Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE @ DIRECT•

Journal of Computational and Applied Mathematics 157 (2003) 169–185

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

An improved error analysis for Newton-like methods under generalized conditions

Ioannis K. Argyros*

Department of Mathematical Sciences, Cameron University, Lawton, OK 73505-6377, USA

Received 19 August 2002; received in revised form 25 December 2002

Abstract

We introduce new semilocal convergence theorems for Newton-like methods in a Banach space setting. Using new and very general conditions we provide different sufficient convergence conditions than before. This way we introduce more precise majorizing sequences, which in turn lead to finer error estimates and a better information on the location of the solution. Moreover for special choices of majorizing functions our results reduce to earlier ones. In the local case we obtain a larger convergence radius (ball). Finally, as an application, we show that in the case of Newton's method the famous Newton–Kantorovich hypothesis can be weakened under the same information.

MSC: 65B05; 65G99; 65J15; 47H17; 49M15; CR:1.5

Keywords: Newton-like method; Banach space; Majorant principle; Newton–Kantorovich hypothesis; Fréchet-derivative; Majorizing sequence; Radius of convergence

1. Introduction

In this study we are concerned with the problem of approximating a locally unique solution x^* of equation

$$F(x) = 0, \tag{1}$$

where F is a Fréchet-differentiable operator defined on a closed ball $\bar{U}(x_0, R) = \{x \in X \mid \|x - x_0\| \leqslant R\}$ ($R \geqslant 0$) which is a subset of a Banach space X with values in a Banach space Y.

* Tel.: +1-580-581-2908; fax:+1-580-581-2616.
E-mail address: ioannisa@cameron.edu (I.K. Argyros).

doi:10.1016/S0377-0427(03)00390-X

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time-invariant system is driven by the equation $\dot{x} = F(x)$, where x is the state. Then the equilibrium states are determined by solving Eq. (1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative—when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

Newton-like methods

$$x_{n+1} = x_n - A(x_n)^{-1}F(x_n) \quad (n \geqslant 0) \tag{2}$$

have been used extensively to generate a sequence approximating x^*. Here $A(x) \in L(X, Y)$ the space of bounded linear operators from X into Y and is an approximation to the Fréchet-derivative $F'(x)$ of operator F. Rheinboldt [12] established a convergence theorem for (2) which includes the Newton–Kantorovich theorem for the Newton method $(A(x) = F'(x))$ as a special case [8]. A further generalization was given by Dennis [6], Miel [9], Moret [10], Yamamoto [14], Chen and Yamamoto [5], Argyros [1–4], Potra [11] and others (for a survey of such results see, e.g., [3,4]) have provided local and semilocal convergence results under various assumptions. In particular, Chen and Yamamoto [5] use conditions (3), (4) (see Theorem 1 that follows). Here motivated from this paper we provide different sufficient convergence conditions. In the special case of the Newton–Kantorovich method we show that the famous Newton–Kantorovich hypothesis (see (46)) can be weakened (see (47)).

2. Semilocal analysis for Newton-like methods

We will need for simplicity the following version of the semilocal convergence theorem for Newton-like methods due to Chen and Yamamoto [5, p. 40]: (The general case can be treated along the same lines.)

Theorem 1. *Let $F : U(x_0, R) \subseteq X \to Y$ $(R \geqslant 0)$ be a Fréchet-differentiable operator and $A(x) \in L(X, Y)$. Assume*

(a) $A(x_0)^{-1} \in L(Y, X)$ *and for any* $x, y \in \bar{U}(x_0, r) \subseteq \bar{U}(x_0, R)$

$$\|A(x_0)^{-1}(A(x) - A(x_0))\| \leqslant \bar{w}_0(\|x - x_0\|) + \bar{b}, \tag{3}$$

$$\begin{aligned}\|A(x_0)^{-1}[F'(x + t(y - x)) - A(x)]\| \leqslant{}& \bar{w}(\|x - x_0\| + t\|y - x\|) \\ & - \bar{w}_0(\|x - x_0\|) + \bar{c}, \quad t \in [0, 1],\end{aligned} \tag{4}$$

where $\bar{w}(r+t)-\bar{w}_0(r)$, $t \geqslant 0$ *is a monotonically increasing function with* $\bar{w}(0)=\bar{w}_0(0)=0$, $\bar{w}_0(r)$ *is differentiable*, $\bar{w}_0'(r)>0$, $r\in[0,R]$, *and constants* $\bar{b}$, $\bar{c}$ *satisfy*

$$\bar{b}\geqslant 0, \quad \bar{c}\geqslant 0, \quad \bar{b}+\bar{c}<1. \tag{5}$$

Set

$$\|A(x_0)^{-1}F(x_0)\|\leqslant \eta. \tag{6}$$

Define functions

$$\bar{w}_2(r)=\eta-r+\int_0^r \bar{w}(t)\,\mathrm{d}t, \tag{7}$$

$$\bar{w}_3(r)=\bar{w}_2(r)+(\bar{b}+\bar{c})r \tag{8}$$

and iteration

$$r_0=0, \quad r_{n+1}=r_n+\frac{\bar{w}_3(r_n)-w^*}{1-\bar{b}-\bar{w}_0(r_n)} \quad (n\geqslant 0), \quad r^*=\lim_{n\to\infty} r_n, \tag{9}$$

where w^* *is the minimal value of* $\bar{w}_3$ *on* $[0,R]$, r^* *is the minimal point and* r_0^* *the unique zero of* $\bar{w}_3$ *in* $(0,r^*]$.

