Ioannis K. Argyros (Lawton, OK)

A CONVERGENCE ANALYSIS OF NEWTON'S METHOD UNDER THE GAMMA-CONDITION IN BANACH SPACES

Abstract. We provide a local as well as a semilocal convergence analysis for Newton's method to approximate a locally unique solution of an equation in a Banach space setting. Using a combination of center-gamma with a gamma-condition, we obtain an upper bound on the inverses of the operators involved which can be more precise than those given in the elegant works by Smale, Wang, and Zhao and Wang. This observation leads (under the same or less computational cost) to a convergence analysis with the following advantages: local case: larger radius of convergence and finer error estimates on the distances involved; semilocal case: larger domain of convergence, finer error bounds on the distances involved, and at least as precise information on the location of the solution.

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

$$(1.1) \qquad\qquad F(x) = 0,$$

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

A large number of problems in applied mathematics and also in engineering are solved by finding solutions of certain equations. For example, dynamical 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 operator Q, where x is the state. Then the equilibrium states are determined by solving an equation of type (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: 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.

Newton's method

$$(1.2) \qquad x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) \quad (x_0 \in D,\ n \geq 0)$$

is undoubtedly the most popular method for generating a sequence $\{x_n\}$ approximating x^*. Here for $x \in D$, $F'(x) \in L(X, Y)$ (the space of bounded linear operators from X into Y) denotes the Fréchet derivative of the operator F [3], [4]. There is an extensive literature on local as well as semilocal convergence results for Newton's method (1.2) under various conditions. A survey of such results can be found in [3].

Below we summarize the most important results:

Semilocal case: Kantorovich [4] using information from the domain D and the Lipschitz condition

$$(1.3) \qquad \|F'(x_0)^{-1}[F'(x) - F'(y)]\| \leq \ell \|x - y\| \quad \text{for all } x, y \in D,$$

provided the sufficient convergence condition

$$(1.4) \qquad K = \ell\beta \leq \frac{1}{2}, \quad \|F'(x_0)^{-1} F(x_0)\| \leq \beta.$$

Using a combination of (1.3) and the center-Lipschitz condition

$$(1.5) \qquad \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq \ell_0 \|x - x_0\| \quad \text{for all } x \in D$$

we weakened (1.4) by replacing it with

$$(1.6) \qquad A = \bar{\ell}\beta \leq \frac{1}{2}, \quad \bar{\ell} = \frac{1}{8}\left(\ell + 4\ell_0 + \sqrt{\ell^2 + 8\ell_0\ell}\right)$$

[1, p. 387, Case 3 for $\delta = \delta_0$], [3].

The estimate

$$(1.7) \qquad \ell_0 \leq \ell$$

holds in general, and ℓ/ℓ_0 can be arbitrarily large [1]–[3].

By comparing condition (1.4) with (1.6) we deduce that

$$(1.8) \qquad K \leq 1/2 \ \Rightarrow \ A \leq 1/2$$

but not vice versa, unless $\ell_0 = \ell$.

The above improvements are made, since we use $\frac{1}{1-\ell_0\|x-x_0\|}$ instead of $\frac{1}{1-\ell\|x-x_0\|}$ as an upper bound for $\|F'(x)^{-1}F'(x_0)\|$ in the convergence analysis of Newton's method.

Local case: Rheinboldt [5] used the Lipschitz condition

$$(1.9) \qquad \|F'(x^*)^{-1}[F'(x) - F'(y)]\| \leq L\|x - y\| \qquad \text{for all } x, y \in D$$

to provide the radius of convergence r_R for Newton's method given by

$$(1.10) \qquad\qquad r_R = \frac{2}{3L}.$$

Using a combination of (1.9) and the center-Lipschitz condition

$$(1.11) \qquad \|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq L_0\|x - x^*\| \qquad \text{for all } x \in D$$

we found [1] that the corresponding radius of convergence is given by

$$(1.12) \qquad\qquad r_A = \frac{2}{2L_0 + L}.$$

Note that again

$$(1.13) \qquad\qquad L_0 \leq L.$$

By comparing (1.10) with (1.12), we see that

$$(1.14) \qquad\qquad r_R \leq r_A.$$

Moreover, if $L_0 < L$, then $r_R < r_A$.

In turns out that the advantages of this approach persist in the case when Lipschitz conditions are replaced by gamma-conditions. These observations together with optimization considerations motivated us to write this paper.

Wang [7] motivated by Smale [6] introduced the gamma-conditions (see also [2]): in the semilocal case,

$$(1.15) \qquad \|F'(x_0)^{-1}F''(x)\| \leq \frac{2\gamma}{(1 - \gamma\|x - x_0\|)^3} \qquad \text{for some } \gamma > 0,$$
$$\text{and all } x \in U(x_0, 1/\gamma) \subseteq D,$$

$$(1.16) \qquad\qquad \alpha = \beta\gamma \leq 3 - 2\sqrt{2}$$

to provide a convergence analysis for Newton's method (1.2).

