NEWTON DERIVED METHODS FOR NONLINEAR EQUATIONS AND INEQUALITIES

by

E. Polak and I. Teodoru

Memorandum No. ERL-M449

31 January 1974

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

NEWTON DERIVED METHODS FOR NONLINEAR EQUATIONS AND INEQUALITIES

by

E. Polak and I. Teodoru

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

1. Introduction

It has been shown by Pshenichnyi [6] and Robinson [7] that Newton's method can be extended to the solution of systems of equations and inequalities, where the number of variables may be larger than the number of equations and inequalities. These extensions of Newton's method converge quadratically when started from a good initial guess, but, just like Newton's method, may diverge when started from a poor initial guess. In a recent, as yet unpublished paper, Huang [2'] described a locally convergent modification of Robinson's algorithm.

Our paper presents an alternative stabilized version of Robinson's algorithm which has a greater rate of convergence and requires less strict assumptions than Huang's algorithm. In addition, we present an "iterated" version which is more efficient. The stabilization is accomplished by using an Armijo-type gradient method [1] until a battery of tests indicates that one is close enough to a solution for Robinson's extension of Newton's method to converge. Thus we obtain algorithms which are globally convergent and have root rate of convergence $r \in (1,2]$ depending on the choice of parameters. We

Research sponsored by the National Science Foundation Grant GK-37672 and the U. S. Army Research Office -- Durham Contract DAHC04-73-C-0025.

show, following Brent [2], that there is a choice of parameters which maximizes the efficiency of the algorithm. Our computational experience has been most encouraging.

2. The Algorithm: Convergence

The algorithm we are about to state solves problems of the form: find $z \in \mathbb{R}^n$ such that

$$g(z) = 0 \qquad f(z) \leqq 0 \tag{1}$$

where $g: \mathbb{R}^n \to \mathbb{R}^m$, $f: \mathbb{R}^n \to \mathbb{R}^q$ are continuously differentiable functions. We use superscripts to denote components of f, g, z, etc.

For some $b >> 1$ and $\|\cdot\|$ denoting the Euclidean norm, we define $\bar{f}: \mathbb{R}^n \to \mathbb{R}^{q+1}$ by

$$\bar{f}^j(z) \triangleq \begin{cases} f^j(z) & j = 1,\ldots,q \\ \|z\|^2 - b & j = q+1 \end{cases} \tag{2}$$

We shall use the following notation:

$$h(z) \triangleq \begin{bmatrix} g(z) \\ \bar{f}(z) \end{bmatrix} \qquad H(z) \triangleq \frac{\partial h(z)}{\partial z} \tag{3}$$

$$G(z) \triangleq \frac{\partial g(z)}{\partial z}, \qquad \bar{F}(z) \triangleq \frac{\partial \bar{f}(z)}{\partial z} \tag{4}$$

$$\bar{f}^{j+}(z) \triangleq \max\{0, \bar{f}^j(z)\} \qquad j=1,2,\ldots,q+1 \tag{5}$$

We also define a cost function $f^o: \mathbb{R}^n \to \mathbb{R}^1$ by

$$f^o(z) \triangleq \frac{1}{2}\|g(z)\|^2 + \frac{1}{2}\|\bar{f}^+(z)\|^2 \tag{6}$$

It is not difficult to see that f^o is continuously differentiable and

$$\nabla f^o(z) = G(z)^T g(z) + \bar{F}(z)^T \bar{f}^+(z) \tag{7}$$

Finally, given any $z_0 \in \mathbb{R}^n$

$$C(z_0) \triangleq \{z \mid f^o(z) \leq f^o(z_0)\} \tag{8}$$

We now state assumptions which ensure that our algorithm is globally convergent.

<u>Assumption 1</u>: The derivative matrices $G(\cdot)$ and $F(\cdot)$ are Lipschitz continuous. ¤

<u>Assumption 2</u>: The pair $(\bar{F}(z), G(z))$ satisfies the Robinson LI condition [7] for all $z \in C(z_0)$, where z_0 is the initial guess to a solution for (1) and b is sufficiently large to ensure that the set $C(z_0)$ contains at least one such solution; i.e. for all $z \in C(z_0)$

$$u^T \bar{F}(z) + v^T G(z) = 0 \tag{9}$$

and $u \geqq 0$, implies that $u = 0$ and $v = 0$. ¤

<u>Algorithm</u>

<u>Data:</u> $z_0 \in \mathbb{R}^n$, $b >> \|z_0\|^2$, $\alpha \in (0,1/2)$, $\beta \in (0,1)$, $\hat{\ell} > 1$, $\gamma \in (0,1)$, $k \in N^+$.

<u>Step 0:</u> Set $i = 0$, $j = 0$, $s = 1$, $p = 0$.

<u>Comment:</u> i is the iteration index,, the Jacobians $\bar{F}$ and G are evaluated at z_j, $s = i - j + 1$ is the number of times the same Jacobians have been used.

Step 1: Compute $g(z_i)$, $\bar{f}(z_i)$, $f^o(z_i)$, $G(z_j)$, and $\bar{F}(z_j)$. Stop if $f^o(z_i) = 0$.

Comment: When actually programming this algorithm, compute $G(z_j)$ and $\bar{F}(z_j)$ only as necessary (i.e. every k steps).

