Journal of Computational and Applied Mathematics 228 (2009) 434–443

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

On the semilocal convergence of inexact Newton methods in Banach spaces

Ioannis K. Argyros *

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

ARTICLE INFO

Article history:
Received 19 April 2007
Received in revised form 23 January 2008

MSC:
65H10
65J15
65G99
47H17
49M15

Keywords:
Inexact Newton method
Banach space
Majorizing sequence
Residual
Semilocal convergence
Lipschitz condition
Center-Lipschitz condition

ABSTRACT

We provide two types of semilocal convergence theorems for approximating a solution of an equation in a Banach space setting using an inexact Newton method [I.K. Argyros, Relation between forcing sequences and inexact Newton iterates in Banach spaces, Computing 63 (2) (1999) 134–144; I.K. Argyros, A new convergence theorem for the inexact Newton method based on assumptions involving the second Fréchet-derivative, Comput. Appl. Math. 37 (7) (1999) 109–115; I.K. Argyros, Forcing sequences and inexact Newton iterates in Banach space, Appl. Math. Lett. 13 (1) (2000) 77–80; I.K. Argyros, Local convergence of inexact Newton-like iterative methods and applications, Comput. Math. Appl. 39 (2000) 69–75; I.K. Argyros, Computational Theory of Iterative Methods, in: C.K. Chui, L. Wuytack (Eds.), in: Studies in Computational Mathematics, vol. 15, Elsevier Publ. Co., New York, USA, 2007; X. Guo, On semilocal convergence of inexact Newton methods, J. Comput. Math. 25 (2) (2007) 231–242]. By using more precise majorizing sequences than before [X. Guo, On semilocal convergence of inexact Newton methods, J. Comput. Math. 25 (2) (2007) 231–242; Z.D. Huang, On the convergence of inexact Newton method, J. Zheijiang University, Nat. Sci. Ed. 30 (4) (2003) 393–396; L.V. Kantorovich, G.P. Akilov, Functional Analysis, Pergamon Press, Oxford, 1982; X.H. Wang, Convergence on the iteration of Halley family in weak condition, Chinese Sci. Bull. 42 (7) (1997) 552–555; T.J. Ypma, Local convergence of inexact Newton methods, SIAM J. Numer. Anal. 21 (3) (1984) 583–590], we provide (under the same computational cost) under the same or weaker hypotheses: finer error bounds on the distances involved; an at least as precise information on the location of the solution. Moreover if the splitting method is used, we show that a smaller number of inner/outer iterations can be obtained.

1. Introduction

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

$$F(x) = 0, \tag{1.1}$$

where F is a Fréchet-differentiable operator defined on a convex subset D of a real Banach space X with values in a real Banach space Y.

A large number of problems in applied mathematics and also in engineering is 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 $\dot{x} = Q(x)$ for some suitable operator Q, where x is the state. Then, the equilibrium states are determined

* Tel.: +1 580 5368754; fax: +1 580 5812616.
E-mail address: iargyros@cameron.edu.

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.10.005

by solving Eq. (1.1). 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.

We shall use the iterative procedure

$$x_{n+1} = x_n + s_n, \quad (n \geq 0), \tag{1.2}$$

where step s_n satisfies

$$F'(x_n)s_n = -F(x_n) + r_n \quad (n \geq 0), \tag{1.3}$$

for some null residual sequence $\{r_n\} \subseteq Y$, to generate a sequence $\{x_n\}$ approximating the solution x^*.

A convergence analysis of inexact Newton method (1.2) has been given by many authors and under various assumptions [1–4,7–10,12,14–17]. If $s_n = 0$ $(n \geq 0)$, we obtain the ordinary Newton's method for solving nonlinear equations. Otherwise iterative procedure (1.2) is called inexact Newton's method. By semilocal convergence we mean that we are seeking a solution x^* inside a ball centered at the initial guess x_0, and of a certain finite radius. We recommend the reading of Chapter XVIII on Newton's method of the Kantorovich and Akilov book [15], especially Theorem 6 in Section 1.5, along with the proof, to see how the majorizing function is constructed there (whose least zero plays an important role) (see also relevant Section 4.2 in [7]).

There are two kinds of methods for the solution of linear equations. The first kind of methods are the so-called direct methods, or elimination methods. In this case the exact solution is determined through a finite number of arithmetic operations in real arithmetic without considering the round-off errors. For a list of difficulties and how to handle them we refer the reader to [9].

