

ON THE SOLUTION OF NONLINEAR EQUATIONS CONTAINING A NON-DIFFERENTIABLE TERM

IOANNIS K. ARGYROS

ABSTRACT. We approximate a locally unique solution of a nonlinear operator equation containing a non-differentiable operator in a Banach space setting using Newton's method. Sufficient conditions for the semilocal convergence of Newton's method in this case have been given by several authors using mainly increasing sequences [1]–[6]. Here, we use center as well as Lipschitz conditions and decreasing majorizing sequences to obtain new sufficient convergence conditions weaker than before in many interesting cases. Numerical examples where our results apply to solve equations but earlier ones cannot [2], [5], [6] are also provided in this study.

1. Introduction

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

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

where, for \mathcal{X} , \mathcal{Y} being Banach spaces, and \mathcal{D} an open convex subset of \mathcal{X} , $F : \mathcal{X} \rightarrow \mathcal{Y}$ is a Fréchet-differentiable operator, and $G : \mathcal{X} \rightarrow \mathcal{Y}$ is a continuous operator.

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 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,

Received March 29, 2008; Accepted May 16, 2008.

2000 *Mathematics Subject Classification.* 65K10, 65G99, 65J99, 49M15, 47J20.

Key words and phrases. Banach space, Fréchet-derivative, majorizing sequence, semilocal convergence, Newton's method, successive approximations.

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.

The method of successive approximations or Newton-type methods:

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

has been used by many authors to generate a sequence $\{x_n\}$ approximating x^* . A survey of convergence results for Newton's method (1.2) under mainly Lipschitz-type conditions and increasing majorizing sequences can be found in [4] (see also [1]–[3], [5], [6], and Section 2).

In the special case when $G = 0$, method (1.2) reduces to the classical Newton's method, while in the case when F is linear and $G'(x_0) = 0$, method (1.2) becomes the modified Newton's method [1]–[6].

The sufficient convergence conditions for method (1.2) are not optimal, only necessary, so that one can provide examples (see Example 3.1) where method (1.2) converges but the existing conditions do not hold.

Motivated by this observation and optimization considerations, we provide a new semilocal convergence analysis for method (1.2) using: center-Lipschitz conditions (instead of the less accurate Lipschitz conditions used in earlier works [1], [2], [5], [6] for the computation of the upper bounds on $\|F'(x_n)^{-1} F'(x_0)\|$; Lipschitz conditions and decreasing instead of increasing majorizing sequences. It turns out that this way weaker sufficient convergence conditions are obtained in many interesting cases.

Numerical examples are also provided to show that our results can apply to solve equations in cases the earlier results referred to above cannot.

2. Semilocal convergence analysis of method (1.2)

We need the following result on majorizing sequences for method (1.2)

Lemma 2.1. *Let $\eta \geq 0$, $l_0 > 0$, $l > 0$, and $l_1 \geq 0$ be given constants. Set*

$$t_0 = \frac{1}{l}.$$

Define functions Δ , A , C on $[0, +\infty)^2$, and B on $[0, +\infty)$ by

$$\Delta(t, \gamma) = \left(l t + l_1 \right)^2 - (l - 2 \gamma l_0) t \left(l t + 2 l_1 \right),$$

$$A(t, \gamma) = 2 \left(l t + l_1 + \sqrt{\Delta(t, \gamma)} \right),$$

$$B(t) = 2 \left(l t + 2 l_1 \right),$$

and

$$C(t, \gamma) = \frac{B(t)}{A(t, \gamma)}.$$

Assume that the function f_t defined by

$$(2.1) \quad f_t(x) = 1 - x - C(t, x)$$

has a non-negative zero $\gamma_0 = \gamma(t_0)$ at $t = t_0$, such that:

$$(2.2) \quad \beta = 2 l_0 \eta \leq \gamma_0 \leq \frac{l}{2 l_0}$$

or if

$$(2.3) \quad f_{t_0}(\beta) \geq 0,$$

and (2.2) hold;

or if function f_t has a non-negative zero γ_0^1 at $t = t_1$, such that

$$(2.4) \quad \gamma_0^1 \leq \frac{l}{2 l_0},$$

and

$$(2.5) \quad \beta < \beta_0,$$

where

$$(2.6) \quad \beta_0 = 4 \left[l_1 + 2 + \sqrt{(l_1 + 2)^2 + 4 \left(\frac{l}{2 l_0} - 1 \right)} \right]^{-1};$$

or if

$$(2.7) \quad f_{t_1}(\beta_1) \leq 0, \quad \beta_1 = \min \left\{ \beta_0, \frac{l}{2 l_0} \right\},$$

$$(2.8) \quad f_{t_1}(\beta) \geq 0,$$

(2.4) and (2.5) hold.

