Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Convex Simplex Method

Willard I. Zangwill,

To cite this article:
Willard I. Zangwill, (1967) The Convex Simplex Method. Management Science 14(3):221-238. http://dx.doi.org/10.1287/mnsc.14.3.221

MANAGEMENT SCIENCE
Vol 14, No 3, November 1967
Printed in U S A

THE CONVEX SIMPLEX METHOD*†

WILLARD I ZANGWILL‡

The University of California
Berkeley

This paper presents a method, called the convex simplex method, for minimizing a convex objective function subject to linear inequality constraints The method is a true generalization of Dantzig's linear simplex method both in spirit and in the fact that the same tableau and variable selection techniques are used With a linear objective function the convex simplex method reduces to the linear simplex method Moreover, the convex simplex method actually behaves like the linear simplex method whenever it encounters a linear portion of a convex objective function Many of the sophisticated techniques designed to enhance the efficiency of the linear simplex method are applicable to the convex simplex method In particular, as an example, a network transportation problem with a convex objective function is solved by using the standard transportation tableau and by only slightly modifying the usual procedure for a linear objective function

Background

The necessity of developing efficient algorithms for minimizing non-linear, i e , convex, objective functions subject to linear constraints has recently been re-emphasized by Lhermitte and Bessiere [6] They were faced with an investment problem to be solved by mathematical programming However, a key difficulty was the non-linearity of the objective function Two tacks for circumventing this difficulty were taken First, the objective function was linearized, and the problem solved by an appropriate modification of the linear simplex method The linearization, of course, required a significant increase in the number of variables and constraints over the original problem In the second approach the problem was solved directly and to the same degree of accuracy by an appropriate algorithm for a non-linear objective function

To obtain the solution by the linearized version required 10 times the computer time as that required by using the non-linear programming algorithm The precise times were $\frac{1}{2}$ hour for the linearized problem versus 3 minutes for solving the original problem directly

There are several algorithms for minimizing a convex objective function subject to linear inequality constraints As this paper deals with an extension of the linear simplex method, abbreviated LSM, we will only discuss those algorithms which, without excessively stretching the imagination, can be considered as similar to the LSM Some of the most well known and most ingenious of these

* Received October 1966 and revised March 1967

† This research was supported by a grant from the Ford Foundation to the Graduate School of Business Administration, and administered through the Center for Research in Management Science

‡ The author wishes to thank Professor R Van Slyke for his comments on a draft of this paper, and Mr David Rutenberg for an improvement in the example

simplex-like algorithms are the Gradient Projection method I of Rosen [8], the Reduced Gradient method of Wolfe [13], the Generalized Simplex method of Wegner [11], and the Convex Programming method of Abraham [1] The first two of these have been successfully coded for computer Wegner's method, while certainly no less a contribution than the other methods, seems, at least according to its developer, to be mostly of theoretical value [12] Abraham's might also be placed in the latter category

It is of interest to note that in the first three cases the original theoretical convergence proofs given by the developers are incomplete For the second and third methods counter-examples are known [12, 14] And, at this moment, it is an open question whether the first method theoretically converges or not [10], although exploitation of the ideas in [9] seems to ameliorate this difficulty Nevertheless, for those algorithms that have been computer tested, no case of non-convergence has ever been observed At least for the first two algorithms, perhaps there is some form of degeneracy or cycling similar to that in the LSM which hinders the theoretical proofs of convergence And, again by analogy with the LSM, perhaps this cycling never seems to occur in practical examples [Dantzig, p 231, 2] It seems to be an extremely challenging theoretical problem to isolate this form of degeneracy or cycling for the above methods, although the problem does seem related to the jamming or zigzagging phenomenon [Zangwill 16, Wolfe 12] As will be seen later in this paper, a similar form of cycling can arise in the Convex Simplex method, abbreviated CSM, and an appropriate assumption is required to insure theoretical convergence

Of the four methods mentioned above, the CSM is most similar to the Reduced Gradient method Although Rosen was the first to develop a recursive procedure to update constraint data, Wolfe suggested the use of the Simplex tableau and the updating of data by the ordinary Simplex pivot step By letting the tableau define a set of independent variables, and obtaining the reduced costs for the objective function at the point under consideration, he was able to study the variation of the objective function in terms of these variables, and provided a particular scheme which has proven very successful in computational practice As he has pointed out, many variations of this general plan are possible, and we study here another which seems to avoid certain theoretical convergence difficulties which have been noted for the Wolfe procedure In addition, this paper's approach seems slightly closer to the Simplex method in that it actually reduces to the Simplex method whenever the objective function is linear or a linear portion of the objective function is encountered, while the Wolfe method may not so reduce [1]

The CSM might also be compared to the Frank and Wolfe [4] method The Frank and Wolfe method solves a complete linear programming problem at each linearization, while the CSM takes one Simplex step Essentially, the Frank and Wolfe method uses the Simplex method as a subroutine, whereas the CSM attempts to stay within the framework of the Simplex tableau as much as possible

[1] The author wishes to thank Dr P Wolfe for an informative discussion regarding the historical development of his and other methods

Overview of the Convex Simplex Method

The CSM will now be motivated by an informal discussion Following this explanation, an exact mathematical statement of the method will be presented Our attack will be to follow the LSM as far as it will go, and modify it only as the convexity of the objective function requires Let the problem to be solved be stated as

$$\text{(P)} \quad \text{Min} \quad f(x) \quad \text{subject to} \quad Ax = b, \quad x \geqq 0,$$

where f is a convex function with continuous first partial derivatives, A is an $m \times n$ matrix, x is an n component column vector, and b is an m component column vector Generally, subscripting a vector denotes a component of the vector, e g , b_i is the i^{th} component of b The prime notation indicates transpose, e g , b'

If f were linear, the problem P would be a linear programming problem Although f is assumed convex, the CSM will also work for f pseudo-convex [Mangasarian, 7] Any x which satisfies the constraints is said to be feasible, while an x which solves problem P is termed optimal The value of f which satisfies the problem is called the optimal value Clearly the optimal value could be $-\infty$

Just as in the LSM, at each iteration a tableau T with elements t_{ij} will be generated When a tableau, matrix, or vector has a particular superscript, the components are assumed to be similarly denoted For example on the, k^{th} iteration of the method the tableau T^k is generated Then t_{ij}^k is the element in the i^{th} row j^{th} column of the tableau generated on the k^{th} iteration

