Math 273a: Optimization
The Simplex method

Instructor: Wotao Yin
Department of Mathematics, UCLA

Fall 2015

material taken from the textbook Chong-Zak, 4th Ed.

Overview: idea and approach

- If a standard-form LP has a solution, then there exists an extreme-point solution.
  Therefore: **just search the extreme points**.
- In a standard-form LP, what is an extreme point?
  Answer: a **basic feasible solution (BFS)**, defined in a linear algebraic form
- Remaining: move through a series of BFSs until reaching the optimal BFS

  We shall

  - recognize an optimal BFS
  - move from a BFS to a *better* BFS
    (realize these in linear algebra for the LP standard form)

Overview: edge direction and reduced cost

- **Edges** connect two neighboring BFSs
- **Reduced cost** is how the objective will change when moving along an edge direction
- How to recognize the optimal BFS?
  - none of the feasible edge directions is improving
  - equivalently, all reduced costs are *nonnegative*

Overview: move from a BFS to a better BFS

- If the BFS is *not* optimal, then some reduced cost is *negative*
- How to move to a better BFS?
  - pick a *feasible* edge direction with a *negative* reduced cost
  - move along the edge direction until reaching another BFS
  - it is possible that the edge direction is unbounded

The Simplex method (abstract)

- **input:** an BFS x
- **check:** reduce costs ≥ 0
- **if yes:** optimal, return x; **stop.**
- **if not:** choose an edge direction corresponding to a negative reduced cost, and then move along the edge direction
  - **if unbounded:** then the problem is **unbounded**
  - **otherwise:** replace x by the new BFS; restart

Basic solution (not necessarily feasible)

$$\begin{aligned} \text{minimize } & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to } & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0. \end{aligned}$$

- common assumption: $\mathrm{rank}(\boldsymbol{A}) = m$, full row rank or $\boldsymbol{A}$ is surjective

  (otherwise, *either* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has no solution *or* some rows of $\boldsymbol{A}$ can be safely eliminated)

- write $\boldsymbol{A}$ as

$$\boldsymbol{A} = [\boldsymbol{B}, \boldsymbol{D}]$$

  where $\boldsymbol{B}$ is a square matrix with full rank (its rows/columns are linearly independent). This might require reordering the columns of $\boldsymbol{A}$.

- We call $\boldsymbol{x} = [\boldsymbol{x}_B^T, \boldsymbol{0}^T]^T$ a **basic solution** if $\boldsymbol{B}\boldsymbol{x}_B = \boldsymbol{b}$. $\boldsymbol{x}_B$ and $\boldsymbol{B}$ are called **basic variables** and **basis**.

Basic feasible solution(BFS)

- "basic" because $\boldsymbol{x}_B$ is uniquely determined by $\boldsymbol{B}$ and $\boldsymbol{b}$
- more definitions:
  - if $\boldsymbol{x} \geq \mathbf{0}$ (equivalently, $\boldsymbol{x}_B \geq \mathbf{0}$), then $\boldsymbol{x}$ is a **basic feasible solution (BFS)**
  - if any entry of $\boldsymbol{x}_B$ is 0, then $\boldsymbol{x}$ is a *degenerate*; otherwise, it is called *nondegenerate*. (Why? it may be difficult to move from a degenerate BFS to another BFS)
- given basic columns, a basic solution is determined, and then we check whether the solution is feasible and/or degenerate

Example (example 15.12)

$$\boldsymbol{A} = \left[\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4\right] = \begin{bmatrix} 1 & 1 & -1 & 4 \\ 1 & -2 & -1 & 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

1. Pick $\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_2]$, obtain $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} = [6, 2]^T$.

   $\boldsymbol{x} = [6, 2, 0, 0]^T$ is a basic feasible solution and is nondegenerate.

2. Pick $\boldsymbol{B} = [\boldsymbol{a}_3, \boldsymbol{a}_4]$, obtain $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} = [0, 2]^T$.

   $\boldsymbol{x} = [0, 0, 0, 2]^T$ is a *degenerate* basic feasible solution.

   In this example, $\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_4]$ and $[\boldsymbol{a}_2, \boldsymbol{a}_4]$ also give the same $\boldsymbol{x}$!

3. Pick $\boldsymbol{B} = [\boldsymbol{a}_2, \boldsymbol{a}_3]$, obtain $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} = [2, -6]^T$.

   $\boldsymbol{x} = [0, 2, -6, 0]^T$ is a basic solution but *infeasible*, violating $\boldsymbol{x} \geq 0$.

4. $\boldsymbol{x} = [3, 1, 0, 1]^T$ is feasible but not basic.

The total number of possible basic solutions is at most

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

For small m and n, e.g., $m = 2$ and $n = 4$, this number is 6. So we can check each basic feasible for feasibility and optimality.

