The Simplex Method

(George B. Dantzig 1947)

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

Facts:

Cor. 2.2.2: If $P \neq \emptyset$, $\exists\ v \in P$.

Thm 2.3: If $P \neq \emptyset$ and $\mathbf{z}$ is not unbounded, then $\exists\ v \in P$ such that $\mathbf{z}^* = \mathbf{c}^T\mathbf{v}$.

Cor. 2.2.2: P has finite number of extreme points (vertices).

Observation:

(1) When $\binom{n}{m}$ is small, we can enumerate through *all* bfs and find the optimal solutions.

(2) When $\binom{n}{m}$ is large, we need a *systematic* and *efficient* way to do this!

– Simplex Method.

Basic Idea of Simplex Method

Phase I: (Starting)

To find an initial extreme point (ep) or $P = \emptyset$.

Phase II:

Step 1: (Check Optimality)

If current ep is optimal, STOP! Otherwise,

Step 2: (Pivoting)

Move to a better ep.

Go to Step 1.

Observation:

The algorithm terminates in finite steps.

Question:

How to identity an exterme point?

Cor. 2.1.1: A point $\mathbf{x} \in P$ is an extreme point of P *iff* $\mathbf{x}$ is a bfs corresponding to some basis $\mathbf{B}$.

Cor. 2.1.2: $\exists$ at most

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

bfs, when $\text{rank}(\mathbf{A}) = m \leq n$ a bfs is obtained by

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

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \hline \mathbf{x}_N \end{bmatrix}$$

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

Basic Idea of Simplex Method

Phase I: (Starting)

To find an initial basic feasible solution (bfs) or $P = \emptyset$.

Phase II:

Step 1: (Check Optimality)

If current bfs is optimal, STOP! Otherwise,

Step 2: (Pivoting)

Move to a better bfs.

Go to Step 1.

Question:

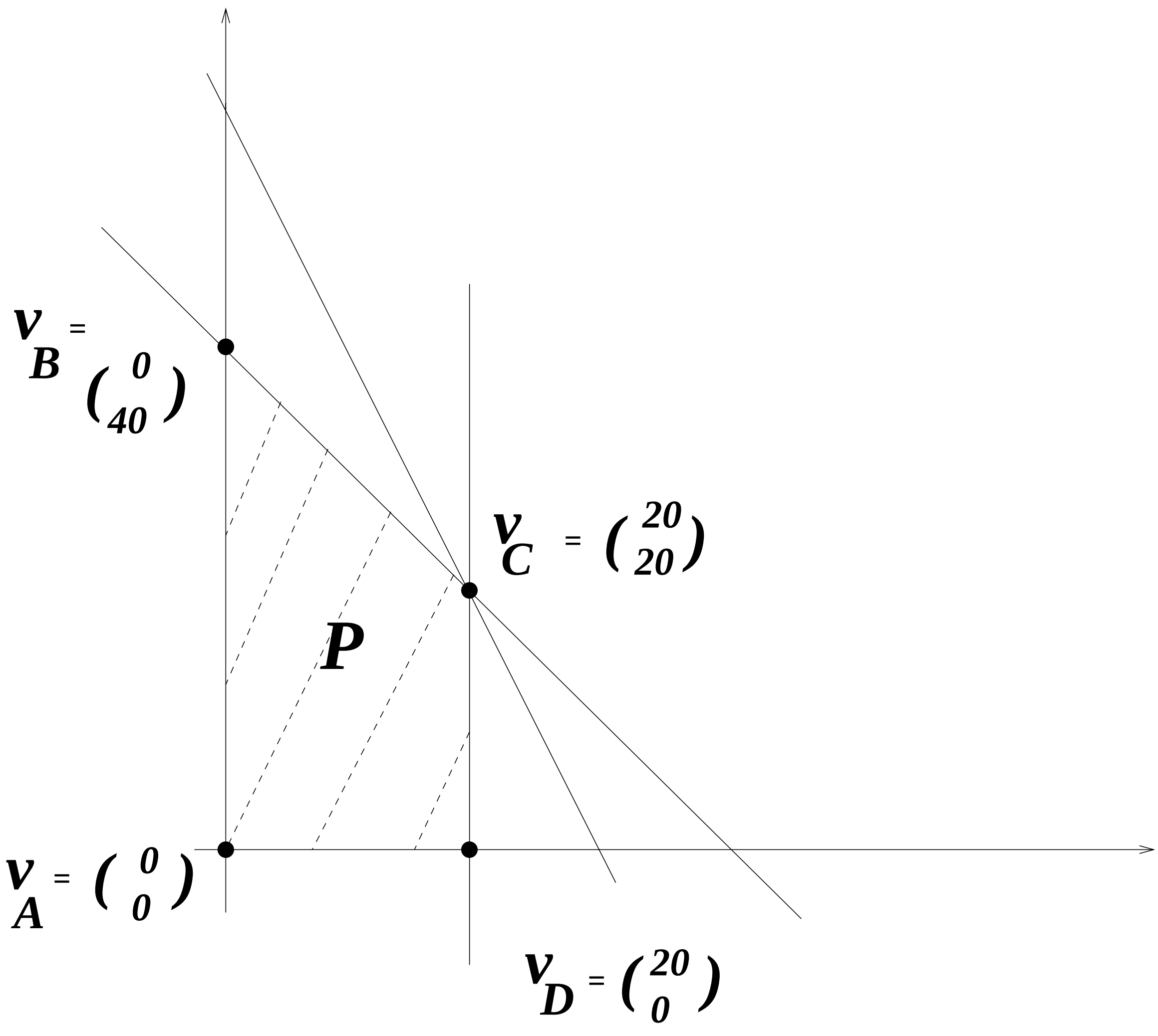
When we move from one bfs to another bfs, do we really move from one extreme point to another extreme point?