Moreover Wang [7], Zhao and Wang [8] used the gamma-condition in the local case,

$$(1.17) \qquad \|F'(x^*)^{-1}F''(x)\| \leq \frac{2\gamma}{(1 - \gamma\|x - x^*\|)^3}$$

for all $x \in U(x^*, s) \subseteq D$, for some $s, \gamma > 0$, where

$$(1.18) \qquad\qquad s_1 = \frac{5 - \sqrt{13}}{6\gamma} \leq s$$

is the convergence radius of Newton's method.

2. Local convergence analysis for Newton's method. It is convenient to introduce a polynomial whose roots play a role in the proof of Theorem 2.2 below.

Let $a \in [0, 1]$ be a given parameter. Set $b = 1 - a$, and define a polynomial p_a by

(2.1) $$p_a(t) = 3a^2 t^3 + a(6b - a)t^2 + (3b^2 - 2ab - 1)t - b^2.$$

By the definition, we get

(2.2) $$p_a(0) = -b^2 \leq 0, \quad p_a(1) = 1.$$

It then follows by the intermediate value theorem that there exists a root of p_a in $[0, 1)$. Let us denote by t_a the minimal number in $[0, 1)$ satisfying $p_a(t_a) = 0$. Let $\gamma > 0$. Define

(2.3) $$s_a = \frac{1 - t_a}{\gamma}.$$

In particular for $a = 1$, $t_1 = (1 + \sqrt{13})/6$, and we get (see also (1.18))

$$s_1 = \frac{5 - \sqrt{13}}{6\gamma}.$$

It is simple algebra to show that for all $a \in [0, 1]$,

(2.4) $$p_a(t_1) \geq 0,$$

which implies

(2.5) $$t_a \leq t_1,$$

and

(2.6) $$s_1 \leq s_a \quad \text{for all } a \in [0, 1].$$

Note also that strict inequality holds in (2.6) for $a \neq 1$.

We introduce some conditions which seem to be the minimum-type hypotheses needed for the study of the local convergence of Newton's method (1.2).

DEFINITION 2.1. Let $F : D \subseteq X \to Y$ be a Fréchet-differentiable operator. Assume that there exists $x^* \in D$ such that

$$F(x^*) = 0 \quad \text{and} \quad F'(x^*)^{-1} \in L(Y, X);$$

and there exist $a \in [0, 1]$ and $\gamma_0, \gamma > 0$ with $\gamma_0 = a\gamma$ such that for all $x \in U(x^*, s) = \{x \in X : \|x - x^*\| < s\}$ where

(2.7) $$s = \min\left\{\frac{1}{\gamma}, \left(1 - \frac{1}{\sqrt{2}}\right)\frac{1}{\gamma_0}\right\}$$

we have

$$(2.8) \qquad \|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq \frac{1}{(1 - \gamma_0\|x - x^*\|)^2} - 1,$$

$$(2.9) \qquad \|F(x) - F(x^*) - F'(x)(x - x^*)\| \leq \frac{\gamma\|x - x^*\|^2}{1 - \gamma\|x - x^*\|}.$$

We can show the following local convergence theorem for Newton's method (1.2).

THEOREM 2.2. *Let $F : D \subseteq X \to Y$ be a Fréchet-differentiable operator. Assume that the conditions of Definition 2.1 hold, and*

$$(2.10) \qquad \overline{U}(x^*, s_a) \subset U(x^*, s) \subseteq D,$$

where s_a is given by (2.3). Then the sequence $\{x_n\}$ generated by Newton's method (1.2) is well defined, remains in $U(x^*, s_a)$ for all $n \geq 0$, and converges to x^* provided that $x_0 \in U(x^*, s_a)$. Moreover, the following error estimates hold for all $n \geq 0$:

$$(2.11) \qquad \|x_{n+1} - x^*\| \leq \frac{\gamma(1 - \gamma_0\|x_n - x^*\|)^2\|x_n - x^*\|^2}{[2(1 - \gamma_0\|x_n - x^*\|)^2 - 1](1 - \gamma\|x_n - x^*\|)}.$$

Proof. We shall first show that

$$F'(x)^{-1} \in L(Y, X) \quad \text{for all } x \in U(x^*, s_a).$$

Indeed, in view of (2.8) and (2.10) we obtain in turn

$$(2.12) \qquad \|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq \frac{1}{(1 - \gamma_0\|x - x^*\|)^2} - 1$$

$$< \frac{1}{(1 - \gamma_0 s)^2 - 1} \leq 1.$$

It follows from (1.12) and the Banach lemma on invertible operators [3], [4] that $F'(x)^{-1} \in L(Y, X)$, and

$$(2.13) \qquad \|F'(x)^{-1}F'(x^*)\| \leq \frac{(1 - \gamma_0\|x - x^*\|)^2}{2(1 - \gamma_0\|x - x^*\|)^2 - 1}.$$

By hypothesis $x_0 \in U(x^*, s_a)$. Assume $x_k \in U(x^*, s_a)$, $k = 0, 1, \ldots, n$. We shall show $x_{k+1} \in U(x^*, s_a)$, and estimate (2.11) holds for all $n \geq 0$.

