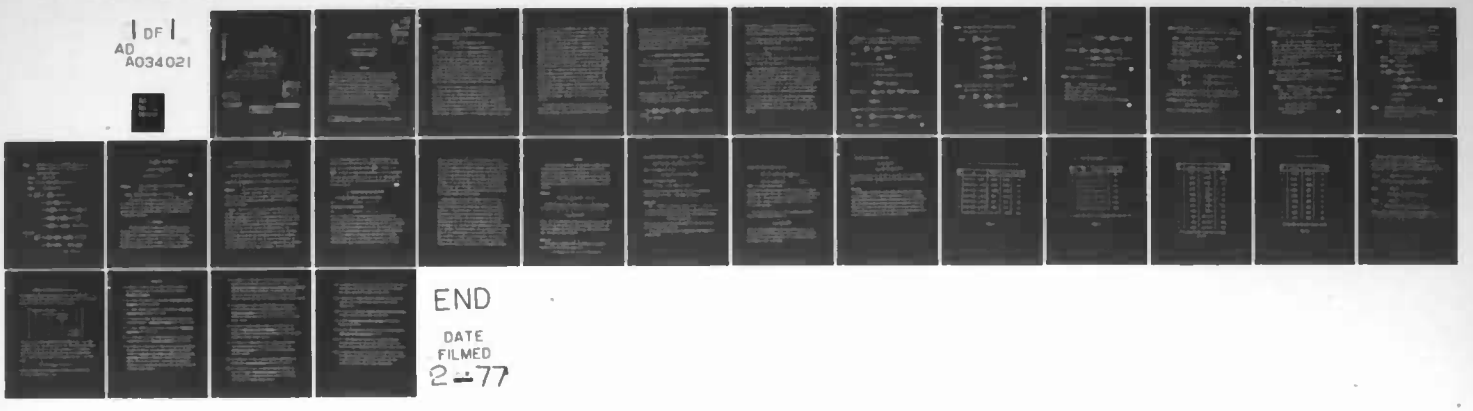
AD-A034 021 CHICAGO UNIV ILL GRADUATE SCHOOL OF BUSINESS F/G 12/1
AN IMPROVED SCALAR GENERATED HOMOTOPY PATH FOR SOLVING F(X) = 0--ETC(U)
SEP 76 C B GARCIA, F J GOULD N00014-75-C-0495
UNCLASSIFIED 7633 NL



END

DATE FILMED 2-77

(1)

(14) Report 7633

(6) AN IMPROVED SCALAR GENERATED HOMOTOPY PATH FOR SOLVING $f(x) = 0$,

by

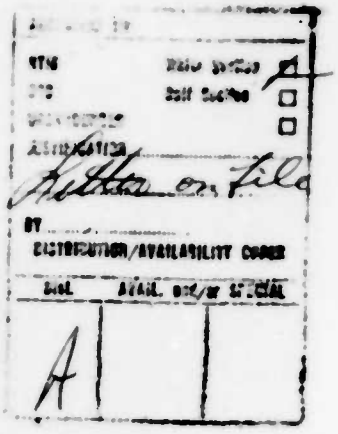
(10) C. B. Garcia and F. J. Gould

(15) N00014-75-C-0495,
NSF-GS-42010

(12) 30p.

(11) September 1976

389 629



AN IMPROVED SCALAR GENERATED HOMOTOPY PATH FOR SOLVING $f(x) = 0$

by

C. B. Garcia and F. J. Gould*

Graduate School of Business
University of Chicago

ABSTRACT

A new scalar labelling algorithm is presented for solving a system of equations by simplicial approximation. The method presented exhibits strong convergence behavior and supercedes previous simplicial pivot algorithms due to the elimination of an extra dimension, the simplification of the pivoting process by using scalar rather than vector labels, and, most importantly, the nature of the homotopy path taken which has the remarkable properties of monotonicity and Jacobian invariance. Examples are presented wherein the new method converges but Newton's method, Euler's method and previously proposed simplicial pivot algorithms fail to converge.

*The work of this author was supported in part by ONR Grant No. N00014-75-C-0495 and NSF Grant No. GS-42010.

1. INTRODUCTION

In this paper we consider the method of simplicial approximation for solving the general problem

$$\text{find } x \in R^n \text{ satisfying } f(x) = 0$$

where $$f: R^n \to R^n \text{ is continuously differentiable.}$$

A number of results on simplicial approximations appear in the literature: [1], [2], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [17], [18], [19], [20]. For example, it is known that a vector labelling [1], [10] method due to Merrill [14] tracks the "homotopy path" $f(x) = \lambda x$, $\lambda \leq 0$ assuming (without loss of generality) that the method is initiated at the origin. In [9] we showed that a "scalar labelling" can be defined on an appropriate triangulation so as to follow precisely the same homotopy path. This is of computational significance since it eliminates the need for an extra "sandwich" dimension and the need to pivot on a linear system.

