

The Simplex Method in Matrix Notation

So far, we have avoided using matrix notation to present linear programming problems and the simplex method. In this chapter, we shall recast everything into matrix notation. At the same time, we will emphasize the close relations between the primal and the dual problems.

1. Matrix Notation

As usual, we begin our discussion with the standard-form linear programming problem:

$$\begin{aligned} &\text{maximize} && \sum_{j=1}^n c_j x_j \\ &\text{subject to} && \sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = 1, 2, \dots, m \\ & && x_j \geq 0 \quad j = 1, 2, \dots, n. \end{aligned}$$

In the past, we have generally denoted slack variables by w_i 's but have noted that sometimes it is convenient just to string them onto the end of the list of original variables. Such is the case now, and so we introduce slack variables as follows:

$$x_{n+i} = b_i - \sum_{j=1}^n a_{ij} x_j, \quad i = 1, 2, \dots, m.$$

With these slack variables, we now write our problem in matrix form:

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

where

$$(6.1) \quad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & 1 & & \\ a_{21} & a_{22} & \dots & a_{2n} & & 1 & \\ \vdots & \vdots & & \vdots & & & \ddots \\ a_{m1} & a_{m2} & \dots & a_{mn} & & & & 1 \end{bmatrix},$$

$$(6.2) \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{and} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{bmatrix}.$$

As we know, the simplex method is an iterative procedure in which each iteration is characterized by specifying which m of the $n + m$ variables are basic. As before, we denote by \mathcal{B} the set of indices corresponding to the basic variables, and we denote by \mathcal{N} the remaining nonbasic indices.

In component notation, the i th component of Ax can be broken up into a basic part and a nonbasic part:

$$(6.3) \quad \sum_{j=1}^{n+m} a_{ij}x_j = \sum_{j \in \mathcal{B}} a_{ij}x_j + \sum_{j \in \mathcal{N}} a_{ij}x_j.$$

We wish to introduce a notation for matrices that will allow us to break up the matrix product Ax analogously. To this end, let B denote an $m \times m$ matrix whose columns consist precisely of the m columns of A that are associated with the basic variables. Similarly, let N denote an $m \times n$ matrix whose columns are the n nonbasic columns of A . Then we write A in a partitioned-matrix form as follows:

$$A = [B \quad N]$$

Strictly speaking, the matrix on the right does not equal the A matrix. Instead, it is the A matrix with its columns rearranged in such a manner that all the columns associated with basic variables are listed first followed by the nonbasic columns. Nonetheless, as long as we are consistent and rearrange the rows of x in the same way, then no harm is done. Indeed, let us similarly rearrange the rows of x and write

$$x = \begin{bmatrix} x_{\mathcal{B}} \\ x_{\mathcal{N}} \end{bmatrix}.$$

Then the following separation of Ax into a sum of two terms is true and captures the same separation into basic and nonbasic parts as we had in (6.3):

$$Ax = [B \quad N] \begin{bmatrix} x_{\mathcal{B}} \\ x_{\mathcal{N}} \end{bmatrix} = Bx_{\mathcal{B}} + Nx_{\mathcal{N}}.$$

By similarly partitioning c , we can write

$$c^T x = \begin{bmatrix} c_{\mathcal{B}} \\ c_{\mathcal{N}} \end{bmatrix}^T \begin{bmatrix} x_{\mathcal{B}} \\ x_{\mathcal{N}} \end{bmatrix} = c_{\mathcal{B}}^T x_{\mathcal{B}} + c_{\mathcal{N}}^T x_{\mathcal{N}}.$$

2. The Primal Simplex Method

A dictionary has the property that the basic variables are written as functions of the nonbasic variables. In matrix notation, we see that the constraint equations

$$Ax = b$$

can be written as

$$Bx_{\mathcal{B}} + Nx_{\mathcal{N}} = b.$$

The fact that the basic variables $x_{\mathcal{B}}$ can be written as a function of the nonbasic variables $x_{\mathcal{N}}$ is equivalent to the fact that the matrix B is invertible, and hence,

$$(6.4) \quad x_{\mathcal{B}} = B^{-1}b - B^{-1}Nx_{\mathcal{N}}.$$

(The fact that B is invertible means that its m column vectors are linearly independent and therefore form a basis for \mathbb{R}^m —this is why the basic variables are called basic, in case you were wondering.) Similarly, the objective function can be written as

$$(6.5) \quad \begin{aligned} \zeta &= c_{\mathcal{B}}^T x_{\mathcal{B}} + c_{\mathcal{N}}^T x_{\mathcal{N}} \\ &= c_{\mathcal{B}}^T (B^{-1}b - B^{-1}Nx_{\mathcal{N}}) + c_{\mathcal{N}}^T x_{\mathcal{N}} \\ &= c_{\mathcal{B}}^T B^{-1}b - ((B^{-1}N)^T c_{\mathcal{B}} - c_{\mathcal{N}})^T x_{\mathcal{N}}. \end{aligned}$$

