Chapter 4, Operations Research (OR)

Kent Andersen

February 7, 2007

1 Linear Programs (continued)

In the last chapter, we introduced the concept of a basis of a linear program (P)

$$\begin{aligned} \text{Minimize } Z &= c^T x \\ \text{s.t.} & \\ Ax &= b, \qquad \text{(P)} \\ x &\geq 0_n, \end{aligned}$$

Recall that a basis of (P) is a subset $B \subseteq \{1, 2, \ldots, n\}$ of the variables. We also introduced the concepts of a basic solution x^B of (P), and a dual basic solution y^B associated with B and the dual (D) of (P). We finished by proving

Lemma 1 *Let B be an arbitrary basis for (P). If B is a feasible basis for both (P) and (D), then x^B is optimal for (P), and y^B is optimal for (D).*

Example 1 *Consider, again, the linear program that models the WGC problem, which we here call (P)*

$$\begin{aligned} \textit{Minimize } Z &= -3x_1 - 5x_2 \\ \textit{subject to} & \\ -x_1 - x_3 &= -4 \qquad (y_1) \\ -2x_2 - x_4 &= -12 \qquad (y_2) \\ -3x_1 - 2x_2 - x_5 &= -18 \qquad (y_3) \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 &\geq 0, x_5 \geq 0 \end{aligned}$$

where the dual variables corresponding to the three constraints have been written next to them. Also consider the dual linear program (D)

$$\begin{aligned} \textit{Maximize } Z' &= -4y_1 - 12y_2 - 18y_3 \\ \textit{subject to} & \\ -y_1 - 3y_3 &\leq -3 \qquad (x_1) \\ -2y_2 - 2y_3 &\leq -5 \qquad (x_2) \\ -y_1 &\leq 0, \qquad (x_3) \\ -y_2 &\leq 0, \qquad (x_4) \\ -y_3 &\leq 0. \qquad (x_5) \end{aligned}$$

Recall that the set $B := \{1,2,3\}$ defines a basis for (P), and that $x^B = (2,6,2,0,0)$. The dual basic solution y^B corresponding to B is given by the unique solution to the system

$$\begin{aligned} -y_1 - 3y_3 &= -3, \\ -2y_2 - 2y_3 &= -5, \\ -y_1 &= 0. \end{aligned}$$

which gives $y^B = (0, 1\frac{1}{2}, 1)$. Since the basic solution x^B is feasible for (P), and the corresponding dual basic solution y^B is feasible for (D), it follows from Lemma 1 that x^B is optimal for (P), and y^B is optimal for (D).

2 The simplex tableau

Consider, now, a linear program (P) in the following form

$$\begin{aligned} &\text{Minimize } Z = c^T x \\ &\text{s.t.} \\ &Ax = b \\ &\ \ x \geq 0_n, \end{aligned}$$

and the corresponding dual (D) of (P)

$$\begin{aligned} &\text{Maximize } Z' = b^T y \\ &\text{s.t.} \\ &(a_{\cdot j})^T y \leq c_j \quad \text{for } j \in \{1,2,\ldots,n\}. \end{aligned}$$

Let $B \subseteq \{1,2,\ldots,n\}$ be a basis for (P), and let $N := \{1,2,\ldots,n\} \setminus B$ denote the non-basic variables. In Chapter 3, we noted that the basis B can be used to rewrite the system "Ax=b", such that the basic variables are expressed in terms of the non-basic variables. We will now achieve this.

The following notation is needed. As earlier, we will let A_B denote the (column induced) submatrix of A obtained from the variables in B. Also, let $e_i \in \mathbb{R}^m$ denote the m-dimensional *unit vector*, or in other words, e_i has the value "1" on the i^{th} coordinate, and the value "0" on all other coordinates.

Since the matrix A_B is non-singular, and the vectors $a_{\cdot j}$ for $j \in B$ are the columns of the matrix A_B, it follows that $A_B^{-1} a_{\cdot j}$ is one of the m-dimensional unit vectors when $j \in B$. In other words, the matrix $(A_B^{-1})A_B$ can be obtained from the m-dimensional identity matrix I_m by simply permuting the columns of I_m. Given $i \in \{1,2,\ldots,m\}$, let $j(i)$ denote the variable $j(i) \in B$ such that $(A_B^{-1})a_{\cdot j(i)}$ is the i^{th} unit vector. Also, for every non-basic variable $j \in N$, let $\bar{a}^B_{\cdot j} := (A_B^{-1})a_{\cdot j}$.