In [1] and [10] a different vector labelling was defined in such a way that under appropriate assumptions the homotopy path followed is of the form $f(x) = \lambda f(0)$, $0 \leq \lambda \leq 1$, assuming once again that the starting point is at the origin. In this latter case an efficient implementation of the algorithm suffers from a difficulty in determining the initial simplex. However, once a satisfactory start is obtained, the ensuing path will in general be distinctly different than the one generated by Merrill's algorithm and it exhibits powerful convergence behavior. In particular, if f has a nonsingular Jacobian at the origin,

and if the origin is the only preimage of $f(0)$, and, finally, if the preimage of the line segment $[0, f(0)]$ (i.e. the segment joining 0 and $f(0)$) is compact then the method is assured to converge to a zero of f. Furthermore, if f is also $1-1$ on the preimage of $[0, f(0)]$ then the method is "norm-reducing" in the sense that $||f(x)||$ decreases as one moves along the path. This is the first known example of a complementary pivoting algorithm which exhibits this more classical monotonicity behavior. Another remarkable feature of the algorithm is that the same path is generated regardless of whether the algorithm is applied to $f(x)$ or to the modified function $g(x)$ obtained by premultiplying f by the inverse of its Jacobian at the starting point i.e. $g(x) = J_f^{-1}(0) \cdot f(x)$. That is, the path defined by $f(x) = \lambda f(0)$ is identical to the path defined by $g(x) = \lambda g(0)$. This invariance appears to be of importance, for with most other complementary pivoting algorithms a different path is obtained by changing from f to g, and in fact the choice of g is indicated (see [8] and [9]) because of much stronger convergence behavior. An obvious way of stating essentially the same point is as follows. Changing the subscripts of the f_i functions can change the course of the path $f(x) = \lambda x$, and hence the subscripting can affect the convergence. The path $f(x) = \lambda f(0)$ is independent of subscripting.

In this paper we present a "scalar labelling" approach that generates the homotopy path $f(x) = \lambda f(0)$, $0 \leq \lambda \leq 1$, and overcomes the above mentioned difficulty in starting. The resulting algorithm is successfully applied to

two examples where Merrill's algorithm appears to fail. This points out the advantage of following the path of the form $f(x) = \lambda f(0)$, $0 \leq \lambda \leq 1$. Use of the scalar labelling to define this path leads to improved speed of convergence because of the ease in pivoting and the elimination of an extra dimension. We also show an example where our algorithm converges but where both Newton's method and Euler's method fail.

2. DESCRIPTION OF THE LABELLING AND THE INITIAL SIMPLEX

Let $f\colon R^n \to R^n$ be continuously differentiable. Furthermore, assume that the following conditions hold:

1) $J_f(0)$ is nonsingular where $J_f(0)$ is the Jacobian of f at $x = 0$.

2) $[J_f^{-1}(0)\, f(0)]_i \neq 0$ all $i = 1, 2, \ldots, n$.

Define $g\colon R^n \to R^n$ by the mapping

$$g(x) = J_f^{-1}(0)\, f(x) \quad .$$

Then by 2) above $g_i(0) \neq 0$, all i. This function g was first used in the complementarity context in [8]. Given any fixed ε such that $0 < \varepsilon < \min_i 4|g_i(0)|$, define the label of x to be

$$L_\varepsilon(x) = \begin{cases} \min \left\{ i \,\middle|\, \dfrac{g_i(x)}{g_i(0)} \leq \dfrac{g_j(x)}{g_j(0)} \ \forall\, j \right\} & \text{if } \dfrac{g_k(x)}{g_k(0)} \in \left(0,\ 1 - \dfrac{\varepsilon}{4|g_k(0)|}\right) \text{ for some } k \\ n+1 & \text{otherwise} \end{cases}$$

The change from f to g is made in order to prove the existence of a unique $(n + 1)$ - complete initial simplex in a neighborhood of the origin. The following series of Lemmas will establish this result.

Let us employ the definition $||x|| = \max_i |x_i|$ throughout the paper. Also, for $\delta > 0$, define the δ-cube

$$D(\delta) = \{x \in R^n | -\delta \leq x_i \leq \delta \text{ , all } i\} \quad .$$

Now, given $\varepsilon > 0$, define

$\sigma_0(\varepsilon) = \{v^0, v^1, \ldots, v^n\}$ where $v^0 = 0$, $v^i = -\varepsilon \operatorname{sgn} g_i(0)\, e^i$, each $i \geq 1$, where e^i is the i^{th} unit vector in R^n.

We shall impose a triangulation on R^n in such a way that $\sigma_0(\varepsilon)$ is an n-simplex of the triangulation and such that each n-simplex intersects only a single orthant of R^n. This triangulation is specified in the appendix. Define a simplex of the triangulation to be $(n + 1)$-complete if the labels on the vertices of the simplex are $1, 2, \ldots n + 1$, and define it to be n-complete if the labels are $1, 2, \ldots, n$. (Thus, only a n-simplex can be $(n + 1)$-complete, and only $(n - 1)$ and n-simplices can be n-complete.)

We now show that there is a $D(\delta)$ such that for all triangulations with grid size small enough (as measured by ε) $\sigma_0(\varepsilon)$ is the unique $(n + 1)$-complete simplex of $D(\delta)$. Recall that ε is used to define the grid size in the sense that the vertices of the initial simplex $\sigma_0(\varepsilon)$ are given by

$$v^0 = 0,\ v^i = -\varepsilon \operatorname{sgn} g_i(0)\, e^i,\ i = 1, \ldots, n \quad .$$

Consider an $x \in D(\delta)$ for some $\delta > 0$. Recalling that, by assumption, $g_i(0) \neq 0$, all i, the Taylor expansion for $g_i(x)$ about the origin yields

$$\frac{g_i(x)}{g_i(0)} = 1 + \frac{x_i}{g_i(0)} + R(||x||), \qquad i = 1, 2, \ldots, n$$

where

$$\lim_{||x|| \to 0} \frac{R(||x||)}{||x||} = 0.$$

We choose $\delta > 0$ in such a way that

i) $\delta < 4/5 \min_i |g_i(0)|$

ii) for each $x \in D(\delta)$, $||x|| \neq 0$, we have

$$\frac{|R(||x||)|}{||x||} < \min_i \frac{1}{4|g_i(0)|}$$

Thus for each i, $\dfrac{-1}{4|g_i(0)|} \leq \max_j \dfrac{-1}{4|g_j(0)|} < \dfrac{R(||x||)}{||x||} < \min_j \dfrac{1}{4|g_j(0)|}$

$$\leq \frac{1}{4|g_i(0)|} \quad .$$

<u>Lemma 1</u> Let $||x|| = \varepsilon \in (0, \delta)$, $x_i = 0$. Then $L_\varepsilon(x) \neq i$.

