

ON THE LOCAL CONVERGENCE OF A NEWTON-TYPE METHOD IN BANACH SPACES UNDER A GAMMA-TYPE CONDITION

IOANNIS K. ARGYROS
CAMERON UNIVERSITY, U.S.A.

and

SAÏD HILOUT
POITIERS UNIVERSITY, FRANCE

Received : November 2007. Accepted : January 2008

Abstract

We provide a local convergence analysis for a Newton-type method to approximate a locally unique solution of an operator equation in Banach spaces. The local convergence of this method was studied in the elegant work by Werner in [11], using information on the domain of the operator. Here, we use information only at a point and a gamma-type condition [4], [10]. It turns out that our radius of convergence is larger, and more general than the corresponding one in [10]. Moreover the same can hold true when our radius is compared with the ones given in [9] and [11]. A numerical example is also provided.

AMS Subject Classification. 65G99, 65K10, 47H17, 49M15.

Key Words. Banach space, Newton-type method, local convergence, gamma-type condition, local convergence, Fréchet-derivative, radius of convergence.

1. Introduction

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

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

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

We revisit the Newton-type method given by $x_0, y_0 \in D$ by

$$(1.2) \quad \begin{aligned} x_{n+1} &= x_n - F'(z_n)^{-1} F(x_n), & z_n &= \frac{x_n + y_n}{2}, \quad (n \geq 0), \\ y_{n+1} &= x_n - F'(z_n)^{-1} F(x_{n+1}), \end{aligned}$$

to generate a sequence $\{x_n\}$, $(n \geq 0)$ approximating x^* [4], [11].

Let us illustrate how this method is conceived:

We start with the identity

$$(1.3) \quad F(x) - F(y) = \int_0^1 F'(y + t(x - y)) dt (x - y) \quad \text{for all } x, y \in D.$$

If x^* is a solution of equation (1.1), then identity (1.3) gives

$$(1.4) \quad F(x) = \int_0^1 F'(x + t(x^* - x)) dt (x^* - x) \quad \text{for all } x \in D.$$

The linear operator in (1.4) can be approximated in different ways [1], [3], [4], [12].

If for example

$$(1.5) \quad \int_0^1 F'(x + t(x^* - x)) dt \simeq F'(x) \quad \text{for all } x \in D,$$

then (1.4) suggests the famous Newton's method [1]–[12]:

$$(1.6) \quad x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) \quad (n \geq 0).$$

Another choice is given by

$$(1.7) \quad \int_0^1 F'(x + t(x^* - x)) dt \simeq F'\left(\frac{x^* + x}{2}\right) \quad \text{for all } x \in D,$$

which leads to the implicit iteration:

$$(1.8) \quad x_{n+1} = x_n - F'\left(\frac{x_n + x_{n+1}}{2}\right)^{-1} F(x_n), \quad (n \geq 0).$$

Unfortunately iterates in (1.8) can only be computed in very restrictive cases, and numerically, the method (1.8) is not a practical procedure.

That is why we consider y_n given in (1.2) as a suitable replacement for x_{n+1} ($n \geq 0$). Hence, we arrive at method (1.2), which requires the computation of two iterates x_n and y_n . The computation of the additional iterate y_n can be seen as a step to calculate the iterate x_{n+1} using Newton's method (1.8).

This shows that iterate x_{n+1} , thus defined is corrected by computing the iterate y_{n+1} using (1.8). Another advantage of method (1.2) is that the particular case $x_0 = y_0$ corresponds to the classical Newton's method (1.8). Procedure (1.2) has a geometrical interpretation similar to the tangent-Secant method in the scalar case, and was introduced by King [8] (see procedure (I, II), p. 299), and extended into Banach space by Werner in [11], where the R-order $1 + \sqrt{2}$ local convergence was established.

