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)

\[
\begin{align*}
\text{Min} & \quad c^T x \\
\text{(LP)} & \quad \text{s. t.} \quad Ax = b \\
& \quad x \geq 0
\end{align*}
\]

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 $\mathbf{x}$ in $\mathcal{P}$ is an extreme point if and only if $\mathbf{x}$ is a basic feasible solution corresponding to some basis $B$.

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

$$
A = \begin{bmatrix} B & N \end{bmatrix}
$$

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

and set $x_N = 0$ to calculate $x_B = B^{-1}b$. 
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

\[
\begin{align*}
\begin{cases}
    x_1 + x_2 & \leq 40 \\
    2x_1 + x_2 & \leq 60 \\
    x_1 & \leq 20 \\
    x_1, x_2 & \geq 0.
\end{cases}
\end{align*}
\]

\[
\begin{align*}
\begin{cases}
    x_1 + x_2 + s_1 & = 40 \\
    2x_1 + x_2 + s_2 & = 60 \\
    x_1 + s_3 & = 20 \\
    x_1, x_2, s_1, s_2, s_3 & \geq 0.
\end{cases}
\end{align*}
\]

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

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

(3) \(BV = \{x_1, x_2, s_3\} \quad NBV = \{s_1, s_2\}\)
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 $\mathbf{x}$ is nondegenerate, then $\mathbf{x}$ is uniquely determined by $n$ hyperplanes.
- Why? $n$ hyperplanes? Where are they?
- Remember that

\[
A = \begin{bmatrix} B & N \end{bmatrix}
\]

\[
x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}
\]

- Let

\[
M = \begin{bmatrix} B & N \\ 0 & I \end{bmatrix}
\]

Then $M$ is nonsingular and

\[
Mx = \begin{bmatrix} B & N \\ 0 & I \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}.
\]

Hence $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = M^{-1} \begin{bmatrix} b \\ 0 \end{bmatrix}$ is uniquely determined by $n$ linearly independent hyperplanes.
Fundamental matrix

• Question:  \( M^{-1} = ? \)

• Answer:  \[
M^{-1} = \begin{bmatrix}
B^{-1} & -B^{-1}N \\
0 & I
\end{bmatrix}
\]

• Hence, \( M^{-1} \) is known when \( B^{-1} \) is known!

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

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

\[
\begin{bmatrix}
B & N \\
0 & I \\
b & 0
\end{bmatrix}
\begin{bmatrix}
x_B \\
x_N
\end{bmatrix} = \begin{bmatrix}
b \\
0
\end{bmatrix}
\]

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
  
  \[
  \binom{n-p}{n-m} = \frac{(n-p)!}{(n-m)!(m-p)!}
  \]

  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 nonbasic variable from zero to positive and decreasing one basic variable from positive to zero. – Pivoting
Pivoting

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

\[ x^1 = x^0 + \lambda d_q \text{ for } \lambda > 0. \]

**edge direction**  **step length**

- **Diagram:**
  
  - **Notation:** \( x^1 \) is the new solution after pivoting, \( x^0 \) is the initial solution, \( d_q \) is a direction vector in \( R^n \).
  
  - **Explanation:** Pivoting by increasing a nonbasic \( x_q \).

- **Equation:**
  
  \[ x^1 = x^0 + \lambda d_q \text{ for } \lambda > 0. \]
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 $d_q \in \mathbb{R}^n$ corresponding the increasing of a nonbasic variable $x_q$.

- 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 \textit{n-m columns} in the part of \[ -B^{-1}N \] .

• Could they be the \textit{edge directions}?
Conjecture

d_q is in the column in M^{-1} corresponding to x_q, i.e.

\[
d_q = \begin{pmatrix}
-B^{-1}A_q \\
0 \\
\vdots \\
0 \\
1 \\
0 \\
\vdots \\
0
\end{pmatrix},
\]

where

\[
A = (A_1 | A_2 | \cdots | A_n).
\]
Example

\[
\begin{align*}
\begin{cases}
  x_1 + x_2 + x_3 &= 40 \\
  2x_1 + x_2 &+ x_4 &= 60 \\
  x_1, x_2, x_3, x_4 &\geq 0.
\end{cases}
\end{align*}
\]

\[
A = \begin{pmatrix}
1 & 1 & 1 & 0 \\
2 & 1 & 0 & 1
\end{pmatrix}.
\]

