SIAM J. NUMER. ANAL.
Vol. 19, No. 2, April 1982

0036-1429/82/1902-0013 $01.00/0

INEXACT NEWTON METHODS*

RON S. DEMBO†, STANLEY C. EISENSTAT‡ AND TROND STEIHAUG§

Abstract. A classical algorithm for solving the system of nonlinear equations $F(x)=0$ is Newton's method:

$$x_{k+1}=x_k+s_k, \quad \text{where } F'(x_k)s_k=-F(x_k), \quad x_0 \text{ given.}$$

The method is attractive because it converges rapidly from any sufficiently good initial guess x_0. However, solving a system of linear equations (the Newton equations) at each stage can be expensive if the number of unknowns is large and may not be justified when x_k is far from a solution. Therefore, we consider the class of *inexact Newton methods*:

$$x_{k+1}=x_k+s_k, \quad \text{where } F'(x_k)s_k=-F(x_k)+r_k, \quad \|r_k\|/\|F(x_k)\|\leqq \eta_k$$

which solve the Newton equations only approximately and in some *unspecified* manner. Under the natural assumption that the forcing sequence $\{\eta_k\}$ is uniformly less than one, we show that all such methods are locally convergent and characterize the order of convergence in terms of the rate of convergence of the relative residuals $\{\|r_k\|/\|F(x_k)\|\}$. Finally, we indicate how these general results can be used to construct and analyze specific methods for solving systems of nonlinear equations.

1. Introduction.

Consider the system of nonlinear equations

$$F(x)=0, \tag{1.1}$$

where $F: R^n \to R^n$ is a nonlinear mapping with the following properties:

(1) There exists an $x^* \in R^n$ with $F(x^*)=0$.
(2) F is continuously differentiable in a neighborhood of x^*.[1]
(3) $F'(x^*)$ is nonsingular.

A classical algorithm for finding a solution to (1.1) is Newton's method. Given an initial guess x_0, we compute a sequence of steps $\{s_k\}$ and iterates $\{x_k\}$ as follows:

(1.2)
FOR $k=0$ STEP 1 UNTIL Convergence DO
  Solve $F'(x_k)s_k=-F(x_k)$
  Set $x_{k+1}=x_k+s_k$.

Newton's method is attractive because it converges rapidly from any sufficiently good initial guess. Indeed, it is a standard with which to compare rapidly convergent methods for solving (1.1), since one way of characterizing superlinear convergence is that the step should approach the Newton step asymptotically in both magnitude and direction (see Dennis and Moré [5]).

One drawback of Newton's method is having to solve the Newton equations (1.2) at each stage. Computing the exact solution using a direct method such as Gaussian elimination can be expensive if the number of unknowns is large and may not be justified when x_k is far from x^*. Therefore, it seems reasonable to use an iterative

* Received by the editors October 9, 1980, and in final revised form April 13, 1981. This research was supported in part by the U.S. Department of Transportation under grant CT-06-0011, the National Science Foundation under grant ENG-78-21615, the Office of Naval Research under grant N0001--76-C-0277 and a grant from the Norwegian Research Council for Science and the Humanities.

† School of Organization and Management, Yale University, New Haven, Connecticut 06520.

‡ Research Center for Scientific Computation, Department of Computer Science, Yale University, New Haven, Connecticut 06520.

§ Department of Mathematical Sciences, Rice University, Houston, Texas, 77001.

[1] For the analysis to come, it would suffice that the Jacobian exist in a neighborhood of x^* and be continuous at x^*.

method and to solve (1.2) only approximately. A natural stopping rule would be based on the size of the relative residual[2] $\|r_k\|/\|F(x_k)\|$, where, if s_k is the step actually computed (i.e., the approximate solution to (1.2)), then the residual r_k is given by

$$r_k \equiv F'(x_k)s_k + F(x_k).$$

Such *Newton-iterative methods* (see [9], [11]) offer a trade-off between the accuracy with which the Newton equations are solved and the amount of work per iteration. An important question is what level of accuracy is required to preserve the rapid local convergence of Newton's method?

