2 Solving LPs: The Simplex Algorithm of George Dantzig

2.1 Simplex Pivoting: Dictionary Format

We illustrate a general solution procedure, called the *simplex algorithm*, by implementing it on a very simple example. Consider the LP

$$
\begin{array}{rrcl}
\max & 5x_1 + 4x_2 + 3x_3 & & \\
\text{s.t.} & 2x_1 + 3x_2 + x_3 & \leq & 5 \\
& 4x_1 + x_2 + 2x_3 & \leq & 11 \\
& 3x_1 + 4x_2 + 2x_3 & \leq & 8 \\
& 0 \leq x_1, x_2, x_3 & &
\end{array} \tag{2.1}
$$

In devising our approach we use a standard mathematical approach; reduce the problem to one that we already know how to solve. Since the structure of this problem is essentially linear, we try to reduce it to a problem of solving a system of linear equations, or perhaps a series of such systems. By encoding the problem as a system of linear equations we bring into play our knowledge and experience with such systems in the new context of linear programming.

In order to encode the LP (2.1) as a system of linear equations we first transform the linear inequalities into linear equations. This is done by introducing a new non-negative variable, called a *slack variable*, for each inequality:

$$
\begin{array}{rcrcll}
x_4 & = & 5 & - & [2x_1 + 3x_2 + x_3] & \geq 0, \\
x_5 & = & 11 & - & [4x_1 + x_2 + 2x_3] & \geq 0, \\
x_6 & = & 8 & - & [3x_1 + 4x_2 + 2x_3] & \geq 0.
\end{array}
$$

To handle the objective, we introduce a new variable z:

$$z = 5x_1 + 4x_2 + 3x_3.$$

Then all of the information associated with the LP (2.1) can be coded as follows:

$$
\begin{array}{rcccccccc}
2x_1 + 3x_2 + x_3 & + & x_4 & & & & & = & 5 \\
4x_1 + x_2 + 2x_3 & & & + & x_5 & & & = & 11 \\
3x_2 + 4x_2 + 2x_3 & & & & & + & x_6 & = & 8 \\
-z + 5x_1 + 4x_2 + 3x_3 & & & & & & & = & 0 \\
\multicolumn{9}{c}{0 \leq x_1, x_2, x_3, x_4, x_5, x_6.}
\end{array} \tag{2.2}
$$

The new variables x_4, x_5, and x_6 are called slack variables since they take up the *slack* in the linear inequalities. This system can also be written using block structured matrix notation:

$$
\begin{bmatrix} 0 & A & I \\ -1 & c^T & 0 \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix},
$$

where

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 4 & 2 \end{bmatrix}, \; I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 5 \\ 11 \\ 8 \end{bmatrix}, \text{ and } c = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}.$$

The augmented matrix associated with the system (2.2) is

$$\text{(2.3)} \qquad \left[\begin{array}{ccc|c} 0 & A & I & b \\ -1 & c & 0 & 0 \end{array}\right]$$

and is referred to as the *initial simplex tableau* for the LP (2.1).

Now return to the system

$$\text{(2.4)} \qquad \begin{array}{rcl} x_4 & = & 5 - 2x_1 - 3x_2 - x_3 \\ x_5 & = & 11 - 4x_1 - x_2 - 2x_3 \\ x_6 & = & 8 - 3x_1 - 4x_2 - 2x_3 \\ z & = & 5x_1 + 4x_2 + 3x_3. \end{array}$$

This system defines the variables x_4, x_5, x_6 and z as linear combinations of the variables x_1, x_2, and x_3. We call this system a *dictionary* for the LP (2.1). More specifically, it is the *initial* dictionary for the the LP (2.1). This initial dictionary defines the objective value z and the slack variables as a linear combination of the initial decision variables. The variables that are "defined" in this way are called the *basic variables*, while the remaining variables are called *nonbasic*. The set of all basic variables is called the *basis*. A particular solution to this system is easily obtained by setting the non-basic variables equal to zero. In this case, we have

$$\begin{array}{rcl} x_1 & = & 0 \\ x_2 & = & 0 \\ x_3 & = & 0 \end{array} \quad \text{giving} \quad \begin{array}{rcl} x_4 & = & 5 \\ x_5 & = & 11 \\ x_6 & = & 8 \\ z & = & 0. \end{array}$$

Note that the solution

$$\text{(2.5)} \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 5 \\ 11 \\ 8 \end{pmatrix}$$

is feasible for the extended system (2.2) since all components are non-negative. We call this solution the *basic feasible solution* (BFS) associated with the dictionary (2.4). Moreover, we call the dictionary (2.4) a *feasible dictionary* for the LP (2.1), and we say that this LP has *feasible origin*.

In general, a dictionary for the LP (2.1) is any system of 4 linear equations that defines three of the variables $x_1, \ldots, x_6$ and z in terms of the remaining 3 variables and has the same solution set as the initial dictionary. The variables other than z that are being defined in the dictionary are called the basis for the dictionary, and the remaining variables are said to be non-basic in the dictionary. Every dictionary identifies a particular solution to the linear system obtained by setting the non-basic variables equal to zero. Such a solution is said to be a *basic feasible solution* (BFS) for the LP (2.1) if it componentwise non-negative, that is, all of the numbers in the vector are non-negative so that the point lies in the feasible region for the LP.

The grand strategy of the simplex algorithm is to move from one feasible dictionary representation of the system (2.2) to another (and hence from one BFS to another) while simultaneously increasing the value of the objective variable z at the associated BFS. In the current setting, beginning with the dictionary (2.4), what strategy might one employ in order to determine a new dictionary whose associated BFS gives a greater value for the objective variable z?