(b)

$$\bar{w}_3(R)\leqslant 0. \tag{10}$$

Then, *Newton-like method* $\{x_n\}$ $(n\geqslant 0)$ *generated by* (2) *is well defined*, *remains in* $\bar{U}(x_0,r^*)$ *for all* $n\geqslant 0$ *and converges to a solution* $x^*\in\bar{U}(x_0,r_0^*)$ *which is unique in*

$$\tilde{U}=\begin{cases}\bar{U}(x_0,R) & \text{if } \bar{w}_3(R)<0 \text{ or } \bar{w}_3(R)=0 \text{ and } r_0^*=R,\\ U(x_0,R) & \text{if } \bar{w}_3(R)=0 \text{ and } r_0^*<R.\end{cases} \tag{11}$$

Moreover, *the following error bounds hold for all* $n\geqslant 0$:

$$\|x_{n+1}-x_n\|\leqslant r_{n+1}-r_n \tag{12}$$

and

$$\|x^*-x_n\|\leqslant r^*-r_n. \tag{13}$$

From now on we assume the more general conditions:

$$\|A(x_0)^{-1}[A(x)-A(x_0)]\|\leqslant w_0(\|x-x_0\|)+b \tag{3'}$$

and

$$\|A(x_0)^{-1}[F'(x+t(y-x))-A(x)]\|\leqslant w(\|x-x_0\|+t\|y-x\|)-w_1(\|x-x_0\|)+c \tag{4'}$$

for $t\in[0,1]$ and $x,y\in\bar{U}(x_0,r)\subseteq\bar{U}(x_0,R)$, where, $w(r+t)-w_1(r)$, $t\geqslant 0$, $w_0(r)$ and $w_1(r)$, $r\in[0,R]$ are monotonically increasing functions with $w(0)=w_0(0)=w_1(0)=0$.

Note that we are using a bar above functions and parameters introduced in earlier results (see [5]) whereas no bar is used for our corresponding functions and parameters (see, e.g., (3′), (4′)).

Moreover, we do not want to use functions $\bar{w}_0$, $\bar{w}$ in (3′) and (4′) instead of w_0 and w, respectively, since this approach limits our choices of functions and parameters (see also Remark 2).

We show a result concerning the convergence of majorizing sequences:

Theorem 2. *Assume there exist $\delta \in [0,1]$, parameters η, b, c and functions w_0, w as in (3)–(6), such that*

$$2\int_0^1 w(\theta\eta)\,\mathrm{d}\theta - 2w_1(0) + 2c + \delta b + \delta w_0(\eta) \leqslant \delta, \tag{14}$$

$$w_0\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right)\right] + b < 1 \tag{15}$$

and

$$2\int_0^1 w\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right) + \theta\left(\frac{\delta}{2}\right)^{n+1}\eta\right]\mathrm{d}\theta - 2w_1\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right)\right]$$
$$+ \delta w_0\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right)\right] \leqslant 2\int_0^1 w(\theta\eta)\,\mathrm{d}\theta - 2w_1(0) + \delta w_0(\eta) \tag{16}$$

for all $n \geqslant 0$.

Then, iteration $\{t_n\}$ $(n \geqslant 0)$ given by

$$t_0 = 0, \quad t_1 = \eta,$$

$$0 \leqslant t_{n+2} = t_{n+1} + \frac{\int_0^1 \{w[t_n + \theta(t_{n+1}-t_n)]\,\mathrm{d}\theta - w_1(t_n) + c\}(t_{n+1}-t_n)}{1 - b - w_0(t_{n+1})} \tag{17}$$

is monotonically increasing, bounded above by

$$t^{**} = \frac{2\eta}{2-\delta}, \tag{18}$$

and converges to some t^* such that

$$0 \leqslant t^* \leqslant t^{**}. \tag{19}$$

Moreover, the following error bounds hold for all $n \geqslant 0$:

$$0 \leqslant t_{n+2} - t_{n+1} \leqslant \frac{\delta}{2}(t_{n+1}-t_n) \leqslant \left(\frac{\delta}{2}\right)^{n+1}\eta. \tag{20}$$

Proof. We must show

$$2\int_0^1 w[t_k + \theta(t_{k+1}-t_k)]\,\mathrm{d}\theta - 2w_1(t_k) + 2c + \delta b + \delta w_0(t_{k+1}) \leqslant \delta, \tag{21}$$

$$0 \leqslant t_{k+1} - t_k \tag{22}$$

and

$$w_0(t_{k+1}) + b < 1 \tag{23}$$

for all $k \geqslant 0$.

Estimate (20) can then follow immediately from (21)–(23). Using induction on the integer k, we get for $k = 0$

$$2\int_0^1 w[t_0 + \theta(t_1 - t_0)]\,\mathrm{d}\theta - 2w_1(t_0) + 2c + \delta b + \delta w_0(t_1) \leqslant \delta$$

$$w_0(t_1) + b = w_0(\eta) + b < 1$$

by the initial conditions. But (14) gives

$$0 \leqslant t_2 - t_1 \leqslant \frac{\delta}{2}(t_1 - t_0).$$

Assume (21)–(23) hold for all $k \leqslant n+1$. Using (14)–(16) we obtain in turn

$$\begin{aligned}
&2\int_0^1 w[t_{k+1} + \theta(t_{k+2} - t_{k+1})]\,\mathrm{d}\theta - 2w_1(t_{k+1}) + 2c + \delta b + \delta w_0(t_{k+1})\\
&\quad\leqslant 2\int_0^1 w\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right) + \theta\left(\frac{\delta}{2}\right)^{k+1}\eta\right]\mathrm{d}\theta\\
&\qquad - 2w_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right)\right] + 2c + \delta b + \delta w_0\left[\frac{2\eta}{2-\delta}\left(1 - \left(\frac{\delta}{2}\right)^{k+1}\right)\right]\\
&\quad\leqslant 2\int_0^1 w(\theta\eta)\,\mathrm{d}\theta - 2w_1(0) + 2c + \delta b + \delta w_0(\eta) \leqslant \delta \quad \text{(by (16) and (20))}.
\end{aligned}$$