<u>Proof:</u> $\dfrac{g_i(x)}{g_i(0)} = 1 + R(||x||) > 1 - \dfrac{||x||}{4|g_i(0)|} = 1 - \dfrac{\varepsilon}{4|g_i(0)|}$ and

hence $L_\varepsilon(x) \neq i$ ■

Lemma 2 Let $x \in D(\delta)$, $|x_i| = ||x|| \neq 0$, and $\operatorname{sgn} x_i = \operatorname{sgn} g_i(0)$.

Then $L_\varepsilon(x) \neq i$ if $0 < \varepsilon < \delta$.

Proof:

$$\frac{g_i(x)}{g_i(0)} = 1 + \frac{x_i}{g_i(0)} + R(||x||)$$

$$= 1 + \frac{|x_i|}{|g_i(0)|} + R(||x||)$$

$$= 1 + \frac{||x||}{|g_i(0)|} + R(||x||)$$

$$= 1 + \frac{3}{4}\frac{||x||}{|g_i(0)|} + \left(\frac{||x||}{4|g_i(0)|} + R(||x||)\right)$$

$$\geq 1 + \frac{3}{4}\frac{||x||}{|g_i(0)|} > 1 \Rightarrow L_\varepsilon(x) \neq i \text{ if } 0 < \varepsilon < \delta$$

■

Lemma 3 Let $x \in D(\delta)$, $|x_i| = ||x|| \neq 0$, $\operatorname{sgn} x_i = -\operatorname{sgn} g_i(0)$.

Then $L_\varepsilon(x) \neq n + 1$ if $0 < \varepsilon \leq ||x||$.

Proof:

$$\frac{g_i(x)}{g_i(0)} = 1 - \frac{|x_i|}{|g_i(0)|} + R(||x||)$$

$$= 1 - \frac{3}{4}\frac{||x||}{|g_i(0)|} + \left(-\frac{||x||}{4|g_i(0)|} + R(||x||)\right)$$

$$\leq 1 - \frac{3}{4}\frac{||x||}{|g_i(0)|} < 1 - \frac{||x||}{4|g_i(0)|} \leq 1 - \frac{\varepsilon}{4|g_i(0)|}$$

if $0 < \varepsilon \leq ||x||$

Furthermore $$\frac{g_i(x)}{g_i(0)} = 1 - \frac{5}{4}\frac{||x||}{|g_i(0)|} + \left(\frac{||x||}{4|g_i(0)|} + R(||x||)\right)$$

$$\geq 1 - \frac{5}{4}\frac{||x||}{|g_i(0)|} \geq 1 - \frac{5}{4}\frac{\delta}{|g_i(0)|} \quad \text{(since } ||x|| \leq \delta\text{)}$$

$$> 0 \quad \text{(since } \delta < \frac{4}{5}\min_i |g_i(0)|\text{)}.$$

Thus $L_\varepsilon(x) \neq n + 1$ ■

Lemma 4 $\sigma_0(\varepsilon)$ is $(n + 1)$-complete if $0 < \varepsilon < \delta$.

Proof: Take any $i \in \{1, 2, \ldots, n\}$.

Note that $|v_i^i| = ||v^i|| = \varepsilon$ and $\operatorname{sgn} v_i^i = -\operatorname{sgn} g_i(0)$.

Hence, by Lemma 3, $L_\varepsilon(v^i) \neq n + 1$

and, by Lemma 1, $L_\varepsilon(v^i) \neq j$ all $j \neq i$.

Hence $L_\varepsilon(v^i) = i$. Since $L(0) = n + 1$, $\sigma_0(\varepsilon)$ is $(n + 1)$-complete.

■

Lemma 5 Given $\varepsilon \in (0, \delta)$, let $\sigma = \{u^0, u^1, \ldots, u^n\} \neq \sigma_0(\varepsilon)$, $\sigma \subset D(\delta)$, and assume σ is in the same orthant as $\sigma_0(\varepsilon)$. Then $n + 1 \notin L_\varepsilon(\sigma)$.

Proof: Take any $j \in \{0, 1, \ldots, n\}$. Note that $||u^j|| \neq 0$.

Hence choose i such that $|u_i^j| = ||u^j|| \neq 0$.

Since $\sigma \subset$ orthant $\sigma_0(\varepsilon)$, sgn $u_i^j = -\text{sgn } g_i(0)$

Since $||u^j|| \geq \varepsilon$, by Lemma 3, $L_\varepsilon(u^j) \neq n + 1$

Thus $n + 1 \notin L_\varepsilon(\sigma)$.

Next, consider an n-simplex $\sigma = \{u^0, u^1, \ldots, u^n\}$, $\sigma \subset D(\delta)$ and σ not in the same orthant as $\sigma_0(\varepsilon)$, given $0 < \varepsilon < \delta$. Let $\bar{u} = (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n)$ be a vector defined by

$$\bar{u}_i = \begin{cases} \min\limits_j u_i^j & \text{if } u_i^j \geq 0 \text{ all } j = 0, 1, \ldots, n \\ \max\limits_j u_i^j & \text{if } u_i^j \leq 0 \text{ all } j = 0, 1, \ldots, n \end{cases}$$

Note that for each i, either $u_i^j \geq 0$ all j or $u_i^j \leq 0$ all j. Also note that for any 2 vertices x, y of σ, $|x_j| - |y_j| \in \{-\varepsilon, 0, \varepsilon\}$ for all j.