Each feasible dictionary is associated with one and only one feasible point. This is the associated BFS obtained by setting all of the non-basic variables equal to zero. This is how we obtain (2.5). To change the feasible point identified in this way, we need to increase the value of one of the non-basic variables from its current value of zero. We cannot decrease the value of a non-basic variable since we wish to remain feasible, that is, we wish to keep all variables non-negative.

Note that the coefficient of each of the non-basic variables in the representation of the objective value z in (2.4) is positive. Hence, if we pick any one of these variables and increase its value from zero while leaving the remaining two at zero, we automatically increase the value of the objective variable z. Since the coefficient on x_1 in the representation of z is the greatest, we can increase z the fastest instantaneous rate by increasing x_1. But choosing the variable with the greatest coefficient is not required and may not yield the greatest increase in z. *Any non-basic variable with a positive coefficient in the representation of z can be used to increase the value of z.*

By how much can we increase x_1 and still remain feasible? For example, if we increase x_1 to 3 then (2.4) says that $x_4 = -1$, $x_5 = -1$, $x_6 = -1$ which is not feasible. So x_1 cannot be increased to 3. To see how much we can increase the value of x_1 we examine the equations in (2.4) one by one. Note that the first equation in the dictionary (2.4),

$$x_4 = 5 - 2x_1 - 3x_2 - x_3,$$

shows that x_4 remains non-negative as long as we do not increase the value of x_1 beyond $5/2$ (remember, x_2 and x_3 remain at the value zero). Similarly, using the second equation in the dictionary (2.4),

$$x_5 = 11 - 4x_1 - x_2 - 2x_3,$$

x_5 remains non-negative if $x_1 \leq 11/4$. Finally, the third equation in (2.4),

$$x_6 = 8 - 3x_1 - 4x_2 - 2x_3,$$

implies that x_6 remains non-negative if $x_1 \leq 8/3$. Therefore, we remain feasible, i.e. keep **all** variables non-negative, if our increase to the variable x_1 remains less than

$$\frac{5}{2} = \min\left\{\frac{5}{2}, \frac{11}{4}, \frac{8}{3}\right\}.$$

If we now increase the value of x_1 to $\frac{5}{2}$, then the value of x_4 is driven to zero. One way to think of this is that x_1 *enters the basis while* x_4 *leaves the basis*. Mechanically, we obtain the new dictionary having x_1 basic and x_4 non-basic by using the defining equation for x_4 in the current dictionary:

$$x_4 = 5 - 2x_1 - 3x_2 - x_3.$$

By moving x_1 to the left hand side of this equation and x_4 to the right, we get the new equation

$$2x_1 = 5 - x_4 - 3x_2 - x_3$$

or equivalently

$$x_1 = \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3.$$

The variable x_1 can now be *eliminated* from the remaining two equations in the dictionary by substituting in this equation for x_1 where it appears in these equations:

$$\begin{aligned}
x_1 &= \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3 \\
x_5 &= 11 - 4\left(\frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3\right) - x_2 - 2x_3 \\
&= 1 + 2x_4 + 5x_2 \\
x_6 &= 8 - 3\left(\frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3\right) - 4x_2 - 2x_3 \\
&= \frac{1}{2} + \frac{3}{2}x_4 + \frac{1}{2}x_2 - \frac{1}{2}x_3 \\
z &= 5\left(\frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3\right) + 4x_2 + 3x_3 \\
&= \frac{25}{2} - \frac{5}{2}x_4 - \frac{7}{2}x_2 + \frac{1}{2}x_3.
\end{aligned}$$

When this substitution is complete, we have the new dictionary and the new BFS:

$$\begin{aligned}
x_1 &= \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3 \\
x_5 &= 1 + 2x_4 + 5x_2 \\
x_6 &= \frac{1}{2} + \frac{3}{2}x_4 + \frac{1}{2}x_2 - \frac{1}{2}x_3 \\
z &= \frac{25}{2} - \frac{5}{2}x_4 - \frac{7}{2}x_2 + \frac{1}{2}x_3,
\end{aligned} \tag{2.6}$$

and the associated BFS is

$$(2.7) \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 5/2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1/2 \end{pmatrix} \qquad \text{with} \qquad z = \frac{25}{2}.$$

This process should seem very familiar to you. It is simply Gaussian elimination. As we know from our knowledge of linear systems of equations, Gaussian elimination can be performed in a matrix context with the aid of the augmented matrix (or, simplex tableau) (2.3). We return to this observation later to obtain a more efficient computational technique.

We now have a new dictionary (2.6) which identifies the basic feasible solution (BFS) (2.7) with associated objective value $z = \frac{25}{2}$. Can we improve on this BFS and obtain a higher objective value? Let's try the same trick again, and repeat the process we followed in going from the initial dictionary (2.4) to the new dictionary (2.6). Note that the coefficient of x_3 in the representation of z in the new dictionary (2.6) is positive. Hence if we increase the value of x_3 from zero, we will increase the value of z. By how much can we increase the value of x_3 and yet keep all the remaining variables non-negative? As before, we see that the first equation in the dictionary (2.6) combined with the need to keep x_1 non-negative implies that we cannot increase x_3 by more than $(5/2)/(1/2) = 5$. However, the second equation in (2.6) places no restriction on increasing x_3 since x_3 does not appear in this equation. Finally, the third equation in (2.6) combined with the need to keep x_6 non-negative implies that we cannot increase x_3 by more than $(1/2)/(1/2) = 1$. Therefore, in order to preserve the non-negativity of all variables, we can increase x_3 by at most

