Publ. Math. Debrecen
43 / 1–2 (1993), 79–85

On the convergence of inexact Newton-like methods

By IOANNIS K. ARGYROS (Lawton)

Abstract. We provide a general theorem for the convergence of inexact Newton-like methods under Yamamoto-type assumptions. Our results extend and improve several situations already in the literature.

I. Introduction

We consider the inexact Newton-like method

$$x_{n+1} = x_n + y_n,\ A(x_n)y_n = -(F(x_n) + G(x_n)) + r_n \quad n \geq 0 \tag{1}$$

for some $x_0 \in U(x_0, R)$, $R > 0$, to approximate a solution x^* of the equation

$$F(x) + G(x) = 0, \text{ in } \bar{U}(x_0, R). \tag{2}$$

Here $A(x), F, G$ denote operators defined on the closed ball $\bar{U}(x_0, R)$ with center x_0 and radius R, of a Banach space E with values in a Banach space $\hat{E}$, whereas r_n are suitable points in $\hat{E}$. The operator $A(x)(\cdot)$ is linear and approximates the Fréchet derivative of F at $x \in U(x_0, R)$. We will assume that for any $x, y \in \bar{U}(x_0, r) \subseteq \bar{U}(x_0, R)$ with $0 \leq \|x - y\| \leq R - r$,

$$\|F'(x + t(x - y)) - A(x)\| \leq B_1(r, \|x - x_0\| + t\|y - x\|),\ t \in [0, 1] \tag{3}$$

and

$$\|G(x) - G(y)\| \leq B_2(r, \|x - y\|). \tag{4}$$

The functions $B_1(r, r')$ and $B_2(r, r')$ defined on $[0, R] \times [0, R]$ and $[0, R] \times [0, R - r]$ are respectively nonnegative, continuous and nondecreasing functions of two variables. Moreover B_2 is linear in the second variable.

1980 *Mathematics Subject Classification* (1985 *Revision*): 47D15, 47H17, 65R20, 65J10.
Keywords: Banach space, Newton-like method.

Note that the Newton method, the modified Newton method and the secant method are special cases of (1) with $A(x_n) = F'(x_n), A(x_n) = F'(x_0)$ and $A(x_n) = S(x_n, x_{n-1})$ respectively.

If we take

$$w(r') + c, \quad c \in [0,1] \tag{5}$$

and

$$e(r'), \tag{6}$$

where w, e are nonnegative, nondecreasing functions on $[0, R-r]$, to be the right hand sides of (3) and (4) respectively, then we obtain the Zabrejko-Nguen-type assumptions considered by Chen and Yamamoto [2]. They provided sufficient conditions for the convergence of the sequence $\{x_n\}$, $n \geq 0$ generated by (1) to solution x^* of equation (2), when $r_n = 0$, $n \geq 0$.

Moret [5] also studied (1), when $G = 0$ and condition (5) is satisfied. Further work on this subject but for even more special cases than the ones considered by the above authors can be found in [1], [3], [4], [5], [6], [7], [8], [9], [10].

In this paper we will derive a criterion for controlling the residuals r_n in such a way that the convergence of the sequence $\{x_n\}$, $n \geq 0$ to a solution x^* of equation (2) is ensured.

We believe that conditions of the form (3)–(4) are useful not only because we can treat a wider range of problems than before, but it turns out that under natural assumptions we can find better error bounds on the distances $\|x_n - x^*\|$, $n \geq 0$.

II. Convergence Theorems

Throughout the paper the notation $\|\cdot\|$ will stand both for norms in E (or in $\hat{E}$) and also for the induced operator norms $L(E, \hat{E})$, where $L(E, \hat{E})$ denotes the space of bounded linear operators from E to $\hat{E}$.

We will need the following proposition.

Proposition. *Let $a \geq 1$, $\sigma > 0$, $0 \leq \mu < 1$, $0 \leq \rho < R$, $s > 0$ be real constants such that the equation*

$$\varphi(t) := a\sigma \left[\int_0^t B_1(R, \rho + \theta) d\theta + B_2(R, t) \right] - t(1 - \mu) + s = 0 \tag{7}$$

has the solutions in the interval $[0, R)$ and let us denote by t^* the least of them.

