The Simplex method

Instructor: Jin Zhang

Department of Mathematics
Southern University of Science and Technology
Spring 2024

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)

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

- common assumption: $\mathrm{rank}(A) = m$, full row rank or A is surjective (otherwise, either $Ax = b$ has no solution or some rows of A can be safely eliminated)
- write A as
$$A = [B, D]$$
where B is a square matrix with full rank (its rows/columns are linearly independent). This might require reordering the columns of A.
- We call $x = [x_B^T, 0^T]^T$ a basic solution if $Bx_B = b$. x_B and B are called basic variables and basis.

Basic feasible solution(BFS)

- "basic" because x_B is uniquely determined by B and b
- more definitions:
  - if $x \geq 0$ (equivalently, $x_B \geq 0$), then x is a basic feasible solution (BFS)
  - if any entry of x_B is 0, then 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)

$$A = [a_1, a_2, a_3, a_4] = \begin{bmatrix} 1 & 1 & -1 & 4 \\ 1 & -2 & -1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

1. Pick $B = [a_1, a_2]$, obtain $x_B = B^{-1}b = [6, 2]^T$.
   $x = [6, 2, 0, 0]^T$ is a basic feasible solution and is nondegenerate.
2. Pick $B = [a_3, a_4]$, obtain $x_B = B^{-1}b = [0, 2]^T$.
   $x = [0, 0, 0, 2]^T$ is a degenerate basic feasible solution.
   In this example, $B = [a_1, a_4]$ and $[a_2, a_4]$ also give the same x!
3. Pick $B = [a_2, a_3]$, obtain $xB = B^{-1}b = [2, -6]^T$.
   $x = [0, 2, -6, 0]^T$ is a basic solution but infeasible, violating $x \geq 0$.
4. $x = [3, 1, 0, 1]^T$ is feasible but not basic.

The total number of possible basic solutions is at most

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

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

Any vector x, basic or not, that yields the minimum value of $c^T x$ over the feasible set $\{x : Ax = b,\ x \geq 0\}$ 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 $P = \{x : Ax = b, x \geq 0\}$ is given by n linearly independent and active (i.e., "$=$") constraints obtained from
  - $Ax = b$ (m linear constraints)
  - $x = 0$ (n linear constraints)
- In a standard form LP, if we assume $\mathrm{rank}(A) = m < n$, then $Ax = b$ give just m linearly independent constraints. Therefore, we need to
  - select addition $(n-m)$ linear constraints from $x_1 \geq 0, \cdots, x_n \geq 0$ 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 x to 0:

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

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

$$\underbrace{\begin{bmatrix} B & D \\ 0 & I \end{bmatrix}}_{M} \begin{bmatrix} x_B \\ x_D \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

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

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

$$x = \begin{bmatrix} x_B \\ x_D \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}$$

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

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

- set some $(n - m)$ components of x as 0
- let x_B denote the remaining components and B denote the corresponding submatrix of the matrix A
- check whether B has full rank. If not, then not getting a point. If yes, then compute

  $$x = \begin{bmatrix} x_B \\ x_D \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}$$

  Furthermore, check if $x_B \geq 0$. If not, then $x \notin P$; if yes, then 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 $x = [x_1, \cdots, x_n]^T$ is a feasible solution. Thus $x \geq 0$.
- WOLG, suppose that only the first p entries of x are positive, so

$$x_1 a_1 + \cdots + x_p a_p = b$$

- Case 1: if $a_1, \cdots, a_p$ are linearly independent, then $p \leq m = \mathrm{rank}(A)$.
  a. if $p = m$, then $B = [a_1, \cdots, a_m]$ forms a basis and x is the BFS
  b. if $p < m$, then we can find $m - p$ columns from $a_{p+1}, \cdots, a_n$ to form basis B and x is the BFS.
- Case 2: if $a_1, \cdots, a_p$ are linearly dependent, then $\exists\ y_1, \cdots, y_p$ such that some $y_i > 0$ and $y_1 a_1 + \cdots + y_p a_p = 0$. From this, we get

$$(x_1 - \epsilon y_1)a_1 + \cdots + (x_p - \epsilon y_p)a_p = b$$

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

$$\epsilon = \min\{\frac{x_i}{y_i} | i = 1, \cdots p, y_i > 0\}$$

  Then, the first p components of $(x - \epsilon y) \geq 0$ 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 $c^T y = 0$.