$$1 = \min\{5, 1\}.$$

When we do this x_6 is driven to zero, so x_3 enters the basis and x_6 leaves. More precisely, first move x_3 to the left hand side of the defining equation for x_6 in (2.6),

$$\frac{1}{2}x_3 = \frac{1}{2} + \frac{3}{2}x_4 + \frac{1}{2}x_2 - x_6,$$

or, equivalently,

$$x_3 = 1 + 3x_4 + x_2 - 2x_6,$$

then substitute this expression for x_3 into the remaining equations,

$$\begin{aligned} x_1 &= \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}[1 + 3x_4 + x_2 - 2x_6] \\ &= 2 - 2x_4 - 2x_2 + x_6 \\ x_5 &= 1 + 2x_4 + 5x_2 \\ z &= \frac{25}{2} - \frac{5}{2}x_4 - \frac{7}{2}x_2 + \frac{1}{2}[1 + 3x_4 + x_2 - 2x_6] \\ &= 13 - x_4 - 3x_2 - x_6\ , \end{aligned}$$

yielding the dictionary

$$
\begin{array}{rrl}
 & x_3 = & 1 + 3x_4 + x_2 - 2x_6 \\
(2.8) & x_1 = & 2 - 2x_4 + 2x_2 + x_6 \\
 & x_5 = & 1 \qquad + 2x_4 + 2x_2 \\
 & z = & 13 - x_4 - 3x_2 - x_6
\end{array}
$$

which identifies the feasible solution

$$
(2.9) \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}
$$

having objective value $z = 13$.

Can we do better? If we try the same trick again on the dictionary (2.8) we find that we are stuck since all of the coefficients of the nonbasic variables in the objective row $z = 13 - x_4 - 3x_2 - x_6$, are non-positive. Hence, increasing any one of their values will not increase the value of the objective. So it seems as though either the method has failed or the associated BFS (2.9) is optimal for the LP (2.1). We claim that this BFS (2.9) yields an optimal solution to the LP (2.1). The optimal solution is given by dropping the entries associated with the slacks,

$$
x = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.
$$

However, we delay proving this fact until we have gathered enough tools for this purpose. Nonetheless, we will call such a dictionary an *optimal dictionary* since its associated BFS yields an optimal solution to the LP. A more formal definition ofr an optimal dictionary will be given later.

The process of moving from one feasible dictionary to the next is called a *simplex pivot*. The process of stringing together a sequence of simplex pivots in order to locate an optimal solution is called the *Simplex Algorithm*. The simplex algorithm is considered one of the ten most important algorithmic discoveries of the 20th century

(http://www.uta.edu/faculty/rcli/TopTen/topten.pdf).

The algorithm was discovered by George Dantzig (1914-2005) who is known as the father of linear programming. In 1984 Narendra Karmarkar published a paper describing a new approach to solving linear programs that was both numerically efficient and had *polynomial complexity*. This new class of methods are called *interior point* methods. These new methods have revolutionized the optimization field since their discovery, and they have led to efficient numerical methods for a wide variety of optimization problems well beyond the confines of linear programming. However, the simplex algorithm continues as an important numerical method for solving LPs, and for many specially structured LPs it is remains the most efficient algorithm.

2.2 Simplex Pivoting: Tableau Format (Augmented Matrix Format)

We now review the implementation of the simplex algorithm by applying Gaussian elimination to the augmented matrix (2.3), also known as the simplex tableau. For this problem, the initial simplex tableau is given by

$$\begin{bmatrix} 0 & A & I & b \\ -1 & c & 0 & 0 \end{bmatrix} = \begin{array}{c} \begin{array}{ccccccc} z & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \\ \left[\begin{array}{ccccccc|c} 0 & 2 & 3 & 1 & 1 & 0 & 0 & 5 \\ 0 & 4 & 1 & 2 & 0 & 1 & 0 & 11 \\ 0 & 3 & 4 & 2 & 0 & 0 & 1 & 8 \\ -1 & 5 & 4 & 3 & 0 & 0 & 0 & 0 \end{array}\right] \end{array} . \tag{2.10}$$

Each simplex pivot on a dictionary corresponds to one step of Gaussian elimination on the augmented matrix associated with the dictionary. For example, in the first simplex pivot, x_1 enters the basis and x_4 leaves the basis. That is, we use the first equation of the dictionary to rewrite x_1 as a function of the remaining variables, and then use this representation to eliminate x_1 form the remaining equations. In terms of the augmented matrix (2.10), this corresponds to first making the coefficient for x_1 in the first equation the number 1 by dividing this first equation through by 2. Then use this entry to eliminate the column under x_1, that is, make all other entries in this column zero (Gaussian elimination):

z	x_1 (Pivot column ↓)	x_2	x_3	x_4	x_5	x_6		ratios ↓	
0	(2)	3	1	1	0	0	5	(5/2)	← Pivot row
0	4	1	2	0	1	0	11	11/4	
0	3	4	2	0	0	1	8	8/3	
-1	(5)	4	3	0	0	0	0		
0	1	3/2	1/2	1/2	0	0	5/2		
0	0	−5	0	−2	1	0	1		
0	0	−1/2	1/2	−3/2	0	1	1/2		
−1	0	−7/2	1/2	−5/2	0	0	−25/2		

In this illustration, we have placed a line above the cost row to delineate its special roll in the pivoting process. In addition, we have also added a column on the right hand side which contains the ratios that we computed in order to determine the pivot row. Recall that we must use the smallest ratio in order to keep all variables in the associated BFS non-negative. Note that we performed the exact same arithmetic operations but in the more efficient matrix format. The new augmented matrix,

$$
\begin{array}{c}
\begin{array}{ccccccc} z & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \\
\left[\begin{array}{ccccccc|c}
0 & 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 \\
0 & 0 & -5 & 0 & -2 & 1 & 0 & 1 \\
0 & 0 & -1/2 & 1/2 & -3/2 & 0 & 1 & 1/2 \\
-1 & 0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2
\end{array}\right]
\end{array}, \tag{2.11}
$$

is the augmented matrix for the dictionary (2.6).