Using (2.9) for $x = x_k$, we get

$$(2.14) \quad \|F'(x^*)^{-1}[F(x_k) - F(x^*) - F'(x_k)(x_k - x^*)]\| \leq \frac{\gamma\|x_k - x^*\|^2}{1 - \gamma\|x_k - x^*\|}.$$

By (1.2), (2.13) (for $x = x_k$), and (2.14) we get in turn

$$(2.15) \quad \|x_{k+1} - x^*\|$$

$$\leq \|F'(x_k)^{-1}F'(x^*)\| \cdot \|F'(x^*)^{-1}[F(x_k) - F(x^*) - F'(x_k)(x_k - x^*)]\|$$

$$\leq \frac{\gamma(1 - \gamma_0\|x_k - x^*\|)^2\|x_k - x^*\|^2}{[2 - (1 - \gamma_0\|x_k - x^*\|)^2 - 1](1 - \gamma\|x_k - x^*\|)}$$

$$\leq \frac{\gamma\|x_k - x^*\|(1 - a\gamma\|x_k - x^*\|)^2\|x_k - x^*\|}{[2(1 - a\gamma\|x_k - x^*\|)^2 - 1](1 - \gamma\|x_k - x^*\|)} < \|x_k - x^*\| < s_a.$$

By definition of the polynomial p_a, and the choice of s_a, which imply $x_{k+1} \in U(x^*, s_a)$, estimate (2.11) holds for $n = k$, and $\lim_{k \to \infty} x_k = x^*$.

   That completes the proof of the Theorem. ∎

   We show that conditions (2.8) and (2.9) can be realized:

   REMARK 2.3. (a) Let us assume that the operator F is twice Fréchet-differentiable, $\gamma_0 = \gamma$, and condition (1.17) holds. It then follows that estimates (2.8) and (2.9) hold true. Indeed, we get

$$(2.16) \quad \|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| = \left\|\int_0^1 F''(x^* + t(x - x^*))(x - x^*)\,dt\right\|$$

$$\leq \int_0^1 \frac{2\gamma\|x - x^*\|\,dt}{(1 - \gamma t\|x - x^*\|)^3} \leq \frac{1}{(1 - \gamma\|x - x^*\|)^2} - 1,$$

and

$$(2.17) \quad \|F'(x^*)^{-1}[F(x) - F(x^*) - F'(x)(x - x^*)]\|$$

$$= \left\|F'(x^*)^{-1}\int_0^1 F''[x^* + t(x - x^*)](1 - t)(x - x^*)^2\,dt\right\|$$

$$\leq \int_0^1 \frac{2\gamma(1 - t)\|x - x^*\|^2 dt}{(1 - \gamma\|x - x^*\|)^3} \leq \frac{\gamma\|x - x^*\|^2}{1 - \gamma\|x - x^*\|}.$$

Note in this case the radius of convergence is s_1 given by (2.3).
   (b) If

$$(2.18) \quad\quad\quad\quad\quad\quad\quad\quad \gamma_0 < \gamma,$$

Theorem 2.2 improves the corresponding theorem by Wang [7], who uses (1.17) instead of weaker conditions (2.8) and (2.9). Indeed, (2.6) holds and our estimates on $\|x_n - x^*\|$ are finer. To avoid repetitions we refer the reader to the remarks and discussions of the next section for possible choices of γ_0 and γ (simply replace x_0 by x^*).