Note that the existence of γ_0 or γ_0^1 follows from the intermediate value theorem applied to function f_{t_0} or f_{t_1} on the interval $\left[\beta, \frac{l}{2 l_0} \right]$ or $[\beta, \beta_1]$ (for $\beta < \beta_1$),

respectively.

Then, scalar sequence $\{t_n\}$ ($n \geq 0$) generated by

$$\begin{aligned}
 t_0 &= \frac{1}{l_0}, & t_1 &= t_0 - \eta, \\
 (2.9) \quad t_{n+1} &= t_n - \frac{\left(l (t_{n-1} - t_n) + 2 l_1 \right) (t_{n-1} - t_n)}{2 l_0 t_n} \quad (n \geq 1),
 \end{aligned}$$

is well defined, decreasing, and converges to some $t^* \in [0, t_0]$.

Proof. If $\eta = 0$, then $t_n = t_0 = t^*$ ($n \geq 1$). Let us assume $\eta \neq 0$. Function Δ is a quadratic polynomial with leading coefficient $2 l_0 \gamma_0$, and whose sign of the discriminant is the same with: $2 \gamma_0 l_0 (2 \gamma_0 l_0 - l)$. It then follows by (2.2) that functions A and C are well defined. It also follows by definition of γ_0 that $\gamma_0 \in (0, 1)$. We can set:

$$\frac{t_{n+1}}{t_n} = 1 - \gamma_n,$$

where,

$$\gamma_n = \gamma(t_n) = \frac{\left(l (t_{n-1} - t_n) + 2 l_1 \right) (t_{n-1} - t_n)}{2 l t_n^2} \quad (n \geq 1).$$

We shall show: $t_k \geq (1 - \gamma_0) t_{k-1}$, which together with $t_{k-1} > 0$ implies implies $0 < t_k < t_{k-1}$. But, $t_k \geq (1 - \gamma_0) t_{k-1}$ holds if $1 - \gamma_k \geq 1 - \gamma_0$ or $\gamma_k \leq \gamma_0$ or

$$(l - 2 l_0 \gamma_0) t_k^2 - 2 \left(l t_{k-1} + l_1 \right) t_k + \left(l t_{k-1} + 2 l_1 \right) t_{k-1} \leq 0$$

or $t_k \geq C(t_{k-1}, \gamma_0) t_{k-1}$. Using (2.2), (2.9), we get:

$$t_1 \geq (1 - \gamma_0) t_0 \implies t_1 \geq C(t_0, \gamma_0) t_0 \iff t_2 \geq (1 - \gamma_0) t_1.$$

Similary, if (2.5)–(2.8) hold, then $t_2 > 0$, and $t_3 \geq (1 - \gamma_0^1) t_2$. By analogy, we show:

$$t_{k-1} \geq (1 - \gamma_0) t_{k-2} \implies t_{k-1} \geq C(t_0, \gamma_0) t_{k-2} \iff t_k \geq (1 - \gamma_0) t_{k-1} \quad (k \geq 1).$$

The induction is completed. Hence, sequence $\{t_n\}$ ($n \geq 0$) is decreasing positive, and as such it converges to some $t^* \in [0, t_0]$.

That completes the proof of Lemma 2.1. □

We can show the following semilocal convergence theorem for method (1.2):

Theorem 2.2. *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be a Fréchet-differentiable operator, and $G : \mathcal{D} \longrightarrow \mathcal{Y}$ be a continuous operator.*

Assume there exist $x_0 \in \mathcal{D}$, and constants $\eta \geq 0$, $l_0 > 0$, $l > 0$, and $l_1 > 0$, such that for $x, y \in \mathcal{D}$, the following conditions hold:

$$(2.10) \quad F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X});$$

$$(2.11) \quad \| F'(x_0)^{-1} (F(x_0) + G(x_0)) \| \leq \eta;$$

$$(2.12) \quad \| F'(x_0)^{-1} (F'(x) - F'(y)) \| \leq l \| x - y \|;$$

$$(2.13) \quad \| F'(x_0)^{-1} (F'(x) - F'(x_0)) \| \leq l_0 \| x - x_0 \|;$$

$$(2.14) \quad \| F'(x_0)^{-1} (G(x) - G(y)) \| \leq l_1 \| x - y \|;$$

$$(2.15) \quad \bar{U}(x_0, t_0 - t^*) = \{x \in \mathcal{X} : \| x - x_0 \| \leq t_0 - t^*\} \subseteq \mathcal{D},$$

where t^* is given in Lemma 2.1, and hypotheses of Lemma 2.1 hold.