It is assumed without loss of generality that a standard linear programming phase I procedure [2, p 100] has insured that all rows of the matrix A are linearly independent, has generated an initial basic feasible solution x^1, a tableau T^1, and a corresponding right-hand side b^1 such that

$$T^1x^1 = b^1, \quad x^1 \geqq 0 \tag{1}$$

Note that T^1 does not include the row consisting of the relative or reduced costs [2, p 95] The CSM will generate a sequence of tableaux each new one obtained by standard pivot techniques from the previous one It may be assumed that each tableau T is an $m \times n$ matrix Let B_j be the number of the column associated with the j^{th} basic variable, $1 \leqq j \leqq m$ The B_j^{th} column in tableau T is a column of all zeros except for a one in the j^{th} place Since x^1 is a basic solution

$$x_{B_j}^1 = b_j^1, \qquad j \in M,$$

where $x_{B_j}^1$ is a basic variable, and M is the set of the first m positive integers All non-basic variables in the vector x^1 are zero

In the LSM the relative cost vector corresponding to tableau T^1 would now be calculated and either x^1 declared optimal or a non-basic variable increased With a convex objective function the relative cost vector can be calculated using the appropriate gradient of f For a given tableau T let

$$\nabla f(x)_B = (\partial f(x)/\partial x_{B_1}, \partial f(x)/\partial x_{B_2}, \quad , \partial f(x)/\partial x_{B_m})' \tag{2}$$

Then the relative costs of x for the given tableau T are

$$c(x) = (\nabla f(x) - T'\nabla f(x)_B) \tag{3}$$

where $\nabla f(x)$ is the gradient of f evaluated at x The tableau from which the relative costs are calculated will always be clear from context It is convenient if a variable x is superscripted, say x^k, to write the relative cost vector similarly superscripted and without its argument, thus $c^k = c(x^k)$ For tableau T^1 the relative cost vector c^1 is obtained In the LSM the value

$$c_s^1 = \min \{c_i^1 \mid i \varepsilon N\}$$

would be calculated, where N is the set of the first n positive integers Should c c_s^1 be non-negative, x^1 is optimal Otherwise, the variable x_s is increased adjusting only basic variables, and the objective function decreases In the LSM, since the objective function is linear, one is assured that the objective function will decrease as long as x_s is increased If $t_{is}^1 \leqq 0$ for all $i \varepsilon M$, then x_s may be indefinitely increased yielding an unbounded objective function Otherwise x_s is increased until a basic variable, say x_{B_r}, becomes zero The tableau is then transformed by pivoting on t_{rs}^1

With a convex objective function if $c_s^1 < 0$, then increasing x_s and adjusting only basic variables will initially produce a decrease in f, but may, should x_s be increased too far, actually cause f to increase The CSM procedure is to increase x_s either until further increase would no longer decrease f or until a basic variable becomes zero The value of x at which the first of these two events occurs is the point x^2 Under an assumption that will be made later, if x_s can be increased indefinitely without forcing a basic variable to zero, and should f continue to decrease as x_s decreases indefinitely, then f will have an optimal value $-\infty$

The precise procedure is as follows Either $t_{is}^1 > 0$ for some i or not If there exists an i such that $t_{is}^1 > 0$, then increasing x_s must force a basic variable to zero Let z^1 be the value of x obtained by increasing x_s until a basic variable becomes zero The value x^2 is obtained from the formula

$$f(x^2) = \min \{f(\lambda x^1 + (1 - \lambda)z^1) \mid 0 \leqq \lambda \leqq 1\} \tag{4}$$

Since f is continuous on the compact interval between x^1 and z^1, some value x^2 must exist Should there be several values of x^2 satisfying (4), choose any one If $x^2 = z^1$, some basic variable, say x_{B_r}, has been forced to zero Then pivot on element t_{rs}^1 to form tableau T^2 and substitute x_s for x_{B_r} in the basis Start iteration step 2 Should $x^2 \neq z^1$, set $T^2 = T^1$, keep the same basis and start iteration 2

On the other hand, it may be that $t_{is}^1 \leqq 0$, $i \varepsilon M$ Then x_s may be increased indefinitely without driving a basic variable to zero, and increasing x_s will geometrically form a ray emanating from x^1 In this case let $z^1 \neq x^1$ be any point on that ray Try to determine a point x^2 on that ray which minimizes f Specifically

$$f(x^2) = \min \{f(x^1 + \lambda(z^1 - x^1)) \mid \lambda \geqq 0\} \tag{5}$$

If x^2 does not exist then, as will be discussed later, f is unbounded below If x^2 does exist set $T^2 = T^1$, keep the same basis and start iteration 2

Iteration 1 is completed, and iteration 2 commences At this point x^2 with respect to tableau T^2 may have a positive non-basic-variable That is, some non-basic variable, say x_t^2, may be positive For generality, assume this is so The CSM now begins to resemble the LSM for variables with upper bounds [2, Ch 18] Again calculate the relative cost vector

$$c^2 = (\nabla f(x^2) - (T^2)'\nabla f(x^2)_B) \tag{6}$$

Just as with the LSM for upper bounds, if the positive non-basic variable has a positive relative cost factor, then decreasing the variable and adjusting only basic variables will decrease f Similarly, if the variable has a negative relative cost factor, increasing it will decrease f Let

$$c_{s_1}^2 = \operatorname{Min} \{c_i^2 \mid i \in N\}, \quad \text{and} \tag{7}$$

$$c_{s_2}^2 \, x_{s_2}^2 = \operatorname{Max} \{c_i^2 \, x_i^2 \mid i \in N\} \tag{8}$$

If $c_{s_1}^2 < 0$, then x_{s_1} is an excellent choice to increase If $c_{s_2}^2 \, x_{s_2}^2 > 0$ then, since $x_{s_2}^2$ must be positive, x_{s_2} is an excellent choice to decrease If both $c_{s_1}^2 < 0$ and $c_{s_2}^2 \, x_{s_2}^2 > 0$, a rule must be given to select which of x_{s_1} or x_{s_2} to alter, while, should $c_{s_1}^2 = c_{s_2}^2 \, x_{s_2}^2 = 0$, then x^2 is optimal More precisely

If $| c_{s_1}^2 | \geqq c_{s_2}^2 x_{s_2}^2$ and $| c_{s_1}^2 | > 0$, increase $x_s \equiv x_{s_1}$

If $| c_{s_1}^2 | \leqq c_{s_2}^2 x_{s_2}^2$ and $c_{s_2}^2 x_{s_2}^2 > 0$, decrease $x_s \equiv x_{s_2}$

