APPLICATIONES MATHEMATICAE
27,3 (2000), pp. 369–375

I. K. ARGYROS (Lawton, OK)

THE EFFECT OF ROUNDING ERRORS ON A CERTAIN CLASS OF ITERATIVE METHODS

Abstract. In this study we are concerned with the problem of approximating a solution of a nonlinear equation in Banach space using Newton-like methods. Due to rounding errors the sequence of iterates generated on a computer differs from the sequence produced in theory. Using Lipschitz-type hypotheses on the mth Fréchet derivative ($m \geq 2$ an integer) instead of the first one, we provide sufficient convergence conditions for the inexact Newton-like method that is actually generated on the computer. Moreover, we show that the ratio of convergence improves under our conditions. Furthermore, we provide a wider choice of initial guesses than before. Finally, a numerical example is provided to show that our results compare favorably with earlier ones.

1. Introduction. In this study we are concerned with approximating a solution of an equation

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

where F is an m times ($m \geq 2$ an integer) continuously differentiable nonlinear operator defined on an open convex subset D of a Banach space E_1 with values in a Banach space E_2.

The Newton method generates a sequence $\{x_n\}$ ($n \geq 0$) which in theory satisfies

$$x_{n+1} = \phi(x_n) \quad (n \geq 0), \tag{2}$$

where

$$\phi(x) = x - F'(x)^{-1}F(x) \quad (x \in D). \tag{3}$$

2000 *Mathematics Subject Classification*: 65B05, 47H17, 49D15.
Key words and phrases: Banach space, Newton-like method, Fréchet derivative, Lipschitz conditions, inexact Newton-like method.

Here, $F'(x)$ denotes the first Fréchet derivative of F evaluated at $x \in D$ (see [1], [3], [5]). Sufficient convergence conditions for Newton methods of the form (2) have been given by several authors. For a survey of such results we refer the reader to [3], [5] and the references there.

We first calculate $F'(x_n)$ and $F(x_n)$ $(n \ge 0)$. Then we need to find a solution $\theta(x_n)$ $(n \ge 0)$ of the equation

$$F'(x_n)(y) = -F(x_n) \quad (n \ge 0), \tag{4}$$

and set

$$\phi(x_n) = x_n + \theta(x_n) \quad (n \ge 0). \tag{5}$$

Due to the presence of rounding errors in numerical computations instead of the sequence $\{x_n\}$ $(n \ge 0)$ we really generate a sequence $\{\overline{x}_n\}$ such that

$$\overline{x}_{n+1} = \bar{\phi}(\overline{x}_n) \quad (n \ge 0), \tag{6}$$

$$\bar{\phi}(x) = [I + E_0(x)]\psi(x), \quad \psi(x) = x + \bar{\theta}(x) \quad (x \in D), \tag{7}$$

where $\bar{\theta}(x_n)$ is the exact solution of the equation

$$[\widehat{A}_n + E_1(\overline{x}_n)](y) = -[F(\overline{x}_n) + E_2(\overline{x}_n)] \quad (n \ge 0) \tag{8}$$

for some $E_0(x), E_1(x), E_2(x) \in L(E_1, E_2)$, the space of bounded linear operators from E_1 into E_2.

In the elegant paper [8] (see also [2], [4], [6], [7]) the convergence of the inexact sequence $\{\overline{x}_n\}$ $(n \ge 0)$ was analyzed, when $E_1 = E_2 = \mathbb{R}^i$ $(i \in \mathbb{N})$ under Lipschitz hypotheses on the first Fréchet derivative. Here we provide sufficient conditions for the local convergence of the inexact sequence $\{\overline{x}_n\}$ $(n \ge 0)$ in the more general setting of a Banach space but using Lipschitz hypotheses on the mth Fréchet derivative. Moreover, we show that the ratio of convergence improves under our conditions. Furthermore, we can provide a wider choice of initial guesses than before. Finally, a numerical example is provided to show that our results compare favorably with earlier ones.

2. Convergence analysis. We need a result whose proof can be found in [8, p. 111].

THEOREM 1. *If both $F'(x_n)$ and $\bar{A}_n$ $(n \ge 0)$ are nonsingular, then $\phi(\overline{x}_n)$ and $\bar{\phi}(\overline{x}_n)$ $(n \ge 0)$ exist and*

$$\|\bar{\phi}(\overline{x}_n) - x^*\| \le \eta_n\|x^*\| + (1 + \eta_n)\{\omega_n\|\overline{x}_n - x^*\| + (1 + \omega_n)\|\phi(\overline{x}_n) - x^*\|\}, \tag{9}$$

$$\eta_n = \|E_0(\overline{x}_n)\|, \quad \omega_n = \|\bar{A}_n^{-1}F'(\overline{x}_n) - I\| + \frac{\|\bar{A}_n^{-1}(\overline{F}_n - F_n)\|}{\|F'(\overline{x}_n)^{-1}F_n\|}. \tag{10}$$

In [2] we proved the following local convergence result for the exact Newton method.

