RR-76-3

SOME ADDITIONAL VIEWS ON THE SIMPLEX METHOD AND THE GEOMETRY OF CONSTRAINT SPACE

WM. ORCHARD-HAYS

MAY 1976

International Institute for Applied Systems Analysis
2361 Laxenburg, Austria

PREFACE

The field of linear programming (LP) has perhaps the longest history among all modern techniques in the decision sciences, at least if attention is restricted to techniques inherently tied to the computer. For over a quarter century, there have been intensive and extensive developments in theory, generalized systems of computer programs, and in applications. However, these efforts have been carried out by three different classes of specialists whose interaction has at times been minimal.

Except for two or three early conceptual developments, LP originated in practical problems at about the same time as electronic computers became a reality, and the growth of the two has been contemporary. While theory tended to be the domain of the academic world, computerized systems were developed by independent consultant organizations and later computer manufacturers, and experience in applications was gained by large commercial and industrial corporations spearheaded by the petroleum industry. The result is that different conceptual approaches, notations and viewpoints have developed that often inhibit the adoption of existing capabilities by new potential users, particularly in academically-oriented organizations.

The scientific staff at IIASA is more from the academic world than from the consulting and commercial sectors. Consequently, there may be some unfamiliarity with the viewpoints and notations in use by the developers of computer systems for mathematical programming applications. This paper summarizes the notation used over a long period by one of the leading developers of such systems and by many of his associates and even competitors. Further, the mathematical viewpoints are more those of an algorithm and software engineer than of a theoretical mathematician, economist, or academic. These viewpoints are extended to geometrical concepts which may help others to understand the somewhat capricious performance of the simplex method on large problems. Since the various projects which IIASA is or will be engaged in will lead to the formulation and solution of large LP models, some understanding of the viewpoints of builders of elaborate systems of programs should be helpful in applying them successfully.

Table of Contents

ABSTRACT

In Part I, the classical statement of an LP problem is compared with the most general form which general-purpose LP software can usually accept. The latter form is then simplified to the form used internally by such software. An extended matrix representation of the conditions used in the simplex method is given, plus a list of the various outcomes of pivot selection. All this is merely a review and summary in consistent notation.

The remainder of Part I views an LP problem as a function of its objective form and parametric algorithms as families of functions. The simplex method, as a process, is also viewed as following a trajectory. The ambiguity of extending this idea to the dual feasible subspace is indicated as well as the difficulty of using this viewpoint for integer programs.

Part II begins with a fairly complete list of notation required in discussing details of the simplex method and its variants. Then a series of definitions, lemmas and theorems are given to make precise such notions as basic solution, distinct solution, adjacency, and dual basis. The main result is a clarification of the phenomena of degeneracy and alternate solutions, in both primal and dual senses. In particular, the complementary nature of ambiguous solutions and multiple solutions is shown. Two trivial examples, easily followed, are sufficient to illustrate these ideas.

Part III applies the ideas of Part II, plus one other, to the old problems of exploring the vicinity of optimality, resolving revised models from an old basis, and a few special problems for which the simplex method is sometimes useful in a non-LP context.

Some Additional Views on the Simplex Method And the Geometry of Constraint Space

INTRODUCTION

There is a wide discrepancy between the terminology and viewpoints used in classical and theoretical presentations of linear programming (LP) and the simplex method, and those used by software engineers who create and extend the systems of computer programs without which LP would be only an abstraction. On the one hand, the classical presentations are too condensed and over-simplified, ignoring practical aspects of real model formulation and solution. On the other hand, the excruciating details of algebra and logic required to perfect a robust system of programs are too tedious to permit overviews and facile manipulation of concepts through manageable terminology.

A similar situation -- perhaps even more disparate -- exists with respect to geometrical concepts which fortify our intuition and make new ideas and hypotheses possible. Theories of convex sets, simplices, supporting hyperplanes, dual spaces, and the like are essential as a foundation to the whole field of optimization. However, these are specialities for the few and certainly algorithm and software engineers are seldom experts. Furthermore, workers in the field tend to make statements such as "a basic solution represents a vertex of the simplex". This is an acceptable ellipsis among knowledgeable professionals, but, taken literally, it is nonsense--an m x m matrix equation cannot represent a point in E^n. Our concepts of the intricate, interlacing elements in constraint space--even in E^2 or E^3 with linear systems--are often inadequate to conceptualize and sort out the algebraic phenomena which we encounter.

Part I of this paper starts with the typical classical statement of an LP problem together with known results of the simplex method. This is extended in similar notation to the most general set of constraints accepted by standard, large Mathematical Programming Systems (MPS). This, in turn, is simplified by the same preliminary transformations used in MPSs to give a workable but general framework for any LP model. Following this, one primal simplex iteration is described with a summary of the results of typical pivot selection routines, which are not amenable to succinct notation. Up to this point, the paper is simply a review and summary in consistent notation.

The remainder of Part I presents some rather unorthodox viewpoints on the role of the various LP quantities, the nature of the "LP function" and the simplex method machinery, and some preliminary discussion of the geometry of constraint spaces.

Part II presents some theory regarding primal and dual basic solutions and their combined geometry in constraint spaces. The meaning of "representations" is clarified. Several definitions and lemmas, and five theorems, create a succinct and rigorous terminology for discussing movements through areas of E^n which have nonsingular representations.

Part II contains a complete list of notation used in discussing simplex transformations and similar operations. This may be useful in itself. One switch in notation is made at this point: superscripts are used to denote rows or row elements in the basis inverse and the transformed LP matrix. Personally, the writer prefers the use of superscripts for all row indices, and has consistently used such notation for many years. However, it is difficult to fight the tide: everyone writes a_{ij} and x_j instead of a_j^i and x^j. Nevertheless, the use of superscripts for A^i, E^j, α^r and α_s^r seems absolutely necessary for clarity in Parts II and III.

Part III exploits the viewpoint of Part II in three areas. First, the old problem of finding all optimal solutions and their adjacent solutions is solved by means of an unambiguous procedure, which is readily programmable. Second, some suggestions are made to reduce the number of iterations when restarting a revised model from an old optimal basis. Finally, a few special model matrices are discussed, which may have some practical value in special circumstances.

One of the motivations for this paper was to try to find a more rational and elementary approach to integer programming. This has not been achieved and only one short section on the subject has been retained. From one viewpoint, the requirement of integrality superimposes a third set of elements in the geometry namely, either a lattice or a kind of "boxwork" of hyperplanes. But this does not seem to help in finding optimal integer solutions, or at least in proving them so. Perhaps someone will yet conceive of a viewpoint which facilitates this. One expects to have to do substantially more work to solve an integer program but it is frustrating for it to be largely guesswork.

PART I: THE SIMPLEX METHOD; TERMINOLOGY AND VIEWPOINTS

Classical Statement of LP Problem and Simplex Solution

Given:

An m x n (real) matrix A, an m x 1 column of constants b, and a 1 x n row of objective coefficients c.

Find:

An n x 1 column of (structural) variables x such that

$$z = cx \text{ is max subject to}$$
$$x \geq 0 \quad , \qquad Ax \leq b \quad .$$

Simplex solution (assuming the problem is feasible and z_{max} is finite):

$x = \bar{x} \geq 0$ such that $A\bar{x} \leq b$ (primal solution),
1 x m row $\bar{\pi}$ such that $\bar{\pi}A \geq c$ (dual solution),
$z_{max} = c\bar{x} = \bar{\pi}b$.

Consequences and subsidiary quantities:
Let $u = b - Ax$, and $d = \pi A - c$. Then

$$\bar{u} = b - A\,\bar{x} \geq 0 \quad , \qquad \bar{d} = \bar{\pi}A - c \geq 0$$

and

$$z = c\bar{x} = \bar{\pi}A\,\bar{x} - \overline{dx} = \overline{\pi A x} + \overline{\pi u} = \bar{\pi}b \quad .$$

But

$$\overline{dx} \geq 0 \quad , \qquad \overline{\pi u} \geq 0 \quad .$$

Therefore $\overline{dx} = \overline{\pi u} = 0$, i.e. $\bar{d}_j = 0$ if $\bar{x}_j > 0$, $\bar{\pi}_i = 0$ if $\bar{u}_i > 0$, and vice versa. Hence the complementarity or Kuhn-Tucker condition is a consequence of the simplex method, not an assumption. This is brought about by the use of a basis which is not indicated in the classical statement.

Most General Form of LP Problem

Given:

An m x n matrix A (All quantities real);
A 1 x n row c of objective coefficients;
Two m x 1 columns $\underline{b}$ and $\overline{b}$ of constant range limits;
Two n x 1 columns of bounds $\underline{L}$ and $\overline{L}$; and
An initial value z_0
where any $\underline{b}_i$, $\overline{b}_i$, $\underline{L}_j$ or $\overline{L}_j$ may be 0, finite or infinite provided

$$\underline{b}_i \leq \overline{b}_i \quad , \qquad \underline{L}_j \leq \overline{L}_j \quad .$$

Find:

$x = \{x_i,\ldots,x_n\}$ such that
$z = cx - z_0$ is max subject to

$$\underline{L}_j \leq x_j \leq \overline{L}_j$$

and

$$\underline{b}_i \leq \sum_j a_{ij}x_j \leq \overline{b}_i \quad .$$

Simplification of Constraints (Rows)

(i) Since z_0 is a constant subtrahend, it can be ignored during the solution process. Note, however, that it may be modified by simplifications of the bounds.

(ii) If $\underline{b}_i = -\infty$, and $\overline{b}_i = +\infty$, the i-th "constraint" is merely a functional and does not affect the solution. It can thus be ignored.

(iii) If $\overline{b}_i$ is finite and $\underline{b}_i = -\infty$, the constraint can be written merely

$$\sum_j a_{ij}x_j \leq \overline{b}_i = b_i \quad .$$

(iv) If $\underline{b}_i$ is finite and $\bar{b}_i = +\infty$, the constraint can be rewritten

$$\sum_j (-a_{ij})x_j \leq -\underline{b}_i = b_i$$

and $\bar{b}_i$ can be ignored.

(v) If $\underline{b}_i = \bar{b}_i$ is finite, the constraint can be written

$$\sum_j a_{ij}x_j = \bar{b}_i \quad .$$

(vi) If $\underline{b}_i < \bar{b}_i$ but both are finite, let $b_i = \bar{b}_i$ and $R_i = \bar{b}_i - \underline{b}_i$. Then the constraint is

$$b_i - R_i \leq \sum_j a_{ij}x_j \leq b_i \quad .$$

Thus the true constraints can always be written

$$b - R \leq Ax \leq b$$

where

$$\begin{aligned} R_i &= +\infty \text{ (iii or iv above) and can be ignored, or} \\ &> 0 \text{ (vi above) , or} \\ &= 0 \text{ (v above) .} \end{aligned}$$

It is assumed that this has been done in discussing simplification of bounds on the structural variables x_j.

Simplification of Bounds on x (Columns)

(vii) If $\underline{L}_j = \bar{L}_j$ is finite, then x_j is fixed. It can be dropped, writing the constraints as

$$(b - A_j\underline{L}_j) - R \leq Ax \leq (b - A_j\underline{L}_j)$$

and adding $c_j\underline{L}_j$ to z_0.

(viii) If $\underline{L}_j$ is finite and $\overline{L}_j = +\infty$, let $\tilde{x}_j = x_j - \underline{L}_j$. Then with $\tilde{x}_j \geq 0$ replacing x_j, write the constraints as

$$(b - A_j\underline{L}_j) - R \leq Ax \leq (b - A_j\underline{L}_j)$$

and add $c_j\underline{L}_j$ to z_0.

(ix) If $\underline{L}_j = -\infty$ and $\overline{L}_j$ is finite, let $\tilde{x} = -x_j + \overline{L}_j$ and $\tilde{A}_j = -A_j$. Then with $\tilde{x}_j \geq 0$ replacing x_j, and $\tilde{A}_j$ replacing A_j in A, write the constraints as

$$(b + A_j\overline{L}_j) - R \leq Ax \leq (b + A_j\overline{L}_j)$$

and subtract $c_j\overline{L}_j$ from z_0.

(x) If $\underline{L}_j < \overline{L}_j$ but both are finite, let $\tilde{x}_j = x_j - \underline{L}_j$ and $L_j = \overline{L}_j - \underline{L}_j$. Then write the constraints ($\tilde{x}_j$ replacing x_j) as

$$(b - A_j\underline{L}_j) - R \leq Ax \leq (\bar{b} - A_j\underline{L}_j)$$

with $0 \leq x_j \leq L_j$ and $c_j\underline{L}_j$ added to z_0.

(xi) If $\underline{L}_j = -\infty$ and $\overline{L}_j = +\infty$, then x_j is a free variable. Free structural variables are rare. In an actual computer code it is probably better to retain them (as is standard practice) but, for simplicity of discussion, it is desirable to eliminate them theoretically. Since this has several possible complications, we will merely assume it has been done. (Free x_j in a valid model are not in fact free but have limits implied by the constraints. Actual elimination, in addition to a considerable amount of fixed work, requires that implied constraints be checked.)

Thus bounds on x can always be written

$$0 \leq x \leq L$$

where any $L_j > 0$ and may be infinite.

All the above rules and transformations (except xi) are commonly performed in LP software systems and the inverse transformations applied to output.

Simplified Generalized LP Problem and Solution

In view of the foregoing, we write an LP problem in the following form.

Given:

An m x n matrix A;

A 1 x n row c of finite objective coefficients;

Two m x 1 columns: b finite, R non-negative but with zero, finite or infinite elements; and

An n x 1 column of upper bounds, L, strictly positive but with finite or infinite elements.

Find:

An n x 1 column $x = \{x_1,...,x_n\}$ such that

$$z = cx \text{ is max subject to}$$
$$0 \leq x_j \leq L_j$$

and

$$b - R \leq Ax \leq b \quad .$$

Known Results with Simplex Method

Assume that feasible x exists and z_{max} is finite. (If not, well-known terminations will so indicate.) Then:

1) An optimal basis $\bar{B}$ is obtained in a finite number of iterations.

2) Letting u = b - Ax, an optimal primal solution

$$I\bar{u} + A\bar{x} = b \quad , \qquad 0 \leq \bar{u} \leq R \quad , \qquad 0 \leq \bar{x} \leq L$$

is obtained in the form

$$\bar{B}\bar{\beta} + AL^* = b - R^* \quad , \qquad \bar{\beta} \geq 0$$

where L^* is an n x 1 column of selected finite values from L, zero elsewhere, and R^* is an m x 1 column of selected finite values from R, zero elsewhere.
The solution vectors $\bar{u}$ and $\bar{x}$ are not directly evident but are composed from non-overlapping segments of R^*, $\bar{\beta}$ and L^*.

3) Letting $\underline{c}$ be a 1 x m row of values from c corresponding to basic x_j and zero elsewhere, an optimal dual structural vector is obtained:

$$\bar{\pi} = \underline{c}\ \bar{B}^{-1}$$

and, letting $d = \pi A - c$, an optimal dual slack vector is obtained:

$$\bar{d} = \bar{\pi}A - c \quad .$$

The signs of $\bar{\pi}_i$ and $\bar{d}_j$ are as follows:

$$\begin{aligned} \bar{\pi}_i &\geq 0 \text{ if } R^*_i = 0 \text{ (i.e., } \bar{u}_i < R_i \\ &= 0 \text{ if } \bar{u}_i \text{ is basic (i.e. in } \bar{\beta}) \\ &\leq 0 \text{ if } R^*_i > 0 \text{ (i.e., } \bar{u}_i = R_i) \end{aligned}$$

$$\begin{aligned} \bar{d}_j &\geq 0 \text{ if } L^*_j = 0 \text{ (i.e., } \bar{x}_j < L_j) \\ &= 0 \text{ if } \bar{x}_j \text{ is basic (i.e. in } \bar{\beta}) \\ &\leq 0 \text{ if } L^*_j > 0 \text{ (i.e., } \bar{x}_j = L_j) \quad . \end{aligned}$$

4) The value of $\bar{z} = z_{max}$ is given by either

$$\bar{z} = c\bar{x} - z_0$$

or

$$= \pi(b - R^* - \sum_j A_j L^*_j) + cL^* - z_0 = \pi b^* + cL^* - z_0 \quad .$$

In the sequel, z_0 will be ignored.

Extended Matrix Representation of Simplex Solution

The well-known results listed in the previous section are virtually intractable in any closed form expressions. Our task now is to represent them in more readily manipulative forms.

First we define expanded primal and dual matrix equations as follows.

