Numerical Optimization

Unit 8 Linear Programming and the Simplex Method

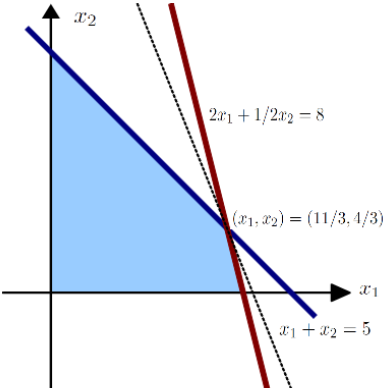
Che-Rung Lee

Department of Computer Science
National Tsing Hua University

November 4, 2022

Example problem

$$\begin{aligned} \min_{x_1, x_2} \quad & z = -4x_1 - 2x_2 \\ \text{s.t.} \quad & x_1 + x_2 \leq 5 \\ & 2x_1 + 1/2x_2 \leq 8 \\ & x_1, x_2 \geq 0 \end{aligned}$$



Matrix formulation

$$
\begin{aligned}
\min_{x_1,x_2} \quad & z = -4x_1 - 2x_2 \\
\text{s.t.} \quad & x_1 + x_2 \leq 5 \\
& 2x_1 + 1/2x_2 \leq 8 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

- Let $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\vec{c} = \begin{pmatrix} -4 \\ -2 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 1 \\ 2 & 1/2 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$.
- The problem can be written as

$$
\begin{aligned}
\min_{\vec{x}} \quad & \vec{c}^T \vec{x} \\
\text{s.t.} \quad & A\vec{x} \leq \vec{b} \\
& \vec{x} \geq 0
\end{aligned}
$$

The standard form

$$
\begin{aligned}
\min_{\vec{x}} \quad & z = \vec{c}^T \vec{x} \\
\text{s.t.} \quad & A\vec{x} = \vec{b} \\
& \vec{x} \geq 0
\end{aligned}
$$

- z : Objective function.
- $\vec{c}$: Cost vector $\in \mathbb{R}^n$
- A : Constraint matrix $\in \mathbb{R}^{m \times n}$, assuming $m \leq n$
- $A\vec{x} = \vec{b}$: Linear equality constraints.
- The i_{th} constraint is $\sum_{j=1}^{n} a_{ij} x_j = b_i$

Converting to the standard form

- Change inequality constraints to equality constraints:

$$\begin{aligned} x_1 + x_2 + x_3 &= 5 \\ 2x_1 + \frac{1}{2}x_2 + x_4 &= 8 \end{aligned}$$

- x_3 and x_4 are called *slack variables*.
- As a result,

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \vec{c} = \begin{pmatrix} -4 \\ -2 \\ 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1/2 & 0 & 1 \end{pmatrix}, \vec{b} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$$

Rules to converting to standard form

1. If $\sum_{j=1}^{n} a_{ij}x_j \leq b_j \quad \Rightarrow \quad$ adding a slack variable $s_i \geq 0$

   $\sum_{j=1}^{n} a_{ij}x_j + s_i = b_i$.

2. If $\sum_{j=1}^{n} a_{ij}x_j \geq b_j \quad \Rightarrow \quad$ adding a surplus variable $e_i \geq 0$

   $\sum_{j=1}^{n} a_{ij}x_j - e_i = b_i$.

3. If $x_i \geq l_i \quad \Rightarrow \quad x_i = \hat{x}_i + l_i \ , \ \hat{x}_i \geq 0$.

4. If $x_i \leq u_i \quad \Rightarrow \quad x_i = u_i - \hat{x}_i \ , \ \hat{x}_i \geq 0$.

5. If $x_i \in \mathbb{R} \quad \Rightarrow \quad x_i = \bar{x}_i - \hat{x}_i \ , \ \bar{x}_i \geq 0 \ , \ \hat{x}_i \geq 0$.