THEOREM 2. *Let F be m times ($m \ge 2$ an integer) continuously Fréchet-differentiable on $U(x^*, \sigma) = \{x \in E_1 \mid \|x^* - x\| < \sigma\} \subseteq D$ for some $\sigma > 0$. Suppose $F'(x^*)$ is nonsingular, $F(x^*) = 0$,*

$$(11)\qquad \alpha_{m+1} = \sup\left\{ \frac{\|F'(x^*)^{-1}[F^{(m)}(x) - F^{(m)}(x^*)]\|}{\|x - x^*\|} \,\Big|\, x \in U(x^*, \sigma),\ x \ne x^* \right\},$$

and

$$(12)\qquad \alpha_i \ge \|F'(x^*)^{-1} F^{(i)}(x^*)\|, \qquad i = 2, \ldots, m.$$

If $x_0 \in U(x^*, \sigma)$ and

$$(13)\qquad \|x_0 - x^*\| < \delta^0,$$

where δ^0 is the positive zero of the equation

$$(14)\qquad \frac{\alpha_{m+1}}{m!} t^m + \ldots + \alpha_2 t - 1 = 0,$$

then

$$(15)\qquad \|x_0 - F'(x_0)^{-1} F(x_0) - x^*\| \le \frac{\frac{m\alpha_{m+1}}{(m+1)!}\|\overline{x}_0 - x^*\|^{m-1} + \frac{(m-1)\alpha_m}{m!}\|\overline{x}_0 - x^*\|^{m-2} + \ldots + \frac{\alpha_2}{2!}}{1 - \alpha_2\|\overline{x}_0 - x^*\| - \ldots - \frac{\alpha_{m+1}}{m!}\|\overline{x}_0 - x^*\|^m} \times \|\overline{x}_0 - x^*\|^2.$$

Moreover, if

$$(16)\qquad \|x_0 - x^*\| < \delta,$$

where δ is the positive zero of the equation

$$(17)\qquad \frac{(2m+1)\alpha_{m+1}}{(m+1)!} t^m + \frac{(2m-1)\alpha_m}{m!} t^{m-1} + \ldots + \frac{3\alpha_2}{2} t - 1 = 0,$$

then the exact Newton method converges quadratically to x^*.

This leads to the following interesting result for the inexact Newton method.

THEOREM 3. *If $\eta_0 = 0$, $\omega_0 < 1$, $\overline{x}_0 \in U(x^*, \sigma)$ with $\overline{x}_0 \ne x^*$, and*

$$(18)\qquad \|\overline{x}_0 - x^*\| < \min\{\delta, \delta_0\},$$

where δ_0 is the positive root of the function

$$(19)\qquad f_0(t) = \frac{\alpha_{m+1}}{(m+1)!}(1 - w_0 + 2m)t^m + \frac{\alpha_m}{m!}[2m - (1 - w_0)]t^{m-1} + \ldots + \frac{\alpha_2}{2!}(3 - w_0)t + w_0 - 1,$$

then

$$\|\bar{\phi}(\overline{x}_0) - x^*\| \le \Bigg\{\omega_0 + (1+\omega_0)\|\overline{x}_0 - x^*\| \times \frac{\frac{m\alpha_{m+1}}{(m+1)!}\|\overline{x}_0 - x^*\|^{m-1} + \frac{(m-1)\alpha_m}{m!}\|\overline{x}_0 - x^*\|^{m-2} + \ldots + \frac{\alpha_2}{2!}}{1 - \alpha_1\|\overline{x}_0 - x^*\| - \ldots - \frac{\alpha_{m+1}}{m!}\|\overline{x}_0 - x^*\|^m}\Bigg\}\|\overline{x}_0 - x^*\| < \|\overline{x}_0 - x^*\|. \tag{20}$$

P r o o f. By hypothesis (18) it follows that $\|\overline{x}_0 - x^*\| < \delta$. If $\phi(\overline{x}_0) = \overline{x}_0 - F'(\overline{x}_0)^{-1}F(\overline{x}_0)$, then inequality (15) gives

$$\|\phi(\overline{x}_0) - x^*\| < \frac{\frac{m\alpha_{m+1}}{(m+1)!}\|\overline{x}_0 - x^*\|^{m-1} + \frac{(m-1)\alpha_m}{m!}\|\overline{x}_0 - x^*\|^{m-2} + \ldots + \frac{\alpha_2}{2!}}{1 - \alpha_2\|\overline{x}_0 - x^*\| - \ldots - \frac{\alpha_{m+1}}{m!}\|\overline{x}_0 - x^*\|^m}\|\overline{x}_0 - x^*\|^2. \tag{21}$$

Hence, the first inequality in (20) follows from (9) by setting $n = 0$ and using (21). Moreover, the term in braces in (20) is less than 1 iff (18) holds.