Then, sequence $\{x_n\}$ ($n \geq 0$), generated by method (1.2) is well defined, remains in $\bar{U}(x_0, t_0 - t^*)$ for all $n \geq 0$, and converges to a solution $x^* \in \bar{U}(x_0, t_0 - t^*)$ of equation $F(x) + G(x) = 0$.

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

$$(2.16) \quad \| x_{n+1} - x_n \| \leq t_n - t_{n+1}$$

and

$$(2.17) \quad \| x_n - x^* \| \leq t_n - t^*.$$

Furthermore, if

$$(2.18) \quad (l + 2 l_0) (t_0 - t^*) + 2 l_1 < 2,$$

the solution is unique in $\bar{U}(x_0, t_0 - t^*)$.

Finally, if there exists $t^{**} > t^*$, such that

$$(2.19) \quad l (t_0 - t^{**}) - 2 l_0 (t^* - t_0) + 2 l_1 \leq 2,$$

then the solution x^* is unique in $U(x_0, t_0 - t^{**})$.

Proof. We shall show by induction on $n \geq 0$ that (2.16) holds. Estimates (2.17) will then follow from (2.16) using standard majorization techniques [3], [4], [6]. Since

$$\| x_1 - x_0 \| \leq \eta = t_0 - t_1,$$

(2.16) holds for $n = 0$, and we have by (2.13):

$$(2.20) \quad \begin{aligned} \| F'(x_0)^{-1} (F'(x_1) - F'(x_0)) \| &\leq l_0 \| x_1 - x_0 \| \\ &\leq l_0 (t_0 - t_1) < 1. \end{aligned}$$

It follows from (2.20), and the Banach lemma on invertible operators [3], [4] that $F'(x_1)^{-1}$ exists, and

$$(2.21) \quad \begin{aligned} \| F'(x_1)^{-1} F'(x_0) \| &\leq (1 - l_0 \| x_1 - x_0 \|)^{-1} \\ &\leq (1 - l_0 (t_0 - t_1))^{-1}. \end{aligned}$$

Assume that for $1 \leq k \leq n$: $\| x_k - x_{k-1} \| \leq t_{k-1} - t_k$. Then

$$\| x_k - x_0 \| \leq t_0 - t_k \leq t_0 - t^*.$$

As in (2.20) for x_k replacing x_1 , we obtain $F'(x_k)^{-1}$ exists, and

$$(2.22) \quad \| F'(x_k)^{-1} F'(x_0) \| \leq (1 - l_0 \| x_k - x_0 \|)^{-1} \leq (1 - l_0 (t_0 - t_k))^{-1}.$$

Using (1.2), we obtain [3], [4]:

$$\begin{aligned}
 (2.23) \quad & \| x_{k+2} - x_{k+1} \| \\
 & \leq \| F'(x_{k+1})^{-1} F'(x_0) \| \left(\int_0^1 \| F'(x_0)^{-1} F'(x_k + t(x_{k+1} - x_k)) \right. \\
 & \quad \left. - F'(x_k) \| dt + \| F'(x_0)^{-1} (G(x_{k+1}) - G(x_k)) \| \right) \| x_{k+1} - x_k \| \\
 & \leq \frac{(l \| x_{k+1} - x_k \| + 2 l_1) \| x_{k+1} - x_k \|}{2 (1 - l_0 \| x_{k+1} - x_0 \|)} \\
 & \leq \frac{(l (t_k - t_{k+1}) + 2 l_1) (t_k - t_{k+1})}{2 (1 - l_0 (t_0 - t_{k+1}))} = t_{k+1} - t_{k+2},
 \end{aligned}$$

which completes the induction for (2.16).

It follows by Lemma 2.1, and (2.23) that $\{x_n\}$ is a Cauchy sequence in a Banach space \mathcal{X} , and as such it converges to some $x^* \in \bar{U}(x_0, t_0 - t^*)$ (since $\bar{U}(x_0, t_0 - t^*)$ is a closed set). By letting $k \rightarrow \infty$ in (2.23) we obtain $F(x^*) + G(x^*) = 0$.

To show uniqueness of the solution first in $\bar{U}(x_0, t_0 - t^*)$, let: $y^* \in \bar{U}(x_0, t_0 - t^*)$ be a solution of equation (1.1).

