Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 18 (1999), No. 3, 785–792

Concerning the Convergence of a Modified Newton-Like Method

Ioannis K. Argyros

Abstract. We provide sufficient convergence conditions for a certain Newton-like method to a locally unique solution of a nonlinear equation in a Banach space. We assume that the Fréchet-derivative of the operator involved satisfies in some sense uniformly continuous conditions, which are weaker than earlier ones. We show that our results apply where earlier ones fail. Finally, we solve a nonlinear integral equation of Uryson-type that cannot be solved using Proposition 2 in [10].

Keywords: *Modified Newton-like method, Banach spaces, Uryson integral equations, uniformly continuity*

AMS subject classification: 65 J 15, 65 B 05, 47 H 17, 49 D 15

1. Introduction

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

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

where F is a Fréchet-differentiable operator defined on a convex subset D of a Banach space E_1 with values in a Banach space E_2. We propose the modified Newton-like method

$$x_{n+1} = x_n - A^{-1}F(x_n) \qquad (n \geq 0, x_0 \in D) \tag{2}$$

to generate a sequence $\{x_n\}_{n\geq 0}$ converging to x^*. Here $A \in L(E_1, E_2)$, the space of bounded linear operators from E_1 into E_2. For $A = F'(x_0)$ we obtain the modified Newton method, whereas for $A = [x_{-1}, x_0; F]$ (divided difference of order one) we obtain the modified Secant method. Several other choices are also possible [5, 6, 8, 9].

Let $x_0 \in D$, $U(x_0, R) = \{x \in E_1 | \, \|x - x_0\| < R\} \subseteq D$, and assume that

$$\|A^{-1}(F'(x) - F'(x_0))\| \leq w(\|x - x_0\|) \qquad (x \in \bar{U}(x_0, r); 0 \leq r < R) \tag{3}$$

for some monotonically increasing function w satisfying

$$\lim_{t\to 0} w(t) = 0. \tag{4}$$

I. K. Argyros: Cameron Univ., Dept. Math., Lawton, OK 73505, U.S.A.

ISSN 0232-2064 / $ 2.50

In the elegant paper [2] stronger conditions of the form

$$\|F'(x_1) - F'(x_2)\| \le v(\|x_1 - x_2\|) \qquad (x_1, x_2 \in U(x_0, R)) \tag{5}$$

or, more generally,

$$\|F'(x_1) - F'(x_2)\| \le v_1(r, \|x_1 - x_2\|) \qquad (x_1, x_2 \in \bar{U}(x_0, r); 0 \le r < R) \tag{6}$$

have been used for some monotonically increasing positive functions v and v_1 with

$$\lim_{t\to 0} v(t) = 0 = \lim_{t\to 0} v_1(r, t) \qquad (0 \le r < R), \tag{7}$$

in connection with Newton's method

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \qquad (n \ge 0, x_0 \in D). \tag{8}$$

Conditions of type (5) and (6) have also been studied in the special cases when $v(t) = kt^\lambda$ or $v_1(r,t) = k(r)t^\lambda$ for $\lambda \in [0,1]$ in connection with (8) or more generally with Newton-like methods of the form

$$x_{n+1} = x_n - A(x_n)^{-1}F(x_n) \qquad (n \ge 0, x_0 \in D \tag{9}$$

[3 - 6]. Here for each fixed $x \in D$, $A(x) \in L(E_1, E_2)$.

A semilocal convergence theorem is provided here for method (2) under the weak condition (3). In order to demonstrate the importance of condition (3), we apply our results to solve a nonlinear integral equation of Uryson type using method (2) for $A = F'(x_0)$. At the same time we show that corresponding results in the above-mentioned papers do not guarantee the convergence of method (2) to a solution of equation (1). Finally, we note that the results obtained here are in an affine invariant form, whereas the ones in [2, 8, 10] are not. The advantages of results given in affine invariant form over corresponding ones not in this form have been explained in [5, 7].

