The Simplex Method

Com S 477x/577x
Oct 22, 2002

To understand how simplex solves a general linear programming problem, we need to know the answers to a number of questions:

• How to determine if a minimum feasible solution has been found?
• Which basic variable should become non-basic at a pivot step?
• Which nonbasic variable is to enter the basis at a pivot step?
• How to find an initial basic feasible solution to start simplex?

We already had an answer to the first question, that is, checking if the coefficients of the new objective function are all nonnegative. Let us now look at the other three questions.

1 Determination of a Variable to Leave Basis

Suppose we know which nonbasic variable is to become basic, a topic we will focus on later. Then we can always determine which basic variable should become non-basic.

For convenience, let us assume the following holds.

Nondegeneracy assumption: In every basic feasible solution, no basic variable has zero value.\(^1\)

Given the constraints

\[
Ax = b, \quad x \geq 0, 
\]

suppose after zero or a few steps we have a tableau of the following form (without cost coefficients):

<table>
<thead>
<tr>
<th>(a'_1)</th>
<th>(a'_2)</th>
<th>(\ldots)</th>
<th>(a'_m)</th>
<th>(a'_{m+1})</th>
<th>(\ldots)</th>
<th>(a'_k)</th>
<th>(\ldots)</th>
<th>(a'_n)</th>
<th>(b')</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>(\ldots)</td>
<td>0</td>
<td>(a'_{1,m+1})</td>
<td>(\ldots)</td>
<td>(a'_{1,k})</td>
<td>(\ldots)</td>
<td>(a'_{1,n})</td>
<td>(b'_1)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>(\ldots)</td>
<td>0</td>
<td>(a'_{2,m+1})</td>
<td>(\ldots)</td>
<td>(a'_{2,k})</td>
<td>(\ldots)</td>
<td>(a'_{2,n})</td>
<td>(b'_2)</td>
</tr>
<tr>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\ddots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\ddots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>(\ldots)</td>
<td>1</td>
<td>(a'_{m,m+1})</td>
<td>(\ldots)</td>
<td>(a'_{m,k})</td>
<td>(\ldots)</td>
<td>(a'_{m,n})</td>
<td>(b'_m)</td>
</tr>
</tbody>
</table>

\(^1\)When this assumption is violated, a degenerate basic variable (with zero value) occurs in a basic feasible solution. Often it can be handled as a nondegenerate basic feasible solution. However, it is possible that at pivoting the new variable will come in at zero value. This implies that the zero-valued basic variable is the one to go out. The objective will not decrease and the new basic feasible solution will also be degenerate. The result is a cycle that could be repeated indefinitely. Methods have been developed to avoid such cycles [1, pp. 78].
The tableau represents a solution with basic variables \( x_1, x_2, \ldots, x_m \). We assume \( b'_1, b'_2, \ldots, b'_m \geq 0 \) so that the corresponding basic solution \( x_1 = b'_1, \ldots, x_m = b'_m \) is feasible.

Suppose we wish to bring into the basis the variable \( x_k \), \( k > m \), and maintain feasibility. Since

\[
\begin{align*}
\sum_{i=1}^{m} b'_i a'_i & = b' \\
\sum_{i=1}^{k} a'_i a'_k + \sum_{i=k+1}^{m} b'_i a'_i & = a'_k
\end{align*}
\]

we multiply (2) by some \( \epsilon \geq 0 \) and subtract from (1), obtaining

\[
(b'_1 - \epsilon a'_1) a'_1 + (b'_2 - \epsilon a'_2) a'_2 + \cdots + (b'_m - \epsilon a'_m) a'_m + \epsilon a'_k = b'.
\]  

There are two cases:

**Case 1**: There exists \( a'_{ik} > 0 \) for some \( 1 \leq i \leq m \). Then we set

\[
\epsilon = \min \left\{ \frac{b'_i}{a'_{ik}} \left| a'_{ik} > 0 \right. \right\}
\]

and pivot at \( a'_{jk} \), where \( j \) is the minimizing index. Then we will have a new basic feasible solution with \( x_k \) replacing \( x_j \). Note that only one single index \( j \) can achieve \( \epsilon \). Otherwise, one basic variable would become zero after the pivoting is performed to bring \( x_k \) into the basis, thereby violating the nondegeneracy assumption.

