Chapter 3, Operations Research (OR)

Kent Andersen

February 7, 2007

1 Linear Programs (continued)

In the last chapter, we introduced the general form of a linear program, which we denote (P)

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

and we derived the dual linear program of (P), which we denote (D)

$$\begin{aligned} \text{Maximize } Z_D &= b^T y \\ &\text{s.t.} \qquad \text{(D)} \\ A^T y &\leq c. \end{aligned}$$

We also argued that the following relationships hold between the linear programs (P) and (D)

(a) If (P) is unbounded, then (D) is infeasible.

(b) If (D) is unbounded, then (P) is infeasible.

(c) If x is feasible for (P), and y is feasible for (D), then $b^T y \leq c^T x$ (Weak duality).

(d) If (P) and (D) are both feasible, then there exists a feasible solution x^* to (P), and a feasible solution y^* to (D), such that $b^T y = c^T x$ (Strong duality).

We proved (a)-(c). Furthermore, we mentioned that part (d) can be be proved by using Farkas's Lemma, which we restate below. As last time, let

$$\mathcal{X} := \{x \in \mathbb{R}^n : Ax = b \text{ and } x \geq 0\}.$$

denote the set of feasible solutions to the problem (P). Farkas's Lemma can be stated as follows.

Lemma 1 *(Farkas Lemma): One of the following two conditions hold, but not both.*

(i) $\mathcal{X} \neq \emptyset$*, or in other words, there exists* $x \in \mathbb{R}^n$ *such that* $x \geq 0_n$ *and* $Ax = b$.

(ii) There exists $y \in \mathbb{R}^m$ *such that* $A^T y \leq 0$ *and* $b^T y > 0$.

There is another relation between (P) and (D), which we now prove using Farkas's Lemma.

Lemma 2 *If (P) is infeasible, then (D) is either unbounded or infeasible.*

Proof: If (P) is infeasible, then it follows from Farkas's Lemma that property (ii) of Farkas's Lemma holds. In other words, there exists $y^r \in \mathbb{R}^m$ such that $A^T y^r \leq 0$ and $b^T y^r > 0$. If (D) is also infeasible, the lemma is clearly true, so we can assume (D) is feasible. Let $\bar{y}$ be an arbitrary feasible solution to (D). We have $A^T \bar{y} \leq c$. Consider the line starting at $\bar{y}$ in the direction y^r, or in other words, the set of points of the form $\bar{y} + \alpha y^r$ for some $\alpha \geq 0$. We claim that any point on this line is feasible for (D). Indeed we have $A^T(\bar{y} + \alpha y^r) = A^T \bar{y} + \alpha A^T y^r \leq c + 0$ (since $A^T \bar{y} \leq c$ and $A^T y^r \leq 0$. Hence every point on the line starting at $\bar{y}$ in the direction y^r is feasible for (D). Finally, since $b^T(\bar{y}+\alpha y^r) = b^T \bar{y} + \alpha(b^T y^r)$ and $b^T y^r > 0$, it follows that the objective value for (D) of the points on this line become larger and larger as α increases. Therefore (D) is unbounded.

■

By using basically the same proof one can also prove:

(e) If (D) is infeasible, then (P) is either unbounded or infeasible.

2 Bases, basic solutions and the simplex method

In preparing for the simplex method, we now introduce the concept of a *basic solution* of a linear program. Consider the equality constraints involved in defining the problem (P)

$$Ax = b,$$

and suppose this system involves more variables than equalities, and that all rows of the matrix A are linearly independent. In other words assume $n \geq m$ and $rank(A) = m$ (recall that A is an $m \times n$ matrix). The intuition of a basis, and a basic solution, is to solve the system $Ax = b$ in terms of a subset of the variables $x_1, x_2, \ldots, x_n$. In other words, the idea is to use the equalities $Ax = b$ to express some of the variables in terms of the remaining variables.

Definition 1 *Let $\{1, 2, \ldots, n\}$ index the variables in the problem (P). A* basis *for the linear program (P) is a subset $B \subseteq \{1, 2, \ldots, n\}$ of the variables such that*

(i) B has size m - that is $|B| = m$.

(ii) The columns of A corresponding to variables in B are linearly independent. In other words, if we let A_B denote the (column induced) submatrix of A indexed by B, then $\text{rank}(A_B) = m$.