6. For the problem $\max_{\vec{x}} \vec{c}^T \vec{x} \quad \Rightarrow \quad -\min_{\vec{x}} -\vec{c}^T \vec{x}$.

Some terminology

- Feasible set: $\mathcal{F} = \{\vec{x} \in \mathbb{R}^n | A\vec{x} = \vec{b}, \vec{x} \geq 0\}$.
- If $\mathcal{F} \neq \emptyset$, the problem is feasible or consistent.
- If $\mathcal{F} = \emptyset$, the problem is infeasible.
- If $\vec{c}^T\vec{x} \geq \alpha$ for all $\vec{x} \in \mathcal{F}$, the problem is bounded.
- If the solution is at infinity, the problem is unbounded.
- The problem may have infinity number of solutions.
- Hyperplane $H = \{\vec{x} \in \mathbb{R}^n | \vec{a}^T\vec{x} = \beta\}$ whose normal is $\vec{a}$
- Closed half space $H = \{\vec{x} \in \mathbb{R}^n | \vec{a}^T\vec{x} \leq \beta\}$ or $H = \{\vec{x} \in \mathbb{R}^n | \vec{a}^T\vec{x} \geq \beta\}$
- Polyhedral set or polyhedron (polygon): A set of the intersection of finite closed half spaces.
- Poly tope: nonempty and bounded polyhedron.

Convex set

Let $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_p \in \mathbb{R}^n$ and $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R}$.

Linear combination	$\vec{y} = \alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 \ldots + \alpha_p\vec{x}_p$
Affine combination	$\vec{y} = \alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 \ldots + \alpha_p\vec{x}_p$ and $\alpha_1 + \alpha_2 + \ldots + \alpha_p = 1$
Convex combination	$\vec{y} = \alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 \ldots + \alpha_p\vec{x}_p$ and $0 \le \alpha_1, \alpha_2, \ldots \alpha_p \le 1$ and $\alpha_1 + \alpha_2 + \ldots + \alpha_p = 1$
Cone combination	$\vec{y} = \alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 \ldots + \alpha_p\vec{x}_p$ and $\alpha_1, \alpha_2, \ldots, \alpha_p \ge 0$

For a set $S \subset \mathbb{R}^n, S \neq \emptyset$, if $\forall \vec{x}_1, \vec{x}_2 \in S$ s.t. the affine(convex) combination of $\vec{x}_1, \vec{x}_2$ are in S, we say S is a affine(convex) set.

The simplex method

Basic idea

1. Find a “vertex” of the poly-tope.
2. Find the best direction and move to the next “vertex” (pricing).
3. Test optimality of the “vertex”.

Basic feasible point

- A vertex $\vec{x}$ in the polytope C is called a basic feasible point.
- Geometrically, $\vec{x}$ is not a convex combination of any other point in C.
- Algebraically, $A\vec{x} = \vec{b}$, the columns of A corresponding to the positive elements of $\vec{x}$ are linearly independent.
- Theorem: at least one of the solution is the basic feasible point.
- Which means we only need to search those basic feasible points.
- For m hyperplanes in an n dimensional space, $m \geq n$, the intersection of any n hyperplanes can be a basic feasible point. Therefore, we have $C_n^m = \dfrac{m!}{n!(m-n)!}$ points to check.
  - For $m = 2n$, $C_n^{2n} > 2^n$. The time complexity of doing so is exponential!
  - We need a systematical way to solve this.

Basic variables and nonbasic variables

- We need to find an intersection of n hyperplanes, whose normal vectors are linearly independent. (why?)
- Partition $A = [B|N]$ where B is invertible.

Example

For $A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1/2 & 0 & 1 \end{pmatrix}$, we let $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $N = \begin{pmatrix} 1 & 1 \\ 2 & 1/2 \end{pmatrix}$

- Partition $\vec{x} = \begin{bmatrix} \vec{x}_B \\ \vec{x}_N \end{bmatrix}$ accordingly.

Example

Based on the above partition, $\vec{x}_B = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}$, $\vec{x}_N = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Compute the basic feasible point

- Let $\vec{x}_N = 0$ and solve $B\vec{x}_B = \vec{b}$
  - $\vec{x}_B$ is called the "basic variables"
  - $\vec{x}_N$ is the "nonbasic variables"
- $\vec{x} = \begin{bmatrix} B^{-1}\vec{b} \\ \vec{0} \end{bmatrix}$ is a basic feasible point. (why?)