If $| c_{s_1}^2 | = c_{s_2}^2 x_{s_2}^2 = 0$, x^2 is optimal

Ties are broken arbitrarily

Should $x_s = x_{s_1}$, increase x_s until either f no longer increases or a basic variable becomes zero, and continue in a manner precisely the same as for iteration 1 Should $x_s = x_{s_2}$, decrease x_s until either f no longer decreases, x_s itself becomes zero, or a basic variable becomes zero Call the corresponding value of the x variable when the first of these three situations occurs x^3 If at x^3 a basic variable, say x_{B_r}, becomes zero, pivot on t_{rs}^2 to obtain T^3, and let x_s replace x_{B_r} in the basis If at x^3 no basic variable is forced to zero, let $T^3 = T^2$ and keep the same basis Successive iterations continue in a manner precisely analogous to iteration 2

Mathematical Statement of the Algorithm

Our first task is to pose conditions under which a given point x^* is optimal Lemma 1 formulates these optimality conditions by restating the Kuhn-Tucker Theorem [5] A theorem similar to Lemma 1 may be found in [3] In the lemma the following notation is helpful Given any vector, putting the vector in brackets with the subscript i denotes the i^{th} component of the vector, e g, $[b]_i = b_i$

Lemma 1 Let T^* be any linear programming tableau with b^* the corresponding right-hand side so that

$$T^*x = b^*,$$
$$x \geqq 0,$$

if and only if x is feasible Let x^* be a particular feasible point and

$$c^* = c(x^*) = (\nabla f(x^*) - (T^*)'\nabla f(x^*)_B) \tag{9}$$

its corresponding relative cost vector Let

$$c^*_{s_1} = \mathrm{Min}_i\{c_i^* \mid i \varepsilon N\}, \quad \text{and}$$
$$c^*_{s_2}x^*_{s_2} = \mathrm{Max}\,\{c_i^*x_i^* \mid i \varepsilon N\}$$

Then if $c^*_{s_1} = c^*_{s_2}x^*_{s_2} = 0$, x^* is optimal

Proof Observe that the problem

$$\text{(P 1)} \quad \mathrm{Min} \quad f(x)$$
$$\text{subject to} \quad T^*x = b^*$$
$$x \geqq 0$$

is precisely equivalent to Problem P as the constraints in both problems are equivalent

Since x^* is feasible, x^* must satisfy

$$T^*x^* = b^* \tag{10}$$
$$x^* \geqq 0$$

Also since $c^*_{s_1} = c^*_{s_2}x^*_{s_2} = 0$, x^* satisfies

$$\begin{aligned}[-\nabla f(x^*) + (T^*)'u]_i &= 0 \quad \text{if } x_i^* > 0\\ &\leqq 0 \quad \text{if } x_i^* = 0\end{aligned} \tag{11}$$

where $u = \nabla f(x^*)_B$

But (10) and (11) are simply the Kuhn-Tucker conditions that x^* be optimal to Problem P 1 Hence x^* is optimal for Problem P Q E D

Lemma 1 provides sufficient conditions that a point x^* be optimal But, what if these conditions do not hold? Consider first the LSM Then, since the objective function is linear, the objective function may be written

$$f(x^* + \Delta x) = f(x^*) + (c^*)'\Delta x, \tag{12}$$

where Δx is a m component column vector indicating a small increment of x With a convex objective function, the right-hand side of equation (12) represents a linear approximation to f at x^* Since, by assumption, the partial derivatives are continuous, in an appropriate neighborhood of x^* the convex function f actually behaves like its linear approximation Thus, if $c_i^* < 0$ for some i, increasing Δx_i and adjusting only basic variables will decrease f, while if $c_i^* > 0$, decreasing Δ_s and adjusting only basic variables will decrease f Because f has continuous first partial derivatives, we can be sure that f actually decreases for sufficiently small, but definite, changes in Δx_s Recall that all basic variables have zero relative costs Thus if $c_i^* \neq 0$ an appropriate change in Δx_i will decrease f But there is another factor to be considered, that of feasibility The increment Δx_i cannot be changed in a manner to drive any x variable negative Certainly, if $x_i^* = 0$ and $c_i^* > 0$, Δx_i cannot be made negative Furthermore, if the solution is degenerate, that is, a basic variable $x_{B_i} = 0$, it may not be possible to change Δx_i at all in the direction indicated

There is another area in which the non-linearity of the objective function causes complication First assume a feasible point exists, for if it does not the phase I procedure would determine this fact It is then known that with a linear objective function either the objective function is unbounded below, or there exists an optimal point x^{op} such that the optimal value is $f(x^{op})$ With a non-linear objective function, in addition to the above two cases, a third case might arise The optimal value may be finite and yet there not exist an optimal point It is universal in non-linear programming algorithms to make some sort of assumption regarding this third case We call our assumption Assumption A

Assumption A Either the optimal value is unbounded from below, i e , $-\infty$, or the set $P_{x^1} = \{x \mid x \text{ is feasible and } f(x) \leqq f(x^1)\}$ is bounded Here x^1 is any feasible point which, without loss of generality, is assumed to exist

One implication of assumption A is as follows Suppose on iteration k the tableau T^k and point x^k are given Also assume that x_s is to be increased adjusting only basic variables, and that $t_{is}^k \leqq 0$ for all i Thus x_s may be increased indefinitely without loss of feasibility By assumption A we are assured that either the optimal value is $-\infty$, or by increasing x_s a point x^{k+1} will be found such that increasing x_s further will produce no additional decrease in f

The precise steps in the CSM will now be stated

The Algorithmic Procedure

In the procedure any tie may be arbitrarily broken

Initialization Step

An appropriate LSM phase I procedure has generated a basic feasible solution x^1 with corresponding tableau T^1 Go to step 1 of iteration k with $k = 1$

Iteration k

The feasible point x^k and tableau T^k are given

Step 1—Calculate the relative cost vector

(13) Let
$$c^k = (\nabla f(x^k) - (T^k)'\nabla f(x^k)_B),$$
$$c_{s_1}^k = \operatorname{Min}\{c_i^k \mid i \in N\},$$
and
$$c_{s_2}^k x_{s_2}^k = \operatorname{Max}\{c_i^k x_i^k \mid i \in N\}$$

If $c_{s_1}^k = c_{s_2}^k x_{s_2}^k = 0$, terminate x^k is optimal Otherwise, go to step 2

Step 2—Determine the non-basic variable to change

If $| c^k_{s_1} | \geqq c^k_{s_2} x^k_{s_2}$, increase $x_s = x_{s_1}$ adjusting only basic variables

If $| c^k_{s_1} | \leqq c^k_{s_2} x^k_{s_2}$, decrease $x_s = x_{s_2}$ adjusting only basic variables