That completes the proof of Theorem 3.

The following result provides sufficient conditions for the local convergence of the inexact Newton method.

THEOREM 4. *If* $\eta_n = 0$, $\omega_n \le \omega < 1$ *for all* $n \ge 0$ *and* $\overline{x}_0 \in U(x^*, \sigma)$ *satisfies*

$$\|\overline{x}_0 - x^*\| < \delta(\omega), \tag{22}$$

where $\delta(\omega)$ *is the positive root of the function* (19) *with* w_0 *being* w,

$$f(t) = \frac{\alpha_{m+1}}{(m+1)!}(1 - w + 2m)t^m + \frac{\alpha_m}{m!}[2m - (1+w)]t^{m-1} + \ldots + \frac{\alpha_2}{2!}(3-w)t + w - 1, \tag{23}$$

then the inexact Newton method (6)–(8) *generates a sequence* $\{\overline{x}_n\}$ $(n \ge 0)$ *which converges to* x^*.

P r o o f. The result follows from Theorem 3 by induction on $n \ge 0$.

REMARK 1. The conditions used in this study are different from the corresponding ones in [6]–[8] unless $\alpha = 0$, and $E_1 = E_2 = \mathbb{R}^i$ $(i \in \mathbb{N})$.

REMARK 2. Theorem 4 provides sufficient conditions for local convergence. However, as noted in [8, p. 113], $\eta_n \ne 0$ in general, which may lead to $\omega_n > 1$, so that convergence breaks down. Therefore, though the theory can predict monotonic decrease of the sequence $\{\|x_n - x^*\|\}$ $(n \ge 0)$, in practice the conditions of the theory fail to hold in some neighborhood of x^*, and within this neighborhood the behavior of $\{\overline{x}_n\}$ $(n \geq 0)$ is unpredictable. We examine the extent of this neighborhood by introducing the notation

$$\sigma_n = \omega_n + (1+\omega_n) \times \frac{\frac{m\alpha_{m+1}}{(m+1)!}\|\overline{x}_n - x^*\|^{m-1} + \frac{(m-1)\alpha_m}{m!}\|\overline{x}_n - x^*\|^{m-2} + \ldots + \frac{\alpha_2}{2!}}{1 - \alpha_2\|\overline{x}_n - x^*\| - \ldots - \frac{\alpha_{m+1}}{m!}\|\overline{x}_n - x^*\|^m}\|\overline{x}_n - x^*\| \tag{24}$$

for $n \geq 0$. Using (9), (15) and (24) we can easily see that $\|\bar{\phi}(\overline{x}_n) - x^*\| < \|\overline{x}_n - x^*\|$ if

$$\frac{\|x_n - x^*\|}{\|x^*\|} > \frac{\eta_n}{1-(1+\eta_n)\sigma_n}, \quad (1+\eta_n)\sigma_n < 1. \tag{25}$$

Thus, the crucial condition is $\sigma_n < 1$, and by (24) this condition implies

$$\omega_n < 1, \quad \|\overline{x}_n - x^*\| < \min\{\delta, \delta_n\} \quad (n \geq 0) \tag{26}$$

where δ_n is the positive root of the function

$$f_n(t) = \frac{\alpha_{m+1}}{(m+1)!}(1 - w_n + 2m)t^m + \frac{\alpha_m}{m!}[2m - (1+w_n)]t^{m-1} + \ldots + \frac{\alpha_2}{2!}(3 - w_n)t + w_n - 1 \quad (n \geq 0). \tag{27}$$

Hence, as in condition (3.7) of [8, p. 113], we conclude that the crucial condition is

$$\|\bar{A}_n^{-1}F'(\overline{x}_n) - I\| + \frac{\|\bar{A}_n^{-1}(\overline{F}_n - F_n)\|}{\|F'(\overline{x}_n)^{-1}F_n\|} < 1. \tag{28}$$

3. Concluding comments—applications. The results obtained here have theoretical and practical value. As an example we consider an operator F that satisfies an autonomous differential equation of the form (see [3], [5])

$$F'(x) = T(F(x)), \quad x \in U(x^*, \sigma), \tag{29}$$

where $T : E_2 \to E_1$ is a known Fréchet-differentiable operator. Using (29) we get $F'(x^*) = T(F(x^*)) = T(0)$, and $F''(x^*) = F'(x^*)Q'(F(x^*)) = Q(0)Q'(F(0))$. That is, without knowing the solution x^* we can use the results obtained here. Below, we consider such an example for $m = 2$.