More generally, we consider the class of *inexact Newton methods* which compute an approximate solution to the Newton equations in some *unspecified* manner such that

$$\frac{\|r_k\|}{\|F(x_k)\|} \leqq \eta_k,$$

where the nonnegative *forcing sequence* $\{\eta_k\}$ is used to control the level of accuracy (cf. Altman [1]). To be precise, an inexact Newton method is any method which, given an initial guess x_0, generates a sequence $\{x_k\}$ of approximations to x^* as follows:

FOR $k=0$ STEP 1 UNTIL Convergence DO

Find *some* step s_k which satisfies

$$F'(x_k)s_k = -F(x_k) + r_k, \quad \text{where } \frac{\|r_k\|}{\|F(x_k)\|} \leqq \eta_k \tag{1.3}$$

Set $x_{k+1} = x_k + s_k$.

Here η_k may depend on x_k; taking $\eta_k \equiv 0$ gives Newton's method.

This paper analyzes the local behavior of such inexact Newton methods. In § 2, we prove that these methods are locally convergent if the forcing sequence is uniformly less than one. In § 3, we characterize the order of convergence and indicate how to choose a forcing sequence (*and thus how to construct an inexact Newton method*) which preserves the rapid convergence of Newton's method. In § 4, we discuss how these results can be used to analyze specific methods for solving (1.1). Generalizations of these results to inexact quasi-Newton methods are given by Steihaug [12].

2. Local convergence of inexact Newton methods. In this section, under the rather weak assumption that the forcing sequence $\{\eta_k\}$ is uniformly less than one, we show that inexact Newton methods are locally convergent, i.e., that the sequence of iterates $\{x_k\}$ converges to x^* from any sufficiently good initial guess x_0. This requirement is natural in that $s_k=0$ satisfies (1.3) if $\eta_k \geqq 1$.

LEMMA 2.1. (Ortega and Rheinboldt [9, 2.3.3]). *For any $\varepsilon>0$, there exists $\delta>0$ such that $F'(y)$ is nonsingular and*

$$\|F'(y)^{-1} - F'(x^*)^{-1}\| < \varepsilon$$

if $\|y-x^*\|<\delta$.

LEMMA 2.2. (cf. Ortega and Rheinboldt [9, 3.1.5]). *For any $\varepsilon>0$, there exists $\delta>0$ such that*

$$\|F(y) - F(x^*) - F'(x^*)(y-x^*)\| \leqq \varepsilon\|y-x^*\|$$

if $\|y-x^*\|<\delta$.

[2] Here $\|\cdot\|$ denotes an arbitrary norm in R^n and the induced operator norm.

THEOREM 2.3. *Assume that* $\eta_k \leqq \eta_{\max} < t < 1$. *There exists* $\varepsilon > 0$ *such that, if* $\|x_0 - x^*\| \leqq \varepsilon$, *then the sequence of inexact Newton iterates* $\{x_k\}$ *converges to* x^*. *Moreover, the convergence is linear in the sense that*

$$\|x_{k+1} - x^*\|_* \leqq t\|x_k - x^*\|_*, \tag{2.1}$$

where $\|y\|_* \equiv \|F'(x^*)y\|$.

Proof. Since $F'(x^*)$ is nonsingular,

$$\frac{1}{\mu}\|y\| \leqq \|y\|_* \leqq \mu\|y\| \quad \text{for } y \in R^n, \tag{2.2}$$

where

$$\mu \equiv \max\,[\|F'(x^*)\|, \|F'(x^*)^{-1}\|].$$

Since $\eta_{\max} < t$, there exists $\gamma > 0$ sufficiently small that

$$(1+\gamma\mu)[\eta_{\max}(1+\mu\gamma) + 2\mu\gamma] \leqq t.$$

Now, choose $\varepsilon > 0$ sufficiently small that

$$\|F'(y) - F'(x^*)\| \leqq \gamma, \tag{2.3}$$

$$\|F'(y)^{-1} - F'(x^*)^{-1}\| \leqq \gamma, \tag{2.4}$$

$$\|F(y) - F(x^*) - F'(x^*)(y - x^*)\| \leqq \gamma\|y - x^*\| \tag{2.5}$$

if $\|y - x^*\| \leqq \mu^2\varepsilon$. Such an ε exists by virtue of the continuity of F' at x^*, Lemma 2.1 and Lemma 2.2.

