Linear Programming in Matrix Form

Appendix B

We first introduce matrix concepts in linear programming by developing a variation of the simplex method called the *revised simplex method*. This algorithm, which has become the basis of all commercial computer codes for linear programming, simply recognizes that much of the information calculated by the simplex method at each iteration, as described in Chapter 2, is not needed. Thus, efficiencies can be gained by computing only what is absolutely required.

Then, having introduced the ideas of matrices, some of the material from Chapters 2,3, and 4 is recast in matrix terminology. Since matrices are basically a notational convenience, this reformulation provides essentially nothing new to the simplex method, the sensitivity analysis, or the duality theory. However, the economy of the matrix notation provides added insight by streamlining the previous material and, in the process, highlighting the fundamental ideas. Further, the notational convenience is such that extending some of the results of the previous chapters becomes more straightforward.

B.1 A PREVIEW OF THE REVISED SIMPLEX METHOD

The revised simplex method, or the simplex method *with multipliers*, as it is often referred to, is a modification of the simplex method that significantly reduces the total number of calculations that must be performed at each iteration of the algorithm. Essentially, the revised simplex method, rather than updating the entire tableau at each iteration, computes *only* those coefficients that are needed to identify the pivot element. Clearly, the reduced costs must be determined so that the entering variable can be chosen. However, the variable that leaves the basis is determined by the minimum-ratio rule, so that only the updated coefficients of the entering variable and the current righthand-side values are needed for this purpose. The revised simplex method then keeps track of only enough information to compute the reduced costs and the minimum-ratio rule at each iteration.

The motivation for the revised simplex method is closely related to our discussion of simple sensitivity analysis in Section 3.1, and we will re-emphasize some of that here. In that discussion of sensitivity analysis, we used the shadow prices to help evaluate whether or not the contribution from engaging in a new activity was sufficient to justify diverting resources from the current optimal group of activities. The procedure was essentially to ''price out'' the new activity by determining the opportunity cost associated with introducing one unit of the new activity, and then comparing this value to the contribution generated by engaging in one unit of the activity. The opportunity cost was determined by valuing each resource consumed, by introducing one unit of the new activity, at the shadow price associated with that resource. The custom-molder example used in Chapter 3 to illustrate this point is reproduced in Tableau B.1. Activity 3, producing one hundred cases of champagne glasses, consumes 8 hours of production capacity and 10 hundred cubic feet of storage space. The shadow prices, determined in Chapter 3, are $\$\frac{11}{14}$ per hour of hour of production time and $\$\frac{1}{35}$ per hundred cubic feet of storage capacity, measured in hundreds of dollars. The resulting opportunity cost of

Tableau B.1

Basic variables	*Current values*	x_1	x_2	x_3	x_4	x_5	x_6	$-z$
x_4	60	6	5	8	1			0
x_5	150	10	20	10		1		0
x_6	8	1	0	0			1	0
$(-z)$	0	5	4.5	6				1

diverting resources to produce champagne glasses is then:

$$\left(\tfrac{11}{14}\right) 8 + \left(\tfrac{1}{35}\right) 10 = \tfrac{46}{7} = 6\tfrac{4}{7}.$$

Comparing this opportunity cost with the \$6 contribution results in a net loss of $\$\frac{4}{7}$ per case, or a loss of $\$57\frac{1}{7}$ per one hundred cases. It would clearly not be advantageous to divert resources from the current basic solution to the new activity. If, on the other hand, the activity had priced out positively, then bringing the new activity into the current basic solution would appear worthwhile. The essential point is that, by using the shadow prices and the original data, it is possible to decide, without elaborate calculations, whether or not a new activity is a promising candidate to enter the basis.

It should be quite clear that this procedure of pricing out a new activity is not restricted to knowing in advance whether the activity will be promising or not. The pricing-out mechanism, therefore, could in fact be used at each iteration of the simplex method to compute the reduced costs and choose the variable to enter the basis. To do this, we need to define shadow prices at each iteration.

In transforming the initial system of equations into another system of equations at some iteration, we maintain the canonical form at all times. As a result, the objective-function coefficients of the variables that are currently basic are zero at each iteration. We can therefore define simplex multipliers, which are essentially the shadow prices associated with a particular basic solution, as follows:

> ***Definition.*** The simplex multipliers $(y_1, y_2, \ldots, y_m)$ associated with a particular basic solution are the multiples of their initial system of equations such that, when all of these equations are multiplied by their respective simplex multipliers and subtracted from the initial objective function, the coefficients of the basic variables are zero.

Thus the basic variables must satisfy the following system of equations:

$$y_1 a_{1j} + y_2 a_{2j} + \cdots + y_m a_{mj} = c_j \qquad \text{for } j \text{ basic.}$$

The implication is that the reduced costs associated with the nonbasic variables are then given by:

$$\bar{c}_j = c_j - (y_1 a_{1j} + y_2 a_{2j} + \cdots + y_m a_{mj}) \qquad \text{for } j \text{ nonbasic.}$$

If we then performed the simplex method in such a way that we knew the simplex multipliers, it would be straightforward to find the largest reduced cost by pricing out all of the nonbasic variables and comparing them.

In the discussion in Chapter 3 we showed that the shadow prices were readily available from the final system of equations. In essence, since varying the righthand side value of a particular constraint is similar to adjusting the slack variable, it was argued that the shadow prices are the negative of the objective-function coefficients of the slack (or artificial) variables in the final system of equations. Similarly, the simplex multipliers at each intermediate iteration are the negative of the objective-function coefficients of these variables. In transforming the initial system of equations into any intermediate system of equations, a number of iterations of the simplex method are performed, involving subtracting multiples of a row from the objective function at each iteration. The simplex multipliers, or shadow prices in the final tableau, then

Tableau B.2

Basic variables	*Current values*	x_4	x_5	x_6
x_2	$4\frac{2}{7}$	$-\frac{1}{7}$	$\frac{3}{35}$	
x_6	$1\frac{4}{7}$	$-\frac{2}{7}$	$\frac{1}{14}$	1
x_1	$6\frac{3}{7}$	$\frac{2}{7}$	$-\frac{1}{14}$	
$(-z)$	$-51\frac{3}{7}$	$-\frac{11}{14}$	$-\frac{1}{35}$	

reflect a summary of all of the operations that were performed on the objective function during this process. If we then keep track of the coefficients of the slack (or artificial) variables in the objective function at each iteration, we immediately have the necessary simplex multipliers to determine the reduced costs $\overline{c}_j$ of the nonbasic variables as indicated above. Finding the variable to introduce into the basis, say x_s, is then easily accomplished by choosing the maximum of these reduced costs.

The next step in the simplex method is to determine the variable x_r to drop from the basis by applying the minimum-ratio rule as follows:

$$\frac{\overline{b}_r}{\overline{a}_{rs}} = \operatorname*{Min}_i \left\{ \frac{\overline{b}_i}{\overline{a}_{is}} \middle| \overline{a}_{is} > 0 \right\}.$$

Hence, in order to determine which variable to drop from the basis, we need both the current righthand-side values and the current coefficients, in each equation, of the variable we are considering introducing into the basis. It would, of course, be easy to keep track of the current righthand-side values for each iteration, since this comprises only a single column. However, if we are to *significantly reduce* the number of computations performed in carrying out the simplex method, we cannot keep track of the coefficients of each variable in each equation on every iteration. In fact, the only coefficients we need to carry out the simplex method are those of the variable to be introduced into the basis at the current iteration. If we could find a way to generate these coefficients after we knew which variable would enter the basis, we would have a genuine economy of computation over the standard simplex method.

It turns out, of course, that there is a way to do exactly this. In determining the new reduced costs for an iteration, we used only the initial data and the simplex multipliers, and further, the simplex multipliers were the negative of the coefficients of the slack (or artificial) variables in the objective function at that iteration. In essence, the coefficients of these variables summarize all of the operations performed on the objective function. Since we began our calculations with the problem in canonical form with respect to the slack (or artificial) variables and $-z$ in the objective function, it would seem intuitively appealing that the coefficients of these variables in any equation summarize all of the operations performed on that equation.

To illustrate this observation, suppose that we are given only the part of the final tableau that corresponds to the slack variables for our custom-molder example. This is reproduced from Chapter 3 in Tableau B.2.

In performing the simplex method, multiples of the equations in the initial Tableau B.1 have been added to and subtracted from one another to produce the final Tableau B.2. What multiples of Eq. 1 have been added to Eq. 2? Since x_4 is isolated in Eq. 1 in the initial tableau, any multiples of Eq. 1 that have been added to Eq. 2 must appear as the coefficient of x_4 in Eq. 2 of the final tableau. Thus, without knowing the actual sequence of pivot operations, we know their *net* effect has been to subtract $\frac{2}{7}$ times Eq. 1 from Eq. 2 in the initial tableau to produce the final tableau. Similarly, we can see that $\frac{2}{7}$ times Eq. 1 has been added to Eq. 3 Finally, we see that Eq. 1 (in the initial tableau) has been scaled by multiplying it by $-\frac{1}{7}$ to produce the final tableau. The coefficient of x_5 and x_6 in the final tableau can be similarly interpreted as the multiples of Eqs. 2 and 3, respectively, in the initial tableau that have been added to each equation of the initial tableau to produce the final tableau.

We can summarize these observations by remarking that the equations of the final tableau must be given in terms of multiples of the equations of the initial tableau as follows:

$$\begin{aligned}
\text{Eq. 1:}\quad & (-\tfrac{1}{7})(\text{Eq. 1}) + (\tfrac{3}{35})(\text{Eq. 2}) + 0(\text{Eq. 3});\\
\text{Eq. 2:}\quad & (-\tfrac{2}{7})(\text{Eq. 1}) + (\tfrac{1}{14})(\text{Eq. 2}) + 1(\text{Eq. 3});\\
\text{Eq. 3:}\quad & (\tfrac{2}{7})(\text{Eq. 1}) + (-\tfrac{1}{14})(\text{Eq. 2}) + 0(\text{Eq. 3}).
\end{aligned}$$

The coefficients of the slack variables in the final tableau thus summarize the operations performed on the equations of the initial tableau to produce the final tableau.

We can now use this information to determine the coefficients in the final tableau of x_3, the production of champagne glasses, from the initial tableau and the coefficients of the slack variables in the final tableau. From the formulas developed above we have:

$$\begin{aligned}
\text{Eq.1:}\quad & (-\tfrac{1}{7})(8) + (\tfrac{3}{35})(10) + (0)(0) = -\tfrac{2}{7};\\
\text{Eq.2:}\quad & (-\tfrac{2}{7})(8) + (\tfrac{1}{14})(10) + (1)(0) = -\tfrac{11}{7};\\
\text{Eq.3:}\quad & (\tfrac{2}{7})(8) + (-\tfrac{1}{14})(10) + (0)(0) = \tfrac{11}{7}.
\end{aligned}$$

The resulting values are, in fact, the appropriate coefficients of x_3 for each of the equations in the final tableau, as determined in Chapter 3. Hence, we have found that it is only necessary to keep track of the coefficients of the slack variables at each iteration.