Example. Let $E_1 = E_2 = \mathbb{R}$. Define functions F, T on $U(0,1)$ by

$$F(x) = e^x - 1 \quad (x \in U(0,1)), \tag{30}$$

$$T(x) = x + 1 \quad (x \in U(0,1)). \tag{31}$$

It follows from (30) and (31) that equation (29) is satisfied.

Using (11), (12), (17), (18), (19) and (30) we find for $\omega_0 = 1/2$ that: $\alpha = e$, $\beta = 1$, $\delta = .411254048$ and $\min\{\delta, \delta_0\} = \delta_0 = .27587332$. That is, conditions (16) and (18) are satisfied provided

$$\|x_0 - x^*\| < .411254048 \tag{32}$$

and

$$\|\overline{x}_0 - x^*\| < .27587332, \tag{33}$$

respectively.

In order to compare our results with the ones in [7], [8], let us first introduce

$$\mu = \sup\left\{ \frac{\|F'(x^*)^{-1}[F'(x) - F'(y)]}{\|x-y\|} \,\middle|\, x, y \in U(x^*, \sigma),\ x \neq y \right\}. \tag{34}$$

Then the conditions in [7], [8] corresponding to (16) and (18) are

$$\|x_0 - x^*\| < \frac{2}{3\mu} \tag{35}$$

and

$$\|\overline{x}_0 - x^*\| < \frac{2(1-\omega_0)}{(3-\omega_0)\mu}, \tag{36}$$

respectively.

It can be easily seen from (30) and (34) that $\mu = e$. Hence, conditions (35) and (36) are satisfied provided that

$$\|x_0 - x^*\| < .245253, \tag{37}$$

$$\|\overline{x}_0 - x^*\| < .1471518, \tag{38}$$

respectively. That is, (32) and (35) provide a wider choice for x_0 and $\overline{x}_0$ than conditions (37) and (38) respectively. It turns out that the ratios of convergence are smaller in our case also. Indeed, (15) and (20) give respectively for $\|x_0 - x^*\| \leq .2$ and $\|\overline{x}_0 - x^*\| \leq .1$ that

$$\begin{aligned} \|x_0 - F'(x_0)^{-1}F(x_0) - x^*\| &\leq .913609703\|x_0 - x^*\|^2 \\ &\leq .182721941\|x_0 - x^*\| \end{aligned} \tag{39}$$

and

$$\|\overline{\phi}(\overline{x}_0) - x^*\| \leq .599944213\|\overline{x}_0 - x^*\|. \tag{40}$$

The corresponding results in [7], [8] are

$$\|x_0 - F'(x_0)^{-1}F(x_0) - x^*\| \leq \frac{\mu\|x_0 - x^*\|^2}{2(1-\mu\|x_0 - x^*\|)} \tag{41}$$

and

$$\|\overline{\phi}(\overline{x}_0) - x^*\| \leq \left\{ \omega_0 + \frac{(1+\omega_0)\mu\|\overline{x}_0 - x^*\|}{2(1-\mu\|\overline{x}_0 - x^*\|)} \right\} \|\overline{x}_0 - x^*\|, \tag{42}$$

respectively. If we use the above values, (41) and (42) give

$$\begin{aligned} \|x_0 - F'(x_0)^{-1}F(x_0) - x^*\| &\leq .913609703\|x_0 - x^*\|^2 \\ &\leq .182721941\|x_0 - x^*\| \end{aligned} \tag{43}$$

and

$$\|\bar{\phi}(\overline{x}_0) - x^*\| \leq .599944213\|\overline{x}_0 - x^*\|, \quad (44)$$

respectively. That is, our ratios of convergence (39) and (40) are smaller than (43) and (44) given in [7], [8]. These observations are important in numerical computations.

Our results can be compared favorably with all the examples given in [8]. However, we leave the details to the motivated reader.

References

[1] I. K. Argyros, *On the convergence of some projection methods with perturbations*, J. Comput. Appl. Math. 36 (1991), 255–258.

[2] —, *Concerning the radius of convergence of Newton's method and applications*, Korean J. Comput. Appl. Math. 6 (1999), 451–462.

[3] I. K. Argyros and F. Szidarovszky, *The Theory and Application of Iteration Methods*, CRC Press, Boca Raton, FL, 1993.

[4] R. S. Dembo, S. C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal. 19 (1982), 400–408.

[5] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1982.

[6] T. J. Ypma, *Numerical solution of systems of nonlinear algebraic equations*, Ph.D. thesis, Oxford, 1982.

[7] —, *Affine invariant convergence results for Newton's method*, BIT 22 (1982), 108–118.

[8] —, *The effect of rounding errors on Newton-like methods*, IMA J. Numer. Anal. 3 (1983), 109–118.

Ioannis K. Argyros
Cameron University
Department of Mathematics
Lawton, OK 73505, U.S.A.
E-mail: ionnisa@cameron.edu

Received on 22.12.1999