CONVERGENCE THEOREMS FOR NEWTON'S AND MODIFIED NEWTON'S METHODS

IOANNIS K. ARGYROS

ABSTRACT. In this study we are concerned with the problem of approximating a locally unique solution of an equation in a Banach space setting using Newton's and modified Newton's methods. We provide weaker convergence conditions for both methods than before [5]-[7]. Then, we combine Newton's with the modified Newton's method to approximate locally unique solutions of operator equations. Finer error estimates, a larger convergence domain, and a more precise information on the location of the solution are obtained under the same or weaker hypotheses than before [5]-[7]. The results obtained here improve our earlier ones reported in [4]. Numerical examples are also provided.

1. Introduction

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

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

where F is a Fréchet-differentiable operator defined on an open convex subset 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 systems.

Received by the editors April 27, 2009.

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

Key words and phrases. Banach space, Newton-Kantorovich method, radius of convergence, Fréchet-derivative, Banach Lemma on invertible operators.

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 — 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 most popular methods for approximation x^* are undoubtedly Newton's method

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

and the modified Newton's method

$$(1.3) y_{n+1} = y_n - F'(y_0)^{-1} F(y_n) (n \ge 0), (y_0 = x_0).$$

There is an extensive literature on the semilocal as well as the local convergence results for both methods under various hypotheses. Such results can be found in [1]-[7], and the references there.

The most popular hypotheses are of Newton-Kantorovich type [4], [6], [7]. Indeed, let $x_0 \in D$. Assume there exist constants $\eta > 0$, $\ell > 0$ such that

(1.4)
$$F'(x_0)^{-1} \in L(\mathcal{Y}, \mathcal{X}),$$

$$||F'(x_0)^{-1}F(x_0)|| \le \eta,$$

(1.6)
$$||F'(x_0)^{-1}[F'(x) - F'(y)]|| \le \ell ||x - y|| \text{ for all } x, y \in D,$$

$$(1.7) h = \ell \ \eta \le \frac{1}{2},$$

and

(1.8)
$$\overline{U}(x_0, s^*) = \{x \in X : ||x - x_0|| \le s^*\} \subseteq D,$$

where,

$$(1.9) s^* = \frac{1 - \sqrt{1 - 2h}}{\ell} \,.$$

Estimate (1.7) is the crucial non-optimum sufficient condition for the semilocal convergence of both methods [2], [4], [6], [7] (see also Theorem 1).

Under condition (1.7) method (1.2) converges quadratically to x^* (if (1.7) holds as a strict inequality) whereas method (1.3) converges linearly to x^* . There are examples in the literature where both methods converge to x^* but condition (1.7) is violated. Therefore one would expect that there may be conditions weaker than (1.7). This is the motivation for our study. Note that in view of the Lipschitz condition (1.6) it follows that there exists $\ell_0 > 0$ such that the center-Lipschitz condition

$$(1.10) ||F'(x_0)^{-1}[F'(x) - F'(x_0)]|| \le \ell_0 ||x - x_0|| for all x \in D$$

holds.

In general

$$(1.11) \ell_0 \le \ell,$$

holds true, and $\frac{\ell}{\ell_0}$ can be arbitrarily large [1]-[2].

Recently, in [3, p. 387, Case 3, for $\delta = \delta_0$], [2], we showed that condition (1.7) can always be replaced by the weaker

(1.12)
$$h_1 = \ell_1 \, \eta \le \frac{1}{2}, \qquad \ell_1 = \frac{1}{8} \left(\ell + 4 \, \ell_0 + \sqrt{\ell^2 + 8 \, \ell_0 \, \ell} \right),$$

in the case of Newton's method (1.2) (see also, Example 10).

Here, we show that in the case of the modified method (1.3), for convergence, condition (1.7) can be replaced by

$$(1.13) h_0 = \ell_0 \,\, \eta \le \frac{1}{2},$$

and (1.6) by weaker condition (1.10). Finer error estimates on the distances involved, a larger convergence domain, and a more precise information on the location of the solution than in earlier results [6] are also obtained this way (see Theorem 3 for method (1.2), and Theorem 5 for method (1.3)).

Using (1.13) (whenever (1.7) (or (1.12) do not hold), we can employ method (1.3) for a finite number of steps, say N until condition (1.7) (or (1.12)) is satisfied for $x_0 = y_N$. Then faster method (1.2) takes over from method (1.3). The results also improve further the corresponding ones in [4], where $\frac{\ell_0 + \ell}{2}$ was used instead of ℓ_1 (see Remark 9).