The coefficients of the slack variables are what is known as the ''inverse'' of the current basis. To see this relationship more precisely, let us multiply the matrix* corresponding to the slack variables by the matrix of columns from the initial tableau corresponding to the current basis, arranged in the order in which the variables are basic. In matrix notation, we can write these multiplications as follows:

$$\begin{bmatrix} -\frac{1}{7} & \frac{3}{35} & 0 \\ -\frac{2}{7} & \frac{1}{14} & 1 \\ \frac{2}{7} & -\frac{1}{14} & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 & 6 \\ 20 & 0 & 10 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which, in symbolic form, is $B^{-1}B = I$. The information contained in the coefficients of the slack variables is then the *inverse* of the current basis, since the multiplication produces the identity matrix. In general, to identify the basis corresponding to the inverse, it is only necessary to order the variables so that they correspond to the rows in which they are basic. In this case, the order is x_2, x_6, and x_1.

In matrix notation, the coefficients $\overline{A}_j$ of any column in the current tableau can then be determined from their coefficients A_j in the initial tableau and the basis inverse by $B^{-1}A_j$. For example,

$$\overline{A}_3 = B^{-1}A_3 = \begin{bmatrix} -\frac{1}{7} & \frac{3}{35} & 0 \\ -\frac{2}{7} & \frac{1}{14} & 0 \\ \frac{2}{7} & -\frac{1}{14} & 0 \end{bmatrix} \begin{bmatrix} 8 \\ 10 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{7} \\ -\frac{11}{7} \\ \frac{11}{7} \end{bmatrix}.$$

If we now consider the righthand-side values as just another column b in the initial tableau, we can, by analogy, determine the current righthand–side values by:

$$\overline{b} = B^{-1}b = \begin{bmatrix} -\frac{1}{7} & \frac{3}{35} & 0 \\ -\frac{2}{7} & \frac{1}{14} & 0 \\ \frac{2}{7} & -\frac{1}{14} & 1 \end{bmatrix} \begin{bmatrix} 60 \\ 150 \\ 8 \end{bmatrix} = \begin{bmatrix} 4\frac{2}{7} \\ 1\frac{4}{7} \\ 6\frac{3}{7} \end{bmatrix}.$$

*A discussion of vectors and matrices is included in Appendix A.

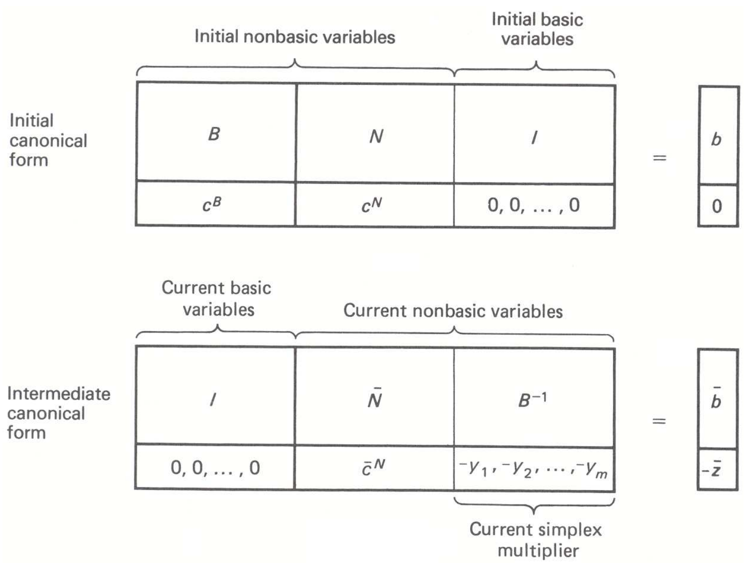

Figure B.1

Hence, had x_3 been a promising variable to enter the basis, we could have easily computed the variable to drop from the basis by the minimum-ratio rule, once we had determined $\overline{A}_3$ and $\overline{b}$.

In performing the revised simplex method, then, we need not compute all of the columns of the tableau at each iteration. Rather, we need only keep track of the coefficients of the slack variables in all the equations including the objective function. These coefficients contain the simplex multipliers y and inverse of the current basis B^{-1}. Using the original data and the simplex multipliers, the reduced costs can be calculated easily and the entering variable selected. Using the original data and the inverse of the current basis, we can easily calculate the coefficients of the entering variable in the current tableau and the current righthand-side values. The variable to drop from the basis is then selected by the minimum-ratio rule. Finally, the basis inverse and the simplex multipliers are updated by performing the appropriate pivot operation on the current tableau, as will be illustrated, and the procedure then is repeated.

B.2 FORMALIZING THE APPROACH

We can formalize the ideas presented in the previous section by developing the revised simplex method in matrix notation. We will assume from here on that the reader is familiar with the first three sections of the matrix material presented in Appendix A.

At any point in the simplex method, the initial canonical form has been transformed into a new canonical form by a sequence of pivot operations. The two canonical forms can be represented as indicated in Fig. B.1.

The basic variables in the initial basis may include artificial variables, as well as variables from the original problem; the initial canonical form is determined by the procedures discussed in Chapter 2. The new basis may, of course, contain variables in common with the initial basis. To make Fig. B.1 strictly correct, we need to imagine including some of these variables *twice*.

We can derive the intermediate tableau from the initial tableau in a straight-forward manner. The initial system of equations in matrix notation is:

$$
\begin{aligned}
x^B \geq 0, \qquad x^N &\geq 0, \qquad x^I \geq 0, \\
Bx^B + Nx^N + Ix^I \qquad &= b, \\
c^B x^B + c^N x^N \qquad &- z = 0,
\end{aligned} \tag{1}
$$

where the superscripts B, N and I refer to basic variables, nonbasic variables, and variables in the initial identity basis, respectively. The constraints of the intermediate system of equations are determined by multiplying the constraints of the initial system (1) on the left by B^{-1}. Hence,

$$Ix^B + B^{-1}Nx^N + B^{-1}x^I = B^{-1}b, \tag{2}$$

which implies that the updated nonbasic columns and righthand-side vector of the intermediate canonical forms are given $\overline{N} = B^{-1}N$ and $\overline{b} = B^{-1}b$, respectively.

Since the objective function of the intermediate canonical form must have zeros for the coefficients of the basic variables, this objective function can be determined by multiplying each equation of (2) by the cost of the variable that is basic in that row and subtracting the resulting equations from the objective function of (1). In matrix notation, this means multiplying (2) by c^B on the left and subtracting from the objective function of (1), to give:

$$0x^B + (c^N - c^B B^{-1}N)x^N - c^B B^{-1}x^I - z = -c^B B^{-1}b. \tag{3}$$

We can write the objective function of the intermediate tableau in terms of the simplex multipliers by recalling that the simplex multipliers are defined to be the *multiples of* the equations in the initial tableau that *produce zero* for the coefficients of the basic variables when *subtracted from* the initial objective function. Hence,

$$c^B - yB = 0 \qquad \text{which implies} \qquad y = c^B B^{-1}. \tag{4}$$

If we now use (4) to rewrite (3) we have

$$0x^B + (c^N - yN)x^N - yx^I - z = -yb, \tag{5}$$

which corresponds to the intermediate canonical form in Fig. B.1. The coefficients in the objective function of variables in the initial identity basis are the negative of the simplex multipliers as would be expected.

Note also that, since the matrix N is composed of the nonbasic columns A_j from the initial tableau, the relation $\overline{N} = B^{-1}N$ states that each updated column $\overline{A}_j$ of $\overline{N}$ is given by $\overline{A}_j = B^{-1}A_j$. Equivalently, $A_j = B\overline{A}_j$ or

$$A_j = B_1\overline{a}_{1j} + B_2\overline{a}_{2j} + \cdots + B_m\overline{a}_{mj}.$$

This expression states that the column vector A_j can be written as a linear combination of columns $B_1, B_2, \ldots, B_m$ of the basis, using the weights $\overline{a}_{1j}, \overline{a}_{2j}, \ldots, \overline{a}_{mj}$. In vector terminology, we express this by saying that the column vector

$$\overline{A}_j = \langle \overline{a}_{1j}, \overline{a}_{2j}, \ldots, \overline{a}_{mj} \rangle$$

is the *representation* of A_j in terms of the basis B.

Let us review the relationships that we have established in terms of the simplex method. Given the current canonical form, the current basic feasible solution is obtained by setting the nonbasic variables to their lower bounds, in this case zero, so that:

$$x^B = B^{-1}b = \overline{b} \geq 0, \qquad x^N = 0.$$

The value of the objective function associated with this basis is then

$$z = yb = \overline{z}.$$

To determine whether or not the current solution is optimal, we look at the reduced costs of the nonbasic variables.

$$\overline{c}_j = c_j - yA_j.$$

If $\overline{c}_j \leq 0$ for all j nonbasic, then the current solution is optimal. Assuming that the current solution is not optimal and that the maximum $\overline{c}_j$ corresponds to x_s, then, to determine the pivot element, we need the representation of the entering column A_s in the current basis and the current righthand side,

$$\overline{A}_s = B^{-1}A_s \qquad \text{and} \qquad \overline{b} = B^{-1}b, \qquad \text{respectively.}$$

If $\overline{c}_s > 0$ and $\overline{A}_s \leq 0$, the problem is unbounded. Otherwise, the variable to drop from the basis x_r is determined by the usual minimum-ratio rule. The new canonical form is then found by pivoting on the element $\overline{a}_{rs}$.

Note that, at each iteration of the simplex method, only the column corresponding to the variable entering the basis needs to be computed. Further, since this column can be obtained by $B^{-1}A_s$, only the initial data and the inverse of the current basis need to be maintained. Since the inverse of the current basis can be obtained from the coefficients of the variables that were slack (or artificial) in the initial tableau, we need only perform the pivot operation on these columns to obtain the updated basis inverse. This computational efficiency is the foundation of the revised simplex method.

Revised Simplex Method

STEP (0) : An initial basis inverse B^{-1} is given with $\overline{b} = B^{-1}b \geq 0$. The columns of B are $[A_{j_1}, A_{j_2}, \ldots, A_{j_m}]$ and $y = c^B B^{-1}$ is the vector of simplex multipliers.

STEP (1) : The coefficients of $\overline{c}$ for the nonbasic variables x_j are computed by pricing out the original data A_j, that is,

$$\overline{c}_j = c_j - yA_j = c_j - \sum_{i=1}^{m} y_i a_{ij} \quad \text{for } j \text{ nonbasic.}$$

If all $\overline{c}_j \leq 0$ then *stop*; we are optimal. If we continue, then there exists some $\overline{c}_j > 0$.

STEP (2) : Choose the variable to introduce into the basis by

$$\overline{c}_s = \operatorname*{Max}_j \{\overline{c}_j | \overline{c}_j > 0\}.$$

Compute $\overline{A}_s = B^{-1}A_s$. If $\overline{A}_s \leq 0$, then *stop*; the problem is unbounded. If we continue, there exists $\overline{a}_{is} > 0$ for some $i = 1, 2, \ldots, m$.

