

# Linear Programming and the Simplex Method

David Gale

This exposition of linear programming and the simplex method is intended as a companion piece to the article in this issue on the life and work of George B. Dantzig in which the impact and significance of this particular achievement are described. It is now nearly sixty years since Dantzig's original discovery [3] opened up this whole new area of mathematics. The subject is now widely taught throughout the world at the level of an advanced undergraduate course. The pages to follow are an attempt at a capsule presentation of the material that might be covered in three or four lectures in such a course.

## Linear Programming

The subject of linear programming can be defined quite concisely. It is concerned with the problem of maximizing or minimizing a linear function whose variables are required to satisfy a system of linear *constraints*, a constraint being a linear equation or inequality. The subject might more appropriately be called linear optimization. Problems of this sort come up in a natural and quite elementary way in many contexts but especially in problems of economic planning. Here are two popular examples.

## The Diet Problem

A list of foods is given and the object is to prescribe amounts of each food so as to provide a meal that has preassigned amounts of various nutrients such as calories, vitamins, proteins, starch,

etc. Further, each food has a cost, so among all such adequate meals the problem is to find one that is least costly.

## The Transportation Problem

A certain good, say steel, is available in given amounts at a set of  $m$  origins and is demanded in specified amounts at a set of  $n$  destinations. Corresponding to each origin  $i$  and destination  $j$  there is the cost of shipping one unit of steel from  $i$  to  $j$ . Find a shipping schedule that satisfies the given demands from the given supplies at minimum cost.

Almost all linear programming applications, including these examples, can be motivated in the following way. There is a given set of *goods*. For the diet problem these are the various nutrients. For the transportation problem there are  $m + n$  goods, these being steel at each of the origins and destinations. There is also a set of processes or *activities* which are specified by amounts of the goods *consumed* as *inputs* or *produced* as *outputs*. In the diet problem, for example, the carrot-consuming activity  $\mathbf{c}$  has an output of  $c_1$  units of calories,  $c_2$  units of vitamin A, etc. For the transportation problem the transportation activity  $\mathbf{t}^{ij}$  has as input one unit of steel at origin  $i$  and as output one unit of steel at destination  $j$ . In general, an activity  $\mathbf{a}$  is column vector whose positive entries are outputs, negative entries inputs. These vectors will be denoted by boldface letters. It is assumed that activity  $\mathbf{a}^j$  may be carried out at any nonnegative level  $x_j$ . The constraints are given by another activity vector  $\mathbf{b}$ , the *right-hand side*, which specifies the amounts of the different goods that are to be produced or consumed. For the diet problem it is the list of amounts of the prescribed

---

David Gale is professor of mathematics at the University of California, Berkeley. His email address is gale@math.berkeley.edu.

nutrients. For the transportation problem it is the given supplies and demands at the origins and destinations. Finally, associated with each activity  $\mathbf{a}^j$  is a cost  $c_j$ . Given  $m$  goods and  $n$  activities  $\mathbf{a}^j$  the linear programming problem (LP) is then to find activity levels  $x_j$  that satisfy the constraints and minimize the total cost  $\sum_j c_j x_j$ . Alternatively,  $c$  may be thought of as the profit generated by activity  $\mathbf{a}$ , in which case the problem is to maximize rather than minimize  $\sum_j c_j x_j$ .

The simplex method is an algorithm that finds solutions of LPs or shows that none exist. In the exposition to follow we will treat only the special case where the constraints are equations and the variables are nonnegative, but the more general cases are easily reduced to this case.

### The Simplex Method

In the following paragraphs we describe the simplex algorithm by showing how it can be thought of as a substantial generalization of standard Gauss-Jordan elimination of ordinary linear algebra. To gain intuition as to why the algorithm works, we will refer to the linear activity model of the previous section. Finally we will mention some interesting discoveries in the analysis of the algorithm, including a tantalizing unsolved problem.

### Pivots and Tableaus

The basic computational step in the simplex algorithm is the same as that in most of elementary linear algebra, the so-called *pivot* operation. This is the operation on matrices used to solve systems of linear equations, to put matrices in echelon form, to evaluate determinants, etc.