Step 2: Compute a vector v_i which solves the problem

$$\text{minimize}\{\|v\|_\infty | g(z_i) + G(z_j)v = 0,\ \bar{f}(z_i) + \bar{F}(z_j)v \leqq 0\} \tag{10}$$

where

$$\|v\|_\infty \triangleq \max_r |v^r| \tag{11}$$

Comment: Due to Assumption 2, the linearized problem (10) always has a solution, obtainable by linear programming techniques.

Step 3: If $\|v_i\|_\infty \le \gamma^p$ and $z_i + v_i \in C(z_0)$, set $z_{i+1} = z_i + v_i$, set $p_i = p$, set $p = p + 1$, set $i = i + 1$ and go to step 14; else go to step 4.

Step 4: Set $w_i = v_i$, $\phi(z_i) = -2f^o(z_i)$.

Step 5: If $j = i$ go to step 11; else go to step 6.

Step 6: Set $\ell = 0$.

Step 7: Compute $f^o(z_i + \beta^\ell w_i)$.

Step 8: If

$$f^o(z_i + \beta^\ell w_i) - f^o(z_i) \le \beta^\ell \alpha\phi(z_i) \tag{12}$$

set $\ell_i = \ell$, set $z_{i+1} = z_i + \beta^{\ell_i} w_i$, set $i = i + 1$ and go to step 14; else go to step 9.

Step 9: If $\ell < \hat{\ell}$ set $\ell = \ell + 1$ and go to step 7; else go to step 10.

Step 10: Compute $\nabla f^o(z_i)$, set $w_i = -\nabla f^o(z_i)$ and set $\phi(z_i) = -\|\nabla f^o(z_i)\|^2$.†

†When the solution of (10) is not costly, replace step 10 with: Compute $G(z_i)$, $F(z_i)$, set $j = i$, set $s = 1$ and go to step 2.

Step 11: Set $\ell = 0$.

Step 12: Compute $f^o(z_i + \beta^\ell w_i)$.

Step 13: If (12) is satisfied, set $\ell_i = \ell$, set $z_{i+1} = z_i + \beta^{\ell_i} w_i$, set $i = i + 1$ and go to step 14; else set $\ell = \ell + 1$ and go to step 12.

Step 14: If $s < k$, set $s = s + 1$ and go to step 1; else, set $s = 1$, $j = i$ and go to step 1. ◻

We shall now establish the convergence properties of the algorithm in four stages. First we shall show that if $\{z_i\}_{i=0}^{\infty}$ is constructed by the algorithm, then it contains a subsequence which converges to a solution. Next we shall show that the relation $z_{i+1} = z_i + v_i$, i.e. $\|v_i\|_\infty \leq \gamma^{p_i}$, $z_i + v_i \in C(z_0)$, holds an infinite number of times. Next we shall show that if there is an i' such that the test $\|v_i\|_\infty \leq \gamma^{p_i}$, $z_i + v_i \in C(z_0)$ in step 3 is satisfied for all $i \geq i'$ and $\|v_i\|$ is sufficiently small, then $z_i \to \hat{z}$, a solution of (1). This result takes the form of a local convergence theorem. We shall complete the proof by exhibiting the existence of an i' such that $z_{i+1} = z_i + v_i$ for all $i \geq i'$.

Proposition 1: Let v_i be computed by the algorithm at z_i, from (10). Then $\phi(z_i) = 0$ if and only if $w_i = 0$ if and only if $\nabla f^o(z_i) = 0$, i.e. if and only if z_i solves (1). ◻

Lemma 1: Suppose that $z \in \mathbb{R}^n$ is such that $\phi(z) < 0$, where $\phi(z)$ was defined in the algorithm (step 4 or step 10). Then there exists an integer $\bar{\ell} \geq 0$, finite, and an $\varepsilon(z) > 0$ such that

$$f^o(z' + \beta^{\bar{\ell}} w') - f^o(z') \leq \beta^{\bar{\ell}} \alpha \phi(z') \tag{13}$$

for all $z' \in B(z, \varepsilon(z)) \triangleq \{z' \,|\, \|z' - z\| \leq \varepsilon(z)\}$, for all w' satisfying for some $Q < \infty$

$$\|w'\| \leq Q \quad \text{and} \quad \langle \nabla f^o(z'), w' \rangle \leq \phi(z') \tag{14}$$

Proof: Because of Assumption 1, $\nabla f^o(\cdot)$ is continuous. Hence $\phi(\cdot)$ is continuous and there exists an $\varepsilon(z) > 0$ such that $\phi(z') \leq \phi(z)/2 < 0$ for all $z' \in B(z,\varepsilon(z))$. Since $B(z,\varepsilon(z))$ is compact, $\nabla f^o(\cdot)$ is uniformly Lipschitz continuous on this set, with constant L, say, and hence, for any $\lambda > 0$, and w' satisfying (14),

$$f^o(z'+\lambda w') - f^o(z') - \lambda \langle \nabla f^o(z'),w' \rangle$$

$$= \int_0^1 \langle \nabla f^o(z'+s\lambda w') - \nabla f^o(z'),\lambda\, w' \rangle ds \leq \frac{L\lambda^2}{2} \|w'\|^2 \tag{15}$$

Consequently,

$$f^o(z'+\lambda w') - f^o(z') \leq \lambda\alpha \langle \nabla f^o(z'), w' \rangle + \lambda \Big[(1-\alpha) \langle \nabla f^o(z'), w' \rangle$$

$$+ \frac{L\lambda}{2} \|w'\|^2 \Big] \leq \lambda\alpha\phi(z') + \frac{\lambda}{2} [(1-\alpha)\, \phi(z) + L\lambda Q^2] \tag{16}$$

Let $\bar{\ell} \geq 0$ be the smallest integer such that $(1-\alpha)\, \phi(z) + L\beta^{\bar{\ell}} Q^2 \leq 0$. Then, clearly $\bar{\ell}$ satisfies (13) for all $z' \in B(z,\varepsilon(z))$, for all w' satisfying (14) and the lemma is proved. ¤

Corollary 1: The algorithm is well defined, i.e., it does not jam up between steps 12 and 13.