(c) As an example showing that (2.6) can indeed be a strict inequality, let $a = 1/2$. Using (2.1) and (2.3), we obtain

$$t_{1/2} = .65185 < t_1 = \frac{1 + \sqrt{13}}{6} = .76759,$$

and

$$s_1 = \frac{.23241}{\gamma} < \frac{.34815}{\gamma} = s_{1/2}.$$

3. Semilocal convergence analysis of Newton's method. Let $\beta, \gamma > 0$. Define a function f on the interval $[0, 1/\gamma)$ by

$$(3.1) \qquad f(w) = \beta - w + \frac{\gamma w^2}{1 - \gamma w}.$$

Wang [7] showed that if condition (1.16) holds, then the function f has two positive roots w^*, w^{**} satisfying

$$(3.2) \qquad \beta \leq w^* \leq \left(1 + \frac{1}{\sqrt{2}}\right)\beta \leq \left(1 - \frac{1}{\sqrt{2}}\right)\frac{1}{\gamma} \leq w^{**} \leq \frac{1}{2\gamma}.$$

Moreover, the scalar sequence $\{w_n\}$ given by

$$(3.3) \qquad w_0 = 0, \quad w_{n+1} = w_n - \frac{f(w_n)}{f'(w_n)}$$

is nondecreasing and converges to w^*. Furthermore, Wang showed the following semilocal convergence result for Newton's method (1.2):

THEOREM 3.1 ([7]). *Let $F : D \subseteq X \to Y$ be a twice Fréchet-differentiable operator. Let $x_0 \in D$ be such that $F'(x_0)^{-1} \in L(Y, X)$. Assume that there exist $\beta, \gamma > 0$ such that for all $x \in U(x_0, 1/\gamma) \subseteq D$ the gamma-conditions (1.15) and (1.16) hold. Then the sequence $\{x_n\}$ generated by Newton's method (1.2) is well defined, remains in $\overline{U}(x_0, w^*)$ for all $n \geq 0$, and converges to a solution x^* of the equation $F(x) = 0$ in $\overline{U}(x_0, w^*)$, which is unique in $U(x_0, w^{**})$. Moreover, the following error estimates hold for all $n \geq 0$:*

$$(3.4) \qquad \|x_{n+1} - x_n\| \leq w_{n+1} - w_n$$

and

$$(3.5) \qquad \|x_n - x^*\| \leq w^* - w_n,$$

where the sequence $\{w_n\}$ is given by (3.3) and

$$(3.6) \qquad w^* = \lim_{n \to \infty} w_n.$$

In the rest of the section we show how to improve these results.

CASE 1. We use hypothesis (1.16), but instead of (1.15) we use the weaker conditions (3.10) and (3.11) given in Definition 3.2 below.

Let $\gamma_0 > 0$ with

(3.7) $$\gamma_0 \leq \gamma.$$

Define a function f_0 on $[0, 1/\gamma_0)$ by

(3.8) $$f_0(w) = \beta - w + \frac{\gamma_0 w^2}{1 - \gamma_0 w},$$

and a scalar sequence $\{t_n\}$ by

(3.9) $$t_0 = 0, \quad t_{n+1} = t_n - \frac{f(t_n)}{f_0'(t_n)}.$$

We need the definition of certain gamma-type conditions:

DEFINITION 3.2. Let $F : D \subseteq X \to Y$ be a Fréchet-differentiable operator. Let $x_0 \in D$ be such that $F'(x_0)^{-1} \in L(Y, X)$. Assume that there exist $\beta, \gamma_0, \gamma > 0$ with $\gamma_0 \leq \gamma$ such that for all $x, y \in U(x_0, \gamma_1) \subseteq D$,

(3.10) $$\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq \frac{1}{(1 - \gamma_0 \|x - x_0\|)^2} - 1$$

and

(3.11) $$\|F'(x_0)^{-1}[F(y) - F(x)] - F'(x)(y - x)\|$$
$$\leq \frac{\gamma \|y - x\|^2}{[1 - \gamma \|x - x_0\|]^2 (1 - \gamma \|y - x_0\|)}.$$

Then we can show the following semilocal convergence theorem for Newton's method (1.2).

THEOREM 3.3. *Assume the conditions of Definition 3.2 hold, (1.16), and* $U(x_0, \gamma_1) \subseteq D$, *where*

(3.12) $$\gamma_1 = \min\left\{\frac{1}{\gamma}, \left(1 - \frac{1}{\sqrt{2}}\right)\frac{1}{\gamma_0}\right\}.$$

Then the sequence $\{x_n\}$ *generated by Newton's method (1.2) is well defined, remains in* $\overline{U}(x_0, t^*)$ *for all* $n \geq 0$, *and converges to a solution* x^* *of the equation* $F(x) = 0$ *in* $\overline{U}(x_0, t^*)$ *which is unique in* $U(x_0, w^{**})$ *where* $t^* = \lim_{n \to \infty} t_n$. *Moreover, the following error estimates hold for all* $n \geq 0$:

(3.13) $$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n \leq w_{n+1} - w_n,$$
(3.14) $$\|x_n - x^*\| \leq t^* - t_n \leq w^* - w_n.$$

Proof. We shall first show that

(3.15) $$0 \leq t_k \leq w_k,$$
(3.16) $$0 \leq t_{k+1} - t_k \leq w_{k+1} - w_k,$$
(3.17) $$0 \leq t^* - t_k \leq w^* - w_k.$$

If we show that (3.15) and (3.16) hold for all $n \geq 0$, then (3.17) follows by using standard majorization techniques [3], [4]. Estimates (3.15) and (3.16) hold true for $k = 0$ by the initial conditions. Assume they hold for all $n \leq k$. The functions f, f_0 are nonincreasing, whereas f', f_0' are nondecreasing. It is therefore simple algebra to show under the induction hypothesis that (3.15) and (3.16) hold for k replaced by $k + 1$. The sequence $\{t_n\}$ is nondecreasing, and bounded above by w^*, and as such it converges to a $t^* \in [\beta, w^*]$.