Answer:

Not necessary!

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

Observation:

(1) 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*.

   – Nondegenerate bfs.

(2) Only when there exists at least one basic variable$= 0$, then the ep may correspond to *more than one* bfs.

   – Degenerate bfs.

Terminology: An *LP is nondegenerate* if all bfs are nondegenerate.

Property 1: If a bfs $\mathbf{x}$ is nondegenerate, then $\mathbf{x}$ is uniquely determined by n hyperplanes.

Why?

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

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

Consider

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

then $\mathbf{M}$ is nonsingular, and

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

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

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

Answer:

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

i.e. $\mathbf{M}^{-1}$ is known when $\mathbf{B}^{-1}$ is known!

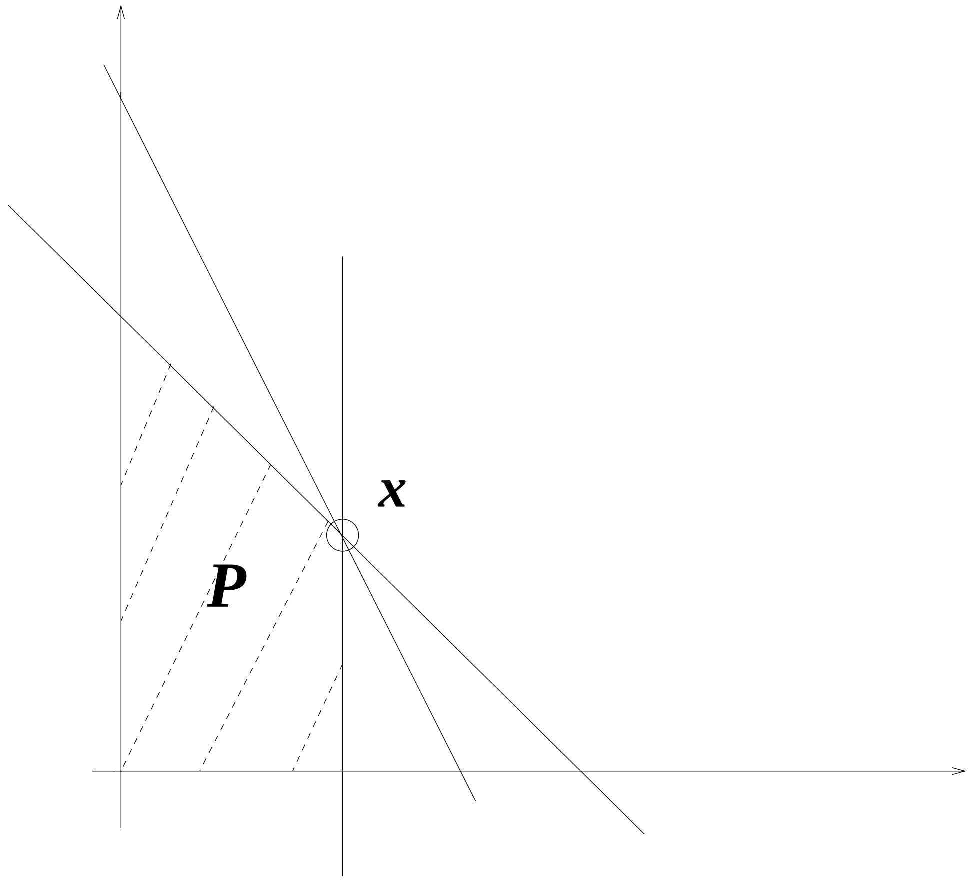
We call it a *Fundamental Matrix*.

Property 2: If a bfs $\mathbf{x}$ is degenerate, then $\mathbf{x}$ is *over-determined* by more than n hyperplanes.

