Numerische Mathematik 11, 324–330 (1968)

On Newton-Like Methods

J. E. Dennis, Jr.*

Received May 2, 1967

1. Introduction

Let X and Y be Banach spaces and let F be a nonlinear mapping of some open subset Ω of X into Y. Let F have a continuous Fréchet second derivative on Ω_0, the closure of an open convex subset of Ω. When $\Gamma(x)=[F'(x)]^{-1}$ exists for every x in Ω_0, Newton's method for finding an approximation to a root of F is

$$x_{n+1}=x_n-\Gamma(x_n)F(x_n), \quad n=0, 1, \ldots,$$

where x_0 is prechosen. We will not assume the existence of $\Gamma(x)$. The purpose of this paper is to study Newton-like methods of the form

$$x_{n+1}=x_n-M(x_n)F(x_n), \quad n=0, 1, \ldots,$$

where x_0 is prechosen and M is some, not necessarily continuous, correspondence between Ω_0 and $L(Y, X)$. For a practical problem, such as the simultaneous solution of nonlinear equations, Newton's method is usually absurd because of the relative impossibility of doing the calculations necessary to find x_{n+1} exactly. Any number of expediencies suggest themselves, such as the use of a matrix of difference approximations in place of the Jacobian matrix. Some very ingenious algorithms for the solution of this problem are given by Broyden [6] and Brown and Conte [7]. All the methods suggested there are of the type listed above. These two papers abound with excellent numerical examples illustrating the imminent practicality of these methods. Many authors have studied methods of this type. Among them [8] and [11] study the method as stated here and the bibliography is composed of references in which some variant of this method is studied. Many more could be listed.

In [8], the author, by this approach, obtained a generalization of a theorem due to Mysovskih [12] on the convergence of Newton's method and a convergence theorem for Newton-like methods. This latter theorem yielded only linear convergence. Here we present sufficient conditions on $M(x)$ for the sequence of error bounds of the iteration to be of order p, $1 \leqq p \leqq 2$.

2. Results

It is convenient to assume that the reader is familiar with the basic facts concerning $L(X, Y)$, the bounded linear operators from X into Y, $B(X, Y)$, the bounded bilinear operators from X into Y, and the Taylor series expansion in terms of the Fréchet derivatives. (See [8, 12].) $N(x, r)$ will denote the open ball centered at x with radius r. cl $N(x, r)$ will be its closure.

* Department of Mathematics, University of Utah, Salt Lake City, Utah.

The following theorem strengthens a theorem in [9].

Theorem 1. Let the following conditions be satisfied:

(1) For every $x \in \Omega_0$, $M(x)$ exists in $L(Y, X)$ and $\|M(x)\| \leqq B$.

(2) For every $x \in \Omega_0$, $\|F''(x)\| \leqq K$.

(3) $0 < \|F(x_0)\| \leqq \eta$, for some $x_0 \in \Omega$.

(4) For every $x \in \Omega_0$, $\|I - F'(x) M(x)\| \leqq \delta < 1$.

(5) $h = \frac{B^2 K \eta}{1-\delta} < 2$.

(6) $N(x_0, r) \subset \Omega_0$, where $r = \frac{B\eta}{1-\alpha}$ and $\alpha = \delta + (1-\delta)\frac{h}{2}$.

Then F has a zero σ, $\|\sigma - x_0\| < r$, to which $x_{n+1} = x_n - M(x_n) F(x_n)$, $n = 0, 1, \ldots$ converges and the speed of convergence is given by

$$\|x_n - \sigma\| < \frac{B \eta \alpha^n}{1-\alpha}.$$

Proof. Set

$$\eta_0 = \eta, \quad h_0 = h, \quad \alpha_0 = \alpha, \quad \eta_{n+1} = \eta_n \alpha_n, \quad h_n = \frac{B^2 K \eta_n}{1-\delta} = h_{n-1}\alpha_{n-1},$$

and

$$\alpha_n = \delta + \frac{(1-\delta) h_n}{2}.$$

Notice that α_n is a convex combination of 1 and $h_n/2$ and is therefore strictly between zero and one unless $x_n = \sigma$. We will then take $1 > \alpha_n > 0$.