Proof: We know from Proposition 1 that it is not possible to reach step 11 with $\phi(z_i) < 0$ and $w_i = 0$. Next, since $f^o(z)$ contains the term $((\|z\|^2 - b)^+)^2$, it is clear that $C(z_0)$ is compact and hence there exists an $M < \infty$ such that $\max\{\|\nabla f^o(z)\|_\infty, \|h(z)\|_\infty\} \leq M$ for all $z \in C(z_0)$. Hence, making use of (56) and the equivalence of norms in $\mathbb{R}^n$, we conclude that there exists a $Q \in [M,\infty)$, such that $\|w_i\| \leq Q$ for any z_i constructed by the algorithm. Next, if $w_i = v_i$ we have from (10) that $< \nabla f^o(z_i), w_i > \leq -2f^o(z_i) = \phi(z_i)$, whereas if $w_i = -\nabla f^o(z_i)$ the same result follows from the definition in Step 10. Hence by Lemma 1 there exists an $\ell_i < \infty$ satisfying (12). ¤

Lemma 2: Suppose that the algorithm has constructed an infinite sequence $\{z_i\}_{i=0}^\infty$. Then $\{z_i\}_{i=0}^\infty$ has at least one accumulation point $\hat{z}$ which solves (1), i.e. $g(\hat{z}) = 0$, $\bar{f}(\hat{z}) \leqq 0$.

Proof: First, suppose that there is an $\hat{i}$ such that for all $i \geq \hat{i}$, z_{i+1} is constructed according to the formula in steps 8 or 13 ($z_{i+1} = z_i + \beta^{\ell_i} w_i$). Then it follows directly from Lemma 1, and theorem 1.3.9 in [5] that every accumulation point $\hat{z}$ of $\{z_i\}_{i=0}^\infty$ satisfies $\nabla f^o(\hat{z}) = 0$ and hence, from Assumption 2, is a solution of (1). Furthermore, since the set $C(z_{\hat{i}}) = \{z | f^o(z) \leq f^o(z_{\hat{i}})\}$ is compact, and $z_i \in C(z_{\hat{i}})$ for all $i \geq \hat{i}$, it follows that $\{z_i\}$ has at least one accumulation point $\hat{z}$ which solves (1).

Next, suppose there is no $\hat{i}$ such that $z_{i+1} = z_i + \beta^{\ell_i} w_i$ for all $i \geq \hat{i}$. Then there must exist an infinite subset $K \subset \{0,1,2,\ldots\}$ such that $\|v_i\|_\infty \leq \gamma^{p_i}$ for all $i \in K$, and $p_i \to \infty$ as $i \to \infty$, $i \in K$. Consequently, $v_i \to 0$ as $i \to \infty$, $i \in K$. This implies that $g(z_i) \to 0$, $\bar{f}(z_i) \to \hat{f} \leqq 0$ as $i \to \infty$, $i \in K$ since the Jacobians $G(\cdot)$ and $\bar{F}(\cdot)$ are bounded over $C(z_0)$. Since $\{z_i\}_{i \in K}$ is compact, there exists an infinite subset $K' \subset K$ and a $\hat{z} \in C(z_0)$ such that $z_i \to \hat{z}$ as $i \to \infty$, $i \in K'$, and therefore $g(\hat{z}) = 0$, $\bar{f}(\hat{z}) = \hat{f} \leqq 0$, i.e. $\{z_i\}_{i=0}^\infty$ has an accumulation point $\hat{z}$ which solves (1). ¤

Corollary 2: Suppose that $\{z_i\}_{i=0}^{\infty}$ is a sequence constructed by the algorithm and suppose that $\hat{z}$ is an accumulation point of $\{z_i\}_{i=0}^{\infty}$ which solves (1). If $K \subset \{0,1,2...\}$ is an index set identifying a subsequence of $\{z_i\}_{i=0}^{\infty}$ which converges to $\hat{z}$, i.e. $z_i \to \hat{z}$ as $i \to \infty$, $i \in K$, then $v_i \to 0$ as $i \to \infty$, $i \in K$.

Proof: Note that

$$\|v_i\|_{\infty} \leq \nu_i \triangleq \min\left\{\|v\|_{\infty} \middle| g(z_i) + G(z_j)\, v = 0,\ \bar{f}^{+}(z_i) + \bar{F}(z_j)\, v \overset{<}{=} 0\right\} \tag{17}$$

and $\nu_i \to 0$ as $i \to \infty$, $i \in K$ by (56) and Lemma 2. Hence $v_i \to 0$ as $i \to \infty$, $i \in K$. ¤

Corollary 3: Suppose that $\{z_i\}_{i=0}^{\infty}$ is a seqpence constructed by the algorithm, and suppose that $\{z_i\}_{i\in K}$ is a subsequence converging to a solution $\hat{z}$ of (1), where $K \subset \{0,1,2,...\}$. Then there is an infinite subset $K' \subset K$ such that $z_{i+1} = z_i + v_i$ for all $i \in K'$.