Furthermore, for any i and j

$$|u_i^j| \text{ is either } |\bar{u}_i| \text{ or } |\bar{u}_i| + \varepsilon.$$

First, let us consider the case where $||\bar{u}|| = 0$.

Lemma 6 Given $0 < \varepsilon < \delta$, let $\sigma = \{u^0, u^1, \ldots, u^n\} \subset D(\delta)$, σ not in the same orthant as $\sigma_0(\varepsilon)$, and $||\bar{u}|| = 0$. Then there is an $i \in \{1, 2, \ldots, n\}$ such that $i \notin L_\varepsilon(\sigma)$.

Proof: If $||\bar{u}|| = 0$ then $||u^j|| \leq \varepsilon$ for all j.
Since $\sigma \not\subset$ orthant of $\sigma_0(\varepsilon)$, there is an i such that $u_i^j \neq 0$ implies sgn u_i^j = sgn $g_i(0)$, for all j. Consider any $j = 0, 1, \ldots, n$. If $||u^j|| = 0$ then $u^j = 0$ so that $L_\varepsilon(0) = n + 1 \neq i$. Otherwise, $||u^j|| = \varepsilon$. If $u_i^j = 0$, Lemma 1 implies that $L_\varepsilon(u^j) \neq i$. If $|u_i^j| = ||u^j|| = \varepsilon$, then Lemma 2 implies $L_\varepsilon(u^j) \neq i$. ∎
Next, consider the case where $||\bar{u}|| \neq 0$.

Lemma 7 Let i be such that $|\bar{u}_i| = ||\bar{u}|| = k\varepsilon$ *for some* positive integer k. Then for all $j = 0, 1, \ldots, n$ either $|u_i^j| = ||u^j||$ or $|u_i^j| = \frac{k}{k+1}||u^j||$. (Of course $||u^j|| \neq 0$, since $||u^j|| \geq ||\bar{u}|| \neq 0$.)

Proof: Take any $j = 0, 1, \ldots, n$.

Then $|u_i^j| \leq ||u^j|| = |u_k^j|$ (some k) $\leq |\bar{u}_k| + \varepsilon \leq ||\bar{u}|| + \varepsilon$

Furthermore $|u_i^j| \geq |\bar{u}_i| = ||\bar{u}||$.

Hence if $|u_i^j| < ||u^j||$ (i.e. $|u_i^j| = ||u^j|| - \varepsilon$) we have

$$|u_i^j| = ||\bar{u}||, \quad ||u^j|| = ||\bar{u}|| + \varepsilon$$

Thus, $|u_i^j| = k\varepsilon$ and $||u^j|| = (k+1)\varepsilon$

$$\Rightarrow |u_i^j| = \frac{k}{k+1}||u^j||$$

∎

Lemma 8 Consider $\sigma = \{u^0, u^1, \ldots, u^n\}$, $\sigma \subset D(\delta)$, σ not in the same orthant as $\sigma_0(\varepsilon)$. Let i be such that $|\bar{u}_i| = ||\bar{u}|| \neq 0$. Then if $\operatorname{sgn} \bar{u}_i = \operatorname{sgn} g_i(0)$ we have $L_\varepsilon(u^j) \neq i$ all $j = 0, 1, \ldots, n$.

Proof: Take any $j \in \{0, 1, \ldots, n\}$. Since $\bar{u}_i \neq 0$, we have $u_i^j \neq 0$ and $\operatorname{sgn} u_i^j = \operatorname{sgn} g_i(0)$. By Lemma 7 either $|u_i^j| = ||u^j||$ or $|u_i^j| = \frac{k}{k+1} ||u^j||$ where k is a positive integer such that $||\bar{u}|| = k\varepsilon$.

Case a. $|u_i^j| = ||u^j||$

Then by Lemma 2 $L_\varepsilon(u^j) \neq i$.

Case b. $|u_i^j| = \frac{k}{k+1} ||u^j||$

Then $$\frac{g_i(u^j)}{g_i(0)} = 1 + \frac{|u_i^j|}{|g_i(0)|} + R(||u^j||)$$

$$= 1 + \frac{k}{k+1} \frac{||u^j||}{|g_i(0)|} + R(||u^j||)$$

$$\geq 1 + \frac{1}{2} \frac{||u^j||}{|g_i(0)|} + R(||u^j||) \quad \left(\text{since } \frac{k}{k+1} \geq \frac{1}{2}\right)$$

$$\geq 1 \text{ since } u^j \in D(\delta)$$

$$\Rightarrow L_\varepsilon(u^j) \neq i$$

■

Lemma 9 Consider σ as above. Let i be such that $|\bar{u}_i| = ||\bar{u}|| \neq 0$. Then if $\operatorname{sgn} \bar{u}_i = -\operatorname{sgn} g_i(0)$ we have $L_\varepsilon(u^j) \neq n + 1$ all $j = 0, 1, \ldots, n$.

Proof: Take any $j \in \{0, 1, \ldots, n\}$. Of course $\operatorname{sgn} u_i^j = -\operatorname{sgn} g_i(0)$. By Lemma 7, $|u_i^j| = ||u^j||$ or $|u_i^j| = \frac{k}{k+1}||u^j||$ for some positive integer k.

Case a. $|u_i^j| = ||u^j||$

Since $||u^j|| \geq \varepsilon$, Lemma 3 implies $L_\varepsilon(u^j) \neq n+1$.

Case b. $|u_i^j| = \frac{k}{k+1}||u^j||$.