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

\[
B = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix},
N = \begin{pmatrix}
1 & 1 \\
2 & 1
\end{pmatrix}.
\]

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

\[
M^{-1} = \begin{pmatrix}
1 & 0 & -1 & -1 \\
0 & 1 & -2 & -1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}.
\]
Example - continue

• From \( v_A \) to \( v_B \),
  \[
  \begin{bmatrix}
    0 \\
    40 \\
    0 \\
    20
  \end{bmatrix} - \begin{bmatrix}
    0 \\
    0 \\
    0 \\
    60
  \end{bmatrix} = 40 \begin{bmatrix}
    1 \\
    -1
  \end{bmatrix}
  \]

• From \( v_A \) to \( v_D \),
  \[
  \begin{bmatrix}
    30 \\
    0 \\
    10 \\
    0
  \end{bmatrix} - \begin{bmatrix}
    0 \\
    0 \\
    0 \\
    60
  \end{bmatrix} = 30 \begin{bmatrix}
    1 \\
    0 \\
    -1
  \end{bmatrix}
  \]

\[
M^{-1} = \begin{pmatrix}
1 & 0 & -1 & -1 \\
0 & 1 & -2 & -1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

\[
v_A = \begin{bmatrix}
0 \\
40 \\
20 \\
60
\end{bmatrix}, \quad v_B = \begin{bmatrix}
0 \\
40 \\
20 \\
60
\end{bmatrix}, \quad v_C = \begin{bmatrix}
20 \\
20 \\
20 \\
30
\end{bmatrix}, \quad v_D = \begin{bmatrix}
0 \\
0 \\
0 \\
0
\end{bmatrix}
\]
General case

In general, for $\lambda \geq 0$

$$x(\lambda) = x + \lambda d_q = \left( \begin{array}{c} x_B \\ x_N \end{array} \right) + \lambda \left( \begin{array}{c} -B^{-1}A_q \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{array} \right)$$

(1) For nonbasic variables, all are kept at zero, except $x_q$ increases by $\lambda$, i.e.

$$x_N(\lambda) = x_N + \lambda \left( \begin{array}{c} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{array} \right)$$

(2) For basic variables, since $Bx_B + Nx_N = b$,

thus $x_B = B^{-1}b - B^{-1}Nx_N$

when $x_q$ increases by $\lambda$ and the rest n.b.v are kept at 0, then $x_B(\lambda) = B^{-1}b - \lambda B^{-1}A_q$, Hence

$$d_q = \left( \begin{array}{c} -B^{-1}A_q \\ e_q \end{array} \right)$$
Question

• Is an edge direction $d_q$ always a feasible direction?
• That means for a small enough step length $\lambda > 0$, we need
  \[ x(\lambda) = x + \lambda d_q \in P. \]
• Must show that $Ax(\lambda) = b$ and $x(\lambda) \geq 0$.
• Equivalently, we need to show that $Ad_q = 0$ and $x(\lambda) \geq 0$. 
Answer - I

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

- Proof:

\[
(1) \; A d_q = 0 \text{ can be derived from } MM^{-1} = I.
\]

\[
(2) \text{ For nondegenerate case, } \\
x(\lambda) = x + \lambda \left( \frac{-B^{-1}A_q}{e_q} \right)
\]

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

• No, an edge direction is not necessarily a feasible direction when the problem is degenerate.

• Proof:

\[
x(\lambda) = x + \lambda \left( \frac{-B^{-1}A_q}{e_q} \right)
\]

say \( x_i = 0 \), no matter how small \( \lambda \) is, \( x_i(\lambda) < 0 \)!!
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:

\[
\begin{align*}
z(x(\lambda)) &= c^T x(\lambda) \\
&= c^T (x + \lambda d_q) \\
&= z(x) + \lambda [c_B^T | c_N^T] \left( \begin{array}{c} -B^{-1} A_q \\ e_q \end{array} \right) \\
&= z(x) + \lambda [c_q - c_B^T B^{-1} A_q] \\
&= z(x) + \lambda r_q
\end{align*}
\]

If \( r_q = c^T d_q = c_q - c_B^T B^{-1} A_q < 0 \), then \( d_q \) is a good direction!
Reduced cost

- Definition: The quantity of
  \[ r_q = c^T d_q = c_q - c^T_B B^{-1} A_q \]
  is called a reduced cost with respect to the variable \( x_q \).