The variables in B are called basic *variables, and the remaining variables (indexed by $N := \{1, 2, \ldots, n\} \setminus B$) are called* non-basic *variables.*

In the following, we let A_S denote the (column induced) submatrix of A induced by the set S, where S is a subset of the variables. The system $Ax = b$ can be "solved" so as to express the variables in B in terms of the variables in N by multiplying the equality system $Ax = b$ with A_B^{-1} on both sides. Also observe that, for a basis B of (P), the solution to the following system is unique

$$Ax = b, \quad (1)$$

$$x_j = 0 \quad \text{for every } j \in N. \quad (2)$$

The unique solution $x^B \in \mathbb{R}^n$ to the system (1)-(2) is called the *basic solution* corresponding to the basis B. Observe that a basic solution is obtained by setting a subset of the non-negativity constraints $x_j \geq 0$ that are involved in defining the problem (P) to equalities ($x_j = 0$). Hence, if a basic solution x^B satisfies $x_j^B \geq 0$ for all $j \in B$, then x^B is a feasible solution to (P). This motivates the following definition.

Definition 2 *Let $B \subseteq \{1, 2, \ldots, n\}$ be a basis for (P). If $x_j^B \geq 0$ for all $j \in B$, then B is called a* feasible basis, *and x^B is called a* basic feasible solution. *Otherwise B is called an* infeasible basis *for (P), and x^B is called a* basic infeasible solution.

Example 1 *Consider, again, the linear program that models the WGC problem. We consider the version of the problem, where slack variables x_3, x_4 and x_5 have been introduced in the constraints (see example 3 of Chapter 2).*

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

Here we have 5 variables, and since the coefficient matrix of the corresponding system "$Ax = b$" includes the identity matrix (the coefficients on the slack variables), the rank of A is 3 for this example. Hence every basis for this problem consists of 3 variables. The set $B^1 = \{1, 2, 4\}$ constitutes a basis, since the solution to the system $x_1+x_3 = 4$, $2x_2+x_4 = 12$, $3x_1+2x_2+x_5 = 18$, $x_3 = 0$ and $x_5 = 0$ is unique. The corresponding basic solution is given by $(x_1, x_2, x_3, x_4, x_5) = (4, 3, 0, 6, 0)$ (see Fig. 1), and this basis is therefore a feasible basis. Observe that, when a slack variable is non-basic, then the corresponding inequality is satisfied with equality by the basic solution.

The set $B^2 = \{1, 3, 4\}$ is also a basis for the WGC problem. The corresponding basic solution is given by $(x_1, x_2, x_3, x_4, x_5) = (6, 0, -2, 12, 0)$. Since x_3 has a negative value in this basic solution, B^2 is an infeasible basis for (P).

■

The simplex method for solving linear programs is motivated by the following fact (which we will not prove - at least today).

Theorem 1 *If the linear program (P) is feasible and bounded, then then there exists a feasible basis B for (P), such that corresponding basic solution x^B is an optimal solution to (P).*

The simplex method can now be described, informally, as follows.

Step 1: Find a feasible basis B for (P).

Step 2: If x^B is optimal - STOP.

Step 3: Find another feasible basis B' of (P) that is "better" than B. Return to Step 2.

The above steps completely describe the simplex method. However, several issues are not clear. How do we find a feasible basis for (P)? How do we verify that a given basic solution is also optimal? If we can answer this last question, and we conclude that the basis is *not* optimal, how do we find a "better" basis, and how is "better" defined?

Although we will not go into the details at this point, we can reveal the stucture of the "new" basis B' obtained in Step 3 from the "old" basis B. The basis B' is obtained from B by simply *interchanging* a basic variable $i \in B$ with a non-basic variable $j \in \{1, 2, \ldots, n\} \setminus B$. In other words, the new basis B' is of the form $B' = (B \setminus \{i\}) \cup \{j\}$, where i is basic and j is non-basic. This update of the basis is called a *pivot*, and the bases B and B' are said to be *adjacent*.

