

STATIONARITY IN THE SIMPLEX METHOD

NEIL CAMERON

(Received 1 October 1985)

Abstract

Degeneracies occur with increasing frequency in some large scale linear programming problems, but with a simple change to the (revised) simplex method, resulting stationarity of the algorithm can be reduced. The method introduced here may also prevent cycling; neither the lexicographic refinement of Dantzig, Orden and Wolfe nor the perturbation technique of Charnes may be required to prevent cycling.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 90 C 05, 90 C 06, 65 K 05.

1. Introduction

Murtagh [6] draws attention to the increasing occurrence of degeneracies in applications of the simplex method as problems get larger; in particular, massive degeneracies can appear in LP problems derived in certain integer programming algorithms (see Tomlin [7]). Charnes [3] developed a perturbation technique to avoid cycling degeneracy (or ‘circling’ as described by Dantzig [4]), while Dantzig, Orden and Wolfe [5] used lexicographic ordering of vectors (as described in Cameron [2]) to show that cycling could always be avoided theoretically and in practice (even if degeneracy could not). In this paper we show how cycling can be avoided more simply as a by-product of avoiding as much degeneracy as possible.

We consider a solvable linear programming problem (LP)

$$\min f(x) \text{ subject to } Ax = b \text{ and } x \geq 0,$$

in standard minimum form as, for example, in Cameron [2]. Here $f(x) = c^T x$, the cost (column) vector c in \mathbb{R}^n is known, as are the vector b in \mathbb{R}^m and the real m by n matrix A of rank m , where m is less than n . An optimal programme vector x is sought. The ordering in \mathbb{R}^n is the usual entrywise partial vector order: $x \geq 0$ means x_j is a non-negative number, $j = 1$ to n . Suppose A is partitioned as $(B|F)$, where B is an invertible (basis) matrix such that the vector $B^{-1}b \geq 0$; such can be found by solving a first phase auxiliary problem, if necessary. If x and c are partitioned correspondingly, for example

$$x = \begin{pmatrix} x_B \\ x_F \end{pmatrix},$$

then $x_B = B^{-1}b - B^{-1}F x_F$ and the *basic solution* x such that $x_B = B^{-1}b$, $x_F = 0$ is feasible for (LP); geometrically this determines a vertex x of the convex constraint polyhedron C in \mathbb{R}^n . The set C is of course the intersection of the m hyperplanes $A_{i*}x = b_i$, $i = 1$ to m , and the first orthant $x \geq 0$ (by A_{i*} is meant the i th row of A ; A_{i*} is a one by n matrix), and the vertex is *degenerate* if less than m of its entries are positive numbers.

In the simplex method a *test vector* t in \mathbb{R}^n is defined by $t^T = c^T B^{-1}A - c^T$, and if $t \leq 0$, then the solution $x_B = B^{-1}b$, $x_F = 0$ is optimal. If this criterion is not satisfied, then a new basis matrix $B(1)$ is found by replacing one column of B , say a k th column A_{*k} of A , by a column of F , say the l th column of A ; this is done in such a way that $B(1)^{-1}b \geq 0$, and, for the (possibly) new basic feasible solution $x(1)$, the objective value is at least as good as for x : $f(x(1)) \leq f(x)$. The l th column is selected so that t_l is a positive number (and in general additional criteria such as ‘choose the largest t_l ’, as in Cameron [2] or Murtagh [6], appear arbitrary or to have justification more pedagogical than mathematical). After the l th column is chosen, the numbers $\beta_i = (B^{-1}b)_i / (B^{-1}A)_{il}$ are calculated for all $i = 1$ to m , where $(B^{-1}A)_{il} > 0$, and where β_k is chosen to be least among these. Pivoting at the (k, l) entry of $B^{-1}A$, expressed as follows, has the desired effect.

$$(1) \quad B(1)^{-1}b = B^{-1}b + \beta_k(e_k - B^{-1}A_{*l}),$$

$$(2) \quad f(x(1)) = f(x) - \beta_k t_l,$$

$$(3) \quad B(1)^{-1} = B^{-1} + \frac{1}{B_{k*}^{-1}A_{*l}}(e_k - B^{-1}A_{*l})B_{k*}^{-1}.$$

Here e_k is the canonical unit vector in \mathbb{R}^m with k th entry one and with all other entries zero; $B_{k*}^{-1}A$ is the k th row and $B^{-1}A_{*l}$ the l th column of $B^{-1}A$.