Step 3—Calculate x^{k+1} and T^{k+1}

There are three cases to consider

Case A x_s is to be increased, and for some i, $t^k_{is} > 0$

Increasing x_s will drive a basic variable to zero Let z^k the x value when that occurs Specifically,

$$z_i^k = x_i^k, \qquad i \varepsilon D - s, \tag{14a}$$

$$z_s^k = x_s^k + \Delta^k, \tag{14b}$$

$$z^k_{B_i} = x^k_{B_i} - t^k_{is}\Delta^k, \qquad i \varepsilon M, \tag{14c}$$

where $D - s$ is the set of indices of the non-basic variables except s, and

$$\Delta^k = x^k_{B_r}/t^k_{rs} = \mathrm{Min}\,\{x^k_{B_i}/t^k_{is} \mid t^k_{is} > 0\} \tag{15}$$

Find x^{k+1} where

$$f(x^{k+1}) = \mathrm{Min}\,\{f(x) \mid x = \lambda x^k + (1 - \lambda)z^k, 0 \leqq \lambda \leqq 1\} \tag{16}$$

If $x^{k+1} \neq z^k$, set $T^{k+1} = T^k$, and go to iteration k with $k + 1$ replacing k Do not change basis If $x^{k+1} = z^k$, pivot on t^k_{rs} forming T^{k+1}, go to iteration k with $k + 1$ replacing k, and replace x_{B_r} with x_r in the basis

Case B x_s is to be increased and $t^k_{is} \leqq 0$ for all i

In this case x_s may be increased indefinitely without driving a basic variable to zero Define z^k the same as in equation (14) except let $\Delta^k = 1$ Then attempt to determine x^{k+1} such that

$$f(x^{k+1}) = \mathrm{Min}\,\{f(x) \mid x = x^k + \lambda(z^k - x^k), \lambda \geqq 0\} \tag{17}$$

If no x^{k+1} exists, terminate The optimal value is $-\infty$ If x^{k+1} does exist set $T^{k+1} = T^k$, go to iteration k with $k + 1$ replacing k and the same basis

Case C x_s is decreased

Determine z^k using equation (14) except defining Δ^k as follows

$$\Delta^k = \mathrm{Max}\,[\Delta_1^k, \Delta_2^k]$$

where

$$\Delta_1^k = x^k_{Br}/t^k_{rs} = \mathrm{Max}\,\{x^k_{B_i}/t^k_{is} \mid t^k_{is} < 0\}$$

and $\Delta_2^k = -x_s^k$ Should $t^k_{is} \geqq 0$, $i \varepsilon M$, let $\Delta_1^k = -\infty$ Here z^k is the x corresponding to the point where, as x_s is decreased, either a basic variable becomes zero or x_s itself becomes zero, whichever occurs first Calculate x^{k+1} using equation (16)

If $x^{k+1} \neq z^k$, or if $x^{k+1} = z^k$ and $\Delta^k > \Delta_1^k$, let $T^{k+1} = T^k$, and do not change basis If $x^{k+1} = z^k$ and $\Delta^k = \Delta_1^k$, obtain T^{k+1} by pivoting on t^k_{rs} and replace x_{B_r} by x_r in the basis

There are some subtleties in the procedure worth noting First of all, if in adjusting x_s to obtain x^{k+1} it turns out that a basic variable, say x_{B_r}, becomes zero, the new tableau T^{k+1} is obtained by pivoting on t^k_{rs} The new basic variables are similar to the old ones except x_s replaces x_{B_r} in the basis Pivoting occurs even if there is a degeneracy and $x^{k+1} = x^k$ If, on the other hand, at x^{k+1} a basic variable does not become zero, tableau T^k with the same basis becomes T^{k+1}

It is also important to observe that the sequence of points generated $\{x^k\}$ produces a monotonic decreasing objective function That is,

$$f(x^k) \geqq f(x^{k+1}) \tag{18}$$

Convergence Proof

A convergence proof for finite termination will be given first

Theorem 1 Assume the procedure terminates on iteration k If the termination occurs in Step 1 so that $c^k_{s_1} = c^k_{s_2} x^k_{s_2} = 0$, then x^k is optimal If the termination occurs in Step 3 of Case B, then under assumption A the optimal value is unbounded from below

Proof If $c^k_{s_1} = c^k_{s_2} \; x^k_{s_2} = 0$, then by lemma 1 x^k is optimal If termination occurs in Step 3 of Case B, the optimal value must be $-\infty$, for otherwise the set P_{x_1} of assumption A would be compact Equation (17) could then be replaced by

$$f(x^{k+1}) = \operatorname{Min} \{f(x) \mid x \varepsilon P_{x_1}, x = x^{k+1} + \lambda(z^k - x^k), \lambda \geqq 0\}$$

The minimization being on a compact set and f being continuous would force x^{k+1} to exist But then termination in Step 3 of Case B would not occur Q E D

The convergence proof should there be an infinite number of iterations is considerably more complicated Since the CSM is precisely the LSM if the objective function is linear, it is evident that the CSM might cycle and not converge Experience indicates that the LSM has never cycled in a practical problem although degeneracies do sometimes occur [1, p 231] We thereby impose the following anti-cycling assumption, abbreviated ACA

The Anti-Cycling Assumption

The algorithm will not cycle

More specifically, two tableaux T^α and T^β are said to be the same if $t^\alpha_{ij} = t^\beta_{ij}$ for all i and j and, in addition, the column numbers associated with basic variables are the same If two tableaux are the same, we write $T^\alpha = T^\beta$, otherwise they are different and $T^\alpha \neq T^\beta$ The ACA then implies the following Assume we are given a tableau T^* and a feasible point x^* and that application of the algorithm generates a sequence x^{*+i} and T^{*+i}, $i = 1, 2,$ The ACA implies that there cannot exist a $j \geqq 1$ such that both $x^* = x^{*+j}$ and $T^* = T^{*+j}$

In the following theorems and lemmas the notation K, perhaps superscripted, will denote an infinite sequence of positive integers

Theorem 3· Assume the CSM generates an infinite sequence of points $\{x^k\}$ Under assumption A and the ACA any cluster point of this sequence is an optimal point Furthermore, if no cluster point exists, the optimal value is unbounded from below

Proof If no cluster point exists the optimal value must be $-\infty$ For if it were not $-\infty$, by assmuption A

$$x^k \,\varepsilon\, P_{x^1} \quad \text{for all} \quad k,$$

since by equation (18) $f(x^k) \leqq f(x^1)$ for all k But P_{x^1} is compact so a cluster point would have to exist