Proof: By Corollary 2, $v_i \to 0$ as $i \to \infty$, $i \in K$. Hence, since $f^o(z_i) \to 0$ as $i \to \infty$, $i \in K$, and $f^o(\cdot)$ is uniformly continuous on $\{z_i\}_{i\in K}$, there exists an integer i'' such that for all $i \in K$, $i \geq i''$, $\|v_i\|_\infty \leq \gamma$,

$$f^o(z_i) \leq \frac{1}{2} f^o(z_0) \tag{18}$$

$$f^o(z_i+v_i) - f^o(z_i) \leq \frac{1}{2} f^o(z_0) \tag{19}$$

i.e., $z_i + v_i \in C(z_0)$ for all $i \geq i''$, $i \in K$; moreover the construction $z_{i+1} = z_i + v_i$, $i \in K$, must occur at least once, so that p_i'' is well defined. Now let p_i, $i \in K$, $i \geq i''$ be an arbitrary integer. Then, since $v_i \to 0$ as $i \to \infty$, $i \in K$, there exists a j such that $i + j \in K$ and $\|v_{i+j}\|_\infty \leq \gamma^{p_i+1}$, which together with (18), (19) implies that $z_{i+j+1} =$

$z_{i+j} + v_{i+j}$. Thus there exists an infinite subset $K' \subset K$ such that $z_{i+1} = z_i + v_i$ for all $i \subset K'$. ∎

<u>Theorem 1 (Local Convergence)</u>: Let $g: \mathbb{R}^n \to \mathbb{R}^m$ and $f: \mathbb{R}^n \to \mathbb{R}^q$ be differentiable functions, with $G(\cdot)$, $\bar{F}(\cdot)$ uniformly Lipschitz continuous with constant L. Let $b > 0$, $\gamma \in (0,1)$ and an integer $k \geq 1$ be given. Suppose there exists a $y_0 \in \mathbb{R}^n$ such that the pair $[\bar{F}(y_0), G(y_0)]$ satisfies the LI condition and (see also (58))

$$h \triangleq \mu^* L \eta \leq \gamma - \gamma^2 , \quad \mu^* \geq \sup_{y \in Y} \mu[\bar{F}(y), G(y)]$$

$$Y = \left\{ y \middle| \|y - y_0\|_\infty \leq \frac{\eta}{1-\delta} \right\}, \quad \delta = \frac{1}{2} - \sqrt{\frac{1}{4} - h} \tag{20}$$

and

$$\eta \triangleq \min\{\|v\|_\infty | g(y_0) + G(y_0)\, v = 0,\ \bar{f}(y_0) + \bar{F}(y_0)\, v \leq 0\} \tag{21}$$

For $i = 0,1,2,\ldots$, let $j(i) \triangleq k[i/k]$, where $[i/k]$ denotes the integer part of i/k. Then the iterative process

$$y_{i+1} \in \text{Arg}\min\{\|y - y_i\|_\infty | g(y_i) + G(y_{j(i)})(y - y_i) = 0,$$

$$\bar{f}(y_i) + \bar{F}(y_{j(i)})\,(y - y_i) \leq 0\} \quad i = 0,1,2,\ldots, \tag{22}$$

results in well defined sequences $\{y_i\}_{i=0}^{\infty}$ such that any such sequence converges to some $\hat{y}$ satisfying $g(\hat{y}) = 0$, $\bar{f}(\hat{y}) \leq 0$.

Proof: First suppose that the process (22) results in a well defined sequence $\{y_i\}_{i=0}^{\infty} \subset Y$. For $i = 0,1,2,\ldots$, consider the following associated linear system:

$$g(y_i) + G(y_{j(i)})(y-y_i) - [g(y_{i-1}) + G(y_{j(i-1)})(y_i-y_{i-1})] = 0 \tag{23}$$

$$\bar{f}(y_i) + \bar{F}(y_{j(i)})(y-y_i) - [\bar{f}(y_{i-1}) + \bar{F}(y_{j(i-1)})(y_i-y_{i-1})] \leq 0 \tag{24}$$

Since, by inspection, any solution to (23)-(24) is a feasible point for (22), we obtain the following bound from (3), (56) and the Lagrange formula:

$$\begin{aligned}\|y_{i+1} - y_i\|_\infty &\leq \mu_{j(i)} \|h(y_i) - h(y_{i-1}) - H(y_{j(i-1)})(y_i - y_{i-1})\|_\infty \\ &\leq \mu_{j(i)} \int_0^1 \|H(y_{i-1} + s(y_i - y_{i-1})) - H(y_{j(i-1)})\|_\infty \, ds \|y_i - y_{i-1}\|_\infty \\ &\leq \mu_{j(i)} L\left[\|y_{i-1} - y_{j(i-1)}\|_\infty + \|y_i - y_{i-1}\|_\infty\right] \|y_i - y_{i-1}\|_\infty \\ &\leq \mu_{j(i)} L \left[\sum_{\nu=j(i-1)}^{i-1} \|y_{\nu+1} - y_\nu\|_\infty\right] \|y_i - y_{i-1}\|_\infty \end{aligned} \tag{25}$$

Next, from (20) it follows that δ is real for any $\gamma \in (0,1)$ and

$$h = \delta - \delta^2 \leq \delta \leq \gamma. \tag{26}$$

We shall now show by induction that the process (22) is well defined for $i = 0,1,2,\ldots$, that $y_i \in Y$ for $i = 0,1,2,\ldots$, and that

$$\|y_{i+1} - y_i\|_\infty \leq \delta\|y_i - y_{i-1}\|_\infty, \quad i = 1,2,\ldots \,. \tag{27}$$

