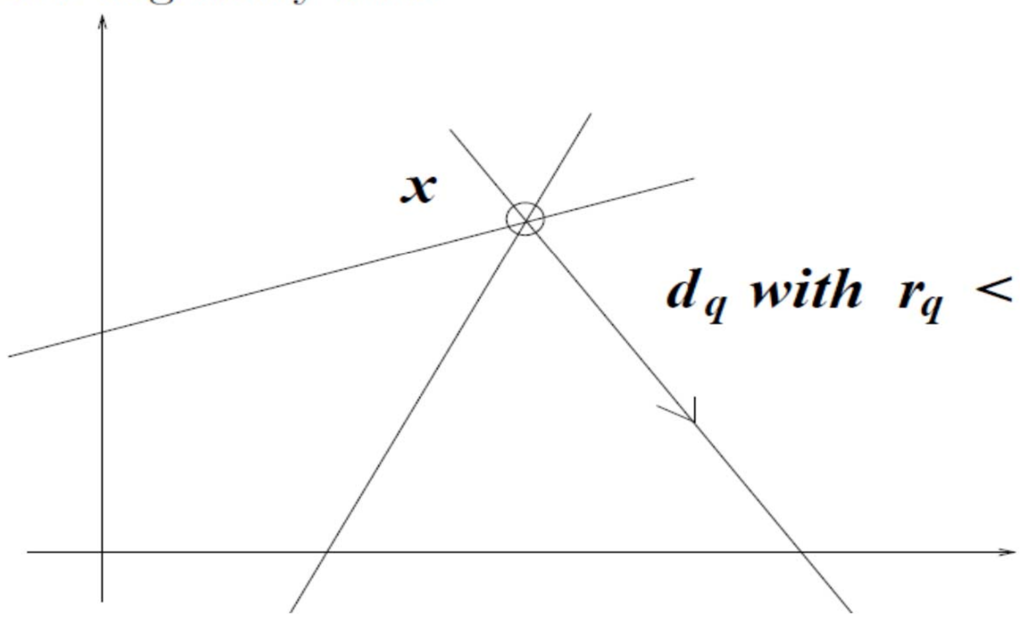
- Is the converse statement of the theorem true? i.e.,

  "If a bfs $\mathbf{x}$ is optimal, then $r_q \geq 0, \ \forall$ n.b.v x_q."

- Answer:

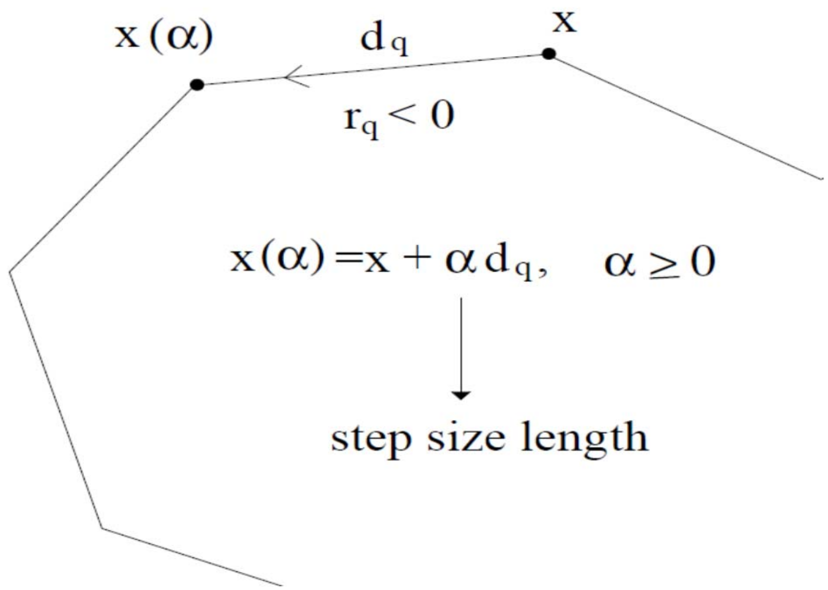
  True only for the nondergeneracy case.