Another kind of methods are the iterative ones, which result in a two-stage method, or sometimes termed as inner/outer iterations for solving nonlinear equation (1.1).

In this study we are motivated by optimization considerations and the elegant works in [12,14,16]. Guo provided semilocal convergence analysis for inexact Newton method (1.2) using Lipschitz conditions on the Fréchet-derivative F' of operator F. He also provided bounds on the number of inner iteration steps.

We use a combination of Lipschitz and center-Lipschitz conditions along the lines of our works on Newton as well as Newton-like methods [5–7] to provide a new convergence analysis for inexact Newton method (1.2) with advantages over earlier works [1–4,7–10,12,14–17] (especially [12,14–17]) as stated in the abstract of this paper.

2. Type I semilocal convergence analysis of inexact Newton method (1.2)

The main new idea is to introduce a center-Lipschitz condition (with constant γ_0), and then use it instead of the Lipschitz condition (with constant γ) employed in [12] to provide more precise upper bounds on the norms $\|F'(x)^{-1} F'(x_0)\|$ in case $\gamma_0 < \gamma$ (see also the proof of Theorem 2.1, and Remark 2.2 that follow). We can show the main semilocal convergence result for the inexact Newton method (1.2):

Theorem 2.1. *Let $F: D \subseteq X \to Y$ be a Fréchet-differentiable operator. Suppose: $F'(x_0)^{-1} \in L(Y, X)$ for some $x_0 \in D$, and there exist parameters $\beta > 0$, $\gamma_0 \geq 0$, $\gamma \geq 0$, and $\eta \in [0, 1)$ such that for all $x, y \in D$:*

$$\|F'(x_0)^{-1}F(x_0)\| \leq \beta,$$
$$\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq \gamma_0\|x - x_0\|,$$
$$\|F'(x_0)^{-1}[F'(x) - F'(y)]\| \leq \gamma\|x - y\|,$$
$$\frac{\|F'(x_0)^{-1}r_n\|}{\|F'(x_0)^{-1}F(x_n)\|} \leq \eta_n, \quad \eta = \max_n\{\eta_n\},$$
$$\beta\gamma \leq p_0,$$

and

$$U_1 = \overline{U}\left(x_0, \frac{s_1}{1-\sigma}\right) = \left\{x \in X : \|x - x_0\| \leq \frac{s_1}{1-\sigma}\right\} \subseteq D,$$

where,

$$p_0 = -\frac{2\eta^2 + 14\eta + 11 - \sqrt{(4\eta + 5)^3}}{(1 + \eta)(1 - \eta)^2},$$

s_1 is the smallest positive zero of function h given by [7,11,13]*:*

$$h(s) = \gamma s^3 + \beta\gamma(1-\eta)s^2 - 2\beta(1-\beta)s + 2\beta^2(1-\eta^2),$$

and

$$\sigma(s_1) = \sigma = \frac{s_1 - \beta(1-\eta)}{s_1 + \beta(1-\eta)} < 1.$$

Then, sequence $\{x_n\}$ generated by the inexact Newton method (1.2) is well defined, remains in U_1 for all $n \geq 0$, and converges to a solution x^* of equation $F(x) = 0$. Moreover the following estimates hold for all $n \geq 0$:

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n,$$
$$\|x^* - x_n\| \leq t^* - t_n \leq \frac{\sigma^n}{1-\sigma} s_1 \quad (n \geq 0),$$

where, scalar sequence $\{t_n\}(n \geq 0)$ is given by

$$t_0 = 0, \qquad t_{n+1} = t_n + f_n(t_n - t_{n-1}),$$
$$f_n(t) = \frac{1-\sigma}{1-\sigma-\gamma_0 s_1}\left[\frac{\gamma}{2}t^2 + (\sigma+1)\sigma^{n-1}\beta\eta\right],$$

and

$$t^* = \lim_{n\to\infty} t_n \leq \frac{s_1}{1-\sigma}.$$

Proof. It follows exactly as in Theorem 2.6 in [12, p. 237]. However we use the center-Lipschitz condition to obtain the more precise estimate

$$\|F'(x)^{-1}F'(x_0)\| \leq \frac{1}{1-\gamma_0\|x-x_0\|}$$

(using the Banach Lemma on invertible operators on U_1 [7]), instead of

$$\|F'(x)^{-1}F'(x_0)\| \leq \frac{1}{1-\gamma\|x-x_0\|}$$

obtained in [12] using the Lipschitz condition.