Example

$$\vec{x} = \begin{pmatrix} x_3 \\ x_4 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \\ 0 \\ 0 \end{pmatrix}.$$ (Where is this point?)

Compute the search direction

- Rewrite the object function z as a function of nonbasic variables.

$$A = [B|N] \text{ and } A\vec{x} = \vec{b}$$

  which implies $B\vec{x}_B + N\vec{x}_N = \vec{b}$.
- Let $\vec{x}_B = B^{-1}(\vec{b} - N\vec{x}_N)$ and substitute it to z.

$$\begin{aligned}
z_{k+1} &= \vec{c}^T\vec{x} \\
&= \vec{c}_B^T\vec{x}_B + \vec{c}_N^T\vec{x}_N \\
&= \vec{c}_B^T B^{-1}(\vec{b} - N\vec{x}_N) + \vec{c}_N^T\vec{x}_N \\
&= (-c_B^T B^{-1}N + \vec{c}_N^T)\vec{x}_N + \vec{c}_B^T B^{-1}\vec{b} \\
&= \vec{p}^T\vec{x}_N + \vec{c}_B^T B^{-1}\vec{b}
\end{aligned}$$

  Now z has only nonbasic variables.

Pricing vector

- The vector $\vec{p} = \vec{c}_N - N^T (B^{-1})^T \vec{c}_B$ is called the *pricing vector*.
- Since all nonbasic variables are zero at this time, if x_i's coefficient (the ith element of $\vec{p}$) is negative, then by increasing x_i's value, we can decrease z's value.
- What if all the elements in $\vec{p}$ are positive?
- If there are more than one elements in $\vec{p}$ are negative, which nonbasic variable x_i should be chosen to increase its value?

Example

At this point, $z = -4x_1 - 2x_2$. We choose to increase x_1.

Search direction

Let the ith element of $\vec{x}_N$, denoted ν_i, be the chosen element to be increased. What is the search direction?

- Since all the constraints need be satisfied, to increase ν_i implies to change some basic variables.
- How to find this relation?

$$\begin{aligned} A\vec{x} &= \vec{b} \\ B\vec{x}_B + N\vec{x}_N &= \vec{b} \\ \vec{x}_B &= B^{-1}(\vec{b} - N\vec{x}_N) \end{aligned}$$

- Let the ith column of N be $\vec{n}_i$.

$$\vec{x}_B = B^{-1}(\vec{b} - \nu_i\vec{n}_i).$$

- When ν_i is increased by 1, the change of $\vec{x}_B$ is $-B^{-1}\vec{n}_i$ (because $B^{-1}\vec{b}$ are their current values.).
- Other $\vec{x}_N$ elements remain the same. (why?)

Search direction

- The search direction is

$$\vec{d} = \begin{pmatrix} -B^{-1}\vec{n}_i \\ \vec{0} \\ 1 \\ \vec{0} \end{pmatrix} \begin{array}{l} \leftarrow \text{Basic variables} \\ \leftarrow \text{Other nonbasic variables} \\ \leftarrow \text{The index of } \nu_i \\ \leftarrow \text{Other nonbasic variables} \end{array}$$

Example

We choose x_1 to increase its value. The 1st column of A is $(1 \;\; 2)^T$. Therefore, $-B^{-1}\vec{n}_1 = (-1 \;\; -2)^T$.

$$\vec{d} = \begin{pmatrix} d_3 \\ d_4 \\ d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \end{pmatrix}$$

Step length

How large can the step length be?

- The only constraint of changing those basic variables is to keep them nonnegative.
- Let α be the step length.

$$\vec{x}_B^{(new)} = \vec{x}_B^{(now)} + \alpha\vec{d} = B^{-1}\vec{b} + \alpha\vec{d} \geq 0$$

- The ratio test: the only basic variables that we care are those whose $\vec{d}$ elements are negative. (why?)

$$\alpha = \min_{x_j \in \vec{x}_B, d_j < 0} |x_j / d_j| \,. \tag{1}$$

- What if all d_js are positive?