**Case 2**: All \( a'_{1k}, \ldots, a'_{mk} \) are negative. In this case, the coefficients in (3) increase as \( \epsilon \) is increased, and no new basic feasible solution is obtained. As a result, the solution of \( Ax = b \) can have arbitrarily large components. For instance, \( x_1 = b'_1 - \epsilon a'_{1k}, \ldots, x_m = b'_m - \epsilon a'_{mk}, x_k = \epsilon \), and \( x_j = 0 \) for \( k + 1 \leq j \leq n \) and \( j \neq k \), as \( \epsilon \) becomes arbitrarily large. So the set of feasible solutions is unbounded.

## 2 Determination of a Variable to Enter Basis

The idea of the simplex method is to select a column to pivot so that the resulting new basic feasible solution will yield a lower value to the objective function than the previous one.

Suppose the basic solution at the current pivot step is

\[
(x^T_B, 0^T) = (b'_1, b'_2, \ldots, b'_m, 0, \ldots, 0),
\]

and the corresponding tableau takes the form

\[
\begin{array}{cccccccc}
 a'_1 & a'_2 & \cdots & a'_m & a'_{m+1} & \cdots & a'_j & \cdots & a'_n & b' \\
 1 & 0 & \cdots & 0 & a'_{1,m+1} & \cdots & a'_{1,j} & \cdots & a'_{1,n} & b'_1 \\
 0 & 1 & \cdots & 0 & a'_{2,m+1} & \cdots & a'_{2,j} & \cdots & a'_{2,n} & b'_2 \\
 \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
 0 & 0 & \cdots & 1 & a'_{m,m+1} & \cdots & a'_{m,j} & \cdots & a'_{m,n} & b'_m \\
\end{array}
\]

Suppose the objective function is

\[
z = c'_1 x_1 + c'_2 x_2 + \cdots + c'_n x_n;
\]  

\[2\]
for the basic solution, it has the value

\[ z_0 = c_B x_B, \]

where \( c_B = (c'_1, c'_2, \ldots, c'_m) \).

If arbitrary values are assigned to \( x_{m+1}, \ldots, x_n \), we can solve for the basic variables as

\[
\begin{align*}
    x_1 &= b'_1 - \sum_{j=m+1}^{n} a'_{1j}x_j, \\
    x_2 &= b'_2 - \sum_{j=m+1}^{n} a'_{2j}x_j, \\
    &\vdots \\
    x_m &= b'_m - \sum_{j=m+1}^{n} a'_{mj}x_j.
\end{align*}
\] (5)

Substitute equations (5) into (4) to eliminate \( x_1, x_2, \ldots, x_n \):

\[
    z = z_0 + (c'_{m+1} - z_{m+1})x_{m+1} + \cdots + (c'_n - z_n)x_n,
\] (6)

where

\[
    z_j = a'_{1j}c'_1 + a'_{2j}c'_2 + \cdots + a'_{mj}c'_m, \quad m+1 \leq j \leq n.
\] (7)

From (6) we can now determine if introducing one of the nonbasic variable would decrease the value of the objective function. More specifically, if \( c'_k - z_k < 0 \) for some \( k, m+1 \leq k \leq n \), then making \( x_k > 0 \) would decrease the cost. The formulas (6) and (7) already take into account the changes of values that would be required in \( x_1, \ldots, x_m \) to accommodate the change in \( x_j \).

**Theorem 1 (Improvement of Basic Feasible Solution)** Given a nondegenerate basic feasible solution with corresponding objective value \( z_0 \), suppose \( c'_k - z_k < 0 \) for some \( k \). Then there is a feasible solution with objective value \( z < z_0 \). If the variable \( x_k \) can be substituted for some variable in the original basis to yield a new basic feasible solution, this new solution will have objective value \( z < z_0 \). If \( x_k \) cannot be substituted to yield a basic feasible solution, then the solution set is unbounded and the objective function can be made arbitrarily small.

In the above theorem, when \( x_k \) with \( c'_k - z_k < 0 \) cannot be substituted into the basis, \( a'_{1k}, \ldots, a'_{mk} \) must be negative. As a result, \( x_k \) can have arbitrarily large value, thereby resulting in arbitrarily small \( z \) according to (6).