Then
$$\frac{g_i(u^j)}{g_i(0)} = 1 - \frac{|u_i^j|}{|g_i(0)|} + R(||u^j||)$$

$$= 1 - \frac{k}{k+1}\frac{||u^j||}{|g_i(0)|} + R(||u^j||)$$

$$\leq 1 - \frac{1}{2}\frac{||u^j||}{|g_i(0)|} + R(||u^j||) \quad \text{since } \frac{k}{k+1} \geq \frac{1}{2}$$

$$\leq 1 - \frac{1}{4}\frac{||u^j||}{|g_i(0)|} + \left(-\frac{||u^j||}{4|g_i(0)|} + R(||u^j||)\right)$$

$$< 1 - \frac{1}{4}\frac{||u^j||}{|g_i(0)|} \leq 1 - \frac{\varepsilon}{4|g_i(0)|} \quad \text{since } ||u^j|| \geq \varepsilon.$$

Furthermore
$$\frac{g_i(u^j)}{g_i(0)} = 1 - \frac{5k+1}{4k+4}\frac{||u^j||}{|g_i(0)|} + \left(\frac{||u^j||}{4|g_i(0)|} + R(||u^j||)\right)$$

$$\geq 1 - \frac{5k+1}{4k+4}\frac{||u^j||}{|g_i(0)|} \geq 1 - \frac{5k+1}{4k+4}\frac{\delta}{|g_i(0)|}$$

(since $||u^j|| \leq \delta$)

$$\geq 1 - \frac{5}{4}\,\frac{\delta}{|g_i(0)|} \quad \text{since} \quad \frac{5k+1}{4k+4} \leq \frac{5}{4}$$

$$> 0 \ (\text{since} \ \ \delta < \tfrac{4}{5} \min_i |g_i(0)|)$$

$$\Rightarrow L_\varepsilon(u^j) \neq n+1. \qquad \blacksquare$$

Theorem 1. $\sigma_0(\varepsilon)$ is the unique $(n+1)$-complete simplex in $D(\delta)$ if $0 < \varepsilon < \delta$.

Proof: This follows from Lemmas 4, 5, 6, 8 and 9. $\blacksquare$

The choice of ε in Theorem 1 is crucial in the speed of convergence of our method. If ε is too large, it is possible that the method cannot be initiated. If ε is too small, the method may require a prohibitively large number of pivots to reach a terminal simplex. Our test examples show that $\varepsilon = \min_i |g_i(0)|$ worked well.

3. CONVERGENCE

The above Theorem 1 is all that is needed to validate our method. The method is now a familiar one in complementarity theory: starting from the simplex $\sigma_0(\varepsilon)$ generate the sequence of distinct simplices $\sigma_1, \sigma_2, \ldots, \sigma_t \ldots$ such that $\sigma_i \cap \sigma_{i+1}$ is n-complete for $i = 0, 1, 2, \ldots,$. Terminate upon reaching for the first time an n-simplex σ_T which is $(n+1)$-complete. We may now use any vertex of σ_T as the next point to restart the method.

A condition under which this method is assured to converge to a zero of f is given by the following Theorem. First define the set H by

$$H = \{x|f(x) = \lambda f(0),\ 0 \leq \lambda \leq 1\} = \{x|g(x) = \lambda g(0),\ 0 \leq \lambda \leq 1\}$$

We define the ρ-neighborhood of H to be $\{y \mid ||y - x|| \leq \rho \text{ for some } x \in H\}$.

Theorem 2. Let $f: R^n \to R^n$, f continuously differentiable on R^n. Suppose $f(x) = f(0)$ iff $x = 0$, $J_f^{-1}(0)$ exists, $[J_f^{-1}(0)\ f(0)]_i \neq 0$ all i, and H is compact. Then there exists a solution x to $f(x) = 0$ and furthermore as the mesh size goes to zero the method will converge to a solution.

Proof: Let $\delta > 0$ be such that Theorem 1 holds and let $\{\varepsilon_i\}^{\infty}$ be a sequence, $0 < \varepsilon_i \leq \delta$, with ε_i decreasing to 0. Let C_i denote the path of simplices generated by the method for a triangulation of mesh size ε_i. Then for every ρ-neighborhood of H, say N_ρ, there exists a $\bar{i}$ such that $i > \bar{i}$ implies $C_i \subseteq N_\rho$. To see this assume to the contrary that there is a neighborhood N_ρ for which $C_i \cap (R^n - N_\rho) \neq \emptyset$ for infinitely many i. Let $x^i \in \partial N_\rho \cap C_i$. Then it follows from the continuity of f on N_ρ, the nature of the labeling, and the fact that each simplex in each path is n-complete, that every cluster point of $\{x^i\}$ is in H, which is a contradiction. It follows from this result that for i sufficiently large each path C_i terminates with a final simplex, say σ_i. Since, by Theorem 1, $\sigma_0(\varepsilon)$ is the unique

$(n + 1)$-complete simplex in the δ-cube, each terminal simplex σ_i is outside this cube. Since this terminal simplex contains labels 1 through $n + 1$, any x in the simplex must satisfy $\frac{g_i(x)}{g_i(0)} \geq 1$ for all i or $\frac{g_i(x)}{g_i(0)} \leq 0$ for all i. Since the unique preimage of $g(0)$ is, by assumption, the origin, $g(x)$ is approximately zero for x in σ_i. Take any sequence $\{s^i\}$ such that $s^i \in \sigma_i$. It is clear that any cluster point of $\{s^i\}$ is a zero of f and g. ∎