Combining (6.5) and (6.4), we see that we can write the dictionary associated with basis \mathcal{B} as

$$(6.6) \quad \begin{aligned} \zeta &= c_{\mathcal{B}}^T B^{-1}b - ((B^{-1}N)^T c_{\mathcal{B}} - c_{\mathcal{N}})^T x_{\mathcal{N}} \\ x_{\mathcal{B}} &= B^{-1}b - B^{-1}Nx_{\mathcal{N}}. \end{aligned}$$

Comparing against the component-form notation of Chapter 2 (see (2.6)), we make the following identifications:

$$\begin{aligned} c_{\mathcal{B}}^T B^{-1}b &= \bar{c} \\ c_{\mathcal{N}} - (B^{-1}N)^T c_{\mathcal{B}} &= [\bar{c}_j] \\ B^{-1}b &= [\bar{b}_i] \\ B^{-1}N &= [\bar{a}_{ij}], \end{aligned}$$

where the bracketed expressions on the right denote vectors and matrices with the index i running over \mathcal{B} and the index j running over \mathcal{N} . The basic solution associated with dictionary (6.6) is obtained by setting $x_{\mathcal{N}}$ equal to zero:

$$(6.7) \quad \begin{aligned} x_{\mathcal{N}}^* &= 0, \\ x_{\mathcal{B}}^* &= B^{-1}b. \end{aligned}$$

As we saw in the last chapter, associated with each primal dictionary there is a dual dictionary that is simply the negative-transpose of the primal. However, to have the negative-transpose property, it is important to correctly associate complementary pairs of variables. So first we recall that, for the current discussion, we have appended the primal slack variables to the end of the original variables:

$$(x_1, \dots, x_n, w_1, \dots, w_m) \longrightarrow (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}).$$

Also recall that the dual slack variables are complementary to the original primal variables and that the original dual variables are complementary to the primal slack variables. Therefore, to maintain the desired complementarity condition between like indices in the primal and the dual, we need to relabel the dual variables and append them to the end of the dual slacks:

$$(z_1, \dots, z_n, y_1, \dots, y_m) \longrightarrow (z_1, \dots, z_n, z_{n+1}, \dots, z_{n+m}).$$

With this relabeling of the dual variables, the dual dictionary corresponding to (6.6) is

$$\begin{aligned} -\xi &= -c_B^T B^{-1}b - (B^{-1}b)^T z_B \\ z_{\mathcal{N}} &= (B^{-1}N)^T c_B - c_{\mathcal{N}} + (B^{-1}N)^T z_B. \end{aligned}$$

The dual solution associated with this dictionary is obtained by setting z_B equal to zero:

$$(6.8) \quad \begin{aligned} z_B^* &= 0, \\ z_{\mathcal{N}}^* &= (B^{-1}N)^T c_B - c_{\mathcal{N}}. \end{aligned}$$

Using (6.7) and (6.8) and introducing the shorthand

$$(6.9) \quad \zeta^* = c_B^T B^{-1}b,$$

we see that we can write the primal dictionary succinctly as

$$(6.10) \quad \begin{aligned} \zeta &= \zeta^* - z_{\mathcal{N}}^{*T} x_{\mathcal{N}} \\ x_B &= x_B^* - B^{-1}N x_{\mathcal{N}}. \end{aligned}$$

The associated dual dictionary then has a very symmetric appearance:

$$(6.11) \quad \begin{aligned} -\xi &= -\zeta^* - (x_B^*)^T z_B \\ z_{\mathcal{N}} &= z_{\mathcal{N}}^* + (B^{-1}N)^T z_B. \end{aligned}$$

The (primal) simplex method can be described briefly as follows. The starting assumptions are that we are given

- (1) A partition of the $n + m$ indices into a collection \mathcal{B} of m basic indices and a collection \mathcal{N} of n nonbasic ones with the property that the basis matrix B is invertible,
- (2) An associated current primal solution $x_B^* \geq 0$ (and $x_{\mathcal{N}}^* = 0$), and
- (3) An associated current dual solution $z_{\mathcal{N}}^*$ (with $z_B^* = 0$)

such that the dictionary given by (6.10) represents the primal objective function and the primal constraints. The simplex method then produces a sequence of steps to “adjacent” bases such that the current value ζ^* of the objective function ζ increases at each step (or, at least, would increase if the step size were positive), updating x_B^* and $z_{\mathcal{N}}^*$ along the way. Two bases are said to be adjacent to each other if they differ in only one index. That is, given a basis \mathcal{B} , an adjacent basis is determined by removing one basic index and replacing it with a nonbasic index. The index that gets removed corresponds to the leaving variable, whereas the index that gets added corresponds to the entering variable.

One step of the simplex method is called an iteration. We now elaborate further on the details by describing one iteration as a sequence of specific steps.