Why? Except

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

$\exists$ at least one more $x_i \in \mathbf{x}_B$ such that $x_i = 0$. (another hyperplane!!)



Property 3: For a degenerate bfs $\mathbf{x}$ 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 exterme point.

Simplex Method Under Nondegeneracy

Basic Idea:

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

Definition:

Two bfs are *adjacent* if they have $m - 1$ basic variables in common.

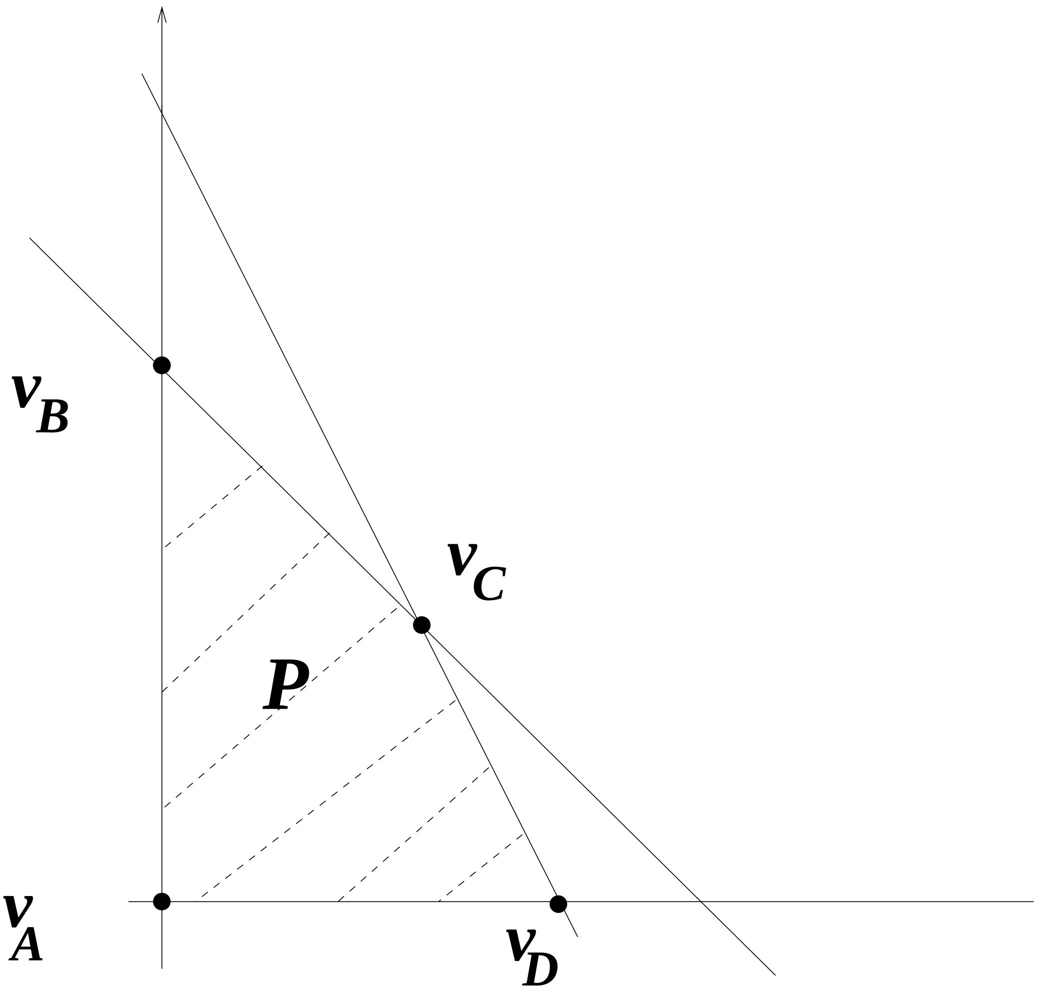
Observation:

(1) Under nondegeneracy, every extreme point have exactly $n - m$ adjacent extreme points.

(2) 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.

[See Example]

Example:

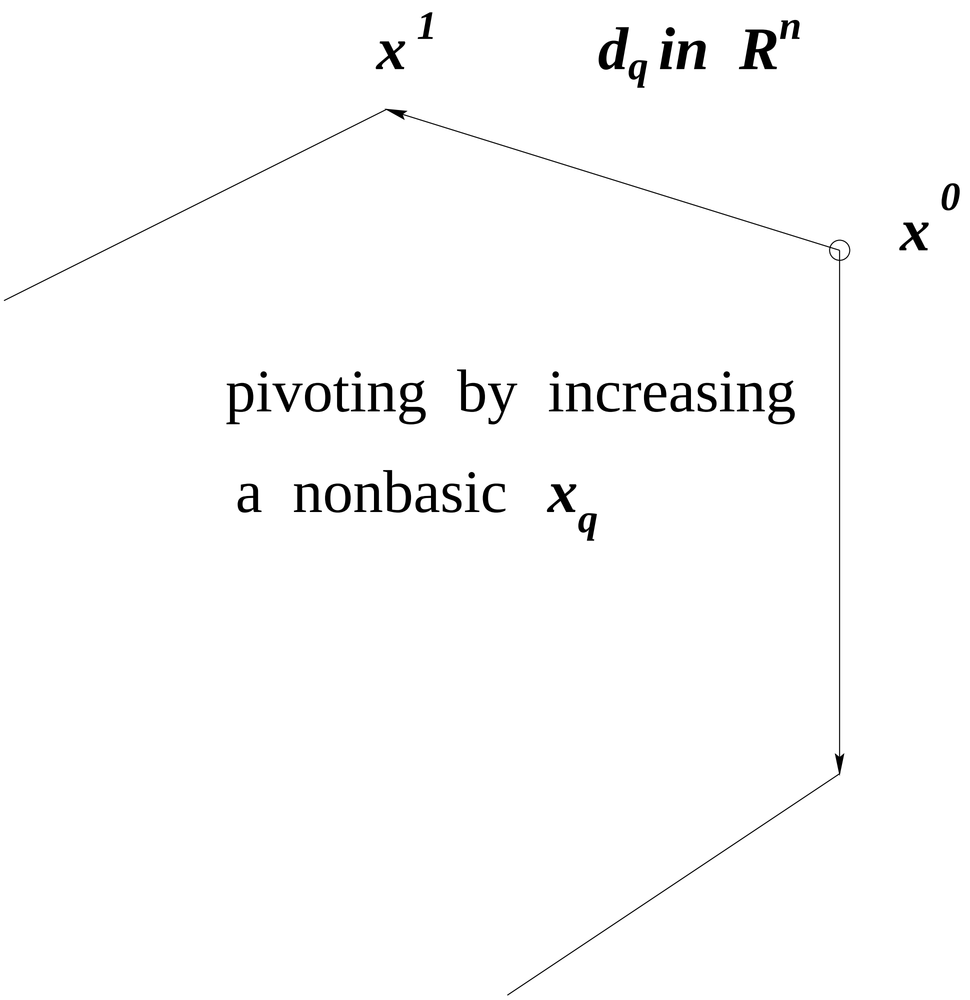


$$v_A = \begin{bmatrix} 0 \\ 0 \\ 40 \\ 60 \end{bmatrix}, v_B = \begin{bmatrix} 0 \\ 40 \\ 0 \\ 20 \end{bmatrix}, v_c = \begin{bmatrix} 20 \\ 20 \\ 0 \\ 0 \end{bmatrix}, v_D = \begin{bmatrix} 30 \\ 0 \\ 10 \\ 0 \end{bmatrix}.$$

Terminology:

pivoting

- One nonbasic variable enters the basis.
- One basic variable leaves the basis.



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

edge direction   step length

Question:

How to find an edge direction?

Hint:

(1) There should be $n - m$ edge directions leading to the adjacent extreme points.

(2) $\mathbf{d}_q \in \mathbf{R}^n$ for $\mathbf{x}_q$ n.b.v.

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

Conjecture:

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

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

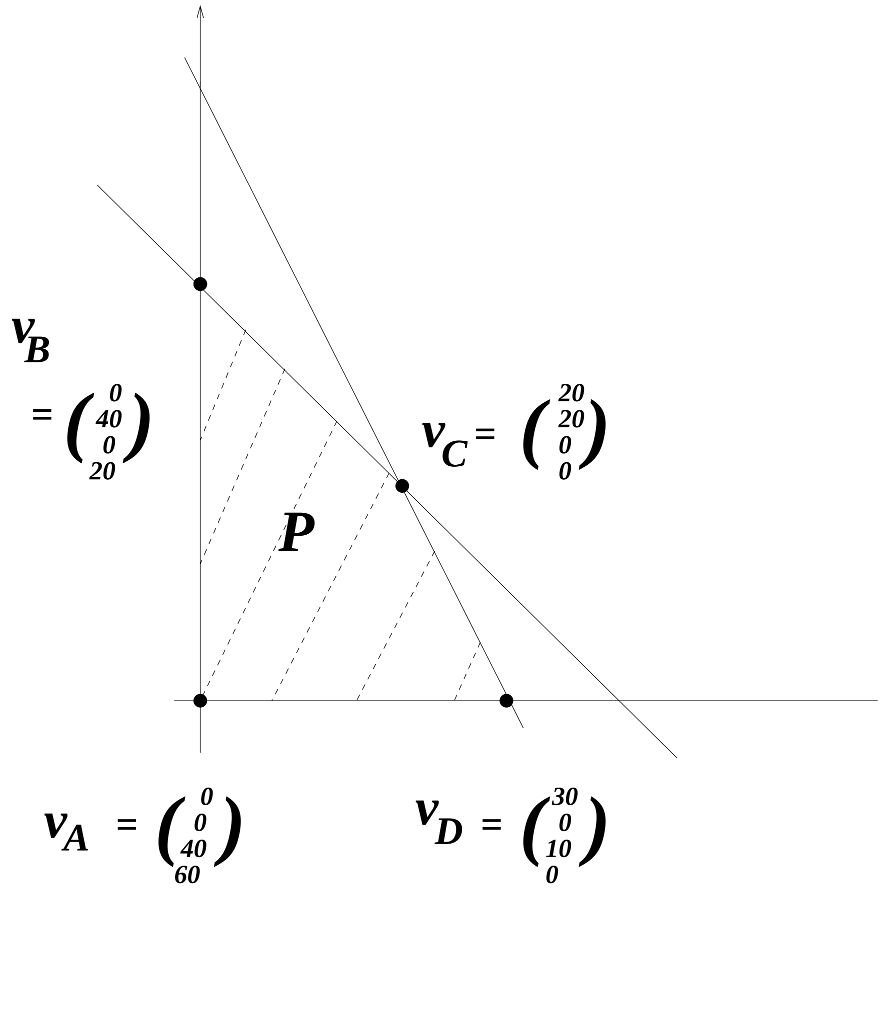
where

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

Example

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

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



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

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

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

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

In general, for $\lambda \geq 0$

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

Proof:

(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*?
i.e., for small enough $\lambda > 0$, $\mathbf{x}(\lambda) = \mathbf{x} + \lambda\mathbf{d}_q \in P$.

[Hint: Has to show that $\mathbf{A}\mathbf{x}(\lambda) = \mathbf{b} \Leftrightarrow \mathbf{A}\mathbf{d}_q = \mathbf{0}$ and $\mathbf{x}(\lambda) \geq 0$]

Facts:

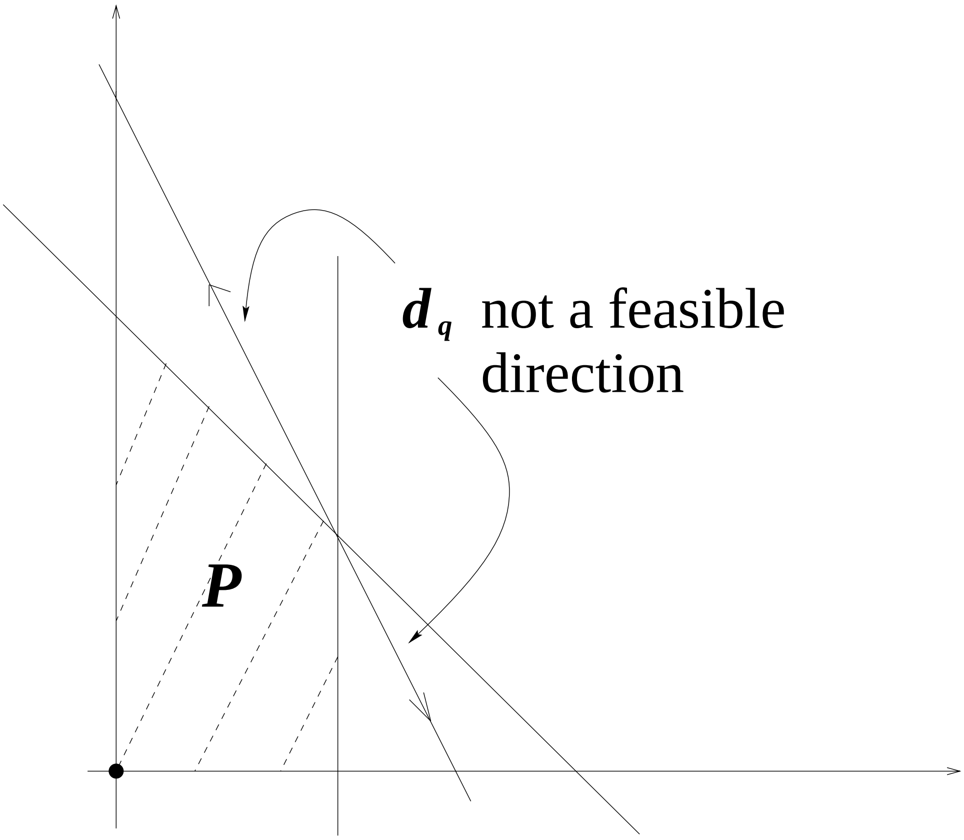
(1) $\mathbf{A}\mathbf{d}_q = \mathbf{0}$ can be derived from $\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$.