Multiplying the linear system "$Ax = b$" with A_B^{-1} on both sides then gives

$$x_{j(i)} + \sum_{j \in N} \bar{a}^B_{i,j} x_j = (A_B^{-1} b)_i \quad \text{for } i = 1,2,\ldots,m,$$

where $(A_B^{-1} b)_i$ denotes the i^{th} component of the m-dimensional vector $A_B^{-1} b$. Note that, since $x_j^B = 0$ for all $j \in N$ in the basic solution x^B corresponding to B, the value $(A_B^{-1} b)_i$ is identical to the value $x^B_{j(i)}$. We can therefore rewrite the above system as

$$x_{j(i)} + \sum_{j \in N} \bar{a}^B_{i,j} x_j = x^B_{j(i)} \quad \text{for } i = 1,2,\ldots,m.$$

The above linear system expresses the basic variables $x_{j(i)}$ (where $j(i) \in B$) in terms of the non-basic variables x_j (where $j \in N$). For each basic variable $x_{j(i)}$, the equation $x_{j(i)} + \sum_{j\in N} \bar{a}^B_{i,j} x_j = x^B_{j(i)}$ explains how the value $x^B_{j(i)}$ of the basic variable $x_{j(i)}$ changes when the value of a non-basic variable x_j is increased from its current value (zero). This fact becomes important when we would like to change from one basis B to another basis B'.

To obtain the so-called simplex tableau, we need to use the above equalities to rewrite the objective function $Z = c^T x$ in terms of the non-basic variables only:

$$
\begin{aligned}
c^T x &= \\
\sum_{j=1}^{n} c_j x_j &= \\
\sum_{i=1}^{m} c_{j(i)} x_{j(i)} + \sum_{j\in N} c_j x_j &= \\
\sum_{i=1}^{m} c_{j(i)} (x^B_{j(i)} - \sum_{j\in N} \bar{a}^B_j x_j) + \sum_{j\in N} c_j x_j &= \qquad \text{(using the above equalities)} \\
\sum_{i=1}^{m} c_{j(i)} x^B_{j(i)} + \sum_{j\in N} (c_j - \sum_{i=1}^{m} c_{j(i)} \bar{a}^B_{i,j}) x_j &= \\
c^T x^B + \sum_{j\in N} (c_j - \sum_{i=1}^{m} ((a_{.j(i)})^T y^B) \bar{a}^B_{i,j}) x_j &= \quad \text{(since } c^T x^B = \sum_{i=1}^{m} c_{j(i)} x^B_{j(i)} \text{ and } (a_{.j(i)})^T y^B = c_{j(i)}) \\
c^T x^B + \sum_{j\in N} (c_j - (y^B)^T \sum_{i=1}^{m} a_{.j(i)} \bar{a}^B_{i,j}) x_j &= \\
c^T x^B + \sum_{j\in N} (c_j - (a_{.j})^T y^B) x_j. & \quad \text{(since } \sum_{i=1}^{m} a_{.j(i)} \bar{a}^B_{i,j} = (A_B)(\bar{a}^B_{.j}) = A_B A_B^{-1} a_{.j} = a_{.j})
\end{aligned}
$$

After these calcutions, we therefore have that the equality $Z = c^T x$ can be rewritten as $Z = c^T x^B + \sum_{j\in N} r^B_j x_j$, where r^B_j is defined by $r^B_j := c_j - (a_{.j})^T y^B$ (recall that y^B denotes the dual basic solution corresponding to B). Observe that the scalars r^B_j for $j \in N$ simply state whether the constraint $(a^T_{j.})y \leq c_j$ in the dual (D) of (P) is satisfied or not by the dual basic solution y^B. The quantities r^B_j are called the *reduced costs* corresponding to B. The *simplex tableau* is the the following equality system related to the basis B

$$Z = c^T x^B + \sum_{j\in N} r^B_j x_j \tag{1}$$

$$x^B_{j(i)} = x_{j(i)} + \sum_{j\in N} \bar{a}^B_{i,j} x_j \quad \text{for } i = 1, 2, \ldots, m. \tag{2}$$

The interpretation of the reduced costs r^B_j for $j \in N$ are as follows. If we increase the value of x_j by one unit, and modify the current values of the basic variables according to (2), then we need to increase the cost Z by r^B_j units in order for (1) and (2) to be satisfied. Observe that, if B is a basis that satisfies the conditions in Lemma 1, or in other words, if x^B is feasible for (P) and y^B is feasible for (D), then $r^B_j \geq 0$ for all $j \in N$, which simply states that the cost $Z = c^T x$ can not be decreased by increasing the value of *any* non-basic variable. This is the economic intuition behind optimality, which we will discuss in further detail in the next section.

