

IOANNIS K. ARGYROS (Lawton, OK)

A WEAKER AFFINE COVARIANT NEWTON–MYSOVSKIKH THEOREM FOR SOLVING EQUATIONS

Abstract. The Newton–Mysovskikh theorem provides sufficient conditions for the semilocal convergence of Newton’s method to a locally unique solution of an equation in a Banach space setting. It turns out that under weaker hypotheses and a more precise error analysis than before, weaker sufficient conditions can be obtained for the local as well as semilocal convergence of Newton’s method. Error bounds on the distances involved as well as a larger radius of convergence are obtained. Some numerical examples are also provided.

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

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

where F is continuously Fréchet differentiable on a closed convex subset D of a Banach space X with values in a Banach space Y .

The most popular method for 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 $F'(x) \in L(X, Y)$ ($x \in D$) denotes the Fréchet derivative of the operator F [2], [5], [6]. The semilocal and local convergence of this method has been examined under various conditions by many authors [1]–[4]. A survey of such results can be found in [2] and the references there.

In particular we are motivated by the affine covariant Newton–Mysovskikh theorem (ACNMT) (see also Theorem 1 below) [5], [6] which basically

2000 *Mathematics Subject Classification*: 65H10, 65G99, 65J15, 47H17, 49M15.

Key words and phrases: Newton’s method, covariant Newton–Mysovskikh theorem, Banach space, local-semilocal convergence, radius of convergence.

states that if the initial guess x_0 is close enough to the solution x^* then Newton’s method (2) converges quadratically to x^* . The basic condition is given by

$$(3) \quad \|F'(z)^{-1}[F'(y) - F'(x)](y - x)\| \leq w\|y - x\|^2 \quad \text{for all } x, y, z \in D.$$

Here we have noticed (see Theorem 1) that condition (3) is not used at full strength to prove convergence of method (2). Indeed, a much weaker hypothesis (see (6)) can replace (3) in the proof of the theorem. This observation has the following advantages (semilocal case):

- (a) the applicability of (ACNMT) is extended;
- (b) finer error bounds on the distances $\|x_{n+1} - x_n\|, \|x_n - x^*\|$ ($n \geq 0$);
- (c) a more precise information on the location of the solution x^* ;
- (d) the Lipschitz constant is at least as small and easier to compute.

Local case:

- (a) a larger radius of convergence;
- (b) finer error bounds on the distances $\|x_n - x^*\|$ ($n \geq 0$).

The above advantages are obtained under not only weaker hypotheses but under less or the same computational cost.

Some numerical examples are also provided for both the semilocal and local case where our results apply while other ones [5], [6] fail.

We can show the following weaker semilocal convergence version of the affine covariant Newton–Mysovskikh theorem (ACNMT):

THEOREM 1. *Let $F: D \subseteq X \rightarrow Y$ be a Fréchet continuously differentiable operator, and let $x_0 \in D$ be an initial guess. Assume further that*

$$(4) \quad F'(x)^{-1} \in L(Y, X) \quad (x \in D),$$

$$(5) \quad \|F'(x_0)^{-1}F(x_0)\| \leq \eta,$$

$$(6) \quad \|F'(y)^{-1}[F'(x + t(y - x)) - F'(x)](y - x)\| \leq w_1\|y - x\|^2t$$

for all $x, y \in D, t \in [0, 1]$,

$$(7) \quad 2h_1 = w_1\|F'(x_0)^{-1}F(x_0)\| \leq w_1\eta < 2,$$

$$(8) \quad \overline{U}(x_0, r_1) \subseteq D,$$

where

$$(9) \quad \overline{U}(x_0, r_1) = \{x \in X \mid \|x - x_0\| \leq r_1\},$$

$$(10) \quad r_1 = \frac{\|F'(x_0)^{-1}F(x_0)\|}{1 - h_1/2}.$$

Then the sequence $\{x_n\}$ ($n \geq 0$) generated by Newton’s method (2) is well defined, remains in $\overline{U}(x_0, r_1)$ for all $n \geq 0$, and there exists a unique $x^* \in$

$\bar{U}(x_0, r_1)$ such that $F(x^*) = 0$ and $x_n \rightarrow x^*$ as $n \rightarrow \infty$ with

$$(11) \quad \|x_{n+1} - x_n\| \leq \frac{1}{2} w_1 \|x_n - x_{n-1}\|^2,$$

$$(12) \quad \|x_{n+1} - x^*\| \leq \varepsilon_n \|x_n - x_{n-1}\|^2 \quad \text{for all } n \geq 1,$$

where

$$(13) \quad \varepsilon_n = \frac{1}{2} w_1 \left[1 + \sum_{j=1}^{\infty} (h_1^{2^n})^{2^j} \right] \leq \frac{1}{2} \frac{w_1}{1 - h_1^{2^n}}.$$

Proof. Using (3) and (6) we obtain in turn

$$(14) \quad \begin{aligned} x_{k+1} - x_k &= -F'(x_k)^{-1} [F(x_k) - F(x_{k-1}) - F'(x_{k-1})(x_k - x_{k-1})] \\ &= -F'(x_k)^{-1} \int_0^1 [F'(x_{k-1} + t(x_k - x_{k-1})) - F'(x_{k-1})](x_k - x_{k-1}) dt \end{aligned}$$

and

$$(15) \quad \begin{aligned} \|x_{k+1} - x_k\| &\leq \left\| F'(x_k)^{-1} \int_0^1 [F'(x_{k-1} + t(x_k - x_{k-1})) - F'(x_{k-1})](x_k - x_{k-1}) dt \right\| \\ &\leq \frac{1}{2} w_1 \|x_k - x_{k-1}\|^2, \end{aligned}$$

which shows (11).