Now assume a cluster point x^∞ exists It must be shown that x^∞ is optimal Let K be the infinite subsequence of the integers such that $x^k \to x^\infty$ for $k \,\varepsilon\, K$ Because there are only a finite number of tableaux, some particular tableau, say T^α, must be repeated infinitely often for $k \,\varepsilon\, K$ Thus, $T^\alpha = T^k$ for $k \,\varepsilon\, K^1 \subset K$ Also, there being only a finite number of variables, a given non-basic variable with respect to tableau T^α will be selected to be changed in a given direction infinitely often Hence, for all iterations $k \,\varepsilon\, K^2 \subset K^1$, the variable x_s is selected to be changed in a given direction That is, x_s is selected either to be increased for all $k \,\varepsilon\, K^2$ or to be decreased for all $k \,\varepsilon\, K^2$

Thus far, an infinite subsequence of the integers has been determined such that for all iterations $k \,\varepsilon\, K^2$, $T^\alpha = T^k$, and x_s is changed in a specified direction Also $x^k \to x^\infty$, $k \,\varepsilon\, K^2$

The proof will now proceed by contradiction The point x^∞ is assumed to be not optimal, and a contradiction will be demonstrated Let the iteration procedure be applied at point x^∞ with respect to tableau T^α Assume that the variable x_s be selected to be altered i e , either increased or decreased, in the same direction as it was for all iterations $k \,\varepsilon\, K^2$ By Lemma 2 which follows this theorem it is known that the iteration procedure would indicate x_s as a candidate to be so altered And since ties may be broken arbitrarily, we may assume without loss of generality that x_s is selected by the iteration procedure Call $x^{\infty+1}$ the point generated by the iteration procedure In the following argument $x^{\infty+1}$ is assumed to exist as only minor modifications in the proof are required should $x^{\infty+1}$ not exist There are two possibilities possibility a) is that $x^{\infty+1} \neq x^\infty$ Then by Lemma 4 which follows this theorem $\lim_{k\to\infty} f(x^k) = -\infty$ But because the objective function is monotonic,

$$-\infty = \lim_{k\to\infty} f(x^k) = \lim_{k\varepsilon K^2} f(x^k)$$

By continuity

$$-\infty = \lim_{k\varepsilon K^2} f(x^k) = f(x^\infty) > -\infty$$

Thus possibility a) cannot occur

Possibility b) is that $x^{\infty+1} = x^\infty$, which may occur if there is degeneracy Should possibility b) occur, by Lemma 3 which follows this theorem, there is a $K^3 \subset K^2$ such that on iteration $k + 1$ for all $k \,\varepsilon\, K^3$, $T^{\alpha+1} = T^{k+1}$, where $T^{\alpha+1}$ is obtained by pivoting from T^α, and also a given non-basic variable, call it x_t, is selected to be changed in a given direction In addition, $\{x^{k+1}\} \to x^\infty$ for $k \,\varepsilon\, K^3$ Again apply the iteration procedure at x^∞, but this time with respect to tableau $T^{\alpha+1}$ and select the variable x^v to be changed Call $x^{\infty+2}$ the new point generated by this iteration There are two possibilities a) $x^{\infty+2} \neq x^\infty$ and b) $x^{\infty+2} = x^\infty$ Just as above, a) cannot occur so that b) must occur Via Lemma 3, there is a $K^4 \subset K^3$ such that on all iterations $k+2$ for $k \varepsilon K^4$, $T^{\alpha+2} = T^{k+2}$ where $T^{\alpha+2}$ is obtained by pivoting from $T^{\alpha+1}$ and a given non-basic variable, call it x_u, is selected to be changed in a given direction In addition, $\{x^{k+2}\} \to x^\infty$, $k \varepsilon K^4$

Applying the iteration procedure at x^∞ with respect to tableau $T^{\alpha+3}$ must, by the same arguments as above, yield $x^{\infty+3} = x^\infty$ Clearly the same argument may be applied to $T^{\alpha+3}$, yielding a $T^{\alpha+4}$, $T^{\alpha+5}$,

It has been shown that by applying the CBM at x^∞ with respect to tableau T^α that the point x^∞ is again generated but a new tableau $T^{\alpha+1}$ is obtained Using the tableau $T^{\alpha+1}$ at x^∞ also generates x^∞ A sequence of tableaux T^α, $T^{\alpha+1}$, , $T^{\alpha+i}$ is obtained by normal application of the CSM, but the point x^∞ is not altered by the iteration procedure Since there are only a finite number of tableaux, there must be some j such that $T^\alpha = T^{\alpha+j}$ But then cycling has occurred The ACA prohibits this, and the contradiction has been demonstrated Q E D

Lemma 2 Assume at iteration k for all $k \varepsilon K$ that $T = T^k$ and that x_s was selected to be changed in a given direction Also let $x^k \to x^\infty$, $k \varepsilon K$ where x^∞ is not optimal Assume the iteration step is applied at x^∞ with tableau T Then x_s will be a candidate for change at x^∞ and the in same direction as the given direction

Proof First assume x_s was selected to be increased for $k \varepsilon K$ Then

$$c(x^k)_s \leqq c(x^k)_i, \qquad i \varepsilon N, \text{ and} \tag{19}$$

$$-c(x^k)_s \geqq c(x^k)_i x_i^k, \qquad i \varepsilon N, \tag{20}$$

by step 2 of the iteration procedure But by equation (3) and the assumption that the partial derivatives of f are continuous, $c(x)$ is continuous in x Since $x^k \to x^\infty$, $k \varepsilon K$,

$$\lim_{k\varepsilon K} c(x^k)_s = c(x^\infty)_s \leqq \lim_{k \varepsilon K} c(x^k)_i = c(x^\infty)_i, \qquad i \varepsilon N, \tag{21}$$

and similarly

$$-c(x^\infty)_s \geqq c(x^\infty)_i x_i^\infty, \qquad i \varepsilon N \tag{22}$$

As x^∞ is not optimal, x_s is a candidate for increase

If x_s were selected for decrease for $k \varepsilon K$, then

$$c(x^k)_s\, x_s^k \geqq c(x^k)_i x_i^k, \qquad i \varepsilon N,$$

and

$$c(x^k)_s\, x_s^k \geqq -c(x^k)_i, \qquad i \varepsilon N$$

Applying the same reasoning as above x_s will be a candidate for decrease at x^∞ with tableau T Q E D

Lemma 3 Assume at iteration k for all $k \varepsilon K$ that $T^\alpha = T^k$ and that x_s was selected to be changed in a specified direction Also let $x^k \to x^\infty$ where x^∞ is not optimal Assume that the iteration procedure was applied at x^∞ with tableau T^α, x_s was selected to be changed in the same direction as the specified direction, and that the point generated by the iteration was $x^{\infty+1}$ Also assume $x^{\infty+1} = x^\infty$ Then under these assumptions there exists a $K^1 \subset K$ such that for all iterations