We shall also show by induction on $k \geq 1$ that

$$(3.18) \qquad \|x_k - x_{k-1}\| \leq t_k - t_{k-1}$$

and

$$(3.19) \qquad \overline{U}(x_k, t^* - t_k) \subseteq \overline{U}(x_{k-1}, t^* - t_{k-1}).$$

For every $z \in \overline{U}(x_1, t^* - t_1)$,

$$\|z - x_0\| \leq \|z - x_1\| \, \|x_1 - x_0\| \leq t^* - t_1 + t_1 = t^* = t^* - t_0,$$

so $z \in \overline{U}(x_0, t^* - t_0)$. We also have

$$\|x_1 - x_0\| = \|F'(x_0)^{-1}F(x_0)\| \leq \beta = t_1 = t_1 - t_0.$$

That is, (3.18) and (3.19) hold for $k = 0$. Assuming they hold for $n = 0, 1, \ldots, k$, we have

$$\|x_k - x_0\| \leq \sum_{i=1}^{k} \|x_i - x_{i-1}\| \leq \sum_{i=1}^{k} (t_i - t_{i-1}) = t_k - t_0 = t_k \leq t^*.$$

It follows from (3.14), and the induction hypotheses that

$$(3.20) \qquad \|F'(x_0)^{-1}[F'(x_k) - F'(x_0)]\| \leq \frac{1}{(1 - \gamma_0\|x_k - x_0\|)^2} - 1$$

$$\leq \frac{1}{(1 - \gamma_0 t_k)^2} - 1 < \frac{1}{(1 - \gamma_0 t^*)^2} - 1 \leq 1,$$

by the choice of t^*.

In view of (3.20), and the Banach lemma on invertible operators [3], [4] we see that $F'(x_k)^{-1}$ exists, and

$$(3.21) \qquad \|F'(x_k)^{-1}F'(x_0)\| \leq \frac{1}{1 - \left[\frac{1}{(1 - \gamma_0\|x_k - x_0\|)^2} - 1\right]}$$

$$\leq -\frac{1}{f_0'(\|x_k - x_0\|)} \leq \frac{(1 - \gamma_0 t_k)^2}{2(1 - \gamma_0 t_k)^2 - 1}$$

$$= -\frac{1}{2f_0(t_k)}.$$

Using the approximation

$$(3.22) \qquad F(x_k) = F(x_k) - F(x_{k-1}) - F'(x_{k-1})(x_k - x_{k-1})$$

and (3.11), we get

$$(3.23) \quad \|F'(x_0)^{-1}F(x_k)\|$$
$$= \|F'(x_0)^{-1}[F(x_k) - F(x_{k-1}) - F'(x_{k-1})(x_k - x_{k-1})]\|$$
$$\leq \frac{\gamma\|x_k - x_{k-1}\|^2}{(1 - \gamma\|x_{k-1} - x_0\|)^2(1 - \gamma\|x_k - x_0\|)}$$
$$\leq \frac{\gamma(t_k - t_{k-1})^2}{(1 - \gamma t_{k-1})^2(1 - \gamma t_k)} = f(t_k).$$

Hence, by (1.2), (3.21), and (3.22) we obtain

$$(3.24) \qquad \|x_{k+1} - x_k\| = \|[F'(x_k)^{-1}F'(x_0)][F'(x_0)^{-1}F(x_k)]\|$$
$$\leq \|F'(x_k)^{-1}F'(x_0)\| \, \|F'(x_0)^{-1}F(x_k)\|$$
$$\leq -\frac{f(t_k)}{f_0'(t_k)} = t_{k+1} - t_k,$$

which shows (3.18) for all k.