Given a matrix  $A$  one chooses a nonzero *pivot entry*  $a_{ij}$  and adds multiples of row  $i$  to the other rows so as to obtain zeros in the  $j$ th column. The  $i$ th row is then normalized by dividing it by  $a_{ij}$ . For solving linear equations a pivot element can be any nonzero entry. By contrast, the simplex method restricts the choice of pivot entry and is completely described by giving a pair of simple rules, the entrance rule that determines the pivot column  $j$  and the exit rule that determines the pivot row  $i$  (in theory a third rule may be needed to take care of degenerate cases). By following these rules starting from the initial data the algorithm arrives at the solution of the linear program in a finite number of pivots. Our purpose here is to present these rules and show why they work.

We shall need one other concept.

**Definition.** The *tableau*  $X$  of a set of vectors  $A = \{\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^n\}$  with respect to a basis  $B = \{\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^m\}$  is the  $m \times n$  matrix

	$\mathbf{a}^1$	$\mathbf{a}^2$	$\dots$	$\mathbf{a}^j$	$\dots$	$\mathbf{a}^n$
$\mathbf{b}^1$	$x_{11}$	$x_{12}$	$\dots$	$x_{1j}$	$\dots$	$x_{1n}$
$\mathbf{b}^2$	$x_{21}$	$x_{22}$	$\dots$	$x_{2j}$	$\dots$	$x_{2n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{b}^i$	$x_{i1}$	$x_{i2}$	$\dots$	$x_{ij}$	$\dots$	$x_{in}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{b}^m$	$x_{m1}$	$x_{m2}$	$\dots$	$x_{mj}$	$\dots$	$x_{mn}$

where  $x_{ij}$  is the coefficient of  $\mathbf{b}^i$  in the expression for  $\mathbf{a}^j$  as a linear combination of the  $\mathbf{b}^i$ .

In matrix terms,  $X$  is the (unique) matrix satisfying

$$(1) \quad BX = A.$$

(Symbols  $A, B$  are used ambiguously to stand either for a matrix or the set of its columns.)

It will often be useful to include the unit vectors  $\mathbf{e}^i$  in the tableau in which case it appears as shown below.

	$\mathbf{e}^1$	$\mathbf{e}^2$	$\dots$	$\mathbf{e}^j$	$\dots$	$\mathbf{e}^m$
$\mathbf{b}^1$	$y_{11}$	$y_{12}$	$\dots$	$y_{1j}$	$\dots$	$y_{1m}$
$\mathbf{b}^2$	$y_{21}$	$y_{22}$	$\dots$	$y_{2j}$	$\dots$	$y_{2m}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{b}^i$	$y_{i1}$	$y_{i2}$	$\dots$	$y_{ij}$	$\dots$	$y_{im}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{b}^m$	$y_{m1}$	$y_{m2}$	$\dots$	$y_{mj}$	$\dots$	$y_{mm}$

	$\mathbf{a}^1$	$\mathbf{a}^2$	$\dots$	$\mathbf{a}^j$	$\dots$	$\mathbf{a}^n$
	$x_{11}$	$x_{12}$	$\dots$	$x_{1j}$	$\dots$	$x_{1n}$
	$x_{21}$	$x_{22}$	$\dots$	$x_{2j}$	$\dots$	$x_{2n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$x_{i1}$	$x_{i2}$	$\dots$	$x_{ij}$	$\dots$	$x_{in}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$x_{m1}$	$x_{m2}$	$\dots$	$x_{mj}$	$\dots$	$x_{mn}$

Note that from (1),  $Y = \{y_{ij}\}$  is the solution of  $BY = I$  so  $Y = B^{-1}$ . Multiplying (1) on the left by  $Y$  gives the important equation

$$(2) \quad YA = X.$$

It is easy to verify that if one pivots on entry  $x_{ij}$ , one obtains a new matrix  $X'$  that is the tableau with respect to the new basis in which  $\mathbf{a}^j$  has replaced  $\mathbf{b}^i$ .

