BULL. AUSTRAL. MATH. SOC. 47D15, 47H17, 65J15, 65B05
VOL. 45 (1992) [415–422]

SHARP ERROR BOUNDS FOR NEWTON-LIKE METHODS UNDER WEAK SMOOTHNESS ASSUMPTIONS

IOANNIS K. ARGYROS

We provide sufficient convergence conditions as well as sharp error bounds for Newton-like iterations which generalise a wide class of known methods for solving nonlinear equations in Banach space.

1. INTRODUCTION

Let F be a nonlinear operator defined on a convex subset E_3 of a Banach space E_1 with values in a Banach space E_2. A lot of methods for solving the equation

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

can be written in the Newton-like form

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

where for each $n \geqslant 0$, $A(x_n)^{-1}$ is a bounded linear operator from E_2 into E_1 (that is, $A(x_n)^{-1} \in L(E_2, E_1)$). Obviously the linear operator $A(x_n)$ must be a consistent approximation to the Fréchet-detivative F' of F. The best known method of type (2) are Newton's methods, where $A(x_n) = F'(x_n)$, and the secant method, where $A(x_n) = \delta F(x_n, x_{n-1})$, $n \geqslant 0$, δF being a consistent approximation of the Fréchet-derivative of F. Other authors (see, for example **[1, 2, 3, 6, 7, 8, 9]** and the references therein) in order to find an approximate solution x^* of equation (1) have imposed various conditions such as

$$\left\| A(x_0)^{-1}(F'(x + t(y - x)) - A(x)) \right\| \leqslant w[(\|x - x_0\| + t\|y - x\|)^p], \tag{3}$$

$$\left\| A(x_0)^{-1}(A(x) - A(x_0)) \right\| \leqslant w_0(\|x - x_0\|^p) \tag{4}$$

for all $x, y \in E_3$ and some $p, t \in [0, 1]$, where $x_0 \in E_3$. Here w, w_0 denote non-decreasing continuous functions from $|\mathbb{R}^+$ into $|\mathbb{R}^+$ with $w(0) = w_0(0) = 0$. Denote by N the class of all such functions. However these conditions do not provide sharp error estimates for Newton-like methods when $0 < p < 1$ (see for example **[1, 2, 3, 6, 7, 8, 9]**. In the elegant paper by Galperin and Waksman [4] sharp error bounds were found for Newton's method using the notion of a w-regularly continuous operator. Here we use a generalised notion of the above definition and provide sharp error bounds for Newton-like methods. Our results can be compared favourably with results already in the literature for various choices of the linear operator $A(x)$.

Received 9 May 1991

2. Convergence Results

Given an operator $G\colon E_3 \subset E_1 \to E_2$, and a linear operator $A(x) : E_3 \to E_2$, we say that G is w, A-continuous at a point $x \in E_3$ if the function w belongs to the class

$$M(G, x, E_3) := \{w \in N \mid \forall y \in E_3 \left\|A(x_0)^{-1}(G(y) - A(x))\right\| \leqslant w(\|x - y\|)\},$$

and that G is w, A-continuous on E_3 if w belongs to

$$M(G, E_3) := \{w \in N \mid \forall x,\, y \in E_3 \left\|A(x_0)^{-1}(G(x) - A(y))\right\| \leqslant w(\|x - y\|)\}.$$

All functions of $M(G, x, E_3)$ are called local continuity moduli of G(at x), whereas those of $M(G, E_3)$ are called (global) continuity moduli of G (on E_3) [3, 6].

Let N^* denote the subclass of N consisting of all $w \in N$ that are concave. Denote

$$H(x, y) = \min\left\{\left\|A(x_0)^{-1}G(x)\right\|,\, \left\|A(x_0)^{-1}A(y)\right\|\right\},\; x,\, y \in E_3.$$

Given $w \in N^*$, we say that G is w, A-regularly continuous on E_3, if

$$\begin{aligned} w^{-1}&\Big(H(x,\, x + t(y-x)) + \left\|A(x_0)^{-1}(G(x + t(y-x)) - A(x))\right\|\Big) \\ &- w^{-1}(H(x,\, x + t(y-x))) \leqslant \|x_0 - x\| + t\,\|y - x\| \end{aligned} \tag{5}$$

for all $x, y \in E_3$ and $t \in [0, 1]$.