STEP (3) : Choose the variable to drop from the basis by the minimum-ratio rule:

$$\frac{\overline{b}_r}{\overline{a}_{rs}} = \text{Min}_i \left\{ \frac{\overline{b}_i}{\overline{a}_{is}} \middle| \overline{a}_{is} > 0 \right\}.$$

The variable basic in row r is replaced by variable s giving the new basis $B = [A_{j_i}, \ldots, A_{j_{r-1}}, A_s, A_{j_{r+1}}, \ldots, A_{j_m}]$.

STEP (4) : Determine the new basis inverse B^{-1}, the new righthand-side vector $\overline{b}$, and new vector of simplex multipliers $y = c^B B^{-1}$, by pivoting on $\overline{a}_{rs}$.

STEP (5) : Go to STEP (1).

We should remark that the initial basis in STEP (0) usually is composed of slack variables and artificial variables constituting an identity matrix, so that $B = I$ and $B^{-1} = I$, also. The more general statement of the algorithm is given, since, after a problem has been solved once, a good starting feasible basis is generally known, and it is therefore unnecessary to start with the identity basis.

The only detail that remains to be specified is how the new basis inverse, simplex multipliers, and righthand-side vector are generated in STEP (4). The computations are performed by the usual simplex pivoting procedure, as suggested by Fig. B.2. We know that the basis inverse for any canonical form is always given by the coefficients of the slack variables in the initial tableau. Consequently, the new basis inverse will be given by pivoting in the tableau on $\overline{a}_{rs}$ as usual. Observe that whether we compute the new

	B^{-1}				$\overline{b}$
$\overline{a}_{1s}$	β_{11}	β_{12}	$\cdots$	β_{1m}	b_1
$\vdots$	$\vdots$			$\vdots$	$\vdots$
$\overline{a}_{rs}$	β_{r1}	β_{r2}		β_{rm}	$\overline{b}_r$
$\vdots$	$\vdots$			$\vdots$	$\vdots$
$\overline{a}_{ms}$	β_{m1}	β_{m2}	$\cdots$	β_{mm}	$\overline{b}_m$
$\overline{c}_s$	$-y_1$	$-y_2$	$\cdots$	$-y_m$	$-\overline{z}$

Figure B.2 Updating the basis inverse and simplex multipliers.

columns $\overline{A}_j$ for $j \neq s$ or not, pivoting has the same effect upon B^{-1} and $\overline{b}$. Therefore we need only use the reduced tableau of Fig. B.2.

After pivoting, the coefficients in place of the β_{ij} and $\overline{b}_i$ will be, respectively, the new basis inverse and the updated righthand-side vector, while the coefficients in the place of $-y$ and $-\overline{z}$ will be the new simplex multipliers and the new value of the objective function, respectively. These ideas will be reinforced by looking at the example in the next section.

B.3 THE REVISED SIMPLEX METHOD—AN EXAMPLE

To illustrate the procedures of the revised simplex method, we will employ the same example used at the end of Chapter 2. It is important to keep in mind that the revised simplex method is merely a modification of the simplex method that performs fewer calculations by computing only those quantities that are essential to carrying out the steps of the algorithm. The initial tableau for our example is repeated as Tableau B.3. Note that the example has been put in canonical form by the addition of artificial variables, and the necessary Phase I objective function is included.

Tableau B.3 Initial data tableau.

										Artificial variables		
Basic variables	*Current values*	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
x_9	4	1	-1	1	-1	-4	2	-1		1		
x_8	6	-3	3	1	-1	-2	0	0	1			
x_{10}	1	0	0	-1	1	0	1	0			1	
x_{11}	0	1	-1	1	-1	-1	0	0				1
$(-z)$	0	-3	3	2	-2	-1	4	0				
$(-w)$	5	2	-2	1	-1	-5	3	-1				

At each iteration of the revised simplex method, the current inverse of the basis, and a list of the basic variables and the rows in which they are basic, must be maintained. This information, along with the initial tableau, is sufficient to allow us to carry out the steps of the algorithm. As in Chapter 2, we begin with the Phase I objective, maximizing the negative of the sum of the artificial variables, and carry along the Phase II objective in canonical form. Initially, we have the identity basis consisting of the slack variable x_8, the artificial variables x_9, x_{10}, and x_{11}, as well as $-z$ and $-w$, the Phase II and Phase I objective values, respectively. This identity basis is shown in Tableau B.4. Ignore for the moment the column labeled x_6, which has been appended.

Now, to determine the variable to enter the basis, find $\overline{d}_s = \text{Max}\, \overline{d}_j$ for j nonbasic. From the initial tableau, we can see that $\overline{d}_s = d_6 = 3$, so that variable x_6 will enter the basis and is appended to the current

Tableau B.4 Initial basis.

Basic variables	*Current values*	x_9	x_8	x_{10}	x_{11}	x_6	*Ratio*
x_9	4	1				2	$\frac{4}{2}$
x_8	6		1			0	
x_{10}	1			1		1	$\frac{1}{1}$
x_{11}	0				1	0	
$(-z)$	0					4	
$(-w)$	5					3	

basis in Tableau B.4. We determine the variable to drop from the basis by the minimum-ratio rule:

$$\frac{\overline{b}_r}{\overline{a}_{r6}} = \operatorname*{Min}_i \left\{ \frac{\overline{b}_i}{\overline{a}_{i6}} \middle| \overline{a}_{i6} > 0 \right\} = \text{Min} \left\{ \frac{4}{2}, \frac{1}{1} \right\} = 1.$$

Since the minimum ratio occurs in row 3, variable x_{10} drops from the basis. We now perform the calculations implied by bringing x_3 into the basis and dropping x_{10}, but only on that portion of the tableau where the basis inverse will be stored. To obtain the updated inverse, we merely perform a pivot operation to transform the coefficients of the incoming variable x_6 so that a canonical form is maintained. That is, the column labeled x_6 should be transformed into all zeros except for a *one* corresponding to the circled pivot element, since variable x_6 enters the basis and variable x_{10} is dropped from the basis. The result is shown in Tableau B.5 including the list indicating which variable is basic in each row. Again ignore the column labeled x_3 which has been appended.

Tableau B.5 After iteration 1.

Basic variables	*Current values*	x_9	x_8	x_{10}	x_{11}	x_3	*Ratio*
x_9	2	1		−2		3	$\frac{2}{3}$
x_8	6		1	0		1	$\frac{6}{1}$
x_6	1			1		−1	
x_{11}	0			0	1	1	$\frac{0}{1}$
$(-z)$	−4			−4		6	
$(-w)$	2			−3		4	

We again find the maximum reduced cost, $\overline{d}_s = \text{Max}\, \overline{d}_j$, for j nonbasic, where $\overline{d}_j = d_j - yA_j$. Recalling that the simplex multipliers are the negative of the coefficients of the slack (artificial) variables in the objective fuction, we have $y = (0, 0, 3, 0)$. We can compute the reduced costs for the nonbasic variables from: $\overline{d}_j = d_j - y_1a_{1j} - y_2a_{2j} - y_3a_{3j} - y_4a_{4j}$, which yields the following values:

$\overline{d}_1$	$\overline{d}_2$	$\overline{d}_3$	$\overline{d}_4$	$\overline{d}_5$	$\overline{d}_7$	$\overline{d}_{10}$
2	−2	4	−4	−5	−1	−3

Since $\overline{d}_3 = 4$ is the largest reduced cost, x_3 enters the basis. To find the variable to drop from the basis, we have to apply the minimum-ratio rule:

$$\frac{\overline{b}_r}{\overline{a}_{r3}} = \operatorname*{Min}_i \left\{ \frac{\overline{b}_i}{\overline{a}_{i3}} \middle| \overline{a}_{i3} > 0 \right\}.$$

Now to do this, we need the representation of A_3 in the current basis. For this calculation, we consider the

Phase II objective function to be a constraint, but never allow $-z$ to drop from the basis.

$$\overline{A}_3 = B^{-1}A_3 = \begin{bmatrix} 1 & & -2 & & \\ & 1 & 0 & & \\ & & 1 & & \\ & & 0 & 1 & \\ & & -4 & & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \\ 6 \end{bmatrix}$$

Hence,

$$\frac{\overline{b}_1}{\overline{a}_{13}} = \frac{2}{3}, \qquad \frac{\overline{b}_2}{\overline{a}_{23}} = \frac{6}{1}, \qquad \frac{\overline{b}_4}{\overline{a}_{43}} = \frac{0}{1},$$

so that the minimum ratio is zero and variable x_{11}, which is basic in row 4, drops from the basis. The updated tableau is found by a pivot operation, such that the column $\overline{A}_3$ is transformed into a column containing all zeros except for a *one* corresponding to the circled in pivot element in Tableau B.5, since x_3 enters the basis and x_{11} drops from the basis. The result (ignoring the column labeled x_2) is shown in Tableau B.6.

Tableau B.6 After iteration 2

Basic variables	*Current values*	x_9	x_8	x_{10}	x_{11}	x_2	*Ratio*
x_9	2	1		−2	−3	2	$\frac{2}{2}$
x_8	6		1	0	−1	4	$\frac{6}{4}$
x_6	1			1	1	−1	
x_3	0			0	1	−1	
$(-z)$	−4			−4	6	9	
$(-w)$	2			−3	−4	2	

Now again find $\overline{d}_s =$ Max $\overline{d}_j$ for j nonbasic. Since $y = (0, 0, 3, 4)$ and $\overline{d}_j = d_j - yA_j$, we have

Since $\overline{d}_2 = 2$ is the largest reduced cost, x_2 enters the basis. To find the variable to drop from the basis, we find the representation of A_2 in the current basis:

$$\overline{A}_2 = B^{-1}A_2 = \begin{bmatrix} 1 & & -2 & -3 & \\ & 1 & 0 & -1 & \\ & & 1 & 1 & \\ & & 0 & 1 & \\ & & -4 & -6 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \\ 0 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -1 \\ -1 \\ 9 \end{bmatrix},$$

and append it to Tableau B.6. Applying the minimum-ratio rule gives:

$$\frac{\overline{b}_1}{\overline{a}_{12}} = \frac{2}{2}, \qquad \frac{\overline{b}_2}{\overline{a}_{22}} = \frac{6}{4},$$

and x_9, which is basic in row 1, drops from the basis. Again a pivot is performed on the circled element in the column labeled x_2 in Tableau B.6, which results in Tableau B.7 (ignoring the column labeled x_5).