Primal Equation:

$$\begin{bmatrix} B & A \\ & -I_n \end{bmatrix} \begin{bmatrix} \beta \\ L^* \end{bmatrix} = \begin{bmatrix} b - R^* \\ - L^* \end{bmatrix} , \quad \text{B is nonsingular, \underline{the basis}}$$

Let $\underline{L}$ be a m x 1 column with elements $\underline{L}_h = R_i$ if the h-th basic variable is u_i, and $\underline{L}_h = L_j$ if the h-th basic variable is x_j. Then the condition for primal feasibility is $\underline{0} \le \underline{\beta} \le \underline{L}$.

Dual Equation:

$$(\pi, d) \begin{bmatrix} B & A \\ & -I_n \end{bmatrix} = (\underline{c}, c) .$$

In order to describe dual feasibility conditions and also to define the basis, five more matrices are needed. Let e_i be the i-th m x 1 unit column (orthonormal), and E_j be the j-th n x 1 unit column. Let P_B be an m x m matrix consisting of selected columns e_i, and otherwise zero. The initial P_B is I_m. Let P_R be an m x m matrix which is all zero except where $R^*_i > 0$, and then the column is $-e_i$. The initial $P_R = 0$. Also, let P_Z be m x m, all zero except $-e_i$, when $R_i = 0$ and u_i in non-basic. The initial $P_Z = 0$.

Similarly, let Q_B be an n x m matrix consisting of selected columns E_j, and otherwise zero. The initial $Q_B = 0$. Let Q_L be an n x n matrix which is all zero except where $L^*_j > 0$, and then the column is $-E_j$. The initial Q_L is an arbitrary selection of this kind for L_j finite.

Let p be the number of e_i in P_B (initially m), and q the number of E_j in Q_B (initially 0). Then at every step the following conditions must hold:

$$p + q = m \quad \text{(complete basis);}$$
$$B = P_B + AQ_B \quad \text{is nonsingular (valid basis);}$$
$$P_B P_R = 0 \quad , \quad P_B P_Z = 0 \quad , \quad P_R P_Z = 0 ;$$
$$Q_L Q_B = 0 \quad \text{(note that } Q_B Q_L \text{ is nonconformable).}$$

The conditions for dual feasibility are then

$$\pi(I_m + P_Z + 2P_R) \geq 0$$
$$d(I_n + 2Q_L) \geq 0 \quad .$$

Let

$$M = \begin{bmatrix} B & A \\ & -I_n \end{bmatrix} \quad .$$

Then

$$M^{-1} = \begin{bmatrix} B^{-1} & \alpha \\ & -I_n \end{bmatrix}$$

where

$$\alpha = B^{-1}A \quad .$$

Hence

$$\begin{bmatrix} \beta \\ L^* \end{bmatrix} = \begin{bmatrix} B^{-1} & \alpha \\ & -I_n \end{bmatrix} \begin{bmatrix} b - R^* \\ - L^* \end{bmatrix}$$

$$(\pi, d) = (\underline{c}, c) \begin{bmatrix} B^{-1} & \alpha \\ & -I_n \end{bmatrix} \quad .$$

One Simplex Iteration

The two matrices

$$S_P = I_m + P_Z + 2P_R \quad , \qquad S_Q = I_N + 2Q_L$$

are pricing selection matrices. If πS_P has any negative elements, the corresponding u_i are candidates to re-enter the basis. Since $\pi_i = 0$ for u_i basic, a basic u_i cannot be selected. Note also that the columns of S_P corresponding to non-zero columns of P_Z are zero. Hence a u_i restricted to the value zero, which has left the basis, is never a candidate to enter again. Similarly, if dS_Q has negative elements, the corresponding x_j are candidates. If x_j is basic, then $d_j = 0$ and cannot be selected. If neither πS_P nor dS_Q has negative elements, and the primal solution is feasible, then the solution is optimal.

Suppose either πS_P or dS_Q has negative elements, and take the algebraically smallest. If this is for π_s, let $\alpha_s = B^{-1}e_s$; if it is for d_s , let $\alpha_s = B^{-1}A_s$. Now a pivot selection routine must be used to determine a value θ. There are seven possible outcomes for a primal feasible solution:

1) $\theta = \dfrac{\beta_r}{\alpha_{rs}}$ with $\alpha_{rs} > 0$. The candidate enters the basis from zero at level θ, replacing the r-th basic variable which goes to zero. Either P_B or Q_B changes in two columns or each changes in one column. $\underline{L}_r$ changes to R_s or L_s.

2) $\theta = \dfrac{\beta_r}{\alpha_{rs}}$ with $\alpha_{rs} < 0$. The candidate enters the basis from upper bound at level $R_s + \theta$ or $L_s + \theta$ replacing the r-th basic variable which goes to zero. P_B or Q_B changes as in 1 above. Also either P_R or Q_L changes to zero in the s-th column; this also changes R* or L* in one element. $\underline{L}_r$ changes as in 1 above.

3) $\theta = \dfrac{\beta_r - \underline{L}_r}{\alpha_{rs}}$ with $\alpha_{rs} < 0$. The results are the same as in 1 above except that the r-th basic variable goes to

upper bound and P_B or Q_L changes to non-zero in one column; this also changes R* or L* in one element.

4) $\theta = \frac{\beta_r - L_r}{\alpha_{rs}}$ with $\alpha_{rs} > 0$. The results are the same as in 2 above except that the r-th basic variable goes to upper bound. One column of P_R or Q_L becomes zero and another becomes non-zero. R* or L* changes in two elements.

5) $\theta = R_s$ or L_s. The candidate goes from lower to upper bound. One column of either P_R or Q_L becomes non-zero with a corresponding change in R* or L*.

6) $\theta = -R_s$ or $-L_s$. The candidate goes from upper to lower bound. One column of either P_R or Q_L becomes zero with a corresponding change in R* or L*.

7) $\theta = +\infty$. Necessarily, R_s or L_s is infinite. A class of unbounded solutions is determined.

Additionally, if the outgoing variable in the above cases 1 or 2 is u_r, for which $R_r = 0$, the r-th column of P_Z becomes $-e_r$. P_Z is never reduced.

If the current solution is infeasible, the number of cases is the same, but the selection rules are more complicated since $\beta_i < 0$ exist; this gives rise to several new β_i, α_{is} sign and magnitude combinations.

The pivot routine is rigidly defined (though variations are possible), and is the heart of the simplex method. Note that the selection of infeasible π_s or d_s of greatest magnitude is merely a rule of thumb. Any infeasible value is usable. However, given a candidate, the above 7 cases are deterministic. This characterizes the simplex method as iterative in nature, and no closed form expression is possible. It is clearly impossible to express the above solution changes in any kind of standard matrix or functional notation. They are only describable by cases which lead to improved solutions.

Candidate and pivot selection maintain the necessary conditions on P_B, P_R, P_Z, Q_B and Q_L and hence on B. Complementarity is maintained automatically.

Characterization of Simplex Variables

In some LP applications, the x_j are regarded as control variables and the u_i as state variables. While this may be appropriate in cases where LP is used as a simulation, in a more general view it makes little sense. It is worth considering just what kind of function an LP model is.

Given a matrix A, at least two geometries are implied; in fact both are used as a framework for the simplex method. Constraint space is E^n and m directions are implied, in addition to the n orthonormal coordinate units. Activity space is E^m and n vectors from the origin are implied in addition to the m orthonormal units. There is a strong relationship between these spaces. (In an elementary model, the dual problem has complementary spaces.)

One is not interested in the entire spaces but in a convex manifold, or simplex, in constraint space. This is defined by the vectors b, R and L, and the constraints

$$b - R \leq Ax \leq b \quad , \quad 0 \leq x \leq L$$

where $x \in E^n$. To begin with, all x of interest are in the positive orthant, possibly further constrained by the hyperplanes $x_j = L_j$. (One can, in fact, start from other x, as in restarting a modified problem from an old basis.) The other facets of the simplex are defined by the directions given by the rows of A and the distances from the origin implied by b and b - R. Hence, given A, b, R and L, a convex manifold $F \subset E^n$ is defined. Any $x \in F$ is said to be feasible. F can be expanded or contracted by changes in b, R and L which have the effect of moving hyperplanes parallel to themselves. Clearly, this can change both the number of facets and the number of vertices of F. Changes in A, of course, can distort F in any way.

The correspondence between F and the vectors in the E^m activity space is as follows. For every vertex of F, there is a nonnegative linear combination of the m + n vectors in E^m, some m of which are linearly independent and the remaining having their variables at a limit. This linear combination gives b, regarded as a point in E^m. A simplex basis is merely a linearly independent subset of the linear combination, the other vectors being combined by IR* and AL* and their variables regarded as temporarily fixed. The converse is not true, that is, given an R* and L* and a basis which gives the point b, this is not necessarily a vertex of F. It is, however, a vertex of a convex manifold which includes F. This fact is used in Phase 1 of the simplex method. (It is assumed that elements of R* and L* are either 0 or upper limits. Otherwise, non-vertex points of F are derived.)

All the foregoing is, of course, well known. The point is, however, that the functional z is not a function of x but of c. For, given A, b, R and L, the simplex F in E^n is completely determined. Only the points $x \in F$ are valid, i.e. feasible. The value of z is specified by $z = cx - z_0$, $x \in F$. Assuming that the purpose of LP is to maximize z over the manifold F, this maximum is determined by the direction c in E^n. Except for possible multiple x on a facet or edge of F orthogonal to c, the value of z over F is uniquely determined once c is specified. The simplex method is a process which, starting from any vertex x in F (or even in E^n) moves toward and eventually reaches an optimal point. It takes advantage of the fact that at least one optimal x is a vertex of F. (In general, ∞ may have to be regarded as a vertex of F.)

Hence one is justified in regarding F as a function in E^n and z = F(c). Given any c, there is a unique (or infinite) value z given by

$$z = \max cx \quad , \qquad x \in F \quad .$$

This assumes, of course, that the manifold F is not void which is equivalent to saying the function F is undefined.

Leaving A fixed, one can generate a family of F functions by varying b, R or L. These can be generated by the well-known parametric RHS algorithm (which can include upper bounds). In the equally well-known parametric objective algorithm, one is really calculating different values of z by parameterizing the argument c for the same function F.

If A is varied numerically but not dimensionally, a more extensive family of F functions are generated. This can be done by the less-utilized structural parametric algorithms. These create hyperbolic changes rather than linear ones and can lead to singularities. That is, let F_1 transform to F_2, and let $x_{(1)}$ be the optimal point in F_1. As $F_1 \rightarrow F_2$, B may become singular; ($x_{(1)}$ becomes infinite but this can also occur because of an unbounded F_2.) If one attempts to transform beyond F_2, say to F_3 for which B^{-1} exists, then $x_{(1)} \rightarrow \bar{x}_{(1)}$ which is not in the simplex F_3.

If the dimension m of A is changed, then the number of hyperplanes defining F is changed. This changes the basis (activity) space E^m but may or may not have any significant effect on F. In general, if m is reduced, F becomes larger and, if m is increased, F becomes smaller and may even vanish. However, in particular cases, there may be no material effect on F.

If the dimension n of A changed, say to $\bar{n}$, then E^n changes to $E^{\bar{n}}$. The mapping of $\bar{F} \subset E^{\bar{n}}$ into $F \subset E^n$ may be many-one, one-one, or one-many. In general, one has a different family of functions, defined in a space of different dimensions.

Simplex Iterations Regarded as a Trajectory; Phases 1 and 2

Given F, c and any vertex $x_o \in E^n$, the simplex method proceeds in one of two phases depending on whether $x_o \in F$ or not. If not, Phase 1 is performed. This amounts to partially solving a series of LP problems over $F_0 \supset F_1 \supset \ldots \supset F_T \supset F$. For each F_t, a different c_t is used. The final F_T is completely solved and the value $z_T = c_T x_T = F_T(c_T)$ is zero unless F is void, in which case no feasible solution exists.

Instead of defining the c_t directly, z_t is defined as follows. Assume that R* and L* are feasible. (There is never any reason why they should not be.) Let f_- be the set $\{i : \beta_i < 0\}$, and f_+ the set $\{i : \beta_i > \underline{L}_i\}$, at any stage t. Then

$$z = \sum_{i \varepsilon f_-} \beta_i - \sum_{i \varepsilon f_+} (\beta_i - \underline{L}_i) \leq 0 \quad .$$

Define $f = (f_1, \ldots, f_m)$ by

$$\begin{aligned} f_i &= 0 \text{ if } \beta_i \text{ is feasible} \\ &= 1 \text{ if } \beta_i < 0 \\ &= -1 \text{ if } \beta_i > \underline{L}_i \end{aligned}$$

and

$$\pi = fB^{-1} \quad .$$

It is easily shown that using this π and πA for normal pricing, and applying a somewhat enhanced pivot routine (as previously indicated), will either increase z (until one or more members of the sets f_- and f_+ drop out, in which case the next stage is commenced) or show that z cannot be increased. (F is void.) When both sets are empty, $z = 0$. (One additional rule must be imposed. Since $\pi B = f$ contains -1's, basic variables must not be priced.)

One can thus consider Phase 1 as defining a piece-wise linear trajectory from some initial point x_o outside F along edges of enclosing simplices leading to a vertex of F. Phase 2 is then a trajectory along edges of F to an optimal vertex. In general, these trajectories are not unique and depend on chance selections and tie-breaking rules.

It is sometimes hypothesized that a shorter total trajectory can be found by combining the phases. This is done by defining a scale factor $\sigma \geq 0$ and pricing in Phase 1 with $\pi A - \sigma c$. The scale is varied in magnitude depending on progress. Experience with this technique is mixed but, on the whole, it appears not to be very effective.

Let X(c) be the set of x on edges of F for which z is maximum. Then X(c) is a region of stability with respect to the simplex method. Any random errors in the algorithm which are not persistent (i.e., correct numbers are recalculated if necessary) will not prevent the trajectory from reaching a vertex of X(c). An arbitrary iteration made while on X(c) followed by a proper iteration will return to a vertex of X(c). If a (nonbasic) π_i or d_j of zero is used to select a candidate, the iteration moves along an edge of X(c) to an adjacent vertex. Thus X(c), as well as z, is a function of c.

Interpretation of Dual Feasibility as Another Simplex

It is possible to have a dual feasible solution which is not primal feasible. We assume this to be a basic solution in the foregoing sense. Although Phase 1 could be applied, the well-known dual algorithm[1] can be used to follow a trajectory to X(c) which is everywhere dual feasible. This may or may not represent a practical advantage but it is interesting in principle.

It seems superfluous to describe the dual algorithm, even briefly. Rather we can consider $F \subset E^n$ and its related simplices. Assuming that X(c) is finite, it lies in a hyperplane G defined by $cx = z_{max}$. For any x on one side of this hyperplane, $cx < z_{max}$; this side contains F. On the other side, $cx > z_{max}$. Hence G divides E^n into two parts, one containing a dual feasible simplex D and the other containing F. $F \cap D = X(c) \subset G$. Other points on G are primally infeasible but not necessarily dually infeasible without qualification. We will further examine this phenomenon in Part II.

---

[1] Although it is not straightforward to dualize a model with ranges and bounds, it is relatively simple to adapt the dual algorithm to a primal format with such conditions.

Basic solutions which are doubly infeasible are common. If there are extraneous primal constraints, it is also possible to have hyperplanes through D that do not form a distinct facet, but on which basic solutions exist that are primally infeasible and dually feasible. The dual algorithm may pass right through such points without stopping, as, for example, when two or more primal infeasibilities are removed at once. The existence of hyperplanes through D for which basic solutions on each side are dually feasible (but not necessarily zero in any dual basic variables) is one of the more disturbing aspects of convex geometry. In effect, certain dual variables are always feasible for any feasible values of the other dual variables in a basic solution. Hence extraneous primal constraints are translated into extraneous dual variables. It is also true that extraneous primal variables translate into extraneous dual constraints, but this is hardly surprising since constraint space has an unnecessarily high dimension. (Note: an "extraneous" constraint is one which is never binding for F, but is not necessarily "redundant" in the sense of linear dependence.)

Difficulties of Integer Programming

If one requires the x_j (or some subset of them) to take on only integer values, then the set of feasible solutions, say W, is not compact but consists of either lattice points or disjoint subsimplices. This causes three difficulties.

a) z_{max} is not, in general, achieved at a vertex of F. Hence basic solutions in continuous variables do not include z_{max}.

b) $F \cap G$ does not, in general, include any part of W. Even if it does, it is difficult to locate or identify. Hence, it is not a region of stability for the simplex method.

c) No continuously feasible trajectory exists which connects two or more disjoint parts of W.