Let $v > 0$, $\mu^1 \geq 0$ such that

$$v(1 - \mu) - (1 - \mu^1) \leq 0. \tag{8}$$

Then, for every s^1 satisfying

$$0 < s^1 \le v\left[\sigma\left(\int_0^s B_1(R,\rho+\theta)d\theta + B_2(R,s)\right) + s\mu\right] \tag{9}$$

and for every ρ^1 such that

$$0 \le \rho^1 \le \rho + s, \tag{10}$$

the equation

$$\varphi^1(t) := av\sigma\left[\int_0^t B_1(R,\rho^1+\theta)d\theta + B_2(R,t)\right] - t(1-\mu^1) + s^1 = 0 \tag{11}$$

has nonnegative solutions and at least one of them, denoted by $t^{\star\star}$, lies in the interval $[s^1, t^* - s]$.

Proof. We first observe that since $\varphi(t^\star) = 0$ and $0 \le \mu < 1$, we obtain from (7) that $s \le t^*$. We will show that

$$\varphi^1(t^* - s) \le 0. \tag{12}$$

Using (7)–(11), we obtain

$$\begin{aligned}
&\varphi^1(t^* - s)\\
&= av\sigma\left[\int_0^{t^*-s} B_1(R,\rho^1+\theta)d\theta + B_2(R,t^*-s)\right] - (t^*-s)(1-\mu^1) + s^1\\
&\le v\left[a\sigma\left(\int_s^{t^\star} B_1(R,\rho+\theta)d\theta + B_2(R,t^\star) - B_2(R,s)\right)\right.\\
&\quad \left. + \sigma\left(\int_0^s B_1(R,\rho+\theta)d\theta + B_2(R,s)\right) + s\mu - \frac{(t^*-s)}{v}(1-\mu^1)\right]\\
&\le v\left[a\sigma\left(\int_0^{t^\star} B_1(R,\rho+\theta)d\theta + B_2(R,t^\star)\right) - t^*(1-\mu) + s\right.\\
&\quad \left. + t^*(1-\mu) - s + s\mu - \frac{(t^*-s)}{v}(1-\mu^1)\right]\\
&\le v(t^*-s)\left[(1-\mu) - \frac{(1-\mu^1)}{v}\right] \le 0,
\end{aligned}$$

by (8). Moreover, by (11) it follows immediately that $\varphi^1(s^1) \ge 0$. Hence, by the above inequality and (12) $\varphi^1(t)$ has nonnegative real roots and for the least of them $t^{\star\star}$, it is

$$s^1 \le t^{\star\star} \le t^* - s.$$

Furthermore, from (11) we get $\mu^1 < 1$.

That completes the proof of the proposition.

We can now prove the following result.

Theorem 1. *Let $\{s_n\}$, $\{\mu_n\}$, $\{\sigma_n\}$, $n \geq 0$ be real sequences, with $s_n > 0$, $\mu_n \geq 0$, $\sigma_n > 0$. Let $\{\rho_n\}$ be a sequence on $[0, R)$, with $\rho_0 = 0$ and*

$$\rho_{n+1} \leq \sum_{j=0,1,2,\dots,n} s_j, \qquad n \geq 0. \tag{13}$$

Suppose that $1 - \mu_0 > 0$ and that, for a given constant $a \geq 1$, the function

$$\varphi_0(t) := a\sigma_0 \left[\int_0^t B_1(R, \rho_0 + \theta)d\theta + B_2(R, t)\right] - t(1 - \mu_0) + s_0 \tag{14}$$

has roots on $[0, R)$.

Assume that for every $n \geq 0$ the following conditions are satisfied

$$s_{n+1} \leq v_n \left[\sigma_n \left(\int_0^{s_n} B_1(R, \rho_n + \theta)d\theta + B_2(R, s_n)\right) + s_n\mu_n\right], \tag{15}$$

$$v_n(1 - \mu_n) - (1 - \mu_{n+1}) \leq 0, \tag{16}$$

where $v_n = \dfrac{\sigma_{n+1}}{\sigma_n}$.

