MAT 3007 – Optimization

The Simplex Method

Lecture 05 *June 16th*

Andre Milzarek iDDA / CUHK-SZ

Repetition

Recap

We studied basic properties of linear optimization problems:

- We observed that optimal solutions tend to be extreme points.
- To connect this geometric observation to an algebraic condition, we defined the notion of basic solutions.

How to find a basic solution:

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

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

Definition:

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

Theorem: Extreme Points and BFS

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

1. x is an extreme point.
2. x is a basic feasible solution.

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

Recap: Basic Feasible Solutions

Fundamental LP Theorem

Given an LP in standard form with A having full row rank m:

1. If the feasible set is nonempty, there is a basic feasible solution.
2. If there is an optimal solution, there is an optimal solution that is a basic feasible solution.

Remark:

- In order to find an optimal solution, we only need to consider basic feasible solutions.
- If an LP with m constraints (in standard form) has an optimal solution, then there must be an optimal solution such that there is no more than m positive entries.
- If an LP has a unique optimal solution, then it must be a BFS.

Recap: Basic Feasible Solutions

We can enumerate and check all the BFS to find the opt. solution. But this is not practical (too slow).

We want to try 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 function value, or 2) stop.

In the last lecture, we defined a neighbor solution (basis differs by one index):

- We still need to develop an efficient way to find the neighbors.

The Simplex Method

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

Define:

$$A_B = \begin{bmatrix} | & | & & | \\ A_{B(1)} & A_{B(2)} & \cdots & A_{B(m)} \\ | & | & & | \end{bmatrix}$$

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

Rearranging the variables, we can write $A = [A_B, A_N]$, $x = [x_B; x_N]$, where x_B are the basic variables, and x_N are the non-basic variables.

By definition, we have:

$$x_B = A_B^{-1} b \qquad x_N = 0.$$

Finding a Neighboring BFS

Now, we want to find a neighboring BFS:

- Finding a neighbor means changing one of the basic indices.
- 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.

We consider moving x (the current BFS) to $x + \theta d$, $\theta \geq 0$, where:

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

What constraints do we have on d?

- We need to guarantee that the resulting step $x + \theta d$ is still feasible, that is:

$$A(x + \theta d) = b = Ax,$$

i.e., $Ad = 0$.

Now, we write $d = [d_B; d_N]$. Since $d_j = 1$ and $d_{j'} = 0$ for all other non-basic indices, we have

$$0 = \begin{bmatrix} A_B & A_N \end{bmatrix} \begin{pmatrix} d_B \\ d_N \end{pmatrix} = A_B d_B + A_j.$$

Therefore

$$d_B = -A_B^{-1} A_j.$$

That means that the direction d is uniquely determined:

$$d = [d_B; d_N] = [-A_B^{-1} A_j; 0; \ldots; 1; \ldots; 0],$$

where the 1 is at the jth entry. We call such d the jth basic direction.

Moving along the jth basic direction currently only guarantees that the equality constraint $Ax = b$ still holds.

We also need to consider the constraint $x \geq 0$. We need to verify that this still holds (i.e., $x + \theta d \geq 0$):

- For non-basic variables i, x_i was 0 and d_i is non-negative (either 1 or 0), therefore, $x_i + \theta d_i$ is still positive.
- For basic variables, as long as they were all strictly positive, there must exist a small θ such that $x + \theta d \geq 0$.

Observation:

- 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 $c^\top x$. We can similarly decompose c into basic and non-basic parts, corresponding to the basic/non-basic indices, i.e.,

$$c = [c_B; c_N]$$

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

If d is the jth basic direction, then we have:

$$c^\top d = c_j - c_B^\top A_B^{-1} A_j =: \bar{c}_j.$$

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

Reduced Costs

The reduced cost

$$\bar{c}_j = c_j - c_B^\top A_B^{-1} A_j$$

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 compensating the change in the basic variables necessitated by the constraints $Ax = b$.

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).

$\rightsquigarrow$ The reduced costs are the indicators of where we want to go!

Question: What is the reduced cost of a basic variable? Or what is $\bar{c}_j$ when $j = B(i)$?

$$\bar{c}_{B(i)} = c_{B(i)} - c_B^\top A_B^{-1} A_{B(i)} = c_{B(i)} - c_B^\top e_i = c_{B(i)} - c_{B(i)} = 0.$$

Therefore, the reduced costs for basic variables are zero (e_i is the ith unit vector with 1 at ith position and 0s otherwise).

Theorem: Stopping Criterion

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

Remark:

- This theorem gives a stopping criterion for the simplex algorithm: We stop when all the reduced costs are nonnegative.
- It also means that if we can not find a neighbor solution that improves the objective function, then we must have already found an optimal solution.

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 and the BFS is optimal.

Example: Reduced Costs – 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.

Observation: At a certain x, if the reduced costs satisfy $\bar{c}_j \geq 0$, then x is optimal. What if some $\bar{c}_j < 0$?

