

A SEMILOCAL CONVERGENCE ANALYSIS FOR A CERTAIN CLASS OF MODIFIED NEWTON PROCESSES

IOANNIS K. ARGYROS

ABSTRACT. An error analysis introduced in [1], [2] is utilized in combination with nondiscrete mathematical induction to provide a finer than before [3]–[7], [11] semilocal convergence analysis for a certain class of modified Newton processes.

1. Introduction

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

$$(1) \quad F(x) = 0,$$

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 .

The most popular method for generating a sequence $\{x_n\}$ ($n \geq 0$) approximating x^* is undoubtedly Newton's method

$$(2) \quad x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad (n \geq 0), \quad (x_0 \in D),$$

where $F'(x) \in L(X, Y)$ the space of bounded linear operators from X into Y [2], [4].

A survey on local as well as semilocal convergence theorems for Newton method (1.2) can be found in [1], [2], [4], [12] and the references there. It is known that if x_0 is sufficiently close to x^* , sequence $\{x_n\}$ ($n \geq 0$) converges quadratically to x^* . A drawback in the application of this method is that at each step the inversion of linear $F'(x)$ is required. Such a task is very difficult in general. We can always replace method (1.2) by the modified Newton method

$$(3) \quad y_{n+1} = y_n - F'(y_0)^{-1}F(y_n) \quad (y_0 = x_0), \quad (n \geq 0).$$

Received June 18, 2007.

2000 *Mathematics Subject Classification.* 65H10, 65G99, 47H17, 49M15.

Key words and phrases. Newton's method, Banach space, Fréchet derivative, Nondiscrete Mathematical Induction, semilocal convergence .

However the convergence is only linear. In practice we prefer something in between. That is we use the iterative procedure of the form

$$\begin{aligned} x_{n+1}^0 &= x_n^m \\ (4) \quad x_{n+1}^{k+1} &= x_{n+1}^k - F'(x_n^m)^{-1}F(x_{n+1}^k), \quad k = 0, 1, \dots, m-1, \quad (n \geq 0). \end{aligned}$$

Clearly at each step the procedure requires the computation of m values of the operator F , the Fréchet derivative of F at one point and only one inversion of a linear operator. It follows from a result by J.W. Schmidt and H. Schwetlick [11] that method (1.4) has order of convergence equal to $m+1$.

Potra and Ptak in [7] using the method of Nondiscrete Mathematical Induction inaugurated by V. Ptak [9], [10] showed a semilocal convergence result for method (1.4) and also provided error bound on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x^*\|$ which are sharp in some sense. The convergence result essentially follows from a theorem of J. Dennis [3] but Potra's error bounds are new. In particular for $m=1$ this contains a theorem given in [6], whereas for $m=2$ his result improves a result obtained in [5].

The following standard conditions are used for all $x, y \in D$

$$\begin{aligned} (5) \quad & \|T_0[F'(x) - F'(y)]\| \leq \ell \|x - y\|, \quad \ell \neq 0, \\ (6) \quad & \|T_0F(x_0)\| \leq r_0, \\ (7) \quad & h_k = 2\ell r_0 \leq 1, \\ (8) \quad & \mu \geq r^1 = \frac{1 - \sqrt{1 - h_k}}{\ell}, \end{aligned}$$

and

$$(9) \quad \bar{U}(x_0, \mu) = \{x \in X \mid \|x - x_0\| \leq \mu\} \subseteq D.$$

Here we are motivated by Potra's paper and we show that by using more precise estimates and the same or even weaker hypotheses we can provide a finer semilocal convergence analysis.

In order for us to achieve this task we use a combination of (1.5) and the corresponding center-Lipschitz condition

$$(10) \quad \|T_0[F'(x) - F'(x_0)]\| \leq \ell_0 \|x - x_0\| \quad \text{for all } x \in D.$$

Note that in general

$$(11) \quad \ell_0 \leq \ell$$

holds and $\frac{\ell}{\ell_0}$ can be arbitrarily large [1], [2]. Note that the case $m=1$ has already been examined by us in [1], [2].

In order for us to make the study as self-contained as possible we provide some needed concepts on nondiscrete mathematical induction [7]–[9].