Moreover, we must show

$$t_k \leqslant t^{**}, \tag{24}$$

$$t_0 = 0 \leqslant t^{**}, \quad t_1 = \eta \leqslant t^{**} \quad \text{and} \quad t_2 \leqslant \eta + \frac{\delta}{2}\eta = \frac{2+\delta}{2}\eta \leqslant t^{**}.$$

Assume (24) holds for all $k \leqslant n+1$. It follows from (20):

$$\begin{aligned}
t_{k+2} &\leqslant t_{k+1} + \frac{\delta}{2}(t_{k+1} - t_k) \leqslant t_k + \frac{\delta}{2}(t_k - t_{k-1}) + \frac{\delta}{2}(t_{k+1} - t_k)\\
&\leqslant \cdots \leqslant \eta + \frac{\delta}{2}\eta + \left(\frac{\delta}{2}\right)^2\eta + \cdots + \left(\frac{\delta}{2}\right)^{k+1}\eta\\
&\leqslant \frac{1 - (\frac{\delta}{2})^{k+1}}{1 - \frac{\delta}{2}}\eta \leqslant \frac{2\eta}{2-\delta} = t^{**}.
\end{aligned}$$

Hence, sequence $\{t_n\}$ $(n \geqslant 0)$ is bounded above by t^{**}. Furthermore, sequence $\{t_n\}$ $(n \geqslant 0)$ is monotonically increasing by (17) and as such it converges to some t^* satisfying (19).

That completes the proof of Theorem 2. □

Remark 1. Conditions (3′) and (4′) reduce to (3) and (4), respectively, if we choose $\bar{w}(r) = w(r)$, $\bar{w}_0(r) = w_0(r) = w_1(r)$ for all $r \in [0, R]$, $\bar{b} = b$ and $\bar{c} = c$. Moreover our conditions (3′) and (4′) allow more flexibility in the choice of functions. Note also that conditions (14) and (15) are of the Newton–Kantorovich-type hypotheses (see also (46)) which are always present in the study of Newton-like methods.

We provide the main result on the semilocal convergence of Newton-like methods.

Theorem 3. *Assume hypotheses of Theorem 2 hold and*

$$\bar{U}(x_0, t^*) \subseteq \bar{U}(x_0, R). \tag{25}$$

Then Newton-like method $\{x_n\}$ $(n \geqslant 0)$ generated by (2) is well defined, remains in $\bar{U}(x_0, t^*)$ for all $n \geqslant 0$ and converges to a solution $x^* \in \bar{U}(x_0, t^*)$ of equation $F(x) = 0$.

Moreover, the following error bounds hold for all $n \geqslant 0$:

$$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n \tag{26}$$

and

$$\|x_n - x^*\| \leqslant t^* - t_n, \tag{27}$$

where iteration $\{t_n\}$ $(n \geqslant 0)$ is given by (17). The solution x^* is unique in $\bar{U}(x_0, t^*)$ if

$$\int_0^1 [w((1+t)t^*) - w_1(t^*)]\,\mathrm{d}t + w_0(t^*) + b + c < 1. \tag{28}$$

Furthermore, if there exists R_0 such that

$$R_0 \in (t^*, R] \tag{29}$$

and

$$\int_0^1 [w(t^* + t(t^* + R_0)) - w_1(t^*)]\,\mathrm{d}t + w_0(t^*) + b + c \leqslant 1, \tag{30}$$

then the solution x^* is unique in $U(x_0, R_0)$.

Proof. We must show estimate (26). For $n = 0$, (26) is obvious, since

$$\|x_1 - x_0\| = \|A(x_0)^{-1}F(x_0)\| \leqslant \eta = t_1 - t_0 \quad \text{(by (6))}.$$

Suppose (26) holds for $n = 0, 1, \ldots, k+1$; this implies in particular that

$$\begin{aligned} \|x_{k+1} - x_0\| &\leqslant \|x_{k+1} - x_k\| + \|x_k - x_{k-1}\| + \cdots + \|x_1 - x_0\| \\ &\leqslant (t_{k+1} - t_k) + (t_k - t_{k-1}) + \cdots + (t_1 - t_0) = t_{k+1} - t_0 = t_{k+1}. \end{aligned}$$

We show that (26) holds for $n=k+2$. Using (3) and (15) we get

$$\|A(x_0)^{-1}[A(x_{k+1})-A(x_0)]\| \leqslant w_0(\|x_{k+1}-x_0\|)+b \leqslant w_0(t_{k+1})+b<1. \tag{31}$$

It follows from (31) and the Banach Lemma on invertible operators [8] that $A(x_{k+1})^{-1}$ exists and

$$\|A(x_{k+1})^{-1}A(x_0)\| \leqslant \frac{1}{1-w_0(\|x_{k+1}-x_0\|)-b} \leqslant \frac{1}{1-b-w_0(t_{k+1})}. \tag{32}$$

By (4) we obtain

$$\begin{aligned}\|A(x_0)^{-1}F(x_{k+1})\| &\leqslant \|A(x_0)^{-1}\{F(x_{k+1})-A(x_k)(x_{k+1}-x_k)-F(x_k)\}\| \\ &\leqslant \int_0^1 \|A(x_0)^{-1}[F'(x_k+t(x_{k+1}-x_k))-A(x_k)]\|\,\|x_{k+1}-x_k\|\,\mathrm{d}t \\ &\leqslant \int_0^1 \{[w(\|x_k-x_0\|+t\|x_{k+1}-x_k\|) \\ &\qquad -w_1(\|x_k-x_0\|)]\,\mathrm{d}t+c\}\|x_{k+1}-x_k\|. \end{aligned} \tag{33}$$