Here $w^{-1}(s)$ stands for the least root of the equation $w(t) = s$. Clearly, w^{-1} is an increasing convex function defined on $[0, w(\infty))$. Because of w^{-1} convexity, the above inequality implies $w \in M(G, E_3)$. As in [4] we can show that the converse is not always true. For $x_0, x, y \in E_3$, assume $A(x_0)$ is invertible and define the numbers α, r, $\overline{a}$, $\overline{a}'$, $\overline{b}$, c, q by

$$\begin{gathered} \left\|A(x_0)^{-1}F(x_0)\right\| \leqslant \alpha,\; r = \|x - y\|,\; \overline{a} = w^{-1}\left(\left\|A(x_0)^{-1}A(x)\right\|\right), \\ \overline{a}' = w^{-1}\left(\left\|A(x_0)^{-1}F'(x)\right\|\right),\; \overline{b} = w^{-1}\left(\left\|A(x_0)^{-1}F'(y)\right\|\right) - r; \\ c = \|x - x_0\|,\; q = \overline{a}' - \overline{b}, \end{gathered}$$

the functions $q(s, t)$, R^+, B, C, D by

$$\begin{gathered} q(s, t) = \min\{t,\, s - t\},\; R^+ = \max\{R, 0\}, \\ B(a, a'b, c, r) = \int_0^r \left[w\Big(\min\{a, (a' - q(s,t))^+\} + c + t\Big) - w\Big(\min\{a, (a' - q(s,t))^+\}\Big)\right] dt, \\ C(r) = B(a(r),\, a'(r),\, b(r),\, r,\, r), \end{gathered}$$

with (for each fixed $r \geqslant 0$)

$$a = a(r) = w^{-1}(1 - w_0(r)),\ b = b(r) = w^{-1}(1 - w_0(r) - w(r)) - r,$$
$$a' = a'(r) = w^{-1}(1 - w_0(r) - w(r))$$

and
$$D(r) = \alpha + \frac{C(r)}{1 - w_0(r)}.$$

Finally, define the iteration $\{t_n\}$, $n \geqslant 0$, by $t_0 = 0$, $t_1 = \alpha$ and

$$t_{n+2} = t_{n+1} + \frac{B(a(t_{n+1} - t_n),\ a'(t_{n+1} - t_n),\ b(t_{n+1} - t_n),\ t_n,\ t_{n+1} - t_n)}{1 - w_0(t_n)},\ n \geqslant 0.$$

We can now state and prove the main result:

THEOREM. *Let* $F\colon E_3 \subset E_1 \to E_2$ *and* $w \in N^*$
Assume:

(i) *There exist* $x_0 \in E_3$ *and a positive number* α *such that* $A(x_0)$ *is invertible and* $\left\|A(x_0)^{-1}F(x_0)\right\| \leqslant \alpha$.

(ii) *There exists a minimum positive number* $r* \in \left(0,\ w_0^{-1}(1)\right)$ *such that*

(6) $$D(r) \leqslant r \text{ and } 1 - w_0(r) - w(r) \geqslant 0 \text{ for all } f \in (0,\ r^*].$$

(iii) $U = U(x_0,\ r^*) = \{x \in E_1 \mid \|x - x_0\| \leqslant r^*\} \subset E_3$.

(iv) *Given* $A(x) \in L(U,\ E_2)$ *satisfying (4) for* $p = 1$ *for all* $x \in U$, *let* F *be Fréchet differentiable on* U *and* F' *be* $w - A$ *regularly continuous on* U.

Then,

(1) *the function* B *does not increase in each of its first three arguments and increases in the other two;*

(2) *the iteration* $\{t_n\}$, $n \geqslant 0$ *is increasing and bounded above by* r^* *with* $t^* = \lim_{n\to\infty} t_n \leqslant r^*$;

(3) *the operator* $A(x)$ *is invertible on* U;

(4) *the Newton-like iterations (2) are well defined, remain in* $U(x_0,\ t^*)$ *and converge to a solution* x^* *of equation (1);*

(5) x^* *is the unique solution of equation (1) in* $U(x_0,\ r^*)$;