However, assuming that an integer solution exists within F, there is a related function $F_W \subseteq F$ for which the maximum point x = w of W is a vertex of F_W. For, suppose we knew w; then we could calculate Aw and adjust b or R accordingly. Alternatively, we could bound x_j on one side or the other by w_j. The latter approach is the one taken by many branch and bound methods. However, it has the disadvantage of introducing new hyperplanes which cause extraneous dual variables and change the nature of both F and D. Some penalty function approaches are essentially an attempt to deform F into F_W by changes in b and R. These methods have the difficulty of identifying a point of W when a hyperplane slides through it, since it is not a vertex unless a sufficient number of hyperplanes are moved together in various precise proportions.

Suppose one solves the continuous problem and arrives at an optimal point $x_o \in X(c)$. (We will ignore the imponderables of unbounded continuous solutions which may have finite integer solutions. This requires irrational coefficients in any event.) We can assume that L^*_j is an integer if x_j is an integer variable. (Why should anyone put a noninteger upper bound on an integer variable?) Therefore any nonbasic x_j can move in only one direction by a minimum of one unit. The cost of such a move is d_j, assuming that it is possible from x_o while remaining in F.

A basic integer variable, on the other hand, can move in two directions (though one may be by an essentially zero amount), either of which gives an integer value, assuming that it is possible. The cost of doing this depends on how it is done. One way is to find a nonbasic variable to change. Suppose x_k is in basis position r. Then letting λ_j = +1 or -1 according as $x_j = 0$ or $x_j = L_j$

$$\bar{\varphi}_k = \min_{\lambda_j \alpha_{rj} < 0} \left\{ \frac{d_j}{-\alpha_{rj}} \right\}$$

gives the minimum rate of cost for movement of x_k up, and

$$\underline{\varphi}_k = \min_{\lambda_j \alpha_{rj} > 0} \left\{ \frac{d_j}{\alpha_{rj}} \right\}$$

gives the minimum rate of cost for movement of x_k down, where d_j and α_{rj} are understood to include π_i and π_{ri} ((r,i)-th element of B^{-1}) over all nonbasic, moveable variables x_j and u_i. This leads to three additional questions:

- whether it is possible to move the nonbasic variable enough to effect the desired change;
- whether the nonbasic variable must itself move by an integer amount; and
- whether the move, even if possible, will adversely affect other basic integer variables already at an integer value or nearly so.

Another main difficulty with integer programs is that complementarity cannot be maintained since an integer point will, in general, be in the interior of F which is not representable by a basic solution. Rather than belaboring these endless questions, we turn to a more fruitful theory, which, however, does not seem to help with the solution of integer programs.

PART II: THE COMBINED GEOMETRY OF PRIMAL AND DUAL BASIC SOLUTIONS

The concept of a dual simplex D introduced in Part I is not really very helpful and imposes awkward concepts such as useless hyperplanes cutting through it. Identification of the vertices in the half-space for which $z > z_{max}$ is necessary but, as now to be developed, a different interpretation of dual solutions is more fruitful.

The examples shown in Figures 2a, 2b, 2c, 2d and 3 will be used to illustrate various points in this section and to motivate certain definitions and theorems.

List of Common Notations, Identities, Conditions and Basis Change Formulae

We establish common notation to be used throughout the following definitions, lemmas and theorems.

The LP problem is assumed to have a finite maximum $z_{max} = z_0$ attained at one or more vertices P_o with basis B_o.
Other points and bases are designated by $P_1, P_2, \ldots$, with bases $B_1, B_2, \ldots$, respectively. Subscripting applies to any quantity associated with a point or a basis; if a subscript already appears, the point subscript is last.
Quantities associated with points and bases are:

- $\{u,x\}$: complete primal solution (column) vector.
- $\{\pi,d\}$: complete dual solution (row) vector.
- R^* : values of nonbasic u_i at R_i, zero elsewhere; an m x 1 column.
- L^* : values of nonbasic x_j at L_j, zero elsewhere; an n x 1 column.
- $\underline{L}_i$: upper limit for i-th basis variable. Note: $\underline{L}$ may sometimes be regarded as an m x 1 column. However, it must then contain infinite values. $\underline{L}_i$ refers to a finite R_i or L_{j_i}.
- β : column vector of basic variables. It will be assumed that any $u_i \in \beta$ (i.e., $e_i \in \beta$) is in its home position to avoid second order subscripts. When necessary, x_{j_i} will denote x_j in i-th basis position.

$\underline{c}$: the subrow of c belonging to the basis, in basis order. Hence $\pi = \underline{c}B^{-1}$.

A_j : any nonbasic column.

A_s : a particular A_j entering the basis or changing bound.

α_s : $B^{-1}A_s$.

π^r : the r-th row of B^{-1}.

α^r : the r-th row of $B^{-1}A$.

α_s^r : the pivot element in a change of basis.

d_j : $\pi A_j - c_j$. Also includes $\pi_i = \pi e_i$ when discussing all dual variables. If s is ambiguous for u_s or x_s, then

$$x_s \quad , \quad d_s \quad , \quad A_s \quad , \quad \alpha_s \quad , \quad \alpha_s^r$$

are read to include

$$u_s \quad , \quad \pi_s \quad , \quad e_s \quad , \quad B_s^{-1} \quad , \quad \pi_s^r \quad .$$

N : the dual basis corresponding to B.

β_r : the r-th basic variable leaving. In general, the index r may refer to any quantity associated with a variable leaving the basis.

B_r : r-th column of B.

b^* ≡ $b - R^* - AL^*$, i.e., current adjusted right hand side (rhs).

γ : row vector of dual basic variables, i.e., the active part of (π, d), ordered in dual basis order.

Also note the following identities for any basic solution, regardless of feasibility condition:

$$B\beta = b^* \qquad \text{(m x 1 column)}$$

$$z = \pi b^* + cL^* \qquad \text{(scalar)}$$

$$\pi A - c = d \text{ and hence } \pi A - d = c \qquad \text{(1 x n row)} \quad .$$

For any point $x \in E^n$, whether a vertex or not,

$$z = cx$$
$$u + Ax = b \quad .$$

Feasibility conditions at a vertex are:

Primal : $0 \leq \beta \leq \underline{L}$. (All nonbasic variables are at feasible values.)

Dual : $\pi_i = 0$ if u_i is basic, by construction,

$d_j = 0$ if x_j is basic, by construction.

If u_i is nonbasic, $\lambda_i = 1$ if $u_i = 0$; and

$\lambda_i = -1$ if $u_i = R_i$.

$\lambda_i \pi_i \geq 0$ for dual feasibility.

If x_j is nonbasic, $\lambda_j = 1$ if $x_j = 0$; and

$\lambda_j = -1$ if $x_j = L_j$.

$\lambda_j d_j \geq 0$ for dual feasibility.

For any change of primal solution starting from a vertex

$$B(\beta - \theta\alpha_s) + \theta A_s = b^* = b - R^* - AL^*.$$

If

(i) x_s changes from 0 to L_s, then $\theta = L_s$ and θA_s is transferred to the new AL^*. λ_s changes from +1 to -1.

(ii) x_s changes from L_s to 0, then $\theta = -L_s$ and θA_s is cancelled on the left and in AL^*. λ_s changes from -1 to +1.

(iii) x_s enters the basis, then some x_{j_r} leaves and

$$\theta = \frac{\beta_r}{\alpha_s^r} \text{ if } x_{j_r} \to 0 \qquad (\lambda_{j_r} \to 1)$$

$$\theta = \frac{\beta_r - \underline{L}_r}{\alpha_s^r} \text{ if } x_{j_r} \to L_{j_r} = \underline{L}_r \qquad (\lambda_{j_r} \to -1),$$

and $L_{j_r} A_{j_r}$ is transferred to AL*. Also,

$x_s \to x_s + \theta$, the new β_r (λ_s now effectively 0).

If $x_s = L_s$, then $L_s A_s$ is cancelled from AL*.

In all cases above:

$z \to z - \theta d_s$.

If x_s is really u_s, read

u_s for x_s , e_s for A_s , π_s for d_s ,

R_s for L_s and R* for AL* .

If x_{j_r} is really u_r, read

u_r for x_{j_r} , e_r for A_{j_r} ,

R_r for L_{j_r} and R* for AL* .

In case (iii) above, the dual values change as follows:

$$\pi \to \pi + \varphi\pi^r$$
$$d \to d + \varphi\alpha^r$$

where

$$\varphi = \frac{d_s}{-\alpha_s^r} .$$

Note that the change in z is

$$-\theta d_s = \varphi\,\beta_r \text{ or } \varphi(\beta_r - \underline{L}_r) .$$

<u>Definitions, Lemmas and Theorems</u>

<u>Definition</u>: A <u>basic solution</u> is one in which some m of the m + n primal variables $\{u_i, x_i\}$ have been identified whose columns of coefficients are linearly independent and form a basis B in E^m, and all the remaining n primal variables

are at finite limits. The finite limits on the nonbasic u_i are represented by $R^* \geq 0$, and those on the x_j by $L^* \geq 0$. The basic variables, both u_i and x_j, are represented in the vector β given by

$$\beta = B^{-1}(B - R^* - \Sigma\, A_j L^*_j) \quad .$$

<u>Definition</u>: An <u>extreme point</u> or <u>vertex</u> in primal constraint space E^n is one which can be represented by a basic solution. It is said to be <u>ambiguous</u> if it can be represented by more than one basic solution.

<u>Lemma 1</u>: A vertex is ambiguous if and only if β contains a limit value, say β_r, for which the representation of some usable nonbasic column in terms of B, say

$$\alpha_s = B^{-1}A_s \quad (\text{or } \alpha_s = B^{-1}e_s \text{ for nonbasic } u_i)$$

has an element $\alpha_s^r \neq 0$.

<u>Proof</u>: If for any i = r, all α_j^r for any nonbasic column are zero, then the r-th basic variable cannot be replaced and β_r is a constant for any solution. If the only

$a_j^r \neq 0$ are for the representation of some logical e_i (one of which then must in fact be e_r) and all such u_i are limited to the value 0, then all these e_i are unusable and β_r is constant.

If $\alpha_j^r \neq 0$ for some usable column and $\beta_r = 0$, then the j-th variable can enter the basis at its current limit value with the r-th basic variable leaving at zero. If $\beta_r = \underline{L}_r$, then the j-th variable can again enter the basis at its current limit value with the r-th basic variable leaving at $\underline{L}_r$. In both cases, the two solutions represent the same point, i.e. they are ambiguous.

If $\alpha_j^r \neq 0$ for some usable column and $0 < \beta_r < \underline{L}_r$, and if the j-th column enters the basis in position r, it must then take on a value which drives β_r to either 0 or $\underline{L}_r$ and hence represents a different vertex.

Finally, if any u_i or x_j changes from one finite limit to another, the new solution represents a different vertex.

Note that a vertex may be ambiguous while the u_i for an equality constraint is in the basis and maybe unambiguous once it has left.

Definition: Let x_1 and x_2 be the normally ordered columns of all primal structural variable values in any representations of vertices P_1 and P_2, respectively. Then P_1 and P_2 are distinct if $x_1 \neq x_2$.

Definition: Two vertices are adjacent if they are distinct and there exist basic solutions for each such that a) either only one variable changes its status in R* or only one in L*; or b) exactly two variables change their status between β and either R* or L* in going from one vertex to the other. The locus of points defined by this change, regarded as a continuous move, is called an edge. The two solutions are called edge ends.

Note that if P_1 and P_2 are adjacent and either is ambiguous, then the pair of edge ends is not unique. Also, if P_1 and P_2 are adjacent, P_2 and $P_3 \neq P_1$ are adjacent; then P_1 and P_3 may also be adjacent. Furthermore, the edge P_2—P_3 may be contained in P_1—P_3. In the illustration, for example, T and B are adjacent, B and P are adjacent and T and P are adjacent, since if x replaces u_3 in the basis for T, one gets the basis for P. The points T and P are unambiguous but B is ambiguous. The possible basic sets are as follows:

T	P	$B_{(1)}$	$B_{(2)}$	$B_{(3)}$
u_1	u_1	u_1	x	x
y	x	x	u_2	y
u_3	y	u_3	u_3	u_3
u_4	u_4	y	y	u_4 .

One can also get from B to P from either $B_{(1)}$ or $B_{(3)}$ with one change, but not from $B_{(2)}$. The same is true from B to T. To get from B to E, either $B_{(1)}$ or $B_{(2)}$ can be used but not $B_{(3)}$. The third combination, either $B_{(2)}$ or $B_{(3)}$ but not $B_{(1)}$, can be used to get from B to either A or C. From B to 0, any of the above three combinations can be used, but 0 is itself triply ambiguous.

The foregoing makes clear the nature of so-called degeneracy. Degeneracy is not a global phenomenon but is a characteristic of ambiguous vertices. Only in very rare cases does it lead to "cycling". However, it makes the choice among multiple bases very uncertain with respect to finding the next edge. Even if one knows the next edge, it may take several basis changes to "turn the corner".

However, note that although either $B_{(1)}$ or $B_{(2)}$ may be used to get from B to E, only $B_{(1)}$ leads to a dual-basic intermediate solution, R. (See Figures.) Note further that the dual solutions for

$$S = 1/3T + Y,$$
$$R = 5/21B_{(1)} + 16/21E,$$
$$P = \text{(doubly basic)},$$

which all lie on G: $x + 2y = 29/6$, all have the same dual solution. Let us compute the solutions for Q to illustrate how the composite solutions are obtained.

The value of z for Q is 29/6 since it lies on G. The values of z for B and C, which straddle Q on the same edge, are 7/2 and 5. Since z changes linearly on the edge, we have the proportions

$$z(B) = 7/2 = 21/6 \quad ; \quad z(Q) = 29/6 \quad ; \quad z(C) = 30/6 \quad .$$

Hence Q is 8/9 of the way between B and C, or

$$Q = 1/9B + 8/9C \quad .$$

There is no difficulty in computing the nonbasic primal solution for Q, viz.:

	u_0	u_1	u_2	u_3	u_4	x	y
1/9B =	7/18	0	0	1/9	0	1/18	3/18
8/9C =	40/9	0	-8/9	0	8/9	8/9	16/9
Q =	29/6	0	-8/9	1/9	8/9	17/18	35/18 .

To compute a valid dual solution, however, we must determine which basis - $B_{(1)}$, $B_{(2)}$, $B_{(3)}$ - to use. Since we already guess that Q should have the same dual solution as P, we can compare the (π,d) rows for P and C.

C:	1	1/2	0	3/2	0	0	0
P:	1	0	1/6	4/3	0	0	0 .

Hence the basis for B must have nonzero values for π_1 and π_2, i.e., u_1 and u_2 out of the basis. This is $B_{(3)}$. It is readily verified that

$$(\pi,d)_Q = 1/9(\pi,d)_{B_{(3)}} + 8/9(\pi,d)_C = (\pi,d)_P$$

and also that this fails for $B_{(1)}$ or $B_{(2)}$.

The preceding illustrates the following definitions and theorems.

Definition: Let B be a primal basis for some vertex P. The corresponding dual basis N in E^n is defined as follows: Let

A^i be the i-th row of A ,

N^j be the j-th row of N , and

E^j be the transpose of E_j .

Assume that the basis B is ordered so that any $e_i \in B$ are in their home positions, i.e., $B_i = e_i$. Then if

$$B_i = A_j \quad , \qquad \text{let } N^j = A^i \quad ;$$

$$\text{if } A_j \notin B \quad , \qquad \text{let } N^j = -E^j \quad .$$

<u>Lemma 2</u>: N is nonsingular.

<u>Proof</u>: By possible reordering of rows and columns, B has the form

$$B = \begin{bmatrix} I_{m-j} & \bar{A} \\ 0 & \underline{A} \end{bmatrix}$$

where J is the number of A_j in B. Then $\underline{A}$ must be nonsingular. Similarly

$$N = \begin{bmatrix} \underline{A} & \underline{A}/ \\ 0 & -I_{n-J} \end{bmatrix}$$

which is also nonsingular.

Note that both B and N are submatrices of M, defined earlier. Also, if $\underline{\pi}$ is the row of π_i for e_i not in B, and $\underline{d}$ the row of d_j for A_j not in B, then, assuming proper ordering, $(\underline{\pi},\underline{d})N = c$ since all other π_i and d_j are zero.

