Parametric Linear Programming and
Anti-Cycling Pivoting Rules

by

T.L. Magnanti*
and J.B. Orlin**
Sloan School of Management
M.I.T.

ABSTRACT

The traditional perturbution (or lexicographic) methods for resolving degeneracy in linear programming impose decision rules that eliminate ties in the simplex ratio rule and, therefore, restrict the choice of exiting basic variables. Bland's combinatorial pivoting rule also restricts the choice of exiting variables. Using ideas from parametric linear programming, we develop anti-cycling pivoting rules that do not limit the choice of exiting variables beyond the simplex ratio rule. That is, any variable that ties for the ratio rule can leave the basis. A similar approach gives pivoting rules for the dual simplex method that do not restrict the choice of entering variables.

The primal simplex method for minimization problems permits an entering variable at each iteration to be any variable with a negative reduced cost and permits the exiting variable to be any variable that satisfies the minimum ratio rule. As is well-known, any implementation of the procedure is guaranteed to converge if the problem is nondegenerate. In addition, there are two well-known methods for resolving degeneracy. The first of these, the perturbation (or equivalently, the lexicographic) method, avoids cycling by refining the selection rule for the exiting variable (Charnes [1952], Dantzig [1951], Wolfe [1963]). The second method, the combinatorial rule, developed by Bland [1979], avoids cycling by refining the selection rule for both the exiting and entering variables. The situation raises the following natural question: Is there a simplex pivoting procedure for avoiding cycling that does not restrict the minimum ratio rule choice of exiting variables? In this note, we answer this question affirmatively by describing an anti-cycling rule based on a "homotopy principle" that avoids cycling by refining the selection rule for only the entering variable. We also describe an analogous dual pivoting procedure that avoids cycling by refining only the choice of exiting variables.

Our procedures are based upon a few elementary observations concerning parametric simplex methods. These observations may be of some importance in their own right, since they may shed light on some theoretical issues encountered in several recent analyses

2

of average case performance of parametric simplex methods

(e.g., Adler [1983], Borgward [1982], Haimovich [1983], Smale

[1983]). In particular, whenever the probability distribution of a

parametric linear program is chosen, as is frequently the case, so

that the problem satisfies a property that we call dual

nondegeneracy, then the parametric algorithm converges finitely even

if it is degenerate.

Parametric Linear Programming

Consider the following parametric linear programming problem:

$$\text{Minimize} \qquad (c + \theta d)x$$
$$\text{subject to} \qquad Ax = b \qquad\qquad P(\theta)$$
$$x \geq 0$$

where A is an mxn constraint matrix with (for notational

convenience) full row rank. For a given value θ, we say that $P(\theta)$

is nearly dual nondegenerate if for each primal feasible basis B

there is at most one nonbasic variable x_i whose reduced cost

$\bar{c}_i + \theta \bar{d}_i$ is 0. We say that the parametric problem P is dual

nondegenerate if $P(\theta)$ is nearly dual nondegenerate for all $\theta \epsilon R$.

Consider the usual parametric simplex algorithm for solving

$P(\theta)$ for all values of θ, starting with a basis that is optimal for

all sufficiently large values of θ. In the case that P is dual

nondegenerate, we show that the procedure will not cycle (without

any perturbations). We then apply this result to give new primal

simplex pivot rules that (1) are guaranteed to avoid cycling, and

3

(2) rely only on the selection of the entering variable; i.e., any basic variable satisfying the minimum ratio rule may leave the basis.

The following procedure is a version of the usual parametric simplex method as applied to a minimization problem.