(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!

(3) For degenerate case, Not necessary!



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

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

Question: Which edge direction is a "good" direction to move?
i.e., which n.b.v. is a good candidate to "pivot in"?

Observation1:

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

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!

Theorem 3.2

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

Observation2:

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

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

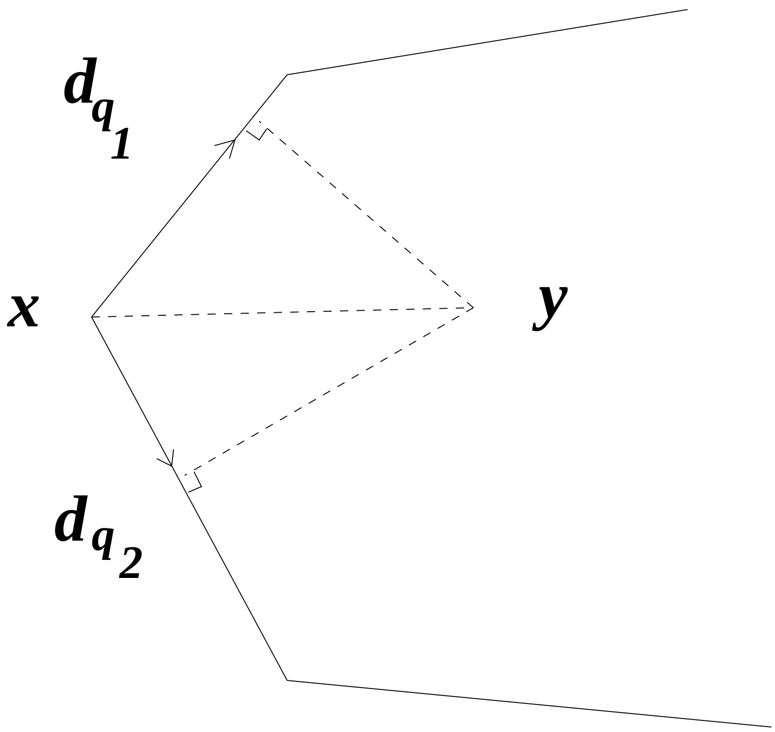
Observation3:

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

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

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

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

Theorem 3.1

Given a bfs $\mathbf{x}^0 = \left(\frac{\mathbf{B}^{-1}\mathbf{b}}{\mathbf{0}} \right)$ with basis $\mathbf{B}$, if $r_q \geq 0, \ \forall$ n.b.v x_q, then $\mathbf{x}$ is optimal.

Proof:

$\forall \ \mathbf{y} \in P, \ \mathbf{y} = \left(\frac{\mathbf{y}_B}{\mathbf{y}_N} \right) \geq 0, \ \mathbf{A}\mathbf{y} = \mathbf{b}$

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

Thus

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

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

with $\mathbf{y}_N = \begin{bmatrix} \vdots \\ y_q \\ \vdots \end{bmatrix} \geq 0$

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

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

Observation1:

A bfs $\mathbf{x}$ is the *unique optimal solution*, if $r_q > 0$, $\forall$ n.b.v. x_q.

Observation2:

If $\mathbf{x}$ is an optimal bfs with
$r_{q_1}, r_{q_2}, \ldots, r_{q_k} = 0$,
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 Thm 3.1 true?