The initial augmented matrix (2.10) has basis x_4, x_5, and x_6. The columns associated with these variables in the initial tableau (2.10) are distinct columns of the identity matrix. Correspondingly, the basis for the second tableau is x_1, x_5, and x_6, and again this implies that the columns for these variables in the tableau (2.11) are distinct columns of the identity matrix. In tableau format, this will always be true of the basic variables, i.e., their associated columns are distinct columns of the identity matrix. To recover the BFS (basic feasible solution) associated with this tableau we first set the non-basic variables equal to zero (i.e. the variables not associated with columns of the identity matrix (except in very unusual circumstances)): $x_2 = 0$, $x_3 = 0$, and $x_4 = 0$. To find the value of the basic variables go to the column associated with that variable (for example, x_1 is in the second column), in that column find the row with the number 1 in it, then in that row go to the number to the right of the vertical bar (for x_1 this is the first row with the number to the right of the bar being 5/2). Then set this basic variable equal to that number ($x_1 = 5/2$). Repeating this for x_5 and x_6 we get $x_5 = 1$ and $x_6 = 1/2$. To get the corresponding value for z, look at the z row and observe that the corresponding linear equation is

$$-z - \frac{7}{2}x_2 + \frac{1}{2}x_3 - \frac{5}{2}x_4 = -\frac{25}{2},$$

but x_2, x_3, and x_4 are non-basic and so take the value zero giving $-z = -25/2$, or $z = 25/2$.

Of course this is all exactly the same information we obtained from the dictionary approach. The simplex, or augmented matrix approach is simply a more efficient computational procedure. For computational purposes, we use the tableau form of the simplex in to solve specific LPs. However, in order to understand the inner workings of the algorithm it is *essential* that you understand how to go back and forth between these two representations, i.e the dictionary representation and its corresponding simplex tableau (or, augmented matrix). Let us now continue with the second simplex pivot.

In every tableau we always reserve the bottom row for encoding the linear relationship between the objective variable z and the currently non-basic variables. For this reason we call this row the *objective row*, and to distinguish its special role, we place a line above it in the tableau (this is reminiscent of the way we place a vertical bar in an augmented matrix to distinguish the right hand side of a linear equation). In the objective row of the tableau (2.11),

$$[-1,\ 0,\ -7/2,\ 1/2,\ -5/2,\ 0,\ 0,\ |-25/2],$$

we see a positive coefficient, 1/2, in the 4th column. Hence the objective row coefficient for the non-basic variable x_3 in this tableau is 1/2. This indicates that if we increase the the value of x_3, we also increase the value of the objective z. This is not true for any of the other currently non-basic variables since their cost row coefficients are all non-positive. Thus, the only way to increase the value of z is to bring x_3 into the basis, or equivalently, pivot on the x_3 column which is the 4th column of the tableau. For this reason, we call the x_3 column the *pivot column*. Now if x_3 is to enter the basis, then which variable leaves? Just as with the dictionary representation, the variable that leaves the basis is that currently basic variable whose non-negativity places the greatest restriction on increasing the value of x_3. This restriction is computed as the smallest ratio of the right hand sides and the positive coefficients in the x_3 column:

$$1 = \min\{(5/2)/(1/2),\ (1/2)/(1/2)\}.$$

The ratios are only computed with the positive coefficients since a non-positive coefficient means that by increasing this variable we do not decrease the valued of the corresponding basic variable and so it is not a restricting equation. Since the minimum ratio in this instance is 1 and it comes from the third row, we find that the *pivot row* is the third row. Looking across the third row, we see that this row identifies x_6 as a basic variable since the x_6 column is a column of the identity with a 1 in the third row. Hence x_6 is the variable leaving the basis when x_3 enters. The intersection of the pivot column and the pivot row is called the *pivot*. In this instance it is the number 1/2 which is the $(3,4)$ entry of the simplex tableau. Pivoting on this entry requires us to first make it 1 by multiplying this row through by 2, and then to apply Gaussian elimination to force all other entries in this column to zero:

			Pivot column ↓					ratios	
0	1	3/2	1/2	1/2	0	0	5/2	5	
0	0	−5	0	−2	1	0	1		
0	0	−1/2	(1/2)	−3/2	0	1	1/2	(1)	← pivot row
−1	0	−7/2	(1/2)	−5/2	0	0	−25/2		

0	1	2	0	2	0	−1	2
0	0	−5	0	−2	1	0	1
0	0	−1	1	−3	0	2	1
−1	0	−3	0	−1	0	−1	−13

This simplex tableau is said to be an *optimal tableau* since it is feasible (the associated BFS is non-negative) and the cost row coefficients for the variables are all non-positive. A BFS obtained from an optimal tableau is called and *optimal basic feasible solution.* The optimal BFS is obtained by setting the non-basic variables equal to zero and setting the basic variables equal to the value on the right hand side corresponding to the one in its column: $x_1 = 2,\ x_2 = 0,\ x_3 = 1,\ x_4 = 0,\ x_5 = 1,\ x_6 = 0$. The optimal objective value is obtained by taking the negative of the number in the lower right hand corner of the optimal tableau: $z = 13$.