Since the value of the Phase I objective is equal to zero, we have found a feasible solution. We end Phase I, dropping the Phase I objective function from any further consideration, and proceed to Phase II. We must now find the maximum reduced cost of the Phase II objective function. That is, find $\overline{c}_s =$ Max $\overline{c}_{j,}$ for j nonbasic, where $\overline{c}_s = c_j - yA_{j.}$ The simplex multipliers to initiate Phase II are the negative of the coefficients

$\overline{d}_1$	$\overline{d}_2$	$\overline{d}_4$	$\overline{d}_5$	$\overline{d}_7$	$\overline{d}_{10}$	$\overline{d}_{11}$
−2	2	0	−1	−1	−3	−4

Tableau B.7 After iteration 3

Basic variables	*Current values*	x_9	x_8	x_{10}	x_{11}	x_5
x_2	1	$\frac{1}{2}$		-1	$-\frac{3}{2}$	$-\frac{1}{2}$
x_8	2	-2	1	4	5	1
x_6	2	$\frac{1}{2}$		0	$-\frac{1}{2}$	$-\frac{3}{2}$
x_3	1	$\frac{1}{2}$		-1	$-\frac{1}{2}$	$-\frac{3}{2}$
$(-z)$	12	$-\frac{9}{2}$		5	$\frac{15}{2}$	$\frac{-9}{2}$
$(-w)$	0	-1		-1	-1	

$\overline{c}_1$	$\overline{c}_4$	$\overline{c}_5$	$\overline{c}_7$	$\overline{c}_9$	$\overline{c}_{10}$	$\overline{c}_{11}$
0	0	$\frac{19}{2}$	$\frac{9}{2}$	$-\frac{19}{2}$	5	$\frac{15}{2}$

of the slack (artificial) variables in the $-z$ equation. Therefore, $y = (\frac{9}{2}, 0, -5, -\frac{15}{2})$, and the reduced costs for the nonbasic variables are: Since $\overline{c}_5$ is the maximum reduced cost, variable x_5 enters the basis. To find the variable to drop from the basis, compute:

$$\overline{A}_5 = B^{-1}A_5 = \begin{bmatrix} \frac{1}{2} & & -1 & -\frac{3}{2} \\ -2 & 1 & 4 & 5 \\ \frac{1}{2} & & 0 & -\frac{1}{2} \\ \frac{1}{2} & & -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} -4 \\ -2 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ -\frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}.$$

Since only $\overline{a}_{25}$ is greater than zero, variable x_8, which is basic in row 2, drops from the basis. Again a pivot operation is performed on the circled element in the column labeled x_5 in Tableau B.7 and the result is shown in Tableau B.8.

Tableau B.8 Final reduced tableau

Basic variables	*Current values*	x_9	x_8	x_{10}	x_{11}
x_2	2	$-\frac{1}{2}$	$\frac{1}{2}$	1	1
x_5	2	-2	1	4	5
x_6	5	$-\frac{5}{2}$	$\frac{3}{2}$	6	7
x_3	4	$-\frac{5}{2}$	$\frac{3}{2}$	5	7
$(-z)$	32	$\frac{29}{2}$	$-\frac{19}{2}$	-33	-40

We again find $\overline{c}_s = \text{Max}\,\overline{c}_j$ for j nonbasic. Since $y = (-\frac{29}{2}, \frac{19}{2}, 33, 40)$ and $\overline{c}_j = c_j - yA_j$, we have: The only positive reduced cost is associated with the artificial variable x_9. Since artificial variables are

$\overline{c}_1$	$\overline{c}_4$	$\overline{c}_7$	$\overline{c}_8$	$\overline{c}_9$	$\overline{c}_{10}$	$\overline{c}_{11}$
0	0	$-\frac{29}{2}$	$-\frac{19}{2}$	$\frac{29}{2}$	-33	-40

never reintroduced into the basis once they have become nonbasic, we have determined an optimal solution $x_2 = 2, x_5 = 2, x_6 = 5, x_3 = 4$, and $z = -32$.

Finally, it should be pointed out that the sequence of pivots produced by the revised simplex method is exactly the same as that produced by the usual simplex method. (See the identical example in Chapter 2).

*B.4 COMPUTER CONSIDERATIONS AND THE PRODUCT FORM

The revised simplex method is used in essentially all commercial computer codes for linear programming, both for computational and storage reasons.

For any problem of realistic size, the revised simplex method makes fewer calculations than the ordinary simplex method. This is partly due to the fact that, besides the columns corresponding to the basis inverse and the righthand side, only the column corresponding to the variable entering the basis needs to be computed at each iteration. Further, in pricing out the nonbasic columns, the method takes advantage of the low density of nonzero elements in the initial data matrix of most real problems, since the simplex multipliers need to be multiplied only by the nonzero coefficients in a nonbasic column. Another reason for using the revised simplex method is that roundoff error tends to accumulate in performing these algorithms. Since the revised simplex method maintains the original data, the inverse of the basis may be recomputed from this data periodically, to significantly reduce this type of error. Many large problems could not be solved without such a periodic reinversion of the basis to reduce roundoff error.

Equally important is the fact that the revised simplex method usually requires less storage than does the ordinary simplex method. Besides the basis inverse B^{-1} and the current righthand-side vector $\overline{b}$, which generally contain few zeros, the revised simplex method must store the original data. The original data, on the other hand, generally contains many zeros and can be stored compactly using the following methods. First, we eliminate the need to store zero coefficients, by packing the nonzero coefficients in an array, with reference pointers indicating their location. Second, often the number of significant digits in the original data is small—say, three or fewer—so that these can be handled compactly by storing more than one coefficient in a computer word. In contrast, eight to ten significant digits must be stored for every nonzero coefficient in a canonical form of the usual simplex method, and most coefficients will be nonzero.

There is one further refinement of the revised simplex method that deserves mention, since it was a fundamental breakthrough in solving relatively large-scale problems on second-generation computers. The *product form* of the inverse was developed as an efficient method of storing and updating the inverse of the current basis, when this inverse has to be stored on a peripheral device.

$$\begin{array}{|c|cccc|c|}
\multicolumn{1}{c}{} & \multicolumn{4}{c}{\overbrace{\hphantom{\beta_{11}\quad\beta_{12}\quad\cdots\quad\beta_{1m}}}^{B^{-1}}} & \multicolumn{1}{c}{\overbrace{\hphantom{\overline{b}_1}}^{\overline{b}}} \\
\hline
\overline{a}_{1s} & \beta_{11} & \beta_{12} & \cdots & \beta_{1m} & \overline{b}_1 \\
\vdots & \vdots & & & \vdots & \vdots \\
\overline{a}_{rs} & \beta_{r1} & \beta_{r2} & \cdots & \beta_{rm} & \overline{b}_r \\
\vdots & \vdots & & & \vdots & \vdots \\
\overline{a}_{ms} & \beta_{m1} & \beta_{m2} & \cdots & \beta_{mm} & \overline{b}_m \\
\hline
\end{array}$$

Fig. B.3 Reduced tableau.

When recomputing B^{-1} and $\overline{b}$ in the revised simplex method, we pivot on $\overline{a}_{rs}$, in a reduced tableau, illustrated by Fig. B.3. The pivot operation first multiplies row r by $1/\overline{a}_{rs}$, and then subtracts $\overline{a}_{is}/\overline{a}_{rs}$ times row r from row i for $i = 1, 2, \ldots, m$ and $i \neq r$. Equivalently, pivoting premultiplies the above tableau by the elementary matrix.

$$E = \begin{array}{c} \begin{bmatrix} 1 & & & & \eta_1 & & & \\ & 1 & & & \eta_2 & & & \\ & & \ddots & & \vdots & & & \\ & & & 1 & & & & \\ & & & & \eta_r & & & \\ & & & & & 1 & & \\ & & & & \vdots & & \ddots & \\ & & & & \eta_m & & & 1 \end{bmatrix} \\ \uparrow \\ \text{Column } r \end{array}$$

where $\eta_r = 1/\overline{a}_{rs}$ and $\eta_i = -\overline{a}_{is}/\overline{a}_{rs}$ for $i \neq r$.

An elementary matrix is defined to be an identity matrix except for one column. If the new basis is B^*, then the new basis inverse $(B^*)^{-1}$ and new righthand-side vector $\overline{b}^*$ are computed by:

$$(B^*)^{-1} = EB^{-1} \quad \text{and} \quad \overline{b}^* = E\overline{b}. \tag{6}$$

After the next iteration, the new basis inverse and righthand-side vector will be given by premultiplying by another elementary matrix. Assuming that the initial basis is the identity matrix and letting E_j be the elementary matrix determined by the jth pivot step, after k iterations the basis inverse can be expressed as:

$$B^{-1} = E_k E_{k-1} \cdots E_2 E_1.$$

This *product form* of the inverse is used by almost all commercial linear-programming codes. In these codes, $\overline{b}$ is computed and maintained at each step by (6), but the basis inverse is not computed explicitly. Rather, the elementary matrices are stored and used in place of B^{-1}. These matrices can be stored compactly by recording only the special column $\langle \eta_1, \eta_2, \ldots, \eta_m \rangle$, together with a marker indicating the location of this column in E_j. Using a pivoting procedure for determining the inverse, B^{-1} can always be expressed as the product of no more than m elementary matrices. Consequently, when k is large, the product form for B^{-1} is recomputed. Special procedures are used in this calculation to express the inverse very efficiently and, consequently, to cut down on the number of computations required for the revised simplex method. The details are beyond the scope of our coverage here.

Since the basis inverse is used only for computing the simplex multipliers and finding the representation of the incoming column in terms of the basis, the elementary matrices are used only for the following two calculations:

$$y = \mathrm{c}^B B^{-1} = c^B E_k E_{k-1} \cdots E_1 \tag{7}$$

and

$$\overline{A}_s = B^{-1} A_s = E_k E_{k-1} \cdots E_1 A_s. \tag{8}$$

Most commercial codes solve problems so large that the problem data cannot be kept in the computer itself, but must be stored on auxiliary storage devices. The product form of the inverse is well suited for sequential-access devices such as magnetic tapes or drums. The matrices $E_k, E_{k-1}, \ldots, E_1$ are stored sequentially on the device and, by accessing the device in one direction, the elementary matrices are read in the order $E_k, E_{k-1}, \ldots, E_1$ and applied sequentially to c^B for computing (7). When rewinding the device, they are read in opposite order $E_1, E_2, \ldots, E_k$ and applied to A_s for computing (8). The new elementary matrix E_{k+1} is then added to the device next to E_k. Given this procedure and the form of the above calculations, (7) is sometimes referred to as the *b-tran* (backward transformation) and (8) as the *f-tran* (forward transformation).

B.5 SENSITIVITY ANALYSIS REVISITED

In Chapter 3 we gave a detailed discussion of sensitivity analysis in terms of a specific example. There the analysis depended upon recognizing certain relationships between the initial and final tableaus of the simplex method. Now that we have introduced the revised simplex method, we can review that discussion and make some of it more rigorous. Since the revised simplex method is based on keeping track of only the original data tableau, the simplex multipliers, the inverse of the current basis, and which variable is basic in each row, the final tableau for the simplex method can be computed from this information. Therefore, all the remarks that we made concerning sensitivity analysis may be derived formally by using these data.

We will review, in this section, varying the coefficients of the objective function, the values of the righthand side, and the elements of the coefficient matrix. We will not need to review our discussion of the existence of alternative optimal solutions, since no simplifications result from the introduction of matrices. Throughout this section, we assume that we have a maximization problem, and leave to the reader the derivation of the analogous results for minimization problems.