Now $r > B\eta_0 \geqq \|M(x_0)\| \, \|F(x_0)\| \geqq \|x_1 - x_0\|$ and so $x_1 \in N(x_0, r)$. Furthermore,

$$\begin{aligned}
\|F(x_1)\| &= \left\| F(x_0) + F'(x_0)(x_1 - x_0) + \int_{x_0}^{x_1} F''(x)(x_1 - x, \cdot)\, dx \right\| \\
&\leqq \|F(x_0) + F'(x_0)(-M(x_0) F(x_0))\| + \frac{B^2 K \eta_0^2}{2} \\
&\leqq \|[I - F'(x_0) M(x_0)](F(x_0))\| + \frac{B^2 K \eta_0^2}{2} \\
&\leqq \delta\eta_0 + \frac{B^2 K \eta_0^2}{2} = \left(\delta + \frac{(1-\delta) h_0}{2}\right)\eta_0 = \alpha_0 \eta_0 = \eta_1.
\end{aligned}$$

Assume, by way of induction, that $\|x_k - x_0\| < r$ and $\eta_j \geqq \|F(x_j)\|$ for every k and j such that $1 \leqq k \leqq n$ and $1 \leqq j < n$.

Exactly as for $n = 1$, one obtains that

$$\|F(x_n)\| \leqq \delta\eta_{n-1} + \frac{B^2 K \eta_{n-1}^2}{2} = \alpha_{n-1}\eta_{n-1} = \eta_n$$

and so

$$\|x_{n+1} - x_n\| \leqq B\eta_n.$$

Now

$$\alpha_k = \delta + \frac{(1-\delta) h_k}{2} < \delta + \frac{(1-\delta) h_{k-1}}{2} = \alpha_{k-1}.$$

Hence

$$\alpha_k < \alpha_0 \quad \text{and so} \quad \eta_{j+1} = \eta_j \alpha_j = \eta_{j-1}\alpha_{j-1}\alpha_j < \alpha^{j+1}\eta_0.$$

We complete the induction by noting that $x_{n+1}\in N(x_0, r)$ since,

$$\|x_{n+1}-x_0\| \leqq \sum_0^n B\eta_j < B\eta_0 \sum_0^n \alpha^j = B\eta_0 \frac{1-\alpha_0^{n+1}}{1-\alpha_0} < r.$$

$\langle x_n\rangle$ is a Cauchy sequence since,

$$\begin{aligned}\|x_{k+m}-x_k\| &\leqq \sum_{i=k}^{k+m-1} B\eta_i < B\eta_0 \sum_{i=k}^{k+m-1} \alpha^i = B\eta_0\left(\sum_{i=0}^{k+m-1}\alpha^i - \sum_{i=0}^{k-1}\alpha^i\right)\\ &= B\eta_0\left[\frac{1-\alpha_0^{k+m}}{1-\alpha_0} - \frac{1-\alpha_0^{k+1}}{1-\alpha_0}\right]\\ &= B\eta_0 \frac{\alpha_0^k\,(\alpha_0^m-\alpha_0)}{1-\alpha_0} \xrightarrow[k\to\infty]{} 0, \quad \text{independant of } m.\end{aligned}$$

X is a Banach space, hence complete in the norm topology and so $\langle x_n\rangle$ has a limit point σ.

$$\|x_n-\sigma\| \leqq B\sum_n^\infty \eta_i < B\eta_0 \sum_n^\infty \alpha_0^i = B\eta_0 \frac{\alpha_0^n}{1-\alpha_0} \quad \text{for any} \quad n \geqq 0.$$

When $n=0$, this reduces to $\|x_0-\sigma\|<r$.

We finish the proof by noting that

$$\|F(x_n)\| \leqq \eta_n < \eta_0\alpha^n \xrightarrow[n\to\infty]{} 0, \quad \text{and} \quad \|F(x_n)\| \xrightarrow[n\to\infty]{} \|F(\sigma)\|.$$

Condition (4) is excessively restrictive. The following is an easy corollary to the proof of Theorem 1.