Numerical examples are also provided. The technique introduced here can extend to other Newton-type iterative methods [2], [3], [6], [7].

2. Semilocal Onvergence Analysis

The following semilocal convergence result for methods (1.2) and (1.3) can be found in [2], [6], [7]:

Theorem 1. Let $F: D \subseteq \mathcal{X} \to \mathcal{Y}$ be a differentiable operator.

Assume there exist $x_0 \in D$, and constants $\ell > 0$, $\eta > 0$ such that

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

$$\|F'(x_0)^{-1}F(x_0)\| \leq \eta,$$

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

$$h = \ell \eta \leq \frac{1}{2},$$

and

$$\overline{U}(x_0, s^*) \subseteq D$$
,

where,

$$s^* = \frac{1 - \sqrt{1 - 2 \ell \eta}}{\ell} \,.$$

Then sequences $\{y_n\}$, $\{x_n\}$ are well defined, remain in $\overline{U}(x_0, s^*)$ for all $n \geq 0$ and converge to a unique solution x^* of equation F(x) = 0 in $\overline{U}(x_0, s^*)$. Moreover the following estimates hold:

$$||y_{n+1} - y_n|| \le q^n ||y_1 - y_0|| \le q^n \eta,$$

$$||y_n - x^*|| \le \frac{q^n}{1 - q} \eta,$$

$$||x_{n+2} - x_{n+1}|| \le \frac{\ell (s_{n+1} - s_n)^2}{2 (1 - \ell s_{n+1})},$$

and

$$||x_n - x^*|| \le s^* - s_n, \quad s^* = \lim_{n \to \infty} s_n,$$

where,

$$s_0 = 0$$
, $s_1 = \eta$, $s_{n+2} = s_{n+1} + \frac{\ell (s_{n+1} - s_n)^2}{2 (1 - \ell s_{n+1})}$ $(n \ge 0)$,

and

$$(2.1) q = 1 - \sqrt{1 - 2 \ell h}.$$

Remark 2. There is a plethora of estimates on the distances $||x_{n+1}-x_n||$, $||x_n-x^*||$, $||y_{n+1}-y_n||$, $||y_n-x^*||$ $(n \ge 0)$ [1]-[7]. However we decided to list only the estimates

related to what we need in this study. In the case of Newton's method (1.2) we showed in [3] the following improvement of Theorem 1.

Theorem 3 ([3]). Let $F: D \subseteq \mathcal{X} \to \mathcal{Y}$ be a differentiable operator.

Assume there exist $x_0 \in D$, and constants $\ell_0 > 0$, $\ell > 0$, $\eta \ge 0$ such that

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

$$\|F'(x_0)^{-1}F(x_0)\| \leq \eta,$$

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

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

$$h_1 = \ell_1 \eta \leq \frac{1}{2},$$

$$\overline{U}(x_0, t^*) \subseteq D,$$

where,

$$t_0 = 0$$
, $t_1 = \eta$, $t_{n+2} = t_{n+1} + \frac{\ell (t_{n+1} - t_n)^2}{2 (1 - \ell_0 t_{n+1})}$ $(n \ge 0)$,

and

$$t^* = \lim_{n \to \infty} t_n \le \frac{2 \eta}{2 - \ell_2}, \quad \ell_2 = \frac{-\frac{\ell}{\ell_0} + \sqrt{\left(\frac{\ell}{\ell_0}\right)^2 + \frac{8 \ell}{\ell_0}}}{2}.$$

Then sequence $\{x_n\}$ $(n \geq 0)$ 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 unique solution x^* of equation F(x) = 0 in $\overline{U}(x_0, t^*)$.

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

(2.2)
$$||x_{n+1} - x_n|| \le t_{n+1} - t_n, ||x_n - x^*|| \le t^* - t_n, t_n \le s_n,$$

$$(2.3) t_{n+1} - t_n \le s_{n+1} - s_n,$$

and

$$(2.4) t^* - t_n \le s^* - s_n.$$

Remark 4. Note also that (2.2) and (2.3) hold as strict inequalities if $\ell_0 < \ell$ [1]–[4]. Moreover we have:

$$(2.5) h \le \frac{1}{2} \Rightarrow h_1 \le \frac{1}{2},$$

but not vice versa unless if $\ell_0 = \ell$. That is under the same computational cost we managed to weaken (1.7) since in practice the computation of ℓ also requires the computation of ℓ_0 . Furthermore, in Example 10, we show that (1.12) holds but condition (1.7) is violated.