Assume that $\|x_0 - x^*\| \leqq \varepsilon$. We prove (2.1) by induction. Note that, by (2.2), the induction hypothesis and (2.2) again,

$$\|x_k - x^*\| \leqq \mu\|x_k - x^*\|_* \leqq \mu t^k \|x_0 - x^*\|_* \leqq \mu^2\|x_0 - x^*\| \leqq \mu^2\varepsilon,$$

so that (2.3)–(2.5) hold with $y = x_k$. Moreover, the kth stage of the inexact Newton method is defined in the sense that there exist s_k which satisfy (1.3). Since

$$\begin{aligned} F'(x^*)(x_{k+1} - x^*) &= [I + F'(x^*)[F'(x_k)^{-1} - F'(x^*)^{-1}]] \\ &\quad \cdot [r_k + [F'(x_k) - F'(x^*)](x_k - x^*) - [F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)]] \end{aligned}$$

taking norms,

$$\begin{aligned} \|x_{k+1} - x^*\|_* &\leqq [1 + \|F'(x^*)\|\,\|F'(x_k)^{-1} - F'(x^*)^{-1}\|] \\ &\quad \cdot [\|r_k\| + \|F'(x_k) - F'(x^*)\|\,\|x_k - x^*\| + \|F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)\|] \\ &\leqq (1+\mu\gamma)[\eta_k\|F(x_k)\| + \gamma\|x_k - x^*\| + \gamma\|x_k - x^*\|], \end{aligned}$$

using the definition of μ, (2.4), (1.3), (2.3) and (2.5). Since

$$F(x_k) = [F'(x^*)(x_k - x^*)] + [F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)],$$

taking norms,

$$\|F(x_k)\| \leqq \|x_k - x^*\|_* + \|F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)\| \leqq \|x_k - x^*\|_* + \gamma\|x_k - x^*\|,$$

using (2.5). Therefore,

$$\begin{aligned} \|x_{k+1} - x^*\|_* &\leqq (1+\mu\gamma)[\eta_k\|x_k - x^*\|_* + \gamma\|x_k - x^*\|] + 2\gamma\|x_k - x^*\|] \\ &\leqq (1+\mu\gamma)[\eta_{\max}(1+\mu\gamma) + 2\mu\gamma]\|x_k - x^*\|_* \end{aligned}$$

using (2.2). The result now follows from the choice of γ. Q.E.D.

Theorem 2.3 shows that inexact Newton methods are locally convergent. Krasnosel'skii and Rutickii [7] proved a Newton–Kantorovich theorem for such methods with the bound on the relative residual replaced by a bound on the relative error in the approximate solution to the Newton equations. Altman [1] proved a similar result for damped inexact Newton methods.

3. Rate of convergence of inexact Newton methods.

In this section, we characterize the order of convergence of the inexact Newton iterates in terms of the rate of convergence of the relative residuals and indicate how the forcing sequence influences that rate of convergence.

DEFINITION. Let $\{x_k\}$ be a sequence which converges to x^*. Then

(1) $x_k \to x^*$ superlinearly if[3]

$$\|x_{k+1}-x^*\| = o(\|x_k - x^*\|) \quad \text{as } k \to \infty;$$

(2) $x_k \to x^*$ with (strong) order at least q $(q>1)$ if

$$\|x_{k+1}-x^*\| = O(\|x_k - x^*\|^q) \quad \text{as } k \to \infty;$$

(3) $x_k \to x^*$ with weak order at least $q(q>1)$ if

$$\limsup_{k\to\infty} \|x_k - x^*\|^{1/q^k} < 1 \quad \text{as } k \to \infty.$$