4. JACOBIAN INVARIANCE AND MONOTONICITY

It is clear that finding a solution to the system

$$f_i(x) = 0, \quad i = 1, \ldots, n \tag{1}$$

is equivalent to solving the system

$$g_i(x) = 0, \quad i = 1, \ldots, n \tag{2}$$

where $g(x) = J_f^{-1}(0) \cdot f(x)$. Most simplicial pivoting algorithms follow different paths depending on whether system (1) or (2) is being solved. Moreover it is not unusual for the simplicial path associated with (1) to be unbounded whereas the path for (2) converges (see [8], [9], [13], [19], and [20]). In other words, the transformation to system (2) is known to improve convergence. For the method presented in this paper the simplicial path is the same, regardless of whether (1) or (2) is attacked. More precisely, suppose the labeling function $L_\varepsilon(x)$ is redefined exclusively in terms of f rather

than g (i.e. replace g_i with f_i in the definition). If this is done, a unique initial $n+1$-complete simplex may not exist, and hence it may be either impossible or difficult to implement the algorithm. However, if a unique first simplex is found, then the path followed will coincide with the path generated by the method of this paper. In other words, the paths for (1) and (2) are the same. We have defined the labeling function in terms of g (i.e. we solve system (2)) only to constructively prove the existence of $\sigma_0(\varepsilon)$. This assures us of a start. The fact that the path is invariant under the Jacobian transformation endows this algorithm with a natural property not shared by others.

Another property that the method possesses is that if f is $1-1$ on the homotopy path $H = \{x | f(x) = \lambda f(0),\ 0 \leq \lambda \leq 1\}$ then $||f(x)||$ decreases monotonically to zero on that path. Monotonicity is assured inasmuch as λ will decrease from 1 to 0 as one traces the path H. This property is especially useful in restarting our method. In other simplicial pivot techniques, there is no assurance that the sequence of approximate solutions, x^k (in the terminal simplex of the iteration corresponding to ε_k) will become "better" approximations to a true solution as k increases. In our method however, for ε_k suitably small, at any point x^k in a terminal simplex it will be true that $||f(x^k)||$ is less than $||f(x^{k-1})||$ i.e. in each iteration the norm of f at a point in the terminal simplex is less than the norm of f at a point in the initial simplex for that iteration. Thus each point x^k is an improved estimate.

5. EXAMPLES

Here we present three examples. In each example our method converged. We also attempted to solve the first two examples by using Merrill's algorithm [14] initiated at the same starting point. In each of these two examples the variables on the homotopy path appeared to grow without bound and hence Merrill's algorithm terminated without convergence. In the third example both Newton's method and Euler's method failed whereas our method converged.

Example I

$$f_i(x) = x_i^3 - \sum_{j \neq i} x_j - 1000 \qquad 1 \leq i \leq n$$

We used our method on this problem for $n = 10$. The starting point was chosen to be $x_i = 20$ all i. The mesh size ε_k was chosen by

$$\varepsilon_{k+1} = \min_i |g_i(x^k)|$$

where x^k is the approximation furnished by the previous iteration. This selection of ε_k was the most ideal among the options we considered. As seen in the examples, such a choice yielded $\varepsilon_{k+1} = \frac{\varepsilon_k}{2}$ or better in every case. The results are shown in Table I.

Example II

This example is a standard $R(h^2)$ discretization of a two point boundary problem (see Moré and Cosnard [15])

$$u''(t) = \frac{1}{2}(u(t) + t + 1)^3 \quad 0 < t < 1, \; u(0) = u(1) = 0 \; .$$

The resulting problem in the unknowns $x_k = u(t_k)$ is defined by

$$f_k(x) = 2x_k - x_{k+1} - x_{k-1} + \frac{h^2}{2}(x_k + t_k + 1)^3, \quad 1 \leq k \leq n$$

where n is taken to be 10, and

$$x_0 = x_{n+1} = 0, \; t_k = kh \quad \text{and} \quad h = \frac{1}{n+1} \; .$$

The method is started at the point

$$x^0 = (\varepsilon_1, \ldots, \varepsilon_n) \quad \text{where} \quad \varepsilon_i = t_i(t_i - 1)$$

and the results are presented in Table II. The system of equations has a unique solution $x^* = (\varepsilon_1^*, \ldots, \varepsilon_n^*)$ where $-0.5 \leq \varepsilon_i^* \leq 0$.

Example III

Let $$f_i(x) = nx_i + \frac{n}{2}\sin x_i - \sum_{j \neq i} x_j - 100 - i \quad , \quad 1 \leq i \leq n.$$

A solution (up to an accuracy of 8.106×10^{-5}) to this problem for $n = 5$ is $x = (101.043, 101.166, 101.293, 101.423, 101.557)$. We tested Newton's method

$$x^{k+1} = x^k - J_f^{-1}(x^k)\, f(x^k)$$

twice on the above problem using initial points $x^0 = (20, 20, 20, 20, 20)$ and $x^0 = (120, 120, 120, 120, 120)$.

In both cases, the method failed to converge due to some component of x growing large.

We also tested Euler's method of the form

$$x^{k+1} = x^k - J_f^{-1}(x^k)\left[f(x^k) + \frac{k - T}{T} f(x^0)\right]$$

$$x^{k+1} = x^k - J_f^{-1}(x^k) f(x^k) \qquad \begin{matrix} k = 0, 1, \ldots, T - 1 \\ k = T, T + 1, \ldots \end{matrix}$$

where T is a given positive integer.

This method is a "continuation method" which may be visualized as that of approximating the path $f(x) = \lambda f(0)$, $0 \leq \lambda \leq 1$ by way of Newton directions ([16], p. 232). The method converged from starting point $x_i^0 = 120$ all i using $T = 100$. However, it failed to converge from $x_i^0 = 20$ using $T = 100$, 1000, and 10000.

With the same initial points, we tested our method on the problem above first using grid sizes computed via

$$\varepsilon_{k+1} = \min_i |g_i(x^k)|$$

where x^k is the approximating solution after major iteration k. For this approach our method moved to points within 10 units of the solution, but failed to come closer because points x^k were encountered for which the preimage of $g(x^k)$ is not unique.