For example, the Newton-Kantorovich theorem guarantees the existence of a solution of equation (1) if

$$2\|x_1 - x_0\|\,\|F'(x_0)^{-1}\|\ell_1 \le 1 \qquad \text{or} \qquad 2\|x_1 - x_0\|\ell \le 1,$$

where

$$\|F'(x) - F'(y)\| \le \ell_1\|x - y\| \quad \text{or} \quad \|F'(x_0)^{-1}(F'(x) - F'(y))\| \le \ell\|x - y\|$$

for all $x, y \in D$. However, for linear operators $L_1, L_2 \in L(E_1, E_2)$ we have $\|L_1L_2\| \le \|L_1\| \cdot \|L_2\|$. Hence if the first inequality holds above so does the second. However, the converse is not necessarily true (see [7: p. 2/Example 1]). Moreover, in some iterative methods, the generated sequence is known in some subset S of all affine transformations with domain E_2. In these cases, it is reasonable to require only S-invariance for the associated convergence theorems. This means that both the assumptions and the statements of the theorems should remain unaltered, when F is replaced by LF for any $L \in S$.

2. Convergence analysis

It is convenient to define the number

$$a = \|A^{-1}F(x_0)\|, \tag{10}$$

the function

$$f(r) = a - r + \int_0^r w(t)\,dt, \tag{11}$$

the equation

$$f(r) = 0 \tag{12}$$

and the iteration

$$\left.\begin{aligned} t_0 &= 0 \\ t_{n+1} &= t_n + f(t_n) \quad (n \geq 0). \end{aligned}\right\} \tag{13}$$

We need the following lemma concerning the convergence of iteration (13). A similar result is shown in [10: p. 675]. However, our proof is slightly difference, since we make the assumption about the existence and the uniqueness of solution r^* of equation (12).

Lemma. Assume that equation (12) has a unique solution $r^* \in [0, R]$, and that $f(R) \leq 0$. Then iteration $\{t_n\}_{n\geq 0}$ generated by (13) is monotonically increasing and converges to r^*.

Proof. The function $g(t) = t + f(t)$ is clearly increasing on $[0, r^*]$ and $f(t) \geq 0$ for $t \in [0, r^*]$. So, if $t_k \in [0, r^*]$ for some k, we get

$$t_k \leq t_k + f(t_k) = t_{k+1} \qquad \text{and} \qquad t_{k+1} = t_k + f(t_k) \leq r^* + f(r^*) = r^*.$$

This proves the lemma by induction ∎

We can now show the main semilocal convergence theorem for method (2).

Theorem. *Assume that F is Fréchet-differentiable with* (3) *and* (4), *and let f be defined by* (11). *Suppose that equation* (12) *has a unique solution $r^* \in [0, R]$ and $f(R) \leq 0$. Then equation* (1) *has a solution $x^* \in \bar{U}(x_0, r^*)$; this solution is unique in $U(x_0, R)$. Moreover, method* (2) *generates a sequence $\{x_n\}_{n\geq 0}$ which converges to x^*. Furthermore, the error bounds*

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n \tag{14}$$

$$\|x^* - x_n\| \leq r^* - t_n \tag{15}$$

hold for all $n \geq 0$.

Proof. We first show estimate (14) using induction on the integer n. For $n = 0$, (2) and (10) give $\|x_1 - x_0\| \leq a = t_1 - t_0$, which shows (14) in this case. Suppose (14) holds for $n = 0, 1, \ldots, k-1$; this implies, in particular, that

$$\begin{aligned} \|x_k - x_0\| &\leq \|x_k - x_{k-1}\| + \ \ldots \ + \|x_1 - x_0\| \\ &\leq (t_k - t_{k-1}) + \ \ldots \ + (t_1 - t_0) \\ &= t_k - t_0 \\ &= t_k \\ &\leq r^*. \end{aligned}$$