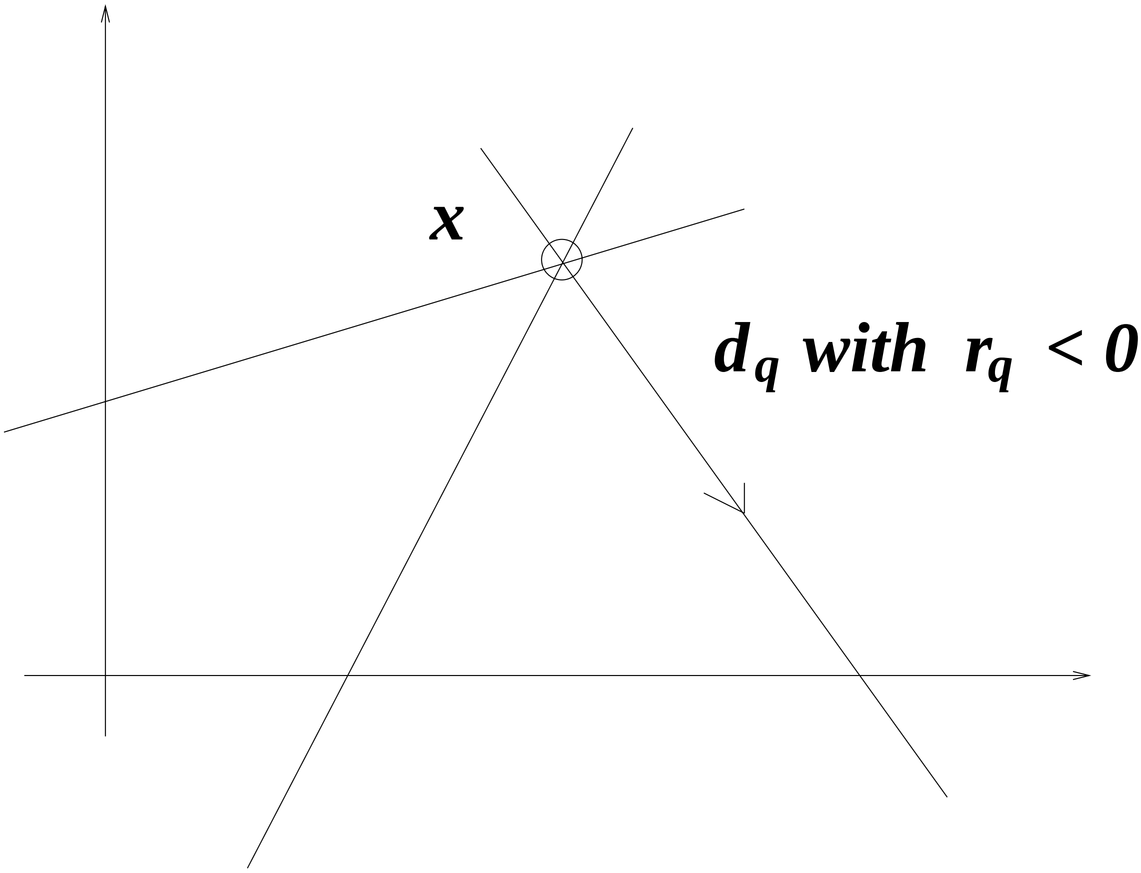
i.e.,

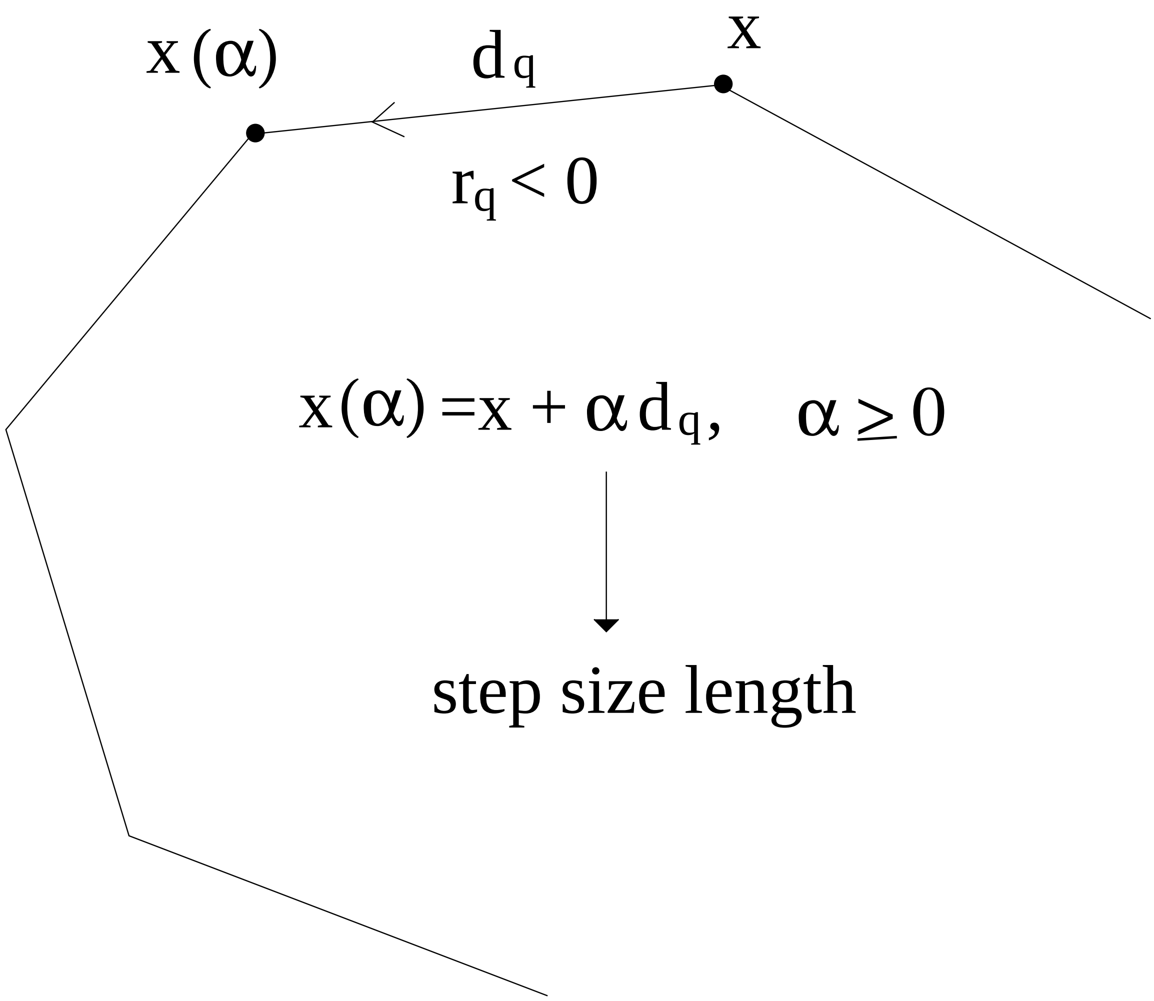
"If a bfs $\mathbf{x}$ is optimal, then $r_q \geq 0, \ \forall$ n.b.v x_q."