Extreme points $\Longleftrightarrow$ BFSs

Theorem

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

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

$$y_1 a_1 + \cdots + y_p a_p = 0,$$
$$x_1 a_1 + \cdots + x_p a_p = b.$$

Since $x_1, \cdots, x_p > 0$ by assumption, for sufficiently small $\epsilon > 0$, we have $x + \epsilon y \in P$ and $x - \epsilon y \in P$, which have x as their middle point. By the definition of extreme point, we must have $x + \epsilon y = x - \epsilon y$ and thus $y = 0$. Therefore, $a_1, \cdots, a_p$ are linearly independent.

"$\Leftarrow$" Let $x \in P$ be an BFS corresponding to the basis $B = [a_1, \cdots, a_m]$. Let $y,\ z \in P$ be such that $x = \alpha y + (1-\alpha)z$ for some $\alpha \in (0,1)$. We show $y = z$ and conclude that x is an extreme point. Since $y \geq 0$ and $z \geq 0$ yet the last $(n-m)$ entries of x are 0, the last $(n-m)$ entries of y, z are 0, too. Hence,

$$\begin{aligned} y_1 a_1 + \cdots + y_p a_m &= b, \\ z_1 a_1 + \cdots + z_p a_m &= b, \end{aligned}$$

and thus $(y_1 - z_1)a_1 + \cdots + (y_m - z_m)a_m = 0$. Since $a_1, \cdots, a_m$ are linearly independent, $y_i - z_i = 0$ for all i and thus $y = z$. ■

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 x has its basic subvector $x_B > 0$. 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 $x = [x_B; 0]^T$.
  let $A = [B, D], B = \{1, \cdots, m\}$ and $D = \{m+1, \cdots, n\}$.
- $\{x\} = \{y : Ay = b, y_i = 0, \forall i \in D\} \subset P$
- pick any $j \in D$, then

$$\{y \geq 0 : Ay = b, y_i = 0, \forall i \in D \backslash j\} \subset P$$

  is the edge connected to 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 $x = [x_B; 0]^T$ corresponding to x_j is

$$\delta^{(j)} = \begin{bmatrix} \delta_B^{(j)} \\ 0 \\ \delta_j^{(j)} \\ 0 \end{bmatrix} = \begin{bmatrix} -B^{-1}a_j \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

  where satisfies $A\delta^{(j)} = 0$.
- we have

$$\{y \geq 0 : Ay = b, y_i = 0 \ \forall i \in D \backslash \{j\}\} = \{y \geq 0 : y = x + \epsilon\delta^{(j)}, \epsilon \geq 0\}$$

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

Reduced cost

- given the BFS $x = [x_B; 0]^T$, the non-basic coordinate j, and the edge direction $\delta(j) = [-B^{-1}a_j, 0, 1, 0]^T$

- the unit change in $c^T x$ along $\delta^{(j)}$ is
$$\bar{c}_j = c^T \delta^{(j)} = c_j - c_B^T B^{-1} a_j$$

- a negative reduced cost $\Rightarrow$ moving along $\delta^{(j)}$ will decrease the objective

- define the reduced cost (row) vector: $\bar{c}^T = [\bar{c}_1, \cdots, \bar{c}_n]$ as
$$\bar{c}^T := c^T - c_B^T B^{-1} A,$$
which includes the reduced costs for all edge directions.
Note that its basic part: $\bar{c}_B^T = c_B^T - c_B^T B^{-1} B = 0$.