To begin with, we compute the ranges on the coefficients of the objective function so that the basis remains unchanged. Since only the objective-function coefficients are varied, and the values of the decision variables are given by $x^B = B^{-1}b$, these values remain unchanged. However, since the simplex multipliers are given by $y = c^B B^{-1}$, varying any of the objective-function coefficients associated with basic variables will alter the values of the simplex multipliers.

Suppose that variable x_j is *nonbasic*, and we let its coefficient in the objective function c_j be changed by an amount Δc_j, with all other data held fixed. Since x_j is currently not in the optimal solution, it should be clear that Δc_j may be made an arbitrarily large negative number without x_j becoming a candidate to enter the basis. On the other hand, if Δc_j is increased, x_j will not enter the basis so long as its new reduced cost $\overline{c}_j^{\text{new}}$ satisfies:

$$\overline{c}_j^{\text{new}} = c_j + \Delta c_j - yA_j \leq 0,$$

which implies that

$$\Delta c_j \leq yA_j - c_j = -\overline{c}_j,$$

or that

$$-\infty < c_j + \Delta c_j \leq yA_j. \tag{9}$$

At the upper end of the range, x_j becomes a candidate to enter the basis.

Now suppose that x_j is a *basic* variable and, further that it is basic in row i. If we let its coefficient in the objective function $c_j = c_i^B$ be changed by an amount Δc_i^B, the first thing we note is that the value of the simplex multipliers will be affected, since:

$$y = (c^B + \Delta c_i^B u^i)B^{-1}, \tag{10}$$

where u^i is a row vector of zeros except for a *one* in position i. The basis will not change so long as the reduced costs of all nonbasic variables satisfy:

$$\overline{c}_j^{\text{new}} = c_j - yA_j \leq 0. \tag{11}$$

Substituting in (11) for y given by (10),

$$\overline{c}_j^{\text{new}} = c_j - (c^B + \Delta c_i^B u^i)B^{-1}A_j \leq 0,$$

and noting that $B^{-1}A_j$ is just the representation of A_j in the current basis, we have

$$\overline{c}_j^{\text{new}} = c_j - (c^B + \Delta c_i^B u_i)\overline{A}_j \leq 0. \tag{12}$$

Condition (12) may be rewritten as:

$$\overline{c}_j - \Delta c_i^B u^i \overline{A}_j \leq 0,$$

which implies that:

$$\Delta c_i^B \geq \frac{\overline{c}_i}{\overline{a}_{ij}} \quad \text{for } \overline{a}_{ij} > 0,$$

and

$$\Delta c_i^B \leq \frac{\overline{c}_i}{\overline{a}_{ij}} \quad \text{for } \overline{a}_{ij} < 0, \tag{13}$$

are no limit for $\overline{a}_{ij} = 0$.

Finally, since (13) must be satisfied for all nonbasic variables, we can define upper and lower bounds on Δc_i^B as follows:

$$\operatorname*{Max}_j \left\{ \frac{\overline{c}_j}{\overline{a}_{ij}} \middle| a_{ij} > 0 \right\} \leq \Delta c_i^B \leq \operatorname*{Min}_j \left\{ \frac{\overline{c}_j}{\overline{a}_{ij}} \middle| \overline{a}_{ij} < 0 \right\}. \tag{14}$$

Note that, since $\overline{c}_j \leq 0$, the lower bound on Δc_i^B is nonpositive and the upper bound is nonnegative, so that the range on the cost coefficient $c_i^B + \Delta c_i^B$ is determined by adding c_i^B to each bound in (14). Note that Δc_i^B may be unbounded in either direction if there are no $\overline{a}_{ij}$ of appropriate sign.

At the upper bound in (14), the variable producing the minimum ratio is a candidate to enter the basis, while at the lower bound in (14), the variable producing the maximum ratio is a candidate to enter the basis. These candidate variables are clearly not the same, since they have opposite signs for $\overline{a}_{ij}$. In order for any candidate to enter the basis, the variable to drop from the basis x_r is determined by the usual minimum-ratio rule:

$$\frac{\overline{b}_r}{\overline{a}_{rs}} \operatorname*{Min}_i \left\{ \frac{\overline{b}_i}{\overline{a}_{is}} \middle| \overline{a}_{is} > 0 \right\}. \tag{15}$$

If $\overline{A}_s \leq 0$, then variable x_s can be increased without limit and the objective function is unbounded. Otherwise, the variable corresponding to the minimum ratio in (15) will drop from the basis if x_s is introduced into the basis.

We turn now to the question of variations in the righthand-side values. Suppose that the righthand-side value b_k is changed by an amount Δb_k, with all other data held fixed. We will compute the range so that the basis remains unchanged. The values of the decision variables will change, since they are given by $x^B = B^{-1}b$, but the values of the simplex multipliers, given by $y = c^B B^{-1}$, will not. The new values of the basic variables, x^{new} must be nonnegative in order for the basis to remain feasible. Hence,

$$x^{\text{new}} = B^{-1}(b + u^k \Delta b_k) \geq 0, \tag{16}$$

Where u^k is a column vector of all zeros except for a *one* in position k. Nothing that $B^{-1}b$ is just the representation of the righthand side in the current basis, (16) becomes

$$\overline{b} + B^{-1}u^k \Delta b_k \geq 0;$$

and, letting β_{ij} be the elements of the basis inverse matrix B^{-1}, we have:

$$\bar{b}_i + \beta_{ik} \Delta b_k \geq 0 \quad (i = 1, 2, \ldots, m), \tag{17}$$

which implies that:

$$\Delta b_i \geq \frac{-\overline{b}_i}{\beta_{ik}} \qquad \text{for } \beta_{ik} > 0,$$

and

$$\Delta b_i \leq \frac{-\overline{b}_i}{\beta_{ik}} \quad \text{for } \beta_{ik} < 0,$$

and no limit for $\beta_{ik} = 0$.

Finally, since (18) must be satisfied for all basic variables, we can define upper and lower bounds on Δb_i as follows:

$$\operatorname*{Max}_i \left\{ \frac{-\overline{b}_i}{\beta_{ik}} \middle| \beta_{ik} > 0 \right\} \leq \Delta b_k \leq \operatorname*{Min}_i \left\{ \frac{-\overline{b}_i}{\beta_{ik}} \middle| \beta_{ik} < 0 \right\}. \tag{19}$$

Note that since $\overline{b}_i \geq 0$, the lower bound on Δb_k is nonpositive and the upper bound is nonnegative. The range on the righthand-side value $b_k + \Delta b_k$ is then determined by adding b_k to each bound in (19).

At the upper bound in (19) the variable basic in the row producing the minimum ratio is a candidate to be dropped from the basis, while at the lower bound in (19) the variable basic in the row producing the maximum ratio is a candidate to be dropped from the basis. The variable to enter the basis in each of these cases can be determined by the ratio test of the dual simplex method. Suppose the variable basic in row r is to be dropped; then the entering variable x_s is determined from:

$$\frac{\overline{c}_s}{\overline{a}_{rs}} = \operatorname*{Min}_j \left\{ \frac{\overline{c}_j}{\overline{a}_{rj}} \middle| \overline{a}_{rj} < 0 \right\}. \tag{20}$$

If there does not exist $\overline{a}_{rj} < 0$, then no entering variable can be determined. When this is the case, the problem is infeasible beyond this bound.

Now let us turn to variations in the coefficients in the equations of the model. In the case where the coefficient corresponds to a nonbasic activity, the situation is straightforward. Suppose that the coefficient a_{ij} is changed by an amount Δa_{ij}. We will compute the range so that the basis remains unchanged. In this case, both the values of the decision variables and the shadow prices also remain unchanged. Since x_j is assumed nonbasic, the current basis remains optimal so long as the new reduced cost $\overline{c}_j^{\text{new}}$ satisfies

$$\overline{c}_j^{\text{new}} = c_j - y\left(A_j + u^i\,\Delta a_{ij}\right) \leq 0, \tag{21}$$

where u^i is a column vector of zeros except for a *one* in position i. Since $\overline{c}_j = c_j - yA_j$, (21) reduces to:

$$\overline{c}_j^{\text{new}} = \overline{c}_j - y_i\,\Delta a_{ij} \leq 0. \tag{22}$$

Hence, (22) gives either an upper or a lower bound on Δa_{ij}. If $y_i > 0$, the appropriate range is:

$$\frac{\overline{c}_j}{y_i} \leq \Delta a_{ij} < +\infty, \tag{23}$$

and if $y_i < 0$, the range is:

$$-\infty < \Delta a_{ij} \leq \frac{\overline{c}_j}{y_i}. \tag{24}$$

The range on the variable coefficient $a_{ij} + \Delta a_{ij}$ is simply given by adding a_{ij} to the bounds in (23) and (24). In either situation, some x_s becomes a candidate to enter the basis, and the corresponding variable to drop from the basis is determined by the usual minimum-ratio rule given in (15).

The case where the coefficient to be varied corresponds to a basic variable is a great deal more difficult and will not be treated in detail here. Up until now, all variations in coefficients and righthand-side values have been such that the basis remains unchanged. The question we are asking here violates this principle. We could perform a similar analysis, assuming that the basic variables should remain unchanged, but the basis and its inverse will necessarily change. There are three possible outcomes from varying a coefficient of a basic variable in a constraint. Either (1) the basis may become singular; (2) the basic solution may become infeasible; or (3) the basic solution may become nonoptimal. Any one of these conditions would define an effective bound on the range of Δa_{ij}. A general derivation of these results is beyond the scope of this discussion.

B.6 PARAMETRIC PROGRAMMING

Having discussed changes in individual elements of the data such that the basis remains unchanged, the natural question to ask is what happens when we make simultaneous variations in the data or variations that go beyond the ranges derived in the previous section. We can give rigorous answers to these questions for cases where the problem is made a function of one parameter. Here we essentially compute ranges on this parameter in a manner analogous to computing righthand-side and objective-function ranges.

We begin by defining three different parametric-programming problems, where each examines the optimal value of a linear program as a function of the scalar (*not* vector) parameter θ. In Chapter 3, we gave examples and interpreted the first two problems.

Parametric righthand side

$$P(\theta) = \text{Max } cx,$$

subject to:

$$Ax = b^1 + \theta b^2, \tag{25}$$
$$x \geq 0.$$

Parametric objective function

$$Q(\theta) = \text{Max } (c^1 + \theta c^2)x,$$

subject to:

$$Ax = b, \tag{26}$$
$$x \geq 0.$$

Parametric rim problem

$$R(\theta) = \text{Max } (c^1 + \theta c^2)x,$$

subject to:

$$Ax = b^1 + \theta b^2, \tag{27}$$
$$x \geq 0.$$

Note that, when the parameter θ is fixed at some value, each type of problem becomes a simple linear program.