Moreover, for every $z \in \overline{U}(x_{k+1}, t^* - t_{k+1})$, we have

$$\|z - x_k\| \leq \|z - x_{k+1}\| + \|x_{k+1} - x_k\| \leq t^* - t_{k+1} + t_{k+1} - t_k = t^* - t_k,$$

which implies

$$(3.25) \qquad\qquad z \in \overline{U}(x_k, t^* - t_k).$$

The sequence $\{t_n\}$ is Cauchy. It follows from (3.18) and (3.19) that $\{x_n\}$ is a Cauchy sequence too, and as such it converges to some $x^* \in \overline{U}(x_0, t^*)$ (since $\overline{U}(x_0, t^*)$ is a closed set). By letting $k \to \infty$ in (3.24) we obtain $F(x^*) = 0$. Estimate (3.14) follows (3.13) by using standard majorization techniques [3], [4]. Finally, the uniqueness of solution x^* in $U(x_0, w^{**})$ has already been shown in [7], where Wang showed $\lim_{n \to \infty} \|x_n - y^*\| = 0$ for any y^* with $F(y^*) = 0$ and $y^* \in U(x_0, w^{**})$.

That completes the proof of the theorem. ∎

REMARK 3.4. (a) In Theorem 3.3 we assumed that the operator F is only once differentiable, and conditions (3.10), (3.11) are weaker than (1.15). Under these conditions we provided finer error estimates on the distances, and at least as precise information on the location of the solution x^*. Note also that: if strict inequality holds in (3.7), it also does in (3.15) and (3.16) (right hand sides only for $k > 1$). In the special case when F is twice Fréchet-differentiable, and (1.15) holds, then (3.10), and (3.11) also hold true for $\gamma_0 = \gamma$.

Indeed, we have

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

$$= \left\| F'(x_0)^{-1} \int_0^1 F''(x_0 + t(x - x_0))(x - x_0)\, dt \right\|$$

$$\leq \int_0^1 \frac{2\gamma\|x - x_0\|\, dt}{[1 - \gamma t\|x - x_0\|]^3} \leq \frac{1}{(1 - \gamma\|x - x_0\|)^2} - 1,$$

which shows (3.10).

We can also get

(3.27) $\|F'(x_0)^{-1}[F(y) - F(x) - F'(x)(y - x)]\|$

$$= \left\| F'(x_0)^{-1} \int_0^1 F''(x + t(y - x))(1 - t)\, dt\, (y - x)^2 \right\|$$

$$\leq \int_0^1 \frac{2\gamma(1 - t)\|y - x\|^2\, dt}{[1 - \gamma\|x + t(y - x) - x_0\|]^3}$$

$$\leq \frac{\gamma\|y - x\|^2}{(1 - \gamma\|x - x_0\|)^2(1 - \gamma\|y - x_0\|)},$$

which implies (3.11).

(b) It also follows from the proof of Theorem 3.3 that the scalar sequence $\{r_n\}$ given by

$$r_0 = 0, \quad r_1 = \beta,$$

(3.28)
$$r_{n+1} = r_n - \frac{\gamma(1 - \gamma_0 r_n)^2(r_n - r_{n-1})^2}{[2(1 - \gamma_0 r_n)^2 - 1](1 - \gamma r_{n-1})^2(1 - \gamma r_n)} \quad (n \geq 1),$$

is majorizing for $\{x_n\}$, nondecreasing, and

(3.29) $\qquad 0 \leq r_n \leq t_n \leq w_n,$

(3.30) $\qquad 0 \leq r_{n+1} - r_n \leq t_{n+1} - t_n \leq w_{n+1} - w_n,$

(3.31) $\qquad 0 \leq r^* - r_n \leq t - t_n \leq w^* - w_n,$

(3.32) $\qquad r^* = \lim_{n \to \infty} r_n \leq t^* \leq w^*.$

CASE 2. We drop condition (1.16) in Theorem 3.3 and replace it by (3.37) given in Lemma 3.6 below.

In view of the advantages given by (3.29)–(3.32), we would like to know if a direct study of the sequence (3.28) can lead to a sufficient convergence condition weaker than (1.16), especially in the case when $\gamma_0 < \gamma$.

In what follows we show that this can be done.

As in Section 2, set $\gamma_0 = a\gamma$, $a \in [0, 1]$. Let $\delta \in (0, 2)$, and define a function h_a by

(3.33) $\qquad h_a(s) = 2a^2(1 + \delta)s^3 - 2a(2\delta + a\delta + 2)s^2 + (\delta + 2 + 4a\delta)s - \delta.$

We get $h_a(0) = -\delta < 0$ and $h_a(1) = 2(a - 1)^2 \geq 0$.

It then follows by the intermediate value theorem that there exists a zero of h_a in $(0, 1]$ for any $\delta \in (0, 2)$ and any $a \in [0, 1]$.

For each fixed $\delta \in (0, 2)$ and $a \in [0, 1]$, denote by α_0 the minimal zero of h_a on $(0, 1]$. Set

(3.34) $\qquad \alpha^* = \min\left\{\alpha_0, \alpha_1 = \left(1 - \dfrac{1}{\sqrt{2}}\right) \dfrac{1}{2a} (2 - \delta)\right\} > 0.$

We need a lemma on majorizing sequences for Newton's method (1.2).