We now recap the complete sequence of pivots in order to make a final observation that will help streamline the pivoting process: pivots are circled,

z	x_1	x_2	x_3	x_4	x_5	x_6	
0	(2)	3	1	1	0	0	5
0	4	1	2	0	1	0	11
0	3	4	2	0	0	1	8
-1	5	4	3	0	0	0	0
0	1	3/2	1/2	1/2	0	0	5/2
0	0	−5	0	−2	1	0	1
0	0	−1/2	(1/2)	−3/2	0	1	1/2
−1	0	−7/2	1/2	−5/2	0	0	−25/2
0	1	2	0	2	0	−1	2
0	0	−5	0	−2	1	0	1
0	0	−1	1	−3	0	2	1
−1	0	−3	0	−1	0	−1	−13

Observe from this sequence of pivots that the z column is never touched, that is, it remains the same in all tableaus. Essentially, it just serves as a place holder reminding us that in the linear equation for the cost row the coefficient of z is -1. Therefore, for the sake of expediency we will drop this column from our simplex computations in most settings, and simply re-insert it whenever instructive or convenient. However, *it is of great importance to always remember that it is there*! Indeed, we will make explicit and essential use of this column in order to arrive at a full understanding of the duality theory for linear programming. After removing this column, the above pivots take the following form:

x_1	x_2	x_3	x_4	x_5	x_6	
(2)	3	1	1	0	0	5
4	1	2	0	1	0	11
3	4	2	0	0	1	8
5	4	3	0	0	0	0
1	3/2	1/2	1/2	0	0	5/2
0	−5	0	−2	1	0	1
0	−1/2	(1/2)	−3/2	0	1	1/2
0	−7/2	1/2	−5/2	0	0	−25/2
1	2	0	2	0	−1	2
0	−5	0	−2	1	0	1
0	−1	1	−3	0	2	1
0	−3	0	−1	0	−1	−13

We close this section with a final example of simplex pivoting on a tableau giving only the essential details.

The LP

$$
\begin{array}{lrcrcrcl}
\text{maximize} & 3x_1 & + & 2x_2 & - & 4x_3 & & \\
\\
\text{subject to} & x_1 & + & 4x_2 & & & \leq & 5 \\
 & 2x_1 & + & 4x_2 & - & 2x_3 & \leq & 6 \\
 & x_1 & + & x_2 & - & 2x_3 & \leq & 2 \\
 & 0 & \leq & x_1, & x_2, & x_3 & &
\end{array}
$$

Simplex Iterations

x_1	x_2	x_3	x_4	x_5	x_6		ratios
1	4	0	1	0	0	5	5
2	4	-2	0	1	0	6	3
(1)	1	-2	0	0	1	2	2
3	2	-4	0	0	0	0	

0	3	2	1	0	-1	3	3/2
0	2	(2)	0	1	-2	2	1
1	1	-2	0	0	1	2	
0	-1	2	0	0	-3	-6	

0	1	0	1	-1	1	1
0	1	1	0	1/2	-1	1
1	3	0	0	1	-1	4
0	-3	0	0	-1	-1	-8

Optimal Solution

$$\begin{bmatrix} x_x \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \qquad \text{optimal value} \ = \ 8$$

This example illustrates a point that needs to be strongly emphasized. The pivot column in the second tableau is chosen to be the x_3 column since its objective row coefficient, "2", is the only positive entry in the objective row. Hence it is the only non-basic variable whose increase will increase the objective since the objective row in the dictionary is

$$z = 6 - x_2 + 2x_3 - 3x_6.$$

To continue pivoting, we now choose the pivot column by forming the ratios as shown, but we did not form the ratio associated with the entry "-2" in the pivot column. To see why, write out the row of the associated dictionary for the "-2" row. This gives

$$x_1 = 2 - x_2 + 2x_3 - x_6. \tag{2.12}$$

Since the pivot column is the third column, x_3 is the variable entering the basis. That is, on this pivot we increase the value of the currently nonbasic variable x_3 from zero to some positive number. The amount of increase is restricted by the need to keep all of the currently basic variables non-negative. This is why we form the ratios. For example, equation (2.12) above defines the basic variable x_1 in terms of the nonbasic variables x_2, x_3 and x_6. If we now increase the value of x_3 from zero, the value of x_1 increases as well. Consequently, this equation places no restriction on increasing the value of x_3. This is why we do not need to form a ratio for this row since this row places no restriction. *In general, any entry in the pivot column that is non-positive does not yield a restriction on the on the value of the incoming variable.* Thus, one does not need to compute the ratios associated with non-positive values.

A final word of advise; when pivoting by hand, it is helpful to keep the tableaus vertically aligned in order to keep track of the arithmetic operations. This allows you to find errors quickly, and errors will occur. Lined paper helps to keep the rows straight. But the columns need to be straight as well. Many students find that it is easy to keep both the rows and columns straight if they do pivoting on graph paper having large boxes for the numbers.