We first consider the parametric righthand-side problem. In this case, the feasible region is being modified as the parameter θ is varied. Suppose that, for $\underline{\theta}$ and $\overline{\theta}$, (25) is a feasible linear program. Then, assuming that $\underline{\theta} < \overline{\theta}$, (25) must be feasible for all θ in the interval $\underline{\theta} \leq \theta \leq \overline{\theta}$. To see this, first note that any θ in the interval may be written as $\theta = \lambda\underline{\theta} + (1 - \lambda)\overline{\theta}$, where $0 \leq \lambda \leq 1$. (This is called a convex combination of $\underline{\theta}$ and $\overline{\theta}$.) Since (25) is feasible for $\underline{\theta}$ and $\overline{\theta}$, there must exist corresponding $\underline{x}$ and $\overline{x}$ satisfying:

$$A\underline{x} = b^1 + \underline{\theta}b^2, \qquad \underline{x} \geq 0. \qquad\qquad A\overline{x} = b^1 + \overline{\theta}b^2, \qquad \overline{x} \geq 0.$$

Multiplying the former by λ and the latter by $(1 - \lambda)$ and adding yields:

$$\lambda A\underline{x} + (1 - \lambda)A\overline{x} = \lambda(b^1 + \underline{\theta}b^2) + (1 - \lambda)(b^1 + \overline{\theta}b^2), \quad \lambda\underline{x} + (1 - \lambda)\overline{x} \geq 0,$$

which may be rewritten as:

$$A(\lambda\underline{x} + (1 - \lambda)\overline{x}) = b^1 + (\lambda\underline{\theta} + (1 - \lambda)\overline{\theta})b^2, \qquad \lambda\underline{x} + (1 - \lambda)\overline{x} \geq 0. \tag{28}$$

Equation (28) implies that there exists a feasible solution for any θ in the interval $\underline{\theta} \leq \theta \leq \overline{\theta}$.

The implication for the parameteric righthand-side problem is that, when increasing (decreasing) θ, once the linear program becomes infeasible it will remain infeasible for any further increases (decreases) in θ.

Let us assume that (25) is feasible for $\theta = \theta_0$ and examine the implications of varying θ. For $\theta = \theta_0$, let B be the optimal basis with decision variables $x^B = B^{-1}\left(b^1 + \theta_0 b^2\right)$ and shadow prices $y = c^B B^{-1}$. The current basis B remains optimal as θ is varied, so long as the current solution remains feasible; that is,

$$\overline{b}(\theta) = B^{-1}(b^1 + \theta b^2) \geq 0,$$

or, equivalently,

$$\overline{b}^1 + \theta\overline{b}^2 \geq 0. \tag{29}$$

Equation (29) may imply both upper and lower bounds, as follows:

$$\theta \geq \frac{-\overline{b}_i^1}{\overline{b}_i^2} \qquad \text{for } \overline{b}_i^2 > 0,$$

$$\theta \leq \frac{-\overline{b}_i^1}{\overline{b}_i^2} \qquad \text{for } \overline{b}_i^2 < 0;$$

and these define the following range on θ:

$$\operatorname*{Max}_i \left\{ \frac{-\overline{b}_i^1}{\overline{b}_i^2} \,\middle|\, \overline{b}_i^2 > 0 \right\} \leq \theta \leq \operatorname*{Min}_i \left\{ \frac{-\overline{b}_i^1}{\overline{b}_i^2} \,\middle|\, \overline{b}_i^2 < 0 \right\}. \tag{30}$$

If we now move θ to either its upper or lower bound, a basis change can take place. At the upper bound, the variable basic in the row producing the minimum ratio becomes a candidate to drop from the basis, while at the lower bound the variable producing the maximum ratio becomes a candidate to drop from the basis. In either case, assuming the variable to drop is basic in row r, the variable to enter the basis is determined by the usual rule of the dual simplex method:

$$\frac{\overline{c}_s}{\overline{a}_{rs}} = \operatorname*{Min}_j \left\{ \frac{\overline{c}_j}{\overline{a}_{rj}} \,\middle|\, \overline{a}_{rj} < 0 \right\}. \tag{31}$$

If $\overline{a}_{rj} \geq 0$ for all j, then no variable can enter the basis and the problem is infeasible beyond this bound. If an entering variable is determined, then a new basis is determined and the process is repeated. The range on θ such that the *new* basis remains optimal, is then computed in the same manner.

On any of these successive intervals where the basis is unchanged, the optimal value of the linear program is given by $P(\theta) = y(b^1 + \theta b^2)$, where the vector $y = c^B B^{-1}$ of shadow prices is not a function of θ. Therefore, $P(\theta)$ is a straight line with slope yb^2 on a particular interval.

Further, we can easily argue that $P(\theta)$ is also a concave function of θ, that is,

$$P(\lambda\underline{\theta} + (1-\lambda)\overline{\theta}) \geq \lambda P(\underline{\theta}) + (1-\lambda)P(\overline{\theta}) \qquad \text{for } 0 \leq \lambda \leq 1.$$

Suppose that we let $\underline{\theta}$ and $\overline{\theta}$ be two values of θ such that the corresponding linear programs defined by $P(\underline{\theta})$ and $P(\overline{\theta})$ in (25) have finite optimal solutions. Let their respective optimal solutions be $\underline{x}$ and $\overline{x}$. We have already shown in (28) that $\lambda\underline{x}+(1-\lambda)\overline{x}$ is a feasible solution to the linear program defined by $P(\lambda\underline{\theta}+(1-\lambda)\overline{\theta})$ in (25). However, $\lambda\underline{x} + (1-\lambda)\overline{x}$ may not be optimal to this linear program, and hence

$$P(\lambda\underline{\theta} + (1-\lambda)\overline{\theta}) \geq c(\lambda\underline{x} + (1-\lambda)\overline{x}).$$

Rearranging terms and noting that $\underline{x}$ and $\overline{x}$ are optimal solutions to $P(\underline{\theta})$ and $P(\overline{\theta})$, respectively, we have:

$$c(\lambda\underline{x} + (1-\lambda)\overline{x}) = \lambda c\underline{x} + (1-\lambda)c\overline{x} = \lambda P(\underline{\theta}) + (1-\lambda)P(\overline{\theta}).$$

The last two expressions imply the condition that $P(\theta)$ is a concave function of θ. Hence, we have shown that $P(\theta)$ is a concave piecewise-linear function of θ. This result can be generalized to show that the optimal value of a linear program is a concave polyhedral function of its righthand-side vector.

Let us now turn to the *parametric objective-function* problem. Assuming that the linear program defined in (26) is feasible, it will remain feasible regardless of the value of θ. However, this linear program may become unbounded. Rather than derive the analogous properties of $Q(\theta)$ directly, we can determine the properties of $Q(\theta)$ from those of $P(\theta)$ by utilizing the duality theory of Chapter 4. We may rewrite $Q(\theta)$ in terms of the dual of its linear program as follows:

$$\begin{aligned} Q(\theta) = \ &\text{Max}(c^1 + \theta c^2)x & = \quad & \text{Min } yb, \\ &\text{subject to:} & & \text{subject to:} \\ &\qquad Ax = b, & & \quad yA \geq c^1 + \theta c^2, \\ &\qquad x \geq 0 & & \end{aligned} \tag{32}$$

and, recognizing that a minimization problem can be transformed into a maximization problem by multiplying by minus one, we have

$$\begin{aligned} Q(\theta) = -\text{Max} - yb \qquad & = -P'(\theta), \\ yA \geq c^1 + \theta c^2. & \end{aligned} \tag{33}$$

Here $P'(\theta)$ must be a concave piecewise-linear function of θ, since it is the optimal value of a linear program considered as a function of its righthand side. Therefore,

$$Q(\lambda\underline{\theta} + (1-\lambda)\overline{\theta}) \leq \lambda Q(\underline{\theta}) + (1-\lambda) Q(\overline{\theta}) \qquad \text{for all } 0 \leq \lambda \leq 1,$$

or

$$P'(\lambda\underline{\theta} + (1-\lambda)\overline{\theta}) \geq \lambda P'(\underline{\theta}) + (1-\lambda) P'(\overline{\theta}) \qquad \text{for all } 0 \leq \lambda \leq 1,$$

which says that $Q(\theta)$ is a convex function.

Further, since the primal formulation of $Q(\theta)$ is assumed to be feasible, whenever the dual formulation is infeasible the primal must be unbounded. Hence, we have a result analogous to the feasibility result for $P(\theta)$. Suppose that for $\underline{\theta}$ and $\overline{\theta}$, (26) is a bounded linear program; then, assuming $\underline{\theta} < \overline{\theta}$, (26) must be bounded for all θ in the interval $\underline{\theta} \leq \theta \leq \overline{\theta}$. The implication for the parametric objective-function problem is that, when increasing (decreasing) θ, once the linear program becomes unbounded it will remain unbounded for any further increase (decrease) in θ.

Let us assume that (26) is bounded for $\theta = \theta_0$ and examine the implications of varying θ. For $\theta = \theta_0$, let B be the optimal basis with decision variables $x^B = B^{-1}b$ and shadow prices $y = (c^{1B} + \theta c^{2B})B^{-1}$. The current basis B remains optimal as θ is varied, so long as the reduced costs remain nonpositive; that is,

$$\overline{c}_j(\theta) = (c_j + \theta c_j^2) - yA_j \leq 0.$$

Substituting for the shadow prices y,

$$\overline{c}_j(\theta) = (c_j^1 + \theta c_j^2) - (c^{1B} + \theta c^{2B})B^{-1}A_j \leq 0;$$

and collecting terms yields,

$$\overline{c}_j(\theta) = c_j^1 - c^{1B}B^{-1}A_j + \theta\left(c_j^2 - c^{2B}B^{-1}A_j\right) \leq 0,$$

or, equivalently,

$$\overline{c}_j(\theta) = \overline{c}_j^1 + \theta\overline{c}_j^2 \leq 0. \tag{34}$$

Equation (34) may imply both upper and lower bounds, as follows:

$$\theta \geq \frac{-\overline{c}_j^1}{\overline{c}_j^2} \qquad \text{for } \overline{c}_j^2 < 0,$$

$$\theta \leq \frac{-\overline{c}_j^1}{\overline{c}_j^2} \qquad \text{for } \overline{c}_j^2 > 0;$$

and these define the following range on θ:

$$\operatorname*{Max}_j \left\{ \frac{-\overline{c}_j^1}{\overline{c}_j^2} \middle| \overline{c}_j^2 < 0 \right\} \leq \theta \leq \operatorname*{Min}_j \left\{ \frac{-\overline{c}_j^1}{\overline{c}_j^2} \middle| \overline{c}_j^2 > 0 \right\}. \tag{35}$$

If we move θ to either its upper or lower bound, a basis change can take place. At the upper bound, the variable producing the minimum ratio becomes a candidate to enter the basis, while at the lower bound, the variable producing the maximum ratio becomes a candidate to enter the basis. In either case, the variable to drop from the basis is determined by the usual rule of the primal simplex method.

$$\frac{\overline{b}_r}{\overline{a}_{rs}} = \operatorname*{Min}_i \left\{ \frac{\overline{b}_i}{\overline{a}_{is}} \middle| \overline{a}_{is} > 0 \right\}. \tag{36}$$

If $\overline{A}_s \leq 0$, then x_s may be increased without limit and the problem is unbounded beyond this bound. If a variable to drop is determined, then a new basis is determined and the process is repeated. The range on θ such that the *new* basis remains optimal is then again computed in the same manner.