Here, we provide a local convergence analysis of the Newton-type method (1.2) using a γ -type condition (see (2.3) and (2.4)). Our radius of convergence r_A (see Theorem 2.2) is larger than the corresponding one denoted by r_W (see (2.28)) given in the elegant work by Wang and Zhao [10]. Note also that a special choice of γ denoted by γ^* (see (2.29)) used in [10].

As it turns out the radius of convergence can be larger than the radii given in [9], [11] where information on a domain is used (see (2.30) and (2.32)) instead of only information at a point used by us.

A numerical example is also provided.

2. Local convergence analysis of the midpoint method (1.2)

Let us define scalar function f on $[0, \frac{1}{\gamma})$ by

$$(2.1) \quad f(t) = b - t + \frac{\gamma t^2}{1 - \gamma t},$$

where $b \geq 0$, and $\gamma > 0$ are given.

It is known [9] that if

$$(2.2) \quad \alpha = b \gamma \leq 3 - 2\sqrt{2},$$

then function f has two roots

$$t^* = \frac{1 + \alpha - \sqrt{(1 + \alpha)^2 - 8\alpha}}{2\gamma}, \quad t^{**} = \frac{1 + \alpha + \sqrt{(1 + \alpha)^2 - 8\alpha}}{2\gamma}$$

satisfying

$$b \leq t^* \leq (1 + \frac{1}{\sqrt{2}}) b \leq (1 - \frac{1}{\sqrt{2}}) \frac{1}{\gamma} \leq t^{**} \leq \frac{1}{2\gamma}.$$

We use throughout this paper the concept of γ -conditions:

Definition 2.1. *An operator $F : D \subseteq X \longrightarrow Y$ satisfies γ -conditions if the following hold:*

(i) *There exists a zero $x^* \in D$ of operator F such that*

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

(ii) *Operator F is thrice-Fréchet-differentiable on D , and for all $x \in D$*

$$(2.3) \quad \| F'(x^*)^{-1} F''(x^*) \| \leq 2\gamma,$$

and

$$(2.4) \quad \| F'(x^*)^{-1} F'''(x^*) \| \leq \frac{6\gamma^2}{(1 - \gamma \|x - x^*\|)^4} = f'''(\|x - x^*\|).$$

In view of (2.1), we have

$$f'(t) = \frac{1 - 2(1 - \gamma t)^2}{(1 - \gamma t)^2},$$

$$f''(t) = \frac{2\gamma}{(1 - \gamma t)^3},$$

and

$$f'''(t) = \frac{6\gamma^2}{(1-\gamma t)^4}.$$

We need the following Lemma:

Lemma 2.2. *Under the γ -conditions given by Definition 2.1, and for all $x \in U(x^*, (1 - \frac{1}{\sqrt{2}}) \frac{1}{\gamma} = r_0) = \{x \in X : \|x - x^*\| < r_0\} \subseteq D$, the following estimates hold:*

$$(2.5) \quad \|F'(x^*)^{-1} F''(x)\| \leq f''(\|x - x^*\|),$$

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

and

$$(2.6) \quad \|F'(x)^{-1} F'(x^*)\| \leq -\frac{1}{f'(\|x - x^*\|)}.$$

Proof. Using the γ -conditions, and the properties of function f , we obtain in turn:

$$\begin{aligned} \|F'(x^*)^{-1} F''(x)\| &\leq \|F'(x^*)^{-1} F''(x^*)\| + \|F'(x^*)^{-1} (F''(x) - F''(x^*))\| \\ &= \|F'(x^*)^{-1} F''(x^*)\| + \\ &\quad \left\| \int_0^1 F'(x^*)^{-1} F''(x^* + t(x - x^*)) (x - x^*) dt \right\| \\ &\leq 2\gamma + \int_0^1 f''(t\|x - x^*\|) \|x - x^*\| dt \\ &= 2\gamma + f''(\|x - x^*\|) - f''(0) = f''(\|x - x^*\|). \end{aligned}$$