  For degeneray case:



How far is my good neighbor?

- Basic concept:



- Question:

How far should we go such that $\mathbf{x}(\alpha)$ is an adjacent bfs?

Analysis of step length

- We have $\mathbf{x}(\alpha) = \mathbf{x} + \alpha d_q, \;\; \alpha > 0.$

  with $r_q = \mathbf{c}^T \mathbf{d}_q = c_q - \mathbf{c}_B^T B^{-1} \mathbf{A}_q < 0$
- Remember that $\mathbf{A}\mathbf{d}_q = \mathbf{0}$, thus $\mathbf{A}\mathbf{x}(\alpha) = \mathbf{A}\mathbf{x} = \mathbf{b}.$

- Case 1: If $\mathbf{d}_q \geq \mathbf{0}$, then $\mathbf{x}(\alpha) \geq \mathbf{0}, \;\; \forall \alpha \geq 0.$

  Hence $\mathbf{x}(\alpha) \in P, \; \forall \; \alpha \geq 0$ and

  $\mathbf{c}^T \mathbf{x}(\alpha) = \mathbf{c}^T \mathbf{x} + \alpha \mathbf{c}^T \mathbf{d}_q \longrightarrow -\infty$, as $\alpha \longrightarrow +\infty.$
- Theorem:

If $\mathbf{x}$ is a bfs with $\mathbf{d}_q \geq \mathbf{0}$ and $r_q < 0$, for some n.b.v. x_q, then the LP is unbounded.

Note: $\mathrm{d}_q = \begin{pmatrix} -\mathrm{B}^{-1}\mathrm{A}_q \\ \hline e_q \end{pmatrix}$. Define $\mathrm{w} \triangleq \mathrm{B}^{-1}\mathrm{A}_q$,

then

$\mathrm{d}_q \geq 0 \iff \mathrm{w} \leq 0$

Analysis - continue

- Case 2: $\mathbf{d}_q$ has at least one component < 0.
  To keep $\mathbf{x}(\alpha) \geq \mathbf{0}$, we have to choose

$$\alpha = \min_{i:\text{basic}} \left\{ \frac{x_i}{-d_{qi}} \mid d_{qi} < 0 \right\}.$$

- Observations:

<u>Note1:</u> $d_{qi} < 0$ can only happen for basic variables $(x_i \in \mathrm{B})$.

<u>Note2:</u> α is determined by the <u>*Minimum ratio test*</u>.

<u>Note3:</u> Under nondegeneracy,

$x_i > 0$ for b.v. x_i

$\Rightarrow \alpha > 0$

$\Rightarrow \mathbf{x}(\alpha)$ is a different extreme point.

For degenerate bfs, it is possible $x_i = 0$, then

$\alpha = 0$

$\Rightarrow \mathbf{x}(\alpha)$ stays at the same extreme point.

Step length by minimum ratio test

- Theorem: If **x** is a bfs, then $\mathbf{x}(\alpha) = \mathbf{x} + \alpha \mathbf{d}_q$ is an adjacent bfs, if the step length α is determined by the minimum ratio test.

- Note that this $\mathbf{x}(\alpha)$ indeed moves to an adjacent extreme point, when the bfs **x** is nondegenerate.

Key steps of Simplex Method

Step1: Find a bfs $\mathbf{x}$ with $\mathbf{A} = [\mathbf{B}|\mathbf{N}]$.

Step2: Check for n.b.v's

$$r_q = \mathbf{c}^T \mathbf{d}_q (= c_q - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{A}_q).$$

If $r_q \geq 0, \ \forall$ nonbasic x_q, then the current bfs is optimal.
Otherwise, pick one $r_q < 0$. Go to next step.