$k+1$ for $k \varepsilon K^1$, $T^{k+1} = T^{\beta}$ where $T^{\beta} \neq T^{\alpha}$, and the variable x_i is selected to be changed in a specified direction In addition, $\{x^{k+1}\} \to x^{\infty}$, $k \varepsilon K^1$

Proof In the proof assume that the specified direction is to increase x_s The proof for x_s being decreased follows the same reasoning as for x_s being increased, and hence, is omitted

By Lemma 2 since x^{∞} is not optimal, x_s will be a candidate for increase, and because all ties may be broken arbitrarily in the iteration procedure, there is no loss in generality by assuming x_s is selected to be increased at x^{∞} As x^{∞} is not optimal $c(x^{\infty})_s < 0$, so that increasing x_s and adjusting the basic variables will actually decrease f But as $x^{\infty+1} = x^{\infty}$, it must be that x_s could not be increased because to do so would drive a basic variable negative Thus z^{∞}, the z generated by the iteration step, must be equal to x^{∞}

$$z^{\infty} = x^{\infty} \tag{23}$$

Moreover, from equation (14)

$$z_i^{\infty} = x_i^{\infty}, \qquad i \varepsilon D - s, \tag{24a}$$

$$z_s^{\infty} = x_s^{\infty} + \Delta^{\infty}, \tag{24b}$$

and

$$z_{Bi}^{\infty} = x_{Bi}^{\infty} - t_{is}^{\alpha}\Delta^{\infty}, \qquad i \varepsilon M, \tag{24c}$$

where

$$\Delta^{\infty} = \mathrm{Min}\,\{x_{Bi}^{\infty}/t_{is}^{\alpha} \mid t_{is}^{\alpha} > 0\} \tag{25}$$

Note that $t_{is}^{\alpha} > 0$ for some i since increasing x_s drives a basic variable negative Also, since $z^{\infty} = x^{\infty}$,

$$\Delta^{\infty} = 0$$

For $k \varepsilon K$ since $T^{\alpha} = T^{k}$, Δ^{k} is calculated by equation (15) Furthermore, since $x^{k} \to x^{\infty}$, $k \varepsilon K$

$$\Delta^{k} \to \Delta^{\infty} = 0, \qquad k \varepsilon K \tag{26}$$

But then by equations (14) and (24)

$$z^{k} \to z^{\infty} = x^{\infty}, \qquad k \varepsilon K \tag{27}$$

In effect, for $k \varepsilon K$ and k large, it takes only a small increase in x_s to drive a basic variable to zero

Recall the variable x^{k+1} is obtained by increasing x_s until either a basic variable becomes zero or until further increase would no longer decrease f, whichever occurs first Since $c(x^{\infty})_s < 0$ and $c(x^{k})_s \to c(x^{\infty})_s$, $k \varepsilon K$, for k large enough and $k \varepsilon K$, there must be a $\delta > 0$ such that x^{k} can be increased at least an amount δ before further increase would no longer decrease f This follows by the continuity of $c(x)$ But by equation (26) for k large enough $\Delta^{k} < \delta$ Hence, the variable x^{k+1} will be obtained by driving a basic variable to zero for k large enough and $k \varepsilon K$

Thus

$$x^{k+1} = z^k \qquad \text{for } k \text{ sufficiently large}$$

But by (27) $z^k \to z^\infty = x^\infty$ for $k \varepsilon K$ so that

$$x^{k+1} \to x^\infty \qquad \text{for } k \varepsilon K$$

In addition, since at x^{k+1} a basic variable becomes zero, T^{k+1} is obtained from $T^k = T^\alpha$ by a pivot operation, so that $T^{k+1} \neq T^\alpha$ Because there is only a finite number of tableaux and a finite number of variables, there must exist a $K^1 \subset K$ such that for all iterations $k+1$ for $k \varepsilon K^1$, $T^\beta = T^{k+1}$ where $T^\beta \neq T^\alpha$ and a given variable x_t is selected to be changed in a given direction Furthermore, $x^{k+1} \to x^\infty$, $k \varepsilon K^1$ Q E D

Lemma 4 Assume that at iteration k for all $k \varepsilon K$ that $T^\alpha = T^k$ and that x_s was selected to be changed in a specified direction Also let $x^k \to x^\infty$, $k \varepsilon K$ Assume the iteration procedure was applied at x^∞ with tableau T^α, x_s was selected to be changed in the same direction as above specified, and that the point generated by the iteration was $x^{\infty+1}$ If $x^{\infty+1} \neq x^\infty$, then

$$\lim_{k\to\infty} f(x^k) = -\infty$$

Proof The proof will be given assuming for all iterations $k \varepsilon K$ that Step 3 of Case A occurred Cases B and C may be proven similarly Since x_s is being increased and $t^\alpha_{is} > 0$ for some i, z^∞ may be expressed as in equations (24) and (25) (Of course, here unlike Lemma 3 $x^{\infty+1} \neq x^\infty$, and $\Delta^\infty \neq 0$)

Furthermore, as $x^k \to x^\infty$, $k \varepsilon K$,

$$\Delta^k \to \Delta^\infty, \qquad k \varepsilon K, \quad \text{and} \tag{29a}$$

$$z^k \to z^\infty, \qquad k \varepsilon K, \tag{29b}$$

using the same reasoning as in Lemma 3 equations (26) and (27) The point x^{k+1} is obtained by equation (16),

$$f(x^{k+1}) = \text{Min}\,\{f(x) \mid x = \lambda x^k + (1-\lambda)z^k, 0 \leqq \lambda \leqq 1\}, \tag{16}$$

and $f(x^{\infty+1})$ is obtained similarly except that x^∞ replacing x^k and z^∞ replacing z^k in the above formual