Example 2 *Consider, again, the linear program that models the WGC problem, where slack variables x_3, x_4 and x_5 have been introduced in the constraints, and "Maximize" is changed to "Minimize".*

$$
\begin{aligned}
&\textit{Minimize } Z = -3x_1 - 5x_2\\
&\textit{subject to}\\
&x_1 + x_3 = 4\\
&2x_2 + x_4 = 12\\
&3x_1 + 2x_2 + x_5 = 18\\
&x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0
\end{aligned}
$$

As noted in Chapter 3, the set $B = \{1, 2, 4\}$ constitutes a feasible basis, and the corresponding basic solution is given by $x^B = (4, 3, 0, 6, 0)$. The corresponding matrix A_B, and its inverse, are given by:

$$
A_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{pmatrix} \textit{ and } A_B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & 0 & \frac{1}{2} \\ 3 & 1 & -1 \end{pmatrix}.
$$

Since $A_B^{-1}a_{.1} = e_1$, $A_B^{-1}a_{.2} = e_2$ and $A_B^{-1}a_{.4} = e_3$, it follows that $j(1) = 1$, $j(2) = 2$ and $j(3) = 4$. The non-basic set is given by $N = \{3, 5\}$, $a_{.3}^T = (1, 0, 0)$ and $a_{.5}^T = (0, 0, 1)$. It follows that $\bar{a}_{.3}^B = (1, -\frac{3}{2}, 3)^T$ and $\bar{a}_{.5}^B = (0, \frac{1}{2}, -1)^T$.

The dual solution y^B corresponding to B can be computed by solving the system $(a_{.1})^T y = c_1$, $(a_{.2})^T y = c_2$ and $(a_{.4})^T y = c_4$. Since $c_1 = -3$, $c_2 = -5$ and $c_4 = 0$, this gives $y^B = (4\frac{1}{2}, 0, -2\frac{1}{2})$. The reduced costs for the non-basic variables x_3 and x_5 are therefore $r_3^B = c_3 - (a_{.3})^T y^B = -4\frac{1}{2}$ and $r_5^B = c_5 - (a_{.5})^T y^B = 2\frac{1}{2}$. The objective value of the basic solution x^B is $c^T x^B = -27$. The simplex tableau associated with B can now be written as

$$
\begin{aligned}
&Z = -27 - 4\frac{1}{2}x_3 + 2\frac{1}{2}x_5,\\
&4 = x_1 + x_3,\\
&3 = x_2 - \frac{3}{2}x_3 + \frac{1}{2}x_5,\\
&6 = x_4 + 3x_3 - x_5.
\end{aligned}
$$

Observe that the reduced cost of x_3 is negative, which indicates that the basis B is not dual feasible, or in other words, the dual constraint $(a_{.3})^T y \leq c_3$, or equivalently $y_1 \leq 0$, is not satisfied by y^B.

3 Finding a better basis: pivoting

Suppose, now, that we have a feasible basis B for (P). The assumpition that B is a feasible basis means that the corresponding basic solution x^B for (P) satisfies $x^B \geq 0_n$ and $\sum_{j \in B} a_{.j} x_j^B = b$. If the the corresponding basic dual solution y^B is feasible for (D), the linear program (P) is solved by the basis B. In this section we assume y^B is *not* feasible for (D), and the question is how to obtain *another* feasible basis B' for (P) that is "closer" to optimality for (P). As noted earlier, we will be looking for bases B' of the form $B' = (B \cup \{\bar{j}\}) \setminus \{\bar{k}\}$, where $\bar{k} \in B$ is a basic variable and $\bar{j} \in N$ is a non-basic variable.

Suppose we have chosen the non-basic variable $\bar{j}$. According to the discussion in the previous section, a good choice for a non-basic variable is a variable with a negative reduced cost $r_{\bar{j}}^B < 0$, since this indicates that we can decrease the cost $Z = c^T x$ by increasing the value of $x_{\bar{j}}$ from its current value (zero). The question is how much $x_{\bar{j}}$ can be increased so that we still have a feasible solution to (P). Also, we would like the remaining non-basic variables in $N \setminus \{\bar{j}\}$ to remain at their current value of zero. In other words, we would like the new basic solution $x^{B'}$ to be a solution to the system

$$Z = c^T x^B + r^B_{\bar{j}} x_{\bar{j}} \tag{3}$$

$$x^B_{j(i)} = x_{j(i)} + \bar{a}^B_{i,\bar{j}} x_{\bar{j}} \quad \text{for } i = 1, 2, \ldots, m. \tag{4}$$