We then used grid sizes of the form

$$\varepsilon_0 = \min_i |g_i(x^0)|$$

$$\varepsilon_{k+1} = \min \{\frac{\varepsilon_k}{2}, \min_i |g_i(x^k)|\}$$

For both initial points $x_i = 20$ all i, $x_i = 120$ all i the method converged to a solution. The results of our tests are shown in Tables III and IV.

Conclusion

In this paper we presented a "scalar labelling" algorithm for solving a system of equations by way of simplicial approximation. The method presented possesses several remarkable properties such as monotonicity and Jacobian invariance. The algorithm compares favorably with previous algorithms of Eaves and Saigal [4] and Merrill [14] due to the elimination of an extra dimension, the simplification of the pivoting process by using scalar rather than vector labels and, most importantly, the nature of the homotopy path taken.

$n = 10 \qquad x^0 = (20, 20, 20, 20, 20, 20, 20, 20, 20, 20)$

ITER. NO.	x^k	MESH SIZE ϵ_k	$\|\|g(x^k)\|\|$	NO. OF PIVOTS TO REACH x^k
1	(8.5, 11.4, ..., 11.4)	2.86314	475.369	758
2	(10.68, 10.34, ..., 10.34)	1.07039	296.865	300
3	(10.13, 10.34, ..., 10.34)	.187178	54.3887	20
4	(10.30, 10.34, ..., 10.34)	.088	15.4807	10
5	(10.28, 10.30, ..., 10.30)	.020	6.3	458
6	(10.29, 10.30, ..., 10.30)	.0085	2.67	10
7	(10.29, 10.3, ..., 10.3)	7.95×10^{-5}	.0246	210

Table I

$n = 10$ with $x_i^0 = \frac{i}{n+1}\left(\frac{i}{n+1} - 1\right)$ $i = 1, 2, \ldots, 10$

ITER. NO.	MESH ϵ_k	$\|\|f(x^k)\|\|$	NO. OF PIVOTS TO REACH x^k
1	4.56×10^{-2}	.1019	158
2	2.24×10^{-2}	3.26×10^{-2}	154
3	1.115×10^{-2}	2.25×10^{-2}	162
4	5.56×10^{-3}	8.863×10^{-3}	150
5	2.77×10^{-3}	3.58×10^{-3}	192
6	1.38×10^{-3}	3.17×10^{-3}	138
7	6.926×10^{-4}	1.58×10^{-3}	108

$x^7 = (-4.3 \times 10^{-2}, -8.2 \times 10^{-2}, -.114, -.142, -.160, -.170, -.170, -.156, -.125, -.075)$

Table II

$n = 5 \qquad x^0 = (20, 20, 20, 20, 20)$

ITER. NO.	$\|\|f(x^k)\|\|$	ϵ_k	NO. OF ITERS. TO REACH x^k
1	41.24	40.24	95
2	6.161	1.2575	1028
3	6.923	.629	46
4	4.105	.3143	75
5	3.691	.1572	77
6	3.084	.0786	94
7	1.044	.0393	115
8	.4968	.0196	72
9	.8970	9.82×10^{-3}	112
10	.4809	4.91×10^{-3}	213
11	.3867	2.46×10^{-3}	274

$x^{11} = (101.095, 101.166, 101.291, 101.421, 101.556)$

Table III

$x^0 = (120, 120, 120, 120, 120)$

ITER. NO.	$\|\|f(x^k)\|\|$	ϵ_k	NO. OF ITERS. TO REACH x^k
1	2.81462	6.3276	95
2	25.4632	1.5819	43
3	5.3331	.79095	28
4	3.3965	.3955	36
5	1.3965	.1977	51
6	.50790	.0989	41
7	.4579	.0494	48
8	.2582	.0247	80
9	.0510	.0124	68

$x^9 = (101.046, 101.165, 101.3, 101.421, 101.55)$

Table IV

The pivot rules are given as follows. Suppose $\{u^0, u^1, \ldots, u^n\}$ and $(s_1, \ldots, s_n)$ are given. If the vertex u^i is dropped we must specify a new initial vertex $\hat{u}^0$ and a new permutation $(\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_n)$. The rule $\hat{u}^{i+1} = \hat{u}^i + Q(\hat{s}_{i+1})$ will then determine the new simplex $(\hat{u}^0, \hat{u}^1, \ldots, \hat{u}^n)$.

Case 1: Drop u^i, $1 \leq i \leq n - 1$

$$\hat{u}^0 = u^0,\ \hat{s}_i = s_{i+1},\ \hat{s}_{i+1} = s_i,\ \hat{s}_j = s_j \text{ for } j \notin \{i, i+1\}.$$

Case 2: Drop u^0

$$\hat{u}^0 = u^1,\ \hat{s}_j = s_{j+1},\ 1 \leq j \leq n - 1,\ \hat{s}_n = s_1$$

Case 3: Drop u^n

$$\hat{u}^0 = u^0 - Q(s_n),\ \hat{s}_1 = s_n,\ \hat{s}_j = s_{j-1},\ 2 \leq j \leq n.$$

Note that $\sigma_0(\varepsilon)$ is determined by $u^0 = 0$ and the permutation $(1, 2, \ldots, n)$ i.e. $\sigma_0(\varepsilon) = \{v^0, v^1, \ldots, v^n\}$, $v^0 = 0$, $v^i = -\varepsilon \operatorname{sgn} g_i(0)\, e^i$, $1 \leq i \leq n$.