Theorem:

If \( x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} \) is a bfs with \( B \) and \( r_q < 0 \) for some n.b.v. \( x_q \), then \( d_q = \begin{pmatrix} -B^{-1}A_q \\ e_q \end{pmatrix} \in \mathbb{R}^n \)
leads to an improved objective value.
Observations

• Observation 1:
  
  For a basic variable $x_q \in B$, 
  
  \[ r_q = c_q - c_B^T B^{-1} A_q \]
  
  \[ = c_q - c_q \]
  
  \[ = 0. \]

• Observation 2:
  
  Any $d_q$ ($x_q$ n.b.v.) with $r_q < 0$ will do for the simplex method. The one with most reduced cost can be found by

  \[ \min_{j:\text{nonbasic}} \left\{ \frac{c^T d_j}{\|d_j\|} \right\}. \]
Optimality check by reduced cost

• Question:

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

• Guess:

\( \forall y \in P, \)

\[ y = x + y_{q_1} d_{q_1} + y_{q_2} d_{q_2}, \quad y_{q_1}, y_{q_2} \geq 0 \]

Hence

\[ c^T y = c^T x + y_{q_1} c^T d_{q_1} + y_{q_2} c^T d_{q_2} \geq c^T x + 0 = c^T x \]
Optimality condition

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

• Proof:

$\forall \ y \in P, \ y = \begin{pmatrix} \frac{y_B}{y_N} \end{pmatrix} \geq 0, \ A\ y = b$

Note $x_N^0 = 0$ and $A\ x^0 = b$

Thus

$M(y - x^0) = \begin{bmatrix} B & N \\ 0 & I \end{bmatrix} \begin{bmatrix} \frac{y_B - x_B^0}{y_N} \end{bmatrix}$

$= \begin{bmatrix} \frac{b-b}{y_N} \\ 0 \\ y_N \end{bmatrix}$

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

Hence $c^T \ y \geq c^T \ x^0, \ \forall \ y \in P.$
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 $y \in P$ such that

$$y = x + \sum_{i=1}^{k} y_{q_i} d_{q_i}$$

is also optimal.
Question

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

  “If a bfs $x$ is optimal, then $r_q \geq 0$, $\forall$ n.b.v $x_q$.”

- Answer:

  True only for the nondegeneracy case.
  For degeneracy case:
How far is my good neighbor?

• Basic concept:

• Question:

How far should we go such that $x(\alpha)$ is an adjacent bfs?
Analysis of step length

- We have \( x(\alpha) = x + \alpha d_q, \ \alpha > 0 \).

  with \( r_q = c^T d_q = c_q - c^T_B B^{-1} A_q < 0 \).

- Remember that \( Ad_q = 0 \), thus \( Ax(\alpha) = Ax = b \).

- Case 1: If \( d_q \geq 0 \), then \( x(\alpha) \geq 0 \), \( \forall \alpha \geq 0 \).

  Hence \( x(\alpha) \in P \), \( \forall \alpha \geq 0 \) and
  \[ c^T x(\alpha) = c^T x + \alpha c^T d_q \rightarrow -\infty, \text{ as } \alpha \rightarrow +\infty. \]

- Theorem:
  If \( x \) is a bfs with \( d_q \geq 0 \) and \( r_q < 0 \), for some n.b.v. \( x_q \), then the LP is unbounded.

  \[ \text{Note: } d_q = \left( \begin{array}{c} -B^{-1} A_q \\ e_q \end{array} \right). \text{ Define } w = B^{-1} A_q, \]

  then \( d_q \geq 0 \iff w \leq 0 \)
Analysis - continue

• Case 2: \( d_q \) has at least one component \(< 0 \).

To keep \( x(\alpha) \geq 0 \), we have to choose

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

• Observations:

**Note 1:** \( d_{qi} < 0 \) can only happen for basic variables \((x_i \in B)\).

**Note 2:** \( \alpha \) is determined by the Minimum ratio test.

**Note 3:** Under nondegeneracy, \( x_i > 0 \) for b.v. \( x_i \)

\( \Rightarrow \alpha > 0 \)

\( \Rightarrow x(\alpha) \) is a different extreme point.