Moreover, we have

$$\begin{aligned} \|F'(x^*)^{-1} (F'(x) - F'(x^*))\| &= \|F'(x^*)^{-1} \int_0^1 F''(x^* + t(x - x^*)) (x - x^*) dt\| \\ &\leq \int_0^1 f''(t\|x - x^*\|) \|x - x^*\| dt \\ &= f'(\|x - x^*\|) - f'(0) = f'(\|x - x^*\|) + 1 < 1. \end{aligned}$$

It follows by the Banach Lemma on invertible operators [4], [12] that $F'(x)^{-1} \in L(Y, X)$, and

$$\begin{aligned} \|F'(x)^{-1} F'(x^*)\| &\leq \frac{1}{1 - \|F'(x^*)^{-1} (F'(x) - F'(x^*))\|} \\ &\leq -\frac{1}{f'(\|x - x^*\|)}. \end{aligned}$$

That complete the proof of the Lemma. \diamond

It is convenient for us to define sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$, $\{d_n\}$ by

$$a_n = \frac{\gamma}{1 - \gamma \|x_n - x^*\|}, \quad b_n = \frac{\gamma^2}{4(1 - \gamma \|x_n - x^*\|)},$$

$$c_n = \frac{(1 - \gamma \|z_n - x^*\|)^2}{2(1 - \gamma \|z_n - x^*\|)^2 - 1},$$

$$d_n = \frac{\gamma}{4(1 - \frac{\gamma}{2} \|x_n - x^*\|)(1 - \frac{\gamma}{2} (\|x_n - x^*\| + \|y_n - x^*\|))};$$

and functions a , b , c , d on $[0, r_0)$ by

$$a(r) = \frac{r}{1 - r}, \quad b(r) = \frac{r^2}{4(1 - r)},$$

$$c(r) = \frac{(1 - r)^2}{2(1 - r)^2 - 1},$$

$$d(r) = \frac{r}{4(1 - \frac{r}{2})(1 - r)}.$$

It is simple algebra to see that system of inequalities

$$(2.7) \quad c(r) [b(r) + 3d(r)] \leq 1$$

is satisfied for all

$$(2.8) \quad r \in [0, \frac{5 - \sqrt{13}}{6}).$$

We shall also use the identities [4]:

$$F(x^*) - F(x) - F'(x)(x^* - x) = \int_0^1 F''(x + t(x^* - x))(1 - t)(x^* - x)^2 dt,$$

(2.9)

$$\begin{aligned}
F(x) - F(y) - F'(z)(x - y) &= \frac{1}{4} \int_0^1 \left[F''\left(z + \frac{t}{2}(x - y)\right) - F''\left(z + \frac{t}{2}(y - x)\right) \right] \\
&\quad (1 - t)(x - y)^2 dt \\
&= \frac{1}{4} \int_0^1 F''' \left(z + \frac{t}{2}(y - x) + s t(x - y) \right) \\
&\quad s(1 - t)(x - y)^3 ds dt,
\end{aligned}
\tag{2.10}$$

$$F'(z) - F'\left(\frac{x + x^*}{2}\right) = \int_0^1 F'' \left(\frac{x + x^*}{2} + t \left(\frac{y - x^*}{2} \right) \right) \left(\frac{y - x^*}{2} \right) dt,
\tag{2.11}$$

and

$$F'\left(\frac{x^* + w}{2}\right) - F'(z) = \int_0^1 F'' \left(z + \frac{t}{2}(x^* + w - x - y) \right) \left(\frac{x^* + w - x - y}{2} \right) dt,
\tag{2.12}$$

for $z = \frac{x + y}{2}$, and all $x, y, w \in D$.