Note that a y_1 satisfying (22) for $i = 0$ exists by hypothesis, and that $y_1 \in Y$. Next, let $i = 1$. If $k > 1$, then $j(1) = 0$ and, by hypothesis, the Jacobians satisfy the LI condition at y_0. Consequently, by Theorem 5, there is at least one y_2 satisfying (22) for $i = 1$. If, on the other hand, $k = 1$, then because of the Lipschitz continuity of the Jacobians we have

$$\mu_0\|H(y_1) - H(y_0)\|_\infty \leq \mu_0 L\|y_1 - y_0\|_\infty \leq \mu^* L\eta = h < 1 \tag{28}$$

and hence by Theorem 5, the LI condition is satisfied for $i = 1$ and there exists at least one y_2 satisfying (22) for $i = 1$. It now follows from (25) that, whether $k = 1$ or $k > 1$,

$$\|y_2 - y_1\|_\infty \leq \mu_{j(1)}L[\|y_1 - y_0\|_\infty]\ \|y_1 - y_0\|_\infty \leq \delta\|y_1 - y_0\|_\infty \tag{29}$$

the last part of (29) holding because of (20) and (26). Also, $\|y_2 - y_0\|_\infty \leq \|y_2 - y_1\|_\infty + \|y_1 - y_0\|_\infty \leq (1+\delta)\eta \leq \frac{1}{1-\delta}\eta$, i.e., $y_2 \in Y$. Thus the sequence is well defined for $i=0,1,2$, it is contained in Y, and (27) holds for $i = 1$. We proceed with the inductive step: for $i = N + 1$, either $j(N+1) = j(N)$ or $j(N+1) = N + 1$. If $j(N+1) = j(N)$, then, since the matrices in the linear programming problem (22) satisfy the LI condition at $i = N$, they also satisfy it at $i = N+1$ and y_{N+2} is well defined. If $j(N+1) \neq j(N)$, then, because of the uniform Lipschitz continuity of the Jacobians,

$$\mu_{j(N)} \|H(y_{N+1}) - H(y_{j(N)})\|_\infty \leq \mu_{j(N)} L \|y_{N+1} - y_{j(N)}\|_\infty$$

$$\leq \mu^* L \sum_{\nu=j(N)}^{N} \|y_{\nu+1} - y_\nu\|_\infty \leq \mu^* L\eta \sum_{\nu=j(N)}^{N} \delta^\nu \leq h/(1-\delta) = \delta. \quad (30)$$

where the last line follows from (26). Hence by Theorem 5, the LI condition is satisfied for $i = N+1$ and thus y_{N+2} is well defined. It now follows from (25) and (26) that

$$\|y_{N+2} - y_{N+1}\|_\infty \leq \delta \|y_{N+1} - y_N\|_\infty \quad (31)$$

i.e., (27) holds for $i = N+1$. Thus, the sequence $\{y_i\}_{i=0}^{\infty}$ is well defined, it satisfies (27) and is contained in Y. It must converge to a $\hat{y}$ because (27) holds for all i. It now follows from the continuity of g, G, $\bar{f}$ and $\bar{F}$, and the fact that $v_i = (y_{i+1} - y_i)$ converges to zero as $i \to \infty$, that $g(\hat{y}) = 0$, $\bar{f}(\hat{y}) \leqq 0$. This completes our proof. ¤

<u>Theorem 2</u>: Suppose that b in the algorithm is sufficiently large so that $\{z | g(z) = 0, \bar{f}(z) \leqq 0\} \neq \phi$, and suppose that Assumptions 1 and 2 are satisfied. Suppose that the algorithm has constructed an infinite sequence $\{z_i\}_{i=0}^{\infty}$. Then (i) there exists an $N \geq 0$ such that $z_{i+1} = z_i + v_i$ for all $i \geq N$ and (ii) $z_i \to \hat{z}$ as $i \to \infty$ with $g(\hat{z}) = 0$, $f(\hat{z}) \leqq 0$.

<u>Proof</u>: According to Lemma 2, there exists an infinite sequence $\{z_i\}_{i\in K}$

such that $z_i \to \hat{z}$, and $f^o(z_i) \to f^o(\hat{z}) = 0$, as $i \to \infty$, $i \in K$, with $\hat{z}$ a solution of (1). By corollaries 2 and 3 there is an infinite subset $K' \subset K$, such that $v_i \to 0$ as $i \to \infty$, $i \in K'$ and $z_{i+1} = z_i + v_i$, for all $i \in K'$.

Next, let $\varepsilon_1 = f^o(z_0)/2$. Then, since $f^o(\cdot)$ is uniformly continuous on the compact set $C(z_0)$, there exists a $\delta_1 > 0$ such that

$$|f^o(z) - f^o(z')| < \varepsilon_1 \tag{32}$$

for all $\|z' - z\|_\infty \leq \delta_1$. Now, since $f^o(z_i) \to 0$, $v_i \to 0$ as $i \to \infty$, $i \in K'$, there exists an $i' \in K'$ such that

$$f^o(z_{i'}) \leq \varepsilon_1 \ , \quad \|v_{i'}\|_\infty \leq (\gamma-\gamma^2)/\mu^* L \tag{33}$$

and

$$\omega \triangleq \frac{\|v_{i'}\|_\infty}{\frac{1}{2} + \sqrt{\frac{1}{4} - \mu^* L \|v_{i'}\|_\infty}} \leq \delta_1 \tag{34}$$