**Theorem 2 (Optimality Condition)** If for some basic feasible solution \( c'_j - z_j \geq 0 \) for all \( j \), then that solution is optimal.

3 **Initializing the Simplex Method**

For LP problems with constraints of the form

\[
\begin{align*}
    Ax &\leq b, \quad \text{with } b \geq 0, \\
    x &\geq 0,
\end{align*}
\]
a basic feasible solution to the corresponding standard form of the problem is provided by slack variables. This provides a means for initiating a simplex procedure.

But initial basic feasible solutions are not always apparent for other types of LP problems. Interestingly, an auxiliary linear program will provide the required initial solution to the original linear program.

We know that the constraints of an LP problem can always be expressed in the form

\[
\begin{align*}
Ax &= b, & \text{with } b &\geq 0, \\
x &\geq 0,
\end{align*}
\]

In order to find an initial basic feasible solution, consider the minimization problem

\[
\begin{align*}
\min \quad & \sum_{i=1}^{m} y_i \\
\text{subject to} \quad & Ax + y = b, \\
& x \geq 0, \\
& y \geq 0,
\end{align*}
\]

where \( y = (y_1, y_2, \ldots, y_m)^T \) is a vector of artificial variables.

It is clear that

- if (8) has a feasible solution, then (9) can achieve minimum objective value zero with \( y = 0 \);
- if (8) has no feasible solution, then (9) has a minimum objective value greater than zero.

Now (9) is an LP problem in variables \( x, y \). It has a trivial basic feasible solution \( y = b \). Use simplex to solve (8) and obtain a basic feasible solution at each step. If the minimum objective value in (9) is zero, then all \( y_i \) must be 0 in the final basic solution, which in the nondegenerate case will have no \( y_i \) in the basis. If some \( y_i \) are in the basis, they can be exchanged for nonbasic \( x_j \) variables (which have zero values) to yield a basic feasible solution involving \( x_j \) variables only.

To reiterate, a general LP problem can be solved by two phases:

**Phase I**: Introduce artificial variables and use simplex to find a basic feasible solution.

**Phase II**: Using the solution found in phase I, run simplex to minimize the original objective function.

**Example 1.** Consider the problem

\[
\begin{align*}
\min \quad & 4x_1 + x_2 + x_3 \\
\text{subject to} \quad & 2x_1 + x_2 + 2x_3 = 4 \\
& 3x_1 + 3x_2 + x_3 = 3 \\
& x_1, x_2, x_3 \geq 0
\end{align*}
\]

We introduce artificial variables \( x_4 \geq 0, x_5 \geq 0 \), and an objective function \( x_4 + x_5 \). The initial tableau is

\[
\begin{array}{cccccc|c}
 a_1 & a_2 & a_3 & a_4 & a_5 & b \\
1 & 2 & 1 & 0 & 0 & 4 \\
3 & 3 & 1 & 0 & 1 & 3 \\
0 & 0 & 0 & 1 & 1 & 0
\end{array}
\]
We first update the last row so that it has zero components under the artificial variables

$$
\begin{array}{cccccc}
2 & 1 & 2 & 1 & 0 & 4 \\
3 & 3 & 1 & 0 & 1 & 3 \\
-5 & -4 & -3 & 0 & 0 & -7 \\
\end{array}
$$

Pivoting at $\begin{bmatrix} 3 \end{bmatrix}$ yields

$$
\begin{array}{cccccc}
0 & -1 & \frac{1}{3} & 1 & -\frac{2}{3} & 2 \\
1 & 1 & \frac{2}{3} & 0 & \frac{1}{3} & 1 \\
0 & 1 & -\frac{4}{3} & 0 & \frac{5}{3} & -2 \\
\end{array}
$$

Next, we pivot at $\begin{bmatrix} \frac{4}{3} \end{bmatrix}$

$$
\begin{array}{cccccc}
0 & -\frac{3}{4} & 1 & \frac{3}{4} & \frac{3}{2} & 0 \\
1 & \frac{4}{3} & 0 & -\frac{1}{4} & \frac{1}{2} & 0 \\
0 & 0 & 0 & 1 & 1 & 0 \\
\end{array}
$$