Step 1. Check for Optimality. If $z_{\mathcal{N}}^* \geq 0$, stop. The current solution is optimal. To see this, first note that the simplex method always maintains primal feasibility and complementarity. Indeed, the primal solution is feasible, since $x_{\mathcal{B}}^* \geq 0$ and $x_{\mathcal{N}} = 0$ and the dictionary embodies the primal constraints. Also, the fact that $x_{\mathcal{N}}^* = 0$ and $z_{\mathcal{B}}^* = 0$ implies that the primal and dual solutions are complementary. Hence, all that is required for optimality is dual feasibility. But by looking at the associated dual dictionary (6.11), we see that the dual solution is feasible if and only if $z_{\mathcal{N}}^* \geq 0$.

Step 2. Select Entering Variable. Pick an index $j \in \mathcal{N}$ for which $z_j^* < 0$. Variable x_j is the *entering variable*.

Step 3. Compute Primal Step Direction $\Delta x_{\mathcal{B}}$. Having selected the entering variable, it is our intention to let its value increase from zero. Hence, we let

$$x_{\mathcal{N}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \swarrow \quad = te_j, \quad \begin{matrix} \\ \\ \\ \text{jth position} \\ \\ \end{matrix}$$

where we follow the common convention of letting e_j denote the unit vector that is zero in every component except for a one in the position associated with index j (note that, because of our index rearrangement conventions, this is not generally the j th element of the vector). Then from (6.10), we have that

$$x_{\mathcal{B}} = x_{\mathcal{B}}^* - B^{-1}Nte_j.$$

Hence, we see that the step direction $\Delta x_{\mathcal{B}}$ for the primal basic variables is given by

$$\Delta x_{\mathcal{B}} = B^{-1}Ne_j.$$

Step 4. Compute Primal Step Length. We wish to pick the largest $t \geq 0$ for which every component of $x_{\mathcal{B}}$ remains nonnegative. That is, we wish to pick the largest t for which

$$x_{\mathcal{B}}^* \geq t\Delta x_{\mathcal{B}}.$$

Since, for each $i \in \mathcal{B}^*$, $x_i^* \geq 0$ and $t \geq 0$, we can divide both sides of the above inequality by these numbers and preserve the sense of the inequality. Therefore, doing this division, we get the requirement that

$$\frac{1}{t} \geq \frac{\Delta x_i}{x_i^*}, \quad \text{for all } i \in \mathcal{B}.$$

We want to let t be as large as possible, and so $1/t$ should be made as small as possible. The smallest possible value for $1/t$ that satisfies all the required inequalities is obviously

$$\frac{1}{t} = \max_{i \in \mathcal{B}} \frac{\Delta x_i}{x_i^*}.$$

Hence, the largest t for which all of the inequalities hold is given by

$$t = \left(\max_{i \in \mathcal{B}} \frac{\Delta x_i}{x_i^*} \right)^{-1}.$$

As always, the correct convention for $0/0$ is to set such ratios to zero. Also, if the maximum is less than or equal to zero, we can stop here—the primal is unbounded.

Step 5. Select Leaving Variable. The leaving variable is chosen as any variable $x_i, i \in \mathcal{B}$, for which the maximum in the calculation of t is obtained.

Step 6. Compute Dual Step Direction $\Delta z_{\mathcal{N}}$. Essentially all that remains is to explain how $z_{\mathcal{N}}^*$ changes. To see how, it is convenient to look at the dual dictionary. Since in that dictionary z_i is the entering variable, we see that

$$\Delta z_{\mathcal{N}} = -(B^{-1}N)^T e_i.$$

Step 7. Compute Dual Step Length. Since we know that z_j is the leaving variable in the dual dictionary, we see immediately that the step length for the dual variables is

$$s = \frac{z_j^*}{\Delta z_j}.$$

Step 8. Update Current Primal and Dual Solutions. We now have everything we need to update the data in the dictionary:

$$\begin{aligned} x_j^* &\leftarrow t \\ x_{\mathcal{B}}^* &\leftarrow x_{\mathcal{B}}^* - t \Delta x_{\mathcal{B}} \end{aligned}$$

and

$$\begin{aligned} z_i^* &\leftarrow s \\ z_{\mathcal{N}}^* &\leftarrow z_{\mathcal{N}}^* - s \Delta z_{\mathcal{N}}. \end{aligned}$$

Step 9. Update Basis. Finally, we update the basis:

$$\mathcal{B} \leftarrow \mathcal{B} \setminus \{i\} \cup \{j\}.$$

We close this section with the important remark that the simplex method as presented here, while it may look different from the component-form presentation given in Chapter 2, is in fact mathematically identical to it. That is, given the same set of pivoting rules and starting from the same primal dictionary, the two algorithms will generate exactly the same sequence of dictionaries.

