Lecture 6: The Simplex Method

Zizhuo Wang

Institute of Data and Decision Analytics (iDDA)
Chinese University of Hong Kong, Shenzhen

Sep 21, 2018

Announcements

- Homework 2 due on Wednesday, Sep 26th

Recap

We started to study structures of linear optimization

- We solved an LP via graph
- Observations are made about solution structures
- One important observation: Optimal solution tends to be at corner points (extreme points)

Definition (Extreme Points)

Let P be a polyhedron. A point $\mathbf{x} \in P$ is said to be an extreme point of P if we cannot find two vectors $\mathbf{y}, \mathbf{z} \in P$, both different from $\mathbf{x}$, and a scalar $\lambda \in [0, 1]$, such that $\mathbf{x} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{z}$

- That is, $\mathbf{x}$ cannot be represented as a convex combination of other points in P. (Here in polyhedron case, two points are equivalent to more points.)

Recap: Basic Solutions

We defined the basic solution and showed how to find them:

1. Choose any m independent columns of A: $A_{B(1)}, \ldots, A_{B(m)}$
2. Let $x_i = 0$ for all $i \neq B(1), \ldots, B(m)$
3. Solve the equation $A\mathbf{x} = \mathbf{b}$ for the remaining $x_{B(1)}, \ldots, x_{B(m)}$.

We call $B = \{B(1), \ldots, B(m)\}$ the basic indices for this basic solution, $A_{B(1)}, \ldots, A_{B(m)}$ the basic columns, A_B the basis matrix and $x_{B(1)}, \ldots, x_{B(m)}$ the basic variables

- In a basic solution, there are no more than m positive entries
- There are finitely many basic solutions for a given LP

Recap: Basic Feasible Solutions

Definition

If a basic solution $\mathbf{x}$ also satisfies that $\mathbf{x} \geq 0$, then we call it a *basic feasible solution* (BFS).

Theorem

For the standard LP polyhedron $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$, *the followings are equivalent:*

1. $\mathbf{x}$ *is an extreme point*
2. $\mathbf{x}$ *is a basic feasible solution*

- This theorem connects the geometric property of LP to the algebraic property

Recap: Basic Feasible Solutions

Theorem

(LP fundamental theorem) Given a linear programming problem where A has full row rank m

- *If there is a feasible solution, there is a basic feasible solution;*
- *If there is an optimal solution, there is an optimal solution that is a basic feasible solution.*

Corollary

In order to find an optimal solution, we only need to look among basic feasible solutions.

Corollary

If an LP with m constraints (in the standard form) has an optimal solution, then there must be an optimal solution such that there is no more than m positive entries.

Recap: Basic Feasible Solutions

In theory, one could enumerate all the BFSs to find the optimal solution. But it is not practical (too slow).

We want something smarter:

- Simplex method
- Inventor: George Dantzig

Idea of the Simplex method:

- Start from one BFS, either 1) find a neighboring BFS that can improve over the current one, or 2) stop.

We need to define more formally what it means by a neighboring solution

Neighboring BFS

Definition

Two basic solutions are *neighboring* (or adjacent) if they differ by exactly one basic (or non-basic) index

For example, a BFS constructed by using columns $\{1, 2, 3\}$ is a neighbor to the BFS constructed by using columns $\{1, 3, 5\}$, but not $\{1, 4, 5\}$.

Example

Consider the production planning problem:

$$\begin{array}{rrrrrrl}
\text{minimize} & -x_1 & -2x_2 & & & & \\
\text{subject to} & x_1 & & +s_1 & & & = 100 \\
 & & 2x_2 & & +s_2 & & = 200 \\
 & x_1 & +x_2 & & & +s_3 & = 150 \\
 & x_1, & x_2, & s_1, & s_2, & s_3 & \geq 0
\end{array}$$

We can list the basic (feasible) solutions

Indices	$\{1,2,3\}$	$\{1,2,4\}$	$\{1,2,5\}$	$\{1,3,4\}$
Solution	(50, 100, 50, 0, 0)	(100, 50, 0, 100, 0)	(100, 100, 0, 0, -50)	(150, 0, -50, 200, 0)
Objective	-250	-200	infeasible	infeasible
Indices	$\{1,4,5\}$	$\{2,3,4\}$	$\{2,3,5\}$	$\{3,4,5\}$
Solution	(100, 0, 0, 200, 50)	(0, 150, 100, -100, 0)	(0, 100, 100, 0, 50)	(0, 0, 100, 200, 150)
Objective	-100	infeasible	-200	0

The other two choices $\{1,3,5\}$ and $\{2,4,5\}$ lead to dependent basic columns (therefore no basic solutions can be obtained)