2.3 Dictionaries: The General Case for LPs in Standard Form

Recall the following standard form for LPs:

$$\begin{array}{rll} \mathcal{P}: & \text{maximize} & c^T x \\ & \text{subject to} & Ax \leq b \\ & & 0 \leq x\ , \end{array}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ and the inequalities $Ax \leq b$ and $0 \leq x$ are to be interpreted componentwise. We now provide a formal definition for a dictionary associated with an LP in standard form. Let

$$(D_I) \qquad \begin{aligned} x_{n+i} &= b_i - \sum_{j=1}^{n} a_{ij} x_j \\ z &= \sum_{j=1}^{n} c_j x_j \end{aligned}$$

be the defining system for the slack variables x_{n+i}, $i = 1, \cdots, n$ and the objective variable z. A dictionary for $\mathcal{P}$ is any system of the form

$$(D_B) \qquad \begin{aligned} x_i &= \widehat{b}_i - \sum_{j \in N} \widehat{a}_{ij} x_j \qquad i \in B \\ z &= \widehat{z} + \sum_{j \in N} \widehat{c}_j x_j \end{aligned}$$

where B and N are index sets contained in the set of integers $\{1, \ldots, n+m\}$ satisfying

(1) B contains m elements,

(2) $B \cap N = \emptyset$

(3) $B \cup N = \{1, 2, \ldots, n+m\}$,

and such that the systems (D_I) and (D_B) have identical solution sets. The set $\{x_j : j \in B\}$ is said to be the basis associated with the dictionary (D_B) (we also refer to the index set B as the basis for the sake of simplicity), and the variables x_i, $i \in N$ are said to be the non-basic variables associated with this dictionary. The *basic solution* identified by this dictionary is

$$\begin{aligned} \hat{x}_i &= \widehat{b}_i \qquad & i \in B \\ \hat{x}_j &= 0 \qquad & j \in N. \end{aligned} \tag{2.13}$$

The dictionary is said to be feasible if $0 \le \widehat{b}_i$ for $i \in B$. If the dictionary D_B is feasible, then the basic solution identified by the dictionary (2.13) is said to be a *basic feasible solution* (BFS) for the LP. A feasible dictionary and its associated BFS are said to be *optimal* if $\widehat{c}_j \le 0 \; j \in N$. The associated optimal solution to the LP is obtained by dropping the slack variable components from the BFS $\hat{x}$, that is the optimal solution to the LP has components $\hat{x}_i, \; i = 1, 2, \dots, n$. At the end of this section, we show that optimal basic feasible solutions are optimal solutions to the linear program $\mathcal{P}$.

Simplex Pivoting by Matrix Multiplication

As we have seen, simplex pivoting can either be performed on dictionaries or on the augmented matrices that encode the linear equations of a dictionary in matrix form. In matrix form, simplex pivoting reduces to our old friend, Gauss-Jordan elimination. In this section, we show that Gauss-Jordan elimination can be represented as a consequence of left multiplication by a specially designed matrix called a *Gauss-Jordan pivot matrix*.

Consider the vectors $e_j \in \mathbb{R}^n, \; j = 1, \dots, n$, where each e_j is is defined to be the vector having a one in the jth position and zeros elsewhere. For example, in $\mathbb{R}^4$, we have

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

The set of vectors $\{e_1, e_2, \dots, e_n\}$ in $\mathbb{R}^n$ is called the *standard unit coordinate basis* for $\mathbb{R}^n$. It is clearly a basis for $\mathbb{R}^n$ in the sense that these vectors are linearly independent and they span $\mathbb{R}^n$. They are called unit vectors since their magnitude is 1. Also observe that they form the columns of the $n \times n$ identity matrix $I_{n \times n}$, i.e

$$I_{n \times n} = [e_1 \, e_2 \, \dots \, e_n].$$

Next consider a vector $v \in \mathbb{R}^m$ block decomposed as

$$v = \begin{bmatrix} a \\ \alpha \\ b \end{bmatrix}$$

where $a \in \mathbb{R}^s$, $\alpha \in \mathbb{R}$, and $b \in \mathbb{R}^t$ with $m = s + 1 + t$. Assume that $\alpha \neq 0$. We wish to determine a matrix G such that

$$Gv = e_{s+1}.$$

We claim that the block matrix

$$G = \begin{bmatrix} I_{s\times s} & -\alpha^{-1}a & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & -\alpha^{-1}b & I_{t\times t} \end{bmatrix}$$

does the trick. Indeed,

$$Gv = \begin{bmatrix} I_{s\times s} & -\alpha^{-1}a & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & -\alpha^{-1}b & I_{t\times t} \end{bmatrix} \begin{pmatrix} a \\ \alpha \\ b \end{pmatrix} = \begin{bmatrix} a-a \\ \alpha^{-1}\alpha \\ -b+b \end{bmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = e_{s+1}.$$

The matrix G is called a *Gauss-Jordan Pivot Matrix*. Note that G is invertible since

$$G^{-1} = \begin{bmatrix} I & a & 0 \\ 0 & \alpha & 0 \\ 0 & b & I \end{bmatrix},$$

and that for any vector of the form $w = \begin{pmatrix} x \\ 0 \\ y \end{pmatrix}$ where $x \in \mathbb{R}^s$ $y \in \mathbb{R}^t$, we have

$$Gw = w.$$

The Gauss-Jordan pivot matrices perform precisely the operations required in order to execute a simplex pivot. That is, each simplex pivot can be realized as left multiplication of the simplex tableau by the appropriate Gauss-Jordan pivot matrix.