For the above system to remain satisfied, if we change the value of $x_{\bar{j}}$ from zero to a positive value of $x_{\bar{j}} = \delta > 0$, we must change the value of $x^B_{j(i)}$ to $x'_{j(i)} = x^B_{j(i)} - \bar{a}^B_{i,\bar{j}}\delta$ for $i = 1, 2, \ldots, m$. Further, if we want the "new" solution x' to be feasible for (P), or in other words, if we want x' to satisfy $x' \geq 0_n$ (the constraints "$Ax = b$" remain satisfied by construction), then we need δ to satisfy

$$\delta \leq \frac{x^B_{j(i)}}{\bar{a}^B_{i,\bar{j}}} \text{ for every } i \in \{1, 2, \ldots, m\} \text{ that satisfies } \bar{a}^B_{i,\bar{j}} > 0. \tag{5}$$

The largest possible value δ^* of δ such that (5) is satisfied provides the largest possible increase in the value of $x_{\bar{j}}$ such that the solution $x'_{j(i)} := x^B_{j(i)} - \bar{a}^B_{i,\bar{j}}\delta^*$ for $i = 1, 2, \ldots, m$, $x'_{\bar{j}} := \delta^*$ and $x'_j := 0$ for $j \in N \setminus \{\bar{j}\}$ is feasible for (P). Furthermore, observe that, if $\delta^* > 0$, then the cost of the new solution x' will be $|r^B_{\bar{j}}|\delta^* > 0$ lower than the cost of the solution x^B to (P) (since $r^B_{\bar{j}} < 0$), and therefore the solution x' is a better solution to (P) than x^B.

Let $\bar{i} \in \{1, 2, \ldots, m\}$ be such that $\bar{a}^B_{\bar{i},\bar{j}} > 0$ and $\delta^* = \frac{x^B_{j(\bar{i})}}{\bar{a}^B_{\bar{i},\bar{j}}}$. Observe that the "new" solution x' satisfies $x'_{j(\bar{i})} = 0$. Therefore, a new feasible basis B' for (P) should be available by simply declaring $j(\bar{i})$ to be non-basic and $\bar{j}$ to be basic, or in other words, updating the basis B to $B' = (B \cup \{\bar{j}\}) \setminus \{j(\bar{i})\}$. We have shown that the solution $x^{B'}$ will be a feasible solution for (P). We will later verify that $x^{B'}$ is indeed a basic solution. The variable $\bar{j}$ is called the *entering* non-basic variable, and the variable $j(\bar{i})$ is called the *leaving* basic variable. The entire operation of changing the basis from B to B' is called a *pivot*, and the element $\bar{a}_{\bar{i},\bar{j}}$ of the simplex tableau is called the *pivot element*.

Example 3 *Consider, again, the simplex tableau for the WGC problem for the basis $B = \{1, 2, 4\}$.*

$$\begin{aligned}
Z &= -27 - 4\frac{1}{2}x_3 + 2\frac{1}{2}x_5, \\
4 &= x_1 + x_3, \\
3 &= x_2 - \frac{3}{2}x_3 + \frac{1}{2}x_5, \\
6 &= x_4 + 3x_3 - x_5.
\end{aligned}$$

The reduced cost of x_3 is negative which, as noted earlier, indicates that it might be possible to improve x^B (getting a better solution to our problem than x^B) by letting x_3 enter the basis. To find the leaving basic variable, we need to figure out what limits an increase in the value of x_3 from its current value of zero.

We have $\bar{a}^B_{1,3} = 1 > 0$ and $\bar{a}^B_{3,3} = 3 > 0$. In other words, if we increase the value of x_3, and modify the values of x^B_1 and x^B_4 accordingly so that the above equalities are satisfied, we will have to decrease the values of the basic variables x_1 and x_4. Since $\frac{x^B_1}{\bar{a}_{1,3}} = 4 > \frac{x^B_4}{\bar{a}_{3,3}} = 2$, it follows that x^B_4 reaches the value zero first, and that the largest possible value for x_3 is $\delta^* = 2$. Therefore the variable $j(3) = 4$ becomes non-basic. The new basis is therefore $B' = \{1, 2, 3\}$, which we know is optimal. Alternatively, the dual basic solution corresponding to B' can be computed, and it can be verified that $y^{B'}$ is dual feasible.