Any vector $\boldsymbol{x}$, basic or not, that yields the minimum value of $\boldsymbol{c}^T\boldsymbol{x}$ over the feasible set $\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$ is called an *optimal (feasible) solution*.

The idea behind "basic solution"

- In $\mathbb{R}^n$, a set of n linearly independent equations define a unique point (special case: in $\mathbb{R}^2$, two crossing lines determine a point)
- Thus, an extreme point of $P = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$ is given by n linearly independent and active (i.e., "$=$") constraints obtained from
  - $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ (m linear constraints)
  - $\boldsymbol{x} = 0$ (n linear constraints)
- In a standard form LP, if we assume $\operatorname{rank}(\boldsymbol{A}) = m < n$, then $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ give just m linearly independent constraints. Therefore, we need to
  - select addition $(n - m)$ linear constraints from $x_1 \geq 0, \ldots, x_n \geq 0$ and make them *active*, that is, set the corresponding components 0
  - ensure that all the n linear constraints are linearly independent

- without loss of generality, we set the last $(n - m)$ components of $\boldsymbol{x}$ to 0:

$$x_{n-m+1} = 0, \ \ldots, \ x_n = 0.$$

- By stacking these equation below $\boldsymbol{Ax} = \boldsymbol{b}$, where $\boldsymbol{A} = [\boldsymbol{B}, \boldsymbol{D}]$, we get

$$\underbrace{\begin{bmatrix} \boldsymbol{B} & \boldsymbol{D} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}}_{\boldsymbol{M}} \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_D \end{bmatrix} = \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}$$

  $\boldsymbol{M} \in \mathbb{R}^{n \times n}$ has n rows.

- If the rows of $\boldsymbol{M}$ are linearly independent (if and only if $\boldsymbol{B} \in \mathbb{R}^{m \times m}$ has full rank), then $\boldsymbol{x}$ is uniquely determined as

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_D \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}.$$

- Therefore, setting some $(n - m)$ components of $\boldsymbol{x}$ to 0 may uniquely determine $\boldsymbol{x}$.

Now, to select an extreme point $\boldsymbol{x}$ of $P = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$, we

- set some $(n - m)$ components of $\boldsymbol{x}$ as 0
- let $\boldsymbol{x}_B$ denote the remaining components and $\boldsymbol{B}$ denote the corresponding submatrix of the matrix $\boldsymbol{A}$
- check whether $\boldsymbol{B}$ has full rank. If not, then not getting a point.
  If yes, then compute
  $$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_D \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}.$$
  Furthermore, check if $\boldsymbol{x}_B \geq 0$. If not, then $\boldsymbol{x} \notin P$; if yes, then $\boldsymbol{x}$ is an extreme point of P.

Fundamental theorem of LP

Theorem

Consider an LP in the standard form.

1. *If it has a feasible solution, then there exists a basic feasible solution (BFS).*
2. *If it has an* ***optimal*** *feasible solution, then there exists an* ***optimal*** *BFS.*

Proof of Part 1

- Suppose $\boldsymbol{x} = [x_1, \ldots, x_n]^T$ is a feasible solution. Thus $\boldsymbol{x} \geq 0$.
- WOLG, suppose that only the first p entries of $\boldsymbol{x}$ are positive, so

$$x_1\mathbf{a}_1 + \cdots + x_p\mathbf{a}_p = \boldsymbol{b} \tag{1}$$

- Case 1: if $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p$ are linearly independent, then $p \leq m = \mathrm{rank}(\boldsymbol{A})$.
  a. if $p = m$, then $\boldsymbol{B} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m]$ forms a basis and $\boldsymbol{x}$ is the BFS
  b. if $p < m$, then we can find $m - p$ columns from $\mathbf{a}_{p+1}, \ldots, \mathbf{a}_n$ to form[1] basis $\boldsymbol{B}$ and $\boldsymbol{x}$ is the BFS.
- Case 2: if $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p$ are linearly dependent, then $\exists\ y_1, \ldots, y_p$ such that some $y_i > 0$ and $y_1\boldsymbol{a}_1 + \cdots + y_p\boldsymbol{a}_p = \mathbf{0}$. From this and (**??**), we get

$$(x_1 - \epsilon y_1)\mathbf{a}_1 + \cdots + (x_p - \epsilon y_p)\mathbf{a}_p = \boldsymbol{b}$$

[1] This is a consequence of the so-called *matroid* structure of linear independence.

- for sufficiently small $\epsilon > 0$, $(x_1 - \epsilon y_1) > 0, \ldots, (x_p - \epsilon y_p) > 0$.
- since there is some $y_i > 0$, set

$$\epsilon = \min \Big\{ \frac{x_i}{y_i} \mid i = 1, \ldots p, \;\; y_i > 0 \Big\}.$$

  Then, the first p components of $(\boldsymbol{x} - \epsilon \boldsymbol{y}) \geq 0$ and at least one of them is 0. Therefore, we have reduced p by at least 1.