2. Degeneracy

From (2) it is immediate that the objective value is unchanged if and only if $\beta_k = 0$. In such a case, (1) gives $B(1)^{-1}b = B^{-1}b$. Since x_k and x_l are each zero in both x and $x(1)$ (although basic and free roles are reversed), and since all other entries are unchanged, $x(1)$ and x are identical. Conversely, if $\beta_k > 0$, then $f(x(1)) < f(x)$, and the objective value at any later basic feasible solution is less than $f(x)$, so x never reappears in the simplex algorithm. Although generally there may be different vertices of the constraint polyhedron with common objective value (in particular there may be different optimal solutions), all vertices defined by any one application of the simplex method have different objective values.

To avoid cycling, rather than use a method such as the lexicographic refinement one could set out to avoid geometrical stationarity or degeneration, that is, to move to another vertex. Such a move often exists even in potentially stationary situations at a degenerate vertex. It is tempting to try this since we thereby avoid the embarrassment of having to update B^{-1} while standing still. (It is in the spirit of trying to reduce the objective value as much as possible that choice of largest t_l is often made; from (2) this is clearly irrelevant if $\beta_k = 0$.) The obvious and simple way to do so is to identify l such that $t_l > 0$, and furthermore such that $(B^{-1}b)_i = 0$ only if $(B^{-1}A)_{il} \leq 0$. For such a choice of l , β_k must be positive, so stationarity is avoided, and not simply the apparently low risk of cycling. If this method is applied in Beale's example, cycling is avoided (see the following example (ii)).

If x defined by $x_B = B^{-1}b$, $x_F = 0$, is not an optimal vertex, then there must exist some basis change to improve $f(x)$. If, whenever $t_l > 0$, there is some k , $1 \leq k \leq m$, such that $(B^{-1}b)_k = 0$ and $(B^{-1}A)_{kl} > 0$, then x is an *essential degeneracy*, and we have to change the basis without changing x . If this procedure is repeated and the lexicographic refinement is applied, then clearly at some first stage x is represented using some basis matrix $B(s)$, say, and there is $l(s)$ such that $t_{l(s)} > 0$, and furthermore such that $(B(s)^{-1}b)_i = 0$ only if $(B(s)^{-1}A)_{il(s)} \leq 0$; pivoting then *moves* x . It would be interesting to find an example of this type where cycling might occur if the lexicographic device is not used. Hoffman's example (see the following example (iii)) is not of this type since there x is an optimal vertex.

There remains the possibility that x is *optimal* but expressed in such a way that the test vector has some positive entries. This is the case of an *essentially degenerate optimal solution*, and changing the basis (applying the lexicographic refinement) will alter the test vector until it is non-negative, thereby identifying x

as optimal. This occurs in the following example (i). Hoffman’s example is also of this type.

3. Example

(i) If many entries in an optimal solution are zero, one might expect some degeneracy leading up to the discovery of the optimal solution (using the test vector criterion), as in the following example.

min x_3 subject to

$$\begin{aligned} x_1 + \frac{2}{5}x_4 - \frac{1}{5}x_5 &= 1, \\ x_2 - \frac{1}{5}x_4 + \frac{4}{5}x_5 &= 2, \\ x_3 - \frac{16}{5}x_4 + \frac{4}{5}x_5 &= 0, \text{ and } x \geq 0. \end{aligned}$$

In this case, using the natural identity matrix I as basis matrix B and calculating t , we find that only t_5 is positive, that $\beta_k = \beta_3 = 0$, and that there is essential degeneracy at $x = (1, 2, 0, 0, 0)^T$ in the basic form expressed. Pivoting at the (3, 5) entry leads to

$$\begin{aligned} x_1 + \frac{1}{4}x_3 + x_4 &= 1, \\ x_2 - x_3 + 3x_4 &= 2, \\ \frac{5}{4}x_3 - 4x_4 + x_5 &= 0, \end{aligned}$$

and the test vector $(0, 0, -1, 0, 0)^T \leq 0$, whence the end-criterion of the simplex method demonstrates the optimal nature of $x = (1, 2, 0, 0, 0)^T$.