Corollary 1.1. Let (1), (2), (3), (5) and (6) of Theorem 1 hold but in place of (4) require only that

(4.1) for every n such that $x_{n+1}=x_n-M(x_n)F(x_n)$ is defined,

$$\|I-F'(x_n)\,M(x_n)\| \leqq \delta < 1.$$

Under these assumptions, the conclusions of the previous theorem hold.

Remark. Corollary 2 might be useful in an *a posteriori* error estimate since the δ in (4.1) might then be much smaller and easier to estimate than the δ in (4).

One of the defects of this theorem is that, although for $M(x)=[F'(x)]^{-1}$ this is a convergence theorem for NEWTON's method, the error bounds only enjoy linear convergence. This is to be expected, however, since under some conditions, one can satisfy the hypotheses on $M(x)$ with a linear operator Γ which is independent of x and so the iteration method itself may be linearly convergent. See [9] for theorems of this type. Theorem 2 will remedy this defect, but first we paraphrase a definition from [15].

Definition 1. Let $\langle x_n\rangle$ be a sequence in X which converges to an element x^* of X. If there is a positive real number p and a nonzero constant C such that

$$\frac{\|x_{n+1}-x^*\|}{\|x_n-x^*\|^p} \xrightarrow[n\to\infty]{} C,$$

then p is called the order of the sequence and C is called the asymptotic error constant.

When we speak of the order of an iteration method, we will always mean the order of the convergent iteration sequence under consideration which was generated by that method. We will usually speak of the order of the sequence of error bounds. Then, of course, X will be the real numbers and x^* will be zero.

Theorem 2. Let (1), (2) and (3) of Theorem 1 hold and in addition let the following conditions be satisfied:

(4.2) $$\delta(x_0) < 1, \quad \text{where} \quad \delta(x) = \|I - F'(x)\, M(x)\|.$$

Let $p \in (1, 2]$ and assume, as long as $x_{n+1} = x_n - M(x_n)\, F(x_n)$ is defined, that $\delta_n^p \geqq \delta_{n+1}$. $(\delta_n = \delta(x_n))$

(5.2) $$h = \frac{B^2 K \eta}{1 - \delta_0} < 2.$$

(6.2) $$r_2 = B\eta\left(1 + \sum_1^\infty \alpha^{2p^{k-1}-1}\right), \quad \alpha = \delta_0 + (1 - \delta_0)\, h/2, \quad \text{and} \quad N(x_0, r_2) \subset \Omega_0.$$

Under these hypotheses, the sequence $\langle x_n \rangle$ is defined and each $\|x_n - x_0\| < r_2$. Furthermore $\langle x_n \rangle$ converges to $\sigma \in N(x_0, r_2)$, $F(\sigma) = 0$ and

$$\|x_n - \sigma\| < B\eta \frac{\alpha^{2p^{n-1}-1}}{1 - \alpha^{(p-1)2p^{n-1}}}.$$

Proof. Since this proof is similiar to the proof of Theorem 1, we will leave out some details. Set

$$\eta_0 = \eta, \quad h_0 = h, \quad \alpha_0 = \alpha, \quad \eta_n = \alpha_{n-1}\eta_{n-1}, \quad h_n = \frac{B^2 K \eta_n}{1 - \delta_n}, \alpha_n = \delta_n + (1 - \delta_n)\frac{h_n}{2}.$$

Exactly as in Theorem 1 one can show $\|F(x_n)\| \leqq \eta_{n-1}\alpha_{n-1} = \eta_n$, under the hypothesis that $x_k \in \Omega_0$, $k \leqq n$. Now

$$\begin{aligned}\alpha_n &= \delta_n + \frac{(1-\delta_n)\, h_n}{2} = \delta_n + \frac{B^2 K \eta_n}{2} = \delta_n + \frac{(1 - \delta_{n-1})\, h_{n-1}\alpha_{n-1}}{2} \\ &= \delta_{n-1}\delta_{n-1}^{p-1} + \frac{(1-\delta_{n-1})\, h_{n-1}\alpha_{n-1}}{2} < \delta_{n-1}\alpha_{n-1}^{p-1} + \frac{(1-\delta_{n-1})\, h_{n-1}\alpha_{n-1}^{p-1}}{2} \\ &= \alpha_{n-1}\alpha_{n-1}^{p-1} = \alpha_{n-1}^p.\end{aligned}$$