- by repeating this process, we either reach Case 1 or $p = 0$. The latter situation can be handled by Case 1 as well.
- therefore, part 1 is proved.
- **Part 2 can be similarly proved except we argue that $\boldsymbol{c}^T \boldsymbol{y} = 0$.**

Extreme points $\iff$ BFSs

Theorem

Let $P = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}, \boldsymbol{x} \geq 0\}$, where $\boldsymbol{A}^{m \times n}$ has full row rank. Then, $\boldsymbol{x}$ is an extreme point of P if and only if $\boldsymbol{x}$ is a BFS to P.

Proof: "$\Longrightarrow$" Let $\boldsymbol{x}$ be an extreme point of P. Suppose, WOLG, its first p components are positive. Let $\boldsymbol{y} = [y_1, \ldots, y_p]$ be such that

$$y_1\mathbf{a}_1 + \cdots + y_p\mathbf{a}_p = \mathbf{0}$$

$$x_1\mathbf{a}_1 + \cdots + x_p\mathbf{a}_p = \boldsymbol{b}.$$

Since $x_1, \ldots, x_p > 0$ by assumption, for sufficiently small $\epsilon > 0$, we have $\boldsymbol{x} + \epsilon\boldsymbol{y} \in P$ and $\boldsymbol{x} - \epsilon\boldsymbol{y} \in P$, which have $\boldsymbol{x}$ as their middle point. By the definition of extreme point, we must have $\boldsymbol{x} + \epsilon\boldsymbol{y} = \boldsymbol{x} - \epsilon\boldsymbol{y}$ and thus $\boldsymbol{y} = 0$. Therefore, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p$ are linearly independent.

"$\Longleftarrow$" Let $\boldsymbol{x} \in P$ be an BFS corresponding to the basis $\boldsymbol{B} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m]$.

Let $\boldsymbol{y}, \boldsymbol{z} \in P$ be such that

$$\boldsymbol{x} = \alpha \boldsymbol{y} + (1 - \alpha)\boldsymbol{z},$$

for some $\alpha \in (0, 1)$. We show $\boldsymbol{y} = \boldsymbol{z}$ and conclude that $\boldsymbol{x}$ is an extreme point.

Since $\boldsymbol{y} \geq 0$ and $\boldsymbol{z} \geq 0$ yet the last $(n - m)$ entries of $\boldsymbol{x}$ are 0, the last $(n - m)$ entries of $\boldsymbol{y}, \boldsymbol{z}$ are 0, too. Hence,

$$y_1 \mathbf{a}_1 + \cdots + y_p \mathbf{a}_m = \boldsymbol{b}$$
$$z_1 \mathbf{a}_1 + \cdots + z_p \mathbf{a}_m = \boldsymbol{b},$$

and thus $(y_1 - z_1)\mathbf{a}_1 + \cdots + (y_m - z_m)\mathbf{a}_m = \mathbf{0}$. Since $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ are linearly independent, $y_i - z_i = 0$ for all i and thus $\boldsymbol{y} = \boldsymbol{z}$. $\square$

Edge directions

- Edges have two functions:
  - reduced costs are defined for edge directions
  - we move from one BFS to another along an edge direction
- From now on, we assume *non-degeneracy*, i.e., any BFS $\boldsymbol{x}$ has its basic subvector $\boldsymbol{x}_B > 0$. The purpose: to avoid *edge of 0 length*.
- An edge has 1 degree of freedom and connects to at least one BFS.

An edge in $\mathbb{R}^n$ is obtained from an BFS by removing one equation from the n linearly independent equations that define the BFS.

- consider the BFS $\boldsymbol{x} = [\boldsymbol{x}_B; 0]^T$.

  let $\boldsymbol{A} = [\boldsymbol{B}, \boldsymbol{D}]$, $B = \{1, \ldots, m\}$ and $D = \{m+1, \ldots, n\}$.

- $\{\boldsymbol{x}\} = \{\boldsymbol{y} : \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}, y_i = 0 \ \forall i \in D\} \subset P$

- pick any $j \in D$, then

$$\left\{\boldsymbol{y} \geq 0 : \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}, y_i = 0 \ \forall i \in D \setminus \{j\}\right\} \subset P$$

  is the edge connected to $\boldsymbol{x}$ corresponding to $x_j \geq 0$

Bottomline: an edge is obtained by releasing a non-basic variable x_j from 0

- pick a non-basic coordinate j
- the **edge direction** for the BFS $\boldsymbol{x} = [\boldsymbol{x}_B; 0]^T$ corresponding to x_j is

$$\boldsymbol{\delta}^{(j)} = \begin{bmatrix} \boldsymbol{\delta}_B^{(j)} \\ \mathbf{0} \\ \boldsymbol{\delta}_j^{(j)} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{B}^{-1}\boldsymbol{a}_j \\ \mathbf{0} \\ 1 \\ \mathbf{0} \end{bmatrix},$$

  where satisfies $\boldsymbol{A}\boldsymbol{\delta}^{(j)} = \mathbf{0}$.