That completes the proof of Theorem 2.1. ∎

Remark 2.2. Let us define sequence $\{\bar{t}_n\}$ as $\{t_n\}$ by simply replacing γ_0 by γ in the definition of $\{t_n\}$, and setting $t_0 = \bar{t}_0 = 0$. Clearly

$$\gamma_0 \leq \gamma$$

and $a^{-1} = \frac{\gamma}{\gamma_0}$ can be arbitrarily large [5–7]. If $\gamma_0 = \gamma$, then our Theorem 2.1 reduces to Theorem 2.6 in [12]. Otherwise it constitutes an improvement since

$$t_n < \bar{t}_n \quad (n \geq 2),$$

and

$$t_{n+1} - t_n < \bar{t}_{n+1} - \bar{t}_n \quad (n \geq 1).$$

Let us also define functions f_1 and $\bar{f}_1$ on interval $\left[0, \frac{1}{\gamma}\right)$ by (for $\sigma = \sigma(s)$):

$$f_1(s) = \frac{\gamma}{2}\left[\frac{1-\sigma}{1-\sigma-\gamma_0 s}\right]^2 (1+\eta)^2\beta + \eta - \sigma + s$$

and

$$\bar{f}_1(s) = \frac{\gamma}{2}\left[\frac{1-\sigma}{1-\sigma-\gamma s}\right]^2 (1+\eta)^2\beta + \eta - \sigma + s.$$

Clearly, we have

$$f_1(s) \leq \bar{f}_1(s) \quad s \in \left[0, \frac{1}{\gamma}\right).$$

It was shown in Lemmas 2.1 and 2.5 [12] that $s_1 \in \left(0, \frac{1}{\gamma}\right)$ is the smallest fixed point of function $\overline{f}_1$. It follows that function f_1 has also a fixed point $s^* \in (0, s_1]$, under a weaker hypothesis on $\beta\gamma$ (i.e. a larger p_0), for $\gamma_0 < \gamma$ (since, then: $s^* \in (0, s_1)$).

That is, under the same computational cost we provide a finer majorizing sequence $\{t_n\}$. Moreover, the upper bound on $\beta\gamma$ can be enlarged, and the information on the location of the solution x^* is at least as precise provided that $\gamma_0 < \gamma$, and s^* replaces s_1 in Theorem 2.1.

Note also that the γ-Lipschitz condition implies the γ_0-center-Lipschitz condition.

Other weaker sufficient convergence conditions can be found along the lines of our works in [4–7] using a direct approach on iteration $\{t_n\}$ and the information on (γ_0, γ), not used in [12]. See also Theorem 3.3 that follows.

Application 2.3. For $X = Y = R^j$, let us split matrix $F'(x_n)$ into

$$F'(x_n) = B_n - C_n,$$

to obtain the inner/outer iteration

$$x_{n+1} = x_n - \left[H_n^{m_n-1} + \cdots + H_n + I\right] B_n^{-1} F(x_n)$$
$$H_n = B_n^{-1} C_n \quad (n \geq 0), \tag{2.1}$$

where m_n is the number of inner iterations. We usually let $m_n = m$, or choose any sequence in advance, like $m_n = n + 1$, $n \geq 0$.

A suitable choice is given in the next result:

Corollary 2.4. *Under the hypotheses of Theorem 2.1, further assume:*

$$m_n \geq \frac{\ell n(\ell_n)}{\ell n(\|H_n\|)},$$
$$\ell_n = \frac{1 - \sigma - \gamma_0 s_0}{1 - \sigma + \gamma_0 s_0} \eta_n$$

where, s_0 is s_1 or s^* (defined in Remark 2.2);

$$B_n^{-1} \quad \textit{exists } (n \geq 0),$$

and the spectral radius of matrix H_n is less than 1. Then, the conclusions of Theorem 2.1 hold true. Note that it follows from the hypothesis spr $(H_n) < 1$, $(I - H_n)^{-1}$ exists, and is given by the Von Neumann series, that is, $(I - H_n)^{-1} = I + H_n + H_n^2 + \cdots$.

Proof. It follows from our Theorem 2.1 and Corollary 3.2 in [12]. ■

Remark 2.5. Let $\{\overline{\ell}_n\}\{\overline{m}_n\}$ be defined as sequences $\{\ell_n\}$, $\{m_n\}$ for (s, γ_0) being replaced by (s_1, γ). It then follows that

$$m_n \leq \overline{m}_n,$$

which is an important observation in computational mathematics [4–10,12,14–17].