LEMMA 3.5 ([1], [3]). *Let $\beta, \ell, \ell_0 \geq 0$ with $\ell_0 \leq \ell$. Assume that condition* (1.16) *holds. Then the scalar sequence*

(3.35) $\qquad z_0 = 0, \quad z_1 = \beta, \quad z_{n+1} = z_n + \dfrac{(z_n - z_{n-1})^2}{2(1 - \ell_0 z_n)} \quad (n \geq 1)$

is nondecreasing, bounded above by

$$z^{**} = \dfrac{2\beta}{2 - \beta_0}$$

and converges to some $z^* \in [\beta, z^{**}]$, where

(3.36) $\qquad \beta = \dfrac{1}{2}\left(-\dfrac{\ell}{\ell_0} + \sqrt{\left(\dfrac{\ell}{\ell_0}\right)^2 + 8\dfrac{\ell}{\ell_0}}\right), \quad \ell_0 \neq 0.$

We also need the following variation of Lemma 3.5 on majorizing sequences:

LEMMA 3.6. *Let $\beta, \gamma_0, \gamma > 0$, with $\gamma_0 = a\gamma$ for some $a \in [0, 1]$, and assume*

(3.37) $\qquad \begin{aligned} \alpha &= \beta\gamma \leq \alpha^* &&\text{if } \alpha_0 < \alpha_1, \\ \alpha &< \alpha_1 &&\text{if } \alpha_1 \leq \alpha_0. \end{aligned}$

Then the scalar sequence $\{r_n\}$ given by (3.28) *is nondecreasing, bounded above by*

(3.38) $\qquad r^{**} = \dfrac{2\beta}{2 - \delta},$

and converges to some $r^* \in [\beta, r^{**}]$. Moreover, the following error estimates hold for all $n \geq 1$:

(3.39) $\qquad 0 \leq r_{n+1} - r_n \leq \dfrac{\delta}{2}(r_n - r_{n-1}).$

Proof. We show by induction on k that

(3.40) $$\frac{\gamma(1 - \gamma_0 t_k)^2 (t_k - t_{k-1})}{[2(1 - \gamma_0 t_k)^2 - 1](1 - \gamma t_{k-1})^2 (1 - \gamma t_k)} \leq \frac{\delta}{2},$$

(3.41) $$0 \leq t_k - t_{k-1},$$

(3.42) $$t_{k-1} < \frac{1}{\gamma},$$

(3.43) $$t_k < \left(1 - \frac{1}{\sqrt{2}}\right)\frac{1}{\gamma_0}.$$

It follows from the choices of α_0, α_1, α^*, and (3.28) that estimates (3.40)–(3.43) hold true for $k = 1$. ∎

Hence, using (3.35) instead condition (1.16) in Theorem 3.3 we arrive at:

THEOREM 3.7. *Under the hypotheses of Theorem 3.3 with (3.37) replacing condition (1.16), and $\{r_n\}$ replacing $\{t_n\}$, the sequence $\{x_n\}$ generated by Newton's method (1.2) is well defined, remains in $\overline{U}(x_0, r^*)$, and converges to a solution x^* of the equation $F(x) = 0$, which is unique in $U(x_0, r^{**})$. Moreover, the following estimates hold for all $n \geq 0$:*

(3.44) $$\|x_{n+1} - x_n\| \leq r_{n+1} - r_n,$$

(3.45) $$\|x_n - x^*\| \leq r^* - r_n.$$

REMARK 3.8. (a) Wang proved in [7] that the γ-condition (1.15) is weaker than the criterion of point estimate given by Smale [6]:

(3.46) $$\gamma(F, x_0) = \sup_{n \geq 2}\left\|F'(x_0)^{-1}\frac{F^{(n)}(x_0)}{n!}\right\|^{1/(n-1)} \leq \gamma,$$

provided that the operator F is analytic on D. It then follows that we can choose

(3.47) $$\gamma_0 = \gamma(F, x_0).$$

That is, inequality (3.7) can be realized.

(b) As an example, let us set $a = 1/2$ and $\delta = .6$. Then using conditions (3.33) and (3.34) we get

$$\alpha_0 = .18955, \quad \alpha_1 = .205025,$$

and

(3.48) $$x^* = .18955 > 3 - 2\sqrt{2} = .17157.$$

Hence, we conclude:

(i) If (1.16) is violated but condition (3.37) holds true, then Theorem 3.7 can be used. That is, the applicability of Newton's method is extended in this case.

(ii) If condition (1.16) holds, then our finer Theorem 3.3 (or Theorem 3.7) can replace Theorem 3.1 given in [7].

CASE 3. We use conditions (3.49) and (3.50) instead of (1.15).