- It means that by bringing the jth variable (non-basic) into the basis, we can decrease the objective value. Thus, we want to go in that direction.

Assume d is the jth basic direction with $\bar{c}_j < 0$. Moving in this direction can reduce the objective. But how far can we go?

- We need to make sure that $x + \theta d \geq 0$.
- We also want to go as far as possible.
- Therefore, we choose:

$$\theta^* = \max\{\theta \geq 0 : x + \theta d \geq 0\}.$$

Choosing a Stepsize

We can move along d as much as

$$\theta^* = \max\{\theta \geq 0 : x + \theta d \geq 0\}$$

without violating any constraints. We consider two cases:

1. If $d \geq 0$, then $\theta^* = \infty$. In this case, one can go unlimitedly far without making the solution infeasible, while keeping the objective decreasing. Therefore, the original LP is unbounded.
2. If $d_i < 0$ for some i, then we can solve:

$$\theta^* = \min_{\{i : d_i < 0\}} -\frac{x_i}{d_i}.$$

Since $d_i \geq 0$ for $i \in N$, we can also write it as:

$$\theta^* = \min_{\{i \in B : d_i < 0\}} -\frac{x_i}{d_i}.$$

Moving to a New Basis

Assume θ^* is finite (otherwise the problem is unbounded), then we can move to another feasible point:

$$y = x + \theta^* d.$$

Let $B(\ell) \in \{B(1), ..., B(m)\}$ be the index with $\theta^* = -x_{B(\ell)}/d_{B(\ell)}$ (it is possible for multiple indices to achieve this $\rightsquigarrow$ we will discuss this later). Then we must have:

$$y_{B(\ell)} = x_{B(\ell)} + \theta^* d_{B(\ell)} = 0.$$

Thus, the basic variable $x_{B(\ell)}$ has become zero, whereas the non-basic variable x_j becomes positive (equal to θ^*). Meaning the basis has changed to

$$B(1), ..., B(\ell - 1), j, B(\ell + 1), ..., B(m).$$

An Iteration of the Simplex Method

Initialization: We start from a BFS x (with corresponding basis B).

1. We first compute the reduced costs $\bar{c}$:

$$\bar{c}_j = c_j - c_B^\top A_B^{-1} A_j.$$

   - If none of the reduced costs is negative, then x is optimal.
   - Otherwise choose some j such that $\bar{c}_j < 0$.

2. Compute the jth basic direction $d = [-A_B^{-1}A_j; 0; ...; 1; ...; 0]$.
   - If $d \geq 0$, then the problem is unbounded, i.e., the optimal value is $-\infty$.
   - Otherwise, compute $\theta^* = \min_{i \in B, d_i < 0}\{-\frac{x_i}{d_i}\}$.

3. Set $y = x + \theta^* d$. Then, the point y is the new BFS with index j replacing $B(\ell)$ in the basis, where $B(\ell)$ is the index attaining the minimum in θ^*. The objective function value is changed by $\theta^* c^\top d = \theta^* \bar{c}_j$.

4. Repeat these procedures.

Simplex Iteration: Discussion

Observation:

- The simplex iteration generates a new feasible point y that has a lower objective value or it stops with an optimal solution (if the reduced costs are nonnegative).
- Our discussion was based on the assumption $x_B > 0$; if there exists i with $x_{B(i)} = 0$ then this iteration might fail and the objective functions might stay the same.

$\rightsquigarrow$ We need to treat these degenerate cases!

Open Questions:

- Is the new point y a BFS? (We only know that it is feasible).
- Given several j with $\bar{c}_j < 0$, which jth basic direction should we choose?
- Suppose there are multiple ℓ with $\theta^* = -x_{B(\ell)}/d_{B(\ell)}$, how should we update the basic indices?

Simplex Iteration: Properties and Convergence

Theorem: Properties of y

Let x be a nondegenerate BFS ($x_B > 0$) with basic indices B and let $y = x + \theta^* d$ be generated by the simplex iteration. Then, y is a basic feasible solution associated with the basic indices

$$\{B(1), ..., B(\ell - 1), j, B(\ell + 1), ..., B(m)\}.$$

Theorem: Convergence

Assume that the feasible set is nonempty and that every BFS is nondegenerate. Then, the simplex method terminates after a finite number of iterations with the following options:

1. The method stops with an BFS which is an optimal solution.
2. We have found a vector d with $Ad = 0$, $d \geq 0$, and $c^\top d < 0$, and the optimal value is $-\infty$.

Degeneracy

We have seen that the objective value will strictly decrease after one simplex method iteration if the initial BFS is nondegenerate.

However, it is possible that the objective stays the same.

Since the change of the objective value (if x_j enters the basis) is $\theta^* \bar{c}_j$ and we have $\bar{c}_j < 0$, this can only happen if $\theta^* = 0$.