3. An Example

In case the reader is feeling at this point that there are too many letters and not enough numbers, here is an example that illustrates the matrix approach to the simplex method. The problem we wish to solve is

$$\begin{aligned} &\text{maximize} && 4x_1 + 3x_2 \\ &\text{subject to} && x_1 - x_2 \leq 1 \\ &&& 2x_1 - x_2 \leq 3 \\ &&& x_2 \leq 5 \\ &&& x_1, x_2 \geq 0. \end{aligned}$$

The matrix A is given by

$$\begin{bmatrix} 1 & -1 & 1 & & \\ 2 & -1 & & 1 & \\ 0 & 1 & & & 1 \end{bmatrix}.$$

(Note that some zeros have not been shown.) The initial sets of basic and nonbasic indices are

$$\mathcal{B} = \{3, 4, 5\} \quad \text{and} \quad \mathcal{N} = \{1, 2\}.$$

Corresponding to these sets, we have the submatrices of A :

$$B = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \quad N = \begin{bmatrix} 1 & -1 \\ 2 & -1 \\ 0 & 1 \end{bmatrix}.$$

From (6.7) we see that the initial values of the basic variables are given by

$$x_{\mathcal{B}}^* = b = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix},$$

and from (6.8) the initial nonbasic dual variables are simply

$$z_{\mathcal{N}}^* = -c_{\mathcal{N}} = \begin{bmatrix} -4 \\ -3 \end{bmatrix}.$$

Since $x_{\mathcal{B}}^* \geq 0$, the initial solution is primal feasible, and hence we can apply the simplex method without needing any Phase I procedure.

3.1. First Iteration. *Step 1.* Since $z_{\mathcal{N}}^*$ has some negative components, the current solution is not optimal.

Step 2. Since $z_1^* = -4$ and this is the most negative of the two nonbasic dual variables, we see that the entering index is

$$j = 1.$$

Step 3.

$$\Delta x_{\mathcal{B}} = B^{-1} N e_j = \begin{bmatrix} 1 & -1 \\ 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Step 4.

$$t = \left(\max \left\{ \frac{1}{1}, \frac{2}{3}, \frac{0}{5} \right\} \right)^{-1} = 1.$$

Step 5. Since the ratio that achieved the maximum in Step 4 was the first ratio and this ratio corresponds to basis index 3, we see that

$$i = 3.$$

Step 6.

$$\Delta z_{\mathcal{N}} = -(B^{-1} N)^T e_i = - \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Step 7.

$$s = \frac{z_j^*}{\Delta z_j} = \frac{-4}{-1} = 4.$$

Step 8.

$$\begin{aligned} x_1^* = 1, \quad x_{\mathcal{B}}^* &= \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 5 \end{bmatrix}, \\ z_3^* = 4, \quad z_{\mathcal{N}}^* &= \begin{bmatrix} -4 \\ -3 \end{bmatrix} - 4 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \end{bmatrix}. \end{aligned}$$

Step 9. The new sets of basic and nonbasic indices are

$$\mathcal{B} = \{1, 4, 5\} \quad \text{and} \quad \mathcal{N} = \{3, 2\}.$$

Corresponding to these sets, we have the new basic and nonbasic submatrices of A ,

$$B = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ 0 & & 1 \end{bmatrix} \quad N = \begin{bmatrix} 1 & -1 \\ 0 & -1 \\ 0 & 1 \end{bmatrix},$$

and the new basic primal variables and nonbasic dual variables:

$$x_{\mathcal{B}}^* = \begin{bmatrix} x_1^* \\ x_4^* \\ x_5^* \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix} \quad z_{\mathcal{N}}^* = \begin{bmatrix} z_3^* \\ z_2^* \end{bmatrix} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}.$$

3.2. Second Iteration. *Step 1.* Since $z_{\mathcal{N}}^*$ has some negative components, the current solution is not optimal.

Step 2. Since $z_2^* = -7$, we see that the entering index is

$$j = 2.$$

Step 3.

$$\Delta x_{\mathcal{B}} = B^{-1} N e_j = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ 0 & & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & -1 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

Step 4.

$$t = \left(\max \left\{ \frac{-1}{1}, \frac{1}{1}, \frac{1}{5} \right\} \right)^{-1} = 1.$$

Step 5. Since the ratio that achieved the maximum in Step 4 was the second ratio and this ratio corresponds to basis index 4, we see that

$$i = 4.$$

Step 6.

$$\begin{aligned} \Delta z_{\mathcal{N}} &= -(B^{-1} N)^T e_i \\ &= - \begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ & 1 & \\ & & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}. \end{aligned}$$

Step 7.

$$s = \frac{z_j^*}{\Delta z_j} = \frac{-7}{-1} = 7.$$

Step 8.

$$\begin{aligned} x_2^* = 1, \quad x_{\mathcal{B}}^* &= \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix} - 1 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \\ z_4^* = 7, \quad z_{\mathcal{N}}^* &= \begin{bmatrix} 4 \\ -7 \end{bmatrix} - 7 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -10 \\ 0 \end{bmatrix}. \end{aligned}$$

Step 9. The new sets of basic and nonbasic indices are

$$\mathcal{B} = \{1, 2, 5\} \quad \text{and} \quad \mathcal{N} = \{3, 4\}.$$

Corresponding to these sets, we have the new basic and nonbasic submatrices of A ,

$$B = \begin{bmatrix} 1 & -1 & 0 \\ 2 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad N = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

and the new basic primal variables and nonbasic dual variables:

$$x_{\mathcal{B}}^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_5^* \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} \quad z_{\mathcal{N}}^* = \begin{bmatrix} z_3^* \\ z_4^* \end{bmatrix} = \begin{bmatrix} -10 \\ 7 \end{bmatrix}.$$

3.3. Third Iteration. *Step 1.* Since $z_{\mathcal{N}}^*$ has some negative components, the current solution is not optimal.