Hence $\alpha_n < \alpha_0^{p^n}$. We will show by induction that

$$\prod_{i=0}^{n-1} \alpha_0^{p^i} \leqq \alpha_0^{2p^{n-1}-1} \quad \text{for} \quad n \geqq 1.$$

Equality for $n = 1$ is obvious. Suppose now that

$$\prod_{i=0}^{k} \alpha_0^{p^i} \leqq \alpha_0^{2p^k - 1}.$$

Then

$$\prod_{i=0}^{k+1} \alpha_0^{p^i} = \alpha_0^{p^{k+1}} \prod_{i=0}^{k} \alpha_0^{p^i} \leqq \alpha_0^{p^{k+1}}\alpha_0^{2p^k-1} = \alpha_0^{(p+2)p^k-1} \leqq \alpha_0^{2pp^k-1} = a_0^{2p^{k+1}-1}$$

Since $1 < p \leqq 2$.

We see now that

$$\eta_n = \eta_{n-1}\alpha_{n-1} = \eta_{n-2}\alpha_{n-2}\alpha_{n-1} = \eta_0 \prod_{i=0}^{n-1} \alpha_i < \eta_0 \prod_{i-0}^{n-1} \alpha_0^{p^i} \leqq \eta_0\alpha_0^{2p^{n-1}-1}, \; n \geqq 1.$$

Hence $\|x_1 - x_0\| \leqq B\eta_0 < r_2$ and

$$\|x_0 - x_{n+1}\| \leqq B \sum_0^n \eta_k < B\eta\left(1 + \sum_1^n \alpha_0^{2p^{k-1}-1}\right) < r_2.$$

$$\|x_{m+k} - x_m\| \leqq B \sum_m^{m+k-1} \eta_i < B\eta \sum_m^{m+k-1} \alpha_0^{2p^{i-1}-1} = B\eta \sum_0^{k-1} \alpha_0^{2p^{m+j-1}} \xrightarrow[m\to\infty]{} 0.$$

Thus $\langle x_n\rangle$ is a Cauchy sequence and so we can choose some element σ to which $\langle x_n\rangle$ converges.

Now $\|x_0 - \sigma\| \leqq B\sum_0^\infty \eta_i < B\eta\left(1 + \sum_1^\infty \alpha^{2p^{i-1}-1}\right) = r_2$. In general, for $n \geqq 1$,

$$\|x_n - \sigma\| < B\eta \sum_n^\infty \alpha^{2p^{i-1}-1} < B\eta\,\alpha^{2p^{n-1}-1} \sum_0^\infty \alpha^{2p^{n-1}(p-1)j} = \frac{B\eta\,\alpha^{2p^{n-1}-1}}{1-\alpha^{2p^{n-1}(p-1)}}.$$

The last inequality is termwise since

$$\alpha^{2p^{n-1}-1+2p^{n-1}(p-1)\cdot j} > \alpha_0^{2p^{n-1}-1+2p^{n-1}(p-1)\sum_0^{j-1} p^i} > \alpha_0^{2p^{n-1}-1+2p^{n-1}(p^j-1)} = \alpha_0^{2p^{n+j-1}-1}.$$

$$F(\sigma) = 0 \quad \text{since} \quad \|F(\sigma)\| = \lim_n \|F(x_n)\| \leqq B \lim_n \eta_n \leqq B\eta \lim_n \alpha^{2p^{n-1}-1} = 0.$$

The proof of this theorem yields the following easy corollary.

Remark. We could have set $r_2 = \sum_0^\infty \alpha^{p^k-1} B\eta$, which is larger.

Corollary 2.1. Let all the hypotheses of Theorem 2 hold except (4.2). In the notation of the proof of Theorem 2 let (4.3) $\delta_{n+1} \leqq \delta_n \alpha_n^{p-1}$, for those values of n for which $x_{n+1} = x_n - M(x_n)F(x_n)$ is defined.

Under these assumptions, the conclusions of Theorem 2 hold.

Remark. As we remarked after Corollary 1.1, Corollary 2.1 may be useful for *a posteriori* error estimates, since we may then be able to satisfy (4.3) with a larger p than (4.2).