3. Type II semilocal convergence analysis of inexact Newton method (1.2)

The fourth hypothesis in Theorem 2.1 uses information on the residual points rn and the values $F(x_n)$, whereas the corresponding hypothesis in the related result that follows uses information on the differences $r_n - r_{n-1}$, and $\|x_n - x_{n-1}\|$, respectively. Note also that the rest of the hypotheses in the two theorems are essentially the same. We can show the following semilocal convergence result for inexact Newton method (1.2):

Theorem 3.1. *Let $F: D \subseteq X \to Y$ be a Fréchet-differentiable operator. Suppose: $F'(x_0)^{-1} \in L(Y, X)$ for some $x_0 \in D$, and there exist parameters $\beta \geq 0$, $\gamma_0 \geq 0$, $\gamma \geq 0$, and $\eta \in [0, 1)$ such that for all $x, y \in D$*

$$\|F'(x_0)^{-1}[F(x_0) - r_0]\| \leq \beta,$$
$$\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq \gamma_0 \|x - x_0\|,$$
$$\|F'(x_0)^{-1}[F'(x) - F'(y)]\| \leq \gamma \|x - y\|, \tag{3.1}$$
$$\|F'(x_0)^{-1}(r_n - r_{n-1})\| \leq \eta_n \|x_n - x_{n-1}\|, \quad \eta = \max_n\{\eta_n\},$$
$$\beta\gamma \leq p_1,$$

and

$$U_2 = \overline{U}(x_0, t^*) \subseteq D,$$

where,

$$p_1 = \frac{(1-\eta)^2}{2},$$

and

$$t^* = \frac{1-\eta-\sqrt{(1-\eta)^2-2\beta\gamma}}{\gamma}.$$

Then, sequence $\{x_n\}$ generated by the inexact Newton method (1.2) is well defined, remains in U_2 for all $n \geq 0$, and converges to a solution x^* of equation $F(x) = 0$.

Moreover the following estimates hold for all $n \geq 0$:

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n,$$
$$\|x^* - x_n\| \leq t^* - t_n,$$

where scalar sequence $\{t_n\}$ is defined by

$$t_0 = 0, \qquad t_{n+1} = t_n + \frac{\frac{\gamma}{2}t_n^2 - (1-\eta)t_n + \beta}{1-\gamma_0 t_n}.$$

Proof. It follows exactly as in Theorem 2.7 [12], but uses the more precise upper bound given in our Theorem 2.1.

That completes the proof of Theorem 3.1. ∎

Remark 3.2. Let us define majorizing sequence $\{\bar{t}_n\}$ as $\{t_n\}$ by replacing γ_0 by γ. Then for the advantages of our Theorem 3.2 over Theorem 2.7 in [12], see Remark 2.2.

As already has been shown in our work in [4–7] a more direct approach leads to an even finer majorizing sequence $\{v_n\}$ and a larger bound on $\beta\gamma$.

Therefore, in particular, we can show:

Theorem 3.3. *Under hypotheses (3.1), further assume:*

$$\begin{aligned} &2\eta \leq \delta < 2(1-\beta\gamma_0), \\ &\beta\gamma \leq p_2, \qquad \beta\gamma_0 < 1-\eta, \end{aligned} \tag{3.2}$$

$$U = \overline{U}(x_0, v^{**}) \subseteq D,$$

where,

$$p_2 = \frac{(2-\gamma)(\delta-2\eta)}{2-\delta+4a},$$
$$v^* = \lim_{n\to\infty} v_n \leq \frac{2\beta}{2-\delta} = v^{**},$$

and

$$v_0 = 0, \qquad v_{n+1} = v_n + \frac{1}{1-\gamma_0 v_n}\left[\frac{\gamma}{2}(v_n - v_{n-1}) + \eta\right](v_n - v_{n-1}).$$

Then, sequence $\{x_n\}$ generated by the inexact Newton method (1.2) is well defined, remains in $\overline{U}(x_0, v^*)$ for all $n \geq 0$, and converges to a solution x^* of equation $F(x) = 0$.