The final tableau above leads to a basic solution to the original problem:

$$x_1 = \frac{1}{2}, \quad x_2 = 0, \quad x_3 = \frac{3}{2}.$$

Beginning Phase II, we use the new cost function in place of the old one and delete the artificial variable columns in the final tableau of Phase I:

$$
\begin{array}{cccc}
a_1 & a_2 & a_3 & b \\
0 & -\frac{3}{4} & 1 & \frac{3}{2} \\
1 & \frac{4}{3} & 0 & \frac{1}{2} \\
4 & 1 & 1 & 0 \\
\end{array}
$$

Again, transform the last row so that zeros appear in the basic columns

$$
\begin{array}{cccc}
0 & -\frac{3}{4} & 1 & \frac{3}{2} \\
1 & \frac{4}{3} & 0 & \frac{1}{2} \\
0 & -\frac{1}{2} & 0 & -\frac{7}{2} \\
\end{array}
$$

Pivoting at $\begin{bmatrix} \frac{5}{3} \end{bmatrix}$ yields the final tableau:

$$
\begin{array}{cccc}
\frac{3}{5} & 0 & 1 & \frac{9}{5} \\
\frac{4}{5} & 0 & 1 & \frac{2}{5} \\
\frac{13}{5} & 0 & 0 & -\frac{11}{5} \\
\end{array}
$$

The optimal solution is

$$x_1 = 0, \quad x_2 = \frac{2}{5}, \quad x_3 = \frac{9}{5}.$$

**Example 2.** In fact, we could also obtain the optimal solution for Example 1 by performing a sequence of “pivoting” on the initial tableau directly:

$$
\begin{array}{cccc}
2 & 1 & 2 & 4 \\
3 & 3 & 1 & 3 \\
4 & 1 & 1 & 0 \\
\end{array}
$$

$$
\begin{array}{cccc}
\boxed{-4} & -5 & 0 & -2 \\
3 & 3 & 1 & 3 \\
1 & -2 & 0 & -3 \\
\end{array}
$$
1 \begin{bmatrix} 2 \frac{2}{5} \frac{1}{5} \\ 0 & -\frac{3}{4} & 1 \\ 0 & -\frac{13}{4} & 0 \end{bmatrix} 0 \begin{bmatrix} \frac{1}{5} \\ \frac{3}{4} \\ \frac{13}{5} \end{bmatrix}

To pivoting from the second tableau to the third tableau, the simplex method would have multiply the first row by $-1$ first.

As we have seen, the above pivoting steps were not determined procedurally (we were a little lucky in following a path to reach the optimal solution shortly). Therefore the two-phase pivoting should be applied on general linear programs.

4 Duality

The linear program

$$\begin{align*}
\min & \quad ce \\
\text{subject to} & \quad Ax \geq b \\
& \quad x \geq 0
\end{align*}$$

is referred to as primal. It has a dual linear program in the form of

$$\begin{align*}
\max & \quad \lambda b \\
\text{subject to} & \quad \lambda A \leq c \\
& \quad \lambda \geq 0
\end{align*}$$

Here $A$ has dimension $m \times n$, and $x, b, c, \lambda$ have dimensions $n \times 1$, $m \times 1$, $1 \times n$, and $1 \times m$, respectively.

The roles of primal and dual LPs can be reversed. To see this, we change the dual above to its equivalent formation

$$\begin{align*}
\min & \quad (-b)^T \lambda^T \\
\text{subject to} & \quad (-A)^T \lambda^T \geq (-c)^T \\
& \quad \lambda \geq 0
\end{align*}$$

The dual of the above LP is clearly equivalent to the original primal.

The dual of any LP can be found by converting the problem to the primal form. For instance, given an LP in standard form:

$$\begin{align*}
\min & \quad cx \\
\text{subject to} & \quad Ax = b \\
& \quad x \geq 0
\end{align*}$$

we write it in the equivalent form

$$\begin{align*}
\min & \quad cx \\
\text{subject to} & \quad Ax \geq b \\
& \quad -Ax \geq -b \\
& \quad x \geq 0
\end{align*}$$
The corresponding dual is then

\[
\begin{align*}
\text{max} & \quad \mathbf{u} \mathbf{b} - \mathbf{v} \mathbf{b} \\
\text{subject to} & \quad \mathbf{u} \mathbf{A} - \mathbf{v} \mathbf{A} \leq \mathbf{c} \\
& \quad \mathbf{u} \geq 0 \\
& \quad \mathbf{v} \geq 0
\end{align*}
\]

We introduce \( \lambda = \mathbf{u} - \mathbf{v} \) to simplify the dual representation to

\[
\begin{align*}
\text{max} & \quad \lambda \mathbf{b} \\
\text{subject to} & \quad \lambda \mathbf{A} \leq \mathbf{c}
\end{align*}
\]

This is the asymmetric form of the duality. The dual vector \( \lambda \) is not restricted to be nonnegative.