Hence, by (2), (32) and (33) we get

$$\begin{aligned}\|x_{k+2}-x_{k+1}\| &\leqslant \|A(x_{k+1})^{-1}A(x_0)\|\cdot\|A(x_0)^{-1}F(x_{k+1})\| \\ &\leqslant \frac{\int_0^1\{[w(t_k+t(t_{k+1}-t_k))-w_0(t_k)]\,\mathrm{d}t+c\}(t_{k+1}-t_k)}{1-b-w_0(t_{k+1})} \\ &= t_{k+2}-t_{k+1}. \end{aligned} \tag{34}$$

Note also,

$$\|x_{k+2}-x_0\| \leqslant \|x_{k+2}-x_{k+1}\|+\|x_{k+1}-x_0\| \leqslant t_{k+2}-t_{k+1}+t_{k+1}-t_0=t_{k+2}\leqslant t^*.$$

Hence, we obtain $x_{k+2}\in\bar U(x_0,t^*)$. It follows from (34) that $\{x_n\}$ $(n\geqslant 0)$ is a Cauchy sequence in a Banach space X and as such it converges to some $x^*\in\bar U(x_0,t^*)$ (since $\bar U(x_0,t^*)$ is a closed set). By letting $k\to\infty$ we obtain $F(x^*)=0$. Moreover, estimate (27) follows from (26) using standard majorization techniques [4,8]. Furthermore, to show uniqueness, let y^* be a solution of equation $F(x)=0$ in $U(x_0,R_0)$. It follows from (2)

$$\begin{aligned}\|y^*-x_{k+1}\| &= \|y^*-x_k+A(x_k)^{-1}F(x_k)-A(x_k)^{-1}F(y^*)\| \\ &\leqslant \|A(x_k)^{-1}A(x_0)\|\left\|\left\{\int_0^1 \|A(x_0)^{-1}[F'(x_k+t(y^*-x_k))-A(x_k)]\right\|\,\|y^*-x_k\|\,\mathrm{d}t \\ &\leqslant \frac{\int_0^1\{[w(\|x_k-x_0\|+t(y^*-x_k))-w_1(\|x_k-x_0\|)]\,\mathrm{d}t+c\}}{1-b-w_0(\|x_k-x_0\|)}\,\|y^*-x_k\| \\ &< \frac{\int_0^1[w(t^*+t(t^*+R_0))-w_1(t^*)]\,\mathrm{d}t+c}{1-b-w_0(t^*)}\|y^*-x_k\|. \end{aligned} \tag{35}$$

Hence, we have

$$\|y^* - x_{k+1}\| < \|y^* - x_k\| \quad (k \geqslant 0). \tag{36}$$

That is by letting $k \to \infty$ in (36) we get

$$\lim_{k\to\infty} x_k = y^*.$$

But we already showed

$$\lim_{k\to\infty} x_k = x^*.$$

Hence, we deduce $x^* = y^*$. The first part of uniqueness uses (28) instead of (30).

That completes the proof of Theorem 3. □

Remark 2. Condition (16) holds in many interesting cases. Assume

$$F'(x) = A(x), \tag{37}$$

$$\|F'(x_0)^{-1}[F'(x) - F'(y)]\| \leqslant \ell\|x - y\|, \tag{38}$$

and

$$\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leqslant \ell_0\|x - x_0\| \tag{39}$$

for all $x, y \in \bar{U}(x_0, r) \subseteq \bar{U}(x_0, R)$. Then we can set

$$\bar{b} = \bar{c} = b = c = 0, \tag{40}$$

$$\bar{w}(r) = w_1(r) = \bar{w}_0(r) = w(r) = \ell r, \tag{41}$$

and

$$w_0(r) = \ell_0 r \tag{42}$$

for all $r \in [0, R]$, and some $\ell \geqslant 0$, $\ell_0 \geqslant 0$ with

$$\ell_0 \leqslant \ell. \tag{43}$$

That is we consider the Newton–Kantorovich method [8]. We must show

$$2\int_0^1 \ell\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{k+1}\right) + \theta\left(\frac{\delta}{2}\right)^{k+1}\eta\right] \mathrm{d}\theta - 2\ell_0\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{k+1}\right)\right]$$
$$+\delta\ell_0\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{k+1}\right)\right] \leqslant 2\ell\int_0^1 \theta\eta\,\mathrm{d}\theta + \delta\ell_0\eta,$$

or

$$\frac{3\delta-2}{2-\delta}\left[1-\left(\frac{\delta}{2}\right)^{k+1}\right] \leqslant 1,$$

which is true for all $k \geqslant 0$ by the choice of δ. Hence we showed (16). Note also that (14) becomes

$$w_0 \left[\frac{2\eta}{2-\delta} \left(1 - \left(\frac{\delta}{2}\right)^{k+1} \right) \right] = \frac{2\ell_0\eta}{2-\delta} \left(1 - \left(\frac{\delta}{2}\right)^{k+1} \right) \leqslant \frac{2\eta\ell_0}{2-\delta}.$$

Function $\bar{w}_3$ reduces to

$$\bar{w}_3(r) = \eta - r + \ell \int_0^r t\,\mathrm{d}t = \eta - r + \frac{\ell r^2}{2}. \tag{44}$$

Hence we obtain

$$w^* = 0, \quad r_0^* = R = r^* = \frac{1-\sqrt{1-h}}{\ell} \tag{45}$$

provided that

$$h = 2\ell\eta \leqslant 1 \quad (\ell \neq 0), \tag{46}$$

which is the famous Newton–Kantorovich hypothesis [8]. However our condition (14) gives for $\delta = 1$:

$$h_0 = (\ell_0 + \ell)\eta \leqslant 1. \tag{47}$$

Note that

$$h \leqslant 1 \Rightarrow h_0 \leqslant 1 \tag{48}$$

but not vice versa unless if $\ell_0 = \ell$. However (43) holds in general. Moreover ℓ/ℓ_0 can be arbitrarily large.