Then the conditions of Theorem 1 are satisfied at $y_0 = z_{i'}$ and any sequence $\{y_j\}_{j=0}^{\infty}$ constructed according to (22) must converge to a $\hat{y}$, which solves (1). Furthermore, from (27), $\|y_j - y_0\|_\infty \leq \omega \leq \delta_1$ for all $j = 1,2,\ldots$, and hence, from (32) and (33),

$$f^o(y_j) \leq f^o(y_0) + \varepsilon_1 \leq f^o(z_0) \tag{35}$$

i.e. $y_j \in C(z_0)$ for $j = 0,1,2,\ldots$. Hence, we must have $z_{i'+1} = y_1$. Now making use of (27) and (26) we obtain

$$\|v_{i'+1}\|_\infty = \|y_2 - y_1\|_\infty \leq \delta\|v_{i'}\|_\infty \leq \gamma\|v_{i'}\|_\infty \leq \gamma^{p_{i'}+1} = \gamma^{p_{i'}+1} \tag{36}$$

and hence $z_{i'+2} = y_2$. Proceeding by induction, we now conclude that $z_{i'+j} = y_j$, $j=0,1,2,\ldots$, and consequently $z_i \to \hat{y}$ a solution of (1). This completes our proof. ¤

3. Rate of Convergence and Efficiency

We derive now the root rate of convergence of the algorithm in Section 2 (see [4] sec. 9.1 for a definition and discussion of root rate). We shall show that the root rate of the iterative process (22), and hence of the process defined in Section 2, is $\geq r \triangleq \sqrt[k]{k+1}$. The proof will proceed in three steps. We shall first show that the sequence $\{s_{ik}\}_{i=0}^{\infty}$ of step lengths ($s_{ik} \triangleq \|y_{ik+1} - y_{ik}\|_\infty$) associated with the k-step process obtained from (22) has an R-rate $\geq$ k+1. (Compare with results of Traub [10] and Shamanskii [9] for multistep methods.) Next we shall obtain a relation between the rate of convergence of the subsequence $\{s_{ik}\}_{i=0}^{\infty}$ and that of the sequence $\{s_i\}_{i=0}^{\infty}$. Finally, we shall show that the R-rate of the process (22) is the same as the R-rate of the sequence $\{s_i\}_{i=0}^{\infty}$.

Define

$$s_i \triangleq \|y_{i+1} - y_i\|_\infty \tag{37}$$

Then we have the following

Lemma 3: Let $g: \mathbb{R}^n \to \mathbb{R}^m$ and $f: \mathbb{R}^n \to \mathbb{R}^q$ be differentiable functions, with $G(\cdot)$ and $\bar{F}(\cdot)$ uniformly Lipschitz continuous with constant L. Let $b > 0$, $\gamma \in (0,1)$, $k \in N^+$ be given. Suppose there exists a $y_0 \in \mathbb{R}^n$ such that the pair $[\bar{F}(y_0), G(y_0)]$ satisfies the LI condition and

$$s_\theta < \min\left\{\frac{1}{[a_k(\mu^*L)^k]^{k+1}}, \frac{\gamma-\gamma^2}{\mu^*L}, 1\right\}, \tag{38}$$

where a_k is defined by the recursive relation

$$a_{\ell+1} = a_\ell \sum_{i=0}^{\ell} a_i, \quad a_0 = 1 \qquad \ell = 0,1,\ldots,k-1 \; , \tag{39}$$

and μ^* satisfies (20).

(The existence of a y_1 solving (22) is ensured by the hypothesis).

Under these conditions, the sequence $\{s_i\}_{i=0}^{\infty}$ associated with the process (22) converges to 0 with root rate of at least $r \triangleq \sqrt[k]{k+1}$.†

Proof: From the local convergence theorem, the iterative process (22) constructs a well-defined sequence $\{y_i\}_{i=0}^{\infty}$. Moreover, this sequence is Cauchy, and hence the sequence $\{s_i\}_{i=0}^{\infty}$ converges to 0.

We begin by showing that for any $\ell \in \{1,2,\ldots k\}$ and for any $i = 0,1,\ldots,$ we can bound $s_{j(i)+\ell}$ by

$$s_{j(i)+\ell} \le a_\ell (\mu^* L)^\ell \, s_{j(i)}^{\ell+1} \tag{40}$$

with a_ℓ as above. The proof will be by induction on ℓ. From (25) we obtain, by replacing i with $j(i)+1$ and using (20) and the definition (37):

$$s_{j(i)+1} \le \mu^* L \, s_{j(i)}^2 \tag{41}$$

which proves (40) for the case when $\ell = 1$. Assume next that (40) holds for $\ell = 1,2,\ldots,\bar{k}$, with $\bar{k} \le k-1$. Making again use of (25) and the inductive hypothesis, we obtain successively:

† By theorem 9.2.7 in [4], r is the root rate of the sequence $\{s_i\}_{i=0}^{\infty}$ only if $0 < \overline{\lim_{i\to\infty}} \, s_i^{\frac{1}{r^i}} < 1$.

$$s_{j(i)+\bar{k}+1} \leq \mu^* L\left(\sum_{\nu=j(i)}^{j(i)+\bar{k}} s_\nu\right) s_{j(i)+\bar{k}}$$

$$= a_{\bar{k}}(\mu^* L)^{\bar{k}+1} s_{j(i)}^{\bar{k}+2} \sum_{\nu=0}^{\bar{k}} a_\nu (\mu^* L\, s_{j(i)})^\nu$$

$$\leq a_{\bar{k}+1}(\mu^* L)^{\bar{k}+1} s_{j(i)}^{\bar{k}+2} \tag{42}$$

where the last line follows from (38) and the fact that the sequence $\{s_i\}_{i=0}^{\infty}$ is a monotone decreasing sequence. The bound in (40) thus holds for any $\ell \in \{1,2,\ldots,k\}$ and, since i was arbitrary, it also holds for any $i \in N^+$.