On any of these successive intervals where the basis is unchanged, the optimal value of the linear program is given by:

$$Q(\theta) = \left(c^{1B} + \theta c^{2B}\right) x^B,$$

where $x^B = B^{-1}b$ and is not a function of θ. Therefore, $Q(\theta)$ is a straight line with slope $c^{2B}x^B$ on a particular interval.

Finally, let us consider the *parametric rim problem*, which has the parameter in both the objective function and the righthand side. As would be expected, the optimal value of the linear program defined in (27) is neither a concave nor a convex function of the parameter θ. Further, $R(\theta)$ is not a piecewise-linear function, either. Let us assume that (27) is feasible for $\theta = \theta_0$, and examine the implications of varying θ. For $\theta = \theta_0$, let B be the optimal basis with decision variables $x^B = B^{-1}(b^1 + \theta b^2)$ and shadow prices $y = \left(c^{1B} + \theta c^{2B}\right) B^{-1}$. The current basis remains optimal so long as the current solution remains feasible and the reduced costs remain nonpositive. Hence, the current basis remains optimal so long as both ranges on θ given by (30) and (35) are satisfied. Suppose we are increasing θ. If the upper bound of (30) is reached before the upper bound of (34), then a dual simplex step is performed according to (31). If the opposite is true, then a primal simplex step is performed according to (36). Once the new basis is determined, the process is repeated. The same procedure is used when decreasing θ.

For a given basis, the optimal value of the objective function is given by multiplying the basic costs by the value of the basic variables x^B; that is,

$$R(\theta) = \left(c^{1B} + \theta c^{2B}\right) B^{-1}(b^1 + \theta b^2),$$

which can be rewritten as

$$R(\theta) = c^{1B}B^{-1}b^1 + \theta\left(c^{2B}B^{-1}b^1 + c^{1B}B^{-1}b^2\right) + \theta^2\left(c^{2B}B^{-1}b^2\right).$$

Hence, the optimal value of the parametric rim problem is a quadratic function of θ for a fixed basis B. In general, this quadratic may be either a concave or convex function over the range implied by (30) and (35).

It should be clear that the importance of parametric programming in all three cases is the efficiency of the procedures in solving a number of different cases. Once an optimal solution has been found for some value of the parameter, say $\theta = \theta_0$, increasing or decreasing θ amounts to successively computing the points at which the basis changes and then performing a routine pivot operation.

B.7 DUALITY THEORY IN MATRIX FORM

The duality theory introduced in Chapter 4 can be stated very concisely in matrix form. As we saw there, a number of variations of the basic duality results can be formulated by slightly altering the form of the primal problem. For ease of comparison with the Chapter 4 results, we will again employ the symmetric version of the dual linear programs:

Primal	*Dual*
Max $z = cx,$	Min $v = yb,$
subject to:	subject to:
$Ax \leq b,$	$yA \geq c,$
$x \geq 0.$	$y \geq 0.$

Note that y and c are row vectors while x and b are column vectors. Let $\overline{z}$ and $\overline{v}$ denote the optimal values of the objective functions for the primal and dual problems, respectively. We will review the three key results of duality theory.

Weak duality. If $\overline{x}$ is a feasible solution to the primal and $\overline{y}$ is a feasible solution to the dual, then $c\overline{x} \leq \overline{y}b$, and consequently, $\overline{z} \leq \overline{v}$.

Strong duality. If the primal (dual) has a finite optimal solution, then so does the dual (primal), and the two extremes are equal, $\overline{z} = \overline{v}$.

Complementary slackness. If $\overline{x}$ is a feasible solution to the primal and $\overline{y}$ is a feasible solution to the dual, then $\overline{x}$ and $\overline{y}$ are optimal solutions to the primal and dual, respectively, if and only if

$$\overline{y}(A\overline{x} - b) = 0 \qquad \text{and} \qquad (\overline{y}A - c)\overline{x} = 0.$$

The arguments leading to these results were given in Chapter 4 but will be briefly reviewed here. *Weak duality* is a consequence of primal and dual feasibility, since multiplying the primal constraints on the left by $\overline{y}$ and the dual constraints on the right by $\overline{x}$ and combining gives:

$$c\overline{x} \leq \overline{y}A\overline{x} \leq \overline{y}b.$$

Since the optimal values must be at least as good as any feasible value,

$$\overline{z} = \operatorname*{Max}_{x} c\overline{x} \leq \operatorname*{Min}_{y} \overline{y}b = \overline{v}.$$

Weak duality implies that if $c\overline{x} = \overline{y}b$ for some primal feasible $\overline{x}$ and dual feasible $\overline{y}$, then $\overline{x}$ and $\overline{y}$ are optimal solutions to the primal and dual, respectively.

The termination conditions of the simplex method and the concept of simplex multipliers provides the *strong duality property*. Suppose that the simplex method has found an optimal basic feasible solution $\overline{x}$ to the primal with optimal basis B. The optimal values of the decision variables and simplex multipliers are $\overline{x}^B = B^{-1}b$ and $\overline{y} = c^B B^{-1}$, respectively. The simplex optimality conditions imply:

$$\overline{z} = c^B x^B = c^B B^{-1} b = \overline{y}b, \tag{37}$$
$$\overline{c}^N = c^N - \overline{y}N \leq 0. \tag{38}$$

The latter condition (38), plus the definition of the simplex multipliers, which can be restated as:

$$\overline{c}^B = c^B - \overline{y}B = 0,$$

imply that $c - \overline{y}A \leq 0$, so that $\overline{y} \geq 0$ is a dual feasible solution. Now $\overline{x}^N = 0$ and (37) implies that:

$$c\overline{x} = c^B\overline{x}^B + c^N\overline{x}^N = \overline{y}b. \tag{39}$$

Since $\overline{x}$ and $\overline{y}$ are feasible solutions to the primal and dual respectively, (39) implies $\overline{z} = c\overline{x} = \overline{y}b = \overline{v}$ which is the desired result.

Finally, *complementary slackness* follows directly from the strong-duality property. First, we assume that complementary slackness holds, and show optimality.

If $\overline{x}$ and $\overline{y}$ are feasible to the primal and dual, respectively, such that

$$\overline{y}(A\overline{x} - b) = 0 \quad \text{and} \qquad (\overline{y}A - c)\overline{x} = 0,$$

then

$$c\overline{x} = \overline{y}A\overline{x} = \overline{y}b. \tag{40}$$

Condition (40) implies the $\overline{x}$ and $\overline{y}$ are optimal to the primal and dual, respectively. Second, we assume optimality holds, and show complementary slackness. Let $\overline{x}$ and $\overline{y}$ be optimal to the primal and dual, respectively; that is,

$$c\overline{x} = \overline{y}b.$$

Then

$$0 = c\overline{x} - \overline{y}b \leq c\overline{x} - \overline{y}A\overline{x} = (c - \overline{y}A)\overline{x} \tag{41}$$

by primal feasibility, and $\overline{y} \geq 0$. Now since $0 \geq c - \overline{y}A$ by dual feasibility, and $\overline{x} \geq 0$, we have

$$0 \geq (c - \overline{y}A)\overline{x}, \tag{42}$$

and (41) and (42) together imply:

$$(c - \overline{y}A)\overline{x} = 0.$$

An analogous argument using dual feasibility and $x \geq 0$ implies the other complementary slackness condition,

$$\overline{y}(A\overline{x} - b) = 0.$$

These results are important for applications since they suggest algorithms other than straightforward applications of the simplex method to solve specially structured problems. The results are also central to the theory of systems of linear equalities and inequalities, a theory which itself has a number of important applications. In essence, duality theory provides a common perspective for treating such systems and, in the process, unifies a number of results that appear scattered throughout the mathematics literature. A thorough discussion of this theory would be inappropriate here, but we indicate the flavor of this point of view.

A problem that has interested mathematicians since the late 1800's has been characterizing the situation when a system of linear equalities and inequalities does not have a solution. Consider for example the system:

$$Ax = b, \qquad x \geq 0.$$

Suppose that y is a row vector. We multiply the system of equations by y to produce a single new equation:

$$(yA)x = yb.$$

If x is feasible in the original system, then it will certainly satisfy this new equation. Suppose, though, that $yb > 0$ and that $yA \leq 0$. Since $x \geq 0$, $yAx \leq 0$, so that no $x \geq 0$ solves the new equation, and thus the original system is infeasible. We have determined a single inconsistent equation that summarizes the inconsistencies in the original system. The characterization that we are working towards states that such a summarizing inconsistent equation always can be found when the system has no solution. It is given by:

Farkas' Lemma. Exactly one of the following two systems has a solution:

$$\begin{array}{llcll} \text{I} & Ax = b, & & \text{II} & yA \leq 0, \\ & & \text{or} & & \\ & x \geq 0. & & & yb > 0. \end{array}$$

The proof of the lemma is straightforward, by considering the Phase I linear-programming program that results from adding artificial variables to System (I):

Minimize et,

subject to:

$$\begin{aligned} Ax + It &= b, \\ x \geq 0, \qquad t &\geq 0, \end{aligned} \tag{43}$$

where e is the sum vector consisting of all ones and t is the vector of artificial variables. Suppose that the simplex method is applied to (43), and the optimal solution is $(\overline{x}, \overline{t})$ with shadow prices $\overline{y}$. By the termination conditions of the simplex method, the shadow prices satisfy:

$$0 - \overline{y}A \geq 0 \qquad (\text{that is, } \overline{y}A \leq 0). \tag{44}$$

By the duality property, we know $e\overline{t} = \overline{y}b$, and since (43) is a Phase I linear program, we have:

$$\begin{aligned} \overline{y}b = e\overline{t} = 0 \qquad & \text{if and only if } I \text{ is feasible,} \\ \overline{y}b = e\overline{t} > 0 \qquad & \text{if and only if } I \text{ is infeasible.} \end{aligned} \tag{45}$$

Equations (44) and (45) together imply the lemma.

We can give a geometrical interpretation of Farkas' Lemma by recalling that, if the inner product of two vectors is positive, the vectors make an acute angle with each other, while if the inner product is negative, the vectors make an obtuse angle with each other. Therefore, any vector y that solves System II must make acute angles with all the column vectors of A and a strictly obtuse angle with the vector b. On the other hand, in order for System I to have a feasible solution, there must exist nonnegative weights x that generate b from the column vectors of A. These two situations are given in Figs. B.4 and B.5.

B.8 RESOLVING DEGENERACY IN THE SIMPLEX METHOD

In Chapter 2 we showed that the simplex method solves any linear program in a finite number of steps if we assume that the righthand-side vector is strictly positive for each canonical form generated. Such a canonical form is called *nondegenerate*.