Example

Indices	$\{1,2,3\}$	$\{1,2,4\}$	$\{1,2,5\}$	$\{1,3,4\}$
Solution	(50, 100, 50, 0, 0)	(100, 50, 0, 100, 0)	(100, 100, 0, 0, -50)	(150, 0, -50, 200, 0)
Objective	-250	-200	infeasible	infeasible
Indices	$\{1,4,5\}$	$\{2,3,4\}$	$\{2,3,5\}$	$\{3,4,5\}$
Solution	(100, 0, 0, 200, 50)	(0, 150, 100, -100, 0)	(0, 100, 100, 0, 50)	(0, 0, 100, 200, 150)
Objective	-100	infeasible	-200	0

Say we start from $\{3,4,5\}$ and only allow to move from one BFS to its neighbor (and also we move only if the objective value decreases). Then

- If we choose $\{1,4,5\}$ first, then we can move along

$$\{3,4,5\} \to \{1,4,5\} \to \{1,2,4\} \to \{1,2,3\}$$

- We can also start from $\{2,3,5\}$, which will give us

$$\{3,4,5\} \to \{2,3,5\} \to \{1,2,3\}$$

- Both get to the optimal solution in finite steps.

Checking Neighbors

The main question is how to find a neighboring BFS that improves the current one (reduces the objective function)

- We only need to show how to do it for one step. Then we can simply iterate the same methods (until we reach optimal).

A naive way is simply to check all neighbors of the current BFS:

- There are $m(n-m)$ potential neighbors for a BFS (choosing one index to leave, choosing one to enter). For each one, one needs to solve a linear equation of m variables
- This works, but is likely to be very slow

We want a more efficient way to do it

First we assume that we have somehow found a BFS whose basis is $B(1), ..., B(m)$.

Define

$$A_B = [A_{B(1)}, A_{B(2)}, ..., A_{B(m)}]$$

and let A_N be the matrix consist of the non-basic columns of A.

By rearranging the order of variables, we can write $A = [A_B, A_N]$, $\mathbf{x} = [\mathbf{x}_B; \mathbf{x}_N]$, where $\mathbf{x}_B$ are the basic variables, and $\mathbf{x}_N$ are the non-basic variables.

By definition, we have

$$\mathbf{x}_B = A_B^{-1}\mathbf{b} \qquad \mathbf{x}_N = 0$$

Find a Neighboring BFS

Now we want to find a neighboring BFS.

- We want to select a non-basic variable x_j to *enter* the basis.
- This means, we want to increase x_j from the current BFS

To do this, we consider moving $\mathbf{x}$ (current BFS) to $\mathbf{x} + \theta \mathbf{d}$ where

1. $d_j = 1$
2. $d_{j'} = 0$ for all other non-basic indices

What constraints do we have on $\mathbf{d}$?

- We need to guarantee that the resulting solution $\mathbf{x} + \theta \mathbf{d}$ is still feasible, that is:

$$A(\mathbf{x} + \theta \mathbf{d}) = \mathbf{b} = A\mathbf{x}$$

i.e., $A\mathbf{d} = 0$.

Basic Directions

Now we write $\mathbf{d} = [\mathbf{d}_B; \mathbf{d}_N]$. Since we know that $d_j = 1$ and $d_{j'} = 0$ for all other non-basic indices. We have

$$A_B \mathbf{d}_B + A_j = 0$$

Therefore

$$\mathbf{d}_B = -A_B^{-1} A_j$$

That means, the entire $\mathbf{d}$ will be uniquely determined:

$$\mathbf{d} = [\mathbf{d}_B; \mathbf{d}_N] = [-A_B^{-1} A_j; 0; ...; 1; ...; 0]$$

where the 1 is at the jth entry. We call such $\mathbf{d}$ the jth basic direction.

Basic Directions

The jth basic direction only guarantees that the equality constraint $A\mathbf{x} = \mathbf{b}$ still holds. We also have the constraint that $\mathbf{x} \geq 0$. We need to verify that this still holds (i.e., $\mathbf{x} + \theta\mathbf{d} \geq 0$):

- For non-basic variables, it was 0 and the change in $\mathbf{d}$ is non-negative (either 1 or 0), therefore, it is still positive
- For basic variables, as long as they were strictly positive, there must exist a small θ such that $\mathbf{x} + \theta\mathbf{d} \geq 0$

Typically the basic variables of a BFS are all positive (i.e., m positive entries), there are cases that some basic variables in a BFS are equal to 0 (degeneracy), we will discuss those cases later

Change in the Objective Value

Remember the objective function of the original LP is $\mathbf{c}^T\mathbf{x}$. We can similarly decompose $\mathbf{c}$ into basic and non-basic parts, corresponding to the basic/non-basic indices, i.e.,

$$\mathbf{c} = [\mathbf{c}_B, \mathbf{c}_N]$$

Now we study what happens to the objective function when we move from the BFS $\mathbf{x}$ to $\mathbf{x} + \theta\mathbf{d}$. The change is $\theta\mathbf{c}^T\mathbf{d}$.

We have (for $\mathbf{d}$ to be the jth basic direction):

$$\mathbf{c}^T\mathbf{d} = c_j - \mathbf{c}_B^T A_B^{-1} A_j \doteq \bar{c}_j$$

We call $\bar{c}_j$ the *reduced cost* of variable x_j.