For degenerate bfs, it is possible \( x_i = 0 \), then \( \alpha = 0 \)

\( \Rightarrow x(\alpha) \) stays at the same extreme point.
Step length by minimum ratio test

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

• Note that this indeed moves to an adjacent extreme point, when the bfs $\mathbf{x}$ is nondegenerate.
Key steps of Simplex Method

Step 1: Find a bfs $x$ with $A = [B|N]$.

Step 2: Check for n.b.v’s

$$r_q = c^T d_q (= c_q - c_B^T B^{-1} A_q).$$

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.

Step 3: If $d_q \geq 0$, 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 $x \leftarrow x + \alpha d_q$ is a new bfs.

Update $B$ and $N$. 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{align*}
\min & \quad -x_1 - x_2 \\
x_1 + x_3 &= 1 \\
x_2 + x_4 &= 1 \\
x_1, x_2, x_3, x_4 &\geq 0
\end{align*}
Example – first iteration

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

\[
\begin{align*}
\mathbf{x} &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\
\mathbf{A} &= [\mathbf{B} | \mathbf{N}] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \\
\mathbf{B}^{-1} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I} \\
\mathbf{B}^{-1}\mathbf{N} &= \mathbf{N} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
\mathbf{M}^{-1} &= \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{align*}
\]
Example – check reduced cost for optimality

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

\[ r_2 = c^T d^2 = [0 \ 0 \ -1 \ -1] \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = -1 < 0 \]
Example – moving to better neighbor

Pick $d^1 \geq 0$, so $x_1$ enters the basis.

$$
\alpha = \min_i \left\{ \frac{x_i}{-d^1_i} \ \middle| \ -d^1_i < 0 \right\} = \frac{x_3}{-d^1_{x_3}} = -\frac{1}{-1} = 1
$$

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

So, $x_3$ leaves the basis.

$\textbf{bfs}\# 2$: b.v. $\{x_1, x_4\}$, n.b.v. $\{x_3, x_2\}$
Example – second iteration

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

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

\[
M^{-1} = \begin{bmatrix}
1 & 0 & -1 & 0 \\
0 & 1 & 0 & -1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

Example – optimality check

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

\[ r_2 = c^T d^2 = [-1 \ 0 \ 0 \ -1] \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = -1 < 0 \]
Example – move to a better neighbor

Pick $d^2 (\geq 0)$, so $x_2$ enters the basis.

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

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

So, $x_4$ leaves the basis.

bfs#3: b.v. $\{x_1, x_2\}$, n.b.v. $\{x_3, x_4\}$
**Example – third iteration**

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

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

\[
M^{-1} = \begin{bmatrix}
1 & 0 & -1 & 0 \\
0 & 1 & 0 & -1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
r_3 = c^T d^3 = [-1 \ -1 \ 0 \ 0]
\]

\[
\begin{bmatrix}
-1 \\
0 \\
1 \\
0
\end{bmatrix}
\]

\[
= 1 > 0
\]

\[
x = \begin{bmatrix}
1 \\
1 \\
0 \\
0
\end{bmatrix}
\]

\[
\text{optimal!}
\]

\[
r_4 = c^T d^4 = [-1 \ -1 \ 0 \ 0]
\]

\[
\begin{bmatrix}
-1 \\
0 \\
1 \\
0
\end{bmatrix}
\]

\[
= 1 > 0
\]
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:**

  \[
  \begin{align*}
  \text{Min} & \quad c^T x \\
  \text{(LP)} & \quad \text{s. t.} \quad Ax = b(\geq 0) \\
  & \quad x \geq 0
  \end{align*}
  \]

- Step 2: **Add} m \text{ artificial variables} for Phase 1 problem:

  \[
  u = \begin{bmatrix}
  u_1 \\
  u_2 \\
  \vdots \\
  u_m
  \end{bmatrix}
  \]

  \[
  \begin{align*}
  \text{Min} & \quad \sum_{i=1}^{m} u_i \\
  \text{(PhI)} & \quad \text{s. t.} \quad Ax + Lu = b(\geq 0) \\
  & \quad x, u \geq 0
  \end{align*}
  \]
What’re special about Phase I problem?