It is perhaps worth noting that if $M(x_n) = M$, independent of n, and satisfies the hypothesis 4.2, then $F'(\sigma)M = I$.

Corollary 2.2. Under the hypotheses of Theorem 2 or Corollary 2.1 the sequence of error bounds of the iteration method $x_{n+1} = x_n - M(x_n)F(x_n)$ is of order p.

Proof. Set $b_n = B\eta \dfrac{\alpha^{2p^{n-1}-1}}{1-\alpha^{(p-1)2p^{n-1}}}$.

$$\frac{b_{n+1}}{b_n^p} = \frac{[1-\alpha^{(p-1)2p^{n-1}}]^p}{1-\alpha^{(p-1)2p^n}} \cdot \frac{\alpha^{2p^n-1}}{\alpha^{2p^n-p}} \cdot (B\eta)^{1-p} \xrightarrow[n\to\infty]{} \alpha^{p-1}(B\eta)^{1-p}.$$

Example. Because $\langle b_n\rangle$ is order p, it does not follow that the true error sequence $\langle e_n\rangle$ is of order p. If K is a positive number and $Kne_n = b_n$, then

$$\frac{e_{n+1}}{e_n^p} = \frac{b_{n+1}}{K(n+1)} \cdot \frac{K^p n^p}{b_n^p}, \text{ which diverges for } p > 1.$$

If b_n/e_n converges to some finite limit, say L, then $L \geqq 1$ and

$$\frac{e_{n+1}}{e_n^p} = \frac{e_{n+1}}{b_{n+1}} \cdot \frac{b_{n+1}}{b_n^p} \cdot \frac{b_n^p}{e_n^p} \xrightarrow[n\to\infty]{} L^{p-1}$$

times the asymptotic error bound on $\langle b_n\rangle$.

It might be said that Corollary 2.2 ensures that the iteration method of Theorem 2 is of *quasi-order* p. The two types of order are, by the remark above, equivalent when $\langle b_n/e_n\rangle$ has a finite limit.

When we let $M(x)=\Gamma_R(x)$, a right inverse for $F'(x)$, we obtain a strengthened version of a theorem due to MYSOVSKIH on the convergence of NEWTON's method. We call the method $x_{n+1}=x_n-\Gamma_R(x_n)F(x_n)$ NEWTON's method even though the name should perhaps be reserved for the case when $M(x_n)F'(x_n)=I$, since in that case the method is easily shown to be of order 2.

The method considered here has a computational advantage, however, since when $x_{n+1}=x_n-\Gamma_R(x_n)F(x_n)$ is defined, properties of the right inverse make it equivalent to the iteration $F'(x_n)(x_{n+1}-x_n)=F(x_n)$.

Corollary 2.3. Let (1), (2), (3) hold and assume that $\Gamma_R(x)$ exists on Ω_0, with $\|\Gamma_R(x)\|\leqq B_1$ uniformly on Ω_0. Assume also:

(5.3) $$h=B_1^2K\eta<2,$$

(6.3) $$r_3=B_1\eta\sum_0^\infty\left[\frac{h}{2}\right]^{2^k-1} \text{ and } N(x_0,r_3)\subset\Omega_0.$$

Under these hypotheses, NEWTON's method, with initial point x_0, defines a sequence $\langle x_n\rangle$ which converges to σ, a root of F. $\|\sigma-x_0\|\leqq r_3$ and

$$\|\sigma-x_n\|<B_1\eta\frac{\left[\frac{h}{2}\right]^{2^n-1}}{1-\left[\frac{h}{2}\right]^{2^n}}.$$

Proof. $\Gamma_R(x)$ clearly satisfies (4.2) with $\delta(x)\equiv 0$ and $p=2$ so the result follows from Theorem 2.

Condition (4) is sufficient to ensure the existence of $\Gamma_R(x)$, but we must augment conditions (5) and (6) to ensure that conditions (5.3) and (6.3) are met.