(6) *the following estimates are true:*

(7) $$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n$$

(8) $$\|x_n - x^*\| \leqslant t^* - t_n \text{ for all } n \geqslant 0$$

(9)
$$\|x_n - x^*\| \leqslant \frac{B(a(\|x_n - x_{n-1}\|), a'(\|x - n - x_{n-1}\|), b(\|x_n - x_{n-1}\|), \|x_{n-1} - x_0\|, \|x_n - x_{n-1}\|)}{1 - w_0(\|x^* - x_0\|)}$$
for all $n \geqslant 1$

(10)
$$\|x_{n+1} - x_n\| \leqslant \|x_n - x^*\| + \frac{B(a(\|x^* - x_n\|), a'(\|x^* - x_n\|), b(\|x^* - x_n\|), \|x_n - x_0\|, \|x^* - x_n\|)}{1 - w_0(\|x_n - x_0\|)}$$

for all $n \geqslant 0$.

(7) *the convergence condition (6) and the estimates (7)–(9) are sharp.*

PROOF: (1) The proof of this part is similar to the corresponding one in [4, Lemma 2.1] and so is omitted.

(2) The first two members of the iteration $\{t_n\}$, $n \geqslant 0$ are such that $t_0 < t_1 \leqslant r^*$. Therefore the denominator of the fraction appearing in the definition of the sequence is positive. That is, $t_1 \leqslant t_2$ (since the numerator is obviously nonnegative). Let us assume that $t_k \leqslant t_{k+1}$, $k = 0, 1, 2, \ldots, n$. Then by the definition of the sequence, $\{t_n\}$, $n \geqslant 0$, $t_{k+1} \leqslant t_{k+2}$. That is, $t_{n+1} \leqslant t_{n+2}$ for $n = k+1$. So far we have shown that the scalar sequence $\{t_n\}$ is increasing for all $n \geqslant 0$. We will show that $t_n \leqslant r^*$ for all $n \geqslant 0$. For $n = 0, 1$ this is true by hypothesis. For $n = 2$, $t_2 \leqslant r^*$, since $t_2 \leqslant D(r^*) \leqslant r^*$. Let us assume that $t_k \leqslant r^*$, $k = 0, 1, 2, \ldots, n$; then

$$C(t_1 - t_0) + C(t_2 - t_1) + \ldots + C(t_{k+1} - t_k) \leqslant C(t_{k+1} - t_0) \leqslant C(t_{t+1}) \leqslant C(r^*),$$

since the function w is increasing and $(t_1 - t_0) + (t_2 - t_1) + \ldots + (t_{k+1} - t_k) = t_{k+1} - t_0$. Hence $t_{k+1} \leqslant C(r^*) \leqslant r^*$, which completes the induction. Therefore the sequence $\{t_n\}$, $n \geqslant 0$ is increasing and bounded above by r^* and as such it converges to some t^* such that $0 < t^* \leqslant r^*$.

(3) Let us observe that the linear operator $A(u)$ is invertible for all $u \in U\left(x_0, w_0^{-1}(1)\right)$. Indeed we obtain

$$\left\|A(x_0)^{-1}(A(u) - A(x_0))\right\| \leqslant w_0(\|u - x_0\|) < 1,$$

so that according to Banach's lemma $A(u)$ is invertible and

$$\left\|A(u)^{-1}A(x_0)\right\| \leqslant (1 - w_0(\|u - x_0\|))^{-1}. \tag{11}$$

Note also that since $\left\|A(x_0)^{-1}A(u)\right\| \cdot \left\|A(u)^{-1}A(x_0)\right\| \geqslant 1$, then $\left\|A(x_0)^{-1}A(u)\right\| \geqslant 1 - w_0(\|u - x_0\|)$.

(4)–(6) It now follows that if (2) is well defined for $n = 1, 2, 3, \ldots, k$ and if (7) holds for $n \leqslant k$ then

$$\|x_0 - x_n\| \leqslant t_n - t_0 \leqslant t^* - t_0 \text{ for } n \leqslant k.$$

This shows that (7) is satisfied for $u = x_i$ with $i \leqslant k$. Thus (2) is well defined for $n = k+1$ too. Also from $\|x_0 - x_k\| \leqslant t_k - t_0 \leqslant t^*$ we obtain $x_k \in U(x_0, t^*)$.