<u>Definition</u>: A <u>valid dual solution</u> is one which corresponds to the solution

$$(\underline{\pi},\underline{d}) = cN^{-1}$$

for some dual basis N. A valid dual solution is <u>feasible</u> if, when $\underline{\pi}$ is embedded in the full form π and $\underline{d}$ in the full form d (zero elsewhere)

$$\pi S_P \geq 0 \quad , \qquad dS_Q \geq 0 \quad .$$

<u>Theorem 1</u>: if z_{max} is finite, there exists at least one feasible dual basis N_o which is valid throughout the hyperplane G defined by

$$cx = z_{max} \quad .$$

If the corresponding $\underline{\pi}$, and $\underline{d}$ contain no zero elements, then N_o is unique.

Proof: Since z_{max} is finite, there is at least one vertex P_1 for which $cx = z_{max}$. If P_1 is ambiguous, then there is at least one basic solution which is both primal and dual feasible. All this follows from the proofs of the simplex method. Suppose P_1 is still ambiguous with doubly feasible, i.e., optimal, basic solutions. Since the primal change value θ between any two of these is zero, the dual change value

$$\varphi = \frac{d_s}{-\alpha_s^r} \qquad (d_s = \pi_s \text{ for incoming } u_s)$$

for a change of basis must also be zero. First note that a change of bound cannot occur or else $\theta = \pm L_s$ or $\pm R_s$, which are not zero. But a change of basis will cause the outgoing variable to have a new $\bar{d}_r = \varphi$ with a sign opposite to optimal. (All sign combinations are easily checked to verify this.) Hence $d_s = \varphi = 0$ if the new basis is dual feasible. But the new π and d vectors are given by

$$\pi + \varphi\pi^r = \hat{\pi}$$

$$d + \varphi\alpha^r = \hat{d}$$

so there is no change.

If there is a second distinct and adjacent optimal vertex P_2, then the θ for going from P_1 to P_2 is not zero. Hence again $d_s = 0$ or z would change by the value $d_s\theta \neq 0$ so that either P_1 or P_2 is not optimal, a contradiction.

Consequently, there is at least one doubly feasible primal basis which has a corresponding N_0. If all usable π_i and d_j are nonzero, N_o is unique. If N_o is not unique, then other such dual bases differ only in rows which do not affect the values of π and d so that any one such dual base produces a valid dual feasible solution for any optimal basic primal solution, except for the assignment of indices.

Let π_o and d_o be the normally ordered rows computed by (any) N_o and $z_o = z_{max}$. The equation

$$\pi_o A - d_o = c$$

does not depend on x. If $\bar{x}$ is any $x \varepsilon G$, i.e., $c\bar{x} = z_o$, then

$$(\pi_o A - d_o)\bar{x} = z_o \quad .$$

If $\bar{x}$ can be represented by an optimal basic solution with basis $\bar{B}$, then $\bar{B}$ implies some $\bar{N}_o$ which differs from N_o only in rows which do not affect π and d. If not, the original N_o gives the correct values for π and d.

Definition: Two vertices P_1 and P_2 are basically distinct if the bases for their representations, B_1 and B_2, are necessarily different. In particular, P_1 and P_2 are not basically distinct if they differ only by a change of bound in some x_j or u_i. If P_1 and P_2 are also adjacent, they are called basically adjacent. If P_1 and P_2 differ by a change of bound, they are limit adjacent.

Lemma 3: If P_1 and P_2 are basically adjacent, there exists a pair of edge end bases, B_1 and B_2, such that the change from (π_2, d_2) to $(\pi_1 d_1)$ involves only n + 1 dual variables.

Proof: By definition of basically adjacent, there exist bases B_1 and B_2 which differ in exactly one column, say

A_{j_r} in position r of B_2 ;

A_s in position r of B_1 .

Then $\alpha_s^r \neq 0$ and $\varphi = \dfrac{d_{s2}}{-\alpha_s^r}$. Both d_{s1} and d_{r2} are zero but d_{s2} may not be zero. In that case

$$d_{r1} = d_{r2} - \frac{d_{s2}}{\alpha_s^r}\alpha_{r2}^r \quad ;$$

but, since $d_{r2} = 0$ and $\alpha_{r2}^r = 1$, then $d_{r1} = \varphi$. Other than d_s and d_r, only n - 1 other variables are affected, or n + 1 in all.

The fact that $d_{r1} = \varphi$ will be used repeatedly.

Definition: The edge end bases B_1 and B_2 in Lemma 3 are called dual adjacent.

Note that a pair of dual adjacent bases may not be unique if P_1 or P_2 is ambiguous. This was already illustrated by the pairs $B_{(1)}$—E, $B_{(2)}$—E. Both pairs are dual adjacent, but $B_{(2)}$—E does not give a valid dual solution at R. (In fact, it gives a valid dual solution at a point between R and E on a line through C parallel to G.)

Theorem 2: Suppose that at an optimal vertex P_o with basis B_o some $d_s \neq 0$, some $0 < \beta_r < L_r$, and $\alpha_s^r \neq 0$. ($\beta_r > 0$ is sufficient for A.)

A.

If $\varphi = \dfrac{d_s}{-\alpha_s^r} > 0$ and

$$\frac{d_s}{-\alpha_s^r} = \min_j \left\{ \frac{d_j}{-\alpha_j^r} : \alpha_j^r \neq 0 \quad , \quad \frac{d_j}{-\alpha_j^r} \geq 0 \right\} ,$$

then there is a basically adjacent vertex P_1 with B_1 formed by replacing B_{ro} with A_s in B_o, for which the dual solution is feasible and $z_1 > z_o$. At least x_{s1} is infeasible.

Proof: Since $\alpha_s^r \neq 0$, A_s can replace B_{ro}. Since $\varphi = \dfrac{d_s}{-\alpha_s^r} > 0$, d_{r1} is feasible if β_{ro} goes to zero. By the assumption on φ, (π_1, d_1) is feasible. The primal change is given by

$$x_{s1} = x_{so} + \theta$$

with

$$\theta = \frac{\beta_r}{\alpha_s^r} \neq 0$$

$$\beta_1 = \beta_0 - \Theta\alpha_s \qquad (\beta_1^r = x_{s1})$$

$$z_1 = z_0 - \Theta d_s \quad .$$

Now if $x_s = 0$, then $d_s > 0$ and $\alpha_s^r < 0$; hence $\Theta < 0$ and $z_1 > z_0$. If $x_s = L_s$, then $d_s < 0$ and $\alpha_s^r > 0$; hence $\Theta > 0$ and again $z_1 > z_0$. Note that either $x_{s1} < 0$ or $x_{s1} > L_s$, which is infeasible in either case.

B. Let $K_i = \underline{L}_i$ if $\underline{L}_i$ is finite; $K_i = 0$ if $\underline{L}_i$ is infinite; If

$$\frac{\beta_r - K_r}{\lambda_s \alpha_s^r} > 0 \text{ and}$$

$$\frac{\beta_r - K_r}{\lambda_s \alpha_s^r} = \min_i \left\{ \frac{\beta_i - K_i}{\lambda_s \alpha_s^i} : \alpha_s^i \neq 0 \quad , \qquad \frac{\beta_i - K_i}{\lambda_s \alpha_s^i} \geq 0 \right\} ,$$

then there is an adjacent vertex P_2, formed either by replacing B_{ro} with A_s in B_o or by changing bound on x_s, for which the primal solution is feasible and $z_2 < z_0$. Either d_{s2} or d_{r2} is infeasible.

<u>Proof</u>: Since $\alpha_s^r \neq 0$, A_s can replace B_{ro}. The assumed minimum ratio is $\lambda_s\theta > 0$. If $x_s = 0$, then $\lambda_s = 1$ and $\theta > 0$. If $x_s = L_s$, then $\lambda_s = -1$ and $\theta < 0$. Hence if $|\theta| \leq L_s$, x_{s2} is feasible; if not, we can set $\lambda_s\theta = L_s$ so x_s goes to opposite limit. By the definition of $\lambda_s\theta$,

$$\beta - \theta\alpha$$

is feasible. If x_s goes to opposite bound without a change of basis, then $d_{s2} = d_{so}$ is now infeasible; all other π_i, d_j remain feasible. If a change of basis is made, and if β_{ro} went to zero, then $\lambda_s\alpha_s^r > 0$ and $d_{r2} = \varphi = d_s/(-\alpha_s^r) < 0$, since d_s and α_s^r have the same sign no matter which sign λ_s has. If β_{ro} went to $\underline{L}_r$, then $\lambda_s\alpha_s^r < 0$ and $d_{r2} = \varphi > 0$, since d_s and α_s^r have opposite signs no matter which sign λ_s has. (To see this, note that $\lambda_s = +1$ implies $d_s > 0$ and hence $\lambda_s\alpha_s^r > 0$ implies $\alpha_s^r > 0$; $\lambda_s\alpha_s^r < 0$ implies $\alpha_s^r < 0$. For $\lambda_s = -1$, $d_s < 0$ and the contrary results occur.) In either case, d_{r2} is infeasible.
In any case, z changes to

$$z_2 = z_o - \theta d_s < z_o$$

since θd_s is always positive.

<u>Corollary 2.1</u>: Vertices may exist with either $z > z_0$ or $z < z_o$ where both primal and dual solutions are infeasible.
<u>Proof</u>: In the proof of Theorem 2, part A., suppose some other nonbasic variable x_t has $\alpha_t^r \neq 0$ and

$$0 \leq \frac{d_t}{-\alpha_t^r} < \frac{d_s}{-\alpha_s^r} = \varphi \quad .$$

Then

$$d_t + \varphi\alpha_t^r = d_t - \frac{d_s}{\alpha_s^r}\,\alpha_t^r \quad .$$

If $d_t > 0$, then $\alpha_t^r < 0$ and $d_t + \varphi\alpha_t^r < 0$. If $d_t < 0$, the opposite holds. Hence d_{t1} is infeasible.
In the proof of Theorem 2 part B., suppose some other basic variable β_q has $\alpha_s^q \neq 0$ and

$$0 \leq \frac{\beta_q - K_q}{\lambda_s\alpha_s^q} < \frac{\beta_r - K_r}{\lambda_s\alpha_s^r} = \theta \quad .$$

Then

$$\beta_q - \theta\alpha_s^q < 0 \quad \text{or} \quad \beta_q - \theta\alpha_s^q > \underline{L}_q$$

as is readily verified from the preceding inequalities.
For example, suppose $\lambda_s = 1$ and $K_q = K_r = 0$. Then

$$\frac{\beta_q}{\alpha_s^q} < \frac{\beta_r}{\alpha_s^r}$$

so

$$\beta_q - \frac{\beta_r}{\alpha_s^r}\,\alpha_s^q < 0 \quad .$$

(Since we thus have shown one case, the corollary is proved.)

One would now like to show that, if P_1 is geometrically adjacent to P_o and $z_1 \neq z_o$, then P_1 is either basically or limit adjacent to P_o for some pair of bases B_o and B_1, at least if B_1 is either primally or dually feasible. Unfortunately, this is not true. If P_o is ambiguous and not all its bases are optimal, there may be no way to get to P_1 from an optimal B_o in one step. (See Figure 3.) The following question thus arises: Suppose P_2 is distinct from P_o and some B_2 is primal (dual?) feasible; then is there always a P_1 which is basically adjacent to P_o and primal (dual?) feasible? Before tackling this question, two related questions must be answered.

(1) On G_o there may be bases not primally feasible and bases not dually feasible. Is it possible for some B_o to be doubly infeasible?

(2) If $P_1 \neq P_o$ has a feasible primal (dual) basis, is it possible for P_1 to also have an infeasible primal (dual) basis? If not, why is P_o special?

Both questions (1) and (2) are improper but it is instructive to answer them. A feasible primal basis B_1 is a $1 \leftrightarrow 1$ transformation in E^m. The solution vector β assumes that the n nonbasic variables are assigned limit values. If some $\beta_i = 0$ or $\underline{L}_i$, any change of basis removing β_i will leave $\{u,x\}$ unchanged. If β_i is not at a limit, the change of basis moves to a different vertex, interpreted in E^n. The same is true of a change in limit. Hence a point P_1 in E^n which is primally feasible is primally feasible for any B_1.

Let γ represent the 1 x n vector of π_i and d_j for which u_i and x_j are non-basic. A feasible dual basis N_1 is a $1 \leftrightarrow 1$ transformation in E^n. The solution vector γ assumes that the m dual-nonbasic (i.e., primal basic) variables have been assigned the value 0. If some $\gamma_k = 0$, any change of basis which removes γ_k (i.e., brings some u_i or x_j into the primal basis) will leave (π,d) unchanged. If $\gamma_k \neq 0$, the change of basis moves to a different dual solution, which is in fact a hyperplane in constraint space, not a vertex. One must be careful to distinguish the E^n of the dual basis from the E^n of constraint space.

Hence the answer to (2) answers (1) also. If B_o is primal feasible for P_o, any $\bar{B}_o$ is also primal feasible, but the corresponding $\bar{N}_o$ need not all be dual feasible. If N_o is dual feasible for the hyperplane $G_o : cx = z_o$, any $\bar{N}_o$ is also dual feasible but the corresponding $\bar{B}_o$ need not all be primal feasible.

The main question is now also seen to be improper with respect to dual bases. The proper question is:

Suppose hyperplane G_2 is distinct from G_o and some N_2 is dual feasible. Then is there also an N_1 which is dual-basically adjacent to G_o and dual feasible?

Both primal and dual questions are subtle. We already know (restricting ourselves to the primal side for the present) that there can be B_o, B_1 and B_2 with $z_o < z_1 < z_2$ such that B_o and B_2 are basically adjacent but B_0 and B_1 are not. However, this is not the proper question. Rather, suppose P_2 is distinct from P_o and B_2 is primally feasible; must there be a P_1 which is basically adjacent or limit adjacent to P_o with $z_1 < z_o$?

If we start from B_2, the simplex method will arrive at an optimal B_o with primal feasibility maintained at each step. If the last iteration changed from some B_1 to B_o with $z_1 < z_o$, then we have found the desired P_1. But suppose the trajectory was as follows:

$$B_2 \to B_1 \to B_o^{(1)} \to B_o^{(2)} \to B_o$$

$$z_2 < z_1 < z_o^{(1)} = z_o^{(2)} = z_o$$

where $N_o^{(1)}$ and $N_o^{(2)}$ are not dual feasible. Must there be a B_3 (possibly B_1 itself) which is basically adjacent to B_o with $z_3 < z_o$?

Now the reason that $B_1 \to B_o^{(1)}$ was not optimal is that either there were ties for θ_1 leading to multiple $\beta_{io}^{(1)} = 0$, or some β_{i1} were already 0 and did not change. In the first case, a different choice for θ_1 would have led directly to $B_o^{(2)}$ or B_o. In the second case, P_1 is ambiguous, and we could have brought the other columns in B_o into B_1 first, provided these columns and the A_s which moves B_1 to $B_o^{(1)}$ are not linearly dependent. In either case, some B_1 would be basically adjacent to B_o. We have now reduced the desired theorem to one question.

<u>Theorem 3</u>: If $P_2 \neq P_o$, $z_2 < z_o$, and P_2 has a feasible basis B_2, then there exists a feasible P_1 with $z_1 < z_o$ which is either basically or limit adjacent to P_o.

<u>Proof</u>: Apply the simplex method starting at B_2 and let B_1 be the last basis for which $z_1 < z_o$. By the preceding discussion, we can assume that, if θ_1 is ambiguous, the most favorable for dual feasibility has been chosen. If the next iteration does not lead to an optimal B_o, let

$B_o^{(t)}$, $t = 1,2,\ldots,T$, be the series of primal feasible, dual infeasible bases between B_1 and B_o. Let A_s, B_r be the columns which enter and leave between B_1 and $B_o^{(1)}$, where B_{r1} may be simply A_s with x_s gone to opposite limit. Let $IN = \{A_t\}$ be the collection of columns which enter, and $OUT = \{B_{qt} : q = q_t\}$ the collection of columns which leave bases $B_o^{(t)}$ in going from $B_o^{(t)}$ to $B_o^{(t+1)}$ where $B_o^{(T+1)} = B_o$.

Suppose first that A_s simply goes to opposite bound. Then the basis does not change and, by withholding A_s, all other basis changes can be made first. The resulting solutions may not be primal feasible. However, assuming that A_s is not itself in IN, then the iteration from $B_o^{(T)}$ to $B_o^{(T+1)}$ produces a basis which differs from B_o by x_s being at the wrong limit. This basis is therefore dual feasible except for d_s. The point for $B_o^{(T+1)}$ is exactly P_1 which is primal feasible. Hence the result holds.

