Algebra Underlying the Revised Simplex Algorithm

Consider an LP problem in the form

(1) maximize $c'x$ subject to $Ax=b$ and $x\geq 0$

Any problem in standard form can be expressed in this manner after the introduction of slack variables.

The simplex algorithm solves (1) by creating a sequence of dictionaries. Each one corresponds to a division of the variables into two classes: basic and nonbasic. By knowing which variables are basic, you have enough information to completely determine the dictionary. We will demonstrate this first on a concrete example, and then on the general abstract problem (1).

First, the example. Consider the problem in standard form

$$\text{to maximize } x_1+2x_2 \text{ subject to}$$
$$x_1+x_2 \leq 6,\quad x_1 \leq 3,\quad x_2 \leq 5$$
$$x_1,\ x_2 \geq 0$$

which can be rewritten, after introducing slack variables, in the form of (1):

(2)
$$\begin{aligned} &\text{maximize } x_1 + 2x_2 + 0x_3 + 0x_4 + 0x_5 \text{ subject to}\\ &x_1 + x_2 + x_3 = 6\\ &x_1 + x_4 = 3\\ &x_2 + x_5 = 5\\ &x_1,\ x_2,\ x_3,\ x_4,\ x_5 \geq 0\end{aligned}$$

Let us compute directly (without pivoting) the dictionary for which x_1, x_2, and x_4 are basic. (This choice of variables is arbitrary; we are not executing the simplex algorithm, only showing how to write down the dictionary corresponding to a given choice of variables). Write the equations in (2) as

(3)
$$\begin{pmatrix}6\\3\\5\end{pmatrix} = \begin{pmatrix}1&1&0\\1&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_4\end{pmatrix} + \begin{pmatrix}1&0\\0&0\\0&1\end{pmatrix}\begin{pmatrix}x_3\\x_5\end{pmatrix}$$

and the objective function as

(4)
$$z = (1\ 2\ 0)\begin{pmatrix}x_1\\x_2\\x_4\end{pmatrix} + (0\ 0)\begin{pmatrix}x_3\\x_5\end{pmatrix}.$$

Using (3), we can solve for x_1, x_2, and x_4, to get:

(5)
$$\begin{pmatrix}x_1\\x_2\\x_4\end{pmatrix} = \begin{pmatrix}1&1&0\\1&0&1\\0&1&0\end{pmatrix}^{-1}\left[\begin{pmatrix}6\\3\\5\end{pmatrix} - \begin{pmatrix}1&0\\0&0\\0&1\end{pmatrix}\begin{pmatrix}x_3\\x_5\end{pmatrix}\right]$$

$$= \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 3 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_5 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_5 \end{pmatrix}$$

Substituting (5) into (4), we can compute the objective function in terms of the nonbasic variables x_3 and x_5:

$$(6) \qquad z = (1\ 2\ 0) \left[\begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_5 \end{pmatrix} \right] + (0\ 0) \begin{pmatrix} x_3 \\ x_5 \end{pmatrix}$$

$$= (1\ 2\ 0) \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} + \left[(0,\ 0) - (1\ 2\ 0) \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} \right] \begin{pmatrix} x_3 \\ x_5 \end{pmatrix}$$

$$= 11 - x_3 - x_5$$

Combining (5) and (6), we see that the desired dictionary is

$$\begin{aligned} x_1 &= 1 - x_3 + x_5 \\ x_2 &= 5 \qquad\ - x_5 \\ x_4 &= 2 + x_3 - x_5 \\ z &= 11 - x_3 - x_5 \end{aligned}$$

Next, we will show how a choice of basis determines the dictionary for the general abstract problem (1). As we proceed, we will relate the abstract problem to the example we have just worked through. Comparing (1) and (2), we see that in the case of the example, we have

$$c = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 6 \\ 3 \\ 5 \end{pmatrix}.$$

Notice that each column of A corresponds to one of the five variables. We want to solve the equation Ax=b for the basic variables. Denote the vector of basic variables as x_B and the vector of nonbasic variables as x_N. Let A_B and A_N be the matrices obtained from the basic columns of A and the nonbasic columns, respectively. Also, let c_N and c_B be the vectors denoting the nonbasic and basic components of c. In the above example, we would have

$$x_N = \begin{pmatrix} x_3 \\ x_5 \end{pmatrix} \quad A_N = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \quad c_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad x_B = \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} \quad A_B = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad c_B = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

(A word about notation: The order in which the variables are written down is not important. We could equally well have written

$$x_N = \begin{pmatrix} x_5 \\ x_3 \end{pmatrix} \qquad A_N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and we could also have written

$$x_B = \begin{pmatrix} x_4 \\ x_1 \\ x_2 \end{pmatrix} \qquad A_B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad c_B = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

but it is important to maintain consistency between the ordering of the variables in x_N and c_N and the ordering of the columns in A_N so that the order of the variables corresponds to the ordering of the columns. The same observation holds for x_B, c_B, and A_B.)

The equation Ax=b can now be written as $A_N x_N + A_B x_B = b$. Solving for x_B, we get

(7) $$x_B = A_B^{-1} b - A_B^{-1} A_N x_N$$

which corresponds to (5) in the example, where $A_B^{-1} b = \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix}$ and where $A_B^{-1} A_N = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix}$. Next we will express the objective function in terms of the nonbasic variables. Substituting from (7), we get

(8) $$\begin{aligned} z = c'x &= c_B' x_B + c_N' x_N \\ &= c_B'(A_B^{-1} b - A_B^{-1} A_N x_N) + c_N' x_N \\ &= c_B' A_B^{-1} b + (c_N' - c_B' A_B^{-1} A_N) x_N . \end{aligned}$$

which corresponds to (6) in the example, where $c_B' A_B^{-1} b = 11$ and $(c_N' - c_B' A_B^{-1} A_N) = (-1\ -1)$. Combining (7) and (8), we obtain the dictionary

$$\begin{array}{l} x_B = A_B^{-1} b - A_B^{-1} A_N x_N \\ \hline z \ = c_B' A_B^{-1} b + (c_N' - c_B' A_B^{-1} A_N) x_N \end{array}$$