Example

Since d_3 and d_4 are all negative, and $x_3 = 5, x_4 = 8$,

$$\alpha = \min(|-5/1|, |-8/2|) = 4.$$

Move to the next location

- If everything goes well, there will be one nonbasic variable ν_i becomes positive, and one basic variable x_j becomes zero.
- We exchange those two variables. Let ν_i be a basic variable and let x_j be a nonbasic variable.
- This process continues until the optimal solution is found. (How to know the optimal solution?)

Example

$x_3 = 5 + (-1) * 4 = 1.$
$x_4 = 8 + (-2) * 4 = 0$ becomes nonbasic and $x_1 = 4$ becomes basic.

The simplex method

The simplex method for linear programming

1. Let $\mathcal{B}, \mathcal{N}$ be the index set of basic variables and nonbasic variables.
2. For $k = 1, 2, \ldots$
    1. $B = A(:, \mathcal{B}), N = A(:, \mathcal{N})$, $\vec{x}_B = B^{-1}b$, and $\vec{x}_N = 0$.
    2. Solve $B^T\vec{v} = \vec{c}_B$
    3. Compute $\vec{p} = \vec{c}_N - N^T\vec{v}$.
    4. If $\vec{p} \geq 0$, stop (the optimal solution found)
    5. Select $i \in \mathcal{N}$ with $\vec{p}(i) < 0$.
    6. Solve $B\vec{s} = A(:, i)$
    7. If $\vec{s} < 0$, stop (unbounded)
    8. Calculate α using (1) and assume the index of zeroed basic variable is j.
    9. Update $\vec{x}_B^+ = \vec{x}_B - \alpha\vec{s}$, $\vec{x}_N = (0, \ldots, \alpha, \ldots, 0)^T$.
    10. Update $\mathcal{B}$ and $\mathcal{N}$ by exchanging index i and j.

Time complexity

- The worst case time complexity of the Simplex method is still exponential. But practically, only $O(n)$ iterations are required.
- This phenomenon has been analyzed by Daniel A. Spielman and Shang-Hua Teng, and they win the Godel prize in 2008.
- See their paper for details: *Smoothed Analysis of Algorithms: Why the Simplex Algorithm Usually Takes Polynomial Time*.
- There are polynomial-time algorithms for the linear programming problems.
  - 1981: Leonid Khachiyan(Ellipsoid method)
  - 1984: Narendra Karamarker(Interior point method), which will be discussed.

Lower bound of the answer

Question: Before we solve the problem, can we use the constraints to estimate the "lower bound" of $z(\vec{x})$?

Example

$$
\begin{aligned}
\min_{x_1, x_2} \quad & z = 5x_1 + 8x_2 & \\
\text{s.t.} \quad & x_1 + 2x_2 \geq 4 & (1)\\
& x_1 + 1/2x_2 \geq 2 & (2)\\
& x_1, x_2 \geq 0 &
\end{aligned}
$$

- From (1), $z_x = 5x_1 + 8x_2 \geq 4x_1 + 8x_2 = 4(x_1 + 2x_2) = 16$
- From (2), $z_x = 5x_1 + 8x_2 \geq 5x_1 + \frac{5}{2}x_2 = 5(x_1 + \frac{1}{2}x_2) = 10$
- From the combination of (1) and (2),
  $z_x = 5x_1 + 8x_2 \geq 5x_1 + 7.75x_2 = 3.5(x_1 + 2x_2) + 1.5(x_1 + \frac{1}{2}x_2) = 17$

Maximum lower bound

- What is the "maximum lower bound" of z from constraints?
- We multiply y_1 to (1) and multiply y_2 to (2), and add them together.

$$
\begin{array}{rrcl}
 & (x_1 + 2x_2)y_1 & \geq & 4y_1 \\
+) & (x_1 + \frac{1}{2}x_2)y_2 & \geq & 2y_2 \\
\hline
\multicolumn{2}{r}{(y_1 + y_2)x_1 + (2y_1 + \frac{1}{2}y_2)x_2} & \geq & 4y_1 + 2y_2
\end{array}
$$

- The problem of maximizing the lower bound becomes

$$
\begin{array}{ll}
\max_{y_1, y_2} & 4y_1 + 2y_2 \\
\text{s.t.} & y_1 + y_2 \leq 5 \\
 & 2y_1 + \frac{1}{2}y_2 \leq 8 \\
 & y_1, y_2 \geq 0
\end{array}
$$

  which is called the *dual problem* of the original problem.
- The original problem is called the *primal problem*.