Example 2 *Consider, again, the WGC problem as defined in Example 1. It is easy to check that the sets $B^1 := \{3, 4, 5\}$, $B^2 := \{2, 3, 5\}$ and $B^3 := \{1, 2, 3\}$ all define feasible bases for the WGC problem. Furthermore, these bases define a sequence of adjacent bases, and the basic solution corresponding to B^3 has $(x_1, x_2) = (2, 6)$, which was demonstrated to be optimal in Chapter 1. Hence, the simplex method can solve the WGC problem with two pivots, if the sequence of bases is chosen to be B^1, B^2, B^3.*

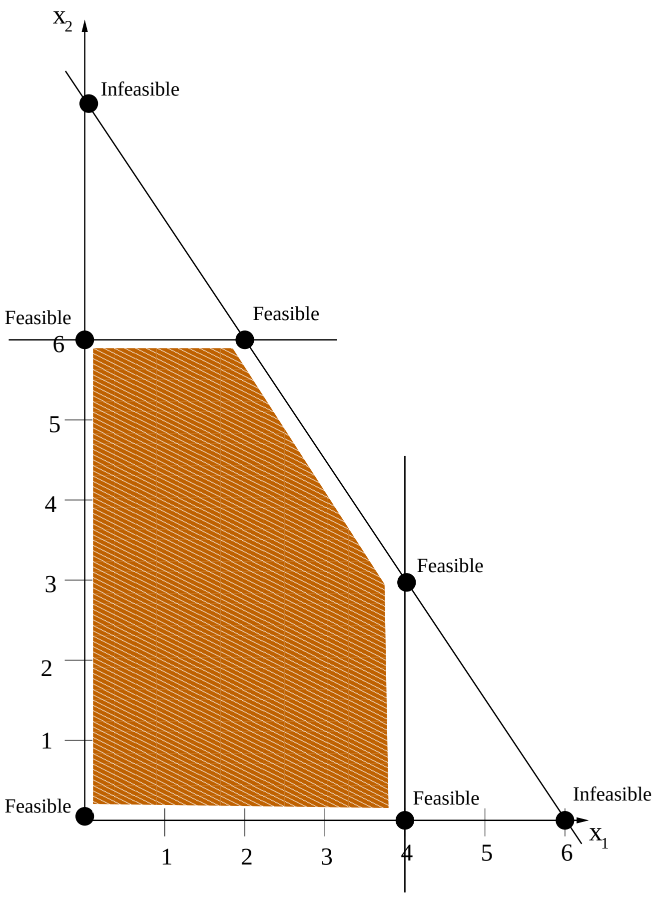

Figure 1: Basic feasible/infeasible solutions for the WGC problem

3 Finding a feasible basis: A two phase approach

In the simplex algorithm described above, it is not clear how to start the algorithm. In other words, it is not clear how to find an initial feasible basis B for (P). We now describe an approach for generating such a feasible basis B for (P). Consider the following linear program (P')

$$\begin{aligned} \text{Minimize } Z^t = &\sum_{i=1}^{m}(t_i^+ + t_i^-) \\ \text{s.t.} & \qquad\qquad \text{(P')} \\ Ax + t^+ - t^- &= b, \\ x &\geq 0_n, \\ t^+, t^- &\geq 0_m. \end{aligned}$$

Observe that the equalities "$Ax + t^+ - t^- = b$" involved in (P') are the same as equalities "$Ax = b$" involved in (P), except that we have added the extra term $(t^+ - t^-)$. Also note that the problem (P') is always feasible (solving the system $t^+ - t^- = b$, $t^+ \geq 0_m$ and $t^- \geq 0_m$ gives a feasible solution with $x = 0_n$). Finally note that (P') is bounded (the objective value is bounded by zero). The variables t^+ and t^- are called *artificial variables*.

Given a feasible solution (x, t^+, t^-) to (P'), we have that x is feasible for (P) if and only if $t^+ = t^- = 0_m$. Hence (P) is feasible if and only if the optimal objective value to (P') is zero. Furthermore, an optimal basis for (P') in which *all* artificial variables are non-basic provides a basic feasible solution to (P). The problem (P') is called the *phase 1* problem associated with (P).

A basic feasible solution to (P') is directly available as follows: for every $i \in \{1, 2, \ldots, m\}$ such that $b_i < 0$, declare t_i^- to be basic, and for every $i \in \{1, 2, \ldots, m\}$ such that $b_i \geq 0$, declare t_i^+ to be basic. Clearly this gives a basic feasible solution to (P'). Given that steps 2 and 3 of the simplex method can be performed (to be discussed later), the simplex method can solve the problem (P') starting from this basis.