A simple calculation will lead now to the desired result. Noting that $s_{j(i)+k} = s_{j(i+k)}$ and applying repeatedly (40) with $\ell=k$, we obtain

$$s_{j(i+k)} \leq a_k(\mu^* L)^k\, s_{j(i)}^{k+1} \leq \left\{a_k(\mu^* L)^k\right\}^{\sum_{\nu=0}^{j(i)/k} (k+1)^\nu} s_0^{r^{j(i+k)}} \,. \tag{43}$$

Observe next that for any $i \in N^+$ and for any integer $k \geq 1$,

$$\sum_{\nu=0}^{\frac{j(i)}{k}} (k+1)^\nu = \frac{(k+1)^{\frac{j(i)}{k}+1} - 1}{k} < (k+1)^{\frac{j(i)}{k}+2} \tag{44}$$

Without loss of generality, we can assume that $a_k(\mu^* L)^k \geq 1$ (since if it is not, we can choose a larger Lipschitz constant). We can then bound (43) using the inequality (44) by:

$$s_{j(i+k)} \leq \left\{a_k(\mu^* L)^k\right\}^{(k+1)^{(\frac{j(i)}{k}+2)}} s_0^{r^{j(i+k)}}$$

$$= \left\{[a_k(\mu^* L)^k]^{k+1} s_0\right\}^{r^{j(i+k)}} \triangleq \theta^{r^{j(i+k)}} \tag{45}$$

Since we can express any $i = 1,2,\ldots$ as $i = j(i-1) + \ell$ for some $\ell \in \{1,2,\ldots,k\}$, we can combine (40) and (45) to obtain:

$$\overline{\lim_{i\to\infty}}\; s_i^{\frac{1}{r^i}} \le \overline{\lim_{i\to\infty}} \left[a_\ell(\mu^* L)^\ell \, s_{j(i-1)}^{\ell+1}\right]^{\frac{1}{r^i}}$$

$$\le \overline{\lim_{i\to\infty}}\; \theta^{\frac{r^{j(i-1)}(\ell+1)}{r^{j(i-1)}\, r^\ell}} = \theta^{\frac{\ell+1}{r^\ell}} < 1 . \tag{46}$$

Hence the sequence $\{s_i\}_{i=0}^\infty$ has a root-rate of convergence $\ge r$ (by Theorem 9.2.7 in [4]), and this completes the proof of the lemma. ¤

Corollary 4: Under the conditions of the preceding lemma, the following bound holds for s_i

$$s_i \le A\, \theta^{\lambda r^i} \qquad \text{for} \quad i = 1,2,\ldots \tag{47}$$

where $\quad A \triangleq a_k(\mu^* L)^k \quad$ and $\quad \lambda \triangleq \min\left\{\frac{\ell+1}{r^\ell} \mid \ell \in \{1,2,\ldots,k\}\right\} < 1$ ¤

Theorem 3: Assume the conditions of Lemma 3 hold. Then the root-rate of any sequence defined by (22) is at least $r = \sqrt[k]{k+1}$.

Proof: According to Theorem 9.2.7 in [4], we only need to show that

$$\overline{\lim_{i\to\infty}}\; \|y_i - \hat{y}\|_\infty^{\frac{1}{r^i}} < 1 \tag{48}$$

Obviously,

$$e_i \triangleq \|y_i - \hat{y}\|_\infty \le \sum_{\nu=i}^{\infty} s_\nu \le \sum_{\nu=i}^{\infty} A\, \theta^{\lambda r^\nu} = A\, \theta^{\lambda r^i} \sum_{\nu=i}^{\infty} \theta^{\lambda(r^\nu - r^i)}$$

$$= A\, \theta^{\lambda r^i} \sum_{\nu=0}^{\infty} \theta^{\lambda r^i (r^\nu - 1)} \tag{49}$$

The series appearing in the last line of (49) converges for any $i \in N^+$, since it satisfies the root test for series. Moreover, denoting the infinite series in (49) by b_i, we observe that $\{b_i\}_{i=0}^{\infty}$ is a decreasing sequence (since $\theta^{\lambda r^{i+1}} < \theta^{\lambda r^i}$). Therefore, the bound in (49) becomes:

$$e_i \leq A\, b_0\, \theta^{\lambda r^i} \tag{50}$$

which is equivalent to (48). ¤

Thus we have proved the following:

Theorem 4: Suppose that b in the algorithm is sufficiently large so that $\{z | g(z) = 0,\ \bar{f}(z) \leqq 0\} \neq \phi$, and suppose that Assumptions 1 and 2 are satisfied. Suppose that the algorithm has constructed an infinite sequence $\{z_i\}_{i=0}^{\infty}$. Then (i) $z_i \to \hat{z}$ as $i \to \infty$, with $g(\hat{z}) = 0$, $\bar{f}(\hat{z}) \leqq 0$,

$$\text{and (ii) } \overline{\lim_{i\to\infty}}\ \|z_i - \hat{z}\|_{\infty}^{\frac{1}{r^i}} < 1 \text{ where } r \triangleq \sqrt[k]{k+1}. \quad ¤$$

We now show how to choose an optimal value of k. Brent [2] defined the efficiency of an algorithm as:

$$E \triangleq \frac{\ln r}{w} \tag{51}$$

where r is the root rate of convergence of the algorithm and w is the average amount of work per iteration (e.g. number of function evaluations, CPU time, etc).