	$\mathbf{a}^1$	$\mathbf{a}^2$	$\dots$	$\mathbf{a}^j$	$\dots$	$\mathbf{a}^n$
$\mathbf{b}^1$	$x'_{11}$	$x'_{12}$	$\dots$	0	$\dots$	$x'_{1n}$
$\mathbf{b}^2$	$x'_{21}$	$x'_{22}$	$\dots$	0	$\dots$	$x'_{2n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{a}^j$	$x'_{i1}$	$x'_{i2}$	$\dots$	1	$\dots$	$x'_{in}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{b}^m$	$x'_{m1}$	$x'_{m2}$	$\dots$	0	$\dots$	$x'_{mn}$

The simplex method solves linear programs by a sequence of pivots in successive tableaus, or, equivalently, by finding a sequence of bases, where each basis differs from its predecessor by a single vector.

### Solving Linear Equations

We start by showing how to solve systems of linear equations using the language of pivots and tableaus. Our fundamental result is the following: **Theorem 1.** *Exactly one of the following systems has a solution:*

$$(3) \quad \mathbf{Ax} = \mathbf{b}$$

or

$$(4) \quad \mathbf{y}^T \mathbf{A} = \mathbf{0}^T, \mathbf{y}^T \mathbf{b} = 1.$$

(Again, boldface letters are column vectors and are converted to row vectors using the superscript  $T$ .)

Note that rather than giving complicated conditions for (3) not to be solvable this formulation puts the solvable and unsolvable case on an equal footing. This *duality* is an essential element of the theory, as will be seen.

### Geometric Interpretation

If (3) has no solution this means that  $\mathbf{b}$  does not lie in the linear subspace generated by the  $\mathbf{a}^j$ . In that case there is a vector  $\mathbf{y}$  orthogonal to the  $\mathbf{a}^j$  but not to  $\mathbf{b}$ .

We shall continue to present such geometric pictures as aids to intuition but the mathematical arguments will not depend on these “pictures”.

It is immediate that (3) and (4) cannot both hold since multiplying (3) by  $\mathbf{y}^T$  and (4) by  $\mathbf{x}$  would imply  $0 = 1$ . The nontrivial part is to show that either (3) or (4) must hold. We do this by giving an algorithm that in at most  $m$  pivots finds a solution to either (3) or (4).

The initial tableau consists of the matrix  $A$  together with the right-hand side  $\mathbf{b}$  all preceded by the identity matrix  $I$ . It is represented schematically in the form below.

$$\begin{array}{ccc|ccc} \{\mathbf{e}^i\} & \{\mathbf{a}^j\} & \{\mathbf{b}\} & & & \\ \{\mathbf{b}^i\} & [I & A & \mathbf{b}] & & \end{array}$$

(The symbols like  $\{\mathbf{e}^i\}$ ,  $\{\mathbf{a}^j\}$  in curly brackets are the row and column headings of the tableau.)

We proceed to try replacing the  $\mathbf{e}^i$  in order by some  $\mathbf{a}^j$ . The tableau at any stage has the form

$$\begin{array}{ccc|ccc} \{\mathbf{e}^i\} & \{\mathbf{a}^j\} & \{\mathbf{b}\} & & & \\ \{\mathbf{b}^i\} & [Y & X & \mathbf{u}] & & \end{array}$$

There are now two possibilities.

Case I. All of the unit vectors  $\mathbf{e}^i$  can be replaced by some of the  $\mathbf{a}^j$ . Then a solution of (3) can be read off from the vector  $\mathbf{u}$  in the  $\mathbf{b}$ -column of the tableau above. Namely, each component of  $\mathbf{u}$ , say

$u_i$ , is the value of a variable  $x_{j_i}$  whose corresponding column  $\mathbf{a}^{j_i} = \mathbf{b}^i$ . Every  $x_j$  whose corresponding column is not in the current basis has the value zero.

Case II. In trying to replace  $\mathbf{e}^k$  there is no available pivot because  $x_{k_j} = 0$  for all  $j$ .