(ii) In Beale’s example (see Beale [1] or Exercise 2.5.9 in Cameron [2]) there are initially two positive entries in the test vector t , and the choice of the larger entry t_1 can lead to cycling. (However, even then, if the lexicographic refinement is applied, the algorithm progresses to optimal in two algebraic iterations with the first degenerate, so geometrically such would be a one-step move to the optimal vertex.) If the smaller positive entry t_3 is chosen, there is a uniquely determined rapid path in two iterations, both algebraic and geometric, to the optimal solution. The initial tableau from Beale [1] is

	1/4	-60	-1/25	9	1	0	0	0
	1/2	-90	-1/50	3	0	1	0	0
	0	0	1*	0	0	0	1	1
t^T	3/4	-150	1/50	-6	0	0	0	0

Using the stationary-avoidance criterion of Section 2, we find a unique pivot choice at the (3, 3) entry, as starred, which moves us immediately to a vertex giving reduced objective value. One more (uniquely determined) iteration leads to the optimal solution (satisfying the test vector end-criterion).

(iii) In Hoffman's cycling example (see page 228 of Dantzig [4]), constructed in 1951, effectively $B^{-1}b$ is 0, and $B^{-1}A$ is a 2 by 10 matrix. This determines 0 to be the basic feasible solution. (Another programme component has constant value 1 and plays no part in the simplex algorithm.) From the updating formula (1), clearly the simplex algorithm can in this case only alter the basis, but no change is possible in the feasible solution 0. What we have is an essentially degenerate but *obviously optimal* solution, and there is no need to use the test vector criterion. Hoffman showed that with very particular pivot choices cycling could occur in his example. Does there exist a less trivial example where cycling is possible at an optimal solution, even with the use of the stationary avoidance criterion?

4. Computational aspects

In the revised simplex method (see Cameron [2] or Murtagh [6]) the vector $B^{-1}b$ and matrix B^{-1} are updated using, for example, formulas (1) and (3), respectively, except that to control rounding error *reversion* is needed periodically. From B^{-1} is calculated $c_B B^{-1}$, and from this the new test vector, etc.

In the method proposed here, $B^{-1}b$ is first updated; if all entries are positive, then the vertex is not degenerate, and so we proceed as usual and move to a new vertex. However, if $B^{-1}b$ has zero entries, say for $i = i_1$ to i_r , then, having found a positive t_j , we check $(B^{-1}A)_{ij} = B_{i_*}^{-1}A_{*j}$ for each $i = i_1$ to i_r . If none of these entries is positive, then we proceed as usual, using $j = l$, and move to a *new* vertex. If any are positive, then we move to another positive t_j and repeat the process (at most $n - (m + 1)$ entries in t need to be checked, and many of these are usually negative or zero). If, finally, no suitable column exists, then we accept the *essential degeneracy*, and we pivot, using any of the earlier columns (such that $t_j > 0$), thereby changing the basis but not the vertex.

References

- [1] E. M. L. Beale, 'Cycling in the dual simplex algorithm', *Naval Res. Logist. Quart.* **2** (1955), 269-276.
- [2] Neil Cameron, *Introduction to Linear and Convex Programming* (Cambridge University Press, Cambridge, 1985).
- [3] A. Charnes, 'Optimality and degeneracy in linear programming', *Econometrica* **20** (1952), 160-170.

- [4] G. B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, N. J., 1963).
- [5] G. B. Dantzig, A. Orden and P. Wolfe, 'The generalized simplex method for minimizing a linear form under linear inequality restraints', *Pacific J. Math.* **5** (1955), 183–195.
- [6] B. A. Murtagh, *Advanced Linear Programming: Computation and Practice* (McGraw-Hill, New York, 1981).
- [7] J. A. Tomlin, 'Branch and bound methods for integer and non-convex programming' (J. Abadie (ed.), *Integer and Nonlinear Programming*, North-Holland, Amsterdam, 1970), pp. 437–450.

Department of Mathematics
Monash University
Clayton, Victoria 3168
Australia