- we have

$$\left\{\boldsymbol{y} \geq 0 : \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}, y_i = 0 \ \forall i \in D \setminus \{j\}\right\} = \{\boldsymbol{y} \geq 0 : \boldsymbol{y} = \boldsymbol{x} + \epsilon\boldsymbol{\delta}^{(j)}, \epsilon \geq 0\}$$

- an non-degenerate BFS has $(n - m)$ different edges
- a degenerate BFS may have fewer edges

Reduced cost

- given the BFS $\boldsymbol{x} = [\boldsymbol{x}_B; 0]^T$, the non-basic coordinate j, and the edge direction $\boldsymbol{\delta}^{(j)} = [-\boldsymbol{B}^{-1}\boldsymbol{a}_j; \boldsymbol{0}; 1; \boldsymbol{0}]^T$
- the unit change in $\boldsymbol{c}^T\boldsymbol{x}$ along $\boldsymbol{\delta}^{(j)}$ is

$$\bar{c}_j = \boldsymbol{c}^T\boldsymbol{\delta}^{(j)} = c_j - \boldsymbol{c}_B^T\boldsymbol{B}^{-1}\boldsymbol{a}_j$$

- a negative reduced cost $\Rightarrow$ moving along $\boldsymbol{\delta}^{(j)}$ will decrease the objective
- define the reduced cost (row) vector: $\bar{\boldsymbol{c}}^T = [\bar{c}_1, \ldots, \bar{c}_n]$ as

$$\bar{\boldsymbol{c}}^T := \boldsymbol{c}^T - \boldsymbol{c}_B^T\boldsymbol{B}^{-1}\boldsymbol{A},$$

  which includes the reduced costs for all edge directions.

  Note that its basic part: $\bar{\boldsymbol{c}}_B^T = \boldsymbol{c}_B^T - \boldsymbol{c}_B^T\boldsymbol{B}^{-1}\boldsymbol{B} = \boldsymbol{0}$.

Optimal BFS

Theorem

Let $\boldsymbol{A} = [\boldsymbol{B}, \boldsymbol{D}]$, where $\boldsymbol{B}$ a basis. Suppose that $\boldsymbol{x} = [\boldsymbol{x}_B; \boldsymbol{0}]^T$ is an BFS. Then, $\boldsymbol{x}$ is optimal if and only if $\bar{\boldsymbol{c}}^T := \boldsymbol{c}^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{A} \geq \boldsymbol{0}^T$.

Proof: Let $\{\boldsymbol{\delta}^{(j)}\}_{j \in J}$ be the set of edge directions of $\boldsymbol{x}$. Then, $P = \{\boldsymbol{y} : \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}, \boldsymbol{y} \geq 0\}$ is a subset of $\boldsymbol{x} + \mathrm{cone}(\{\boldsymbol{\delta}^{(j)}\}_{j \in J})$, that is, any $\boldsymbol{y} \in P$ can be written as

$$\boldsymbol{y} = \boldsymbol{x} + \sum_{j \in J} \alpha_j \boldsymbol{\delta}^{(j)}$$

where $\alpha_j \geq 0$. Then

$$\boldsymbol{c}^T \boldsymbol{y} = \boldsymbol{c}^T \boldsymbol{x} + \sum_{j \in J} \alpha_j \boldsymbol{c}^T \boldsymbol{\delta}^{(j)} = \boldsymbol{c}^T \boldsymbol{x} + \sum_{j \in J} \alpha_j \bar{c}_j \geq \boldsymbol{c}^T \boldsymbol{x}.$$

□

The Simplex method (we have so far)

- **input:** a basis $\boldsymbol{B}$ and the corresponding BFS $\boldsymbol{x}$
- **check:** $\bar{\boldsymbol{c}}^T := \boldsymbol{c}^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{A} \geq \boldsymbol{0}^T$
- **if yes:** optimal, return $\boldsymbol{x}$; **stop.**
- **if not:** choose j such that $\bar{c}_j < 0$, and then move along $\boldsymbol{\delta}^{(j)}$
  - **if unbounded:** then the problem is **unbounded**
  - **otherwise:** replace $\boldsymbol{x}$ by the first BFS reached; restart

Next: unbounded and bounded edge directions

Unbounded edge direction and cost

- suppose at a BFS $\boldsymbol{x}$, we select $\bar{c}_j < 0$ and now move along

$$\boldsymbol{\delta}^{(j)} = \begin{bmatrix} \boldsymbol{\delta}_B^{(j)} \\ \mathbf{0} \\ \boldsymbol{\delta}_j^{(j)} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{B}^{-1}\boldsymbol{a}_j \\ \mathbf{0} \\ 1 \\ \mathbf{0} \end{bmatrix},$$