Appendix: The Triangulation and the Pivot Rules

The triangulation and pivot rules are given for R^n. Let Q denote the $n \times n$ matrix whose i^{th} column is denoted by $Q(i)$, where $Q(i) = v^i - v^{i-1}$, and where $v^0 = 0$, $v^i = -\varepsilon \operatorname{sgn} g_i(0)\, e^i$, $1 \le i \le n$. That is

$$Q = \begin{bmatrix} -\varepsilon \operatorname{sgn} g_1(0) & \varepsilon \operatorname{sgn} g_1(0) & 0 & 0 \\ 0 & -\varepsilon \operatorname{sgn} g_2(0) & \varepsilon \operatorname{sgn} g_2(0) & 0 \\ 0 & 0 & -\varepsilon \operatorname{sgn} g_3(0) & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ & & & \varepsilon \operatorname{sgn} g_{n-1}(0) \\ 0 & 0 & 0 & -\varepsilon \operatorname{sgn} g_n(0) \end{bmatrix}$$

Let $(s_1, \ldots, s_n)$ denote a permutation of the integers $(1, 2, \ldots, n)$. That is, each $s_i \in \{1, \ldots, n\}$ and $s_i = s_j \iff i = j$. Now consider a triangulation of mesh ε. And let $\{u^0, \ldots, u^n\}$ be an n-simplex in the triangulation, where u^0 is any point in the ε-lattice (each u_i^0 is an integral multiple of ε). Associated with this simplex is a permutation $(s_1, s_2, \ldots, s_n)$ such that

$$u^{i+1} = u^i + Q(s_{i+1}), \quad 0 \le i \le n - 1.$$

That is, any simplex in the triangulation is defined by the initial vertex u^0 and the permutation $(s_1, \ldots, s_n)$.

REFERENCES

[1] A. Charnes, C. B. Garcia, and C. E. Lemke, "Constructive proofs of theorems relating to $f(x) = y$ with applications," to appear in Mathematical Programming.

[2] B. C. Eaves, "Computing Kakutani fixed points," SIAM Journal of Applied Mathematics, 21 (1971), pp. 236-244.

[3] B. C. Eaves, "Solving piecewise linear convex equations," Mathematical Programming Study 1 (November 1974) 96-119.

[4] B. C. Eaves and R. Saigal, "Homotopies for computation of fixed points on unbounded regions, " Mathematical Programming 3, No. 2 (1972) 225-237.

[5] B. C. Eaves, "Homotopies for computation of fixed points," Mathematical Programming 3, No. 1 (1972) 1-22.

[6] M. L. Fisher and F. J. Gould, "A simplicial algorithm for the nonlinear complementarity problem," Mathematical Programming 6 (1974) 281-300.

[7] M. L. Fisher, F. J. Gould and J. W. Tolle, "A new simplicial approximation algorithm with restarts: Relations between convergence and labelling," to appear in Proceedings of the Conference on Computing Fixed Points with Applications, Department of Mathematical Sciences, Clemson University, Clemson, South Carolina (June 1974). Also in Center for Mathematical Studies in Business and Economics, Report No. 7427, July, 1974, University of Chicago, Chicago, Illinois.

[8] M. L. Fisher, F. J. Gould and J. W. Tolle, "A simplicial approximation algorithm for solving systems of nonlinear equations," to appear in the Proceedings of the Conference on Mathematical Programming and Its Applications, National Institute of Higher Mathematics, City University, Rome, Italy (April 1974), also in Center for Mathematical Studies in Business and Economics, Report No. 7421, May, 1974, University of Chicago, Chicago, Illinois.

[9] C. B. Garcia and F. J. Gould, "An algorithm based on the equivalence of vector and scalar labels in simplicial approximation," submitted to Mathematical Programming, also in Center for Mathematical Studies in Business and Economics Report No. 7626, July, 1976, University of Chicago, Chicago, Illinois.

[10] C. B. Garcia, "A global existence theorem for the equation $Fx = y$," submitted to SIAM Journal of Applied Mathematics, also in Center for Mathematical Studies in Business and Economics Report No. 7527, June, 1975, University of Chicago, Chicago, Illinois.

[11] C. B. Garcia, C. E. Lemke and H. J. Luethi, "Simplicial approximation of an equilibrium point for non-cooperative N-person games," in Mathematical Programming, eds., T. C. Hu and S. M. Robinson (Academic Press 1973) 227-260.

[12] F. J. Gould and J. W. Tolle, "A unified approach to complementarity in optimization," Discrete Mathematics, 7, Nos. 3-4 (1974), 225-271.

[13] F. J. Gould and J. W. Tolle, "An existence theorem for solutions to $f(x) = 0$," to appear in Mathematical Programming, also in Center for Mathematical Studies in Business and Economics Report No. 7515, March, 1975, University of Chicago, Chicago, Illinois.

[14] O. H. Merrill, "Applications and extensions of an algorithm that computes fixed points of certain upper semi-continuous point to set mappings," Ph.D. Thesis, University of Michigan, Ann Arbor, 1972.

[15] J. Moré and M. Y. Cosnard, "Numerical comparisons of three nonlinear equation solvers," Argonne National Laboratory, Report TM-286, February, 1976.

[16] J. Ortega and W. Rheinboldt, Iterative Solutions of Nonlinear Equations in Several Variables,(Academic Press, New York 1970).

[17] H. Scarf, The computation of economic equilibria (Yale University Press, New Haven 1973).

[18] H. Scarf, "The approximation of fixed points of a continuous mapping," SIAM Journal of Applied Mathematics 15 (1967).

[19] M. J. Todd, Improving the convergence of fixed-point algorithms," Technical report No. 276, October, 1975, Department of Operations Research, College of Engineering, Cornell University, Ithaca, New York.

[20] L. A. Wolsey, "Convergence, simplicial paths and acceleration methods for simplicial approximation algorithms for finding a zero of a system of nonlinear equations," CORE Discussion Paper No. 7427, CORE, Universite Catholique de Louvain, Belgium (December 1974).