(i) If in addition  $u_k = 0$ , then leave  $\mathbf{e}^k$  in the basis and go on to replace  $\mathbf{e}^{k+1}$ .

(ii) If  $u_k \neq 0$ , then the  $k$ th row  $\mathbf{y}_k$  of  $Y$  gives a solution of (4). To see this, note that from (2)  $[Y][A \ \mathbf{b}] = [X \ \mathbf{u}]$ . Since the  $k$ th row of  $X$  is zero, we get  $\mathbf{y}_k A = 0$ , whereas  $\mathbf{y}_k \mathbf{b} = u_k \neq 0$ , so  $\mathbf{y}_k / u_k$  solves (4).

The algebra becomes considerably more complicated if one requires the solutions of (3) to be nonnegative. In this case we have the following existence theorem (known as Farkas's Lemma):

**Theorem 2.** *Exactly one of the following systems has a solution.*

$$(5) \quad \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$$

or

$$(6) \quad \mathbf{y}^T \mathbf{A} \leq \mathbf{0}^T, \mathbf{y}^T \mathbf{b} > 0.$$

Again it is not possible for both systems to have solutions since multiplying (5) by  $\mathbf{y}^T$  gives  $\mathbf{y}^T \mathbf{Ax} = \mathbf{y}^T \mathbf{b} > 0$  whereas multiplying (6) by  $\mathbf{x}$  gives  $\mathbf{y}^T \mathbf{Ax} \leq 0$ .

### Geometric Interpretation

Equation (5) says that  $\mathbf{b}$  lies in the convex cone generated by the columns  $\mathbf{a}^j$  of  $A$ . If  $\mathbf{b}$  does not lie in this cone, then there is a “separating hyperplane” whose normal makes a non-acute angle with the  $\mathbf{a}^j$  and an acute angle with the vector  $\mathbf{b}$ .

The simplex method finds a solution of either (5) or (6) in a finite number of pivots. However, there is no useful upper bound on the number of pivots that may be needed as we shall see shortly.

We may assume to start out that the vector  $\mathbf{b}$  is nonnegative (if not, change the signs of some of the equations).

**Definition.** A basis  $B$  will be called *feasible (strongly feasible)* if  $\mathbf{b}$  is a nonnegative (positive) linear combination of vectors of  $B$ . This is equivalent to the condition that the  $\mathbf{b}$  column  $\mathbf{u}$  of the tableau is nonnegative (positive).

Imitating the algorithm of the previous section we start from the feasible basis  $\{\mathbf{e}^i\}$  of unit vectors and try to bring in the  $\{\mathbf{a}^j\}$  by a sequence of pivots that maintain feasibility, but to do this the pivot rule of the previous section must be restricted.

Suppose  $\mathbf{a}^j$  is to be brought into a new feasible basis. Then in order for the new basis to be feasible the basis vector it replaces must satisfy the following easily derived condition.

### Exit Pivot Rule

The pivot element  $x_{ij}$  must satisfy

- (i)  $x_{ij} > 0$ ,
- (ii)  $u_i/x_{ij} \leq u_k/x_{kj}$  for all  $k$  with  $x_{kj} > 0$ .

Because of the above restriction it is no longer possible to simply replace the basis vectors  $e^i$  one at a time as in the previous section. Some further rule for choosing the pivot column must be given.

We will postpone the proof of Theorem 2 as it will be shown to be a special case of a much more general theorem, which we now describe.

### The Standard Minimum Problem

We consider the following important special case of a linear programming problem.