- since $\boldsymbol{A}\boldsymbol{\delta}^{(j)} = 0$ and $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}$, we have for any α

$$\boldsymbol{A}(\boldsymbol{x} + \alpha\boldsymbol{\delta}^{(j)}) = \boldsymbol{b}$$

- since $\boldsymbol{x} \geq 0$, if $\boxed{\boldsymbol{\delta}_B^{(j)} = -\boldsymbol{B}^{-1}\boldsymbol{a}_j \geq 0}$, then we have for any $\alpha \geq 0$,

$$\boldsymbol{x} + \alpha\boldsymbol{\delta}^{(j)} \geq 0$$

- therefore, $\boldsymbol{x} + \alpha\boldsymbol{\delta}^{(j)}$ is **feasible** for any $\alpha \geq 0$
- however, the cost is (lower) unbounded: $\boldsymbol{c}^T(\boldsymbol{x} + \alpha\boldsymbol{\delta}^{(j)}) = \boldsymbol{c}^T\boldsymbol{x} + \alpha\bar{c}_j$

Bounded edge direction

- if $\boxed{\boldsymbol{\delta}_B^{(j)} = -\boldsymbol{B}^{-1}\boldsymbol{a}_j \not\geq 0}$, then α must be sufficiently small; otherwise

$$\boldsymbol{x} + \alpha\boldsymbol{\delta}^{(j)} \not\geq 0$$

- ratio test

$$\alpha_{\min} = \min\left\{\frac{x_i}{-\delta_i^{(j)}} \mid i \in B,\ \delta_i^{(j)} < 0\right\}$$

- for some $i' \in B$, we have $x_{i'} + \alpha_{\min}\delta_{i'}^{(j)} = 0$
- the *nondegeneracy* assumption: $\boldsymbol{x}_B > 0$, thus $\alpha_{\min} > 0$
- let $\boldsymbol{x}' = \boldsymbol{x} + \alpha_{\min}\boldsymbol{\delta}^{(j)}$:
  - $\boldsymbol{x}'_B \geq 0$ but $x'_{i'} = 0$
  - $x'_j > 0$
  - $x'_i = 0$ for $i \notin B \cup \{j\}$
- **updated basis:** $B' = B \cup \{j\} \setminus \{i'\}$ and **BFS:** $\boldsymbol{x}' = \boldsymbol{x} + \alpha_{\min}\boldsymbol{\delta}^{(j)}$

The Simplex method (we have so far)

- **input:** a basis $\boldsymbol{B}$ and the corresponding BFS $\boldsymbol{x}$
- **check:** $\bar{\boldsymbol{c}}^T := \boldsymbol{c}^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{A} \geq \boldsymbol{0}^T$
- **if yes:** optimal, return $\boldsymbol{x}$; **stop.**
- **if not:** choose j such that $\bar{c}_j < 0$, and then move along $\boldsymbol{\delta}^{(j)}$
  - **if $\boldsymbol{\delta}_B^{(j)} \geq 0$:** then the problem is **unbounded**
  - **otherwise:**
    - $\alpha_{\min} = \min \left\{ \frac{x_i}{-\delta_i^{(j)}} \mid i \in B,\ \delta_i^{(j)} < 0 \right\}$ achieved at index i'
    - updated basis: $B \leftarrow B \cup \{j\} \setminus \{i'\}$
    - updated BFS: $\boldsymbol{x} \leftarrow \boldsymbol{x} + \alpha_{\min} \boldsymbol{\delta}^{(j)}$

Example 16.2

$$\begin{array}{rl} \text{maximize} & 2x_1 + 5x_2 \\ \text{subject to} & x_1 \le 4 \\ & x_2 \le 6 \\ & x_1 + x_2 \le 8 \\ & x_1, x_2 \ge 0. \end{array}$$

- introduce slack variables and reformulate to the standard form

$$\begin{array}{rcccccccc} \text{minimize} & -2x_1 & -5x_2 & -0x_3 & -0x_4 & -0x_5 & & \\ \text{subject to} & x_1 & & +x_3 & & & = & 4 \\ & & x_2 & & +x_4 & & = & 6 \\ & x_1 & +x_2 & & & +x_5 & = & 8 \\ & x_1, & x_2, & x_3, & x_4, & x_5 & \ge & 0. \end{array}$$

- **the starting basis** $B = \{3, 4, 5\}$, $\boldsymbol{B} = [\boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5]$, and **BFS**

$$\boldsymbol{x} = \boldsymbol{B}^{-1}\boldsymbol{b} = [0, 0, 4, 6, 8]^T$$

- reduced costs

$$\bar{\boldsymbol{c}}^T := \boldsymbol{c}^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1}\boldsymbol{A} = [-2, -5, 0, 0, 0]$$

  since $\bar{\boldsymbol{c}}_B^T = 0^T$, only need to compute $\bar{c}_j$ for $j \notin B$.