Step3: If $\mathbf{d}_q \geq 0$, then LP is unbounded.
Otherwise, find

$$\alpha = \min_{i:\text{basic}} \left\{ \frac{x_i}{-d_{q_i}} \mid d_{q_i} < 0 \right\}.$$

Then $\mathbf{x} \leftarrow \mathbf{x} + \alpha \mathbf{d}_q$ is a new bfs.

Update $\mathbf{B}$ and $\mathbf{N}$. Go to Step 2.

Main result

- Theorem: Under the nondegeneracy assumption, simplex method terminates in a finite number of iterations with either an unbounded minimum, or an optimal solution to a given LP.

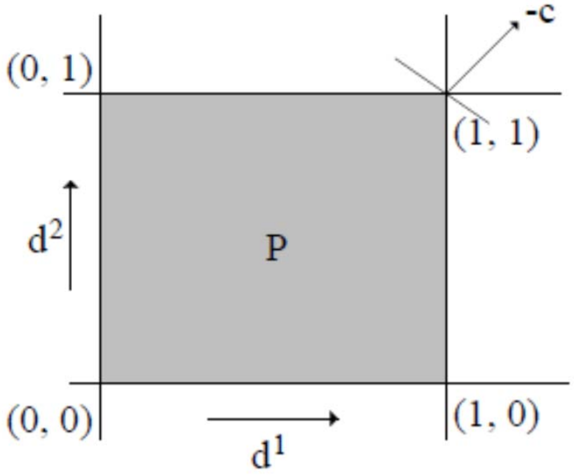
Example

Min $-x_1 - x_2$

s.t. $x_1 \leq 1$

$x_2 \leq 1$

$x_1, x_2 \geq 0$



$$
\begin{aligned}
\min \quad & -x_1 - x_2 \\
& x_1 \quad + x_3 = 1 \\
& \quad x_2 \quad + x_4 = 1 \\
& x_1, x_2, x_3, x_4 \geq 0
\end{aligned}
$$

Example – first iteration

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ & x_1 \qquad + x_3 = 1 \\ & \quad x_2 \qquad + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

bfs#1: b.v. $\{x_3, x_4\}$, n.b.v. $\{x_1, x_2\}$

$$\mathbf{x} = \begin{bmatrix} 1 \\ \underline{1} \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{A} = [\mathbf{B}|\mathbf{N}] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\mathrm{B}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \qquad \mathrm{B}^{-1}\mathrm{N} = \mathrm{N} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathrm{M}^{-1} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Example – check reduced cost for optimality

$$r_1 = \mathbf{c}^T \mathbf{d}^1 = [0\ 0 - 1 - 1] \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = -1 < 0$$

$$r_2 = \mathbf{c}^T \mathbf{d}^2 = [0\ 0\ \ -1\ \ -1] \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = -1 < 0$$

Example – moving to better neighbor

Pick $\mathbf{d}^1 (\not\geq 0)$, so x_1 enters the basis.

$$\alpha = \min_i \{ \frac{x_i}{-d_i^1} \mid d_i^1 < 0 \} = \frac{x_3}{-d_{x_3}^1} = -\frac{1}{-1} = 1$$

$$\mathbf{x} \longleftarrow \mathbf{x} + \alpha \mathbf{d}^1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

So, x_3 leaves the basis.

bfs#2: b.v. $\{x_1, x_4\}$, n.b.v. $\{x_3, x_2\}$

Example – second iteration

bfs#2: b.v. $\{x_1, x_4\}$, n.b.v. $\{x_3, x_2\}$

$$\mathbf{A} = [\mathbf{B}|\mathbf{N}] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\mathbf{M}^{-1} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Example – optimality check

$$r_3 = \mathbf{c}^T \mathbf{d}^3 = [-1\ 0\ 0\ \ -1] \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = 1 > 0$$

$$r_2 = \mathbf{c}^T \mathbf{d}^2 = [-1\ 0\ 0\ \ -1] \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = -1 < 0$$