Optimal BFS

Theorem

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

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

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

where $\alpha_j \geq 0$. Then

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

■

The Simplex method (we have so far)

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

Next: unbounded and bounded edge directions

Unbounded edge direction and cost

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

$$\delta^{(j)} = \begin{bmatrix} \delta_B^{(j)} \\ 0 \\ \delta_j^{(j)} \\ 0 \end{bmatrix} = \begin{bmatrix} -B^{-1}a_j \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

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

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

- since $x \geq 0$, if $\delta_B^{(j)} = -B^{-1}a_j \geq 0$, then we have for any $\alpha \geq 0$,

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

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

Bounded edge direction

- if $\delta_B^{(j)} = -B^{-1}a_j \ngeq 0$, then α must be sufficiently small; otherwise

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

- ratio test

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

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

The Simplex method (we have so far)

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

Example 16.2

$$
\begin{aligned}
\max \quad & 2x_1 + 5x_2 \\
\text{subject to } & x_1 \leq 4 \\
& x_2 \leq 6 \\
& x_1 + x_2 \leq 8 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

- introduce slack variables and reformulate to the standard form

$$
\begin{aligned}
\min \quad & -2x_1 - 5x_2 - 0x_3 - 0x_4 - 0x_5 \\
\text{subject to } & x_1 + x_3 = 4 \\
& x_2 + x_4 = 6 \\
& x_1 + x_2 + x_5 = 8 \\
& x_1, x_2, x_3, x_4, x_5 \geq 0
\end{aligned}
$$

- the starting basis $B = \{3, 4, 5\},\ B = [a_3, a_4, a_5]$, and BFS

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

- reduced costs

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

since $\bar{c}_B^T = 0^T$, 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: $\delta^{(j)} = [0, 1, -B^{-1}a_j]^T = [0, 1, 0, -1, -1]^T$
- $\delta_B^{(j)} = [0, -1, -1]^T \ngeq 0$, so this edge direction is not unbounded
- ratio test

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

  the "min" is achieved at $i' = 4$, so remove 4 from basis
- updated basis: $B = \{2, 3, 5\}, B = [a_2, a_3, a_5]$ and BFS:

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

- current basis: $B = \{2, 3, 5\}, B = [a_2, a_3, a_5]$ and BFS $x = [0, 6, 4, 0, 2]^T$
- reduced costs

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

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

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

  the "min" is achieved at $i' = 5$, so remove 5 from basis
- updated basis: $B = \{1, 2, 3\}, B = [a_1, a_2, a_3]$ and BFS:

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

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

Finite convergence

Theorem

Consider a LP in the standard form

$$\min\ c^T x \text{ subject to } Ax = b, x \geq 0.$$

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 δ such that $A\delta = 0, \delta \geq 0$, and $c^T\delta < 0$. In the former case, the optimal cost is finite, and in the latter case it has unbounded optimal cost of $-\infty$.*

Degeneracy

- x is degenerate if some components of x_B equals 0
- geometrically, more than n active linear constraints at x
- consequence: the ratio test

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

  may return $\alpha_{\min} = 0$, then causing
  - $x + \alpha_{\min}\delta^{(j)} = x$, 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

The Simplex method (we have so far)

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

$$Ax \leq b, \;\; x \geq 0,$$

  where $b \geq 0$.

- add slack variables

$$[A, I] \begin{bmatrix} x \\ s \end{bmatrix} = b, \;\; \begin{bmatrix} x \\ s \end{bmatrix} \geq 0.$$

- an obvious basis is I
- corresponding BFS

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

Summary

The Simplex method:

- traverses through a set of BFS (vertices of the feasible set)
- each BS corresponds to a basis B and has the form $[x_B^T, 0^T]^T$
- $x_B \geq 0 \Rightarrow$BFS. moreover, $x_B > 0 \Rightarrow$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 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 $[B^{-1}, \delta B]$ at a low cost
- dual Simplex method: maintaining nonnegative reduced cost $\bar{c} \geq 0$ and work toward the feasibility $x \geq 0$
- column generation: in large-scale problem, add c_i and a_i on demand
- sensitivity analysis: answer "what if" questions: c, b, A change slightly
- network flow problems: combinatorial problems with exact LP relaxation and fast algorithms