The motivation for this assumption is to ensure that the new value of the entering variable x_s^*, which is given by:

$$x_s^* = \frac{\overline{b}_r}{\overline{a}_{rs}}$$

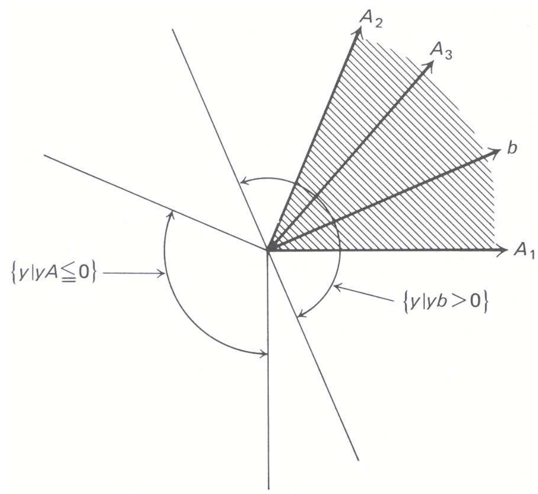

Fig. B.4 System I has a solution.

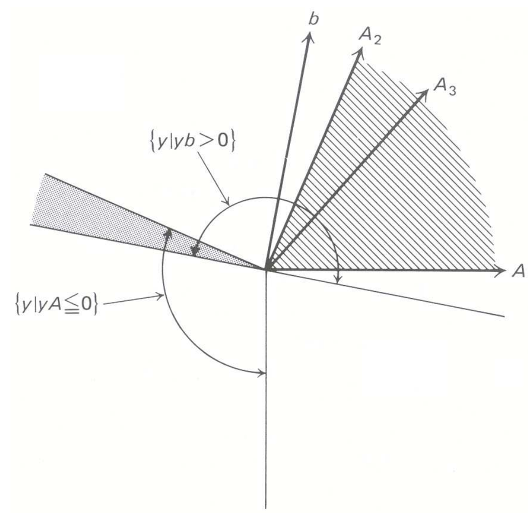

Fig. B.5 System II has a solution.

is strictly positive; and hence, that the new value of the objective function z^*, given by:

$$z^* = \overline{z} + \overline{c}_s x_s^*,$$

shows a strict improvement at each iteration. Since the minimum-ratio rule to determine the variable to drop from the basis ensures $\overline{a}_{rs} > 0$, the assumption that $\overline{b}_i > 0$ for all i implies that $x_s^* > 0$. Further, introducing the variable x_s requires that $\overline{c}_s > 0$ and therefore that $z^* > \overline{z}$. This implies that there is a strict improvement in the value of the objective function at each iteration, and hence that no basis is repeated. Since there is a finite number of possible bases, the simplex method must terminate in a finite number of iterations.

The purpose of this section is to extend the simplex method so that, without the nondegeneracy assumption, it can be shown that the method will solve any linear program in a finite number of iterations. In Chapter 2 we indicated that we would do this by perturbing the righthand-side vector. However, a simple perturbation of the righthand side by a scalar will not ensure that the righthand side will be positive for all canonical forms generated. As a result, we introduce a vector perturbation of the righthand side and the concept of lexicographic ordering.

An m-element vector $a = (a_1, a_2, \ldots, a_m)$ is said to be *lexico-positive*, written $a \succ 0$, if at least one element is nonzero and the first such element is positive. The term lexico-positive is short for lexicographically positive. Clearly, any positive multiple of a lexico-positive vector is lexico-positive, and the sum of two lexico-positive vectors is lexico-positive. An m-element vector a is *lexico-greater* than an m-element vector b, written $a \succ b$, if $(a - b)$ is lexico-positive; that is, if $a - b \succ 0$. Unless two vectors are identical, they are lexicographically ordered, and therefore the ideas of *lexico-max* and *lexico-min* are well-defined.

We will use the concept of lexicographic ordering to modify the simplex method so as to produce a unique variable to drop at each iteration and a strict lexicographic improvement in the value of the objective function at each iteration. The latter will ensure the termination of the method after a finite number of iterations.

Suppose the linear program is in canonical form initially with $b \geq 0$. We introduce a unique perturbation of the righthand-side vector by replacing the vector b with the $m \times (m+1)$ matrix $[b, I]$. By making the vector x an $n \times (m+1)$ matrix X, we can write the initial tableau as:

$$
\begin{array}{llll}
X^B \geq 0, & X^N \geq 0, & X^I \geq 0, & \\
BX^B + NX^N & + IX^I & & = [b, I], \\
c^B X^B + c^N X^N & & -Z & = [0, 0],
\end{array}
$$

where X^B, X^N, and X^I are the matrices associated with the basic variables, the non-basic variables, and the variables in the initial identity basis, respectively. The intermediate tableau corresponding to the basis B is determined as before by multiplying the constraints of (46) by B^{-1} and subtracting c^B times the resulting constraints from the objective function. The intermediate tableau is then:

$$
\begin{array}{llll}
X^B \geq 0, & X^N \geq 0, & X^I \geq 0, & \\
IX^B + \overline{N}X^N + B^{-1}X^I & & & = [\overline{b}, B^{-1}] = \overline{B}, \\
\qquad \overline{c}^N X^N - & yX^I - Z & & = [-\overline{z}, -y] = -\overline{Z},
\end{array}
\tag{47}
$$

where $\overline{N} = B^{-1}N$, $y = c^B B^{-1}$, $\overline{c}^N = c^N - yN$, and $\overline{z} = yb$ as before. Note that if the matrices X^N and X^I are set equal to zero, a basic solution of the linear program is given by $X^B = [\overline{b}, B^{-1}]$, where the first column of X^B, say X_1^B, gives the usual values of the basic variables $X_1^B = \overline{b}$.

We will formally prove that the simplex method, with a lexicographic resolution of degeneracy, solves any linear program in a finite number of iterations. An essential aspect of the proof is that at each iteration of the algorithm, each row vector of the righthand-side perturbation matrix is lexico-positive.

In the initial tableau, each row vector of the righthand-side matrix is lexico-positive, assuming $b \geq 0$. If $b_i > 0$, then the row vector is lexico-positive, since its first element is positive. If $b_i = 0$, then again the row vector is lexico-positive since the first nonzero element is a plus *one* in position $(i + 1)$.

Define $\overline{B}_i$ to be row i of the righthand-side matrix $[\overline{b}, B^{-1}]$ of the intermediate tableau. In the following proof, we will assume that each such row vector $\overline{B}_i$ is lexico-positive for a particular canonical form, and argue inductively that they remain lexico-positive for the next canonical form. Since this condition holds for the initial tableau, it will hold for all subsequent tableaus.

Now consider the simplex method one step at a time.

STEP (1): If $\overline{c}_j \leq 0$ for $j = 1, 2, \ldots, n$, then *stop*; the current solution is optimal.

Since $X^B = \overline{B} = [\overline{b}, I]$, the first component of X^B is the usual basic feasible solution $X_1^B = \overline{b}$. Since the reduced costs associated with this solution are $\overline{c}^N \leq 0$, we have:

$$z = \overline{z} + \overline{c}^N X_1^N \leq \overline{z},$$

and hence $\overline{z}$ is an upper bound on the maximum value of z. The current solution $X_1^B = \overline{b}$ attains this upper bound and is therefore optimal. If we continue the algorithm, there exists some $\overline{c}_j > 0$.

STEP (2): Choose the column s to enter the basis by:

$$\overline{c}_s = \operatorname*{Max}_j \left\{ \overline{c}_j \middle| \overline{c}_j > 0 \right\}.$$

If $\overline{A}_s \leq 0$, then *stop*; there exists an unbounded solution. Since $\overline{A}_s \leq 0$, then

$$X_1^B = \overline{b} - \overline{A}_s x_s \geq 0 \qquad \text{for all } x_s \geq 0,$$

which says that the solution X_1^B is feasible for all nonnegative x_s. Since $\overline{c}_s > 0$,

$$z = \overline{z} + \overline{c}_s x_s$$

implies that the objective function becomes unbounded as x_s is increased without limit. If we continue the algorithm, there exists some $\overline{a}_{is} > 0$.

STEP (3): Choose the row to pivot in by the following modified ratio rule:

$$\frac{\overline{B}_r}{\overline{a}_{rs}} = \text{lexico-}\min_i \left\{ \frac{\overline{B}_i}{\overline{a}_{is}} \middle| \overline{a}_{is} > 0 \right\}.$$

We should point out that the lexico-min produces a *unique* variable to drop from the basis. There must be a unique lexico-min, since, if not, there would exist two vectors $\overline{B}_i/\overline{a}_{is}$ that are identical. This would imply that two rows of $\overline{B}$ are proportional and, hence, that two rows of B^{-1} are proportional, which is a clear contradiction.

STEP (4): Replace the variable basic in row r with variable s and re-establish the canonical form by pivoting on the coefficient $\overline{a}_{rs}$.

We have shown that, in the initial canonical form, the row vectors of the righthand-side matrix are each lexico-positive. It remains to show that, assuming the vectors $\overline{B}_i$ are lexico-positive at an iteration, they remain lexico-positive at the next iteration. Let the new row vectors of the righthand-side matrix be $\overline{B}_i^*$. First,

$$\overline{B}_r^* = \frac{\overline{B}_r}{\overline{a}_{rs}} \succ 0,$$

since $\overline{B}_r \succ 0$, by assumption, and $\overline{a}_{rs} > 0$. Second,

$$\overline{B}_i^* = \overline{B}_i - \overline{a}_{is}\frac{\overline{B}_r}{\overline{a}_{rs}}. \tag{48}$$

If $\overline{a}_{is} \leq 0$, then $\overline{B}_i^* \succ 0$, since it is then the sum of two lexico-positive vectors. If $\overline{a}_{is} > 0$, then (48) can be rewritten as:

$$\overline{B}_i^* = \overline{a}_{is}\left[\frac{\overline{B}_i}{\overline{a}_{is}} - \frac{\overline{B}_r}{\overline{a}_{rs}}\right]. \tag{49}$$

Since $\overline{B}_r/\overline{a}_{rs}$ is the unique lexico-min in the modified ratio rule, (49) implies $\overline{B}_i^* \succ 0$ for $\overline{a}_{is} > 0$, also.

STEP (5): Go to STEP (1)

We would like to show that there is a strict lexicographic improvement in the objective-value vector at each iteration. Letting $-\overline{Z} = [-\overline{z}, -y]$ be the objective-value vector, and $\overline{Z}^*$ be the new objective-value vector, we have:

$$\overline{Z}^* = \overline{Z} + \overline{c}_s X_s^*;$$

then, since $\overline{c}_s > 0$ and $X_s^* = \overline{B}_r^* \succ 0$, we have

$$\overline{Z}^* = \overline{Z} + \overline{c}_s X_s^* \succ \overline{Z},$$

which states that $\overline{Z}^*$ is lexico-greater than $\overline{Z}$. Hence we have a strict lexicographic improvement in the objective-value vector at each iteration.

Finally, since there is a strict lexicographic improvement in the objective-value vector for each new basis, no basis can then be repeated. Since there are a finite number of bases and no basis is repeated, the algorithm solves any linear program in a finite number of iterations.