We can show the local convergence theorem for the Newton-type method (1.2):

Theorem 2.3. *Under the γ -conditions given by Definition 2.1 for $x \in \overline{U}(x^*, r^* = \frac{5 - \sqrt{13}}{6\gamma}) \subseteq D$, sequences $\{x_n\}, \{y_n\}$ generated by the Newton-type method (1.2) are well defined, remain in $U(x^*, r^*)$ for all $n \geq 0$, and converge to the unique zero of equation $F(x) = 0$ in $\overline{U}(x^*, r^*)$ provided that $x_0, y_0 \in U(x^*, r^*)$.*

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

$$\|x_{n+1} - x^*\| \leq c_n \left[b_n \|x_n - x^*\|^2 + d_n \|y_n - x^*\| \right] \|x_n - x^*\|, \quad (2.13)$$

and

$$\|y_{n+1} - x^*\| \leq c_n \left[b_{n+1} \|x_{n+1} - x^*\|^2 + d_n (\|x_{n+1} - x^*\| + \|y_n - x^*\| + \|x_n - x^*\|) \right] \|x_{n+1} - x^*\|. \quad (2.14)$$

Proof. By hypotheses $x_0, y_0 \in U(x^*, r^*)$, and for $x = \frac{x_0 + y_0}{2}$ in (2.6) we get $F'(z_0)^{-1}$ exists, and

$$\|F'(z_0)^{-1} F'(x^*)\| \leq -\frac{1}{g'(\|z_0 - x^*\|)}. \quad (2.15)$$

Let us assume that $x_k, y_k \in U(x^*, r^*)$ for $k = 0, 1, \dots, n$. Then by (2.6) $F'(z_k)^{-1}$ exists, and

$$\|F'(z_k)^{-1} F'(x^*)\| \leq -\frac{1}{g'(\|z_k - x^*\|)}. \quad (2.16)$$

We shall show that $x_{k+1}, x_{k+1} \in U(x^*, r^*)$, and estimates (2.13), (2.14) hold true.

Using (1.2) we obtain the identity

$$\begin{aligned} x_{k+1} - x^* &= x_k - F'(z_k)^{-1} F(x_k) - x^* \\ &= F'(z_k)^{-1} [F'(z_k)(x_k - x^*) - F(x_k) + F(x^*)] \\ (2.17) \quad &= F'(z_k)^{-1} [F'(\frac{x_k + x^*}{2})(x_k - x^*) - F(x_k) + F(x^*)] + \\ &\quad F'(z_k)^{-1} [F'(z_k) - F'(\frac{x_k + x^*}{2})](x_k - x^*). \end{aligned}$$

In view of (2.4), (2.5), (2.7), (2.10), (2.11), (2.16) and (2.17) we obtain

$$\begin{aligned}
(2.18) \quad \|x_{k+1} - x^*\| &\leq c_k [b_k \|x_k - x^*\|^2 + d_k \|y_k - x^*\|] \|x_k - x^*\| \\
&\leq c(r) [b(r) + d(r)] \|x_k - x^*\| \\
&< \|x_k - x^*\| < r^*,
\end{aligned}$$

which shows (2.13) for $n = k$ and $x_{k+1} \in U(x^*, r^*)$.

By (1.2) we obtain the identity

$$\begin{aligned}
(2.19) \quad y_{k+1} - x^* &= F'(z_k)^{-1} \left[[F(x^*) - F(x_{k+1}) - F'(\frac{x^* + x_{k+1}}{2})(x^* - x_{k+1})] + \right. \\
&\quad \left. [F'(\frac{x^* + x_{k+1}}{2}) - F'(\frac{x_k + y_k}{2})](x^* - x_{k+1}) \right].
\end{aligned}$$

Using (2.4), (2.5), (2.7), (2.11), (2.12), (2.16) and (2.19) we get

$$\begin{aligned}
(2.20) \quad \|y_{k+1} - x^*\| &\leq c_k \left[b_{k+1} \|x_{k+1} - x^*\|^2 + d_k (\|x_{k+1} - x^*\| + \|y_k - x^*\| + \right. \\
&\quad \left. \|x_k - x^*\|) \right] \|x_{k+1} - x^*\| \\
&\leq c(r) [b(r) + 3d(r)] \|x_{k+1} - x^*\| \\
&< \|x_{k+1} - x^*\| < r^*,
\end{aligned}$$

which shows (2.14) for $n = k$ and $y_{k+1} \in U(x^*, r^*)$. Moreover by letting $k \rightarrow \infty$ in (2.17), and (2.19) we get $\lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} y_k = x^*$.

