JOURNAL OF THE
CHUNGCHEONG MATHEMATICAL SOCIETY
Volume **25**, No. 2, May 2012

LOCAL CONVERGENCE RESULTS FOR NEWTON'S METHOD

Ioannis K. Argyros* and Saïd Hilout**

Abstract. We present new results for the local convergence of Newton's method to a unique solution of an equation in a Banach space setting. Under a flexible gamma–type condition [12], [13], we extend the applicability of Newton's method by enlarging the radius and decreasing the ratio of convergence. The results can compare favorably to other ones using Newton–Kantorovich and Lipschitz conditions [3]–[7], [9]–[13]. Numerical examples are also provided

1. Introduction

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

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

where F is a Fréchet–differentiable operator defined on a convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$.

A large number of problems in applied mathematics and also in engineering are 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 by solving equation (1.1). Similar equations are used in the case of discrete

Received January 20, 2012; Accepted April 17, 2012.
2010 Mathematics Subject Classification: Primary 65G99, 65K10, 47H17; Secondary 49M15.
Key words and phrases: Newton's method, Banach space, local convergence, radius of convergence, ratio of convergence, gamma–type condition, Lipschitz condition, Fréchet–derivative.

Correspondence should be addressed to Ioannis K. Argyros, iargyros@cameron.edu.

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). Excpet 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.

Newton's method

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

is undoubtedly the most popular method for generating a sequence approximating $x^\star$. Here $F'(x) \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$, denotes the Fréchet–derivative of operator F [6], [10]. A survey on local as well as semilocal convergence theorems for Newton's method (1.2) under Newton–Kantorovich–type or γ–type conditions can be found in [6], [10] and the references there (see [1]–[5], [7]–[9], [11]–[13]).

In Section 2, we provide new semilocal convergence theorem under a γ–type condition (see Definition 2.1). We also compare semilocal and local results on Newton's method in order for us to answer to the question: (which is the motivation for writing this paper)

Can you find conditions under which the largest convergence radius and the smaller ratio can be obtained for Newton's method ?

Numerical examples are also provided.

2. Local convergence analysis of Newton's method (1.2)

Let $b \geq 0$, $\gamma_0 > 0$ and $\gamma > 0$ be given constants. It is convient for us to define constant a and functions f_0, f, g_0, g on interval $[0, \frac{1}{\gamma})$ by

$$a = \frac{\gamma_0 - \gamma}{\gamma},$$

$$f(t) = b - t + \frac{\gamma\, t^2}{1 - \gamma\, t},$$

$$f_0(t) = f(t) + (\gamma - \gamma_0)\, t^2,$$

$$g(t) = 3\,t^2 - t - 1$$

and

$$g_0(t) = 2\,a\,(1-t)\,t^2 - g(t).$$

We shall assume:

$$\gamma_0 < \gamma.$$

Set

$$t_1 = \frac{1+\sqrt{13}}{6}.$$

It then follows that $g(t_1) = 0$ and $g_0(t_1) < 0$.

Therfore, by the intermediate value theorem, there exists a real $t_0 \in (0, t_1)$ such that:

$$g_0(t_0) = 0.$$

Denote also by t_0 to the minimal such number in $(0, t_1)$. Set

$$t_0 = 1 - \gamma\, r_0$$

and

$$t_1 = 1 - \gamma\, r_1.$$

It then follows that

$$r_1 = \frac{5-\sqrt{13}}{6\,\gamma} < r_0.$$

Note also that if $\gamma = \gamma_0$, then $r_0 = r_1$, $t_0 = t_1$, $f(t) = f_0(t)$ and $g(t) = g_0(t)$ on $[0, t_0)$. We need the following definition of a γ–type condition:

Definition 2.1. *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be a thrice–Fréchet–differentiable operator. We say that operator F satisfies the (γ_0, γ)–condition at $x^\star \in \mathcal{D}$ if*

$$F'(x^\star)^{-1} \in L(\mathcal{Y}, \mathcal{X});$$

$$\| \; F'(x^\star)^{-1} \; F''(x^\star) \; \| \leq 2 \; \gamma_0,$$

$$\begin{aligned} &\| \; F'(x^\star)^{-1} \; F'''(x^\star) \; \| \\ &\quad \leq \frac{6 \; \gamma^2}{(1-\gamma \;\; \| \; x - x^\star \; \|)^4} = f_0'''(\| \; x - x^\star \; \|) = f'''(\| \; x - x^\star \; \|). \end{aligned}$$

for all $x \in \mathcal{D}$ and

$$\overline{U}(x^\star, r_0) = \{ x \in \mathcal{X} \; : \| \; x - x^\star \; \| \leq r_0 \} \subseteq \mathcal{D},$$

Note that a suitable choice (but not the only one) for γ_0 and γ is: $\gamma > \gamma_0$ and

$$\gamma_0 = \sup_{k \geq 2} \| \frac{F'(x^\star)^{-1} F^{(k)}(x^\star)}{k!} \|^{\frac{1}{k-1}},$$

provided that operator F is analytic on $\overline{U}(x^\star, r_0)$ [12], [13] and the supremum is finite.