If, after solving the phase 1 problem (P'), it is impossible to eliminate the artificial variables from the basis, the problem (P) is infeasible and therefore solved. Otherwise, a basic feasible solution for (P) is available, and the simplex algoritm can be started from this basic feasible solution (this second step is also called the *phase 2* problem).

4 Dual basic solutions and optimality

Consider a basis $B \subseteq \{1, 2, \ldots, n\}$ for the problem (P). We now give an interpretation of the basis B for (P) in the dual linear program (D) of (P). This interpretation provides the key for determinating whether or not a feasible basis for (P) is also optimal for (P) (step 2 of the simplex algorithm). Also, the concept of a dual basic solution provides the key for determinating a "better" basis, if the current basic solution is not optimal (step 3 of the simplex algorithm). Consider, again, the dual (D) of (P)

$$\begin{aligned} \text{Maximize } Z_D = b^T y& \\ \text{s.t.}& \qquad\qquad \text{(D)} \\ A^T y \leq c.& \end{aligned}$$

For every variable $j \in \{1, 2, \ldots, n\}$, let $a_{.j}$ denote the j^{th} column of the matrix A. Since B is a basis for (P), the columns $a_{.j}$ for $j \in B$ are linearly independent. The constraints $A^T y \leq c$ of (D) can be written as

$$(a_{.j})^T y \leq c_j \text{ for all } j \in \{1, 2, \ldots, n\}. \tag{3}$$

The *dual basic solution* $y^B \in \mathbb{R}^m$ associated with the basis B of (P) is defined to be the unique solution to the system

$$(a_{.j})^T y = c_j \text{ for all } j \in B.$$

Observe that the dual basic solution y^B is obtained from the subset of the constraints (3) of (D) indexed by B (by "setting" these inequalities to equalities). The basis B is called *dual feasible*, if y^B satisfies the remaining constraints of (D), or in other words, if $(a_{.j})^T y^B \leq c_j$ for every $j \in \{1, 2, \ldots, n\} \setminus B$. Otherwise B is called a *dual infeasible* basis. The following lemma shows that the problem (P) can be solved by finding a basis B for (P), which is feasible for both the linear program (P), and the linear program (D).

Lemma 3 *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).*

Proof: Since B is a feasible basis for both (P) and (D), x^B is feasible for (P), and y^B is feasible for (D) (by definition). Hence, we only need to verify optimality. From weak duality (item (c) on page 1), this is equivalent to showing $c^T x^B = b^T y^B$. We have

$$\begin{aligned}
c^T x^B &= \\
\sum_{j \in B} c_j x_j^B &= \\
\sum_{j \in B} ((a_{.j})^T y^B) x_j^B &= \qquad \text{(since } (a_{.j})^T y^B = c_j \text{ for } j \in B) \\
(y^B)^T \sum_{j \in B} a_{.j} x_j^B &= \\
(y^B)^T b & \qquad \text{(since } \sum_{j \in B} a_{.j} x_j^B = b).
\end{aligned}$$

■

Lemma 3 provides the desired certificate for optimality of a basis B of (P). Observe that we have the following consequences of Lemma 3.

(i) If B is a feasible basis for (P), and B is *not* an optimal basis for (P), then B is an infeasible basis for (D). In other words, there exists a variable x_j of (P), where $j \in \{1, 2, \ldots, n\} \setminus B$, such that $c_j - (a_{.j})^T y^B < 0$ (the j^{th} constraint of (D) is violated by y^B).

(ii) If B is a feasible basis for (P), and the basis B satisfies $c_j - (a_{.j})^T y^B \geq 0$ for all $j \in \{1, 2, \ldots, n\}$, then B is an optimal basis for the problem (P) (the basis B is feasible for both (P) and (D).

Hence, to check whether or not a feasible basis B for (P) provides an optimal solution to (P) (step 2 of the simplex algoritm/method), we can simply compute y^B (solving a non-singular system), and then check whether y^B satisfies $c_j - (a_{.j})^T y^B \geq 0$ for all $j \in \{1, 2, \ldots, n\}$ (check whether B is dual feasible). If the answer is yes, we have solved the problem (P). If the answer is no, we can find some non-basic variable x_k of (P) that satisfies $c_k - (a_{.k})^T y^B < 0$. Such a variable x_k will, as we will see later, be a candidate for entering the basis (step 3 of the simplex method).