Moreover the following estimates hold true for all $n \geq 0$

$$\|x_{n+1} - x_n\| \leq v_{n+1} - v_n \leq \left(\frac{\delta}{2}\right)^n \beta$$

and

$$\|x^* - x_n\| \leq v^* - v_n \leq v^{**} - v_n.$$

Proof. Sequence $\{v_n\}$ is monotonically increasing and bounded above by v^{**}, since by (3.2) for all $k \geq 0$

$$\gamma(t_k - t_{k-1}) + \delta\gamma_0 t_k + 2\eta \leq \gamma \left(\frac{\delta}{2}\right)^{k-1} \beta + 2\gamma_0 \frac{2\beta}{2-\delta} + 2\eta \leq \delta,$$

and

$$\gamma_0 t_k \leq \gamma_0 \frac{1-\left(\frac{\delta}{2}\right)^k}{1-\frac{\delta}{2}} \beta \leq \gamma_0 \frac{2\beta}{2-\delta} < 1$$

(see also Lemma 2.1 in [5], or Lemma 1 in [6]).

Therefore,

$$0 \leq v_{k+1} - v_k \leq \frac{\delta}{2}(v_k - v_{k-1}),$$

and

$$v_k \leq \frac{2\beta}{2-\delta}$$

hold for all $k \geq 0$. Hence, there exists $v^* \in [\beta, v^{**}]$ such that $v^* = \lim_{k\to\infty} v_k$. We shall show

$$\|x_{k+1} - x_k\| \leq v_{k+1} - v_k \quad (k \geq 0).$$

Since, $\|x_k - x_{k-1}\| \leq v_k - v_{k-1}$, we can get in turn

$$\begin{aligned}
\|x_{k+1} - x_k\| &= \|F'(x_k)^{-1}[F(x_k) - r_k]\| \\
&\leq \|F'(x_k)^{-1}F'(x_0)\| \left\{\|F'(x_0)^{-1}[F(x_k) - F(x_{k-1}) - F'(x_{k-1})(x_k - x_{k-1})]\| + \|F'(x_0)^{-1}(r_{k-1} - r_k)\|\right\} \\
&\leq \frac{1}{1-\gamma_0\|x_k - x_0\|}\left[\frac{\gamma}{2}\|x_k - x_{k-1}\|^2 + \eta_k\|x_k - x_{k-1}\|\right] \\
&\leq \frac{1}{1-\gamma_0 v_k}\left[\frac{\gamma}{2}(v_k - v_{k-1}) + \eta\right](v_k - v_{k-1}) = v_{k+1} - v_k. \qquad (3.3)
\end{aligned}$$

That is we showed the first error estimate for all $k \geq 0$.

The second estimate follows from the first by standard majorization techniques [4–7].

In view of the fact that sequence $\{v_n\}$ is Cauchy, it follows that $\{x_n\}$ is a Cauchy sequence too in a Banach space X and as such it converges to some $x^* \in \overline{U}(x_0, v^*)$ (since $\overline{U}(x_0, v^*)$ is a closed set). By letting $k \to \infty$ in (3.3) and since $\lim_{v\to\infty} r_k = 0$, we get $F(x^*) = 0$.

That completes the proof of Theorem 3.3. ∎

Remark 3.4. (a) Let us define related majorizing sequence $\{\overline{v}_n\}$ by simply replacing γ_0 by γ in the definition of sequence $\{v_n\}$. Then again as in Remark 2.2 we have (under the hypotheses of Theorem 2.7 in [12], and our Theorem 3.3):

$$v_n \leq \overline{v}_n \quad (n \geq 2),$$
$$v_n - v_{n-1} \leq \overline{v}_n - \overline{v}_{n-1} \quad (n \geq 2),$$

and

$$v^* - v_n \leq t^* - \overline{v}_n \quad (n \geq 0).$$

Note that strict inequality holds in the first two error estimates provided that $\gamma_0 < \gamma$. If $\gamma_0 = \gamma$ the results in Theorem 3.3 reduce to the ones in Theorem 2.7 in [12].

(b) We are also interested to see, when

$$p_1 < p_2. \qquad (3.4)$$

It is simple algebra to show that (3.4) holds provided that:

$$\delta \in (\delta_1, 2(1 - \beta\gamma_0)),$$

where,

$$\delta_1 = \frac{p_1 + 2\eta + 2 - \sqrt{(p_1 + 2\eta)^2 + 4}}{2}.$$

Note that $\delta_1 \in [0, 1)$. Set

$$\delta^* = \min\{2\eta, \delta_1\}.$$

Then according to Theorem 3.3

$$\delta \in (\delta^*, \delta_0) \quad \text{for } \delta^* \leq \delta_0$$

in order to obtain the results favorable to our comparison.