Given an  $m$ -vector  $\mathbf{b}$ , an  $n$ -vector  $\mathbf{c}^T$  and an  $m \times n$  matrix  $A$ , find an  $n$ -vector  $\mathbf{x}$  so as to

$$(7) \quad \begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{array}$$

Corresponding to this problem (and using exactly the same data) is a *dual problem*, find an  $m$ -vector  $\mathbf{y}^T$  so as to

$$(8) \quad \begin{array}{ll} \text{maximize} & \mathbf{y}^T \mathbf{b} \\ \text{subject to} & \mathbf{y}^T A \leq \mathbf{c}^T. \end{array}$$

### Terminology

The linear function  $\mathbf{c}^T \mathbf{x}$  and  $\mathbf{y}^T \mathbf{b}$  are called the *objective functions* of (7) and (8), respectively. The vectors  $\mathbf{x}, \mathbf{y}$  that solve these problems are called *optimal solutions*; the numbers  $\mathbf{c}^T \mathbf{x}$  and  $\mathbf{y}^T \mathbf{b}$  are *optimal values*.

Let  $\mathbb{X}$  and  $\mathbb{Y}$  be the sets (possibly empty) of all vectors that satisfy the constraints of (7) and (8), respectively. Such vectors are said to be *feasible*.

### Key Observation

If  $\mathbf{x} \in \mathbb{X}$ ,  $\mathbf{y}^T \in \mathbb{Y}$ , then  $\mathbf{y}^T \mathbf{b} \leq \mathbf{c}^T \mathbf{x}$ .

**Proof.** Multiply  $A\mathbf{x} = \mathbf{b}$  by  $\mathbf{y}^T$  and  $\mathbf{y}^T A \leq \mathbf{c}^T$  by  $\mathbf{x}$ .

An important consequence of this observation is

**Corollary.** If  $\mathbf{x} \in \mathbb{X}$  and  $\mathbf{y} \in \mathbb{Y}$  satisfy  $\mathbf{y}^T \mathbf{b} = \mathbf{c}^T \mathbf{x}$ , then  $\mathbf{x}$  and  $\mathbf{y}$  are optimal solutions, and the numbers  $\mathbf{y}^T \mathbf{b} = \mathbf{c}^T \mathbf{x}$  are optimal values for their respective problems.

The converse of this fact makes up the **Fundamental Duality Theorem**. The dual problems above have optimal solutions  $\mathbf{x}, \mathbf{y}^T$  if and only if  $\mathbb{X}$  and  $\mathbb{Y}$  are nonempty in which case the optimal values  $\mathbf{c}^T \mathbf{x}$  and  $\mathbf{y}^T \mathbf{b}$  are equal.

### Historical Note

The first explicit statement of the duality theorem is due to von Neumann [7] in a manuscript that was privately circulated but never formally published in his lifetime. Moreover it has been hard to

verify the validity of the von Neumann proof. The first formally published proof is due to Gale, Kuh, and Tucker [4].

To see how Theorem 2 is a special case of the duality theorem we take  $[I \ A \ \mathbf{b}]$  as the given initial tableau of a standard minimum problem and the vector  $\mathbf{c}^T = (1, \dots, 1, 0, \dots, 0)$  whose first  $m$  entries are 1, the rest 0, as the objective function. The set  $\mathbb{X}$  is nonempty since it contains the nonnegative vector  $(b_1, \dots, b_m, 0, \dots, 0)$ , and the set  $\mathbb{Y}$  is nonempty since it contains the zero vector. If the optimal value is zero, then a subvector of the optimal solution  $\mathbf{x}$  solves (5). If the optimal value is positive, then the dual optimal vector  $\mathbf{y}$  solves (6).

We will now see how the simplex method solves problems (7), (8) and gives a constructive proof of the duality theorem.

As a first step we incorporate the objective function  $\mathbf{c}^T \mathbf{x}$  as the 0th row of the tableau. That is, we define the *augmented* column vector  $\hat{\mathbf{a}}^j$  to be  $(-c_j, \mathbf{a}^j)$ , and we introduce the 0th unit vector  $\mathbf{e}^0$ . The *augmented* initial tableau is then,

$$\begin{array}{cccc} \mathbf{e}^0 & \{\mathbf{e}^i\} & \{\hat{\mathbf{a}}^j\} & \hat{\mathbf{b}} \\ \mathbf{e}^0 & \left[ \begin{array}{ccc|c} 1 & 0 & -\mathbf{c}^T & 0 \\ 0 & I & A & \mathbf{b} \end{array} \right] \end{array}$$