If the operator F is thrice Fréchet-differentiable on D, we can introduce the following conditions:

DEFINITION 3.9. Let $F : D \subseteq X \to Y$ be thrice Fréchet-differentiable. Assume that

$$(3.49) \qquad \|F'(x_0)^{-1} F''(x_0)\| \le 2\gamma_0,$$

and

$$(3.50) \qquad \|F'(x_0)^{-1} F'''(x)\| \le \frac{6\gamma^2}{(1 - \gamma_0\|x - x_0\|)^4} = f'''(\|x - x_0\|)$$

for all $x \in D$.

REMARK 3.10. According to the lemma below conditions (3.49) and (3.50) can replace (1.15) (for $\gamma_0 = \gamma$) in Theorem 3.1, and (3.10) and (3.11) in Theorems 3.3 and 3.7. In this case inequality (3.7) can also be realized for γ_0 given by (3.47) or not. Finally, note that clearly as in (3.34) there exists $\delta \in (0, 2)$ such that

$$\alpha \le 3 - 2\sqrt{2} \;\Rightarrow\; \alpha \le \alpha^*,$$

but not necessarily vice versa, unless $\gamma_0 = \gamma$.

LEMMA 3.11. *Assume that the operator F satisfies the conditions of Definition 3.9 on $U(x_0, \gamma_1) \subseteq D$. Then the following estimates hold on $U(x_0, \gamma_1)$:*

$$(3.51) \qquad \|F'(x_0)^{-1} F''(x)\| \le f''(\|x - x_0\|) + 2(\gamma_0 - \gamma),$$

$$(3.52) \qquad\qquad F'(x) \in L(Y, X),$$

and

$$(3.53) \qquad \|F'(x)^{-1} F'(x_0)\| \le -\frac{1}{f'(\|x - x_0\|) - 2(\gamma_0 - \gamma)\|x - x_0\|}.$$

Proof. Using (3.49) and (3.50) we get

$$
\begin{aligned}
(3.54) \quad & \|F'(x_0)^{-1} F''(x)\| \\
& \le \|F'(x_0)^{-1} F''(x_0)\| + \|F'(x_0)^{-1}[F''(x) - F''(x_0)]\| \\
& = \|F'(x_0)^{-1} F''(x_0)\| + \left\| \int_0^1 F'(x_0)^{-1} F'''(x_0 + t(x - x_0))(x - x_0)\, dt \right\| \\
& \le 2\gamma_0 + \int_0^1 f'''(t\|x - x_0\|)\|x - x_0\|\, dt \\
& = 2\gamma_0 + f''(\|x - x_0\|) - f''(0) = f''(\|x - x_0\|) + 2(\gamma_0 - \gamma),
\end{aligned}
$$

which shows (3.51).

Moreover, using (3.49) and (3.51), we obtain

$$(3.55) \quad \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\|$$

$$= \left\| F'(x_0)^{-1} \int_0^1 F''(x_0 + t(x - x_0))(x - x_0)\, dt \right\|$$

$$\leq \int_0^1 [f''(t\|x - x_0\|) + 2(\gamma_0 - \gamma)]\|x - x_0\|\, dt$$

$$= f'(\|x - x_0\|) - f'(0) + 2(\gamma_0 - \gamma)\|x - x_0\|$$

$$= f'(\|x - x_0\|) + 1 + 2(\gamma_0 - \gamma)\|x - x_0\| < 1,$$

which together with the Banach lemma on invertible operators implies (3.52) and (3.53). ∎

References

[1] 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.

[2] —, *On the convergence of the secant method under the gamma condition*, Central Eur. J. Math. 5 (2007), 205–214.

[3] —, *Computational Theory of Iterative Methods*, Stud. Comput. Math. 15, Elsevier, New York, 2007.

[4] L. V. Kantorovich and G. P. Akilov, *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.

[5] W. C. Rheinboldt, *An adaptive continuation process for solving systems of nonlinear equations*, in: Mathematical Models and Numerical Methods, A. N. Tikhonov et al. (eds.), Banach Center Publ. 3, Polish Sci. Publ., 1978, 129–142.

[6] S. Smale, *Newton's method estimates from data at one point*, in: The Merging of Disciplines in Pure, Applied and Computational Mathematics, Springer, New York, 1986, 185–186.

[7] X. H. Wang, *Convergence on the iteration of Halley family in weak conditions*, Chinese Sci. Bull. 42 (1997), 552–555.

[8] F. Zhao and D. Wang, *The theory of Smale's point estimation, and its applications*, J. Comput. Appl. Math. 60 (1995), 253–259.

Department of Mathematical Sciences
Cameron University
Lawton, OK 73505, U.S.A.
E-mail: iargyros@cameron.edu