<u>Begin</u>

    let B^0 be an optimal basis for $P(\theta)$ for all $\theta \geq \theta^0$;

    let $i=1$;

    <u>while</u> $\bar{d} \not\leq 0$ <u>do</u>

    <u>begin</u>

        let $B = B^{i-1}$;

        let $\bar{c}$, $\bar{d}$ denote the reduced costs for the vectors c and d with respect to basis B;

        select index s so that $-\bar{c}_s/\bar{d}_s = \max\ (-\bar{c}_j/\bar{d}_j : \bar{d}_j > 0)$;

        let $\theta^i = -\bar{c}_s/\bar{d}_s$;

        <u>if</u> $B^{-1}A_s \leq 0$ <u>then</u> quit (since $P(\theta)$ is unbounded from below for all $\theta \leq \theta'$);

        <u>else</u> let B^i be obtained from B by pivoting in x_s and pivoting out any variable chosen by the usual simplex pivot rule (i.e., ties in the ratio rule can be broken arbitrarily);

        let $i = i+1$;

    <u>end</u>.

<u>end</u>.

4

We note that if $\bar{d} \leq 0$ at any point in this algorithm, then $\bar{c} + \theta\bar{d} \geq \bar{c} + \theta^{i-1}\bar{d}$ for all $\theta \leq \theta^{i-1}$ and hence B is an optimal basis for all $\theta \leq \theta^{i-1}$. The iterative modification of θ in the parametric programming procedure can be conceptualized differently as expressed in the following observation.

REMARK 1. Suppose that $B = B^{i-1}$ is an optimal basis for problem $P(\theta^*)$ for some θ^* and let θ^i be selected as in the "while loop" of the parametric algorithm when applied to this basis B (consequently, $\bar{c}$ and $\bar{d}$ are defined by B). Then $\theta^i = \min(\theta: B$ is optimal for $P(\theta))$.

PROOF. If $\theta < \theta^i$, then $\bar{c}_s + \theta\,\bar{d}_s < 0$ and B is non-optimal. Also, by our choice of θ^i,

$$\bar{c}_j + \theta^i\,\bar{d}_j \geq \begin{cases} \bar{c}_s + \theta^i\bar{d}_s & \text{if } \bar{d}_j > 0 \\[2mm] \bar{c}_j + \theta^*\bar{d}_j & \text{if } \bar{d}_j \leq 0. \end{cases}$$

But then each $\bar{c}_j + \theta^i\bar{d}_j \geq 0$, since $\bar{c}_s + \theta^i\bar{d}_s = 0$ and $\bar{c}_j + \theta^*\bar{d}_j \geq 0$ because B is an optimal basis for $P(\theta^*)$. Since $\bar{c}_j + \theta^i\bar{d}_j = 0$ for every variable j corresponding to a column from B, B is optimal for $P(\theta^i)$. ∎

Let B^0 be an optimal basis for $P(\theta^0)$ and let θ^i and B^i for $i \geq 1$ be defined recursively as in the parametric procedure. Moreover, let

$\bar{c}_s^i$ = reduced cost with respect to the basis B^{i-1} of the variable pivoted into B^{i-1} to obtain B^i.

PROPOSITION 1. For all $i \geq 0$,

    (i) B^i is optimal in $[\theta^{i+1}, \theta^i]$;

    (ii)   If $P(\theta)$ is dual nondegenerate, then $\theta^{i+1} < \theta^i$;

    (iii) $\bar{c}_s^i < 0$ for all i such that $\theta^i > 0$.

PROOF: Part (i) is a consequence of our previous remark, the fact that both B^i and B^{i-1} are optimal at θ^i, and the fact that $\{\theta: B^i$ is optimal for $P(\theta)\}$ is an interval. We prove (ii) via a contradiction. Let i be selected so that $\theta^{i+1} = \theta^i$ and either $i = 0$ or else $\theta^i < \theta^{i-1}$. Let x_t be the variable pivoted into basis B^i to obtain B^{i+1} and let x_p be the variable pivoted out of basis B^{i-1} to obtain B^i. Also let $\bar{d}$ be the reduced cost for d with respect to B^i. Then $x_t \neq x_p$ because $\bar{d}_t > 0$ and $\bar{d}_p < 0$ ($\bar{d}_p < 0$ since $\bar{c}_p + \theta^i \bar{d}_p = 0$ and $\bar{c}_p + \theta \bar{d}p > 0$ for $\theta < \theta^i$). Moreover, the assumption $\theta^{i+1} = \theta^i$ implies that $\bar{c}_t + \theta^{i+1} \bar{d}_t = \bar{c}_t + \theta^i \bar{d}_t = \bar{c}_p + \theta^i \bar{d}_p = 0$, contradicting the near dual degeneracy of $P(\theta^i)$.