2. Nondiscrete Mathematical Induction

If m is a natural integer and S is a given set we denote by A^m the Cartesian product of m copies of A . If $y \in A^m$, $P_j y$ denotes its j -th component, so that

$$y = (P_1(y), P_2(y), \dots, P_j(y), \dots, P_m(y)).$$

For each $n \geq 0$ we define

$$g^{(0)}(s) = s,$$

and

$$g^{(n+1)}(s) = g(g^{(n)}(s)) \quad (n \geq 0), \quad (s \in S).$$

Let T be either $(0, +\infty)$ or $(0, b)$ with b positive. Let w be a function mapping T into T^m . We shall, sometimes write w_j for $p_j w$. For $w^{[1]} = w$, define $w^{[n]}$ by

$$w^{[n+1]} = w^{[n]} \circ w_m \quad (n \geq 1).$$

We need the definition:

Definition 2.1. A function $w: T \rightarrow T^m$ is called a rate of convergence of type $(1, m)$ on T if

$$(1) \quad \sum_{n=1}^{\infty} \left(\sum_{j=1}^m P_j w^{[n]}(r) \right) < \infty \quad \text{for all } r \in T.$$

Set $w_0(r) = r$ and for all $r \in T$ define

$$f = w_0 + w_1 + \dots + w_{m-1}.$$

Mapping w is a rate of convergence of type $(1, m)$ on T if and only if the series

$$\sigma(r) = f(r) + f(w_m(r)) + f(w_m^{(2)}(r)) + \dots$$

is convergent for all $r \in T$. Clearly we have

$$(2) \quad \sigma(w_m(r)) = \sigma(r) - f(r).$$

Using this notion Potra [7] constructed the multidimensional analogue of method (1.5) as follows, first let

$$\sigma_j = \sigma - (w_0 + w_1 + \dots + w_{j-1}), \quad j = 2, 3, \dots, m.$$

Let X be a Banach space, Q a mapping of X into X^m and let a starting point $x_0 \in X$ be given. Define a sequence $\{x_n\} \in X^m$ by:

$$(3) \quad x_1 = Q(x_0)$$

$$(4) \quad x_{n+1} = Q(P_m(x_n)) \quad n \geq 1.$$

We denote the components $P_j, Q(x)$ of $Q(x)$ by $Q_j(x)$ and set $Q_0 = I$. We also use the notation

$$x_0^m = x_0, \quad x_n^j = P_j(x_n) \quad (n \geq 1), \quad j = 1, 2, \dots, m.$$

The proofs of the Proposition and Lemma that follow needed for our main semilocal convergence theorem can be found in [7].

Proposition 2.2. *If we can attach to the pair (F, x_0) a $(1, m)$ rate of convergence w and a family of sets $Z(r) \subset X$, $r \in T$ such that the following conditions hold:*

$$(5) \quad x_0 \in Z(r_0) \text{ for some } r_0 \in T$$

$$(6) \quad x \in Z(r) \text{ implies } Q_m(x) \in Z(w_m(r))$$

and

$$(7) \quad d(Q_{j+1}(x), Q_j(x)) \leq w_j(r), \quad r \in T, \quad j = 0, 1, 2, \dots, m-1,$$

then there exists a point $x^* \in X$ with the following properties:

- (a) each of the m sequences $\{P_j x_n\}$, $j = 1, 2, \dots, m$ converges to x^* ;
- (b) the following estimates hold for all $n \geq 0$:

$$(8) \quad x_n^m \in Z(w_m^{[n]}(r_0))$$

$$(9) \quad d(x_{n+1}^{j+1}, x_{n+1}^j) \leq P_j w^{[n+1]}(r_0), \quad j = 1, 2, \dots, m-1$$

$$(10) \quad d(x_{n+1}^1, x_n^m) \leq w_m^{[n]}(r_0),$$

$$(11) \quad d(x_n^m, x^*) \leq \sigma(w_m^{(n)}(r_0)),$$

$$(12) \quad d(x_{n+1}^j, x^*) \leq \sigma_j, \quad j = 1, 2, \dots, m-1;$$

(c) If

$$(13) \quad x_{n-1}^m \in Z(d(x_n^1, x_{n-1}^m))$$

for some n , then, for this n , the following hold

$$(14) \quad d(x_n^j, x^*) \leq \sigma_j(d(x_n^1, x_{n-1}^m)), \quad j = 1, 2, \dots, m.$$

Lemma 2.3. *Let $a \geq 0$ and m a natural integer be given. Let T be the set of all positive real numbers. Define functions f , w_j , $j = 0, 1, \dots, m$ for $r \in T$ by*

$$(15) \quad f(r) = r + \sqrt{r^2 + a^2}$$

$$(16) \quad w_0(r) = r$$

$$(17) \quad w_{k+1} = \frac{w_k}{2f} [w_k + 2(w_0 + \dots + w_{k-1})], \quad 0 \leq k \leq m-2$$

and

$$(18) \quad w_m = \frac{w_{m-1}[w_{m-1} + 2(w_0 + \dots + w_{m-2})]}{2[f - (w_0 + \dots + w_{m-1})]},$$

respectively.