Using the approximation

$$\begin{aligned}F(x_k) &= F(x_k) - F(x_{k-1}) - A(x_k - x_{k-1})\\ &= \int_0^1 \Big[F'\big(x_{k-1} + t(x_k - t_{k-1})\big) - A\Big](x_k - x_{k-1})\,dt,\end{aligned} \tag{16}$$

hypothesis (3), (13), and the induction hypothesis, we obtain in turn

$$\begin{aligned}\|x_{k+1} - x_k\| &= \|A^{-1}F(x_k)\|\\ &= \left\|\int_0^1 A^{-1}\big[F'\big(x_{k-1} + t(x_k - x_{k-1})\big) - A\big](x_k - x_{k-1})\,dt\right\|\\ &\le \int_0^1 \|A^{-1}\big[F'(x_{k-1} + t(x_k - x_{k-1})\big) - A\big] \cdot \|x_k - x_{k-1}\|\,dt\\ &\le \int_0^1 w\big(\|x_{k-1} + t(x_k - x_{k-1}) - x_0\|\big)\|x_k - x_{k-1}\|\,dt\\ &\le \int_0^1 w\big((1-t)\|x_{k-1} - x_0\| + t\|x_k - x_0\|\big)\|x_k - x_{k-1}\|\,dt\\ &\le \int_0^1 w\big((1-t)t_{k-1} + t_k\big)(t_k - t_{k-1})\,dt\\ &= \int_{t_{k-1}}^{t_k} w(t)\,dt\\ &= f(t_k)\\ &= t_{k+1} - t_k.\end{aligned} \tag{17}$$

Hence, estimate (14) is true for all $n \ge 0$. Moreover, estimate (15) follows from (14) by using standard majorization techniques [5, 8, 10].

To show uniqueness, let us assume that $y \in U(x_0, R)$ with $F(y) = 0$. Using (2), (3) and the approximation

$$\begin{aligned}x_{k+1} - y^* &= x_k - y^* - A^{-1}F(x_k)\\ &= A^{-1}\big[A(x_k - y^*) - \big(F(x_k) - F(y)\big)\big]\\ &= -A^{-1}\left[\int_0^1 \big(F'\big(y + t(x_k - y)\big) - A\big)\right](x_k - y)\,dt, \}cr\end{aligned} \tag{18}$$

we get as in (17)

$$\|x_{k+1} - y^*\| \le \int_0^1 w\big((1-t)\|x_0 - y\| + t\|x_k - x_0\|\big)\|x_k - y\|\,dt. \tag{19}$$

If $y \in \bar{U}(x_0, r^*)$, then $\|x_0 - y^*\| \le r^*$ and (19) gives

$$\|x_{k+1} - y\| \le w(r^*)\|x_k - y^*\| \le w(r^*)^{k+1}r^*. \tag{20}$$

If $y \in U(x_0, R)$ and $r^* \neq R$, then $\|y - x_0\| = \mu R \ (0 < \mu < 1)$, and (19) can give

$$\|x_{k+1} - y\| \leq \left(\frac{1}{R - r^*} \int_{r^*}^{R} w(t)\, dt \right)^{k+1} R. \tag{21}$$

It follows from the Lemma and (11) that $0 \leq w(t) < 1$ for $t \in [0, R]$. Hence it follows from (20) and (21) that $\lim_{k\to\infty} x_k = y$. That completes the proof of the theorem ∎

Remark 1. It was explained in [2], why sometimes it is useful to pass from the function w to the function

$$\bar{w}(r) = \sup \big\{ w(u) + w(v) : \ u + v = r \big\}. \tag{22}$$

Consider the function

$$\bar{f}(r) = \bar{a} - r + \int_0^r \bar{\bar{w}}(t)\, dt \qquad (r \in [0, R])$$

where $\bar{a} = \frac{a}{1+w(a)}$ and $\bar{\bar{w}}(r) = \frac{1}{1+w(a)} \bar{w}(r) \ \ (r \in [0, R])$, the equation

$$\bar{f}(r) = 0$$

and the iteration

$$\left.\begin{aligned} \bar{t}_0 &= 0 \\ \bar{t}_{n+1} &= \int_0^{\bar{t}_n} \bar{w}(t)\, dt - w(a)\bar{t}_n + a \ \ (n \geq 0). \end{aligned}\right\}$$

Replace f by $\bar{f}$ in the above theorem. Then it can easily be seen by following the proof of the theorem that the conclusions obtained there hold in this setting also.