We may assume that we have found an initial feasible basis, by applying the simplex method to the nonnegative solution problem (5), (6) as described. Then with respect to the general augmented basis  $\{\mathbf{e}^0, \hat{\mathbf{b}}^1, \dots, \hat{\mathbf{b}}^m\}$  the tableau has the form below.

$$\begin{array}{cccc} \mathbf{e}^0 & \{\mathbf{e}^i\} & \{\hat{\mathbf{a}}^j\} & \hat{\mathbf{b}} \\ \mathbf{e}^0 & \left[ \begin{array}{ccc|c} 1 & \mathbf{y}^T & \mathbf{z}^T & w \\ 0 & Y & X & \mathbf{u} \end{array} \right] \end{array}$$

Note that by definition of the tableau,  $w\mathbf{e}^0 + \sum_i u_i \hat{\mathbf{b}}^i = \hat{\mathbf{b}}$  so for the 0th entry we have  $w - \sum_i u_i c_i = 0$ , so  $w = \sum_i u_i c_i$ , which is the value of the objective function for the basic feasible basis  $\{\mathbf{b}^i\}$ .

Also from the tableau  $z_j = \sum_i x_{ij} c_i - c_j$ . This number compares the cost associated with activity vector  $\mathbf{a}^j$  to that of the corresponding linear combination of the basis vectors  $\mathbf{b}^i$ . If  $z_j$  is positive it means that one could reduce the total cost by bringing the vector  $\mathbf{a}^j$  into the next basis. This leads to the following two important observations:

I. If the row  $\mathbf{z}^T$  is nonpositive, then  $\mathbf{u}$  can be extended to a solution  $\mathbf{x}$  of (7) and  $\mathbf{y}^T$  solves (8). Namely, using (2) again we have

$$\left[ \begin{array}{ccc|c} 1 & \mathbf{y}^T & \mathbf{z}^T & w \\ 0 & Y & X & \mathbf{u} \end{array} \right] \left[ \begin{array}{ccc|c} -\mathbf{c}^T & 0 & & \\ A & \mathbf{b} & & \end{array} \right] = \left[ \begin{array}{cc|c} \mathbf{z}^T & w & \\ X & \mathbf{u} & \end{array} \right]$$

so  $-c_j + \mathbf{y}^T \mathbf{a}^j = z_j \leq 0$ . Hence  $\mathbf{y}^T A \leq \mathbf{c}^T$  so  $\mathbf{y}^T$  satisfies the dual constraints. Similarly we have  $\mathbf{y}^T \mathbf{b} = w = \mathbf{c}^T \mathbf{x}$ , so from the corollary above, the

Duality Theorem is verified.

II. If for some  $j$  we have  $z_j > 0$  and  $x_{ij} \leq 0$  for all  $i$ , then problem (7) has no minimum, for from the tableau we have  $\mathbf{a}^j - \sum_i x_{ij} \mathbf{b}^i = \mathbf{0}$ ; and since the  $x_{ij} \leq 0$ , we have

$$\lambda(\hat{\mathbf{a}}^j - \sum_i x_{ij} \hat{\mathbf{b}}^i) + \sum_i x_{ij} \hat{\mathbf{b}}^i = \mathbf{b}$$

giving a new feasible (nonnegative) solution for any  $\lambda > 0$ . The corresponding change in the objective function is  $\lambda(c_j - \sum_i x_{ij} c_i) = -\lambda z_j$ ; so since  $z_j > 0$  the function has no minimum.

In view of I above, one is looking for a feasible solution with  $z$  nonpositive, so the following pivot rule suggests itself.

### Entrance Pivot Rule

*Bring into the basis any column  $\hat{\mathbf{a}}^j$  where  $z_j$  is positive.*

Note that as described the second pivot rule may, in general, require making a choice among many possible positive  $z_j$ . A natural choice would be for example to choose the column with the largest value of  $z_j$ , a choice originally suggested by Dantzig.