We now observe that (7) is true for $n = 0$. Assume that it is true for $k = 0, 1, 2, \ldots, n$. Then by (2)

(12)

$$\begin{aligned}
&\|x_{k+2} - x_{k+1}\| = \left\|A(x_{k+1})^{-1}F(x_{k+1})\right\| \\
&= \left\|A(x_{k+1})^{-1}(F(x_{k+1}) - F(x_k) - A(x_k)(x_{k+1} - x_k))\right\| \\
&= \left\|A(x_{k+1})^{-1}A(x_0)\right\| \cdot \left\|A(x_0)^{-1}(F(x_{k+1}) - F(x_k) - A(x_k)(x_{k+1} - x_k))\right\| \\
&= \left\|A(x_{k+1})^{-1}A(x_0)\right\| \cdot \left\|A(x_0)^{-1}\left[\int_0^1 F'(x_h + F(x_{k+1} - x_k)) - A(x_k)\right](x_{k+1} - x_k)dt\right\|.
\end{aligned}$$

We now apply (11) for $u = x_{k+1}$, (5) for $x = x_k$, $y = x_{k+1}$, $G = F'$ to obtain

(13)

$$\begin{aligned}
&\|x_{K+2} - x_{K+1}\| \\
&\leqslant \frac{1}{1 - w_0(t_k)} \int_0^1 [w(w^{-1}(H(x_k, x_k + t(x_{k+1} - x_k)) + \|x_k - x_0\| + \|x_{k+1} - x_k\| t)) \\
&\qquad - w(w^{-1}(H(x_k, x_k + t(x_{k+1} - x_k))))] \|x_{k+1} - x_k\| \, dt \\
&\leqslant \frac{B(a(t_{k+1} - t_k), a'(t_{k+1} - t_k), b(t_{k+1} - t_k), t_k, t_{k+1} - t_k)}{1 - w_0(t_k)} = t_{k+2} - t_{k+1}.
\end{aligned}$$

This shows (7) for $n = k+1$. Hence, $\{x_n\}$, $n \geqslant 0$ is a Cauchy sequence in a Banach space and as such it converges to a point $x^* \in U$. By (12) and (13) we observe that the numerator of (13) is an upper bound for $\left\|A(x_0)^{-1}F(x_{k+1})\right\|$ which tends to 0 as $k \to \infty$. Hence, by continuity, $F(x^*) = 0$. The estimate (8) now follows easily from (7).

To show uniqueness, let us assume that there exist two solutions x^* and y^* in $U(x_0, r^*)$ and consider the estimate $F(x^*) - F(y^*) = L^*(x^* - y^*)$ with $L^* = \int_0^1 F'(y^* + t(x^* - y^*))dt$.

Then as before (see (11)) we can show $\left\|I - A(x_0)^{-1}L^*\right\| < 1$. That is, L^* is invertible, which shows $x^* = y^*$.

Set $L = \int_0^1 F'(x^* + t(x_n - x^*))dt$ and use (11), and the estimates

$$\|x_n - x^*\| \leqslant \left\| \left(A(x_0)^{-1}L\right)^{-1} \right\| \cdot \left\| A(x_0)^{-1}F(x_n) \right\|,$$

$$\|x_{n+1} - x_n\| = (x^* - x_n) + \left(A(x_0)^{-1}A(x_n)\right)^{-1}$$
$$[A(x_0)^{-1}\left(F(x^*) - F(x_n) - A(x_n)^{-1}(x^* - x_n)\right)]$$

to obtain (9) and (10) respectively.

(7) This follows exactly as in part (5) Theorem 2.1 in [4], which completes the proof of the theorem.

It can easily be seen that if $w(t) = \gamma t$, $w_0(t) = \beta t$ for some $\beta, \gamma > 0$ and the sequence $\|x^* - x_n\|$ is monotone then (9) and (10) can provide an upper and a lower bound on $\|x^* - x_n\|$ respectively expressed in terms of the rest of the norms. Moreover define the numbers r_1, r_2, r_3, Δ and the intervals I_1, I_2, I by

$$r_1 = \frac{1}{\beta + \gamma}, \quad r_2 = \frac{1 + \alpha\beta - \sqrt{\Delta}}{3\gamma + 2\beta}, \quad r_3 = \frac{1 + \alpha\beta + \sqrt{\Delta}}{3\gamma + 2\beta},$$

with $\Delta = (1 + \alpha\beta)^2 - 2\alpha(3\gamma + 2\beta)$, $I_1 = (0, r_1)$, $I_2 = [r_2, r_3]$, and $I = I_1 \cap I_2$.