(c) The estimates on $\beta\gamma$ and δ can be improved even further with some extra labor. Indeed from the proof of Theorem 3.3, we can show instead the weaker estimate:

$$\gamma(t_k - t_{k-1}) + \delta\gamma_0 t_k + 2\eta \leq \gamma\beta\left(\frac{\delta}{2}\right)^{k-1} + 2\delta\gamma_0 \frac{1-\left(\frac{\delta}{2}\right)^k}{2-\delta}\beta + 2\eta \leq \delta$$

or

$$2\beta\gamma\left[1 - \frac{\delta}{2} - 2a\left(\frac{\delta}{2}\right)^2\right]\left(\frac{\delta}{2}\right)^{k-1} + \delta^2 - 2\delta(1+\eta) + 4\eta \leq 0$$

or

$$g(\delta) = c_2\delta^2 + c_2\delta + c_0 \leq 0,$$

where,

$$c_2 = 1 - a\beta\gamma, \qquad c_1 = -[2(1+\eta) + \beta\gamma], \qquad c_0 = 2(\beta\gamma + 2\eta).$$

If

$$g(\delta_2) < 0, \qquad \delta_2 = \min\left\{\delta_0, \frac{-1+\sqrt{1+8a}}{2a}\right\}$$

then function g has a unique zero δ_3 in $(0, \delta_2)$. Therefore, we should choose $\delta \in [\overline{\delta}_3, \delta_2)$ (for $\overline{\delta}_3 < \delta_2$), where $\overline{\delta}_3 = \max\{2\eta, \delta_3\}$. The above condition is weaker than (3.2), since here we replace t_k by $\frac{1-\left(\frac{\delta}{2}\right)^k}{1-\frac{\delta}{2}}$ which is smaller than $\frac{2\beta}{2-\delta}$ used in the proof of Theorem 3.3. However, the above condition is more difficult to verify than (3.2) (see, however, Example 3.5).

If

$$\delta \geq \max\left\{2\eta, \frac{-1+\sqrt{1+8a}}{2a}\right\} = \overline{\delta},$$

then, we should also have

$$\delta^2 - 2\delta(1+\eta) + 4\eta \leq 0.$$

Set

$$\delta_4 = \max\left\{\overline{\delta}, 1+\eta-\sqrt{(1+\eta)^2-4\eta}\right\},$$

$$\delta_5 = \min\left\{\delta_0, 1+\eta+\sqrt{(1+\eta)^2-4\eta}\right\}.$$

If $\delta_4 < \delta_5$, we must choose $\delta \in [\delta_4, \delta_5]$.

(d) Let us assume F is a twice Fréchet-differentiable operator and the Lipschitz conditions in (3.1) are replaced by

$$\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq 1 + h_0'(\|x - x_0\|)$$

and

$$\|F'(x_0)^{-1}F''(x)\| \leq h''(\|x - x_0\|),$$

where for some $d_0 > 0$, $d > 0$ with $d_0 \leq d$, scalar functions h_0 and h are defined on interval $[0, r_0)$, $r_0 = \left(1 - \frac{1}{\sqrt{2}}\right)\frac{1}{d_0}$ by

$$h_0(r) = \beta - t + \frac{\gamma_0 t^2}{1-\gamma_0 t}$$

and

$$h(r) = \beta - (1-\eta)t + \frac{\gamma t^2}{1-\gamma t}.$$

Define also scalar sequence $\{w_n\}$ by

$$w_0 = 0, \qquad w_n = w_{n-1} - \frac{h(w_{n-1})}{h_0'(w_{n-1})} \quad (n \geq 1).$$

Let us also assume

$$\alpha = \beta d \leq 3 - \eta - \sqrt{(3-\eta)^2 - (1+\eta)^2}\,.$$

The conclusions of Theorem 3.1 hold true with the above changes and $\{w_n\}$, w^*, replacing $\{t_n\}$, t^* respectively, where

$$w^* = \frac{1 - \eta + \alpha - \sqrt{(1-\eta+\alpha)^2 - 4\alpha(2-\eta)}}{2d(2-\eta)}\,.$$

Note that if $\eta = 0$, and $d_0 = d$ our results reduce to the ones given in [15]. Moreover, if $\eta \neq 0$, and $d_0 = d$ our results reduce to the ones given in [16]. Otherwise, our results constitute an improvement (under the same computational cost and hypotheses) since our scalar sequences and distances $w_n - w_{n-1}$ are smaller than the corresponding ones in [14, 15] (since $\gamma_0 < \gamma$ and $\frac{\gamma}{\gamma_0}$ can be arbitrarily large [5,7]).