1. \( \mathbf{u} = b, \mathbf{x} = 0 \) is a bfs of \((\text{PhI})\).

2. \((\text{PhI})\) is bounded below by 0.

3. \((\text{LP})\) is feasible if and only if \( \mathbf{z}^*_{\text{PhI}} = 0 \)

4. Under nondegeneracy, if \( \mathbf{z}^*_{\text{PhI}} = 0 \), then an optimal solution of \((\text{PhI})\) is a bfs of \((\text{LP})\).
How about degenerate case?

5. If $z^*_p = 0$ 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 $e_k^T B^{-1} A_q \neq 0$ for a n.b.v. $x_q$, then $u_i$ can be replaced by $x_q$ to form a starting basis.

2. If $e_k^T B^{-1} A_q = 0$, $\forall$ n.b.v. $x_q$, then the $k$-th row of $A x = b$ 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:

$$
\text{Min} \quad \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} M u_i \\
\text{s. t.} \quad Ax + Iu = b (\geq 0) \\
x, u \geq 0
$$
What’re special about big-M problem

1. $x = 0, u = b$, is a bfs.

2. $z^*$ can be finite at an optimal solution $(\frac{x^*}{u^*})$ or unbounded below.

3. Suppose $z^*$ is finite at $(\frac{x^*}{u^*})$. If
   (i) $u^* = 0$,
   then $\forall x$ feasible to (LP), $(\frac{x}{0})$ is feasible to (big-M). Thus
   \[
   c^T x + M \times 0 \geq c^T x^* + M \sum_{i=1}^{m} u^*_i
   \]
   \[
   c^T x \geq c^T x^* + 0
   \]
   i.e., $x^*$ is optimal to (LP).

   (ii) $u^* \neq 0$,
   then for $x$ feasible to (LP), $(\frac{x}{0})$ is feasible to (big-M) and
   \[
   c^T x + M \times 0 \geq c^T x^* + M \sum_{i=1}^{m} u^*_i
   \]
   But this is impossible for $M$ is large enough. Hence $P = \emptyset$.

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

Min $\sum_{j=1}^{n} c_jx_j + \sum_{i=1}^{m} Mu_i$

s. t. $Ax + Iu = b(\geq 0)$

$x, u \geq 0$
Big-M problem

• Question: How big should M be?

• Example:

\[
\begin{align*}
\text{(LP)} & \quad \text{Min } x_1 \\
\text{s. t.} & \quad \epsilon x_1 - x_2 - x_3 = \epsilon \quad (\epsilon > 0) \\
& \quad x_1, x_2, x_3 \geq 0.
\end{align*}
\]

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 \( z^* = 1 \).
How big should $M$ be?

- Big-M problem:
  \[
  \begin{align*}
  \text{Min} & \quad x_1 + Mu \\
  \text{s. t.} & \quad \epsilon x_1 - x_2 - x_3 + u = \epsilon \\
  & \quad x_1, x_2, x_3, u \geq 0.
  \end{align*}
  \]

- Observations:
  1. \[
  \begin{bmatrix}
  0 \\ 0 \\ 0 \\ \epsilon
  \end{bmatrix}
  \] is a bfs with $z = M\epsilon$.
  2. \[
  \begin{bmatrix}
  1 \\ 0 \\ 0 \\ 0
  \end{bmatrix}
  \] is a bfs with $z = 1$.
  3. To make sure (Big-M) generates a bfs to (LP), we need $M\epsilon > 1$ or $M > 1/\epsilon$.
  
  But remember that $\epsilon$ can be arbitrarily small!
Consequence

- Commercial LP solvers prefer using the two-phase method.
Prevent cycling for finite termination

• Problem: When LP is degenerate,
  \[ x_p = 0 \text{ for some basic variable } x_p \]
  \[ \Rightarrow \text{step-length } \alpha = 0 \]
  \[ \Rightarrow z = c^T x = c_B^T B^{-1} b \text{ is not strictly decreasing!} \]

• Key idea: Keep something strictly monotone.
  1. Brand’s rule: Leaving and entering in order.
  2. Lexicographic rule (1955):

\[ [c_B^T B^{-1} b \mid c_B^T B^{-1}] \]