Now we look at the underlying interpretation of a dual LP relative to the original LP. This is illustrated on the transportation problem.

**Example 2. (Dual of the transportation problem)** Recall that the transportation problem asks to select the pattern of product shipments between a number of origins and a number of destinations so as to minimize transportation cost while meeting the demand of each destination.

To interpret the dual problem, imagine an entrepreneur who, feeling that he can ship more efficiently, comes to the manufacturer with the offer to buy his product at the plant sites (origins) and sell it at the warehouses (destinations). The product prices to be used in these transactions vary from point to point. He must choose these prices so that his offer will be attractive to the manufacturer.

So he select prices \( \lambda_1, \ldots, \lambda_m \) for the \( m \) origins and \( \mu_1, \ldots, \mu_n \) for the \( n \) destinations. To compete with the usual transportation, his prices must satisfy the constraints \( \mu_j - \lambda_i \leq c_{ij} \), for all \( i, j \), which represents the net amount the manufacturer must pay to sell a unit of product at origin \( i \) and buy it back at destination \( j \). Then he needs to solve the LP problem

\[
\begin{align*}
\text{max} & \quad \sum_{j=1}^{n} \mu_j b_j - \sum_{i=1}^{m} \lambda_i a_i \\
\text{subject to} & \quad \mu_j - \lambda_i \leq c_{ij} & i = 1, \ldots, m \\
& \quad j = 1, \ldots, n
\end{align*}
\]

whereas the original transportation problem is

\[
\begin{align*}
\text{min} & \quad \sum_{i,j}^{n} c_{ij} x_{ij} \\
\text{subject to} & \quad \sum_{j=1}^{m} x_{ij} = a_i & i = 1, \ldots, m \\
& \quad \sum_{i=1}^{m} x_{ij} = b_j & j = 1, \ldots, n \\
& \quad x_{ij} \geq 0 & i = 1, \ldots, m \\
& \quad j = 1, \ldots, n
\end{align*}
\]

There exists deeper connection between a problem and its dual than just their forms and underlying interpretations. Let us consider the primal problem in standard form

\[
\begin{align*}
\text{min} & \quad \mathbf{c} \mathbf{x} \\
\text{subject to} & \quad \mathbf{A} \mathbf{x} = \mathbf{b} \\
& \quad \mathbf{x} \geq 0
\end{align*}
\]  

(10)
and it dual
\[
\max \quad \lambda b \\
\text{subject to} \quad \lambda A \leq c
\] (11)

**Lemma 3** If \( x \) and \( \lambda \) are feasible for the primal problem (10) and dual problems (11), respectively, then \( cx \geq \lambda b \).

The above lemma directly follows from that
\[
\lambda b = \lambda Ax \leq cx.
\]

**Corollary 4** If \( x^* \) and \( \lambda^* \) are feasible solutions of the primal and dual problems, respectively, and if \( cx^* = \lambda^*b \), then \( x^* \) and \( \lambda^* \) are optimal.

**Theorem 5 (Duality Theorem)** If either the primal problem (10) or the dual problem (11) has a finite optimal solution, so does the other and the corresponding values of the objective functions are equal. If either problem has an unbounded objective, the other problem has no feasible solution.

The following theorem relates the duality theorem to the simplex procedure:

**Theorem 6** Let the primal problem (10) have an optimal basic feasible solution \( x = (x_B, 0)^T \), without loss of generality, and let the corresponding columns in \( A \) form a submatrix \( B \), and the corresponding cost coefficients form a row vector \( c_B \). Then the vector \( \lambda = c_B B^{-1} \) is an optimal solution to the dual problem (11). The optimal values of both problems are equal.

**References**