Step 2. Since $z_3^* = -10$, we see that the entering index is

$$j = 3.$$

Step 3.

$$\Delta x_{\mathcal{B}} = B^{-1} N e_j = \begin{bmatrix} 1 & -1 & 0 \\ 2 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}.$$

Step 4.

$$t = \left(\max \left\{ \frac{-1}{2}, \frac{-2}{1}, \frac{2}{4} \right\} \right)^{-1} = 2.$$

Step 5. Since the ratio that achieved the maximum in Step 4 was the third ratio and this ratio corresponds to basis index 5, we see that

$$i = 5.$$

Step 6.

$$\begin{aligned} \Delta z_{\mathcal{N}} &= -(B^{-1} N)^T e_i \\ &= - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}. \end{aligned}$$

Step 7.

$$s = \frac{z_j^*}{\Delta z_j} = \frac{-10}{-2} = 5.$$

Step 8.

$$\begin{aligned} x_3^* = 2, \quad x_{\mathcal{B}}^* &= \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} - 2 \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}, \\ z_5^* = 5, \quad z_{\mathcal{N}}^* &= \begin{bmatrix} -10 \\ 7 \end{bmatrix} - 5 \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}. \end{aligned}$$

Step 9. The new sets of basic and nonbasic indices are

$$\mathcal{B} = \{1, 2, 3\} \quad \text{and} \quad \mathcal{N} = \{5, 4\}.$$

Corresponding to these sets, we have the new basic and nonbasic submatrices of A ,

$$B = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad N = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and the new basic primal variables and nonbasic dual variables:

$$x_{\mathcal{B}}^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 2 \end{bmatrix} \quad z_{\mathcal{N}}^* = \begin{bmatrix} z_5^* \\ z_4^* \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

3.4. Fourth Iteration. *Step 1.* Since $z_{\mathcal{N}}^*$ has all nonnegative components, the current solution is optimal. The optimal objective function value is

$$\zeta^* = 4x_1^* + 3x_2^* = 31.$$

It is undoubtedly clear at this point that the matrix approach, as we have presented it, is quite a bit more tedious than the dictionary manipulations with which we are quite familiar. The reason is that, with the dictionary approach, dictionary entries get updated from one iteration to the next and the updating process is fairly easy, whereas with the matrix approach, we continually compute everything from scratch and therefore end up solving many systems of equations. In the next chapter, we will deal with this issue and show that these systems of equations don't really have to be solved from scratch each time; instead, there is a certain updating that can be done that is quite analogous to the updating of a dictionary. However, before we take up such practical considerations, let us finish our general discussion of the simplex method by casting the dual simplex method into matrix notation and discussing some related issues.

4. The Dual Simplex Method

In the presentation of the primal simplex method given in the previous section, we tried to make the symmetry between the primal and the dual problems as evident as possible. One advantage of this approach is that we can now easily write down the dual simplex method. Instead of assuming that the primal dictionary is feasible

($x_{\mathcal{B}}^* \geq 0$), we now assume that the dual dictionary is feasible ($z_{\mathcal{N}}^* \geq 0$) and perform the analogous steps:

Step 1. Check for Optimality. If $x_{\mathcal{B}}^* \geq 0$, stop. The current solution is optimal. Note that for the dual simplex method, dual feasibility and complementarity are maintained from the beginning, and the algorithm terminates once a primal feasible solution is discovered.

Step 2. Select Entering Variable. Pick an index $i \in \mathcal{B}$ for which $x_i^* < 0$. Variable z_i is the *entering variable*.

Step 3. Compute Dual Step Direction $\Delta z_{\mathcal{N}}$. From the dual dictionary, we see that

$$\Delta z_{\mathcal{N}} = -(B^{-1}N)^T e_i.$$

Step 4. Compute Dual Step Length. We wish to pick the largest $s \geq 0$ for which every component of $z_{\mathcal{N}}$ remains nonnegative. As in the primal simplex method, this computation involves computing the maximum of some ratios:

$$s = \left(\max_{j \in \mathcal{N}} \frac{\Delta z_j}{z_j^*} \right)^{-1}.$$

If s is not positive, then stop here—the dual is unbounded (implying, of course, that the primal is infeasible).

Step 5. Select Leaving Variable. The leaving variable is chosen as any variable $z_j, j \in \mathcal{N}$, for which the maximum in the calculation of s is obtained.