The primal and the dual problem.

The primal and the dual

Primal problem	Dual problem
$\begin{array}{ll} \min_{\vec{x}} & \vec{c}^T\vec{x} \\ \text{s.t.} & A\vec{x} \geq \vec{b} \\ & \vec{x} \geq 0 \end{array}$	$\begin{array}{ll} \max_{\vec{y}} & \vec{b}^T\vec{y} \\ \text{s.t.} & A^T\vec{y} \leq \vec{c} \\ & \vec{y} \geq 0 \end{array}$

Example

Primal problem	Dual problem
$\begin{array}{ll} \min_{x_1,x_2} & 5x_1 + 8x_2 \\ \text{s.t.} & x_1 + 2x_2 \geq 4 \\ & x_1 + \frac{1}{2}x_2 \geq 2 \\ & x_1, x_2 \geq 0 \end{array}$	$\begin{array}{ll} \max_{y_1,y_2} & 4y_1 + 2y_2 \\ \text{s.t.} & y_1 + y_2 \leq 5 \\ & 2y_1 + \frac{1}{2}y_2 \leq 8 \\ & y_1, y_2 \geq 0 \end{array}$

Duality

Theorem (The weak duality)

If $\vec{x}$ is feasible for the primal problem and $\vec{y}$ is feasible for the dual problem , then

$$\vec{y}^T\vec{b} \leq \vec{y}^T A\vec{x} \leq \vec{c}^T\vec{x}.$$

Theorem (The strong duality)

If $\vec{x}^*$ is the optimal solution of the primal. If $\vec{y}^*$ is the optimal solution of the dual. Then

$$\vec{c}^T\vec{x}^* = \vec{b}^T\vec{y}^*$$

Moreover, if the primal (dual) problem is unbounded, the dual (primal) is infeasible.

Properties of the optimal solution

Example

Primal problem	Dual problem
$\min_{x_1,x_2}$ $5x_1 + 8x_2$ s.t. $x_1 + 2x_2 \geq 4$ $x_1 + \frac{1}{2}x_2 \geq 2$ $x_1 + \frac{1}{5}x_2 \geq 1$ $x_1, x_2 \geq 0$	$\max_{y_1,y_2}$ $4y_1 + 2y_2 + y_3$ s.t. $y_1 + y_2 + y_3 \leq 5$ $2y_1 + \frac{1}{2}y_2 + \frac{1}{5}y_3 \leq 8$ $y_1, y_2, y_3 \geq 0$

- The optimal solution of the primal is $52/3$, which happens at $(x_1^*, x_2^*) = (4/3, 4/3)$;
- At the primal optimal solution, the first two constrains hold the equality. But the last constrain does not.
- The optimal solution of the dual is at the point $(y_1^*, y_2^*, y_3^*) = (11/3, 4/3, 0)$;

Complementary slackness

Given a feasible point, an inequality constraint is called active if its equality holds. Otherwise it is called inactive.

Theorem (Complementary slackness)

$\vec{x}^*$ and $\vec{y}^*$ are optimal solution of the primal and the dual problem if and only if

1. *For $j = 1, 2, \ldots, n$, $A(;, j)^T \ \vec{y}^* = c_j$ or $x_j^* = 0$*
2. *For $i = 1, 2, \ldots, m$, $A(i, ;)\vec{x}^* = b_i$ or $y_i^* = 0$*

If we add slack variables $\vec{s}$ to $A\vec{x} + \vec{s} = \vec{b}$, the above theorem can be rewritten as

- If a constraint i is active , $s_i = 0$.
- If a constraint i is inactive , $s_i > 0$.
- The complementarity slackness condition is $y_i^* s_i^* = 0$ for all i.