Then,

(a) for every $n \geq 0$, the equation

$$\varphi_n(t) := av_n\sigma_n \left[\int_0^t B_1(R, \rho_n + \theta)d\theta + B_2(R, t)\right] - t(1 - \mu_n) + s_n \tag{17}$$

has solutions in $[0, R)$ and, denoting by t_n^* the least of them, we have

$$\sum_{j=n,\dots,\infty} s_j \leq t_n^{\star}. \tag{18}$$

(b) Let $\{x_n\}$, $n \geq 0$ be a sequence in a Banach space such that $\|x_{n+1} - x_n\| \leq s_n$. Then, it converges and denoting its limit by x^*, the error bounds

$$\|x^* - x_n\| \leq t_n^* \tag{19}$$

and

$$\|x^* - x_{n+1}\| \leq t_n^* - s_n \tag{20}$$

are true for all $n \geq 0$.

(c) If there exists $h_0 \in [0, R)$ such that

$$\varphi_0(h_0) \leq 0, \tag{21}$$

then $\varphi_0(t)$ has roots on $[0, R)$.

Proof. (a) We use induction on n. Let us assume that for some $n \geq 0$, $1 - \mu_n > 0$, $\varphi_n(t)$ has roots on $[0, R)$ and t_n^* is the least of them. This is true for $n = 0$. Then, by (13), (15), (16) and the proposition, by setting $s = s_n$, $s^1 = s_{n+1}$, $\mu = \mu_n$, $\mu^1 = \mu_{n+1}$ and $v = v_n$, it follows that t_{n+1}^* exists, with

$$s_{n+1} \leq t_{n+1}^* \leq t_n^* - s_n$$

and $1 - \mu_{n+1} > 0$.

That completes the induction and proves (a).

(b) This part follows easily from part (a).

(c) Using (21), we deduce immediately that $\varphi_0(t)$ has roots on $[0, R)$.

That completes the proof of theorem.

We can now prove the main result.

Theorem 2. *Let* (1) *hold. Assume that for $s_0 > 0$, $\sigma_0 > 0$, $0 \leq \mu_0 < 1$ and $a \geq 1$,* (21) *is true. Then, the function $\varphi_0(t)$ defined by* (14) *has roots on $[0, R)$. Denote by t_0^* the least of them and suppose that*

$$t_0^* < R_0 \leq R. \tag{22}$$

Let $s_n > 0$, $\mu_n \geq 0$, $\sigma_n > 0$, $n \geq 0$ be such that $\liminf \sigma_n > 0$ as $n \to \infty$ and condition (15) *is true for all $n \geq 0$.*

Assume that, for all $n \geq 0$, it is

$$\|y_n\| \leq s_n \leq \sigma_n \|F(x_n) + G(x_n)\| \tag{23}$$

and

$$\|r_n\| \leq \frac{\mu_n s_n}{\sigma_n}. \tag{24}$$

Then the sequence $\{x_n\}$, $n \geq 0$ generated by (1) *remains in $U(x_0, t_0^\star)$ and converges to a solution x^* of equation (2). Moreover, the error bounds* (19) *and* (20) *are true for all $n \geq 0$, where t_n^* is the least root in $[0, R)$ of the function $\varphi_n(t)$ defined by* (17)*, with $\rho_n = \|x_n - x_0\|$, $n \geq 0$.*

Proof. The existence of t_0^* is guaranteed by (21). Let us assume that x_n, $x_{n+1} \in U(x_0, t_0^\star)$. We will show that for every $n \geq 0$, condition (15) is true. Since $\|y_0\| \leq s_0$, this is true for $n = 0$.

Using the identity

$$F(x_{n+1}) + G(x_{n+1}) = \int_0^1 [F'(x_n + t(x_{n+1} - x_n)) - A(x_n)](x_{n+1} - x_n)dt \\ + (G(x_{n+1} - G(x_n)) + r_n,$$

(3), (4), (23), (24), setting $\rho_n = \|x_n - x_0\|$ and by taking norms in the above identity we get

$$s_{n+1} \le \sigma_{n+1}\|F(x_{n+1}) + G(x_{n+1})\| \\ \le v_n \left[\sigma_n \left(\int_0^{s_n} B_1(R, \rho_n + \theta)d\theta + B_2(R, s_n)\right) + s_n\mu_n\right]$$

which shows (15) for all $n \ge 0$.