Step 6. Compute Primal Step Direction $\Delta x_{\mathcal{B}}$. To see how $x_{\mathcal{B}}^*$ changes in the dual dictionary, it is convenient to look at the primal dictionary. Since in that dictionary x_j is the entering variable, we see that

$$\Delta x_{\mathcal{B}} = B^{-1}N e_j.$$

Step 7. Compute Primal Step Length. Since we know that x_i is the leaving variable in the primal dictionary, we see immediately that the step length for the primal variables is

$$t = \frac{x_i^*}{\Delta x_i}.$$

Step 8. Update Current Primal and Dual Solutions. We now have everything we need to update the data in the dictionary:

$$\begin{aligned} x_j^* &\leftarrow t \\ x_{\mathcal{B}}^* &\leftarrow x_{\mathcal{B}}^* - t \Delta x_{\mathcal{B}}. \end{aligned}$$

and

$$\begin{aligned} z_i^* &\leftarrow s \\ z_{\mathcal{N}}^* &\leftarrow z_{\mathcal{N}}^* - s \Delta z_{\mathcal{N}} \end{aligned}$$

Step 9. Update Basis. Finally, we update the basis:

$$\mathcal{B} \leftarrow \mathcal{B} \setminus \{i\} \cup \{j\}.$$

To further emphasize the similarities between the primal and the dual simplex methods, Figure 6.1 shows the two algorithms side by side.

Primal Simplex	Dual Simplex
<p>Suppose $x_{\mathcal{B}}^* \geq 0$ while $(z_{\mathcal{N}}^* \not\geq 0)$ { pick $j \in \{j \in \mathcal{N} : z_j^* < 0\}$ $\Delta x_{\mathcal{B}} = B^{-1} N e_j$ $t = \left(\max_{i \in \mathcal{B}} \frac{\Delta x_i}{x_i^*} \right)^{-1}$ pick $i \in \operatorname{argmax}_{i \in \mathcal{B}} \frac{\Delta x_i}{x_i^*}$ $\Delta z_{\mathcal{N}} = -(B^{-1} N)^T e_i$ $s = \frac{z_j^*}{\Delta z_j}$ $x_j^* \leftarrow t$ $x_{\mathcal{B}}^* \leftarrow x_{\mathcal{B}}^* - t \Delta x_{\mathcal{B}}$ $z_i^* \leftarrow s$ $z_{\mathcal{N}}^* \leftarrow z_{\mathcal{N}}^* - s \Delta z_{\mathcal{N}}$ $\mathcal{B} \leftarrow \mathcal{B} \setminus \{i\} \cup \{j\}$ }</p>	<p>Suppose $z_{\mathcal{N}}^* \geq 0$ while $(x_{\mathcal{B}}^* \not\geq 0)$ { pick $i \in \{i \in \mathcal{B} : x_i^* < 0\}$ $\Delta z_{\mathcal{N}} = -(B^{-1} N)^T e_i$ $s = \left(\max_{j \in \mathcal{N}} \frac{\Delta z_j}{z_j^*} \right)^{-1}$ pick $j \in \operatorname{argmax}_{j \in \mathcal{N}} \frac{\Delta z_j}{z_j^*}$ $\Delta x_{\mathcal{B}} = B^{-1} N e_j$ $t = \frac{x_i^*}{\Delta x_i}$ $x_j^* \leftarrow t$ $x_{\mathcal{B}}^* \leftarrow x_{\mathcal{B}}^* - t \Delta x_{\mathcal{B}}$ $z_i^* \leftarrow s$ $z_{\mathcal{N}}^* \leftarrow z_{\mathcal{N}}^* - s \Delta z_{\mathcal{N}}$ $\mathcal{B} \leftarrow \mathcal{B} \setminus \{i\} \cup \{j\}$ }</p>

FIGURE 6.1. The primal and the dual simplex methods.

5. Two-Phase Methods

Let us summarize the algorithm obtained by applying the dual simplex method as a Phase I procedure followed by the primal simplex method as a Phase II. Initially, we set

$$\mathcal{B} = \{n+1, n+2, \dots, n+m\} \quad \text{and} \quad \mathcal{N} = \{1, 2, \dots, n\}.$$

Then from (6.1) we see that $A = [N \quad B]$, where

$$N = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

and from (6.2) we have

$$c_{\mathcal{N}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad \text{and} \quad c_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Substituting these expressions into the definitions of $x_{\mathcal{B}}^*$, $z_{\mathcal{N}}^*$, and ζ^* , we find that

$$\begin{aligned}x_{\mathcal{B}}^* &= B^{-1}b = b \\z_{\mathcal{N}}^* &= (B^{-1}N)^T c_{\mathcal{B}} - c_{\mathcal{N}} = -c_{\mathcal{N}} \\\zeta^* &= 0.\end{aligned}$$

Hence, the initial dictionary reads:

$$\begin{aligned}\zeta &= c_{\mathcal{N}}^T x_{\mathcal{N}} \\x_{\mathcal{B}} &= b - Nx_{\mathcal{N}}.\end{aligned}$$

If b has all nonnegative components and $c_{\mathcal{N}}$ has all nonpositive components, then this dictionary is optimal—the problem was trivial. Suppose, however, that one of these two vectors (but not both) has components of the wrong sign. For example, suppose that b is okay (all nonnegative components) but $c_{\mathcal{N}}$ has some positive components. Then this dictionary is primal feasible, and we can start immediately with the primal simplex method. On the other hand, suppose that $c_{\mathcal{N}}$ has all nonpositive components but b has some negative ones. Then the starting dictionary is dual feasible, and we can commence immediately with the dual simplex algorithm.