    To see (iii), note that $\theta^i = -\bar{c}_s^i/\bar{d}_s^i$ and $\bar{d}_s^i > 0$.

This proposition implies that the sequence of pivots generated by the parametric algorithm defines a simplex algorithm for the nonparametric problem minimize $\{cx: Ax = b, x \geq 0\}$. We record this result as follows:

    COROLLARY. Let $B^0, \ldots, B^t$ be a sequence of bases obtained by the parametric algorithm, and let t be first index for which θ^{t+1} is less than 0. Then

(i) B^t is optimal for min $\{cx: Ax = b, x \geq 0\}$;

(ii)   the bases $B^0, \ldots, B^t$ are distinct;

(iii)   the pivot sequence satisfies the usual pivot selection
        rule, viz., the entering variable has a negative reduced
        cost for cost vector c.   ∎

REMARK 2. Note that Proposition 1 remains valid if we replace dual
nondegeneracy with the slightly weaker assumption that the argmax
$(-\bar{c}_j | \bar{d}_j : \bar{d}_j > 0)$ is unique and hence the index s is unique at each
iteration of the parametric algorithm.

An Application to a Primal Pivot Rule

We next show how to apply the previous proposition to define
primal pivoting rules that, without any special provision for
choosing the exiting variable, avoid cycling.

Our previous results demonstrate that the parametric simplex
method defines an anti-cycling pivot sequence for

        minimize   cx

        subject to Ax = b   (P)

                    $x \geq 0$

whenever we can choose the objective function coefficients c and d so
that

(i)   some basis B of A is optimal for all sufficiently large
        values of θ, and

(ii)   P is dual nondegenerate.

To establish these two criteria, let B be <u>any</u> feasible basis for (P)

7

corresponding to the columns of B and with positive components corresponding to the columns of N will satisfy property (i). To ensure that P is dual nondegenerate, however, requires that we avoid ties in computing the entering variables from the parametric ratio test $-\bar{c}_s/\bar{d}_s = \max \{-\bar{c}_j/\bar{d}_j : \bar{d}_j > 0\}$. One natural approach is to use perturbations of c or d. For example, following the lead from the usual (nonparametric) perturbation theory of linear programming, we might perturb the nonbasic columns of c or d by distinct powers of ε for some small $\varepsilon > 0$. Alternatively, we might use a variant of the familiar big M method: choose the nonbasic cost coefficients of d as distinct powers of M for some large constant M. These procedures lead to the following parametric objective functions:

(1) $c + \varepsilon_N + \theta 1_N$,

(2) $c + \theta(\varepsilon_N + 1_N)$, and

(3) $c + \theta(1/\varepsilon)_N$.

In these expressions, ε_N denotes a vector with zero components corresponding to columns in B and with components ε^1, ε^2, ..., ε^{n-m} for all columns in N in, say, their natural order from A. Also $1_N = \varepsilon_N$ with $\varepsilon = 1$, and $(1/\varepsilon)_N = \delta_N$ with $\delta = 1/\varepsilon$. The first two of these objective functions perturb c and d. The third objective function is the polynomial big M method alluded to prior to the last display with the choice $M = 1/\varepsilon$.

For each of these objective functions, for all $0 < \varepsilon < 1$, B is a unique optimal basis for sufficiently large values of θ. Therefore, each one defines an anti-cycling pivot rule if P is dual nondegenerate. To demonstrate this property, we first establish a preliminary result by a modification of the usual perturbation argument.