Reduced Costs

The reduced cost

$$\bar{c}_j = c_j - \mathbf{c}_B^T A_B^{-1} A_j$$

is a very important concept in the simplex method. It corresponds to the change of the objective value if one tries to change the basis.

- The first term is the cost per unit increase in the variable x_j
- The second term is the cost of the compensating change in the basic variables necessitated by the constraints $A\mathbf{x} = \mathbf{b}$

Reduced Costs

Given the current basis and the index that one wants to add to the basis (j), the reduced costs can be easily computed:

- A positive reduced cost means that incorporating j into the current basis will increase the objective function (not good);
- A negative reduced cost means that incorporating j will reduce the objective function (we want to go in that direction)

Thus the reduced costs are the indicators of where we want to go.

Reduced Costs

We have found the reduced costs are:

$$\bar{c}_j = c_j - \mathbf{c}_B^T A_B^{-1} A_j$$

This is the cost change when we try to add j to the new basis.

What if j is a basic variable, i.e., $j = B(i)$?

$$\bar{c}_{B(i)} = c_{B(i)} - \mathbf{c}_B^T A_B^{-1} A_{B(i)} = c_{B(i)} - c_B^T \mathbf{e}_i = c_{B(i)} - c_{B(i)} = 0$$

Therefore, the reduced costs for basic variables are zero ($\mathbf{e}_i$ is a vector with 1 at ith position and 0s otherwise).

Reduced Costs

Theorem (Stopping Criterion)

Consider a basic feasible solution $\mathbf{x}$ *associated with the basis* $B(1), ..., B(m)$ *and let* $\bar{\mathbf{c}}$ *be the corresponding vector of reduced costs. If* $\bar{\mathbf{c}} \geq 0$*, then* $\mathbf{x}$ *must be optimal.*

Remark

This theorem gives a stopping criterion to the simplex algorithm: We stop when all the reduced costs are non-negative.

- It also means that if one could not find a neighbor solution that is better, then one must have already achieved optimal solution.

Proof

Suppose at a certain basic feasible solution $\mathbf{x}$, all the reduced costs $\bar{\mathbf{c}} \geq 0$. Then consider any feasible solution $\mathbf{y}$, and $\mathbf{d} = \mathbf{y} - \mathbf{x}$. We know that since both $\mathbf{x}$ and $\mathbf{y}$ are feasible, we must have $A\mathbf{d} = 0$. We can further write it as:

$$A_B \mathbf{d}_B + \sum_{i \in N} A_i d_i = 0.$$

Here N is the set of non-basic indices. Thus

$$\mathbf{d}_B = -\sum_{i \in N} A_B^{-1} A_i d_i$$

Now we study the change of objective value, we have

$$\mathbf{c}^T \mathbf{d} = \mathbf{c}_B^T \mathbf{d}_B + \sum_{i \in N} c_i d_i = \sum_{i \in N} (c_i - c_B^T A_B^{-1} A_i) d_i = \sum_{i \in N} \bar{c}_i d_i$$

We know that all the $\bar{c}_i$'s are nonnegative, also all the d_i must be nonnegative too since $\mathbf{y} \geq 0$. Therefore, the objective value at $\mathbf{y}$ must be at least as large as the objective value at $\mathbf{x}$. $\square$

Example of Reduced Cost

Consider the production problem in standard form:

$$\begin{array}{rrrrrrl} \text{minimize} & -x_1 & -2x_2 & & & & \\ \text{subject to} & x_1 & & +s_1 & & & = 100 \\ & & 2x_2 & & +s_2 & & = 200 \\ & x_1 & +x_2 & & & +s_3 & = 150 \\ & x_1, & x_2, & s_1, & s_2, & s_3 & \geq 0 \end{array}$$

If we are at basis $\{1, 2, 3\}$ with BFS $(50, 100, 50, 0, 0)$. Then the reduced costs are

$$\bar{c}_4 = 0 - [-1, -2, 0] \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0.5$$

Similarly, $\bar{c}_5 = 1$. Therefore, the reduced costs are all nonnegative, thus the BFS is optimal.

Example of Reduced Cost Continued

$$\begin{array}{rrrrrrl} \text{minimize} & -x_1 & -2x_2 & & & & \\ \text{subject to} & x_1 & & +s_1 & & & = 100 \\ & & 2x_2 & & +s_2 & & = 200 \\ & x_1 & +x_2 & & & +s_3 & = 150 \\ & x_1, & x_2, & s_1, & s_2, & s_3 & \geq 0 \end{array}$$

If we are at basis $\{1, 4, 5\}$ with BFS $(100, 0, 0, 200, 50)$. Then the reduced costs are:

$$\bar{c}_2 = -2 - [-1, 0, 0] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = -2$$

Similarly, $\bar{c}_3 = 1$. Therefore including x_2 in the basis in the next step will reduce the objective value.