LEMMA 3.1. *Let*

$$\alpha \equiv \max\left[\|F'(x^*)\| + \frac{1}{2\beta}, 2\beta\right],$$

where $\beta \equiv \|F'(x^*)^{-1}\|$. *Then*

$$\frac{1}{\alpha}\|y-x^*\| \leqq \|F(y)\| \leqq \alpha \|y-x^*\|$$

for $\|y-x^*\|$ *sufficiently small.*

Proof. By Lemma 2.2, there exists $\delta>0$ sufficiently small that

$$\|F(y)-F(x^*)-F'(x^*)(y-x^*)\| \leqq \frac{1}{2\beta}\|y-x^*\|$$

if $\|y-x^*\|<\delta$. Since

$$F(y) = [F'(x^*)(y-x^*)] + [F(y)-F(x^*)-F'(x^*)(y-x^*)],$$

taking norms,

$$\begin{aligned}\|F(y)\| &\leqq \|F'(x^*)\|\,\|y-x^*\| + \|F(y)-F(x^*)-F'(x^*)(y-x^*)\| \\ &\leqq \left[\|F'(x^*)\| + \frac{1}{2\beta}\right]\|y-x^*\|,\end{aligned}$$

[3] Let $\{x_k\}$ be any sequence which converges to x^*. Given continuous nonnegative real-valued functions g and h, we write

$$g(x_k) = o(h(x_k)) \quad \text{as } k \to \infty \quad \text{if } \limsup_{k\to\infty} \frac{g(x_k)}{h(x_k)} = 0,$$

and

$$g(x_k) = O(h(x_k)) \quad \text{as } k \to \infty \quad \text{if } \limsup_{k\to\infty} \frac{g(x_k)}{h(x_k)} < +\infty.$$

and

$$\|F(y)\| \geqq \|F'(x^*)^{-1}\|^{-1}\|y-x^*\| - \|F(y)-F(x^*)-F'(x^*)(y-x^*)\|$$

$$\geqq \left[\|F'(x^*)^{-1}\|^{-1} - \frac{1}{2\beta}\right]\|y-x^*\| = \frac{1}{2\beta}\|y-x^*\|$$

if $\|y-x^*\|<\delta$. Q.E.D.

DEFINITION. F' is Hölder continuous with exponent p $(0<p\leqq 1)$ at x^* if there exists $L\geqq 0$ such that

$$\|F'(y)-F'(x^*)\| \leqq L\|y-x^*\|^p$$

for $\|y-x^*\|$ sufficiently small.

LEMMA 3.2 (cf. Ortega and Rheinboldt [9, 3.2.12]). *If F' is Hölder continuous with exponent p at x^*, then there exists $L'\geqq 0$ such that*

$$\|F(y)-F(x^*)-F'(x^*)(y-x^*)\| \leqq L'\|y-x^*\|^{1+p}$$

for $\|y-x^*\|$ sufficiently small.

THEOREM 3.3. *Assume that the inexact Newton iterates $\{x_k\}$ converge to x^*. Then $x_k \to x^*$ superlinearly if and only if*

$$\|r_k\| = o(\|F(x_k)\|) \quad \textit{as } k\to\infty.$$

Moreover, if F' is Hölder continuous with exponent p at x^*, then $x_k\to x^*$ with order at least $1+p$ if and only if

$$\|r_k\| = O(\|F(x_k)\|^{1+p}) \quad \textit{as } k\to\infty.$$

Proof. Assume that $x_k\to x^*$ superlinearly. Since

$$r_k = [F(x_k)-F(x^*)-F'(x^*)(x_k-x^*)] - [F'(x_k)-F'(x^*)](x_k-x^*)$$
$$+[F'(x^*)+[F'(x_k)-F'(x^*)]](x_{k+1}-x^*),$$