For the algorithm in Section 2 we have, by Theorem 4

$$E_k = \frac{\ln (k+1)}{k\, w(k)} \tag{52}$$

It can be easily seen that using any reasonable definition for w, E_k will attain a maximum value for some $k \in N^+$. The value of the maximizer, k_{opt}, can be obtained either by experimentally evaluating w(k), or by making use of an explicit expression for $w(k)$ when available. For example, if the number of function evaluations is the dominant factor in an iteration (as is the case in boundary value problems), then E_k has the form:

$$E_k = \frac{\ell n(k+1)}{k\,\frac{(m+q)(n+k)}{k}} = \frac{\ell n(k+1)}{(m+q)(n+k)} \tag{53}$$

and the value of k_{opt} can be determined a priori.

Conclusion

A limited amount of experimental evidence indicates that the algorithm works well even when the LI condition, which is sufficient to ensure global convergence of the algorithm, is not satisfied at all points of $C(z_0)$.

Finally, it should be observed that the algorithm can be easily adapted for solving boundary value problems with ordinary differential equations.

Appendix

We present below a summary of the results concerning the solution of mixed systems of equations and inequalities which have been used in this paper.

Consider the system

$$Az \leq b \qquad Cz = d \tag{54}$$

where $A \in \mathbb{R}^{(q+1)\times n}$, $C \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^{q+1}$, $d \in \mathbb{R}^{m}$, and let $\mathcal{F}$ be the set of right-hand side vectors for which (54) has at least one solution:

$$\mathcal{F} \triangleq \{\binom{b}{d} \mid \exists\ z \in \mathbb{R}^n \text{ s.t. } Az \leq b,\ Cz = d\} \tag{55}$$

It has been shown (e.g. [8]) that $\mu(A,C)$ defined by:

$$\mu(A,C) \triangleq \max\{\min\{\|z\|_\infty | Az \leq b,\ Cz = d\} \mid \|{b \atop d}\|_\infty < 1,\ ({b \atop d}) \in \mathcal{F}\} \qquad (56)$$

is finite.

Theorem 5 [7]: Assume that the pair (A',C') satisfies the LI condition (see Assumption 2), and let $\mu = \mu(A',C')$. Let $\Delta A \in \mathbb{R}^{(q+1)\times n}$, $\Delta C \in \mathbb{R}^{m\times n}$ with $\delta \triangleq \|{\Delta A \atop \Delta C}\|_\infty$. If $\mu\delta < 1$, then the system (54), with $A = A' + \Delta A$, $C = C' + \Delta C$, has a solution for any right-hand side $({b \atop d})$. Moreover,

$$\mu(A,C) \leq \frac{\mu}{1-\mu\delta}\ . \qquad ¤ \qquad (57)$$

Proposition 2: Let $\mu: \mathbb{R}^{(q+1)\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}^{+}$ be defined by (56). Then μ is an upper-semicontinuous function. ¤

Theorem 6: Let X be a compact subset of $\mathbb{R}^n$. Let $\bar{F} : X \to \mathbb{R}^{(q+1)\times n}$, $G : X \to \mathbb{R}^{m\times n}$ be continuous functions. Assume that for all $z \in X$, the pair $(\bar{F}(z), G(z))$ satisfies the LI condition. Then

$$\mu^* \triangleq \max_{z\in X} \mu[\bar{F}(z),\ G(z)] \qquad (58)$$

exists and $\mu^* > 0$.

REFERENCES

[1] L. Armijo, "Minimization of Functions Having Continuous Partial Derivatives," Pacific J. Math., 16, 1-3, 1966.

[2] R. P. Brent, "Some Efficient Algorithms for Solving Systems of Nonlinear Equations," SIAM J. Num. Anal., 10, 2, 327-344, 1973.

[2'] T. J. Huang, "Algorithms for Solving Systems of Equations and Inequalities with Applications in Nonlinear Programming," Computer Sciences Technical Report #191, Univ. of Wisc., 1973.

[3] L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces, McMillan, New York, 1964.

[4] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, 1970.

[5] E. Polak, Computational Methods in Optimization: A Unified Approach, Academic Press, 1971.

[6] B. N. Pshenichnyi, "Newton's Method for the Solution of Systems of Equalities and Inequalities," Math. Notes Acad. Sc. USSR, 8, 827-830, 1970.

[7] S. M. Robinson, "Extension of Newton's Method to Mixed Systems of Nonlinear Equations and Inequalities," Technical Summary Report No. 1161, Mathematics Research Center, Univ. of Wisc., 1971.

[8] S. M. Robinson, "Bounds for Errors in the Solution Set of a Perturbed Linear System," Lin. Alg. Appl., 6, 69-81, 1971.

[9] V. Shamanskii, "On a Modification of the Newton Method," Ukrain. Mat. J., 19, 133-138, 1967.

[10] J. Traub, Iterative Methods for the Solution of Equations, Prentice Hall, New Jersey, 1964.