Then functions w_j , $j = 1, 2, \dots, m$ constitute a rate of convergence of type $(1, m)$ on T and the corresponding σ function is given by

$$(19) \quad \sigma(r) = r + \sqrt{r^2 + a^2} - a.$$

In [1] we provided a series of weaker conditions for the convergence of method (1.2). For simplicity we only state one of them

$$(20) \quad h_A^\delta = (\ell + \delta \ell_0)r_0 \leq \delta, \quad \delta \in [0, 1].$$

In view of (1.11) note that e.g.

$$(21) \quad h_k \leq 1 \Rightarrow h_A^1 \leq 1$$

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

We can state the following semilocal result for Newton's method (1.2) [1], [2]:

Theorem 2.4. *If conditions (1.5), (1.6), (1.9) (for $\mu \geq r^*$), (1.10), and (2.19) hold, then Newton's method (1.2) converges to a unique solution x^* of equation $F(x) = 0$ in $\bar{U}(x_0, r^*)$, where*

$$(22) \quad r^* = \lim_{n \rightarrow \infty} \bar{r}_n \leq \frac{2r_0}{2 - \delta},$$

and

$$(23) \quad \bar{r}_0 = 0, \quad \bar{r}_1 = r_0, \quad \bar{r}_{n+2} = \bar{r}_{n+1} + \frac{\ell(\bar{r}_{n+1} - \bar{r}_n)^2}{2(1 - \ell_0 \bar{r}_{n+1})} \quad (n \geq 0).$$

Moreover if (1.7) holds then

$$(24) \quad r^* \leq r^1,$$

where r^1 is given by (1.8).

We can show the main semilocal convergence theorem for method (1.4):

Theorem 2.5. *Under hypotheses of Theorem 2.4 for $\delta = 1$, $\mu \geq r^*$, further assume there exists $a \in (0, \frac{1}{\ell}]$ such that*

$$(25) \quad \ell(\sigma(r_0) + a) \leq 1,$$

and

$$(26) \quad \sigma(r_0) \leq r^*.$$

Then sequence $\{x_n\}$ generated by method (1.4) is well defined, remains in $\bar{U}(x_0, r^*)$ for all $n \geq 0$ and converges to a unique solution x^* of equation $F(x) = 0$ in $\bar{U}(x_0, r^*)$.

Moreover the following estimates hold:

$$(27) \quad \|x_n^j - x^*\| \leq \sigma_j(w_m^{(n-1)}(r_0)),$$

and

$$(28) \quad \|x_n^j - x^*\| \leq \sigma_j(\|x_n^1 - x_{n-1}^m\|)$$

for all $j = 1, 2, \dots, m$ and $n \geq 1$.

Proof. Iterative method (1.4) can be written in the form (2.3) provided that mapping Q is defined:

$$\begin{aligned}
 Q_0(x) &= x, \\
 Q_1(x) &= x - F'(x)^{-1}F(x) \\
 Q_2(x) &= Q_1(x) - F'(x)^{-1}F(Q_1(x)) \\
 (29) \quad &\dots \quad \dots \quad \dots \\
 Q_m(x) &= Q_{m-1}(x) - F'(x)^{-1}F(Q_{m-1}(x)) \\
 Q(x) &= (Q_1(x), Q_2(x), \dots, Q_m(x)).
 \end{aligned}$$

We also define $Z(\cdot)$ the family of subsets of X by

$$(30) \quad Z(r) = \{x \in X \mid \|x - x_0\| \leq \sigma(r_0) - \sigma(r), F'(x) \text{ is invertible and } \|F'(x)^{-1}F(x)\| \leq r\}.$$

We shall show that all hypotheses of Proposition 2.2 hold true. In view of the initial condition (1.6) and (2.29) we deduce $x_0 \in Z(r_0)$. Let us assume $x \in Z(r)$. We must show

$$(31) \quad \|Q_i(x) - Q_{i+1}(x)\| \leq w_i(r), \quad i = 0, 1, \dots, m-1.$$

For $i = 0$, (2.30) becomes $\|F'(x)^{-1}F(x)\| \leq r$, which follows from (2.29). For $m > 1$ assume (2.30) holds true for $i = 0, 1, 2, \dots, k, k \leq m-2$.