Now as $x^k \to x^\infty$ and $z^k \to z^\infty$, $k \varepsilon K$, and since in (16) f is continuous on a compact set,

$$f(x^{k+1}) \to f(x^{\infty+1}), \qquad k \varepsilon K \tag{30}$$

Because x^∞ is not optimal, $c(x^\infty)_s < 0$, so that increasing x_s produces a definite decrease in f Note that x_s is actually increased a positive amount since $x^{\infty+1} \neq x^\infty$ by assumption Thus for some $\epsilon > 0$

$$f(x^{\infty+1}) = f(x^\infty) - 2\epsilon \tag{31}$$

Also, by continuity,

$$f(x^k) \to f(x^\infty), \qquad k \varepsilon K \tag{32}$$

Equations (30), (31) and (32) imply that for k large enough and $k \varepsilon K$, say $k \varepsilon K^1$,

$$(33) \qquad f(x^{k+1}) \leqq f(x^k) - \epsilon, \qquad k \varepsilon K^1$$

$$\text{Lim}_{i\to\infty} f(x^i) = \text{Lim}_{i\to\infty} \left(\textstyle\sum_{k=1}^{i-1} [f(x^{k+1}) - f(x^k)]\right) + f(x^1)$$

$$\leqq \text{Lim}_{i\to\infty} \left(\textstyle\sum_{k\varepsilon K^1, 1 \leqq k \leqq i-1} (f(x^{k+1}) - f(x^k))\right) + f(x^1)$$

and by (33)

$$\leqq \text{Lim}_{i\to\infty} - \textstyle\sum_{k\varepsilon K^1\, 1\leqq k \leqq i-1} \epsilon + f(x^1) = -\infty \qquad \text{Q E D}$$

Some Remarks

a) In step 1 of the iteration procedure the variable x_{s_2}, a good candidate to decrease, is determined by

$$(34) \qquad c_{s_2}^k x_{s_2}^k = \text{Max}\,\{c_i^k x_i^k \mid i \varepsilon N\}$$

It might be reasonable to consider another criterion for selecting x_{s_2}, one that more resembles the ordinary upper bounded LSM criterion, namely,

$$(35) \qquad c_{s_2}^k = \text{Max}\,\{c_i^k \mid x_i^k > 0\}$$

One reason criterion (34) was selected instead of (35) is that (34) weights each c_i^k by a limit on how far x_i may be decreased That is, $x_i > 0$ cannot be decreased an amount more than x_i, or x_i will go negative If $x_i > 0$ is close to zero, (34) considers this fact The second and unquestionably more important reason for using criterion (34) is that Lemma 2 could not be proven with (35) Let $x^k \to x^\infty$, $k \varepsilon K$ and assume for all $k \varepsilon K$ that

$$c_{s_2}^k\, x_{s_2}^k = \text{Max}\,\{c_i^k\, x_i^k \mid i \varepsilon N\} \qquad \text{and}$$

$$c_t^k = \text{Max}\,\{c_i^k \mid x_i^k < 0\}$$

Letting $c_i^k \to c_i^\infty$, $k \varepsilon K$ and using criterion (34) we are assured that

$$c_{s_2}^\infty x_{s_2}^\infty = \text{Max}\,\{c_i^\infty x_i^\infty \mid i \varepsilon N\},$$

but there is no assurance that

$$c_t^\infty = \text{Max}\,\{c_i \mid x_i^\infty > 0\}$$

This point was critical to the proof of Lemma 2

b) At each iteration of the algorithm either formula (16) or (17) has to be evaluated This evaluation is an ordinary 1-dimensional search Note that only $m + 1$ of the x_i change, while the remainder are fixed The search is then considerably easier than if all n variables changed Moreover, even if the function f is highly complicated, the restricted function f, determined when only $m + 1$ of the x_i vary, may be very well behaved, perhaps linear or quadratic For example, if the restricted function is linear, the CSM at that point takes precisely the same step that the LSM would, no complex calculation of x^{k+1} is necessary For a restricted f that is quadratic again the calculation of x^{k+1} is quite easy

c) The ACA was certainly vital in proving convergence Instead of the ACA, other assumptions more akin to the LSM non-degeneracy assumption could have been made The corresponding convergence proofs would be only slightly different from the proof in this paper The ACA was selected instead of a non-degeneracy assumption because of the LSM experience of degenerating, but never cycling, in a practical problem As mentioned previously some technique or assumption is necessary in order to avoid the jamming or cycling phenomenon A special anti-jamming perturbation procedure that will eliminate the ACA is now under development

d) It should be obvious that many of the variants of the LSM also apply to the CSM It is hoped that a large number of the specialized techniques that increase the efficiency of the LSM in specific cases will also apply to the CSM As a case in point consider the following example

An Example

A classical transportation problem will be solved but with a convex objective function The procedure is to apply the CSM to this problem in the same manner as the LSM is applied to the linear transportation problem [2, p 300] The problem to be solved is, for purposes of clarity, quite simple Obviously a computer could solve far more complex problems

$$\operatorname{Min} f(x) = x_{11} + 2x_{12} + x_{13}^2 + x_{21}^2 + 3x_{22} + 2x_{23}^2 + e^{x_{11}x_{21}}$$

$$\begin{array}{cccccccl}
x_{11} + & x_{12} + & x_{13} & & & & = 3 = a_1 \\
 & & & x_{21} + & x_{22} + & x_{23} & = 2 = a_2 \\
x_{11} & & & + x_{21} & & & = 1 = b_1 \\
 & x_{12} & & & + x_{22} & & = 2 = b_2 \\
 & & x_{13} & & & + x_{23} & = 2 = b_3 \\
 & & & x_{ij} \geqq 0 & & &
\end{array}$$

The form of the tableau is

x_{11} $\dfrac{\partial f}{\partial x_{11}}$	x_{12} $\dfrac{\partial f}{\partial x_{12}}$	x_{13} $\dfrac{\partial f}{\partial x_{13}}$	a_1 u_1
x_{21} $\dfrac{\partial f}{\partial x_{21}}$	x_{22} $\dfrac{\partial f}{\partial x_{22}}$	x_{23} $\dfrac{\partial f}{\partial x_{23}}$	a_2 u_2
b_1 v_1	b_2 v_2	b_3 v_3	implicit ↑ ← prices

The tableau is precisely analogous to the tableau for a linear objective function [2, p 310] except that $\partial f/\partial x_{ij}$ has replaced c_{ij}, the linear cost coefficient

In the calculations a superscript B will indicate a variable in the basis The notation, $x = (x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23})$, $z = (z_{11}, z_{12}, z_{13}, z_{21}, z_{22}, z_{23})$, etc, will be used

As an initial basic solution let

$$x^1 = (1, 2, 0, 0, 0, 2)$$

Then

$$\partial f(x^1)/\partial x_{11} = 1, \quad \partial f(x^1)/\partial x_{12} = 2, \quad \partial f(x^1)/\partial x_{13} = 0,$$
$$\partial f(x^1)/\partial x_{21} = 2, \quad \partial f(x^1)/\partial x_{22} = 3, \quad \partial f(x^1)/\partial x_{23} = 8$$