Finally, to show uniqueness let $y^* \in \overline{U}(x^*, r^*)$ be a solution of equation (1.1).

Using the identity

$$(2.21) \quad F(x^*) - F(y^*) = \mathcal{L}(x^* - y^*),$$

where,

$$(2.22) \quad \mathcal{L} = \int_0^1 F'(y^* + t(x^* - y^*)) dt,$$

and Lemma 2.2 for x replaced by $y^* + t(x^* - y^*)$ that \mathcal{L}^{-1} exists. Hence, by (2.20), we deduce $x^* = y^*$. That completes the proof of the theorem. \diamond

In order for us to determine the R -order of the Newton-type method (1.2) we need the Lemma :

Lemma 2.4. [4], [11]

Let $0 < \delta_0, \delta_1 < 1$, $p > 1$, $q \geq 0$, $c \geq 0$. If scalar sequence $\{\delta_n\}$ ($n \geq 0$) satisfies

$$(2.23) \quad 0 < \delta_{n+1} \leq c \delta_n^p \delta_{n-1}^q \quad (n \geq 1)$$

then it converges to zero with R -order of convergence given by

$$(2.24) \quad R(p, q) = \frac{p}{2} + \sqrt{\frac{p^2}{4} + q}.$$

Let us define functions g_1 , g_2 and g_3 on $[0, \frac{1}{\gamma})$ by

$$g_1(r) = \frac{(1-r)^2 r^2}{4(2(1-r)^2 - 1)(1-r)},$$

$$g_2(r) = \frac{r(1-r)^2}{4(2(1-r)^2 - 1)(1 - \frac{r}{2})(1-r)},$$

and

$$g_3(r) = g_1(r)r + \frac{3}{4} \frac{r(1-r)}{(1 - \frac{r}{2})(2(1-r)^2 - 1)}.$$

Set

$$(2.25) \quad \lambda_1 = g_1(r^*), \lambda_2 = g_2(r^*), \text{ and } \lambda_3 = g_3(r^*).$$

In view of (2.7), (2.13), (2.14) and (2.23) we get

$$(2.26) \quad \|x_{n+1} - x^*\| \leq \lambda_1 \|x_n - x^*\|^3 + \lambda_2 \|y_n - x^*\| \|x_n - x^*\|,$$

and

$$(2.27) \quad \|y_{n+1} - x^*\| \leq \lambda_3 \|x_{n+1} - x^*\| \|x_n - x^*\|.$$

It then follows from (2.24) and (2.25) that there exists $c > 0$ such that (2.21) holds true for $\delta_n = \|x_n - x^*\|$, $p = 1$ and $q = 1$. Hence, we arrived at:

Corollary 2.5. Under the hypotheses of Theorem 2.3, the Newton-type method (1.2) is of R -order of convergence $1 + \sqrt{2}$.

Remark 2.6. As noted in [1], [3], [4], [5], [7], [12] the local results obtained here can be used for projection method such as 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]:

$$(2.28) \quad F'(x) = P(F(x)),$$

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

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

$$(2.29) \quad F(x) = e^x - 1.$$

Then, note that we can set $P(x) = x + 1$ in (2.26).