Using (1.2), we obtain the identity:

$$\begin{aligned}
 (2.24) \quad & x_{k+1} - y^* = x_k - y^* - F'(x_k)^{-1} (F(x_k) + G(x_k)) \\
 & = -(F'(x_k)^{-1} F'(x_0)) \left(F'(x_0)^{-1} (F(x_k) - F(y^*)) \right. \\
 & \quad \left. - F'(x_k) (x_k - y^*) + F'(x_0)^{-1} (G(x_k) - G(y^*)) \right).
 \end{aligned}$$

By (2.12), (2.14), (2.22), and (2.24), we get in turn:

$$\begin{aligned}
 (2.25) \quad & \| x_{k+1} - y^* \| \\
 & \leq \| F'(x_k)^{-1} F'(x_0) \| \\
 & \quad \left(\| F'(x_0)^{-1} (F(x_k) - F(y^*) - F'(x_k) (x_k - y^*)) \| \right. \\
 & \quad \left. + \| F'(x_0)^{-1} (G(x_k) - G(y^*)) \| \right). \\
 & \leq \frac{(l \| x_k - y^* \| + 2 l_1) \| x_k - y^* \|}{2 (1 - l_0 \| x_k - x_0 \|)} \\
 & < \| x_k - y^* \| \quad (\text{by (2.18)}).
 \end{aligned}$$

In view of (2.25), we obtain $\lim_{k \rightarrow \infty} x_k = y^*$. But we also showed $\lim_{k \rightarrow \infty} x_k = x^*$. hence, we deduce $x^* = y^*$.

Finally, if $y^* \in U(x_0, t^*)$, using (2.19), we again deduce $\|x_{k+1} - y^*\| < \|x_k - y^*\|$, which also implies $x^* = y^*$.

That completes the proof of Theorem 2.2. □

Remark 2.3. *The conclusions on the uniqueness of the solution x^* hold if in (2.18) or (2.19), t^* is replaced by t_0 .*

We state the following semilocal convergence theorem for method (1.2) for comparison purposes. The proof can be found in [3], [4], [6]:

Theorem 2.4. *Let $F : \mathcal{D} \subseteq \mathcal{X} \rightarrow \mathcal{Y}$ be a Fréchet-differentiable operator, and $G : \mathcal{D} \rightarrow \mathcal{Y}$ be a continuous operator.*

Assume there exist $x_0 \in \mathcal{D}$, and constants $\eta \geq 0$, $l > 0$, and $l_1 > 0$, such that (2.10), (2.11), (2.12), (2.14), and the following hold:

$$\bar{U}(x_0, r^*) \subseteq \mathcal{D},$$

and

$$(2.26) \quad (l + l_1) \eta \leq \frac{1}{2},$$

where

$$(2.27) \quad r^* = \lim_{n \rightarrow \infty} r_n,$$

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

$$(2.28) \quad r_{n+1} = r_n - \frac{\left(l (r_n - r_{n-1}) + 2 l_1 \right) (r_n - r_{n-1})}{2 l r_n} \quad (n \geq 1).$$

Then, sequence $\{x_n\}$ ($n \geq 0$), generated by method (1.2) is well defined, remains in $\bar{U}(x_0, r^)$ for all $n \geq 0$, and converges to a unique solution of equation (1.1) in $\bar{U}(x_0, r^*)$, so estimates (2.16) and (2.17) hold with sequence $\{t_n\}$ replaced by $\{r_n\}$.*

Remark 2.5. *Using Mathematica, we found that the value of γ_0 is given by:*

$$(2.29) \quad \gamma_0 = a + \left(\frac{b}{32^{2/3} (c+d)^{1/3}} + \frac{(c+d)^{1/3}}{62^{1/3}} \right) \frac{1}{l_0 t_0},$$

where

$$a = \frac{l + 4 l_0}{6 l_0},$$

$$b = \left((l + 4 l_0)^2 t_0 + 12 l_0 (l_1 - l_0 t_0) \right) t_0,$$

$$c = 2 \left(18 (l t_0 + l_1) l_0 l + 54 (l t_0 + 2 l_1) l_0^2 - 36 (l t_0 + l_1) l_0^2 + l^3 t_0 - 6 l^2 l_0 t_0 + 12 l l_0^2 t_0 - 8 l^3 t_0 \right) t_0,$$

and

$$d = (c^2 - 4 b^3)^{2/3}.$$

In the interesting case when $l_1 = 0$ (i.e. $G = 0$) (Newton's method), we get use (2.29)

$$(2.30) \quad \gamma_A := \gamma_0 = \gamma_0^1 = \frac{\sqrt{l^2 + 8 l_0 l} - l}{\sqrt{l^2 + 8 l_0 l} + l}.$$

Note that in this case $\gamma_0 \in (0, \frac{1}{2}]$, with $\gamma_0 = \frac{1}{2}$, when $l_0 = l$.