For example, consider the following initial feasible tableau:

$$\left[\begin{array}{cccccc|c} 1 & 4 & 2 & 1 & 0 & 0 & 11 \\ 3 & \textcircled{2} & 1 & 0 & 1 & 0 & 5 \\ 4 & 2 & 2 & 0 & 0 & 1 & 8 \\ \hline 4 & 5 & 3 & 0 & 0 & 0 & 0 \end{array}\right]$$

where the $(2,2)$ element is chosen as the pivot element. In this case,

$$s = 1, \quad t = 2, \quad a = 4, \quad \alpha = 2, \quad \text{and} \quad b = \begin{bmatrix} 2 \\ 5 \end{bmatrix},$$

and so the corresponding Gauss-Jordan pivot matrix is

$$G_1 = \begin{bmatrix} I_{1\times 1} & -\alpha^{-1}a & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & -\alpha^{-1}b & I_{2\times 2} \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{-5}{2} & 0 & 1 \end{bmatrix}.$$

Multiplying the simplex on the left by G_1 gives

$$\begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{-5}{2} & 0 & 1 \end{bmatrix} \left[\begin{array}{cccccc|c} 1 & 4 & 2 & 1 & 0 & 0 & 11 \\ 3 & 2 & 1 & 0 & 1 & 0 & 5 \\ 4 & 2 & 2 & 0 & 0 & 1 & 8 \\ \hline 4 & 5 & 3 & 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ \frac{3}{2} & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{5}{2} \\ 1 & 0 & \textcircled{1} & 0 & -1 & 1 & 3 \\ \hline -\frac{7}{2} & 0 & \frac{1}{2} & 0 & \frac{-5}{2} & 0 & \frac{-25}{2} \end{array}\right].$$

Repeating this process with the new pivot element in the $(3,3)$ position yields the Gauss-Jordan pivot matrix

$$G_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{2} & 1 \end{bmatrix},$$

and left multiplication by G_2 gives

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{2} & 1 \end{bmatrix} \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ \frac{3}{2} & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{5}{2} \\ 1 & 0 & 1 & 0 & -1 & 1 & 3 \\ \hline \frac{-7}{2} & 0 & \frac{1}{2} & 0 & \frac{-5}{2} & 0 & \frac{-25}{2} \end{array}\right] = \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & \frac{-1}{2} & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 & 3 \\ \hline -4 & 0 & 0 & 0 & \frac{-3}{2} & \frac{-1}{2} & -14 \end{array}\right]$$

yielding an optimal tableau.

If

$$T_0 := \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} \tag{2.4}$$

is the initial tableau, then

$$G_2 G_1 T_0 = \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & \frac{-1}{2} & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 & 3 \\ \hline -4 & 0 & 0 & 0 & -2 & \frac{-1}{2} & -14 \end{array}\right]$$

That is, we would be able to go directly from the initial tableau to the optimal tableau if we knew the matrix

$$G = G_2 G_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{-5}{2} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -2 & \frac{-1}{2} & 1 \end{bmatrix}$$

beforehand. Moreover, the matrix G is invertible since both G_1 and G_2 are invertible:

$$G^{-1} = G_1^{-1}G_2^{-1} = \begin{bmatrix} 1 & 4 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 5 & 3 & 1 \end{bmatrix}$$

(you should check that $GG^{-1} = I$ by doing the multiplication by hand). In general, every sequence of simplex pivots has a representation as left multiplication by a single invertible matrix since since pivoting corresponds to left multiplication of the tableau by a Gauss-Jordan pivot matrix, and Gauss-Jordan pivot matrices are always invertible. We now examine the consequence of this observation more closely in the general case. In this discussion, it is essential that we include the column associated with the objective variable z which we have largely ignored up to this point.

Recall the initial simplex tableau, or augmented matrix associated with the system (D_I):

$$T_0 = \begin{bmatrix} 0 & A & I & b \\ -1 & c^T & 0 & 0 \end{bmatrix} .$$

Observe that we include the first column, i.e. the column associated with the objective variable z in the augmented matrix. Let the matrix

$$T_k = \begin{bmatrix} 0 & \widehat{A} & R & \widehat{b} \\ -1 & \widehat{c}^T & -\widehat{y}^T & \widehat{z} \end{bmatrix}$$

be another simplex tableau obtained from the initial tableau after a sequnce of k simplex pivots. The first column remains unchanged. since simplex pivots do not alter the first column. This is the reason why it does not appear in our hand computations. However, in this discussion, its presence and the fact that it remains unchanged by simplex pivoting is key! Since T_k is another simplex tableau the $m \times (n+m)$ matrix $[\widehat{A} \quad R]$ must posses among its columns the m columns of the $m \times m$ identity matrix. These columns of the identity matrix correspond to the basic variables associated with this tableau (except in the very unusual case when there are more than m columns of the identity present).

Our prior discussion on Gauss-Jordan pivot matrices tells us that T_k can be obtained from T_0 by multiplying T_0 on the left by some nonsingular $(m+1) \times (m+1)$ matrix G where G is the product of a sequence of Gauss-Jordan pivot matrices. In order to better understand the action of G on T_0 we need to decompose G into a block structure that is conformal with that of T_0:

$$G = \begin{bmatrix} M & u \\ v^T & \beta \end{bmatrix} ,$$

where $M \in \mathbb{R}^{m\times m}$, $u, v \in \mathbb{R}^m$, and $\beta \in \mathbb{R}$. Then

$$\begin{aligned}
\begin{bmatrix} 0 & \widehat{A} & R & \widehat{b} \\ -1 & \widehat{c}^T & -y^T & \widehat{z} \end{bmatrix} &= T_k \\
&= GT_0 \\
&= \begin{bmatrix} M & u \\ v^T & \beta \end{bmatrix} \begin{bmatrix} 0 & A & I & b \\ -1 & c^T & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} -u & MA + uc^T & M & Mb \\ -\beta & v^T A + \beta c^T & v^T & v^T b \end{bmatrix} .
\end{aligned}$$