Now assume that A_s is as above, but that A_s is in IN but not in OUT (or in IN once more than in OUT). A_s must still be at the wrong limit in $B_o^{(T+1)}$, and some other column B_{q_s} must be out at a limit. Since this is the same point as before, we must be able to replace A_s with B_{q_s} and have the same situation as above. But then, if x_s changes limit the solution is still not dual feasible since A_s should be in the basis. But now it goes in at the proper limit and must replace B_{q_s} which must itself have gone to an opposite limit. Hence x_{q_s} must also have an upper bound, and changing its limit in the first place would have achieved the desired result.

Suppose now that A_s enters the basis from B_1 to $B_o^{(1)}$. We can outlaw the case that A_s goes into the basis at opposite limit. Then A_s enters at a nonlimit value and can never leave since all succeeding $\theta_t = 0$. Hence B_{r1} is out of the basis and made a nonzero change (β_{r1} or $\underline{L}_{r1} - \beta_{r1}$). If x_{r1} is not in IN, then we can return to P_1 from B_o by replacing A_s with B_{r1}. This must be possible since B_{r1} has a nonzero component for A_s in terms of $B_o^{(1)}$; since all other columns in B_o are linearly independent of A_s, this component cannot vanish. If B_{r1} is in IN, then some other column B_{q_r} is out of B_o with a nonzero component for A_s; but the component in the corresponding β for B_{r1} is at a limit. In $B_o^{(1)}$, B_{q_r} or some surrogate for it had a limit component in β. If we invoke the assumption that the most favorable θ_1 for dual feasibility was selected, then the component of B_{q_r} in β_1 was also at a limit. Hence A_s had a nonzero component at B_1 for both B_{r1} and B_{q_r}; but β_{r1} was nonlimit while the component in β for B_{q_r} was at a limit. Hence P_1 was ambiguous and replacing B_{r1} with B_{q_r} gives an alternate basis. But we can get back to this basis from B_o by replacing B_{r1} with B_{q_r} there. Any such change of basis leads to a basically adjacent P_1 with $z_1 < z_o$.

Theorem 4: Let the hyperplane $G_2 : cx = z_2$, be distinct from G_o, with $z_2 > z_o$ and a dual feasible basis N_2. Then there exists a dual feasible G_1 with $z_1 > z_0$ and a pair of dual-basically adjacent bases N_o and N_1.

We will not give a formal proof of this proposition. The arguments of Theorem 3 can be used almost intact by reading "hyperplane" for "point", "dual" for "primal", γ for β, etc. A change of bound does not arise for d_j or π_i in the same sense as it did for x_j (and, by extension, u_i). However, the analogue is as follows. Instead of d changing to $d + \varphi\alpha^r$ on such an iteration, one can regard $-E_s$ as changing to $+E_s$ or vice versa. The handling of change of limit on a u_i can be regarded as A^i changing to $-A^i$ or vice versa.

We are now getting a fairly good view of the topography of constraint space and particularly around the optimal facet. The greatest complications come from ambiguous points in the primal half-space (PHS) and ambiguous hyperplanes in the dual half-space (DHS). The picture emerging is as shown in Figure 1 for E^2, where primal basic solutions in PHS represent vertices, dual basic solutions in DHS represent hyperplanes.

> Definition: A functional cut or slice in primal constraint space E^n is one which can be represented by a basic dual solution. It is said to be ambiguous if it can be represented by more than one dual basic solution. Such a slice is a hyperplane parallel to $cx = z_o$.

It is worth a few words to clarify what is meant by the statements that a primal basis represents a vertex and a dual basis represents a slice, however well-known they may be. A basis in E^m does not, of course, really represent a point in E^n. What is meant is that if one selects a linearly independent set of m coefficient columns (which must exist by construction) out of the total of m+n e_i and A_j, then, for an arbitrary assignment of values to the other n variables, a unique solution exists for the selected m. If the n arbitrary values are set to limit values, then the resulting solution must lie on corresponding bounding hyperplanes of the manifold, represented by the equations $A^i x = b_i^*$ and $x_j = 0$ or L_j. Lemma 2 says that the set of rows for these equations must also be linearly independent and, by definition, form the dual basis. Since a set of n linearly independent hyperplanes in E^n define a point, we call this a vertex. The m basic variables represent unnormalized distances from the other hyperplanes. The fact that they represent a unique linear combination of vectors in E^m is only a convenient viewpoint.

The viewpoint for the dual basis is in terms of n linearly independent hyperplanes for which a unique solution exists, given that the other m dual variables are set to zero. Actually, both viewpoints are equivalent. Hence we shall not try to prove a lemma to the effect that every slice contains at least one vertex; this is really only a tautology since a dual solution is only defined by reference to a primally-defined vertex. However, we will show the uniqueness of the complementary definitions.

Lemma 4: Every B_1, R_1^* and L_1^* (taken as a set) imply a unique N_1 and conversely.

Proof: For summary purposes, we restate the definition of a dual basis. Given a primal m x m basis B_1 and an assignment of limit values R_1^* for nonbasic u_i, L_1^* for nonbasic x_j, the complementary dual basis is the following n x n matrix N_1:

$$\text{If } B_{i1} = A_j \text{ and } u_i = 0 \quad , \qquad \text{then } N_1^j = A^i \quad ;$$

$$\text{if } B_{i1} = A_j \text{ and } u_i = R_i \quad , \qquad \text{then } N_1^j = -A^i \quad ;$$

$$\text{if } A_j \not\in B_1 \text{ and } x_j = 0 \quad , \qquad \text{then } N_1^j = -E^j \quad ;$$

$$\text{if } A_j \not\in B_1 \text{ and } x_j = L_j \quad , \qquad \text{then } N_1^j = +E^j \quad .$$

Clearly the assignments are unambiguous and complete, and therefore reversible, as follows.

Given a dual n x n basis N_1, the complementary primal basis B_1, together with assignment of limit values for nonbasic primal variables, is:

$$\text{If } N_1^j = -E^j \quad , \qquad \text{then } A_j \not\in B_1 \text{ and } x_j = 0 \quad ;$$

$$\text{if } N_1^j = E^j \quad , \qquad \text{then } A_j \not\in B_1 \text{ and } x_j = L_j \quad ;$$

$$\text{if } N_1^j = A^i \quad , \qquad \text{then } B_{i1} = A_j \text{ and } u_i = 0 \quad ;$$

$$\text{if } N_1^j = -A^i \quad , \qquad \text{then } B_{i1} = A_j \text{ and } u_i = R_i \quad .$$

If any B_i remain unassigned, they are filled with e_i.

This characterizes B_1, R_1^* and L_1^* completely and unambiguously (except for ordering of A_j in B positions not filled with e_i).

Lemma 5: If there are two distinct optimal vertices, then G_o is ambiguous.

Proof: Let $P_o^{(1)}$ and $P_o^{(2)}$ have optimal bases B_1 and B_2 together with their R_1^*, L_1^* and R_2^*, L_2^*. Since $x_1 \neq x_2$, there are corresponding dual bases N_1 and N_2, also distinct. However, since $(\pi_1, d_1) = (\pi_2, d_2)$, as computed by B_1 and B_2, $\gamma_1 = \gamma_2$ and contain common zero values which span the differences in rows between N_1 and N_2. Hence G_o can be represented by either N_1 or N_2, both of which are feasible, i.e., G_o is ambiguous.

It is difficult to attach a meaning to "two distinct optimal slices". However, there is an analogous phenomenon, as has already been illustrated. Note first that, if $m \geq n$ and all A_j are linearly independent, all A_j may be in B. In this case $-I_{n-J}$ disappears. However, N is still nonsingular and unambiguous, provided proper identification of the i-indices is maintained. The assumption that all e_i in B are in their home positions is critical.

Note further that if $m > n$ (see Figures 2a to 3), then the rows A^i must be linearly dependent. However, the phenomenon under discussion does not depend on the linear dependence of the A^i. Even if they are independent, the total number of rows available to N for exchange must be dependent. Let us count them.

		min	max
(a)	m rows A^i less number of equality constraints	0	m
(b)	as many rows $-A^i$ as there are finite R_i	0	m
(c)	n rows $-E^j$	n	n
(d)	as many rows E^j as there are finite L_j	0	n
	Total	n	2(m + n)

If there are n, then only n - m rows of N can be exchanged. This can only occur if $m \leq n$ and all A^i are independent (or if $m > n$ with $m - n$ redundant but consistent equality constraints).

In all cases, however, the number of rows available to N is greater than the number of available positions.

Consequently, not only can more than one N represent the same hyperplane as shown in Lemma 5, but dual solutions may also be distinct (not merely ambiguous) and some may be feasible and some not. This is different from the situation noted with respect to vertices and primal bases, and deserves further analysis.

Multiple Sheets for Functional Cuts

The reason that distinct dual solutions represent the same hyperplane is clear enough algebraically: we have more than n hyperplanes given in E^n with which to represent a specific hyperplane. If we return to the extended dual matrix equation

$$(\pi, d)\begin{bmatrix} B & A \\ & -I_n \end{bmatrix} = (\underline{c}, c) \quad ,$$

the ambiguity disappears formally, since we have a nonsingular system of m + n equations in m + n variables; this has a unique solution. If we multiply both sides on the right by the column $\{0, \mathbf{x}\}$ for corresponding **x**, we get the identity

$$(\pi A - d)x = cx = z \quad .$$

However, π is determined by the definition of B (which implies $\underline{c}$) and x is determined by B, R* and L*, whose determination is, of course, the whole problem. Since the specification of B, R* and L* is uniquely equivalent to the specification of N, we learn very little from the above identities except as a means of remembering complementarity. While we are at it, we may as well display the dual identities:

$$\begin{bmatrix} B & A \\ & -I_n \end{bmatrix}\begin{bmatrix} \beta \\ L^* \end{bmatrix} = \begin{bmatrix} b - R^* \\ - L^* \end{bmatrix} \quad .$$

Multiplying both sides on the left by the corresponding (π,d), we have

$$(\underline{c},c)\begin{bmatrix}\beta \\ L^*\end{bmatrix} = (b^* + AL^*) - dL^*$$

or

$$\underline{c}\beta + cL^* = (b^* + AL^*) - (\pi A - c)L^* = \pi b^* + cL^* = z \quad .$$

Note that here we had to use both π and d, whereas above we used only x. The reason for not using u is that it is subsumed in β and R* and keeps hiding, so to speak. But it is exactly the alternate possibilities for u which cause distinct dual solutions to represent the same hyperplane. In other words, it is not sufficient to specify which x_j go <u>in</u> the basis; we must specify which u_i go <u>out</u>. Anyone who thinks the primal slacks are not an essential part of the problem is mistaken.

One way to resolve the ambiguity of distinct dual solutions is to allow the possibility that a hyperplane may consist of multiple sheets.

<u>Definition</u>: A slice has <u>multiple sheets</u>, or briefly is <u>laminated</u>, if it can be represented by two or more dual bases whose solution vectors are nontrivially distinct, i.e., have different numerical values. A sheet is feasible if its dual basis has a feasible solution vector. A laminated slice is feasible if all its sheets are feasible; it is potentially feasible if some sheets are feasible.

<u>Lemma 6</u>: If any slice G is laminated, some vertex P on G is ambiguous.

<u>Proof</u>: Let N_1 and N_2 be two adjacent dual bases which represent G with $(\pi_1,d_1) \neq (\pi_2,d_2)$. These imply two primal solutions represented by the sets B_t, R_t^*, L_t^* (for $t = 1,2$) which have the same value for z. These cannot differ by a change of bound since then the (π_t,d)

differ only in one sign reversal, say for d_s, and z would change by $\pm d_s L_s \neq 0$. Hence $B_1 \neq B_2$. But since the dual solutions change, $\varphi \neq 0$ and so $d_s \neq 0$. Therefore $\theta = 0$ and B_1 and B_2 are ambiguous.

In the preceding proof, the concept of adjacent dual bases was used together with the assumption that these imply primal bases differing in only one column. These seem obvious without formal definition and proof since one can only define dual bases for multiple sheets by going from one such B_1 to another such as B_2. However, the lemma is somewhat more than a tautology since the proof also gives the following non-obvious result.

Lemma 7: Multiple sheets are never implied by two primal solutions which differ only by a change of bound.

Furthermore, the converse of Lemma 6 is not true since P may be ambiguous without implying a change in (π,d). All we can say is:

Lemma 8: If some P on G is ambiguous, then G is either laminated or ambiguous, according as the bases for P give different dual solutions or not.

Clearly individual sheets may themselves be ambiguous. Suppose G contains two distinct point P_1 and P_3 but P_1 is ambiguous with bases B_1 and B_2 which give different dual solutions. Suppose further that B_3 gives the same dual solution as B_2. Then N_1 and N_2 represent different sheets, but N_2 and N_3 represent the same sheet. We also have the following.

Lemma 9: If G is laminated but not ambiguous, then G contains exactly one ambiguous vertex.

Proof: Since G is laminated, there are two or more N_t which give γ_t which differ. Since G is not ambiguous, no two γ_t are equal. Each N_t implies a primal set B_t, R_t^*, L_t^*. By Lemmas 5, 6, 7, all primal sets represent the same point P and do not differ by change of bound. Hence the B_t are different and P is ambiguous.

Thus, in summary, G is ambiguous if and only if it has multiple vertices. If G is laminated, it has ambiguous points. If it has ambiguous vertices, then it is either laminated or ambiguous, or both.

We have still to explain (1) why we cannot get from an ambiguous point to all adjacent points from all representations in one step, and (2) why some pairs of dual solutions combine linearly along an edge and others do not. Actually the two questions are one and the same. In Figure 2a for example, the point B is ambiguous and has three bases. The points P, C, E are unambiguous but each pairs up with one of the B bases for dual solutions, and with two of the B bases for primal solutions. The points Q and R cannot be represented by basic primal solutions but, since they all lie on G_O, they all have the same dual solution. If one tries to compute this by interpolating between B and C or between B and E, one must use the proper pair of edge ends.

If G_B is the slice through B, then it is laminated, with three sheets, and also is unambiguous since no other vertex lies on it. Each of the B bases can be regarded as representing superimposed points on the three sheets. Each sheet has a dual basis which is adjacent to the dual bases for G_O containing P, for G_C containing C, and for G_E containing E. Using the primal basis designations for identification, we have the following adjacency relationships:

$B_{(1)}$ is dual adjacent to E, u_4 is fixed (nonbasic), and u_1 and u_2 can interchange;

$B_{(2)}$ is dual adjacent to C, u_1 is fixed (nonbasic), and u_2 and u_4 can interchange;

$B_{(3)}$ is dual adjacent to P, u_2 is fixed (nonbasic), and u_1 and u_4 can interchange.

Hence R is between $B_{(1)}$ and E, but not between either $B_{(2)}$ or $B_{(3)}$ and E. Note that this adjacency is characterized by which u_i is <u>out</u> of both bases, i.e., which A^i remains in both dual bases.

Nevertheless, $B_{(2)}$, for example, is adjacent to C also since there is a point between R and E on G_C. The dual solution for this point can be interpolated between $B_{(2)}$ and E, i.e., it is the dual solution for C and G_C.

Several other interesting relationships can be gleaned from this simple example. We point out just one more. Consider the vertex Y. It is unambiguous and has a dual basis for G_Y. Y is primally adjacent to E, C and P, and in fact to X. Can we extrapolate from $B_{(1)}$ past E for a point on the intersection of this line and G_Y? The answer is no in any meaningful sense. The (x,y) coordinates are (6/7,18/7) = 10/7E - 3/7B. If one uses these proportions for the dual solutions, a nonbasic dual solution is obtained entirely different from the dual solution at Y. For one thing, there is no way to make d_x go to 1, as it is at Y. The nonbasic dual solution has d = 0 and

$$\pi = (0 \ -3/10 \ 5/2 \ 7/10) \quad ,$$

which multiplied onto b gives the correct value of 6. Thus the difficulty does not arise from the extrapolation. For, suppose there was a vertex beyond E with z greater than 6. We still could not interpolate and obtain (π,d) for Y since d_x would remain zero. The trouble is that one cannot interpolate unless there is cancellation of one of the n + 1 dual variables of Lemma 3. Consequently, dual adjacency, to be fully useful, has a somewhat stronger requirement than has hitherto been noted. The edge ends behave properly since d_s and d_r must be zero at opposite ends. But unless d_s was zero initially (and hence both are zero along the entire edge), neither is zero at any intermediate point. Hence a third dual variable must go through zero, i.e., either be zero or change signs, for a valid interpolation. This cannot occur if both edge ends are dual feasible, unless both are zero in the same positions. In general, one can only hope for valid interpolation if one edge is dual feasible in DHS and the other is in PHS and hence not dual feasible. This is particularly true if one stays on a trajectory which is either primal or dual feasible. Since extrapolation requires one weight to be negative, the opposite is roughly true: both edge ends must be feasible or infeasible. (No such situation occurs in Figure 2a.) However, extrapolation incurs other considerations which we will not discuss.