Example 1. Let $X = Y = \mathbf{R}$, $x_0 = 0$ and define function F on $\mathbf{R}$ by

$$F(x) = c_0 x + c_1 x + c_2 \sin e^{c_3 x}, \tag{49}$$

where c_i, $i = 0, 1, 2, 3$ are given parameters. It can easily be seen by (49) that for c_3 large and c_2 sufficiently small ℓ/ℓ_0 may be arbitrarily large. That is (47) may hold but not (46).

Example 2. Let $X = Y = \mathbf{R}$, $\bar{U}(x_0, R) = \bar{U}(\sqrt{2}, 1)$ and define function F on $\bar{U}$ by

$$F(x) = \frac{1}{6}x^3 - \left(\frac{2^{3/2}}{6} + 0.23 \right). \tag{50}$$

It can easily be seen by (50) that

$$\eta = 0.23, \;\; \ell = 2.4142136, \;\; \ell_0 = 1.914213562,$$

$$h = 1.1105383 > 1 \quad \text{and} \quad h_0 = 0.995538247 < 1.$$

That is there is no guarantee that Newton's method starting at x_0 converges to $x^* = 1.614507018$ since (46) is violated. However since (47) holds our results guarantees $\lim_{n\to\infty} x_n = x^*$.

Furthermore, notice that

$$t^* \in [\eta, 2\eta] \tag{51}$$

and under the hypotheses of Theorem 1,

$$t^* \in [\eta, r^*]. \tag{52}$$

In the next two results we show that our error bounds (26) are more precise than (12).

Theorem 4. *Under hypotheses of Theorems* 1, 3 *and the choices of Remark* 2 *the following error bounds hold*:

$$t_{n+1} < r_{n+1} \quad (n \geqslant 1), \tag{53}$$

$$t_{n+1} - t_n < r_{n+1} - r_n \quad (n \geqslant 1), \tag{54}$$

$$t^* - t_n \leqslant r^* - r_n \quad (n \geqslant 0), \tag{55}$$

$$t^* \leqslant r^*, \tag{56}$$

$$0 \leqslant t_{n+1} - t_n \leqslant \alpha^{2^{n-1}}(r_{n+1} - r_n) \quad (n \geqslant 0), \quad \alpha = \frac{1-\ell\eta}{1-\ell_0\eta} \in [0,1) \tag{57}$$

and

$$0 \leqslant t^* - t_n \leqslant \alpha^{2^{n-1}}(r^* - r_n) \quad (n \geqslant 1). \tag{58}$$

Moreover, *we have* $t_n = r_n$ $(n \geqslant 0)$ *if* $\ell = \ell_0$.

Proof. We use induction on the integer k to show (53) and (54) first. For $n = 0$ in (17) we obtain

$$t_2 - \eta = \frac{\ell\eta^2}{2(1-\ell_0\eta)} \leqslant \frac{\ell\eta^2}{2(1-\ell\eta)} = r_2 - r_1$$

and

$$t_2 \leqslant r_2.$$

Assume

$$t_{k+1} < r_{k+1}, \quad t_{k+1} - t_k < r_{k+1} - r_k \quad (k \leqslant n+1).$$

Using (17) and (9) we get

$$t_{k+2} - t_{k+1} = \frac{\frac{\ell}{2}(t_{k+1} - t_k)^2}{1 - \ell_0 t_{k+1}} < \frac{\frac{\ell}{2}(r_{k+1} - r_k)^2}{1 - \ell r_{k+1}} = r_{k+2} - r_{k+1}$$

and

$$t_{k+2} - t_{k+1} < r_{k+2} - r_{k+1}.$$

Let $m \geqslant 0$, we can obtain

$$\begin{aligned} t_{k+m} - t_k &< (t_{k+m} - t_{k+m-1}) + (t_{k+m-1} - t_{k+m-2}) + \cdots + (t_{k+1} - t_k) \\ &< (r_{k+m} - r_{k+m-1}) + (r_{k+m-1} - r_{k+m-2}) + \cdots + (r_{k+1} - r_k) \\ &< r_{k+m} - r_k. \end{aligned} \tag{59}$$

By letting $m \to \infty$ in (59) we obtain (55). For $n = 1$ in (55) we get (56).

Finally, (57) and (58) follow easily from (17) and (9). Note also that (57) holds as a strict inequality if $n \geqslant 2$.

That completes the proof of Theorem 4. □

Theorem 5. *Assume*

$$w_0(r) \leqslant \bar{w}_0(r) \leqslant w_1(r), \quad w(r) \leqslant \bar{w}(r), \; r \in [0, R] \tag{60}$$

and

$$\eta b - w^* > 0. \tag{61}$$

Then, under the hypotheses of Theorems 1 *and* 3 *the following error bounds hold for all* $n \geqslant 0$:

$$t_{n+1} < r_{n+1}, \tag{62}$$

$$t_{n+1} - t_n < r_{n+1} - r_n, \tag{63}$$

$$t^* - t_n \leqslant r^* - r_n \tag{64}$$

and

$$t^* \leqslant r^*. \tag{65}$$

Proof. We use induction on the integer k to show (50) and (51). For $n = 0$ in (9) and (17) we obtain $t_1 < r_1$ by (61). Moreover, we have