Answer:

True only for the nondergeneracy case.

For degeneray case:





Question:

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

Analysis:

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

Remember that $\mathbf{A}\mathbf{d}_q = \mathbf{0}$, thus $\mathbf{A}\mathbf{x}(\alpha) = \mathbf{A}\mathbf{x} = \mathbf{b}$.

Case1: $\mathbf{d}_q \geq \mathbf{0}$, 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 3.3

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: $\mathbf{d}_q = \begin{pmatrix} \dfrac{-\mathbf{B}^{-1}\mathbf{A}_q}{e_q} \end{pmatrix}$. Define $\mathbf{w} \triangleq \mathbf{B}^{-1}\mathbf{A}_q$, then

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

Case2: $\mathbf{d}_q$ has at least one component < 0.
To keep $\mathbf{x}(\alpha) \geq \mathbf{0}$, we have to choose

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

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

Note2: α is determined by the
Minimum ratio test.

Note3: 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.

Theorem 3.3

Given that $\mathbf{x}$ is a bfs, then $\mathbf{x}(\alpha) = \mathbf{x} + \alpha \mathbf{d}_q$ is an adjacent bfs, if α is determined by the Minimum Ratio Test.

[*Under nondegeneracy*, $\mathbf{x}(\alpha)$ moves to a neighboring extreme point.]

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.

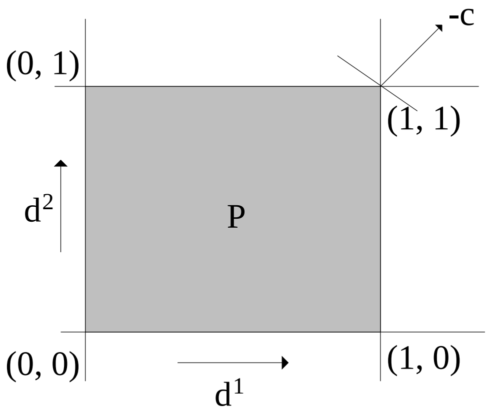
Step3: If $\mathbf{d}_q \geq 0$, then LP is unbounded.
Otherwise, find

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

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

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

Theorem 3.5 Under nondegeneracy, Simplex Method terminates in finite iterations.



$$\begin{aligned} \text{Min} \quad & -x_1 - x_2 \\ \text{s.t.} \quad & x_1 \leq 1 \\ & x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

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

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

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

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

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

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

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

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

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

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. $\{x_3, x_2\}$

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

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

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

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

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. $\{x_3, x_4\}$

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

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

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

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

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

How to start the Simplex Method?

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

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

$$
\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \text{(artificial variables)}
$$

Observations

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).
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 $\mathbf{Ax} = \mathbf{b}$ is redundant. We remove it and start again.

Big-M Method: For a large $M > 0$,

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

Observations

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

   (i) $u^* = 0$,

   then $\forall$ $\mathbf{x}$ feasible to (LP), $\binom{\mathbf{x}}{0}$ is feasible to (big-M). Thus

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

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

   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, $P = \emptyset$.

Problem How big is big?

Example

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

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

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

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

Observation

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

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

(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!

Prevent Cycling

Problem: When degenerate, $x_p = 0$ for some b.v. x_p

$\Rightarrow$ step-length $\alpha = 0$

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

Idea: Keep something strictly monotone.

Lexicographic Rule (1955)

$$[\mathbf{c}_B^T\mathbf{B}^{-1}b \mid \mathbf{c}_B^T\mathbf{B}^{-1}]$$

Bland's Rule

leaving and entering order