The hypothesis (b) of Theorem 1 can now easily be verified by induction and thus, by (18) and (23), the sequence $\{x_n\}$, $n \ge 0$ remains in $U(x_0, t_0^\star)$, converges to x^* and (19) and (20) hold. Finally, from the inequality

$$\|F(x_n) + G(x_n)\| \le \|A(x_n) - F'(x_0)\|\|y_n\| + \|F'(x_0)\|\|y_n\| + \|r_n\|,$$

(3), (24) and the continuity of F and G, as $\liminf \sigma_n > 0$ and $s_n \to 0$, as $n \to \infty$ it follows that $F(x^\star) + G(x^\star) = 0$.

That completes the proof of the theorem.

Remark. (a) In the special case when B_1 and B_2 are given (5) and (6) respectively, then our results can be reduced to the ones obtained by Moret [5, p. 359] (when $G = 0$).

(b) Let $G = 0$ and define the functions $\bar{\varphi}_0(t)$, $\bar{\varphi}_n(t)$ by

$$\bar{\varphi}_0(t) = a\sigma_0 \int_0^t (t - \theta)k(\theta)d\theta - t(1 - \mu_0) + s_0,$$
$$\bar{\varphi}_n(t) = av_n\sigma_n \int_0^t (t - \theta)k(\rho_n + \theta)d\theta - t(1 - \mu_n) + s_n,$$

where k is a nondecreasing function on $[0, R]$ such that

$$\|F'(x) - F'(y)\| \le k(r)\|x - y\|, \quad x, y \in \bar{U}(x_0, r) \quad (r < R_0).$$

Assume that B_1 can be chosen in such a way that

$$\varphi_n(t) \le \bar{\varphi}_n(t), \qquad n \ge 0. \tag{25}$$

Then under the hypotheses of Theorem 2 above and Proposition 1 in [5, p. 359], using (25) we can show

$$\|x^* - x_n\| \le t_n^* \le m_n^\star, \qquad n \ge 0$$

and

$$\|x^* - x_{n+1}\| \le t_n^* - s_n \le m_n^* - s_n, \qquad n \ge 0$$

where by m_n^*, we denote the least solutions of the equations

$$\bar{\varphi}_n(t) = 0, \quad n \geq 0 \text{ in } [0, R).$$

References

[1] P.T. Brown, A local convergence theory for inexact-Newton finite-difference projection methods, *SIAM J. Numer. Anal.* **24** (1987), 407–437.

[2] X. Chen and T. Yamamoto, Convergence domains of certain iterative methods for solving nonlinear equations, *Numer. Funct. Anal. and Optimiz.* **10,** (1 and 2) (1989), 37–48.

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

[4] M.A. Krasnosel'skii and Y.B. Rutickii, Some approximate methods of solving nonlinear equations based on linearization, *Soviet Math. Dokl.* **2** (1961), 1542–1546.

[5] I. Moret, A Kantorovich-type theorem for inexact Newton methods, *Numer. Funct. Anal. and Optimiz.* **10,** (3 and 4) (1989), 351–365.

[6] F.A. Potra and V. Ptàk, Nondiscrete induction and iterative processes, *Pitman Advanced Publishing Program, London*, 1984.

[7] A.H. Sherman, On Newton-iterative methods for the solution of systems of equations, *SIAM J. Number. Anal.* **15** (1978), 755–771.

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

[9] T. Yamamoto, A note on a posterior error bound of Zabrejko and Nguen for Zincenko's iteration, *Numer. Funct. Anal. and Optimiz.* **9** (1987), 987–994.

[10] T.J. Ypma, Local convergence of inexact Newton methods, *SIAM J. Numer. Anal.* **21** (1984), 583–590.

IOANNIS K. ARGYROS
DEPARTMENT OF MATHEMATICS
CAMERON UNIVERSITY
LAWTON, OK 73505

(Received September 15, 1991; revised March 30, 1992)