The last, and most common, case is where both b and $c_{\mathcal{N}}$ have components of the wrong sign. In this case, we must employ a two-phase procedure. There are two choices. We could temporarily replace $c_{\mathcal{N}}$ with another vector that is nonpositive. Then the modified problem is dual feasible, and so we can apply the dual simplex method to find an optimal solution of this modified problem. After that, the original objective function could be reinstated. With the original objective function, the optimal solution from Phase I is most likely not optimal, but it is feasible, and therefore the primal simplex method can be used to find the optimal solution to the original problem.

The other choice would be to modify b instead of $c_{\mathcal{N}}$, thereby obtaining a primal feasible solution to a modified problem. Then we would use the primal simplex method on the modified problem to obtain its optimal solution, which will then be dual feasible for the original problem, and so the dual simplex method can be used to finish the problem.

6. Negative Transpose Property

In our discussion of duality in Chapter 5, we emphasized the symmetry between the primal problem and its dual. This symmetry can be easily summarized by saying that the dual of a standard-form linear programming problem is the negative transpose of the primal problem. Now, in this chapter, the symmetry appears to have been lost. For example, the basis matrix is an $m \times m$ matrix. Why $m \times m$ and not $n \times n$? It seems strange. In fact, if we had started with the dual problem, added slack variables to it, and introduced a basis matrix on that side it would be an $n \times n$ matrix. How are these two basis matrices related? It turns out that they are not themselves related in any simple way, but the important matrix $B^{-1}N$ is still the negative transpose of the analogous dual construct. The purpose of this section is to make this connection clear.

Consider a standard-form linear programming problem

$$\begin{aligned} & \text{maximize} && c^T x \\ & \text{subject to} && Ax \leq b \\ & && x \geq 0, \end{aligned}$$

and its dual

$$\begin{aligned} & \text{minimize} && b^T y \\ & \text{subject to} && A^T y \geq c \\ & && y \geq 0. \end{aligned}$$

Let w be a vector containing the slack variables for the primal problem, let z be a slack vector for the dual problem, and write both problems in equality form:

$$\begin{aligned} & \text{maximize} && c^T x \\ & \text{subject to} && Ax + w = b \\ & && x, w \geq 0, \end{aligned}$$

and

$$\begin{aligned} & \text{minimize} && b^T y \\ & \text{subject to} && A^T y - z = c \\ & && y, z \geq 0. \end{aligned}$$

Introducing three new notations,

$$\bar{A} = [A \ I], \quad \bar{c} = \begin{bmatrix} c \\ 0 \end{bmatrix}, \quad \text{and} \quad \bar{x} = \begin{bmatrix} x \\ w \end{bmatrix},$$

the primal problem can be rewritten succinctly as follows:

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

Similarly, using “hats” for new notations on the dual side,

$$\hat{A} = [-I \ A^T], \quad \hat{b} = \begin{bmatrix} 0 \\ b \end{bmatrix}, \quad \text{and} \quad \hat{y} = \begin{bmatrix} z \\ y \end{bmatrix},$$

the dual problem can be rewritten in this way:

$$\begin{aligned} & \text{minimize} && \hat{b}^T \hat{y} \\ & \text{subject to} && \hat{A} \hat{y} = c \\ & && \hat{y} \geq 0. \end{aligned}$$

Note that the matrix $\bar{A} = [A \ I]$ is an $m \times (n+m)$ matrix. The first n columns of it are the initial nonbasic variables and the last m columns are the initial basic columns. After doing some simplex pivots, the basic and nonbasic columns get jumbled up but we can still write the equality

$$[A \ I] = [\bar{N} \ \bar{B}]$$

with the understanding that the equality only holds after rearranging the columns appropriately.

On the dual side, the matrix $\hat{A} = [-I \ A^T]$ is an $n \times (n+m)$ matrix. The first n columns of it are the initial basic variables (for the dual problem) and the last m

columns are the initial nonbasic columns. If the same set of pivots that were applied to the primal problem are also applied to the dual, then the columns get rearranged in exactly the same way as they did for the primal and we can write

$$[-I \ A^T] = [\hat{B} \ \hat{N}]$$

again with the proviso that the columns of one matrix must be rearranged in a specific manner to bring it into exact equality with the other matrix.