By equating the blocks in the matrices on the far left and far right hand sides of this equation, we find from the first column that

$$u = 0 \qquad \text{and} \qquad \beta = 1 \ .$$

Here we see the key role played by our knowledge of the structure of the first column, i.e. the z or objective variable column. From the $(1, 3)$ and the $(2, 3)$ terms on the far left and right hand sides of (2.5), we also find that

$$M = R, \qquad \text{and} \qquad v = -y \ .$$

Putting all of this together gives the following representation of the k^{th} tableau T_k:

$$T_k = \begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} 0 & A & I & b \\ -1 & c^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & RA & R & Rb \\ -1 & c^T - \widehat{y}^T A & -y^T & -\widehat{y}^T b \end{bmatrix} , \tag{2.5}$$

where the matrix R is necessarily invertible since the matrix

$$G = \begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix}$$

is invertible (prove this!):

$$G^{-1} = \begin{bmatrix} R^{-1} & 0 \\ y^T R^{-1} & 1 \end{bmatrix} . \qquad \text{(check this out by computing the product } GG^{-1}\text{)}$$

The matrix R is called the *record* matrix for the tableau as it keeps track of almost all of the transformations required to obtain the new tableau. Again, the variables associated with the columns of the identity correspond to the basic variables. The tableau T_k is said to be *primal feasible*, or just feasible, if $\widehat{b} = Rb \geq 0$.

The beautiful structure revealed by equation (2.5) is perhaps the most important equation to be given in our discussion of linear programming. It is of central importance to linear programming duality theory and sensitivity analysis. *Its importance cannot be overstated.* We call equation (2.5) the *basic pivoting equation for the simplex algorithm.*

As a first step toward understanding the central significance of the equation (2.5), consider the case where the tableau T_k on the right hand side of (2.5) is an optimal tableau, i.e.

$$\begin{bmatrix} 0 & RA & R & Rb \\ -1 & c^T - y^T A & -y^T & -y^T b \end{bmatrix} \tag{2.6}$$

is an optimal tableau for $\mathcal{P}$. Recall that (2.6) is optimal if and only if it is feasible, $Rb \geq 0$, and all of the variable coefficients in the objective row are non-positive,

$$A^T y \geq c \quad \text{and} \quad 0 \leq y \ , \tag{2.7}$$

in which case we claim that the associated BFS, say $\tilde{x}$, is an optimal solution to the LP with optimal value $c^T \hat{x} = z = b^T y$. The vector $\widehat{x}$ corresponds to the vector x with all of the slack variables removed. In particular, $\widehat{x}$ is feasible for $\mathcal{P}$ with objective value $c^T \widehat{x} = z = b^T y$. Now observe that the system (2.7) says that y is feasible for the dual problem $\mathcal{D}$ with dual objective value $b^T y$. **This is absolutely amazing** since the Weak Duality Theorem now tells us that $\widehat{x}$ solves $\mathcal{P}$ and y solves $\mathcal{D}$!!! That is, any optimal tableau similtaneously gives optimal solutions to *both* the primal and dual problems! We have just proved the following theorem.

Theorem 2.1 (Optimal Tableau Theorem) *Let x be the basic feasible solution for the tableau* (2.6)*. If* (2.6) *is an optimal tableau for the linear program $\mathcal{P}$, then y is an optimal solution to the dual problem $\mathcal{D}$ and the vector $\widehat{x} \in \mathbb{R}^n$ given by $\widehat{x}_j = x_j, \ j = 1, 2, \dots, n$ is an optimal solution to $\mathcal{P}$.*

The Optimal Tableau Theorem motivates the following definition.

Definition 2.1 (Primal and Dual Feasible Dictionaries and Tableaus) *Consider the linear program*

$$\begin{array}{rll} \mathcal{P}: & \textit{maximize} & c^T x \\ & \textit{subject to} & Ax \leq b \\ & & 0 \leq x \ , \end{array}$$

and let

$$(D_B) \qquad \begin{aligned} x_i &= \widehat{b}_i - \sum_{j \in N} \widehat{a}_{ij} x_j \qquad i \in B \\ z &= \widehat{z} + \sum_{j \in N} \widehat{c}_j x_j \end{aligned}$$

be a dictionary for this linear program with associated augmented matrix, or equivalently, simplex tableau

$$T := \begin{bmatrix} 0 & RA & R & Rb \\ -1 & c^T - y^T A & -y^T & -y^T b \end{bmatrix}$$

This tableau is said to be primal feasible for $\mathcal{P}$ if $Rb \geq 0$. It is said to be dual feasible if $c^T - y^T A \leq 0$ (or equivalently $A^T y \geq c$) and $0 \leq y$. If a simplex tableau is both primal and dual feasible, then it is said to be optimal.

The Optimal Tableau Theorem tells us that if the simplex algorithm works, in the sense that it arrives at an optimal tableau after a finite number of simplex pivots, then we have a method for solving all LPs! Unfortunately, the situation is not as simple as this. First, not every LP is feasible, so a solution obviously cannot exist. Second, even if an LP is feasible, it may be unbounded, so, again, a solution does not exist. Finally, even if a solution exists, we have no guarantee that the simplex algorithm can find it after a finite number of pivots, or for that matter an infinite number of pivots. To understand the relationship between linear programs and the simplex algorithm, we need a much deeper understanding of the algorithm itself.