Recall that

$$\theta^* = \min_{\{i \in B \,|\, d_i < 0\}} -\frac{x_i}{d_i}.$$

If for some i's with $d_i < 0$, there is $x_i = 0$ then we obtain $\theta^* = 0$.

$\rightsquigarrow$ This may happen when there are 0s in the BFS x.

Degeneracy

We call a basic feasible solution x (non)degenerate if (none) some of the basic variables are 0.

- Degeneracy can happen. We need to consider what consequences it may have in the algorithm!

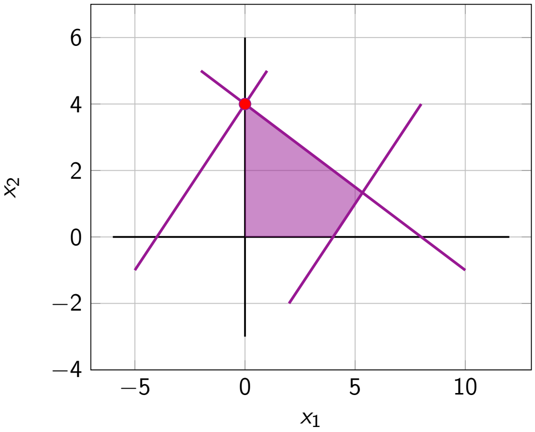
An example:

$$\begin{array}{rrrrrcl} x_1 & +2x_2 & +s_1 & & & = & 8 \\ x_1 & -x_2 & & +s_2 & & = & 4 \\ -x_1 & +x_2 & & & +s_3 & = & 4 \\ x_1, & x_2, & s_1, & s_2, & s_3 & \geq & 0 \end{array}$$

If we choose the basic indices to be $\{1, 2, 4\}$, then the corresponding basic solution is $(0, 4, 8)$. It is degenerate.

$\rightsquigarrow$ The number of non-zeros at the BS is strictly less than m.

Illustration



- More than 2 lines intersect at one point.
- In general, at a degenerate point, more than k planes intersect at a k-dimensional space.

Degeneracy

Assume we encounter degeneracy at some point:

- Given a BFS x with negative reduced cost $\bar{c}_j < 0$ and $\theta^* = 0$. And i is the index that achieves $\min_{\{i \in B, d_i < 0\}} (-x_i/d_i)$. Thus, $x_i = 0$.

We can still change the basic index from i (i leaving the basis) to j (j entering the basis) and proceed to the next iteration.

- Although the iterate and the objective value do not change, the basis changes. Therefore, the reduced costs vector will change in the next iteration – issue seems resolved?

We need to guarantee that there is not any cycle, i.e., we will not visit the same BFS more than once:

- This can only occur in case of degeneracy, since otherwise the objective value will strictly decrease!

Example of Cycling

If not dealt properly, cycling can happen. Consider the LP:

$$A = \begin{pmatrix} -2 & -9 & 1 & 9 & 1 & 0 \\ 1/3 & 1 & -1/3 & -2 & 0 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$c^\top = (-2, -3, 1, 12, 0, 0).$$

If we set $B = \{5, 6\}$ initially, then the sequence shown below leads to a cycle (objective value doesn't change, and there is always an index with negative reduced cost):

Step #	1	2	3	4	5	6
Exiting	x_6	x_5	x_2	x_1	x_4	x_3
Entering	x_2	x_1	x_4	x_3	x_6	x_5
Basis	$\{5, 2\}$	$\{1, 2\}$	$\{1, 4\}$	$\{3, 4\}$	$\{3, 6\}$	$\{5, 6\}$

We will show that cycle can be avoided by designing how to choose incoming/outgoing basis when there are multiple choices.

In the description of the algorithm, we said that at each feasible point, we can choose any j with negative reduced cost to enter the basis in the next iteration.

Sometimes, there are more than one j with $\bar{c}_j < 0$. In this case, we need to introduce some rules to choose the entering basis.

Here are several possible rules:

1. Smallest Index Rule: Choose the smallest index j with $\bar{c}_j < 0$.
2. Most Negative Rule: Choose the smallest $\bar{c}_j$.
3. Most Decrement Rule: Choose j with the smallest $\theta^* \bar{c}_j$.

Recall that

$$\theta^* = \min_{\{i \in B \mid d_i < 0\}} -\frac{x_i}{d_i}.$$

We choose one index that attains this minimum to leave the basis.

It is possible that there are two or more indices that attain the minimum (tie). Then we also need a rule to decide the new basis.

- The most commonly used rule is the smallest index rule.

When this tie happens, the next BFS will be degenerate. (Why?)

Theorem: Bland's Rule

If we use the smallest index rule for choosing both the entering basis and the exiting basis, then no cycle will occur in the simplex algorithm.

- Using the Bland's rule when applying the simplex method, we can guarantee to stop within a finite number of iterations at an optimal solution.

Questions?