The two pivot rules completely describe the simplex algorithm. If for some  $j$ ,  $z_j$  is positive and some  $x_{ij}$  is positive, then bring  $\mathbf{a}^j$  into the basis (by a pivot). It remains to show that eventually either I or II will occur. The argument is simple if we make the following

### Nondegeneracy Assumption

*Every feasible basis is strongly feasible.*

Note if this were not the case, then  $\mathbf{b}$  would be a positive linear combination of fewer than  $m$  of the vectors  $\mathbf{a}^j$ , a degenerate situation. Although the degenerate case would seem (on mathematical grounds) to be rare, this is not so in practice.

Suppose  $x_{ij}$  is the (positive) pivot. Then after pivoting  $w' = w - z_j u_i / x_{ij} < w$  since  $u_i > 0$  from nondegeneracy, so the value of the objective function decreases with each pivot. This means that no basis can recur (since the basis determines the objective value), and since there are only finitely many bases, the algorithm must terminate in either state I or II.

If the nondegeneracy assumption is not satisfied some further argument is necessary. Indeed, examples have been constructed by Hoffman [5] (for one) using the Dantzig pivot choice rule which can lead to "cycling," meaning the sequence of feasible bases recurs indefinitely. It turns out, however, that the following simple rule due to Bland [1] guarantees that no basis will recur.

### Bland's Pivot Selection Rule

Among eligible entering vectors choose the one with lowest index and let it replace the eligible basis vector with lowest index.

While the condition is simple to state, the proof that it avoids cycling is quite subtle. Moreover, it is not as efficient as the customary (Dantzig) pivot choice rule.

Finally, there is a natural economic interpretation of the dual problem. In the model of goods and activities, the vector  $\mathbf{y}^T = (y_1, \dots, y_m)$  can be thought of as a *price vector* where  $y_i$  is the price of one unit of the  $i$ th good. Then  $\mathbf{y}^T \mathbf{a}^j$  is the revenue generated by operating activity  $\mathbf{a}^j$  at unit level. The condition of dual feasibility,  $\mathbf{y}^T \mathbf{a}^j - c_j \leq 0$  states that *profit* (revenue minus cost) cannot be positive. This is an economically natural requirement, for if an activity generated positive profits producers would want to operate it at arbitrarily high levels and this would clearly not be feasible. Duality then says that prices are such that the price vector  $\mathbf{y}^T$  maximizes the value  $\mathbf{y}^T \mathbf{b}$  of the right-hand side activity  $\mathbf{b}$  subject to the condition of no positive profits.

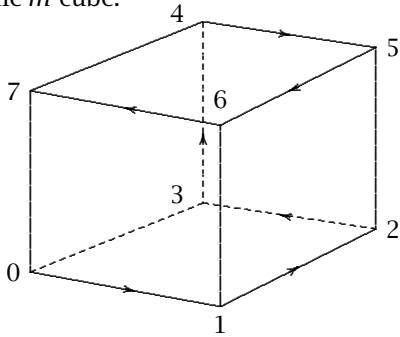
### Some Properties of the Simplex Method

Since its creation by Dantzig in 1947, there has been a huge amount of literature on variations and extensions of the simplex method in many directions, several devised by Dantzig himself. We will mention here only one of these, revolving around the question of the number of pivot steps required to solve an LP in the worst case. Based on a vast amount of empirical experience, it seemed that an  $m \times n$  program was typically solved in roughly  $3m/2$  pivots. However, in 1972, Klee and Minty [6] gave an example of an  $m \times 2m$  standard problem that using the Dantzig choice rule required  $2^m - 1$  pivots. To appreciate the example, first consider the matrix  $A$  consisting of the  $m$  unit vectors  $\mathbf{e}^i$  and  $m$  vectors  $\mathbf{a}^j = \mathbf{e}^j$  where the right-hand side is  $\mathbf{e}$ , the vector all of whose entries are one. By inspection, (see for example the tableau below) the set of feasible solutions  $\mathbb{X}$  decomposes into the direct product of  $m$  unit intervals, thus, it is a unit  $m$ -cube.