The following result is a consequence of the contraction mapping principle [4, 7].

Proposition. *Let*

$$\theta = \sup_{x \in D_0} \|I - A^{-1}F'(x)\| < 1$$

where $D_0 = U(x_0, r_1) \subseteq D$ *and* $r_1 = \frac{a}{1-\theta}$. *Then equation* (1) *has a unique solution* x^* *in* $U(x_0, r_1)$. *Moreover, iteration* $\{x_n\}_{n\geq 0}$ *generated by* (2) *is well defined, remains in* $U(x_0, r_1)$ *for all* $n \geq 0$ *and converges to* x^*. *Furthermore, the error bounds*

$$\|x_n - x^*\| \leq \frac{\theta}{1-\theta}\|x_n - x_{n-1}\| \leq \frac{\theta^n}{1-\theta}\|x_1 - x_0\| = \frac{\theta^n a}{1-\theta}$$

hold for all $n \geq 1$.

To compare our results with the ones obtained in [10: Proposition 2] or in [2: Theorem 1], we set $A = F'(x_0)$, and assume that (6) holds for $v_1(r,t) = k(r)t$ for some non-decreasing function k on $[0, R]$. As in [2, 10] define the function

$$\chi(r) = a + b \int_0^r (r - t)k(t)\, dt - r \tag{23}$$

where

$$b = \|F'(x_0)^{-1}\| \tag{24}$$

and the equation

$$\chi(r) = 0. \tag{25}$$

We now provide a favorable example for our theorem.

Example. Let $E_1 = E_2 = D = C = C[0,1]$, the space of continuous functions on $[0,1]$ equipped with the sup-norm. Consider the nonlinear integral equation of Uryson type given by

$$F(x)(t) = x(t) - \int_0^1 p(t,s,x(s))\,ds. \tag{26}$$

Let $x_0 = 0$ and $A = F'(0)$. Suppose $p(t,s,u) = p_1(t)p_2(s)p_3(u)$ with two continuous functions p_1, p_2 and $p_3 \in C^2$. Setting

$$\gamma = \int_0^1 p_2(s)\,ds \qquad \text{and} \qquad \delta = \int_0^1 p_1(s)p_2(s)\,ds, \tag{27}$$

using formulas (23), (25) - (29) in [1: p. 278], (10), and (24), we get the function

$$k(r) = \|p_1\|_C \cdot \gamma \cdot \sup_{\|u\|\le r} \|p_3''(u)\| \tag{28}$$

and the constants

$$a = \frac{\gamma p_3(0)}{1-\delta p_3'(0)}\|p_1\|_C \tag{29}$$

$$b = 1 + \frac{\gamma p_3'(0)}{1-\delta p_3'(0)}\|p_1\|_C \tag{30}$$

provided that

$$\delta p_3'(0) < 1. \tag{31}$$

Choose

$$p_1(s) = \frac{1}{(s+1)\sqrt{s+1}}, \qquad p_2(s) = \beta\sqrt{s+1},\ \beta = \frac{3}{20(2\sqrt{2}-1)}, \qquad p_3(s) = e^s.$$

Then by (27) we get $\gamma = \frac{2}{3}\beta(2\sqrt{2}-1)$ and $\delta = \beta \ln 2$. Condition (31) becomes

$$\delta p_3'(0) = \beta \ln 2 = .5686421 < 1,$$

which is true. Moreover, by (28) - (30) and the above values we get

$$a = \frac{2\beta(2\sqrt{2}-1)}{3(1-\beta\ln 2)} = .0027389 \qquad \text{and} \qquad b = 1 + \frac{2\beta(2\sqrt{2}-1)}{3(1-\beta\ln 2)} = 1.0027389$$

and $k(r) = \frac{b}{10}e^r$. By (23) and the above values we get

$$\chi(r) = .1106029e^r - 1.1106029r - .0045736.$$

It is simple calculus to show that $\min_{r\in[0,+\infty)} \chi(r) = .1060293 > 0$, χ is increasing on $[2.306712, +\infty)$ and decreasing on $[0, 2.306712]$. Proposition 2 in [10: p. 674] fails to apply since $\chi(r)$ has no zero in the interval $[0, R]$ and $\chi(R) > 0$ for all $R \in [0, +\infty)$. However, our theorem applies.