(e) Under the combinations of the center-Lipschitz condition in (3.1), and the d-condition on the second Fréchet-derivative of F given in (d), above, further define scalar sequence $\{z_n\}$ by

$$z_0 = 0, \qquad z_1 = \beta, \qquad z_{n+1} = z_n + \frac{\gamma(z_n - z_{n-1})^2}{(1 - dz_n)(1-\gamma_0 z_n)} + \frac{\eta(z_n - z_{n-1})}{1 - \gamma_0 z_n}\,.$$

It then follows by the proof of Theorems 3.1 and 3.3 that we can arrive at:

$$\|x_n - x_{n-1}\| \leq z_n - z_{n-1}.$$

Set $d_1 = \max\{\gamma_0, d\}$, and $d_2 = d_2(\delta) = \frac{2\beta}{2-\delta}$. Assume there exists $\delta \in [0, 2)$ such that

$$d_1 \frac{2\beta}{2-\delta} < 1,$$

and

$$\frac{\gamma\beta}{(1 - dd_2)(1-\gamma_0 d_2)} + \frac{\eta}{1-\gamma_0 d_2} \leq \frac{\delta}{2}\,.$$

Then, the induction hypotheses of Theorem 3.3 go through to show

$$dz_n < 1, \qquad \gamma_0 z_n < 1 \quad \text{and} \quad 0 \leq z_{n+1} - z_n \leq \frac{\delta}{2}(z_n - z_{n-1}).$$

Then, the conclusions of Theorem 3.3 hold true with the above changes and setting.

We present an example to justify the elaborate arguments in Remark 3.4:

Example 3.5. Let $X = Y = \mathbf{R}$, $D = [q, 2-q]$, $q \in [0, \frac{1}{2})$, $x_0 = 1$, and define function F on D by

$$F(x) = x^3 - q.$$

It can then be easily seen that $\beta = \frac{1}{3}(1-q)$, $\gamma_0 = 3 - q$, $\gamma = 2(2-p)$, $d_0 = d = 1$. Set $q = .47$. Then no matter how we choose η in [12,14] or [15] (for $\eta = 0$) convergence conditions are violated since

$$\beta\gamma > \frac{1}{2}(1-\eta)^2$$

and

$$\alpha > 3 - \eta - \sqrt{(3-\eta)^2 - (1+\eta)^2}\,.$$

However, the conditions in Remark 3.4(c) hold true. Indeed we have, say for $\eta = .01$,

$$a = .826997386, \qquad c_0 = 1.1212, \qquad c_1 = -2.5606,$$
$$c_2 = .5530\overline{3}, \qquad \delta_0 = .9188 = \delta_2, \qquad \delta_3 = \overline{\delta}_3 = .491928775.$$

Therefore, we should choose $\delta \in (\overline{\delta}_3, \delta_2)$. Then the conclusions of Theorem 3.3 hold true for $\delta \in (\overline{\delta}_3, \delta_2)$. A suitable choice is $\delta = \frac{1}{2}$.

The convergence conditions in Remark 3.4(e) also hold true, say for $\delta = \frac{1}{2}$. Indeed, we have

$$\beta = .17\overline{6}, \qquad d_1 = \gamma_0 = 2.53, \qquad d = 1, \qquad \gamma = 3.06,$$

and

$$\frac{2\beta}{2-\beta} = .235555556.$$

The first condition becomes

$$.595955556 < 1,$$

and the second

$$.19977511 < .25 = \frac{\delta}{2}.$$

Application 3.6. Let us assume $m_n = m$ in iteration (2.1). We can obtain the following parallel with Corollary 2.4 results concerning the estimation of the number of inner iterations under the conditions of Theorem 3.1 (or Theorem 3.3):