In view of (2.28) we have

$$\begin{aligned}
 (32) \quad \|Q_{k+2}(x) - Q_{k+1}(x)\| &= \|F'(x)^{-1}F(Q_{k+1}(x))\| \\
 &\leq \|(I - T_0(F'(x_0) - F'(x)))^{-1}\| \cdot \|T_0F(Q_{k+1}(x))\|.
 \end{aligned}$$

Using (1.10) and (2.24) we get

$$\begin{aligned}
 (33) \quad \|T_0(F'(x_0) - F'(x))\| &\leq \ell_0\|x - x_0\| \\
 &\leq \ell_0(\sigma(r_0) - \sigma(r)) \leq \ell(\sigma(r_0) - \sigma(r)) < 1.
 \end{aligned}$$

It follows from (2.32) and the Banach Lemma on invertible operators [4] that $[I - T_0(F'(x_0) - F'(x))]^{-1}$ exists and

$$\begin{aligned}
 (34) \quad \|[I - T_0(F'(x_0) - F'(x))]^{-1}\| &\leq \frac{1}{1 - \ell_0(\sigma(r_0) - \sigma(r))} \leq \frac{1}{\ell f(r)} \text{ by (2.24)}.
 \end{aligned}$$

We can also write

$$\begin{aligned}
 (35) \quad F(Q_{k+1}(x)) &= F(Q_{k-1}(x)) - F(Q_k(x)) - F'(x)[Q_{k+1}(x) - Q_k(x)],
 \end{aligned}$$

which leads to

$$\begin{aligned}
 & \|T_0 F(Q_{k+1}(x))\| \\
 & \leq \|T_0(F(Q_{k+1}(x)) - F(Q_k(x)) - F'(Q_k(x))(Q_{k+1}(x) - Q_k(x)))\| \\
 & \quad + \|T_0(F'(Q_k(x)) - F'(x))(Q_{k+1}(x) - Q_k(x))\| \\
 & \leq \frac{1}{2}\ell\|Q_{k+1}(x) - Q_k(x)\|^2 + \ell\|Q_k(x) - x\|\|Q_{k+1}(x) - Q_k(x)\| \\
 (36) \quad & = \frac{1}{2}\ell w_k(r)[w_k(r) + 2(w_0(r) + \dots + w_{k-1}(r))].
 \end{aligned}$$

It follows from (2.31), (2.34) and (2.35) that (2.30) holds true for $i = k + 1$.

We shall show $Q_m(x) \in Z(w_m(r))$. In view of (2.30) we obtain

$$(37) \quad \|Q_m(x) - x\| \leq \sum_{i=0}^{m-1} w_i(r) = f(r)$$

and by (2.2)

$$(38) \quad \|Q_m(x) - x_0\| \leq \|Q_m(x) - x\| + \|x - x_0\| \leq \sigma(r_0) - \sigma(w_m(r)).$$

As in (2.32) we have

$$\begin{aligned}
 & \|T_0(F'(x_0) - F'(Q_m(x)))\| \leq \ell_0\|Q_m(x) - x_0\| \\
 (39) \quad & \leq \ell_0(\sigma(r_0) - \sigma(w_m(r))) < 1.
 \end{aligned}$$

That is $F'(Q_m(x))$ is invertible. By (2.35) and (2.37) we get

$$\begin{aligned}
 & \|F'(Q_m(x))^{-1}F(Q_m(x))\| \\
 & \leq \|[I - T_0(F'(x_0) - F'(Q_m(x)))]^{-1}\| \|T_0F(Q_m(x))\| \\
 (40) \quad & \leq \frac{\ell w_{m-1}(r)[w_{m-1}(r) + 2(w_0(r) + \dots + w_{m-2}(r))]}{2\ell f(r)} = w_m(r),
 \end{aligned}$$

which shows (2.5) and (2.6).

According to (b) of Proposition 2.2 we have

$$(41) \quad x_{n-1}^m \in Z(w_m^{(n-1)}(r_0)),$$

and

$$(42) \quad \|x_{n-1}^m - x_n^1\| \leq w_m^{(n-1)}(r_0)$$

for all $n \geq 1$.