taking norms,

$$\|r_k\| \leqq \|F(x_k)-F(x^*)-F'(x^*)(x_k-x^*)\| + \|F'(x_k)-F'(x^*)\|\,\|x_k-x^*\|$$
$$+[\|F'(x^*)\|+\|F'(x_k)-F'(x^*)\|]\|x_{k+1}-x^*\|$$
$$= o(\|x_k-x^*\|) + o(1)\|x_k-x^*\| + [\|F'(x^*)\|+o(1)]o(\|x_k-x^*\|),$$

by Lemma 2.2, the continuity of F' at x^* and the assumption that $x_k\to x^*$ superlinearly. Therefore,

$$\|r_k\| = o(\|x_k-x^*\|) = o(\|F(x_k)\|) \quad \text{as } k\to\infty,$$

by Lemma 3.1.

Conversely, assume that $\|r_k\| = o(\|F(x_k)\|)$. As in the proof of Theorem 2.3.

$$\|x_{k+1}-x^*\| \leqq [\|F'(x^*)^{-1}\| + \|F'(x_k)^{-1}-F'(x^*)^{-1}\|]$$
$$\cdot[\|r_k\| + \|F'(x_k)-F'(x^*)\|\,\|x_k-x^*\| + \|F(x_k)-F(x^*)-F'(x^*)(x_k-x^*)\|]$$
$$= [\|F'(x^*)^{-1}\|+o(1)]\cdot[o(\|F(x_k)\|) + o(1)\|x_k-x^*\| + o(\|x_k-x^*\|)]$$

by Lemma 2.1, the assumption that $\|r_k\| = o(\|F(x_k)\|)$, the continuity of F' at x^* and Lemma 2.2. Therefore,

$$\|x_{k+1}-x^*\| = o(\|F(x_k)\|) + o(\|x_k-x^*\|) = o(\|x_k-x^*\|) \quad \text{as } k\to\infty$$

by Lemma 3.1.

If F' is Hölder continuous at x^*, then the proof is essentially the same, using the Hölder continuity of F' and Lemma 3.2 instead of the continuity of F' and Lemma 2.2. Q.E.D.

Theorem 3.3 characterizes the order of convergence of the inexact Newton iterates in terms of the rate of convergence of the relative residuals. Dennis and Moré [4], [5] proved an equivalent result in terms of the steps $\{s_k\}$ for quasi-Newton methods. Expressed in our notation, $x_k \to x^*$ superlinearly if and only if

$$\|r_k\| = o(\|s_k\|) \quad \text{as } k \to \infty,$$

and $x_k \to x^*$ with order at least $1+p$ if and only if

$$\|r_k\| = O(\|s_k\|^{1+p}) \quad \text{as } k \to \infty,$$

provided F' is Hölder continuous with exponent p at x^*. It should be noted that these conditions were not originally stated in terms of the residuals r_k and that, since they are not scale-invariant, they do not suggest a natural criterion for when to accept an approximate solution to the Newton equations.

The following result gives an analogous characterization of weak order of convergence.

THEOREM 3.4. *Assume that the inexact Newton iterates $\{x_k\}$ converge to x^*. If F' is Hölder continuous with exponent p at x^*, then $x_k \to x^*$ with weak order at least $1+p$ if and only if $r_k \to 0$ with weak order at least $1+p$.*

Proof. Let L be the Hölder constant, and let α and L' be the constants given in Lemma 3.1 and Lemma 3.2 respectively. Pick $\varepsilon > 0$ sufficiently small that $F'(y)$ is nonsingular and

$$\|F(y)\| \leqq \alpha \|y - x^*\|, \tag{3.1}$$

$$\|F'(y)\| \leqq \alpha, \tag{3.2}$$

$$\|F'(y)^{-1}\| \leqq \alpha, \tag{3.3}$$

$$\|F'(y) - F'(x^*)\| \leqq L \|y - x^*\|^p, \tag{3.4}$$

$$\|F(y) - F(x^*) - F'(x^*)(y - x^*)\| \leqq L' \|y - x^*\|^{1+p} \tag{3.5}$$

if $\|y - x^*\| \leqq \varepsilon$. Such an ε exists by Lemma 2.1, Lemma 3.1, the Hölder continuity of F' at x^* and Lemma 3.2.