The first tableau is

1^B 1	2^B 2	0 0	3 0
0 2	0^B 3	2^B 8	2 1
1 1	2 2	2 7	

The implicit prices were calculated in precisely the same manner as for the linear transportation problem The reduced costs are

$$c_{ij}(x) = \partial f(x)/\partial x_{ij} - u_i - v_j$$

so that

$$c(x^1) = (0, 0, -7, 0, 0, 0)$$

and letting $c(x)\,x = (c(x)_{11} x_{11}, c(x)_{12}\, x_{12}, \quad , c(x)_{23}\, x_{23})$,

$$c(x^1)\, x^1 = (0, 0, 0, 0, 0, 0)$$

Thus $c^1_{s_1} = c^1_{13} = -7$ and $c^1_{s_2} = 0$ so that we must increase x_{13} For clarity the upper left of the tableau is rewritten

1	$2 - \theta$	$0 + \theta$
0	$0 + \theta$	$2 + \theta$

The θ's indicate the changes to be made The value z^1 is calculated by making θ as large as possible without driving a variable negative The value z^1 is the same value as would be the x^2 with a linear transportation problem Thus $z^1 = (1, 0, 2, 0, 2, 0)$ To find x^2 use equation (16)

$$\begin{aligned} f(x^2) &= \operatorname{Min}\{f(x) \mid x = \lambda x^1 + (1-\lambda)z^1, 0 \leqq \lambda \leqq 1\} \\ &= \operatorname{Min}\{\lambda\, 1 + (1-\lambda)\, 1 + 2(\lambda\, 2 + (1-\lambda)\, 0) + (\lambda\, 0 + (1-\lambda)2)^2 \\ &\qquad + (\lambda\, 0 + (1-\lambda)\, 0)^2 + 3(\lambda\, 0 + (1-\lambda)\, 2) + 2(\lambda\, 2 + (1-\lambda)\, 0)^2 \\ &\qquad + e^{(\lambda\, 1 + (1-\lambda)\, 1)(\lambda\, 0 + (1-\lambda)\, 0)} \mid 0 \leqq \lambda \leqq 1\} \\ &= \operatorname{Min}\{11 - 10\lambda + 12\lambda^2 \mid 0 \leqq \lambda \leqq 1\} \end{aligned}$$

yielding $\lambda = \frac{5}{12}$ and $x^2 = (1, \frac{5}{6}, \frac{7}{6}, 0, \frac{7}{6}, \frac{5}{6})$

Tableau 2 is

$1^B - \theta$ 1	$\frac{5}{6}^B + \theta$ 2	$\frac{7}{6}$ $2\frac{1}{3}$	3 0
$0 + \theta$ 1	$\frac{7}{6}^B - \theta$ 3	$\frac{5}{6}^B$ $\frac{10}{3}$	2 1
1 1	2 2	2 $\frac{7}{3}$	

where in Tableau 2 the partial derivatives are evaluated at x^2 Then

$$c^2 = (0, 0, 0, -1, 0, 0)$$

$$c^2\, x^2 = (0, 0, 0, 0, 0, 0)$$

Thus $c^2_{s_1} = c^2_{21} = -1$ and $c^2_{s_2}\, x^2_{s_2} = 0$ The variable x_{21} is to be increased By use of the θ's we obtain $z^2 = (0, \frac{11}{6}, \frac{7}{6}, 1, \frac{1}{6}, \frac{5}{6})$ From equation (16) it is determined that $x^3 = z^2$ The variable x_{21} becomes basic while x_{11} leaves the basis Tableau 3 is

0 1	$\frac{11}{6}^B$ 2	$\frac{7}{6}$ $\frac{7}{3}$	3 0
1^B 2	$\frac{1}{6}^B$ 3	$\frac{5}{6}^B$ $\frac{10}{3}$	2 1
1 1	2 2	2 $\frac{7}{3}$	

The relative costs are $c^3 = (1, 0, 0, 0, 0, 0)$ Thus

$$x^3 = (0, \tfrac{11}{6}, \tfrac{7}{6}, 1, \tfrac{1}{6}, \tfrac{5}{6})$$

is optimal

This example illustrates quite clearly that except for some side calculations the CSM proceeds like the LSM

References

1 ABRAHAM, J , "An Approximate Method of Convex Programming," *Econometrica*, Vol 29, 1961, pp 700–703

2 DANTZIG, G B , *Linear Programming and Extensions*, Princeton University Press, Princeton, N J , 1963

3 FAURE, P AND HUARD, P , "Resolution de Progremmas Mathematiques a Fonction Non Linéaire par la Method du Gradient Reduit," *Revue Francaise de Recherche Opérationnelle*, 36, 1965, pp 167–206

4 FRANK, M AND WOLFE, P , "An Algorithm for Quadratic Programming," *Nav Res Log Quart* , Vol 3, 1956, pp 95–110

5 KUHN, H W AND TUCKER, A W , "Non-Linear Programming," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J Neyman (ed), University of California Press, Berkeley, 1951, pp 481–493

6 LHERMITTE, P AND BESSIERE, F , "Sur les Possibilités de la Programmation Non Lineaire Appliquée au Choix des Investissements," Actes de la 3e Conférence International de Recherche Operationnelle, Oslo, 1963 (Duond), pp 597–609

7 MANGASARIAN, O L , "Pseudo-Concave Functions," *J SIAM Control*, Ser A, Vol 3, No 2, 1965, pp 281–290

8 ROSEN, J B , "The Gradient Projection Method for Nonlinear Programming Part I Linear Constraints," *J Soc Indust Appl Math* , Vol 8, No 1, March 1960, pp 181–217

9 ROSEN, J B , "The Gradient Projection Method for Nonlinear Programming Part II Nonlinear Constraints," *J Soc Indust Appl Math* , Vol 9, No 4, Dec 1961, pp 514–532

10 ROSEN, J B , Private communication, April 1966

11 WEGNER, P , "A Non-Linear Extension of the Simplex Method," *Management Science*, Vol 7, No 1, Oct 1960, pp 43–50

12 WEGNER, P , Footnote page 158, *Non Linear Mathematics*, by Saaty, T and Bram, J , McGraw-Hill, New York, N Y , 1964

13 WOLFE, P , "An Extended Simplex Method," *Notices of the Amer Math Soc* , Vol 9, No 4, August 1962, p 308 (Abstract)

14 WOLFE, P , "On the Convergence of Gradient Methods Under Constraints," *IBM Research Report RZ-204*, March 1, 1966, IBM Zurich Research Laboratories, Ruschliken, Zurich, Switzerland

15 WOLFE, P , Private communication, August 1966

16 ZANGWILL, W I , "Some Decomposable Large Step Non-Linear Programming Approaches—with an Application to the Dantzig Quadratic Programming Algorithm," Working Paper No 172, Center for Research in Management Science, University of California, Berkeley, June 1966