PROPOSITION 2

Let A_1, ..., A_m be the columns of some basis B. Suppose that
$D = (D_1, D_2, \ldots, D_m)$ is any other basis, possibly containing
some columns of B. Suppose that $A = [B,N]$, that $h = \varepsilon_N$ and
that $\bar{h} = h - h_D D^{-1} A$ (h_D is the subvector of h corresponding to
the columns in D.) Then $\bar{h}_i$ and $\bar{h}_j$ are distinct nonzero
polynomials in ε with zero constant terms whenever i and j are
indices corresponding to distinct columns of A other than
D_1, D_2, ..., D_m.

PROOF:

We use a linear independence argument. First, let us duplicate
the columns of D and consider the following partitioned matrix:

$$M = \begin{bmatrix} 0 & I & C_D \\ B & N & D \end{bmatrix}$$

containing an (n-m) x (n-m) identity matrix I, and a submatrix C_D of
[0,I] corresponding to the columns D from [B,N]. Therefore, each
column of C_D is either a copy of a column of the submatrix 0 or a
copy of a column of the identity matrix I. We observe that the
matrix M has linear rank of n. Let $\underline{\varepsilon}$ denote the vector (ε^1, ε^2, ...,
ε^{n-m}). Let Q = [0,I]. Then $h = \underline{\varepsilon}Q$. Next consider the following
matrix obtained by pivoting on the basis D in M:

$$M' = \begin{bmatrix} \bar{C}_B & \bar{C}_N & 0 \\ D^{-1}B & D^{-1}N & I \end{bmatrix}.$$

Let $Q' = [\bar{C}_B, \bar{C}_N]$. Then Q' has full row rank $n-m$ because M has full row rank n, and M' is obtained from M by pivoting. Moreover, the last m columns of M' are copies of columns of the first n columns. The column of Q' corresponding to each of these "original columns" must be a 0 vector. Deleting these n zero vectors leaves an $(n-m) \times (n-m)$ nonsingular submatrix Q''.

Finally, observe that $\bar{h} = h - h_D D^{-1} A = \underline{\varepsilon} Q - \underline{\varepsilon} C_D D^{-1} A = \underline{\varepsilon} Q''$. Consequently, the two polynomials $\bar{h}_i$ and $\bar{h}_j$ refered to in the statement of the proposition may be obtained from two distinct elements of the vector $\underline{\varepsilon} Q''$ and are thus distinct polynomials in ε with zero constant terms. ∎

Now consider the selection rule for the incoming variable at any point in the parametric simplex method. Assume that the current basis is D and that $\bar{c}$ and $\bar{d}$ are the reduced costs of c and $d = 1_N$ with respect to this basis. As in Proposition 2, let $h = \varepsilon_N$. Then the ratio rule for choosing the incoming variable for the three objective functions (1), (2) and (3) becomes:

(1') $\max \{-(\bar{c}_j + \bar{h}_j)/\bar{d}_j: \bar{d}_j > 0\}$

(2') $\max \{-\bar{c}_j/(\bar{d}_j + \bar{h}_j): \bar{d}_j > 0\}$, and

(3') $\max \{-\bar{c}_j/\bar{h}_j(1/\varepsilon): \bar{h}_j(1/\varepsilon) > 0\}$.

In (3'), $\bar{h}_j(1/\varepsilon)$ denotes the polynomial in $1/\varepsilon$ obtained by replacing ε by $1/\varepsilon$ in $\bar{h}_j$. By the usual purturbation argument, if ε is a sufficiently small constant (i.e., $\varepsilon < \varepsilon(D)$ for some constant $\varepsilon(D) < 1$), then a single index j gives the unique minimum in each of these ratios. Consequently, for all $\varepsilon \leq \min \{\varepsilon(D): D$ is a basis of $A\}$, the

10

entering variable is unique and Proposition 1 (see Remark 2 as well) and its Corollary apply. (Similarly, if we express the reduced cost for D as a function of ϵ and θ, then by Proposition 2, each reduced cost is a different nonconstant polynomial in ϵ and is linear in θ. Therefore, if ϵ is sufficiently small, then the value of θ for which the reduced of x_j is 0 is different for all nonbasic variables j. That is, P is dual nondegenerate.)