$$\begin{array}{ccc} \{\mathbf{e}^i\} & \{\mathbf{a}^j\} & \mathbf{b} \\ \left[ \begin{array}{cccccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right] & \left[ \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right] \end{array}$$

It is easy to see that every feasible basis, thus every vertex of the cube, must contain either  $\mathbf{a}^i$  or  $\mathbf{e}^i$ , but not both, for all  $i$ . This property will be preserved if the  $\mathbf{a}^i$  are slightly perturbed. By a clever choice of this perturbation and a suitable objective function, the authors show that the Dantzig pivot rule causes the algorithm to visit all of the

$2^m$  vertices following a well known Hamiltonian path on the  $m$ -cube.



**Figure 1. A Hamiltonian path on a 3-cube.**

This means, for example, that the vector  $\mathbf{a}^1$  enters the basis on the first pivot, stays in the basis for one pivot, then exits and stays out for one pivot, then reenters and stays, then exits and stays out, etc. throughout the computation.

So we know that in the worst case the simplex method may require an exponential number of pivots, although, as mentioned earlier, no naturally occurring problem has ever exhibited such behavior. There are also results on the expected number of pivots of a “random” LP.

Since the Klee-Minty example shows that it is possible for the simplex algorithm to behave badly, it is natural to ask whether there may be some other pivoting algorithms that are guaranteed to find an optimal solution in some number of pivots bounded, say, by a polynomial in  $n$ . Let us first consider a lower bound on the number of pivots. If the initial feasible basis is  $A = \{\mathbf{a}^1, \dots, \mathbf{a}^m\}$  and the optimal basis is a disjoint set  $B = \{\mathbf{b}^1, \dots, \mathbf{b}^m\}$  then clearly it will require at least  $m$  pivots to get from  $A$  to  $B$ . The very surprising fact is that if the set  $\mathbb{X}$  of feasible solutions is bounded, then in all known examples it is possible to go from any feasible basis to any other in at most  $m$  pivots. Recall that the set  $\mathbb{X}$  is an  $m$ -dimensional convex polytope whose vertices are the feasible bases. The above observation is equivalent to the statement that any two vertices of this polytope are connected by an edge path of at most  $m$  edges. The conjecture originally posed by Warren Hirsch has been proved by Klee and Walkup through dimension 5 but remains unresolved in higher dimensions.

For Dantzig this “Hirsch conjecture” was especially intriguing since a constructive proof of the conjecture would have meant that there might be an algorithm that solves linear programs in at most  $m$  pivots. However, even if such an algorithm existed, it might be that the amount of calculation involved in selecting the sequence of pivots would make it far less computationally efficient than the simplex method. Indeed this remarkable algorithm and its many refinements remains to this

day the method of choice for solving most linear programming problems.

## References

- [1] R. G. BLAND New finite pivoting rules for the simplex method, *Mathematics of Operations Research* 2 (1977) 103–107.
- [2] J. B. J. FOURIER, Solution d’une question particulière du calcul des inégalités (1826), and extracts from Histoire de l’Académie (1823, 1824), *Oeuvres* II, French Academy of Sciences, pp. 317–328.
- [3] G. B. DANTZIG, Maximization of a linear function subject to linear inequalities, in T. C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, John Wiley & Sons, New York, 1951, pp. 339–347.
- [4] D. GALE, H. W. KUHN, and A. W. TUCKER, Linear programming and the theory of games, in T. C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, John Wiley & Sons, New York, 1951, pp. 317–329.
- [5] A. J. HOFFMAN, Cycling in the simplex algorithm, National Bureau of Standards, Report No. 2974, Washington, D.C., December 16, 1953.
- [6] V. KLEE and G. J. MINTY, How good is the simplex algorithm? in (O. Shisha, ed.) *Inequalities III*, New York, Academic Press, 1972, 159–175.
- [7] J. VON NEUMANN, Discussion of a maximum problem, working paper dated November 15–16, 1947, reproduced in A. H. Taub (ed.), *John von Neumann: Collected Works, Vol. VI*, Pergamon Press, Oxford, 1963, pp. 89–95.