Therefore we obtain by the monotonicity of function σ , (2.40) and (2.41):

$$(43) \quad \|x_{n-1}^m - x_0\| \leq \sigma(r_0) - \sigma(w_m^{(n-1)}(r_0)) \leq \sigma(r_0) - \sigma(\|x_{n-1}^m - x_n^1\|),$$

which shows

$$(44) \quad x_{n-1}^m \in Z(\|x_{n-1}^m - x_n^1\|) \quad (n \geq 1).$$

Moreover (2.40) implies that (2.35) holds for $x = x_{n-1}^m$ and $k = m - 1$. Furthermore using (2.1), the invertibility of T_0 and the continuity of F on $\bar{U}(x_0, r^*)$ we get $F(x^*) = 0$.

Finally the uniqueness part is omitted as identical to the one in the corresponding theorem in [1], [2].

That completes the proof of the theorem. \square

Remark 2.6. (a) If $\ell_0 = \ell$, conditions (2.24) and (2.25) hold true for

$$(45) \quad r^* = r^1$$

and

$$(46) \quad a = \frac{\sqrt{1 - h_k}}{\ell}$$

and our Theorem 2.5 reduces to Theorem 4.1 in [7, p. 115].

(b) If (1.7) is violated but (2.19) (for $\delta = 1$ or not) holds then our Theorems 2.4 and 2.5 can be used in cases where Theorem 4.1 in [7] cannot.

(c) If (1.7) and (2.19) (for say $\delta = 1$) hold but (1.11) is true as a strict inequality then it follows by (2.23) and (2.32) (or (2.34)) that our estimates on the distances $\|x_n - x^*\|$, $\|x_{n+1} - x_n\|$ ($n \geq 0$) and the information on the location of the solution x^* are at least as precise as the corresponding ones in [7]. In practice we may want to replace r^* by $\frac{2r_0}{2-\delta}$ although it can be compared exactly using (2.22).

(d) Theorem 4.1 in [7] can also be improved if $\ell_0 < \ell$ since the second from the right holds as strict inequality in (2.32). The inequalities that follow using (2.32) also hold as strict inequalities which implies that under the hypotheses of this theorem at least estimate (2.41) holds as a strict inequality.

Note that our results are obtained under the same computational cost since in practice the computation of Lipschitz constant ℓ requires that of ℓ_0 .

Hence we have justified the claims made in the Introduction about the usefulness of our approach.

References

- [1] Argyros, I. K., *A unifying local-semilocal convergence analysis and applications for two-point Newton like methods in Banach space*, J. Math. Anal. Applic. **298** (2004), 374–397.
- [2] Argyros, I.K., *Approximate Solution of Operator Equations with Applications*, World Scientific Publ. Co., River Edge, NJ, 2005.
- [3]] Dennis, J., *On the Kantorovich hypothesis for Newton's method*, SIAM J. Numer. Anal. **6** (1969), 493–507.
- [4] Kantorovich, L.V. and Akilov, G.P., *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.
- [5] Nastas, N.K., *The convergence of a certain third order process for the solution of non-linear operator equations* (in Russian), Approximate Solution of Equations, "Stiinca" Kishinev, 1973, pp. 70–75.
- [6] Potra, F.A. and Ptak, V., *Sharp error bounds for Newton's method*, Numer. Math. **34** (1980), 63–72.
- [7] Potra, F.A. and Ptak, V., *On a class of modified Newton processes*, Numer. Funct. Anal. and Optimiz. **2**, 1 (1980), 107–120.
- [8] Potra, F.A. and Ptak, V., *Nondiscrete Induction and Iterative Processes*, Pitman Publ., London, 1984.

- [9] Ptak, V., *Nondiscrete mathematical induction and iterative existence proofs*, Linear Algebra and its Applications, **13** (1976), 223–236.
- [10] Ptak, V., *The rate of convergence of Newton's process*, Numer. Math. **25** (1976), 279–285.
- [[11]] Schmidt, J.W. and Schwetlick, H., *Ableitungsfreie Verfahren mit höherer Konvergenzgeschwindigkeit*, Computing, **3** (1968), 215–226.
- [11] Yamamoto, T., *A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions*, Numer. Math. **44** (1986), 203–220.

CAMERON UNIVERSITY, DEPARTMENT OF MATHEMATICAL SCIENCES, LAWTON, OK 73505,
USA

E-mail address: `iargyros@cameron.edu`