Assume that $x_k \to x^*$ with weak order at least $1+p$. Then there exist constants $0 \leqq \gamma < 1$ and $k_0 \geqq 0$ such that

$$\|x_k - x^*\| \leqq \gamma^{(1+p)^k} \quad \text{for } k \geqq k_0. \tag{3.6}$$

Now choose $k_1 \geqq k_0$ sufficiently large that

$$\|x_k - x^*\| \leqq \varepsilon \quad \text{for } k \geqq k_1.$$

From the definition of r_k,

$$\|r_k\| \leqq \|F'(x_k)\| \, \|x_{k+1} - x_k\| + \|F(x_k)\| \leqq \|F'(x_k)\| [\|x_{k+1} - x^*\| + \|x_k - x^*\|] + \alpha \|x_k - x^*\|$$

by the triangle inequality and (3.1). Using (3.2) and (3.6),

$$\|r_k\| \leqq \alpha [\gamma^{(1+p)^{k+1}} + \gamma^{(1+p)^k}] + \alpha \gamma^{(1+p)^k} = [\alpha \gamma^{p(1+p)^k} + 2\alpha] \gamma^{(1+p)^k},$$

so that

$$\|r_k\| \leqq 3 \alpha \gamma^{(1+p)^k} \quad \text{for } k \geqq k_1.$$

The result now follows immediately from the definition of weak order of convergence.

Conversely, assume that $r_k \to 0$ with weak order at least $1+p$. Then there exist constants $0 \leqq \gamma < 1$ and $k_0 \geqq 0$ such that

$$\|r_k\| \leqq \gamma^{(1+p)^k} \quad \text{for } k \geqq k_0. \tag{3.7}$$

Let $c = [1/(2\alpha(L+L'))]^{1/p}$ and choose $k_1 \geqq k_0$ sufficiently large that

$$\|x_k - x^*\| \leqq \min\{\varepsilon, c\gamma\} \quad \text{for } k \geqq k_1, \tag{3.8}$$

$$\frac{\alpha}{c}\gamma^{(1+p)^k[1-(1+p)^{1-k_1}]} \leqq \frac{1}{2} \quad \text{for } k \geqq k_1. \tag{3.9}$$

From the definition of weak order of convergence it suffices to prove that

$$\|x_k - x^*\| \leqq c\gamma^{(1+p)^{k-k_1}} \quad \text{for } k \geqq k_1.$$

The proof is by induction. The result follows trivially from (3.8) when $k = k_1$. As in the proof of Theorem 2.3,

$$\begin{aligned}\|x_{k+1} - x^*\| &\leqq \|F'(x_k)^{-1}\| [\|r_k\| + \|F'(x_k) - F'(x^*)\| \, \|x_k - x^*\| \\ &\qquad + \|F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)\|] \\ &\leqq \alpha[\|r_k\| + L\|x_k - x^*\|^p \|x_k - x^*\| + L'\|x_k - x^*\|^{1+p}]\end{aligned}$$

by (3.3), (3.4) and (3.5). Using (3.7) and the induction hypothesis,

$$\begin{aligned}\|x_{k+1} - x^*\| &\leqq \alpha\gamma^{(1+p)^k} + \alpha(L+L')[c\gamma^{(1+p)^{k-k_1}}]^{1+p} \\ &= \left[\frac{\alpha}{c}\gamma^{(1+p)^k[1-(1+p)^{1-k_1}]} + \alpha(L+L')c^p\right] c\gamma^{(1+p)^{k+1-k_1}}\end{aligned}$$

By (3.9) and the definition of c,

$$\|x_{k+1} - x^*\| \leqq \left[\frac{1}{2} + \frac{1}{2}\right] c\gamma^{(1+p)^{k+1-k_1}} = c\gamma^{(1+p)^{k+1-k_1}},$$

and the induction is complete. Q.E.D.