Assume:

$$\Delta > 0 \text{ and } I \neq \emptyset \text{ and set } I_3 = [r_2, \min(r_1, r_3)].$$

It can easily be seen then that condition (6) is satisfied for all $r \in I_3$.

Similar conditions can be obtained when $w_0(t) = \beta t^p$, $w(t) = \gamma t^p$ for $p \in [0, 1)$. In the latter case the results in [**3, 5, 7, 8, 9**] cannot apply (since $p = 1$ there). Moreover it can easily be seen that our results compare favourably with the ones in [**1, 2, 3, 6**] in this case.

Finally, consider the equation

$$F(x) + Q(x) = 0 \tag{14}$$

where F is as before and Q is a nonlinear operator defined on E_3 with values on E_2 such that

$$\left\| A(x_0)^{-1}(Q(x) - Q(y)) \right\| \leqslant w_1(\|x - y\|) \text{ for all } x, y \in E_3 \tag{15}$$

for some nondecreasing real function w_1 defined on R^+ with $w_1(0) = 0$. Note that the differentiability of Q is not assumed. Define the function $D_1(r)$ by

$$D_1(r) = \alpha + \frac{C_1(r)}{1 - w_0(r)}, \quad C_1(r) = C(r) + w_1(r)$$

and the iteration $\{s_n\}$, $n \geqslant 0$, by $s_0 = 0$, $s_1 = \alpha$ and

$$s_{n+2} = s_{n+1} + \frac{B(a(s_{n+1} - s_n), a'(s_{n+1} - s_n), b(s_{n+1} - s_n), s_n, s_{n+1} - s_n) + w_1(s_{n+1} - s_n)}{1 - w_0(s_n)},$$
$$n \geqslant 0.$$

Then with the rest of the notation as before we can immediately state and prove a theorem for approximating a solution x^* of equation (14) similar to the one above. Just replace $D(r)$ by $D_1(r)$ and $\{t_n\}$ by $\{s_n\}$, $n \geqslant 0$, in the above theorem and take into account hypothesis (15).

Note that the iteration (2) will become

$$z_{n+1} = z_n - A(z_n)^{-1}(F(z_n) + Q(z_n)),\ z_0 \in E_3,\ n \geqslant 0.$$

The new theorem will cover the case when the operator appearing in equation (1) is not Fréchet-differentiable but it can be decomposed into one that is and one that is not (see also [2] and the references therein). □

References

[1] I.K. Argyros, 'Newton-like methods under mild differentiability conditions with error analysis', *Bull. Austral. Math. Soc.* **37** (1988), 131–147.

[2] X. Chen and T. Yamamoto, 'Convergence domains of certain iterative methods for solving nonlinear equations', *Numer. Funct. Anal. Optim.* **10** (1989), 37–48.

[3] J.E. Dennis, 'Toward a unified convergence theory for Newton-like methods', in *Nonlinear functional analysis and applications*, Editor L.B. Rall, pp. 425–472 (Academic Press, New York, 1971).

[4] A. Galperin and Z. Waksman, 'Newton's method under a weak smoothness assumption', *J. Comput. Appl. Math.* (to appear).

[5] L.V. Kantorovich and G.P. Akilov, *Functional analysis in normed spaces* (Pergamon Press, Oxford, 1964).

[6] G.J. Miel, 'Unified error analysis for Newton-type methods', *Numer. Math.* **33** (1979), 391–396.

[7] F.A. Potra, 'On the convergence of a class of Newton-like methods', in *Iterative solution of nonlinear systems of equations: Lecture notes in Math. 953*, Editors R. Ansorge, T. Meis and W. Torning, pp. 125–137 (Springer-Verlag, Berlin, Heidelberg, New York, 1982).

[8] W.C. Rheinboldt, 'A unified convergence theory for a class of iterative process', *SIAM J. Numer. Anal.* **5** (1968), 42–63.

[9] T. Yamamoto, 'A convergence theorem for Newton-like methods in Banach spaces', *Numer. Math.* **51** (1987), 545–557.

Department of Mathematics
Cameron University
Lawton OK 73505
United States of America