If $\gamma_0 = \gamma$, then we replace r_0 by r_1 in Definition 2.1.

We also need the following lemma connecting operator F with majorizing function f_0.

LEMMA 2.2. *Suppose that F satisfies the (γ_0, γ)–condition. Then the following hold*

$$\| F'(x^\star)^{-1} \; F''(x) \| \leq f_0''(\| x - x^\star \|),$$

$$F'(x)^{-1} \in L(\mathcal{Y}, \mathcal{X})$$

and

$$\| F'(x)^{-1} \; F'(x^\star) \| \leq -\frac{1}{f_0'(\| x - x^\star \|)}.$$

for all $x \in U(x^\star, r_0)$.

Proof. Using Defintion 2.1, we obtain in turn:

$$\begin{aligned}
&\| F'(x^\star)^{-1} \; F''(x) \| \\
&\quad = \| F'(x^\star)^{-1} \; F''(x^\star) \| + \| F'(x^\star)^{-1} \; (F''(x) - F''(x^\star)) \| \\
&\quad \leq 2 \; \gamma_0 + \| \int_0^1 F'(x^\star)^{-1} \; F'''(x^\star + t\,(x - x^\star))\,(x - x^\star)\,dt \| \\
&\quad \leq 2 \; \gamma_0 + \int_0^1 f_0'''(t \; \| x - x^\star \|) \; \| x - x^\star \| \; dt \\
&\quad = 2 \; \gamma_0 + f_0''(\| x - x^\star \|) - f''(0) = f_0''(\| x - x^\star \|).
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
&\| F'(x^\star)^{-1} \; (F'(x) - F'(x^\star)) \| \\
&\quad = \| F'(x^\star)^{-1} \int_0^1 F''(x^\star + t\,(x - x^\star))\,(x - x^\star)\,dt \| \\
&\quad \leq \int_0^1 f_0''(t \; \| x - x^\star \|) \; \| x - x^\star \| \; dt \\
&\quad = f_0'(\| x - x^\star \|) - f_0'(0) = f_0'(\| x - x^\star \|) + 1 < 1.
\end{aligned}$$

It follows from the Banach Lemma on invertible operators [3], [7] that $F'(x)^{-1} \in L(\mathcal{Y}, \mathcal{X})$ and

$$\| F'(x)^{-1} \; F'(x^\star) \| \leq \frac{1}{1- \| F'(x^\star)^{-1} \; (F'(x) - F'(x^\star)) \|}$$
$$\leq - \frac{1}{f_0'(\| x - x^\star \|)}.$$

That complete the proof of the lemma. □

We can show the following local convergence theorem for Newton's method:

Theorem 2.3. *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be a thrice–Fréchet–differentiable operator. Then under hypotheses of Definition 2.1, with $\gamma_0 = \gamma$, sequence $\{x_n\}$ generated by the Newton–type method (1.2) is well defined, remains in $U(x^\star, r_1)$ for all $n \geq 0$ and converges to the unique zero of equation $F(x) = 0$ in $\overline{U}(x^\star, r_1)$ provided that $x_0 \in U(x^\star, r_1)$. Moreover the following estimates hold for all $n \geq 0$:*

$$\| x_{n+1} - x^\star \| \leq a_n \, b_n \; \| x_n - x^\star \|^2, \tag{2.1}$$

where

$$a_n = a_n(\| x_n - x^\star \|) = -g'(\| x_n - x^\star \|)^{-1}$$

and

$$b_n = b_n(\| x_n - x^\star \|) = \frac{\gamma}{1 - \gamma \; \| x_n - x^\star \|}.$$

Proof. By hypothesis, we see that $x_0 \in U(x^\star, r_1)$. Using induction on $k \geq 0$, we shall show that $x_{k+1} \in U(x^\star, r_1)$, so that (2.1) holds true. By Lemma 2.2, for $x = x_k$, we get $F'(x_k)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ and

$$\| F'(x_k)^{-1} \; F'(x^\star) \| \leq -f'(\| x_k - x^\star \|)^{-1}. \tag{2.2}$$