Theorem 3.4 characterizes the weak order of convergence of the inexact Newton iterates in terms of the rate of convergence of the residuals. The sufficiency of forcing the residual to zero was proved for the case $p = 1$ by Pereyra [10].

The following result indicates how the forcing sequence influences the rate of convergence of the inexact Newton iterates.

COROLLARY 3.5. *Assume that the inexact Newton iterates* $\{x_k\}$ *converge to* x^*. *Then*:

(a) $x_k \to x^*$ *superlinearly if* $\lim_{k\to\infty} \eta_k = 0$;

(b) $x_k \to x^*$ *with order at least* $1+p$ *if* F' *is Hölder continuous with exponent* p *at* x^* *and*

$$\eta_k = O(\|F(x_k)\|^p) \quad \text{as } k \to \infty;$$

(c) $x_k \to x^*$ *with weak order at least* $1+p$ *if* F' *is Hölder continuous with exponent* p *at* x^* *and* $\{\eta_k\}$ *converges to* 0 *with weak order at least* $1+p$.

Corollary 3.5 indicates how the forcing sequence $\{\eta_k\}$ influences the rate of convergence of the inexact Newton iterates $\{x_k\}$. Part (a) was proved in the context of unconstrained optimization by McCormick and Ritter [8].

Corollary 3.5 is *constructive* in the sense that is specifies how to devise a method with a given rate of convergence for smooth functions F: simply choose the forcing sequence appropriately and solve the Newton equations (e.g., using an iterative method) sufficiently accurately that (1.3) holds. For example, one could choose $\eta_k = \min\{c\|F(x_k)\|^p, \frac{1}{2}\}$ $(0 < p \leqq 1)$ so that the closer one is to the solution, the better one must solve the Newton equations. Such a choice allows the method to adapt to the problem being solved. An application of this idea to unconstrained optimization is given by Dembo and Steihaug [2].

Given Corollary 3.5(c), one might conjecture that $\{x_k\}$ converges to x^* with order at least $1+p$ if $\{\eta_k\}$ converges to 0 with order at least $1+p$. This is false! In fact the $\{x_k\}$ need not converge any faster than superlinearly even if $\{\|r_k\|/\|F(x_k)\|\}$ converges to 0 with order at least $1+p$ (see the example given in the appendix). The problem is that a step which makes the error very small need not result in a correspondingly small relative residual.

4. Applications. The class of inexact Newton methods is sufficiently general as to encompass most existing methods for solving (1.1). Moreover, the results presented extend immediately to an arbitrary Banach space. But the importance of the inexact Newton approach lies in how easily these results can be applied to analyze specific methods. In this section, we discuss several examples.

Consider using Newton's method on a problem in which F and F' are not known exactly but can be evaluated to any specified accuracy (e.g., F and F' might be obtained by integrating a differential equation numerically). Then Corollary 3.5 specifies how accurately F and F' must be evaluated in order to ensure a given rate of convergence.

Dennis [3] proves a Newton–Kantorovich theorem for Newton-like methods of the form

$$x_{k+1} = x_k - M(x_k)F(x_k), \qquad x_0 \text{ given},$$

where $M(x_k)$ satisfies

$$\|I - F'(x_k)M(x_k)\| \leqq \eta_{\max} < 1.$$

But with this assumption,

$$\|r_k\| = \|F'(x_k)[-M(x_k)F(x_k)] + F(x_k)\| \leqq \|I - F'(x_k)M(x_k)\|\,\|F(x_k)\| \leqq \eta_{\max}\|F(x_k)\|,$$

so that x_{k+1} is an inexact Newton iterate, and the same result can be proved under this weaker assumption (cf. Altman [1]).

Sherman [11] and Dennis and Walker [6] prove that any Newton-iterative method can be made to converge with weak order at least 2 (and superlinearly) if $O(2^k)$ inner iterations (of the linear iterative method) are applied to the Newton equations at the kth outer iteration. The same results follows directly from Corollary 3.5 when one observes that, in the appropriate norm, each inner iteration reduces the norm of the residual by a constant factor less than one, so that the relative residuals corresponding to outer iterations converge to 0 with weak order at least 2.