Now, the primal dictionary involves the matrix $\bar{B}^{-1}\bar{N}$ whereas the dual dictionary involves the matrix $\hat{B}^{-1}\hat{N}$. It probably doesn't seem at all obvious that these two matrices are negative transposes of each other. To see that it is so, consider what happens when we multiply \bar{A} by \hat{A}^T in both the permuted notation and the unpermuted notation:

$$\bar{A}\hat{A}^T = [\bar{N} \ \bar{B}] \begin{bmatrix} \hat{B}^T \\ \hat{N}^T \end{bmatrix} = \bar{N}\hat{B}^T + \bar{B}\hat{N}^T$$

and

$$\bar{A}\hat{A}^T = [A \ I] \begin{bmatrix} -I \\ A \end{bmatrix} = -A + A = 0.$$

These two expressions obviously must agree so we see that

$$\bar{N}\hat{B}^T + \bar{B}\hat{N}^T = 0.$$

Putting the two terms on the opposite sides of the equality sign and multiplying on the right by the inverse of \hat{B}^T and on the left by the inverse of \bar{B} , we get that

$$\bar{B}^{-1}\bar{N} = -(\hat{B}^{-1}\hat{N})^T,$$

which is the property we wished to establish.

Exercises

6.1 Consider the following linear programming problem:

$$\begin{aligned} &\text{maximize} && -6x_1 + 32x_2 - 9x_3 \\ &\text{subject to} && -2x_1 + 10x_2 - 3x_3 \leq -6 \\ & && x_1 - 7x_2 + 2x_3 \leq 4 \\ & && x_1, x_2, x_3 \geq 0. \end{aligned}$$

Suppose that, in solving this problem, you have arrived at the following dictionary:

$$\begin{array}{r} \zeta = -18 - 3x_4 + 2x_2 \\ \hline x_3 = 2 - x_4 + 4x_2 - 2x_5 \\ x_1 = 2x_4 - x_2 + 3x_5. \end{array}$$

- Which variables are basic? Which are nonbasic?
- Write down the vector, x_B^* , of current primal basic solution values.
- Write down the vector, z_N^* , of current dual nonbasic solution values.
- Write down $B^{-1}N$.
- Is the primal solution associated with this dictionary feasible?

- (f) Is it optimal?
- (g) Is it degenerate?

6.2 Consider the following linear programming problem:

$$\begin{aligned} \text{maximize} \quad & x_1 + 2x_2 + 4x_3 + 8x_4 + 16x_5 \\ \text{subject to} \quad & x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 \leq 2 \\ & 7x_1 + 5x_2 - 3x_3 - 2x_4 \leq 0 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{aligned}$$

Consider the situation in which x_3 and x_5 are basic and all other variables are nonbasic. Write down:

- (a) B ,
 - (b) N ,
 - (c) b ,
 - (d) c_B ,
 - (e) c_N ,
 - (f) $B^{-1}N$,
 - (g) $x_B^* = B^{-1}b$,
 - (h) $\zeta^* = c_B^T B^{-1}b$,
 - (i) $z_N^* = (B^{-1}N)^T c_B - c_N$,
 - (j) The dictionary corresponding to this basis.
- 6.3** Solve the problem in Exercise 2.1 using the matrix form of the primal simplex method.
- 6.4** Solve the problem in Exercise 2.4 using the matrix form of the dual simplex method.
- 6.5** Solve the problem in Exercise 2.3 using the two-phase approach in matrix form.
- 6.6** Find the dual of the following linear program:

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{subject to} \quad & a \leq Ax \leq b \\ & l \leq x \leq u. \end{aligned}$$

- 6.7** (a) Let A be a given $m \times n$ matrix, c a given n -vector, and b a given m -vector. Consider the following max-min problem:

$$\max_{x \geq 0} \min_{y \geq 0} (c^T x - y^T Ax + b^T y).$$

By noting that the inner optimization can be carried out explicitly, show that this problem can be reduced to a linear programming problem. Write it explicitly.

- (b) What linear programming problem do you get if the min and max are interchanged?

Notes

In this chapter, we have accomplished two tasks: (1) we have expressed the simplex method in matrix notation, and (2) we have reduced the information we carry from iteration to iteration to simply the list of basic variables together with current values of the primal basic variables and the dual nonbasic variables. In particular, it is not necessary to calculate explicitly all the entries of the matrix $B^{-1}N$.

What's in a name? There are times when one thing has two names. So far in this book, we have discussed essentially only one algorithm: the simplex method (assuming, of course, that specific pivot rules have been settled on). But this one algorithm is sometimes referred to as the simplex method and at other times it is referred to as the *revised simplex method*. The distinction being made with this new name has nothing to do with the algorithm. Rather it refers to the specifics of an implementation. Indeed, an implementation of the simplex method that avoids explicit calculation of the matrix $B^{-1}N$ is referred to as an implementation of the revised simplex method. We shall see in Chapter 8 why it is beneficial to avoid computing $B^{-1}N$.