- since $\bar{c}_2 < 0$, current solution is no optimal and bring $j = 2$ into basis

- compute edge direction: $\boldsymbol{\delta}^{(j)} = [0, 1, -\boldsymbol{B}^{-1}\boldsymbol{a}_j]^T = [0, 1, 0, -1, -1]^T$
- $\boldsymbol{\delta}_B^{(j)} = [0, -1, -1]^T \not\geq 0$, so this edge direction is *not* unbounded
- ratio test

$$\alpha_{\min} = \min\left\{ \frac{x_i}{-\delta_i^{(j)}} \mid i \in B,\ \delta_i^{(j)} < 0 \right\} = \min\left\{\frac{6}{1}, \frac{8}{1}\right\} = 6$$

  the "min" is achieved at $i' = 4$, so remove 4 from basis

- **updated basis**: $B = \{2, 3, 5\}$, $\boldsymbol{B} = [\boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_5]$ and **BFS**:

$$\boldsymbol{x} = \boldsymbol{B}^{-1}\boldsymbol{b} = [0, 6, 4, 0, 2]^T$$

- current basis: $B = \{2, 3, 5\}$, $\boldsymbol{B} = [\boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_5]$ and BFS $\boldsymbol{x} = [0, 6, 4, 0, 2]^T$
- reduced costs

$$\bar{\boldsymbol{c}}^T := \boldsymbol{c}^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{A} = [-2, 0, 0, 5, 0]$$

- since $\bar{c}_1 < 0$, current solution is not optimal, and bring $j = 1$ into basis
- compute edge direction: $\boldsymbol{\delta}^{(j)} = [1, 0, -1, 0, -1]$
- ratio test:

$$\alpha_{\min} = \min \left\{ \frac{x_i}{-\delta_i^{(j)}} \mid i \in B,\ \delta_i^{(j)} < 0 \right\} = \min \left\{ \frac{4}{1}, \frac{2}{1} \right\} = 2$$

  the "$\min$" is achieved at $i' = 5$, so remove 5 from basis
- **updated basis**: $B = \{1, 2, 3, \}$, $\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3]$ and **BFS**:

$$\boldsymbol{x} = \boldsymbol{B}^{-1} \boldsymbol{b} = [2, 6, 2, 0, 0]^T$$

- updated reduced costs: $\bar{\boldsymbol{c}}^T = [0, 0, 0, 3, 2] \geq 0$, the solution is optimal.

Row operations on the canonical augmented matrix

- Starting

$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\boldsymbol{a}_3$	$\boldsymbol{a}_4$	$\boldsymbol{a}_5$	$\boldsymbol{b}$
1	0	1	0	0	4
0	1	0	1	0	6
1	1	0	0	1	8

- Itr 1: $j = 2$ in and $i' = 4$ out. Make "new $\boldsymbol{a}_j$ = old $\boldsymbol{a}_{i'}$" through row ops.

$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\boldsymbol{a}_3$	$\boldsymbol{a}_4$	$\boldsymbol{a}_5$	$\boldsymbol{b}$
1	0	1	0	0	4
0	1	0	1	0	6
1	0	0	-1	1	2

- Itr 2: $j = 1$ in and $i' = 5$ out. Make "new $\boldsymbol{a}_j$ = old $\boldsymbol{a}_{i'}$" through row ops.

$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\boldsymbol{a}_3$	$\boldsymbol{a}_4$	$\boldsymbol{a}_5$	$\boldsymbol{b}$
0	0	1	1	-1	2
0	1	0	1	0	6
1	0	0	-1	1	2

Finite convergence

Theorem

Consider a LP in the standard form

$$\text{minimize}\, \boldsymbol{c}^T \boldsymbol{x} \quad \text{subject to } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}.$$

Suppose that the LP is feasible and all BFSs are nondegenerate. Then,

- *the Simplex method terminates after a finite number of iterations*
- *at termination, we either have an optimal basis B or a direction $\boldsymbol{\delta}$ such that $\boldsymbol{A}\boldsymbol{\delta} = \boldsymbol{0}$, $\boldsymbol{\delta} \geq 0$, and $\boldsymbol{c}^T\boldsymbol{\delta} < 0$. In the former case, the optimal cost is finite, and in the latter case it has unbounded optimal cost of $-\infty$.*