Each of the perturbations (1), (2), and (3) lead to different pivot rules that can be interpreted as certain lexicographic selection procedures. Since our purpose in this note has been merely to establish the possibility of simplex point rules that do not restrict the leaving variable, we will not specify the details of these procedures, nor do we discuss their computation requirements (or claim that they are efficient).

In concluding this section, we might note that the parametrics (or homotopies) seem essential for the results given in this paper. Simply perturbing the objective function to avoid ties in the selection of an entering variable will not suffice. For example, in standard examples of cycling in the simplex method, there is no dual degeneracy and the choice of an entering variable is unique.

Dual Pivot Rules

Arguments similar to those used previously apply to right-hand side parametrics in a dual simplex algorithm. That is, consider the parametric problem

$$\text{minimize} \qquad cx$$
$$\text{subject to} \qquad Ax = b + \theta g$$
$$x \geq 0.$$

11

We say that this problem is <u>primal nondegenerate</u> if for any dual feasible basis B (i.e., $c - c_B B^{-1}A \geq 0$) and parameter value θ', there is at most one basic variable x_j whose value is zero. Then, assuming that B^o is an optimal basis for sufficiently large values of the parameter θ and assuming <u>primal nondegeneracy</u>, we can mimic our earlier arguments to show that the usual parametric dual simplex method will

(i)   compute optimal solutions to P'(θ) for all $\theta \leq \theta^o$,

(ii)   for values of $\theta > 0$, defines a dual simplex method

(i.e., at any step (i), the leaving variable x_r

satisfies $\bar{b}_r = [(B^i)^{-1}b]_r < 0$).

Consequently, the primal nondegeneracy assumption results in a finitely convergent dual simplex algorithm that permits any variable satisfying the dual minimum ratio rule to leave the basis. In order to ensure primal nondegeneracy, we can introduce right hand side perturbations; that is consider right hand sides such as the following [here $\underline{\varepsilon}$ is a column vector defined by $\underline{\varepsilon} = (\varepsilon^1, \varepsilon^2, \ldots, \varepsilon^m)$]:

(i) $b + \theta g + \underline{\varepsilon}$, or

(ii) $b + \theta(g + \underline{\varepsilon})$.

Choosing the m-vector g so that $(B^o)^{-1} g > 0$ will ensure that B^o is a unique optimal basis for sufficiently large values for θ and that this perturbation will ensure primal nondegeneracy.

Acknowledgment

References

Adler, I. (1983). "The Expected number of pivots needed to solve parametric linear programs and the efficency of the self-dual simplex method," Technical Report, Department of Industrial Engineering and Operations Research, University of California, Berkeley, California.

Bland, R.G. (1977). New finite pivoting rules for the simplex method. Math. of Oper. Research 2, 103-107.

Borgwardt, K.H. (1982). "The average number of steps required by the simplex method is polynomial," Zeitschrift fur Operations Research 26, 157-177.

Charnes, A. (1952). Optimality and degeneracy in linear programming. Econometrica 20, 160-170.

Dantzig, G.B. (1951). Maximization of a linear function of variables subject to linear inequalities. In T.C. Koopmans (ed.), Activity Analysis of Production and Allocation, John Wiley and Sons, Inc., New York.

Haimovich, M. (1983). "The simplex method is very good! - On the expected number of pivot steps and related properties of random linear programs," Technical Report, Columbia University, New York.

Smale, S. (1983). "On the average number of steps of the simplex method of linear programming," Mathematical Programming 27, 241-262.

Wolfe, P. (1963). A technique for resolving degeneracy in linear programming. J. SIAM 11, 205-211.