A converse question now arises which is more interesting. If a third dual variable, nonzero at either end, vanishes along an edge, must this point be on a slice? The answer is yes!

<u>Theorem 5</u>: If B_1 and B_2 are edge end bases for P_1–P_2, dual adjacent, where A_s replaces B_{r1} to form B_2, and a third dual variable d_t (or π_t) is nonzero at each end and vanishes on the edge, then the point at which d_t (or π_t) vanishes is on a slice defined by a vertex basically adjacent to both P_1 and P_2.

<u>Proof</u>: Since $d_{t1} \neq 0$ and $d_{t2} \neq 0$, but for some $0 < p < 1$ and $q = 1 - p$, $pd_{t1} + qd_{t2} = 0$, clearly d_{t1} and d_{t2} have opposite signs. Since $d_{t2} - d_{t1} \neq 0$, then $\varphi \neq 0$, and hence $d_{s1} \neq 0$. Since B_1 and B_2 (i.e., P_1 and P_2) are distinct, then $\theta \neq 0$ and hence β_{r1} is not at a limit. Clearly $\alpha^r_{s1} \neq 0$ and, since d_t changes by

$$d_{t2} = d_{t1} - \frac{d_{s1}}{\alpha^r_{s1}} \alpha^r_{t1} \quad ,$$

then $\alpha^r_{t1} \neq 0$. Now consider $\varphi_t = \dfrac{d_{t1}}{-\alpha^r_{t1}} \neq 0$. If A_t replaces B_{r1} in B_1 to form B_3, a different vertex P_3 is obtained in which $d_{t3} = 0$. P_1 and P_3 must be distinct

since $\theta = \dfrac{\beta_r - K_r}{\alpha^r_{t1}} \neq 0$. Hence P_1 and P_3 are basically

adjacent. B_3 has a dual basis N_3 for a slice containing $pP_1 + qP_2$. It remains only to show that P_2 and P_3 are

basically adjacent. Now $\alpha^r_{s2} = \dfrac{\alpha^r_{s1}}{\alpha^r_{t1}} \neq 0$.

Hence A_s can replace A_t in B_3 to give B_2. Since $z_3 = pz_1 + qz_2 \neq z_2$, the θ from B_3 to B_2 is not zero, and so P_1 and P_3 are distinct and hence basically adjacent.

PART III: SOME ADDITIONAL SIMPLEX PROCEDURES AND APPLICATIONS

Exploring the Vicinity of Optimality

Definition: The vicinity of optimality, or V_o, is the set of all solutions which are either optimal or adjacent to an optimal solution, with either dual feasibility maintained ($z > z_o$) or primal feasibility maintained ($z < z_o$).

The term "neighborhood" is to be particularly avoided since V_o may encompass a substantial part of E^n.

We give a procedure for stepping through all solutions in V_o, each just once, although the path to any particular one may not be the most efficient. It is assumed that an optimal solution has been obtained at some vertex P_{oo} with primal basis B_{oo}, R^*_{oo}, L^*_{oo}. This basis must be saved so that the procedure can return to it from time to time. This will be termed "returning to B_{oo}". We first define certain acts (i.e., subroutines) which are used repeatedly. The identifying indices - f, g, p, q - can be assigned or defined in any convenient way.

Act 1: Given some dual feasible primal basis B_{gp}, etc.

1) In the β for B_{gp}, search in order on i for β_i at a limit. If there is none, G_g is not laminated with respect to the current P_{gp}; exit. Else set $r = i$ for each such β_i in sequence.
2) For $i = r$, determine φ_1. If $\varphi_1 = 0$, proceed to next i. Else φ_1 determines some $j = s$, and π^r was (or could be) generated in the process.
3) For $j = s$, $(\pi,d)_{gp} + (\pi^r,\alpha^r)$ gives a feasible dual solution on a sheet G_{gq} at the same vertex P_{gp}.

Note that the row r is not reusable since, if the change of basis were actually made, $d_s = 0$ while β_r remains at a limit. Also, z does not change since $\theta = 0$.

Act 2: Given B_{oo}, etc.

1) In the β for B_{oo}, search in order on i for β_i not at a limit. In this process, bounded basic

variables must be treated as follows:
if $\beta_i = 0$, treat it as not at $\underline{L}_i$;
if $\beta_i = \underline{L}_i$, treat it as not at 0;
if $0 < \beta_i < \underline{L}_i$, two cases exist:

$$\beta_i \text{ not at } 0 \quad , \qquad \beta_i \text{ not at } \underline{L}_i \quad .$$

If no such β_i is found and there are no upper limits in the problem, it is trivial and V_o is the origin. If there are nonbasic variables with upper limits, but no basic ones, then all $\beta_i = 0$ and the constraints are equalities. Resolve the problem with u_i so specified. Hence we assume that some β_i not at a limit is found. Set $r = i$ for each such β_i in sequence. Regard this as some p-case (possibly two).

2) For $i = r$, determine φ_2. If $\varphi_2 \neq 0$, add the triplet $(\Delta z,r,s)$ to a g-list where

$$\Delta z = \frac{\beta_r}{-\alpha_s^r}\, ds \qquad \text{or} \qquad \Delta z = \frac{\beta_r - \underline{L}_r}{-\alpha_s^r}\, ds$$

or both if β_r is intermediate between θ, $\underline{L}_r$. Then proceed to the next i.

3) If no $\varphi_2 = 0$ is found, G_o is unambiguous with respect to P_{op}. Else, if $\varphi_2 = 0$, make the change (or changes) of basis indicated: A_s replaces r-th basic column (with possibly two θ_s). This solution is itself an adjacent primal solution on G_o.

4) Execute Act 1 for B_{op}. When through, return to B_{oo}.

<u>Act 3</u>: Given some primal feasible basis B_{fq}, etc.

1) In the γ for N_{fq}, search in order on k for $\gamma = 0$. If there is none, P_f is not ambiguous with respect to the current G_{fq}, exit. Else set $s = j_k$ for each such γ_k in sequence. (j_k is some primally non-basic u_i or x_j index.)

2) For column s, determine θ_1. If $\theta_1 = 0$, proceed to the next k. Else θ_1 determines some $i = r$ (α_s was necessarily generated). Note: if x_s changes bound, $r = 0$.

3) For $i = r$, $\beta - \theta\alpha_s$ ($\beta_r = x_s + \theta$ if $r \neq 0$) gives a feasible primal solution at some vertex P_{fp} on the same sheet G_{fq}.

Act 4: Given B_{oo}, etc.

1) In the γ for N_{oo}, search in order on k for $\gamma_k \neq 0$. If none is found, the problem is trivial: any primal feasible solution is optimal. (More accurately, V_o is the entire simplex F.) Hence we assume that some $\gamma_k \neq 0$. Let $s = j_k$ for each such γ_k in sequence. Regard this as some q-case.

2) For column s, determine θ_2. If $\theta_2 \neq 0$, add the triplet (Δz,r,s) to an f-list, where $\Delta z = -\theta d_s$, and d_s is $\pm\gamma_k$ ($r = 0$ if a change of bound). In either case proceed to the next k.

Steps of the Main Procedure:

1) Execute Act 1 for B_{oo}. This gives the optimal solution on all other feasible sheets G_{oq} (if any) at vertex P_{oo}.

2) Execute Act 2. This gives all optimal solutions on all feasible sheets G_{oq} at all other vertices P_{op} (if any). It also generates a g-list.

3) Sort the g-list in ascending order on Δz. (This is optional.) Assign g-indices to items of the resulting list. ($g \neq 0$, already used.) If the list is empty, then no vertices exist in DHS, i.e., no feasible slices with $z > z_o$. (This could conceivably happen if one constraint was itself $cx \leq z_0$.) Else, for each g, make the indicated change of basis and call this B_{go}, at vertex P_{go}. This solution is itself an adjacent solution to an optimal solution and is dual feasible. Now execute Act 1. This gives all other dual feasible solutions, i.e., on sheets G_{gq} (if any), at vertex P_{go}.

4) Execute Act 4. This generates an f-list.

5) Sort the f-list in descending order on Δz. (This is optional.) Assign f-indices to items of the resulting list. ($f \neq 0$, already used.) If the list is empty, then no feasible vertices exist in PHS. (This could conceivably happen if one constraint was itself $cx \geq z_o$.) Else, for each f, make the indicated change of basis and call this B_{fo}, at vertex P_{fo}. This solution is itself an adjacent solution to an optimal solution and is primal feasible. Now execute Act 3. This gives all other primal feasible solutions at vertices P_{fp} on the same sheet G_{fo}.

Note: If in step 3, two or more successive g-indices have the same value of z, these can be doubly indexed as gp. Thus p runs through all sheets of slice g. The solutions are then identified by gpq. If in step 5, two or more successive f-indices have the same value of z, these can be doubly indexed as fq. Thus q runs through all sheets of slice f. The solutions are then identified by fqp.

Resolving a Revised Model from an Old Basis

Suppose we have an old basis B_1 for a revised model and, from other considerations, we are nearly sure $z_1 \geq z_o$, where these are values for the revised model at B_1 and optimality. Four situations can occur.

A. β_1 and γ_1 are both feasible. Solution is optimal.

B. β_1 is feasible, and γ_1 is infeasible. We were wrong and $z_1 < z_o$. We can continue with Phase 2 of the simplex method. However, maybe we were not completely wrong. If we can find a B_2 adjacent to B_1 such that β_2 is infeasible and γ_2 is feasible, then by Theorem 5, we can determine a B_o immediately from the last γ_k which changed to a feasible value in moving from B_1 to B_2.

C. β_1 is infeasible and γ_1 is feasible. We were right but the solution is not optimal. We can continue with the dual algorithm. However, maybe we are closer to the solution than we know. If we can find a B_2 adjacent to B_1 such that β_2 is feasible and γ_2 is infeasible, then we can again use Theorem 5.

D. β_1 and γ_1 are both infeasible. This is most likely unless the revision was a simple change in the rim (c or b). We could apply Phase 1 of the simplex method to get primal feasible and then Phase 2 to get optimal, which is the standard approach. But maybe we can do better.

The use of Theorem 5 in situations B and C above is not likely to be very helpful since the B_o that is found is adjacent to B_1 anyway. However, in situation D above, if our guess is right that $z_1 > z_o$ (it could happen that $z_1 = z_o$ even though doubly infeasible), then Phase 1 and Phase 2 go from DHS to PHS and back to G_o. It would be helpful to know how we are infeasible. We might, for example, be on an infeasible sheet of G_o outside the bounds of F, in which case we are fairly close to optimal. If this were true, then we should look for both P_1 and G_1 to be ambiguous, in complementary ways. More precisely,

(1) some β_i is at a limit, some d_j (or π_i) is infeasible, $\alpha_j^i \neq 0$ and $\frac{d_j}{-\alpha_j^i} > 0$;

(2) some β_i is infeasible by $\delta_i (= \beta_i$ if < 0, $= \beta_i - \underline{L}_i$ if $\beta_i > \underline{L}_i)$, some d_j (or π_i) is zero, $\alpha_j^i \neq 0$ and $\frac{\delta_i}{\alpha_j^i} > 0$.

Note that (1) can occur for situation B, and (2) for situation C. It may be worth two BTRANs, two pricing passes, and an FTRAN to check this out. It can be done as follows.

1) Select $i = r$ by max $|\delta_i|$. If all $\delta_i = 0$, set $r = 0$. Compute π and, if $r \neq 0$, then π^r.

2) Price the entire matrix for two columns, A_s using π and, if $r \neq 0$, then A_t using π and π^r. (If $r = 0$, A_t is ignored.) A_s is selected normally, i.e., largest

dual infeasibility. A_t is selected by dual pricing modified as follows,

a) If $\delta_r < 0$, then $\lambda_t \alpha_t^r < 0$ and $\varphi_t \geq 0$ are required. Hence $\lambda_t d_t \geq 0$ required, i.e., dual feasible.

b) If $\delta_r > 0$, then $\lambda_t \alpha_t^r > 0$ and $\varphi_t \leq 0$ are required. Hence $\lambda_s d_s \geq 0$ is required, i.e. dual feasible.

Thus A_s is selected among dual infeasibilities and A_t among dual feasibilities with the sign requirement on the denominator.

If either A_s or A_t is found, FTRAN them to α_s and α_t.

3) a) If no A_s is found, the solution is dual feasible. If $r = 0$, the solution is also optimal. If no A_s is found but $r \neq 0$, revert to dual algorithm starting with A_t.

b) If no A_t is found and $r \neq 0$, there is no feasible solution since $|\delta_r|$ cannot be reduced.

c) If A_t is found but $\varphi_t \neq 0$, then:

(i) if no A_s is found, continue with dual algorithm; or

(ii) if no A_s is found, continue but set flag for step 6 below.

d) If A_t is found and $\varphi_t = 0$, continue.

4) If here, then A_s should exist (or else the procedure terminated or reverted to the dual algorithm). Compute θ_s. If θ_s does not exist, then the solution must be primal feasible and therefore unbounded. If θ_s exists but is not zero, continue but set flag for step 6 below. Else let $i = q$ be the row on which θ_s won (cannot be a change of bound since $\theta = 0$).

5) Let $\theta_t = \dfrac{\beta_r}{\alpha_t^r}$. ($\alpha_t$ exists or should not be here.)

Compute the reduction in primal infeasibility if A_t replaces r-th basic vector. Call this ΔF, negative if there is a reduction, or positive if there is an increase.

6) Compute π^q(BTRAN). Now do another pseudo-pricing

pass to compute the reduction in dual infeasibility if A_s replaces q-th basic vector. Call this ΔG, negative if there is reduction, or positive if there is an increase.

FLAG: Step 6 is omitted if either $\varphi_t \neq 0$ or $\theta_s \neq 0$, and decisions are made as follows:

a) if $\theta_s = 0$, and $\varphi_t \neq 0$, use A_s (as in primal algorithm);
b) if $\theta_s \neq 0$, and $\varphi_t = 0$, use A_t (as in dual algorithm);
c) if $\theta_s \neq 0$, and $\varphi_t \neq 0$, select A_s or A_t by choosing the min $(|\theta_s d_s|, |\varphi_t \delta_r|)$ which changes z as little as possible;
d) if $\theta_s = \varphi_t = 0$ but both ΔF and ΔG are nonnegative, revert to Phase 1;
e) else, choose $\min(\Delta F, \Delta G)$ and make the corresponding change of basis.

7) Repeat steps 1 to 6 above as long as necessary or possible unless 6(c) repeats. Then revert to Phase 1.

Note: The reduction in dual infeasibility in step 6 is carried out as follows:

Let $d_j = \pi A_j$, and $d_j^r = \pi^r A_j$. Compute the amount of infeasibility in d_j (possibly 0) and the amount in $d_j + d_j^r$ (possibly 0), using absolute values; the second minus the first is the change. Sum over all j (including i).

Examples: (1) X_j at 0. $d_j = -3$. $d_j + d_j^r = 5$. Reduction is $|0| - |-3| = -3$.

(2) X_j at L_j. $d_j = 2$, $d_j + \alpha_j^r = 4$. Increase is $|4| - |2| = 2$.

Special Models with a Symmetric or Nonsingular A-Matrix

A. Maximally Independent Columns

Suppose m = n. Let w be a column of all ones and ´ denote transpose. Now set $b = Aw$, and $c = W'A$. Suppose A is also nonsingular. Then an optimal solution is

$$Aw = b \quad , \qquad u = 0$$

$$w'A = c \quad , \qquad d = 0 \quad .$$

But this is the only solution. For suppose there is an adjacent vertex with e_s replacing A_r. Then to maintain primal feasibility we must have $\alpha_s^r > 0$ so that

$$\theta = \frac{1}{\alpha_s^r} > 0$$

but to maintain dual feasibility we must have $\alpha_s^r < 0$ so that

$$\varphi = \frac{1}{-\alpha_s^r} > 0 \quad .$$

Obviously this is impossible.