$$\begin{aligned} t_2 - t_1 &= \frac{\int_0^1 \{w[(t_0 - t_0) + \theta(t_1 - t_0)]\,\mathrm{d}\theta - w_1(t_0 - t_0) + c\}(t_1 - t_0)}{1 - b - w_0(t_1 - t_0)} \\ &< \frac{\int_0^1 \{w[(r_0 - r_0) + \theta(r_1 - r_0)]\,\mathrm{d}\theta - w_0(r_0 - r_0) + c\}}{1 - b - \bar{w}_0(s_1 - s_0)} \\ &= \frac{\bar{w}_3(r_1) - \bar{w}_3(r_0) + (1 - w_0(s_0) - b)(r_1 - r_0)}{1 - b - \bar{w}_0(r_1 - r_0)} \\ &= \frac{\bar{w}_3(r_1)}{1 - b - \bar{w}_0(r_1 - r_0)} = r_2 - r_1 \end{aligned}$$

and

$$t_2 < r_2.$$

Assume

$$t_{k+1} < r_{k+1} \tag{66}$$

and

$$t_{k+1} - t_k < r_{k+1} - r_k \tag{67}$$

for all $k < n$.

Using (9), (17), (66) and (67) we obtain in turn:

$$\begin{aligned} t_{k+2} - t_{k+1} &= \frac{\int_0^1 \{w[t_k + \theta(t_{k+1} - t_k)]\,\mathrm{d}\theta - w_1(t_k) + c\}}{1 - b - w_0(t_{k+1})}(t_{k+1} - t_k) \\ &< \frac{\int_0^1 \{w[r_k + \theta(r_{k+1} - r_k)]\,\mathrm{d}\theta - w_0(r_k) + c\}}{1 - b - \bar{w}_0(r_{k+1})}(r_{k+1} - r_k) \\ &= \frac{\bar{w}_3(r_{k+1}) - \bar{w}_3(r_k) + (1 - w_0(r_k) - b)(r_{k+1} - r_k)}{1 - b - \bar{w}_0(r_{k+1})} = r_{k+2} - r_{k+1} \end{aligned}$$

and

$$t_{k+2} < r_{k+2},$$

which shows (66), (67) for all $n \geqslant 0$.

Let $m \geqslant 0$, we can have

$$\begin{aligned} t_{k+m} - t_k &= (t_{k+m} - t_{k+m-1}) + (t_{k+m-1} - t_{k+m-2}) + \cdots + (t_{k-1} - t_k) \\ &< (r_{k+m} - r_{k+m-1}) + (r_{k+m-1} - r_{k+m-2}) + \cdots + (r_{k+1} - r_k) \\ &= r_{k+m} - r_k. \end{aligned} \tag{68}$$

By letting $m \to \infty$ in (68) we obtain (64). Finally, set $n = 1$ in (64) to obtain (65).

That completes the proof of Theorem 5. □

Note that Theorems 4 and 5 justify the claims made at the introduction.

Remark 3. Conditions (15) and (16) can be replaced by the stronger but easier to check:

$$w_0\left(\frac{2\eta}{2-\delta}\right) + b < 1 \tag{69}$$

and

$$\begin{aligned} &2\left[\int_0^1 w\left(\frac{2\eta}{2-\delta} + \frac{\theta\delta\eta}{2}\right)\mathrm{d}\theta - w_1\left(\frac{2\eta}{2-\delta}\right)\right] + \delta w_0\left(\frac{2\eta}{2-\delta}\right) \\ &\quad \leqslant 2\int_0^1 w(\theta\eta) - 2w_1(0) + \delta w_0(\eta) \end{aligned} \tag{70}$$

Remark 4. (a) The results obtained here can be extended to include non-differentiable operators. Indeed, consider the Newton-like method

$$x_{n+1} = x_n - A(x_n)^{-1}(F(x_n) + G(x_n)) \quad (n \geqslant 0), \tag{71}$$

for approximating solutions x^* of equation

$$F(x) + G(x) = 0, \tag{72}$$

where F is as before but the differentiability of $G : \bar{U}(x_0, R) \to Y$ is not assumed. Suppose

$$\|A(x_0)^{-1}(G(x) - G(y))\| \leqslant q(r)\|x - y\| \tag{73}$$

for all $x, y \in \bar{U}(x_0, r) \subseteq \bar{U}(x_0, R)$, $q(0) = 0$, and q a monotonically increasing function with $q(0) = 0$ on $[0, R]$. Simply add the terms: $\int_0^r q(t)\,\mathrm{d}t$ in (7); $2q(\eta)$ at the left-hand side of (14) and the right-hand side of (16); $2q[2\eta/(2-\delta)(1-(\delta/2)^{n+1})]$ at the left-hand side of (16); $\int_{t_n}^{t_{n+1}} q(t)\,\mathrm{d}t$ at the numerator in (17), and $q(t^*)$, $q(t^* + R_0)$ at the right-hand side of (28), (30) respectively to obtain the corresponding results in this case.

(b) Finally, our results can be extended to the more general case considered in [5], where the convergence of the Newton-like method

$$y_0 \in U(x^0, R), \;\; y_{n+1} = y_n - A(y_n)^{-1}(F(y_n) + G(y_n)) \quad (n \geqslant 0) \tag{74}$$

is studied under conditions (3), (4), (73) (with x^0 replacing x_0) to solve Eq. (52). However we leave the details to the motivated reader.