Indeed, by setting

$$w(r) = \frac{b}{10} r, \tag{32}$$

using (11) and the above values we get $f(r) = \frac{b}{20} r^2 - r + a$. The hypothesis of our theorem is now satisfied if we set

$$r^* = .0027395 \qquad \text{and} \qquad R = 19.942632.$$

Hence, according to our theorem iteration (2) converges to a solution $x^* \in \bar{U}(0, r^*)$ of equation $F(x)(t) = 0$ where F is given by (26). Moreover, x^* is the unique solution of the same equation in $U(0, R)$. Furthermore, estimates (14) and (15) hold in this case for all $n \geq 0$.

Remark 2. Using the choices of the w-functions given in the example we see that $\bar{k}(r) = \frac{1}{10} \leq k(r)$ for $r \in [0, R]$ and by (11) and (23) we get

$$f(r) \leq \chi(r) \qquad \text{for all } r \in [0, R].$$

If $\chi(R) \leq 0$, then $f(R) \leq 0$ also. That means that whenever Proposition 2 in [10: p. 674] applies so does our theorem. The converse is not true as we showed in the above example.

Remark 3. Assume that the k-functions in Remark 2 are constants (Lipschitz) denoted by $\bar{k}$ and k, respectively. The Newton-Kantorovich conditions in this case become

$$\bar{h} = \bar{k} \cdot \eta \leq \frac{1}{2} \qquad \text{and} \qquad h = k \cdot \eta \leq \frac{1}{2}.$$

Since in general $\bar{k} \leq k$ if the second inequality is satisfied so does the first one. However, the converse is not true in general. That means that our conditions are weaker than the corresponding ones in Theorems 1 - 3 and 4 (1.XVIII) of [8].

References

[1] Appell, J., De Pascale, E. and P. P. Zabrejko: *On the application of the Newton-Kantorovich method to nonlinear integral equations of Uryson type.* Numer. Funct. Anal. Optimiz. 12 (1991), 271 – 283.

[2] Appell, J., De Pascale, E., Lysenko, J. V. and P. P. Zabrejko: *New results on Newton-Kantorovich approximations with applications to nonlinear integral equations.* Numer. Funct. Anal. Optimiz. 18 (1997), 1 – 17.

[3] Argyros, I. K.: *A convergence theorem for Newton-like methods under generalized Chen-Yamamoto-type assumptions.* Appl. Math. Comp. 62 (1994), 1 – 15.

[4] Argyros, I. K. and F. Szidarovszky: *On the convergence of modified contractions.* J. Comp. Appl. Math. 55 (1994), 183 – 189.

[5] Argyros, I. K. and F. Szidarovszky: *The Theory and Applications of Iteration Methods.* Boca Raton (Florida, USA): C.R.C. Press Inc. 1993.

[6] Chen, X. T. and Yamamoto: *Convergence domains of certain iterative methods for solving nonlinear equations.* Numer. Funct. Anal. Optimiz. 10 (1989), 37 – 48.

[7] Deuflhard, P. and G. Heindl: *Affine invariant convergence theorems for Newton's method and extensions to related methods.* SIAM J. Numer. Anal. 16 (1979), 1 – 10.

[8] Kantorovich, L. V. and G. P. Akilov: *Functional Analaysis in Normed Spaces.* New York: Pergamon Press 1964.

[9] Szidarovszky, F. and O. Palusinski: *A special matrix equation and its applications in microelectronics.* Appl. Math. Comp. 64 (1994), 115 – 119.

[10] Zabrejko, P. P. and D. F. Nguyen: *The majorant method in the theory of Newton-Kantorovich approximations and the Ptak error estimates.* Numer. Funct. Anal. Optimiz. 9 (1987), 671 – 684.

Received 22.01.1998; in revised form 01.02.1999