Theorem 3.7. *Under the hypotheses of Theorem 3.1 (or Theorem 3.3) further assume:*

$$\|B_0^{-1}F'(x_0)\| \le q,$$
$$a_0 h^m + mbh^{m-1} \le \eta_n, \qquad \sup_n \|H_n\| \le h < 1,$$

where,

$$a_0 = \frac{3 - 2\eta + 2\beta\gamma^n}{\eta^2},$$
$$b = \frac{2-\eta}{\eta} \frac{\delta(\delta+1)\gamma_0}{[1-(1-\eta)\gamma_0\delta]^2}\left[\frac{(1-\eta)^2}{2\gamma} + \frac{1-\eta}{\gamma} + \beta\right] \tag{3.5}$$

and the matrix norm has the property:

$$\|F'(x_0)^{-1}E\| \le \|F'(x_0)^{-1}D\|$$

with E any submatrix of D;

$$\beta\gamma \le p_1 \frac{1}{1+h^m}, \quad \left(\text{or } \beta\gamma \le p_2 \frac{1}{1+h^m}\right)$$

and

$$\overline{U}(x_0, t^*) \subseteq D, \quad (\text{or } \overline{U}(x_0, v^*) \subseteq D).$$

Then the conclusions of Theorem 3.1 hold true.

Proof. It follows exactly as in Corollary 3.3 in [12], and our Theorem 3.1. Here are the changes (with γ_0 replacing γ in the proof):

$$\|F'(x_0)^{-1}F'(x_n)\| \le 1 + \gamma_0\|x_n - x_0\|,$$
$$\|F'(x_n)^{-1}F'(x_0)\| \le \frac{1}{1-\gamma_0\|x_n - x_0\|},$$
$$\|F'(x_0)^{-1}F(x_n)\| \le \frac{\gamma}{2}\|x_n - x_0\|^2 + \|x_n - x_0\| + \beta,$$
$$\|F'(x_0)^{-1}(B_n - B_{n-1})\| \le \gamma\|x_n - x_{n-1}\|,$$

and

$$\|B_n^{-1}F'(x_0)^{-1}\| \le \frac{q}{1-\gamma_0\|x_n - x_0\|q}.$$

The constant $\overline{b}$ defined in [12] (for $\gamma_0 = \gamma$) is larger than b, which is the advantage of our approach for the selection of a smaller η. ■

References

[1] I.K. Argyros, Relation between forcing sequences and inexact Newton iterates in Banach spaces, Computing 63 (2) (1999) 134–144.
[2] I.K. Argyros, A new convergence theorem for the inexact Newton method based on assumptions involving the second Fréchet-derivative, Comput. Appl. Math. 37 (7) (1999) 109–115.
[3] I.K. Argyros, Forcing sequences and inexact Newton iterates in Banach space, Appl. Math. Lett. 13 (1) (2000) 77–80.
[4] I.K. Argyros, Local convergence of inexact Newton-like iterative methods and applications, Comput. Math. Appl. 39 (2000) 69–75.
[5] I.K. Argyros, On the Newton–Kantorovich hypothesis for solving equations, J. Comput. Appl. Math. 169 (2004) 315–332.
[6] I.K. Argyros, A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space, J. Math. Anal. Appl. 298 (2004) 374–397.

[7] I.K. Argyros, Computational Theory of Iterative Methods, in: C.K. Chui, L. Wuytack (Eds.), in: Studies in Computational Mathematics, vol. 15, Elsevier Publ. Co., New York, USA, 2007.
[8] Z.Z. Bai, A class of two-stage iterative methods for systems of weakly nonlinear equations, Numer. Algorithms 14 (4) (1997) 295–319.
[9] Z.Z. Bai, D.R. Wang, On the convergence of the factorization update algorithm.
[10] R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal. 19 (2) (1982) 400–408.
[11] J.M. Gutierrez, A new semilocal convergence theorem for Newton's method, J. Comput. Appl. Math. 79 (1997) 131–145.
[12] X. Guo, On semilocal convergence of inexact Newton methods, J. Comput. Math. 25 (2) (2007) 231–242.
[13] Z.D. Huang, A note on the Kantorovich theorem for Newton iteration, J. Comput. Appl. Math. 49 (1993) 211–217.
[14] Z.D. Huang, On the convergence of inexact Newton method, J. Zheijiang University, Nat. Sci. Ed. 30 (4) (2003) 393–396.
[15] L.V. Kantorovich, G.P. Akilov, Functional Analysis, Pergamon Press, Oxford, 1982.
[16] X.H. Wang, Convergence on the iteration of Halley family in weak condition, Chinese Sci. Bull. 42 (7) (1997) 552–555.
[17] T.J. Ypma, Local convergence of inexact Newton methods, SIAM J. Numer. Anal. 21 (3) (1984) 583–590.