Remark 5. The conclusions of Theorem 1 hold if (4) is replaced by the more general condition:

$$\|A(x_0)^{-1}[F'(x + t(y-x)) - A(x)]\| \leqslant \bar{w}(\|x - x_0\| + t\|y - x\|) - \bar{w}_1(\|x - x_0\|) + \bar{c}_0, \tag{75}$$

where function $\bar{w}_1$ and $\bar{c}_0$ have the properties of $\bar{w}_0$ and $\bar{c}$ respectively, provided that

$$\bar{w}_0(r) \leqslant \bar{w}_1(r) \quad r \in [0, R], \tag{76}$$

and condition (3) by (3′). If $\bar{w}_1(r) = \bar{w}_0(r)$, $r \in [0, R]$ and $\bar{c}_0 = \bar{c}$, condition (75) reduces to (4). Moreover, if strict inequality holds in (76) we obtain more precise error bounds. Indeed, let us denote by $\{s_n\}$ the sequence using (75). That is $\{s_n\}$ is given by

$$s_0 = r_0 = 0, \quad s_1 = r_1, s_{n+1} - s_n$$
$$= \frac{u(s_n) - u(s_{n-1}) + (1 - \bar{w}_1'(s_{n-1}) - \bar{b})(s_n - s_{n-1})}{g(s_n)} \quad (n \geqslant 1). \tag{77}$$

It can easily be seen using induction on n (see also the proof of Theorems 4 and 5) that

$$\|x_{n+1} - x_n\| \leqslant s_{n+1} - s_n < r_{n+1} - r_n \quad (n \geqslant 1) \tag{78}$$

and

$$\|x_n - x^*\| \leqslant s^* - s_n \leqslant r^* - r_n \quad (n \geqslant 0), \quad s_n < r_n \quad (n \geqslant 2), \quad s^* \leqslant r^*, \quad s^* = \lim_{n\to\infty} s_n. \tag{79}$$

Furthermore, condition (75) allows us more flexibility in choosing the functions and constants. As an example let us consider the Newton–Kantorovich method and assume (37)–(39). Then we can choose: $w_0(r)=\ell_0 r$, $\bar{w}(r)=\bar{w}_1(r)=\bar{w}_0(r)=\ell r$, $r\in[0,R]$ and $\bar{b}=\bar{c}=\bar{c}_0=0$. Assuming that (46) and (47) hold we have

$$r_{n+1}=r_n+\frac{\ell(r_n-r_{n-1})^2}{2(1-\ell r_{n-1})} \quad \text{and} \quad s_{n+1}=s_n+\frac{\ell(s_n-s_{n-1})^2}{2(1-\ell_0 s_{n-1})} \quad (n\geqslant 1). \tag{80}$$

Condition (76) becomes $\ell_0\leqslant\ell$, and in case $\ell_0<\ell$ estimates (78) and (79) hold (see also the proof of Theorem 4).

3. Local convergence of Newton-like methods

In order to cover the local case, let us assume x^* is a simple zero of Eq. (72), $A(x^*)^{-1}$ exists and for any $x,y\in\bar{U}(x^*,r)\subseteq\bar{U}(x^*,R)$:

$$\|A(x^*)^{-1}[A(x)-A(x^*)]\|\leqslant v_0(\|x-x^*\|)+\beta, \tag{81}$$

$$\begin{aligned}\|A(x^*)^{-1}[F'(x+t(y-x))-A(x)]\|\leqslant{}& v(\|x-x^*\|+t\|y-x\|)\\ &-v_1(\|x-x^*\|)+\gamma, \quad t\in[0,1]\end{aligned} \tag{82}$$

and

$$\|A(x^*)^{-1}[G(x)-G(y)]\|\leqslant v_2(r)\|x-y\|, \tag{83}$$

where, v_0, β, v, v_1, γ, v_2 are as w_0, b, w, w_1, c, q, respectively. Exactly as in (35) but using (81)–(83) we can show the following local results for Newton-like methods:

Theorem 6. *Assume there exists a minimum solution $\alpha\in[0,R]$ of equation*

$$f(h)=0, \tag{84}$$

where

$$f(h)=\int_0^1[v((1+t)h)-v_1(h)]\,\mathrm{d}t+v_2(h)+v_0(h)+\beta+\gamma-1. \tag{85}$$

Then, Newton-like method $\{x_n\}$ $(n\geqslant 0)$ generated by (71) is well defined, remains in $\bar{U}(x^*,\alpha)$ and converges to a solution x^* of Eq. (72), provided that $x_0\in U(x^*,\alpha)$.

Moreover the following error bounds hold for all $n\geqslant 0$

$$\|x^*-x_{n+1}\|\leqslant p_{n+1}, \tag{86}$$

where

$$p_{n+1}=\frac{\int_0^1[v((1+t)\|x_n-x^*\|)-v_1(\|x_n-x^*\|)]\,\mathrm{d}t\|x_n-x^*\|+\gamma\|x_n-x^*\|+\int_0^{\|x_n-x^*\|}v_2(t)\,\mathrm{d}t}{1-\beta-v_0(\|x_n-x^*\|)}, \tag{87}$$

$(n\geqslant 0)$.

Remark 6. Note that Theorem 6 can be proved using the weaker conditions

$$\|A(x^*)^{-1}[F'(x+t(x^*-x))-A(x)]\| \leqslant \bar{v}(\|x-x^*\|(1+t))-\bar{v}_1(\|x-x^*\|)+\bar{\gamma}, \quad t\in[0,1] \tag{88}$$

and

$$\|A(x^*)^{-1}[G(x)-G(x^*)]\| \leqslant \bar{v}_2(r)\|x-x^*\| \tag{89}$$

for all $x\in\bar{U}(x^*,r)\subseteq\bar{U}(x^*,R)$, instead of (82) and (83), respectively, where $\bar{v}$, $\bar{v}_1$, $\bar{\gamma}$, $\bar{v}_2$ are as v, v_1, γ, and v_2.

Remark 7. As an application let us again consider Newton's method, i.e., $F'(x)=A(x)$, $G(x)=0$, and assume

$$\|F'(x^*)^{-1}[F'(x)-F'(x^*)]\| \leqslant q_0\|x-x^*\| \tag{90}$$

and

$$\|F'(x^*)^{-1}[F'(x)-F'(y)]\| \leqslant q\|x-y\| \tag{91}$$

for all $x,y\in\bar{U}(x_0,R)$.