Theorem 3. Let conditions (1), (2), (3), (4) hold and in addition assume the following:

(5.4) $$h=\frac{B^2K\eta}{(1-\delta)^2}<2.$$

(6.4) $$r_4=\frac{B\eta}{1-\delta}\sum_0^\infty\left[\frac{h}{2}\right]^{2^k-1} \text{ and } N(x_0,r_4)\subset\Omega_0.$$

Under these hypotheses $\Gamma_R(x)$ exists and NEWTON's method, with initial point x_0, defines a sequence of elements of $N(x_0,r_4)$ that converges to σ a root of F. $\|\sigma-x_0\|\leqq r_4$ and

$$\|\sigma-x_n\|<\frac{B_1\eta}{1-\delta}\frac{\left[\frac{h}{2}\right]^{2^n-1}}{1-\left[\frac{h}{2}\right]^{2^n}}.$$

Proof. $\|I-F'(x)M(x)\|\leqq\delta<1$ and so, by a well-known theorem [12, p. 171] $[F'(x)M(x)]^{-1}$ exists with norm not more than $1/(1-\delta)$. Hence

$$\Gamma_R(x)=M(x)[F'(x)M(x)]^{-1}$$

exists with norm not more than $B_1=B/(1-\delta)$. The result now follows from Corollary 2.3.

3. Concluding Remarks

In general, it seems difficult to say much regarding the uniqueness of the root σ that our theorems ensure. If $M(\sigma)$ has the property that it is almost a left inverse for $F'(\sigma)$, i.e., $\|I - M(\sigma)F'(\sigma)\| < 1$, then, we can show that σ is unique in the largest set that is star-like with respect to σ and on which F' is continuous and $\|I - M(\sigma)F'(x)\| < 1$. Here, we mean that a set S is star-like with respect to σ, if for every $p \in S$ the segment joining p and σ is contained in S. Brown and Conte [7] have a numerical example in which Newton's method and a method of the type proposed here converge to different roots from the same initial estimate.

The author wishes to thank Professor Kenneth M. Brown of the Cornell University Department of Computer Science for pointing out an error in the original statement of Corollary 2.2, as well as other helpful comments.

References

1. Altman, M.: On the approximate solution of non-linear functional equations. Bull. Acad. Polon. Sci. Cl. III **5**, 457—460 (1957).
2. Bartle, R. G.: Newton's method in Banach spaces. Proc. Amer. Math. Soc. **6**, 827—831 (1955).
3. Ben-Israel, A.: A modified Newton-Raphson method for the solution of systems of equations. Israel J. Math., **3** 94—97 (1965).
4. — On Newton's method for the solution of systems of equations. J. Math. Anal. Appl. **15**, 243—252 (1966).
5. — On Newton's method. Not. Amer. Math. Soc. **14**, 677—260 (1967).
6. Broyden, C. G.: A class of methods for solving nonlinear simultaneous equations. Math. Comp. **19**, 577—593 (1965).
7. Brown, K. M., and S. D. Conte: The solution of simultaneous nonlinear equations. Proceedings 22nd Nat'l Conference A.C.M., pp. 111—114.
8. Collatz, L.: Functional analysis and numerical mathematics. New York: Academic Press 1966.
9. Dennis, J. E. Jr.: On Newton's method and nonlinear simultaneous displacements. J. SIAM Numer. Anal., Ser. B. **4**, 103—108 (1967).
10. — A stationary Newton method for nonlinear functional equations. J. SIAM Numer. Anal. Ser. B. **4**, 222—232 (1967).
11. Ghinea, M.: Sur la reolution des equations operationnelles dans les espaces de Banach. R.F.T.I. Chiffres **8**, 3—22 (1965).
12. Kantorovich, L. V., and G. P. Akilov: Functional Analysis in Normed spaces. New York: Pergammon Press 1964.
14. Moore, R. H.: Newton's method and variations, Nonlinear Integral Equations. P. Anselone, ed., pp. 65—98. Madison: U. of Wisc. Press 1964.
15. Traub, J. F.: Iterative Methods for the Solution of Equations. Englewood Cliffs, N. J.: Prentice-Hall 1964.

Dr. J. E. Dennis, Jr., Assistent Professor
Dept. of Mathematics
University of Utah
206 Mathematics Building
Salt Lake City, Utah 84112, USA