3. Special cases and applications

A direct comparison between Theorems 2.2 and 2.4 is not possible in general, since the sufficient conditions are not the same and majorizing sequence $\{t_n\}$ is decreasing, whereas $\{r_n\}$ is increasing. However a comparison is possible in some interesting special cases.

1. Case $G = 0$. The sufficient convergence conditions of Theorems 2.2 and 2.4 becomes respectively

$$(3.1) \quad h_A = l_0 \eta \leq \gamma_A,$$

and

$$(3.2) \quad h_K = l \eta \leq \frac{1}{2},$$

Note that

$$(3.3) \quad l_0 \leq l$$

hold in general, and $\frac{l}{l_0}$ can be arbitrarily large [3], [4]. It also follows from (2.30), (3.1)–(3.3) that conditions (3.1), and (3.2) coincide if $l_0 = l$. Otherwise (3.1) improves (3.2), since

$$(3.4) \quad \frac{l}{l_0} \gamma_A > \frac{1}{2} \quad \text{for } l_0 \neq l.$$

Note that this improvement is obtained under the computational cost, since in practice the computation of constant l also requires the computation of l_0 .

2. Case $G \neq 0$. In this case we can only compare Theorem 2.2 with Theorem 2.4 using numerical examples.

Example 3.1. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, $x_0 = 1$, $\mathcal{D} = [\delta, 2 - \delta]$, $\delta \in [0, \frac{1}{2})$, and define function F and G on \mathcal{D} by

$$(3.5) \quad F(x) = x^3 - \delta \quad \text{and} \quad G(x) = \epsilon |x - 1|,$$

where, ϵ is a given real number.

Using (2.11)–(2.14), and (3.5), we obtain:

$$\eta = \frac{1}{3} (1 - \delta), \quad l_0 = 3 - \delta, \quad l = 2 (2 - \delta), \quad \text{and} \quad l_1 = |\epsilon|.$$

Note that function G is not differentiable at $x_0 = 1$.

Hypothesis (3.2) is violated, since

$$(3.6) \quad 4 (2 - \delta) \left(\frac{1}{3} (1 - \delta) + |\epsilon| \right) > 1 \quad \text{for all } \epsilon, \quad \text{and} \quad \delta \in [0, \frac{1}{2}).$$

That is there is no guarantee that sequence $\{x_n\}$ converges to x^* .

However, our Theorem 2.2 can apply to solve equation $F(x) + G(x) = 0$.

Let us consider two cases:

1. Case $\epsilon = 0$. The conditions of Theorem 2.2 are satisfied when (3.1) holds, where γ_0 is given by (2.30). It then follows that (3.1) hold for $\delta \in [.450339002, \frac{1}{2}]$, which coincides with the range found by us in [3] for this case but using a different approach.

2. Case $\epsilon \neq 0$. Choose e.g.: $\epsilon = .1$, and $\delta = .49$. Then, we get:

$$\eta = .17, \quad l_0 = 2.51, \quad l = 3.02, \quad l_1 = .1, \quad t_0 = .398406374, \\ t_1 = .22840637, \quad \gamma_0 = .410812, \quad \gamma_0^1 = .369936,$$

$$\beta = .8534, \quad \frac{l}{2 l_0} = .601593625, \quad \text{and} \quad \beta_0 = 1.058703597.$$

Hence hypotheses (2.5) and (2.6) are satisfied. That is the conclusions of Theorem 2.2 apply to solve equation (1.1), which F , G given by (3.5).

References

- [1] J. Appel, E. De Pascale, N.A. Evkuta, P.P. Zabrejko, *On the two-step Newton method for the solution of nonlinear operator equations*, Math. Nachr. **172** (1995), 5–14.
- [2] I.K. Argyros, *On the solution of equations with non differentiable operators and the Pták error estimates*, BIT **30** (1990), 752–754.
- [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** (2004), 374–397.
- [4] I.K. Argyros, *Computational theory of iterative methods*, Studies in Computational Mathematics **15**, Elsevier, 2007, New York, U.S.A.
- [5] P.P. Zabrejko, P.P. Zlepko, *Generalization of the Newton–Kantorovich method for an equation with non-differentiable operator (Russian)*, Ukr. Mat. Zhurn. **34** (1985), 365–369.

- [6] P.P. Zabrejko, D.F. Nguen, *The majorant method in the theory of Newton–Kantorovich approximations, and the Pták error estimates*, Numer. Funct. Anal. Optimiz. **9** (5 and 6) (1987), 671–684.

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

E-mail address: iargyros@cameron.edu