Dennis and Walker [6] prove that a modified Jacobi–secant method is locally and linearly convergent. The same result follows from Theorem 2.3 when one observes that, in the appropriate norm, (1.3) is satisfied at every iteration for some $\eta_k \equiv \eta_{\max} < 1$.

Appendix: A counterexample. Consider the function of one variable $F(x) = x - x^2$ with root $x^* = 0$. Let $x_0 = \frac{1}{3}$ and define the sequences $\{x_k\}$ and $\{\eta_k\}$ as follows:

$$x_{2i+1} = x_{2i}^{i+1}, \quad x_{2i+2} = x_{2i}^{i+3}, \quad \eta_{2i} = 5x_{2i}, \quad \eta_{2i+1} = 5x_{2i}^2 \quad \text{for } i \geqq 0.$$

Then $x_k \to x^*$ and $s_k \equiv x_{k+1} - x_k$ satisfies

$$\frac{1}{5}\eta_k \leqq \frac{\|r_k\|}{\|F(x_k)\|} \leqq \eta_k, \quad \text{where } r_k = F'(x_k)s_k + F(x_k),$$

so that $\{x_k\}$ is a sequence of inexact Newton iterates corresponding to the forcing sequence $\{\eta_k\}$. Moreover,

$$\frac{\eta_{k+1}}{\eta_k^2} \leqq 1 \quad \text{for } k \geqq 0,$$

so that $\{\eta_k\}$ converges to zero quadratically (as does $\{\|r_k\|/\|F(x_k)\|\}$). However, for any $q > 1$,

$$\limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^q} \geqq \limsup_{i\to\infty} \frac{|x_{2i+2}|}{|x_{2i+1}|^q} = \infty,$$

so that $\{x_k\}$ is not converging to x^* with order q.

REFERENCES

[1] MIECZYSLAW ALTMAN, *Iterative methods of contractor directions*, Nonlinear Analysis: Theory, Methods and Applications, 4 (1980), pp. 761–772.

[2] RON S. DEMBO AND TROND STEIHAUG, *Truncated-Newton algorithms for large-scale unconstrained optimization*, Series #48, School of Organization and Management, Yale University, New Haven, CT, September 1980.

[3] J. E. DENNIS, JR., *On Newton-like methods*, Numer. Math., 11 (1968), pp. 324–330.

[4] J. E. DENNIS, JR. AND JORGE J. MORÉ, *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Comp., 28 (1974), pp. 549–560.

[5] ———, *Quasi-Newton methods, motivation and theory*, SIAM Rev., 19 (1977), pp. 46–89.

[6] J. E. DENNIS, JR. AND HOMER F. WALKER, *Local convergence theorems for quasi-Newton methods* Tech. Rep. 476-(141-171-163)-1, Dept. Mathematical Sciences, Rice University, Houston, TX, 1979.

[7] M. A. KRASNOSEL'SKII AND YA. B. RUTICKII, *Some approximate methods of solving nonlinear operator equations based on linearization*, Soviet Math. Dokl., 2 (1961), pp. 1542–1546.

[8] GARTH P. MCCORMICK AND KLAUS RITTER, *Methods of conjugate directions versus quasi-Newton methods*, Math. Prog., 3 (1972), pp. 101–116.

[9] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[10] VICTOR PEREYRA, *Accelerating the convergence of discretization algorithms*, this Journal, 4 (1967), pp. 508–533.

[11] ANDREW H. SHERMAN, *On Newton-iterative methods for the solution of systems of nonlinear equations*, this Journal, 15 (1978), pp. 755–771.

[12] TROND STEIHAUG, *Quasi-Newton methods for large scale nonlinear problems*, Ph.D. Thesis, School of Organization and Management, Yale University, New Haven, CT, 1981. Also available as Series #49.