We must have that conditions (2.3) and (2.4) hold for some $\gamma \geq 0$. It can easily be seen that we can set $\gamma = \frac{1}{2}$. Hence the radius of convergence is

$r^* = r_A = 2\left(\frac{5 - \sqrt{13}}{6}\right) = .464816242$. The radius of convergence r_W in [10] is given by

$$(2.30) \quad r_W = \frac{1}{2\gamma^*} (3 - 2\sqrt{2})$$

with

$$(2.31) \quad \gamma^* = \sup_{k \geq 2} \| F'(x^*)^{-1} F^{(k)}(x^*) \|^{1/(k-1)} \leq \frac{1}{2}.$$

Therefore, (2.27) gives

$$r_W \leq \sqrt{3} - 2\sqrt{2} = .171573.$$

Moreover, Rheinboldt radius [9] r_R is given by

$$(2.32) \quad r_R = \frac{2}{3l},$$

where l is the Lipschitz constant in condition:

$$(2.33) \quad \| F'(x^*)^{-1} (F'(x) - F'(y)) \| \leq l \| x - y \| \quad \text{for all } x, y \in D.$$

Using (2.27) and (2.30) we get: $l = e$. That is

$$r_R = .245252961.$$

The radius r_{WW} given by Werner in [11] is defined by

$$(2.34) \quad r_{WW} = \frac{2}{\Gamma l_1},$$

where

$$(2.35) \quad \| F'(x^*)^{-1} \| \leq \Gamma$$

and

$$(2.36) \quad \| F'(x) - F'(y) \| \leq l_1 \| x - y \|$$

hold true for all $x, y \in D$.

Hence, since $\Gamma = l_1 = e$, we get by (2.32)

$$(2.37) \quad r_{WW} = .270670566.$$

Hence, we deduce

$$(2.38) \quad r_W < r_R < r_{WW} < r_A.$$

By comparing r_A and r_W we see that it is always true that

$$(2.39) \quad r_W < r_A.$$

Moreover note that under (2.2) the existence of x^* in $U(x_0, \frac{1}{\gamma}(1 - \frac{1}{\sqrt{2}}))$ is guaranteed. However, in practice the existence of x^* may have been established by another way that avoids condition (2.2). Finally note that enlarging the radius of convergence is very important in computational

mathematics since in this case we can obtain a wider range of initial guesses x_0 .

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, pp. 160–169, (1986).
- [2] S. Amat, S. Busquier, Convergence and numerical analysis of a family of two-step Steffensen’s method, Comput. and Math. with Appl., 49, pp. 13–22, (2005).
- [3] I.K. Argyros, A unifying local–semilocal convergence analysis and applications for two-point Newton-like methods in Banach space, J. Math. Anal. and Appl., 298, pp. 374–397, (2004).
- [4] I.K. Argyros, Approximate solution of operator equations with applications, World Scientific Publ. Comp., New Jersey, U.S.A., (2005).
- [5] P.N. Brown, A local convergence theory for combined inexact–Newton/finite difference projection methods, SIAM J. Numer. Anal., 24, pp. 407–434, (1987).
- [6] J.M. Gutiérrez, M.A. Hernánadez, M.A. Salanova, Accessibility of solutions by Newton’s method, Inter. J. Comput. Math., 57, pp. 239–241, (1995).
- [7] L.V. Kantorovich, G.P. Akilov, Functional analysis in normed spaces, Pergamon Press, Oxford, (1982).
- [8] R.F. King, Tangent methods for nonlinear equations, Numer. Math., 18, pp. 298–304, (1972).
- [9] W.C. Rheinboldt, An adaptive continuation process for solving systems of nonlinear equations, Banach Center Publ., 3, pp. 129–142, (1975).

- [10] D. Wang, F. Zhao, The theory of Smale's point estimation and its applications, J. Comput. Appl. Math., 60, pp. 253–269, (1995).
- [11] W. Werner, Über ein verfahren ordnung $1 + \sqrt{2}$ zur Nullstellenbestimmung, Num. Math., 32, pp. 333–342, (1979).
- [12] T.J. Ypma, Local convergence of inexact Newton Methods, SIAM J. Numer. Anal., 21, pp. 583–590, (1984).

Ioannis K. Argyros

Cameron university
Department of Mathematics Sciences
Lawton, OK 73505
U. S. A.
e-mail : ioannisa@cameron.edu

and

Saïd Hilout

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