In view of the identity

$$\begin{aligned} x_{k+1} - x^\star \;&=\; -(F'(x_k)^{-1} \; F'(x^\star)) \, F'(x^\star)^{-1} \\ & \quad \int_0^1 F''(x^\star + t \, (x_k - x^\star)) \, (1 - t) \, (x_k - x^\star)^2 \, dt \end{aligned} \tag{2.3}$$

and Definition 2.1, we get in turn

$$\begin{aligned} & \| F'(x^\star)^{-1} \int_0^1 F''(x^\star + t \, (x_k - x^\star)) \, (1 - t) \, (x_k - x^\star)^2 \, dt \, \| \\ & \quad \leq \frac{\gamma \, \| x_k - x^\star \|^2}{1 - \gamma \; \| x_n - x^\star \|} = b_k \; \| x_k - x^\star \|^2 \, . \end{aligned} \tag{2.4}$$

It then follows from (2.2)–(2.4) that

$$\begin{aligned}\| x_{k+1} - x^\star \| &\leq a_k\, b_k \; \| x_k - x^\star \|^2 \\ &\leq a_k(r_1)\, b_k(r_1)\, r_1 \; \| x_k - x^\star \| \leq \| x_k - x^\star \| < r_1\end{aligned}$$

which shows $x_{k+1} \in U(x^\star, r_1)$ and $\lim_{k\to\infty} x_k = x^\star$. Finally, to show uniqueness, let $y^\star$ be a solution of equation (1.1) in $\overline{U}(x^\star, r_1)$. Then, we get by Definition 2.1

$$\begin{aligned}&\| F'(x^\star)^{-1} \int_0^1 (F'(x^\star + t\,(y^\star - x^\star)) - F'(x^\star))\, dt \, \| \\ &\leq \| F'(x^\star)^{-1} \int_0^1 \int_0^1 F''(x^\star + s\,t\,(y^\star - x^\star))\, ds\, dt\, (t\,(y^\star - x^\star)) \, \| \\ &\leq \int_0^1 \int_0^1 f''(s\,t \; \| y^\star - x^\star \|)\, t\, ds\, dt \;\; \| y^\star - x^\star \| \\ &= \int_0^1 f'(t \; \| y^\star - x^\star \|)\, dt + 1 < 1.\end{aligned} \tag{2.5}$$

It follows from (2.5) and the Banach lemma of invertible operators that $\mathcal{M} = \int_0^1 F'(x^\star + t\,(y^\star - x^\star))\, dt$ is invertible. In view of the identity

$$F(y^\star) - F(x^\star) = \mathcal{M}\,(y^\star - x^\star),$$

we get $x^\star = y^\star$. That completes the proof of the theorem. □

Under Definition 2.1 for $\gamma_0 < \gamma$, we can show the following improvement of Theorem 2.3:

THEOREM 2.4. *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be a thrice–Fréchet–differentiable operator. Then under hypotheses of Definition 2.1, with $\gamma_0 < \gamma$, sequence $\{x_n\}$ generated by the Newton–type method (1.2) is well defined, remains in $U(x^\star, r_0)$ for all $n \geq 0$ and converges to the unique zero of equation $F(x) = 0$ in $\overline{U}(x^\star, r_0)$ provided that $x_0 \in U(x^\star, r_0)$. Moreover the following estimates hold for all $n \geq 0$:*

$$\| x_{n+1} - x^\star \| \leq \overline{a}_n\, \overline{b}_n \; \| x_n - x^\star \|^2, \tag{2.6}$$

where

$$\overline{a}_n = -g_0'(\| x_n - x^\star \|)^{-1}$$

and

$$\overline{b}_n = b_n + \gamma_0 - \gamma.$$

Proof. It follows exactly as in Theorem 2.3 with g_0, f_0, r_0 replacing g, f and r_1 respectively. □

In order for us to compare the ratios in the above theorems, let

$$c_n = a_n\, b_n$$

and

$$\bar{c}_n = \bar{a}_n\, b_n.$$

Then using induction on $k \geq 0$, we arrive at:

Proposition 2.5. *Under hypotheses of Theorems 2.3 and 2.4,*

$$\bar{c}_n \leq c_n \tag{2.7}$$

and

$$r_1 \leq r_0 \tag{2.8}$$

hold for $n \geq 0$. Moreover if $\gamma_0 < \gamma$, then (2.7) and (2.8) hold as strict inequalities.

That is, in Theorem 2.4, a larger radius and a smaller ratio of convergence are obtained than in Theorem 2.3. In this way, the number of initial guesses is enlarged and the number of steps required to obtain a desired error tolerance $\epsilon > 0$ is smaller. These observations are important in computational mathematics [1], [6], [7], [14].