Then we can set

$$\beta=\gamma=0, \quad v_2(r)=0, \quad v_0(r)=q_0r, \quad v(r)=v_1(r)=qr, \quad r\in[0,R]. \tag{92}$$

Using (85) we get

$$\alpha=\frac{2}{2q_0+q}. \tag{93}$$

Local results were not given in [5]. However, Rheinboldt [13] showed that under only (90) the convergence radius is given by

$$q_1=\frac{2}{3q}. \tag{94}$$

Since in general

$$q_0\leqslant q, \tag{95}$$

we conclude

$$q_1\leqslant\alpha. \tag{96}$$

The corresponding error bounds [13] are:

$$\|x_{n+1}-x^*\|\leqslant\delta_n, \tag{97}$$

$$\|x_{n+1}-x^*\|\leqslant\delta_n^1, \tag{98}$$

where

$$\delta_n=\frac{q\|x_n-x^*\|^2}{2[1-q_0\|x_n-x^*\|]} \tag{99}$$

and

$$\delta_n^1 = \frac{q\|x_n - x^*\|^2}{2[1 - q\|x_n - x^*\|]}. \tag{100}$$

That is,

$$\delta_n \leqslant \delta_n^1. \tag{101}$$

If strict inequality holds in (95) then (96) and (101) hold as strict inequalities also (see also Example 3 that follows).

Remark 8. As noted in [1–4,7,15] the local results obtained here can be used for projection methods such as Arnoldi's, the generalized minimum residual method (GMRES), the generalized conjugate residual method (GCR), for combined Newton/finite-difference projection methods and in connection with the mesh independence principle in order to develop the cheapest mesh refinement strategies.

Remark 9. The local results obtained here can also be used to solve equations of the form $F(x)=0$, where F' satisfies the autonomous differential equation [4,8]:

$$F'(x) = T(F(x)), \tag{102}$$

where $T: Y \to X$ is a known continuous operator. Since $F'(x^*) = T(F(x^*)) = T(0)$, we can apply the results obtained here without actually knowing the solution x^* of Eq. (1).

We complete this section with a numerical example.

Example 3. Let $X = Y = \mathbf{R}$, $D = U(0,1)$, $G = 0$, $A(x) = F'(x)$, and define function F on D by

$$F(x) = e^x - 1. \tag{103}$$

Then it can easily be seen that we can set $T(x) = x + 1$ in (102). Hence we set $q = e$. Moreover since $x^* = 0$ we obtain in turn

$$\begin{aligned} F'(x) - F'(x^*) = e^x - 1 &= x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots \\ &= \left(1 + \frac{x}{2!} + \cdots + \frac{x^{n-1}}{n!} + \cdots\right)(x - x^*) \end{aligned}$$

and for $x \in U(0,1)$, $\|F'(x) - F'(x^*)\| \leqslant (e-1)\|x - x^*\|$. That is, $q_0 = e - 1$. Using (90) and (91) we obtain, respectively,

$$q_1 = 0.245252961$$

and

$$\alpha = 0.254028662 > q_1.$$

That is, our convergence radius α is larger than the corresponding one q_1 due to Rheinboldt [13]. This observation is very important in computational mathematics (see also Remark 8).

References

[1] I.K. Argyros, On a new Newton–Mysovskii-type theorem with applications to inexact Newton-like methods and their discretizations, IMA J. Numer. Anal. 18 (1997) 37–56.
[2] I.K. Argyros, Local convergence of inexact Newton-like iterative methods and applications, Comput. Math. Appl. 39 (2000) 69–75.
[3] I.K. Argyros, Advances in the Efficiency of Computational Methods and Applications, World Scientific Publ. Co., River Edge, NJ, 2000.
[4] I.K. Argyros, F. Szidarovszky, The Theory and Applications of Iteration Methods, C.R.C. Press, Boca Raton, FL, 1993.
[5] X. Chen, T. Yamamoto, Convergence domains of certain iterative methods for solving nonlinear equations, Numer. Funct. Anal. Optim. 10 (1,2) (1989) 37–48.
[6] J.E. Dennis, Toward a unified convergence theory for Newton-like methods, in: L.B. Rall (Ed.), Nonlinear Functional Analysis and Applications, Academic Press, New York, 1971, pp. 425–472.
[7] P. Deuflhard, G. Heindl, Affine invariant convergence theorems for Newton's method and extensions to related methods, SIAM J. Numer. Anal. 16 (1979) 1–10.
[8] L.V. Kantorovich, G.P. Akilov, Functional Analysis, Pergamon Press, Oxford, 1982.
[9] G.J. Miel, Majorizing sequences and error bounds for iterative methods, Math. Comput. 34 (149) (1980) 185–202.
[10] I. Moret, A note on Newton-type iterative methods, Computing 33 (1984) 65–73.
[11] F.A. Potra, On the convergence of a class of Newton-like methods, in: R. Ansarge, W. Toening (Eds.), Iterative Solution of Nonlinear Systems of Equations, Lecture Notes in Mathematics, Vol. 953, Springer, Berlin, Heidelberg, New York, 1982, pp. 125–137.
[12] W.C. Rheinboldt, A unified convergence theory for a class of iterative processes, SIAM J. Numer. Anal. 5 (1968) 42–63.
[13] W.C. Rheinboldt, An adaptive continuation process for solving systems of nonlinear equations, Polish Acad. Sci. Banach Control Publ. 3 (1977) 129–142.
[14] T. Yamamoto, A convergence theorem for Newton-like methods in Banach spaces, Numer. Math. 51 (1987) 545–557.
[15] T.J. Ypma, Local convergence of inexact Newton methods, SIAM J. Numer. Anal. 21 (3) (1984) 583–590.