Concerning the semilocal convergence of the modified Newton's method we show that (1.13) replaces condition (1.7).

Theorem 5. Let $F: D \subseteq \mathcal{X} \to \mathcal{Y}$ be a differentiable operator.

Assume there exist $x_0 \in D$ and constants $\ell_0 > 0$, $\eta > 0$, such that

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

$$\|F'(x_0)^{-1}F(x_0)\| \le \eta,$$

$$\|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \le \ell_0 \|x - x_0\| \text{ for all } x \in D,$$

$$h_0 = \ell_0 \eta < \frac{1}{2},$$

and

$$(2.6) \overline{U}(x_0, s_0^*) \subseteq D,$$

where,

$$s_0^* = \frac{2 \eta}{1 + \sqrt{1 - 2 \ell_0 \eta}}.$$

Then sequence $\{y_n\}$ $(n \geq 0)$ generated by the modified Newton's method (1.3) is well defined, remains in $\overline{U}(x_0, s_0^*)$ for all $n \geq 0$ and converges to a unique solution x^* of equation F(x) = 0 in $\overline{U}(x_0, s_0^*)$.

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

$$||y_{n+1} - y_n|| \le q_0^n ||y_1 - y_0||$$

and

$$||y_n - x^*|| \le \frac{q_0^n}{1 - q_0} \eta,$$

where

$$q_0 = 1 - \sqrt{1 - 2 \, \ell_0 \, \eta} \,.$$

Proof. We shall show that the assumptions of the contraction mapping principle [2], [6] are satisfied for the operator

(2.7)
$$P(x) = x - F'(x_0)^{-1} F(x) \text{ on } \overline{U}(x_0, s_0^*).$$

Let $x, y \in \overline{U}(x_0, s_0^*)$. Then we can obtain the identity

$$P(x) - P(y) = x - y - F'(x_0)^{-1}(F(x) - F(y))$$
$$= F'(x_0)^{-1} \int_0^1 \{F'(x_0) - F'[y + t(x - y)]\}(x - y) dt.$$

This identity together with (1.10) implies the estimate

$$||P(x) - P(y)|| \le \ell_0 \int_0^1 [(1-t) ||x - x_0|| + t ||y - x_0||] dt$$

$$\le \ell_0 s_0^* ||x - y|| = q_0 ||x - y||.$$

Consequently, P is a contraction operator in the ball $\overline{U}(x_0, s_0^*)$. To complete the proof, it remains to show that

$$P \ \overline{U}(x_0, s_0^*) \subseteq \overline{U}(x_0, s_0^*).$$

Let $x \in \overline{U}(x_0, s_0^*)$. Then by (2.7) we can obtain in turn

$$||P(x) - x_0|| \le ||P(x) - P(x_0)|| + ||P(x_0) - x_0||$$

$$\le ||F'(x_0)^{-1} \int_0^1 \{F'(x_0) - F'[x_0 + t(x - x_0)]\}(x - x_0) dt|| + \eta$$

$$\le \ell_0 \int_0^1 t dt ||x - x_0||^2 + \eta \le \frac{\ell_0}{2} (s_0^*)^2 + \eta = s_0^*,$$

by the choice of s_0^* . That completes the proof of Theorem 5.

Remark 6. Note that by (2.8) the operator P satisfies a Lipschitz condition with constant q_0 in the ball $\overline{U}(x_0, s_0^*)$. The modified Newton's method thus converges at the rate of a geometric progression with quotient q_0 .

The above analysis of method (1.3) relates to the simplest case. More subtle arguments (see, e.g. Kantorovich and Akilov [6]) show that Theorem 5 remains valid if the sign < in (2.6) is replace by \le . Therefore from now on we can replace (2.6) by (1.13) in Theorem 5.