Example – move to a better neighbor

Pick $\mathbf{d}^2 (\not\geq 0)$, so x_2 enters the basis.

$$\alpha = \frac{x_4}{-d^2_{x_4}} = -\frac{1}{-1} = 1$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

So, x_4 leaves the basis.

bfs#3: b.v. $\{x_1, x_2\}$, n.b.v. $\{x_3, x_4\}$

Example – third iteration

bfs#3: b.v. $\{x_1, x_2\}$, n.b.v. $\{x_3, x_4\}$

$$\mathbf{A} = [\mathbf{B}|\mathbf{N}] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad \mathbf{M}^{-1} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$r_3 = \mathbf{c}^T \mathbf{d}^3 = [\,-1 \quad -1 \;\; 0 \;\; 0] \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = 1 > 0$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{(optimal!)}$$

$$r_4 = \mathbf{c}^T \mathbf{d}^4 = [\,-1 \quad -1 \;\; 0 \;\; 0] \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = 1 > 0$$

How to start the simplex method ?

- How to get an initial basic feasible solution?

  -- eye inspection

  -- randomly generate (test of luck)

  -- systematic approach

  1. Two-phase method (Phase I problem)
  2. big-M method

Two-phase method

- Step 1. Make the right hand side vector nonnegative:

$$\begin{array}{lll} & \text{Min} & \mathbf{c}^T\mathbf{x} \\ \text{(LP)} & \text{s. t.} & \mathbf{A}\mathbf{x} = \mathbf{b}(\geq 0) \\ & & \mathbf{x} \geq 0 \end{array}$$

- Step 2: Add *m* artificial variables for Phase 1 problem:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \qquad \begin{array}{lll} & \text{Min} & \sum_{i=1}^{m} u_i \\ \text{(PhI)} & \text{s. t.} & \mathbf{A}\mathbf{x} + \mathbf{I}\mathbf{u} = \mathbf{b}(\geq 0) \\ & & \mathbf{x}, \mathbf{u} \geq 0 \end{array}$$

What're special about Phase I problem ?

1. $\mathbf{u} = b, \mathbf{x} = 0$ is a bfs of (PhI).
2. (PhI) is bounded below by 0.
3. (LP) is feasible if and only if $\mathbf{z}^*_{PhI} = 0$
4. Under nondegeneracy, if $\mathbf{z}^*_{PhI} = 0$, then an optimal solution of (PhI) is a bfs of (LP).

How about degenerate case ?

5. If $\mathbf{z}^*_{PhI} = 0$ at an optimal bfs which is degenerate with at least one artificial variable u_i in the basis.

   Suppose that $u_i = 0$ is the k-th basic variable in the current basis, then

   (1) if $e_k^T \mathbf{B}^{-1} \mathbf{A}_q \neq 0$ for a n.b.v. x_q, then u_i can be replaced by x_q to form a starting basis.

   (2) if $e_k^T \mathbf{B}^{-1} \mathbf{A}_q = 0$, $\forall$ n.b.v. x_q, then the k-th row of $\mathbf{A}\mathbf{x} = \mathbf{b}$ is redundant. We remove it and start again.

Implication

- Finding a starting basic feasible solution is as difficult as finding an optimal solution with a given basic feasible solution.

Big-M method

- Add a big penalty $M > 0$ to each artificial variable.
- Combine phase I problem with the original problem to consider a big-M problem:

$$
\begin{aligned}
\text{Min} \quad & \textstyle\sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} M u_i \\
\text{s. t.} \quad & \mathbf{A}\mathbf{x} + I\mathbf{u} = \mathbf{b} (\geq 0) \\
& \mathbf{x}, \mathbf{u} \geq 0
\end{aligned}
$$