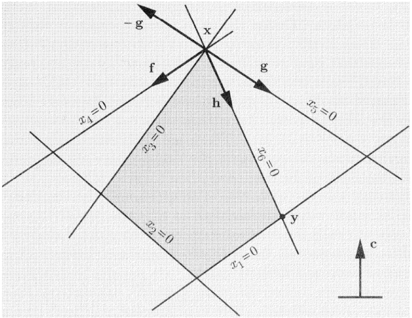
Degeneracy

- $\boldsymbol{x}$ is degenerate if some components of $\boldsymbol{x}_B$ equals 0
- geometrically, more than n active linear constraints at $\boldsymbol{x}$
- consequence: the ratio test

$$\alpha_{\min} = \min\left\{\frac{x_i}{-\delta_i^{(j)}} \mid i \in B,\ \delta_i^{(j)} < 0\right\}$$

  may return $\alpha_{\min} = 0$, then causing
  - $\boldsymbol{x} + \alpha_{\min}\boldsymbol{\delta}^{(j)} = \boldsymbol{x}$, BFS remains unchanged
  - no improvement in cost
  - the basis *does* change
- then, finite termination is no longer guaranteed, revisiting a previous basis is possible. there are a number of remedies to avoid cycling.
- further, a tie in the ratio test will cause the updated BFS to be degenerate

Degeneracy illustration



- $n - m = 2$, we are looking at the 2D affine set $\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$
- $B = \{1, 2, 3, 6\}$, BFS $\boldsymbol{x}$ determined by $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $x_4, x_5 = 0$; $\boldsymbol{x}$ is degenerate
- x_4, x_5 are non-basic, have edge directions $\mathbf{g}$ and $\mathbf{f}$, respectively, both make the ratio tests return 0 and do not improve cost
- after change to basis $\{1, 2, 3, 5\}$, then non-basic x_4, x_6 have edge directios $\mathbf{h}$ and $\mathbf{f}$. $\mathbf{h}$ can improve cost, $\mathbf{f}$ cannot.

The Simplex method (we have so far)

- **input:** a basis $\boldsymbol{B}$ and the corresponding BFS $\boldsymbol{x}$
- **check:** $\bar{\boldsymbol{c}}^T := \boldsymbol{c}^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{A} \geq \boldsymbol{0}^T$
- **if yes:** optimal, return $\boldsymbol{x}$; **stop.**
- **if not:** choose j such that $\bar{c}_j < 0$, and then move along $\boldsymbol{\delta}^{(j)}$
  - **if $\boldsymbol{\delta}_B^{(j)} \geq 0$:** then the problem is **unbounded**
  - **otherwise:**
    - $\alpha_{\min} = \min \left\{ \frac{x_i}{-\delta_i^{(j)}} \mid i \in B,\ \delta_i^{(j)} < 0 \right\}$ achieved at index i'
    - updated basis: $B \leftarrow B \cup \{j\} \setminus \{i'\}$
    - updated BFS: $\boldsymbol{x} \leftarrow \boldsymbol{x} + \alpha_{\min} \boldsymbol{\delta}^{(j)}$

      (if $\alpha_{\min} = 0$, anti-cycle schemes are applied)

Remaining question: how to find the initial BFS?

An easy case

- suppose the original LP has the constraints

$$\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \quad \boldsymbol{x} \geq \boldsymbol{0},$$

  where $\boldsymbol{b} \geq \boldsymbol{0}$.

- add slack variables

$$\begin{bmatrix} \boldsymbol{A}, & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{bmatrix} = \boldsymbol{b}, \quad \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{bmatrix} \geq \boldsymbol{0}.$$

- an obvious basis is $\boldsymbol{I}$

- corresponding BFS

$$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{b} \end{bmatrix}.$$

Two-phase Simplex method

- generally, given a standard-form LP, neglecting $\boldsymbol{c}^T\boldsymbol{x}$ for the moment:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x} \geq \boldsymbol{0}$$

- step 1. ensure $\boldsymbol{b} \geq 0$, by $\times(-1)$ to rows with $b_i < 0$ (this would fail to work for inequality constraints)
- step 2. add *artificial variables*, $y_1, \ldots, y_m$ and minimize their sum

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{y}}{\text{minimize}} \ & y_1 + \cdots + y_m \\ \text{subject to} \ & \begin{bmatrix} \boldsymbol{A}, & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \boldsymbol{b}, \quad \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} \geq \boldsymbol{0}. \end{aligned}$$

- step 3. Phase I: solve this LP, starting with the obvious basis $\boldsymbol{I}$ and BFS
- step 4. Phase II: when $\boldsymbol{y} = 0$ and becomes non-basic, we get a basis from $\boldsymbol{A}$ and a BFS in $\boldsymbol{x}$. then, drop $\boldsymbol{y}$, run the Simplex method with $\min \boldsymbol{c}^T\boldsymbol{x}$.