Now suppose A is singular, let us say that only m-1 of the columns are independent. Without loss of generality we can assume that a basis B contains $A_1, \ldots, A_{m-1}$ and e_m. Then if $B^{-1}A_m = \alpha_m$, $\alpha_m^m = 0$. We still know a feasible primal solution, $Aw = b$, and a dual feasible solution, $w'A = c$, so that a basic optimal solution must exist. Furthermore, $z_{max} = w'Aw$. The representation of column A_m in terms of the basis is the following:

$$\sum_{j=1}^{m-1} A_j \alpha_m^j - A_m = 0$$

$$\sum_{j=1}^{m-1} A_j + A_m = b \quad .$$

If all $\alpha_m^j \geq -1$, we can add the equations to get

$$\sum_{j=1}^{m-1} A_j (1 + \alpha_m^j) = b$$

with all $\beta_j = 1 + \alpha_m^j \geq 0$. If some $\alpha_m^j < -1$, let $\alpha_m^r = \min\{\alpha_m^j\}$ and replace A_r with A_m in B. Then

$$\beta = w - \frac{1}{\alpha_m^r} \alpha_m \geq 0 \quad , \qquad x_r = 0 \text{ and is nonbasic} \quad .$$

We can assume this has been done so that A_m is not in B, but e_m is in B. Hence we have a feasible basic solution.

Now since A_m is a linear combination of the other A_j, A^m is a linear combination of the other A^i. Since A^m is itself in the basis except for a_m^m, we have

$$\sum_{i=1}^{m-1} A^i \pi_i^m + A^m = 0$$

$$\sum_{i=1}^{m-1} A^i + A^m = c \quad .$$

If all $\pi_i^m \leq 1$, we can subtract the first equation from the second and have $\pi_i = 1 - \pi_i^m \geq 0$. If not, let $\pi_s^m = \max\{\pi_i^m\}$ and replace e_m with e_s. Then

$$\pi = w' - \frac{1}{\pi_s^m} \pi^m \geq 0 \quad , \qquad \pi_s = 0 \text{ and is dual nonbasic.}$$

Note that this does not affect β since $\beta_m = u_m = 0$. Hence, the simplex method will put m-1 linearly independent columns in the basis. It is easy to show that, in fact, any set of m-1 columns constitute an optimal basis with an appropriate e_i if normal simplex rules are followed. Hence one can explore the multiple vertices on G_o (which is single sheeted but highly ambiguous) to adjust the x_j to a desirable pattern.

Multiple applications to reduced systems show that the process works for any rank of A. In fact it is just an application of Theorem 1 of the simplex method applied to both the primal and dual problems.

B. Symmetric A-Matrices

Suppose A is not only square but symmetric. Also assume that A is nonsingular and $c = b'$. Then if the entire A is the basis,

$$z = \pi Ax = x'Ax \quad .$$

If $x = A^{-1}b \geq 0$, then this is an immediate solution. All we have is an unconstrained optimization problem except for $x \geq 0$ and, if this is automatically satisfied, we are through. Suppose x has negative elements; then we are doubly infeasible. For, suppose $x_r < 0$, then $\pi_r < 0$ also, since $\pi_r = x_r$ (assuming normal ordering of the basis A). Now if $\pi_r^r < 0$, then bringing in u_r gives $\theta = \frac{x_r}{\pi_r^r} > 0$; but $\varphi = \frac{\pi_r}{-\pi_r^r} < 0$ so d_r becomes infeasible. If $\pi_r^r > 0$, then $\varphi > 0$ but $u_r < 0$. Suppose x_r is the only infeasibility and we bring in u_r feasibly. After the change of basis, $\alpha_r^r < 0$, $d_r < 0$, $x_r = \pi_r = 0$, and $u_r > 0$. If we now try to bring in x_r again it must replace some other x_q for which $\alpha_r^q > 0$. One must exist for otherwise x_r is unbounded and $z \to \infty$. But $x'Ax$ without constraints is an upper bound. Before the change of basis $\pi_r^q > 0$ already since pivoting on a negative element does not change signs elsewhere in the column. So we should have brought in u_q in the first place.

If we had to bring in u_r infeasibly, then we must replace it with some u_s for which $\pi_s^r < 0$. (This must also exist.) But before the pivot, it had an opposite sign (row divided by negative pivot) so we should have brought in u_s in the first place. In such a situation it is difficult to know, a priori, whether one should use the primal or dual algorithm. (Good old Phase 1 will always work.) But, in any event, symmetry cannot be maintained.

If A is singular, and we attempt to invert it, we will be stopped at some point for lack of pivots. In this case we cannot assume that e_r is out because A_r is in. But, in general, symmetry can be maintained only if a nonsingular, symmetric submatrix gives feasible values automatically. In effect, we divide the system into three parts:

$\pi_1 A \geq b_1'$ for A_i out

$Ax_2 \leq b_2$ for e_i in

double equality for A_i in and e_i out

C. The Matrix $I + HH^T$

Lemma 10: Let H be any real p x q matrix. Then both $I_p + HH^T$ and $I_q + H^TH$ are nonsingular.

Proof: Let $J = I_p + HH^T$ and $K = I_q + H^TH$. It will be sufficient to prove that J is nonsingular since K will follow from symmetry. Let x be any real p x 1 column and $y = H^Tx$. Then y is also real. Suppose $Jx = 0$. Then, a fortiori, $x'Jx = 0$. But $x'Jx = x'x + y'y$ and both terms are nonnegative. Hence

$$x'x = y'y = 0$$

and hence $x = 0$. Therefore the columns of J are linearly independent, i.e., J^{-1} exists.

Lemma 11: Let H be any real p x q matrix. Then the matrix

$$S = \begin{bmatrix} I_p & H \\ H^T & -I_q \end{bmatrix}$$

is nonsingular.

Proof:

$$S = \begin{bmatrix} I_p & -H \\ & I_q \end{bmatrix} \begin{bmatrix} I_p + HH^T & \\ H^T & -I_q \end{bmatrix}.$$

The first factor is clearly nonsingular and, by Lemma 10, so, is the second
For the S of Lemma 11,

$$SS = \begin{bmatrix} J & 0 \\ 0 & K \end{bmatrix}$$

and so

$$S^{-2} = \begin{bmatrix} J^{-1} & \\ & K^{-1} \end{bmatrix}$$

and

$$S^{-1} = S^{-2}S = \begin{bmatrix} J^{-1} & J^{-1}H \\ K^{-1}H^T & -K^{-1} \end{bmatrix}$$

but

$$S^{-1} = SS^{-2} = \begin{bmatrix} J^{-1} & HK^{-1} \\ H^TJ^{-1} & -K^{-1} \end{bmatrix} .$$

Hence

$$J^{-1} = HK^{-1} \quad .$$

Multiplying on the right by H^T

$$J^{-1}HH^T = I_p - J^{-1} = HK^{-1}H^T$$

or

$$J^{-1} = I_p - HK^{-1}H^T \quad .$$

Similarly, multiplying on the left by H^T gives

$$K^{-1} = I_q - H^T J^{-1} H \quad .$$

These relationships are apparently rediscovered frequently. Jewell[2] gives various references. He states that the existence of one inverse implies the existence of the other and both exist when either does. But, <u>both always do</u>, which seems to have escaped notice.

Suppose for some given b and c, we want to maximize cx subject to the following:

$$x \geq 0$$

$$y = Hx$$

$$x + H^T y \leq b \quad .$$

If $c = b'$ and we ignore $x \geq 0$, then $x = K^{-1}b$ gives

$$z = \pi Ax = x'Kx = x'x + y'y = \max$$

$$y = Hx$$

$$x + H^T y = Kx = b \quad .$$

One interpretation might be, for example:

x is a change in independent variables X;

y is a change in dependent variables Y;

H is the matrix of partials $\frac{\partial Y_i}{\partial X_j}$ at X;

---

[2] W.S. Jewell, *Two Classes of Covariance Matrices Giving Simple Linear Forecasts*, RM-75-17, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975.

b is some upper limit, as shown, on the combined changes. If $K^{-1}b \geq 0$, we are satisfied. If not, and we solve the implied LP problem and assuming that it is feasible, we get some π and x such that

$$x \geq 0 \quad , \qquad \pi \geq 0$$

$$y = Hx$$

$$x + H^T y = Kx \leq b$$

$$\pi K \geq b'$$

$$z = \pi Kx = \pi x + (\pi H^T) y = \max \quad .$$

The strange looking term $H^T y$ might arise as follows. Suppose we want to limit the sum of squares of changes in Y, i.e., $y'y$, to no more then some multiple μ of the sum of squares of changes in X, i.e., $x'x$, assuming some average change ξ_j for x_j. Let b be a column of constants $(1 + \mu)\xi_j$. Then

$$x'x + y'y \leq x'b$$

is implied by the constraint

$$x + H^T y \leq b$$

since

$$\pi Kx \leq x'Kx \quad .$$

D. Eliminating the Null Space

Suppose we have a square matrix A which is singular and we want to extract the "best" set of linearly independent columns. Observe first that, if we attempt to invert A, we will be stopped after p basis changes from I because of a lack of pivots. We can assume (I,A) has transformed to the following, where q = m - p

$$\begin{array}{|c|c|c|c|}\hline A_p^{-1} & 0 & I_p & \alpha_q \\ \hline \alpha^q & I_q & 0 & 0 \\ \hline \end{array} \leftarrow \begin{array}{|c|c|c|c|}\hline I_p & 0 & A^p & A_q \\ \hline 0 & I_q & A^q & A_q^q \\ \hline \end{array} \quad .$$

Here

$$\alpha_q = A_p^{-1} A_q \quad , \qquad \alpha^q = -A^q A_p^{-1}$$

and

$$\alpha^q A^q = 0 \quad .$$

If we form the q columns

$$\xi_q = \begin{array}{|c|}\hline \alpha_q \\ \hline -I_q \\ \hline \end{array}$$

these are linearly independent and span the null space of the transformation A. In other words, these are eigenvectors for the q eigenvalues of 0. For

$$A\xi_q = 0 \cdot \xi_q = 0 \quad .$$

The trouble is, there are $\binom{m}{q}$ ways of selecting the null space. We should like to do so in sucn a way that the columns in A_p are as "linearly independent as possible", in other words so that the columns in A_q are as similar as possible, and the rows in A^q are as similar as possible; or better, the columns of A_p are as orthogonal as possible. It is perhaps worth pointing out that the 0 block below α_q will not compute identically zero but will be "noise". This means that the exact value of p may be uncertain. Furthermore, I_p may not be clean.

Suppose we had the best choice and formed the matrix

$$Q = \begin{bmatrix} I_p & \alpha_q \\ 0 & -I_q \end{bmatrix} .$$

Then, noting that $Q^{-1} = Q$,

$$Q^{-1}AQ = \begin{bmatrix} I_p & \alpha_q \\ 0 & -I_q \end{bmatrix} \begin{bmatrix} A_p & A_q \\ A^q & A_q^q \end{bmatrix} \begin{bmatrix} I_p & \alpha_q \\ 0 & -I_q \end{bmatrix}$$

$$= \begin{bmatrix} A_p + \alpha_q A^q & A_q + \alpha_q A_q^q \\ -A^q & -A_q^q \end{bmatrix} \begin{bmatrix} I_p & \alpha_q \\ 0 & -I_q \end{bmatrix}$$

$$= \begin{bmatrix} A_p + \alpha_q A^q & A_p \alpha_q + \alpha_q A^q \alpha_q - A_q - \alpha_q A_q^q \\ -A^q & -A^q \alpha_q + A_q^q \end{bmatrix} .$$

Since $A_p \alpha_q = A_q$, $A^q \alpha_q = A_q^q$, the upper right term is

$$A_q + \alpha_q A_q^q - A_q - \alpha_q A_q^q = 0$$

and the lower right term is 0. Hence

$$Q^{-1}AQ = \begin{bmatrix} A_p + \alpha_q A^q & 0 \\ -A^q & 0 \end{bmatrix} .$$

Let $\underline{A} = Q^{-1}AQ$ and $\underline{A}_p = A_p + \alpha_q A^q = A_p + A^{-1}A_q A^q$. There must be another nonsingular matrix P such that

$$P^{-1}\underline{A}P = \begin{bmatrix} D_p & 0 \\ 0 & 0 \end{bmatrix}$$

where D_p is a diagonal of the nonzero eigenvalues. Furthermore, P must be of the form

$$P = \begin{bmatrix} P_1 & \\ P_2 & I \end{bmatrix} , \qquad P_1 \text{ nonsingular} \quad .$$

Then

$$P^{-1}\underline{A}P = \begin{bmatrix} P_1^{-1}\underline{A}_p P_1 & 0 \\ -(P_2 P_1^{-1}\underline{A}_p - A^q)P_1 & 0 \end{bmatrix}$$

where the lower left corner is zero. Since P_1 and $\underline{A}_p$ are both nonsingular,

$$P_2 P_1^{-1}\underline{A}_p = A^q$$

$$P_2 = A^q \underline{A}_p^{-1} P_1 \quad .$$

Finding P_1 is a standard eigenvector problem which we will not pursue. However, note that all the above manipulations depend heavily on errors introduced by the two matrix products

$$\alpha_q A^q = A_p^{-1} A_q A^q$$

$$A^q \alpha_q = A^q A_p^{-1} A_q \quad .$$

The first is p x p and the second is q x q; both have the same rant $r \leq \min(p,q)$. However, the assumption that A_p is the best choice means that the rows of A^q and the columns of A^q are the worst choices and the rank r is fuzzy. Moreover, if we imbed A in an LP model, the rows A^q are the ones with zero values of π and the columns A^q are the ones with zero values of x so they contribute nothing to Z.

If we set up an LP problem which is very easy to solve and rely on the robustness of MPSs, then we should get a fairly good selection for A_p. As in Section A, set $b = Aw$. However, if any $b_i < 0$, multiply the row through by -1. We are free to do this since we do not care about the solution and it will not affect rank. Call the result $\bar{A}$ and $\bar{b}$. Then set $c = w'\bar{A}$. Since $x = w$ is feasible, $z \geq w'Aw$. Since $\pi = w'$ is feasible, $z \leq w'\bar{A}w$. Hence we know z should go to $w'\bar{A}w$, which is computed automatically by applying w' to $\bar{b}$ when forming c. In fact by presetting $-z_0 = -w'\bar{A}w$, we know z should approach zero from below. Now solve

$$\max cx - z_0 \quad \text{, subject to}$$

$$\bar{A}x \leq \bar{b} \quad , \qquad x \geq 0 \quad .$$

Starting with B = I, the solution: $u = \bar{b}$, and $x = 0$ is immediately feasible so there is no Phase 1. Assuming that no row of A is all zeros (which should be discarded a priori), some rows may still sum to zero. (If all rows, or all columns, sum to zero, there is no hope for a solution, and the problem will have to be partitioned or scaled.) Hence zero θ may occur. Dual iterations should be applied to such rows first, as though they were infeasible, in order to bring in $\bar{A}_j$ with negative elements in these positions (which must exist). This should be done until either all such rows are pivoted or no good pivot can be found.

As z approaches 0, all basic u_i (in β) will approach zero and all d_j for nonbasic x_j will approach zero. This means that almost p columns are in the basis. When all these values are ambiguous, i.e., within standard tolerances of zero, the remaining columns can be transformed, i.e., α_q and α_q^q computed. If any element in α_q^q provides a reasonable pivot, the column should be brought in. If not, we can assume p columns are in. The matrix Q is then essentially in hand.

E. Minimizing Sum of Absolute Values of Deviations

For completeness, we describe the use of LP to minimize the sum of absolute values - rather than sum of squares - in fitting a matrix-vector product to a given column. This has been done off and on for many years, sometimes awkwardly.

Suppose we have a matrix A with $m \geq n$, and a column b, and wish to find a column x such that

$$\sum_i |A^i x - b_i| = \min \quad .$$

This can be stated as follows:

$$\text{minimize } w'u + w'v \quad \text{subject to}$$

$$u - v + Ax = b$$

$$u \geq 0 \quad , \qquad v \geq 0 \quad .$$

Note that x is unconstrained. If the MPS can handle free x_j, this is the place to use them. Otherwise, the constraints must be written

$$u - v + Ax - Ay = b$$

$$u \geq 0 \quad , \qquad v \geq 0 \quad , \qquad x \geq 0 \quad , \qquad y \geq 0 \quad .$$

Note that $u'v = 0$ and $x'y = 0$ since not both e_i and $-e_i$ and not both A_j and $-A_j$ can be in the basis. If free x_j are used, then $x - y$ is combined in one vector x. We cannot do this for u and v since free logical variables can never leave the basis. There are advantages and disadvantages to using free x_j and also to not using them.