What're special about big-M problem

$$\begin{aligned} \text{Min} \quad & \textstyle\sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} M u_i \\ \text{s. t.} \quad & \mathbf{A}\mathbf{x} + I\mathbf{u} = \mathbf{b}(\geq 0) \\ & \mathbf{x}, \mathbf{u} \geq 0 \end{aligned}$$

1. $\mathbf{x} = 0, \mathbf{u} = b$, is a bfs.
2. $\mathbf{z}^*$ can be finite at an optimal solution $\binom{\mathbf{x}^*}{\mathbf{u}^*}$ or unbounded below.
3. Suppose $\mathbf{z}^*$ is finite at $\binom{\mathbf{x}^*}{\mathbf{u}^*}$. If

   (i) $u^* = 0$,

   then $\forall$ $\mathbf{x}$ feasible to (LP), $\binom{\mathbf{x}}{0}$ is feas

   (big-M). Thus

   $$\mathbf{c}^T\mathbf{x} + M \times 0 \geq \mathbf{c}^T\mathbf{x}^* + M \sum_{i=1}^{m}$$

   $$\mathbf{c}^T\mathbf{x} \geq \mathbf{c}^T\mathbf{x}^* + 0$$

   i.e., $\mathbf{x}^*$ is optimal to (LP).

   (ii) $u^* \neq 0$,

   then for $\mathbf{x}$ feasible to (LP), $\binom{\mathbf{x}}{0}$ is feasible to (big-M) and

   $$\mathbf{c}^T\mathbf{x} + M \times 0 \geq \mathbf{c}^T\mathbf{x}^* + M \sum_{i=1}^{m} u_i^*$$

   But this is impossible for M is large enough. Hence $P = \emptyset$.

4. If $\mathbf{z}^* \to -\infty$ with all $u_i = 0$, then (LP) is unbounded below. Otherwise, $P = \emptyset$.

Big-M problem

- Question: How big should M be ?

- Example:

$$\begin{aligned} &\text{Min} && x_1 \\ &(\text{LP})\ \text{s. t.} && \epsilon x_1 - x_2 - x_3 = \epsilon\ (\epsilon > 0) \\ & && x_1, x_2, x_3 \geq 0. \end{aligned}$$

Observe the constraint

$$x_1 = \frac{\epsilon + x_2 + x_3}{\epsilon}$$

Hence, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is the optimal bfs with $\mathbf{z}^* = 1$

How big should M be ?

- Big-M problem:

$$\begin{aligned} \text{Min} \quad & x_1 + Mu \\ \text{s. t.} \quad & \epsilon x_1 - x_2 - x_3 + u = \epsilon \\ & x_1, x_2, x_3, u \geq 0. \end{aligned}$$

- Observations:

1. $\begin{bmatrix} 0 \\ 0 \\ 0 \\ \epsilon \end{bmatrix}$ is a bfs with $\mathbf{z} = M\epsilon$.

2. $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is a bfs with $\mathbf{z} = 1$.

3. To make sure (Big-M) generates a bfs to (LP), we need $M\epsilon > 1$ or $M > 1/\epsilon$.
   But remeber that ϵ can be arbitrarily small!

Consequence

- Commercial LP solvers prefer using the two-phase method.

Prevent cycling for finite termination

- Problem: When LP is degenerate,

$$x_p = 0 \text{ for some b.v. } x_p$$

$$\Rightarrow \text{step-length } \alpha = 0$$

$$\Rightarrow \mathbf{z} = \mathbf{c}^T\mathbf{x} = \mathbf{c}_B^T\mathbf{B}^{-1}b \text{ is not strictly decreasing!}$$

- Key idea: Keep something strictly monotone.
  1. Brand's rule: Leaving and entering in order.
  2. Lexicographic rule (1955): $[\mathbf{c}_B^T\mathbf{B}^{-1}b \mid \mathbf{c}_B^T\mathbf{B}^{-1}]$

*R.G. Bland, New finite pivoting rules for the simplex method, Math. Oper. Res. 2 (1977) 103–107.