Remark 7. If $\ell_0 = \ell$, Theorems 3 and 5 reduce to Theorem 1. Otherwise these theorems constitute improvements of it. Indeed see (2.2)–(2.5), and notice that

$$q_0 < q$$

and

Notice also that (1.7) (or (1.12)) implies (1.13) and if $t^* \leq s_0^*$ the quadratic convergence of method (1.2) is guaranteed. Moreover s_0^* given in closed form can then in practice replace t^* . Furthermore if $s_0^* < t^*$ then there exists N > 1 such that $x_n \in \overline{U}(x_0, s_0^*)$ for $n \geq N$ and then again s_0^* can replace t^* .

Next we show that we can start with method (1.3) and after a finite number of steps continue with faster method (1.2):

Proposition 8. Under hypotheses (4)-(16), (10), (13), and (19), for $x_0 = y_0$, define

$$\alpha = \frac{1}{1 - \ell_0 \ s_0^*},$$
$$L = \alpha \ \ell.$$

for a fixed integer N

$$\ell_{2} = \sup_{x \in \overline{U}(y_{N}, r_{N})} \frac{\parallel F'(y_{0})^{-1} [F'(x) - F'(y_{N})] \parallel}{\parallel x - y_{N} \parallel},$$

$$L_{0} = \alpha \ \ell_{2} \leq L,$$

$$I_{0} = L$$

$$\frac{1}{8} (L + 4 L_{0} + \sqrt{L^{2} + 8 L_{0} L}) \quad \text{if} \quad L_{0} < L,$$

$$N = \left[-\frac{\ln 2 \overline{L} \alpha^{2} \eta}{\ln q_{0}} \right] + 1,$$

$$r_{N} = \begin{cases} \frac{1 - \sqrt{1 - 2 \overline{L} \eta_{N}}}{\overline{L}} \quad \text{if} \quad L_{0} = L \end{cases}$$

$$\eta_{N} = \alpha \ q_{0}^{N} \ \eta,$$

and

$$L_{2N} = \frac{-\frac{L}{L_0} + \sqrt{\left(\frac{L}{L_0}\right)^2 + \frac{8L}{L_0}}}{2}, \quad \text{for} \quad L_0 \neq 0,$$

where [r] denotes the integer part of real number r. Set

$$\overline{x_0} = y_N$$
.

Moreover, assume:

$$(2.9) \overline{U}(y_N, r_N) \subseteq D.$$

Then the following hold:

$$(2.13) H_N = \overline{L} \, \eta_N \le \frac{1}{2};$$

Newton's method (1.2), starting at $x_0 = \overline{x_0}$ converges to a unique solution x^{**} of equation F(x) = 0 in $\overline{U}(y_N, r_N)$, and

$$N_0 \leq N_1$$

where

$$N_0 = N$$
 for $L_0 = L$,

and

$$N_1 = N$$
 for $L_0 < L$.

Moreover, if the inclusion

$$(2.14) \overline{U}(y_N, r_N) \subseteq \overline{U}(y_0, s_0^*),$$

hold, then

$$x^* = x^{**}$$

Note that parameter L_0 is independent of N if $L_0 = L$, and the inclusion (2.14) holds id and only if

$$||y_N - y_0|| + r_N \le s_0^*$$

Proof. Using Theorem 5, and the estimates

we obtain

$$H_N = \overline{L} \ \eta_N = \alpha^2 \ \overline{L} \ q_0^N \ \eta \le \frac{1}{2}$$

and

$$N_0 < N_1$$

by the choice of N.

It follows by Theorems 1 and 3, with L_0 , L, η_N , r_N , r_N replacing ℓ_0 , ℓ , η , s^* , t_0^* respectively, that there exists a unique solution x^{**} of equation F(x) = 0 in $\overline{U}(y_N, r_N)$.

Moreover, if inclusion (2.14) holds by the uniqueness of the solution x^* in $\overline{U}(y_0, s_0^*)$, we deduce $x^* = x^{**}$. That completes the proof of Proposition 8.

Remark 9. In our related work [4], we used h_2 , ℓ_2 , $\overline{N_0}$ instead of h_1 , ℓ_1 , N_0 , respectively given by:

$$(2.15) h_2 = \ell_2 \ \eta \le \frac{1}{2},$$

$$\ell_2 = \frac{\ell_0 + \ell}{2},$$

and

(2.17)
$$\overline{N_0} = \left[\frac{-\ln((\ell_0 + \ell) \, \eta \, \alpha^2)}{\ln q_0} \right] + 1.$$

However, we have

$$(2.18) \qquad \qquad \ell_2 \leq \ell_1 \qquad \text{for all} \quad \ell_0 > 0 \quad \text{and} \quad \ell > 0.$$

Therefore, the following implication holds

$$(2.19) h_2 \le \frac{1}{2} \Longrightarrow h_1 \le \frac{1}{2}$$

but not necessarily vice verca, and

$$(2.20) N_0 \le \overline{N_0}.$$

Note also that strict inequality can hold in (2.18) and (2.20). Hence, we have improved our earlier results in [4], which in turn improved the corresponding results in [5]-[7] (see also the example that follows).