Using free x_j

a) A-matrix need not be duplicated with a change of sign.

b) Any basic x_j can change sign on any iteration and can enter the basis with either sign.

c) Disadvantage: once x_j is basic, it can never leave even though some other x_k should replace it for a better (i.e., more robust) solution.

Not using free x_j

a) A-matrix must be duplicated with a change of sign.

b) Distinct iteration is required for x_j to change its sign, either $-A_j$ for A_j or vice versa.

c) Advantage: a basic x_j can leave the basis.

We will assume that free x_j are used.

Now the cost coefficients for u create a problem since the cost row is included as a row 0 in most computer codes and $\{e_0 + e_i\}$ does not form an orthonormal set. This is easily fixed by subtracting all rows $i > 0$ from row 0. We then have

$$\text{minimize } 2w'v - \sum_j (\sum_i a_{ij}) \text{ subject to}$$

$$-v + Ax \geq b$$

$$v \geq 0 \quad , \qquad x \text{ free} \quad .$$

Graphically, the model is

$$\begin{array}{ccc|c|c} (\ 0\ ) & (2\,2\ldots2) & (-\Sigma\ldots-\Sigma) & & = \min \\ \hline I_m & -I_m & A & u & = b \\ & & & v & \\ & & & x & \end{array} \quad .$$

The constant $-z_0$ is $-\sum_i b_i$. Noting that this is a minimization, the result will be some $u \geq 0$, $v \geq 0$, and x, and a π satisfying

$$0 \leq \pi_i \leq 2 \quad , \qquad \text{all } i,$$

$$\pi A_j = \sum_i a_{ij} \quad , \qquad \text{all } j \quad .$$

If $m = n$ and A is nonsingular, then $u = v = 0$ and we have an exact solution. If n were greater than m, then the condition on πA_j cannot hold (in general) and the procedure would terminate with an "unbounded solution" declared, or some other nonsense. Using both A and $-A$ would probably lead to trouble also.

The value of z, regarded as a maximization, is

$$z = -\pi v + \pi A x - \sum_i b_i \leq 0 \quad .$$

At an exact solution, $v = 0$, $\pi = w'$, and $z = 0$.

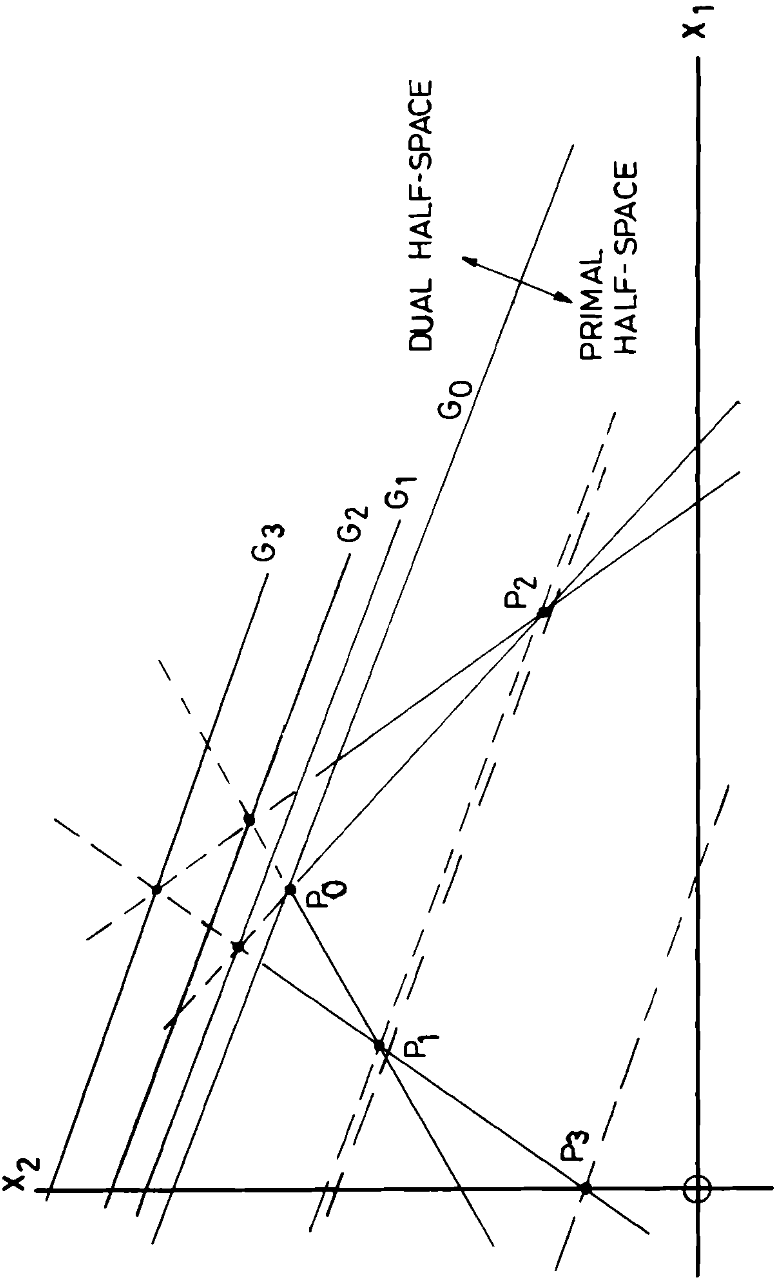
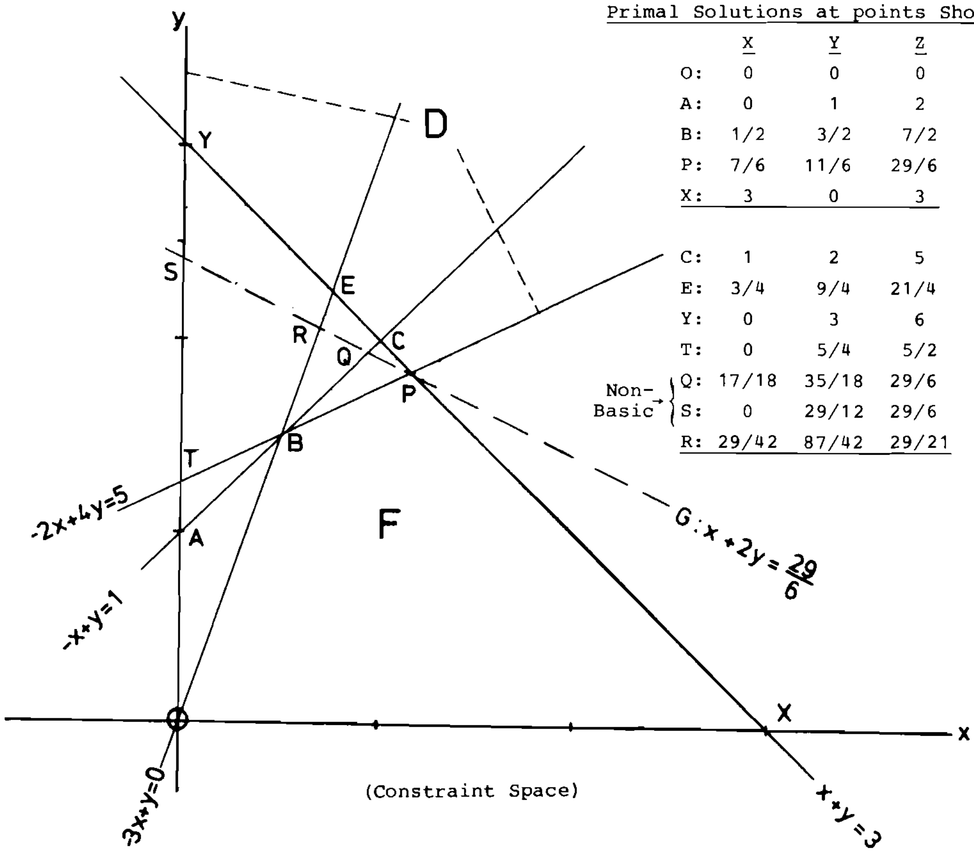


Figure 1.



Primal Solutions at points Shown

	X	Y	Z
O:	0	0	0
A:	0	1	2
B:	1/2	3/2	7/2
P:	7/6	11/6	29/6
X:	3	0	3
C:	1	2	5
E:	3/4	9/4	21/4
Y:	0	3	6
T:	0	5/4	5/2
Non-Basic → Q:	17/18	35/18	29/6
Non-Basic → S:	0	29/12	29/6
R:	29/42	87/42	29/21

max $z = x + 2y$ subject to

$$-x + y \le 1$$
$$-2x + 4y \le 5$$
$$x + y \le 3$$
$$-3x + y \le 0$$
$$x, y \ge 0$$

F = Primal Feasible Simplex:
Vertices: 0,B,P,X
$z_{max} = \frac{29}{6}$ at P

D = Dual Feasible Simplex: the line through S and P plus open domain above.

Dual Problem

min $y_1 + 5y_2 + 3y_3$ subject to

$$y_1 + 2y_2 - y_3 + 3y_4 \le -1$$
$$y_1 + 4y_2 + y_3 + y_4 \ge 2$$

If integer Y_i are required, the solution is Y.

Figure 2a. Illustrated Model.

Primal Feasible

Origin 0

u_0	u_1	u_2	u_3	u_4	x	y	b
1					-1	-2	0
	1				-1	1	1
		1			-2	4	5
			1		1	(1)	3
				1	-3	(1)	0

y replaces u_4 at 0

u_0	u_1	u_2	u_3	u_4	x	y	b	
1				2	-7	0	0	
	1			-1	(2)	0	1	tie
		1		-4	10	0	5	tie
			1	-1	(4)	0	3	
				1	-3	1	0	

$B_{(2)}$ (degeneracy on u_2)

u_0	u_1	u_2	u_3	u_4	x	y	b
1	7/2			-3/2	0		7/2
	1/2			-1/2	1		1/2
	-5	1		(1)	0		0
	-2		1	(1)	0		1
	3/2			-1/2	0	1	3/2

$B_{(3)}$ (degeneracy on u_4)

u_0	u_1	u_2	u_3	u_4	x	y	b
1	-4	3/2		0			7/2
	-2	1/2		0	1		1/2
	-5	1		1			0
	(3)	-1	1	0			1
	-1	1/2		0		1	3/2

Dual Feasible

Y (arbitrarily from 0)

u_0	u_1	u_2	u_3	u_4	x	y	b
1			2		1	0	6
	1		-1		-2	0	-2
		1	-4		-6	0	-7
			1		1	1	3
			-1	1	(-4)	0	-3

E (basic, dual feasible)

u_0	u_1	u_2	u_3	u_4	x	y	b	
1			7/4	1/4	0		21/4	z
	1		-1/2	(-1/2)	0		-1/2	u_1
		1	-5/2	-3/2	0		-5/2	u_2
			3/4	1/4	0	1	9/4	y
			1/4	-1/4	1		3/4	x

C (basic, dual feasible)

u_0	u_1	u_2	u_3	u_4	x	y	b	
1	1/2		3/2	0			5	z
	-1/2		1/2	0	1		1	x
	(-3)	1	-1	0			-1	u_2
	-2		1	1			1	u_4
	1/2		1/2	0		1	2	y

P (optimal)

u_0	u_1	u_2	u_3	u_4	x	y	b	
1	0	1/6	4/3				29/6	z
	0	-1/6	2/3		1		7/6	x
	0	-2/3	5/3	1			5/3	u_4
	1	-1/3	1/3				1/3	u_1
	0	1/6	1/3			1	11/6	y

Figure 2b. Basic Solutions.

Origin 0

u_0	u_1	u_2	u_3	u_4	x	y	b
1					-1	-2	0
	1				-1	1	1
		1			-2	(4)	5
			1		(1)	1	3
				1	-3	1	0

X

u_0	u_1	u_2	u_3	u_4	x	y	b
1			1		0	-1	3
	1		1		0	2	4
		1	2		0	(6)	11
			1		1	1	3
			3	1	0	0	4

P

u_0	u_1	u_2	u_3	u_4	x	y	b	
1		1/6	4/3			0	29/6	z
	1	-1/3	1/3			0	1/3	u_1
		1/6	1/3			1	11/6	y
		-1/6	2/3		1	0	7/6	x
		-2/3	5/3	1		0	5/3	u_4

T (both basic)

u_0	u_1	u_2	u_3	u_4	x	y	b
1		1/2			-2	0	5/2
	1	-1/4			-1/2	0	-1/4
		1/4			-1/2	1	5/4
		-1/4	1		3/2	0	7/4
		-1/4		1	-5/2	0	-5/4

$B_{(1)}$

u_0	u_1	u_2	u_3	u_4	x	y	b
1	0	7/10		-4/5			7/2
	0	1/10		-2/5	1		1/2
	1	-1/5		-1/5			0
	0	-2/5	1	3/5			1
	0	3/10		-1/5		1	3/2

Figure 2c. Short basic, primal feasible path.

S = 1/3T + 2/3Y

1/3(1	0	1/2	0	0	-2	0	) +	
2/3(1	0	0	2	0	1	0	)	
= (1	0	1/6	4/3	0	0	0	)	= (π,d)
1/3{ 5/2	-1/4	0	7/4	-5/4	0	5/4	} +	
2/3{ 6	-2	-7	0	-3	0	3	}	
= {29/6	-17/12	-14/3	7/12	-29/12	0	29/12	}	= {u,x}

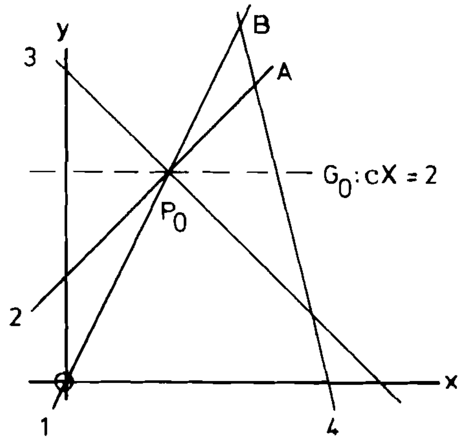
R = 5/21B + 16/21E(B = $B_{(2)}$)

5/21(1	7/2	0	0	-3/2	0	0	) +	
16/21(1	0	0	7/4	1/4	0	0	)	
= (1	5/6	0	4/3	-1/6	0	0	)	Infeasible

(B = $B_{(1)}$)

5/21(1	0	7/10	0	-4/5	0	0	) +	
16/21(1	0	0	7/4	1/4	0	0	)	
= (1	0	1/6	4/3	0	0	0	)	= (π,d)
5/21{ 7/2	0	0	1	0	1/2	3/2	} +	
16/21{21/4	-1/2	-5/2	0	0	3/4	9/4	}	
= {29/6	-8/21	-40/21	5/21	0	29/42	29/14	}	= {u,x}

Figure 2d. Dual-basic, primal-nonbasic solutions.



obj: $\quad y = \max$

1 : $-2x + y \leq 0$

2 : $-x + y \leq 1$

3 : $x + y \leq 3$

4 : $4x + y \leq 10$

B

u_0	u_1	u_3	u_4 (B)	u_4	x	y	β
1	2/3	0		1/3			10/3
	2/3	0		1/3		1	10/3
	-1/6	0		1/6	1		5/3
	-1/2	0	1	-1/2			-2
	-5/6	1		-1/6			-2/3

$0^{(0)}$

u_0	u_1	u_2	u_3	u_4	x	y	β
1					0	-1	0
	1				-2	(1)	0
		1			-1	1	1
			1		1	1	3
				1	4	1	10

A

1	0	4/5		1/5			14/5
	0	4/5		1/5		1	14/5
	0	-1/5		1/5	1		9/5
	0	-3/5	1	-2/5			-8/5
	1	(-6/5)		1/5			4/5

$0^{(1)}$ y in at 0

1	1				-2	0	0
	1				-2	1	0
	-1	1			(1)	0	1
	-1		1		(3)	0	3
	-1			1	6	0	10

$P_o^{(3)}$

1	1/3	0	2/3				2
	1/3	0	2/3			1	2
	-1/3	0	1/3		1		1
	-2/3	1	-1/3				0
	1	0	(-2)	1			4

$P_o^{(1)}$

1	-1	2			0		2
	-1	2			0	1	2
	-1	1			1		1
	(2)	-3	1		0		0
	(5)	(6)		1	0		4

A B

$P_o^{(2)}$ B

1	0	1/2	1/2				2
	0	1/2	1/2			1	2
	0	-1/2	1/2		1		1
	1	(-3/2)	1/2				0
	0	3/2	-5/2	1			4

Figure 3.