Proposition (16.1)

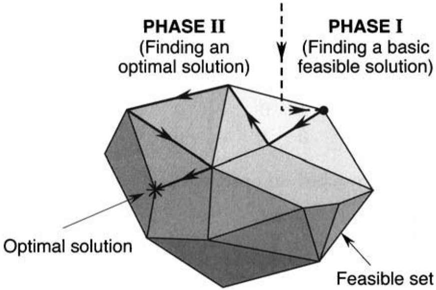
The original LP has a BFS if and only if the phase-I LP has an optimal BFS with 0 objective.

Proof.

"$\Rightarrow$" If the original LP has a BFS $\boldsymbol{x}$, the $[\boldsymbol{x}^T, \mathbf{0}^T]^T$ is a BFS to the phase-I LP. Clearly, this solution gives a 0 objective. Since the phase-I LP is feasible and lower bounded by 0, this solution is optimal to the phase-I LP.

"$\Leftarrow$" Suppose the phase-I LP has an optimal BFS. Then, this solution must have the form $[\boldsymbol{x}^T, \mathbf{0}^T]^T$, where $\boldsymbol{x} \geq \mathbf{0}$. Hence, we have $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq 0$. The original LP has a feasible solution. By the fundamental theorem of LP, there exists a BFS. □

Two-phase Simplex method: illustration



$n - m = 3$, each 2D face corresponds to some $x_i = 0$

Summary

The Simplex method:

- traverses through a set of BFS (vertices of the feasible set)
- each BS corresponds to a basis $\boldsymbol{B}$ and has the form $[\boldsymbol{x}_B^T, \boldsymbol{0}^T]^T$
- $\boldsymbol{x}_B \geq 0 \Rightarrow$ BFS. moreover, $\boldsymbol{x}_B > 0 \Rightarrow$ non-degenerate BFS
- leaves an BFS along an edge direction with a negative reduced cost
- an edge direction may be unbounded, or reaches $x_{i'} = 0$ for some $i' \in B$
- each iteration, some j enters basic and some i' leaves basis
- with anti-cycle schemes, the Simplex stops in a finite number of iterations
- the phase-I LP checks feasibility and, if feasible, obtains a BFS in $\boldsymbol{x}$

Uncovered topics

- **big-M method:** another way to obtain BFS or check feasibility
- **anti-cycle scheme**: special pivot rules that introduce an order to the basis
- Simplex method on the **tableau**
- **revised Simplex method**: maintaining $[\boldsymbol{B}^{-1}, \boldsymbol{\delta}_B]$ at a low cost
- **dual Simplex method**: maintaining nonnegative reduced cost $\bar{\boldsymbol{c}} \geq 0$ and work *toward* the feasibility $\boldsymbol{x} \geq \mathbf{0}$
- **column generation**: in large-scale problem, add c_i and $\boldsymbol{a}_i$ on demand
- **sensitivity analysis**: answer "what if" questions: $c, b, \boldsymbol{A}$ change slightly
- **network flow problems**: combinatorial problems with exact LP relaxation and fast algorithms

Quiz

Consider a general polyhedron $Q = \{x \in \mathbb{R}^n : Ax = b, Bx \geq d\}$.

Definition 1: x is a basic feasible solution of Q if $x \in Q$ and there are n linearly independent active constraints at x.

Consider the standard-form polyhedron $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, where A has linearly-independent rows.

Definition 2: x is a basic feasible solution of P if $A = [B\ D]$ where B is a full-rank square matrix and $x_B = B^{-1}b \geq 0$ and $x_D = 0$.

Show that the two definitions are equivalent for P.

Quiz

Consider the standard-form polyhedron $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, where A has linearly-independent rows. Prove:

1. If two bases B_1 and B_2 lead to the same basic solution x, then x must be degenerate.
2. Suppose x is a degenerate basic solution. Then, it corresponds to two or more distinct bases? Prove or give a counter example.
3. Suppose x is a degenerate basic solution. Then, there exists an adjacent basic solution that is generate. Prove or give a counter example.

Quiz

Consider the standard-form polyhedron $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, where $A \in \mathbb{R}^{m \times n}$ has linearly-independent rows. Prove the following statements or provide counter examples:

1. If $m + 1 = n$, then P has at most two basic feasible solutions.
2. The set of all solutions, if exist, is bounded.
3. At every solution, no more than m variables can be positive.
4. If there is more than one solution, then there are uncountably many solutions.
5. Consider minimizing $\max\{c^T x, d^T x\}$ over P. If this problem has a solution, it must have an solution that is an extreme point of P.

Quiz

Q1: Let $x \in P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$. Show that $d \in \mathbb{R}^n$ is a feasible direction at x if and only if $Ad = 0$ and $d_i \geq 0$ for every i such that $x_i = 0$.

Q2: Let $x \in P = \{x \in \mathbb{R}^n : Ax = b, Dx \leq f, Ex \leq g\}$ such that $Dx = f$ and $Ex < g$. Show that $d \in \mathbb{R}^n$ is a feasible direction at x if and only if $Ad = 0$ and $Dd \leq 0$.