Example 10. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, D = [a, 2-a], $a \in [0, \frac{1}{2})$ and define scalar function F on D by

$$(2.21) F(x) = x^3 - a.$$

Choose $y_0 = 1$. Using (1.5), (1.6), (1.10) and (2.21), we obtain

(2.22)
$$\eta = \frac{1}{3} (1-a), \quad \ell_0 = 3-a < \ell = 2 (2-a) \text{ for all } a \in [0, \frac{1}{2}).$$

The Newton-Kantorovich hypothesis (1.7) becomes

(2.23)
$$h = \frac{2}{3} (1-a) (2-a) > \frac{1}{2}$$

for all $a \in [0, \frac{1}{2})$. That is according to Theorem 1 there is no guarantee that either methods (1.2) or (1.3) starting at $x_0 = y_0 = 1$ converge to x^* .

However, according to (2.15):

$$(2.24) h_2 = \frac{1}{6} (1-a) [3-a+2(2-a)] \le \frac{1}{2},$$

provided that

(2.25)
$$a \in \left[\frac{5-\sqrt{13}}{3}, \frac{1}{2}\right), \quad \text{where} \quad \frac{5-\sqrt{13}}{3} = .46481624.$$

By (1.12), we get

$$(2.26) h_1 \le \frac{1}{2}$$

for all $a \in \left[.450339002, \frac{1}{2}\right)$.

Using condition (1.13) we can do even better since

(2.27)
$$h_0 = \frac{1}{3} (1-a) (3-a) \le \frac{1}{2},$$

provided that

(2.28)
$$a \in \left[\frac{4-\sqrt{10}}{2}, \frac{1}{2}\right)$$
 where $\frac{4-\sqrt{10}}{2} = .41886117$.

which improves the choice for a given by (2.25). However only linear and not quadratic convergence is guaranteed.

Let us now use a=.49. In particular (1.7) does not hold since for $\eta=.17$, $\ell=3.02$

$$h = .5134 > \frac{1}{2} \,.$$

However, (1.13) holds, since for $\ell_0 = 2.51$:

$$h_0 = .4267 \le \frac{1}{2} \,.$$

We get

$$q_0 = .617116205, \quad \alpha = 2.61175848, \quad s_0^* = .24586303,$$

and

$$N = [4.0325] + 1 = 5.$$

Moreover, we obtain

$$\overline{x}_0 = y_4 = .78911736.$$

Finally, note that $x^* = .788373516$.

The ideas presented here can be extented to other Newton-type iterative methods [1], [2], [3], [5] along the same lines (see also [4]).

REFERENCES

- 1. I.K. Argyros: On the Newton-Kantorovich hypothesis for solving equations. J. Comput. Appl. Math. 169 (2004), 315-332.
- Computational theory of iterative methods. Studies in Computational Mathematics, 15 (Editors: C.K. Chui and L. Wuytack), Elsevier Publ. Co., New York, 2007.
- 3. ______: A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space. J. Math. Anal. Appl. 298 (2004), no. 2, 374-397.
- 4. _____: A comparison between convergence theorems for Newton's and modified Newton's methods. submitted for publication.
- J.M. Gutiérrez, M.A. Hernández & M.A. Salanova: Accessibility of solutions by Newton's method. *Intern. J. Comput. Math.* 57 (1995), 239-247.
- 6. L.V. Kantorovich & G.P. Akilov: Functional Analysis. Pergamon Press, Oxford, 1982.
- M.A. Krasnosel'skii, G.M. Vainikko, P.P. Zabreiko, Ya.B. Rutitskii & V.Ya. Stetsenko: Approximate Solution of Operator Equations. Wolters-Noordhoff Publ., Groningen, 1972.

CAMERON UNIVERSITY, DEPARTMENT OF MATHEMATICS SCIENCES, LAWTON, OK 73505, USA Email address: iargyros@cameron.edu