Under the selection of $\gamma^\star$, Wang and Zhao in [13] provided the radius of convergence given by

$$r_2 = \frac{3 - 2\sqrt{2}}{2\,\gamma^\star} < r_1 < r_0,$$

for $\gamma^\star = \gamma_0$ and $\gamma \geq \gamma_0$. That is, Theorem 2.4 provides the largest radius and the smallest ratio of convergence under the γ–condition. It turns out that Theorem 2.4 can compare favorably with others theorems using Lipschitz–type conditions

$$\parallel F'(x^\star)^{-1}\,(F'(x) - F'(y)) \parallel \leq l\ \parallel x - y \parallel$$

and

$$\parallel F'(x^\star)^{-1}\,(F'(x) - F'(x^\star)) \parallel \leq l_0\ \parallel x - x^\star \parallel$$

for all x, $y \in \mathcal{D}$. Indeed Rheinboldt's [11] ball r_W is given by

$$r_W = \frac{2}{3\,l}$$

and Argyros' [3, 6] ball r_A is given by

$$r_A = \frac{2}{2\,l_0 + l}.$$

Note that since $l_0 \leq l$ and $\dfrac{l}{l_0}$ can be arbitrarily large [4]–[6], we obtain

$$r_W < r_A,$$

unless if $l_0 = l$.

Remark 2.6. *As noted in* [1], [5], [6], [7], [10], [14] *the local results obtained here can be used for projection method such us Arnold's, the generalized minimum residual method (GMRES), the generalized conjugate residual method (GCR), for combined Newton/finite projection methods and in connection with the mesh independence principle to develop the cheapest and most efficient mesh refinement strategies.*

Remark 2.7. *The local results obtained can also be used to solve equation of the form $F(x) = 0$, where F' satisfies the autonomous differential equation* [4]*:*

$$F'(x) = P(F(x)), \tag{2.9}$$

where $P : \mathcal{Y} \longrightarrow \mathcal{X}$ is a known continuous operator. Since $F'(x^\star) = P(F(x^\star)) = P(0)$, we can apply our results without actually knowing the solution of $x^\star$ of equation (1.1).

We complete this study with a simple numerical example where we compare the radii introduced above.

Example 2.8. *Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, $\mathcal{D} = U(0,1)$ and define function h on $\mathcal{D}$ by*

$$F(x) = e^x - 1.$$

Note that we can set $P(x) = x + 1$ in (2.9). It can easily be seen that $l_0 = e - 1$, $l = e$, $\gamma_0 = \gamma = \dfrac{1}{2}$ and $\gamma^\star < \dfrac{1}{2}$. Therefore, we get

$$r_R = .245252961, \quad r_A = .324947231 \quad \text{and} \quad r_0 = r_1 = .46481624.$$

which justify the claims made above this example.

References

[1] E. L. Allgower, K. Böhmer, F. A. Potra, W.C. Rheinboldt, *A mesh independence principle for operator equations and their discretizations*, SIAM J. Numer. Anal. **23** (1986), 160-169.

[2] S. Amat, S. Busquier, *Convergence and numerical analysis of a family of two–step Steffensen's method*, Comput. Math. Appl. **49** (2005), 13-22.

[3] I. K. Argyros, *On the Newton–Kantorovich hypothesis for solving equations*, J. Comput. Appl. Math. **169** (2004), 315-332.

[4] I. K. Argyros, *On the secant method for solving nonsmooth equations*, J. Math. Anal. Appl. **322** (2006), 146-157.

[5] 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.

[6] I. K. Argyros, *Computational Theory of Iterative Methods*, Studies in Computational Mathematics, **15**, Elsevier, 2007, New York, U.S.A.

[7] P. N. Brown, *A local convergence theory for combined inexact–Newton/finite difference projection methods*, SIAM J. Numer. Anal. **24** (1987), 407-434.

[8] S. Chandrasekhar, *Radiative Transfer*, Dover Publ. New York, 1960.

[9] J. M. Gutiérrez, M. A. Hernánadez, M. A. Salanova, *Accessibility of solutions by Newton's method*, Inter. J. Comput. Math. **57** (1995), 239-241.

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

[11] W. C. Rheinboldt, *An adaptive continuation process for solving systems of nonlinear equations*, Banach Center Publ. **3** (1975), 129-142.

[12] X. H. Wang, *Convergence on the iteration of Halley family in weak conditions*, Chinese Science Bulletin **42** (1997), 552-556.

[13] D. Wang, F. Zhao, *The theory of Smale's point estimation and its applications*, J. Comput. Appl. Math. **60** (1995), 253-269.

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

*

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

**

Poitiers University
Laboratoire de Mathématiques et Applications
Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179
86962 Futuroscope Chasseneuil Cedex, France
E-mail: said.hilout@math.univ-poitiers.fr