We now prove $\{x_n\}$ is a Cauchy sequence in $\bar{U}(x_0, r_1)$. By induction on k we will show

$$(16) \quad \|x_{k+1} - x_k\| \leq \frac{2}{w_1} h_1^{2^k}, \quad k \in \mathbb{N}_0.$$

For $k = 0$, we have by (2), (5) and (7),

$$\|x_1 - x_0\| = \|F'(x_0)^{-1} F(x_0)\| \leq \eta \leq \frac{2}{w_1} h_1.$$

Assuming (16) to be true for some $k \in \mathbb{N}$, we get

$$(17) \quad \|x_{k+2} - x_{k+1}\| \leq \frac{1}{2} w_1 \|x_{k+1} - x_k\|^2 \leq \frac{1}{2} w_1 \left(\frac{2}{w_1} h_1^{2^k} \right)^2 = \frac{2}{w_1} h_1^{2^{k+1}}.$$

It follows from (16) that $x_{k+1} \in \bar{U}(x_0, r_1)$:

$$(18) \quad \begin{aligned} \|x_{k+1} - x_0\| &\leq \|x_{k+1} - x_k\| + \dots + \|x_1 - x_0\| \\ &\leq \frac{2}{w_1} (h_1^{2^k} + \dots + h_1) = \frac{2}{w_1} h_1 \sum_{j=0}^{\infty} h_1^{2^j - 1} = r_1. \end{aligned}$$

Similarly it can be shown that

$$(19) \quad \|x_{m+k} - x_m\| \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

That is, $\{x_n\}$ is a Cauchy sequence in $\overline{U}(x_0, r_1)$ and as such it converges to some $x^* \in \overline{U}(x_0, r_1)$ (since $\overline{U}(x_0, r_1)$ is a closed set).

Hence, we get

$$x_{k+1} - x_k = -F'(x_k)^{-1}F(x_k) \xrightarrow{k \rightarrow \infty} -F'(x^*)^{-1}F(x^*) = 0$$

and thus $F(x^*) = 0$. Then we get in turn:

$$(20) \quad \begin{aligned} \|x_k - x^*\| &= \lim_{k < m \rightarrow \infty} \|x_k - x_m\| \\ &\leq \lim_{k < m \rightarrow \infty} [\|x_m - x_{m-1}\| + \dots + \|x_{k+1} - x_k\|] \\ &\leq \frac{2}{w_1} \lim_{m \rightarrow \infty} \sum_{i=k+1}^m h_1^{i-1} = \frac{2h^k}{w_1} \lim_{m \rightarrow \infty} \sum_{i=k+1}^m h_1^{i-1-k}. \end{aligned}$$

We also have

$$(21) \quad h_1^k \leq (w_1/2)^2 \|x_{k-1} - x_{k-2}\|^2 = (h_1^{k-1})^2,$$

whence

$$(22) \quad h_1^{k+m} \leq (h_1^k)^{2^m}, \quad k \in \mathbb{N}_0.$$

That is,

$$(23) \quad \begin{aligned} \|x_k - x^*\| &\leq \frac{1}{2} w_1 \|x_{k-1} - x_{k-2}\|^2 [1 + (h^k)^2 + \dots] \\ &\leq \frac{1}{2} w_1 \left[1 + \sum_{j=1}^{\infty} (h_1^k)^{2^j} \right] \|x_{k-1} - x_{k-2}\|^2, \end{aligned}$$

which shows (12).

Finally, to show uniqueness, let $y^* \in \overline{U}(x_0, r_1)$ be a solution of equation (1). Then for all $t \in [0, 1]$ we have

$$x^* + t(y^* - x^*) - x_0 = (1 - t)(x^* - x_0) + t(y^* - x_0)$$

and

$$\begin{aligned} \|x^* + t(y^* - x^*) - x_0\| &\leq (1 - t)\|x^* - x_0\| + t\|y^* - x_0\| \\ &\leq (1 - t)r_1 + tr_1 = r_1. \end{aligned}$$

That is,

$$(24) \quad x^* + t(y^* - x^*) \in \overline{U}(x_0, r_1) \quad (t \in [0, 1]).$$

Using the identity

$$(25) \quad \begin{aligned} 0 = F(y^*) - F(x^*) &= \int_0^1 F'(x^* + t(y^* - x^*))(y^* - x^*) dt \\ &= M(y^* - x^*), \end{aligned}$$

where

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

and hypothesis (4) we deduce that for $x_t = x^* + t(y^* - x^*)$ the linear operator M is invertible. Hence, we conclude $x^* = y^*$. That completes the proof of Theorem 1. ■

REMARK 2. Clearly condition (3) implies (6) but not vice versa. That is, (6) is a weaker condition than (3). Note that

$$(26) \quad w_1 \leq w$$

holds in general and w/w_1 can be arbitrarily large [1].

The classical affine covariant Newton–Mysovskikh theorem has been shown using (3), h , r instead of (6), h_1 and r_1 respectively, where

$$(27) \quad 2h = w \|F'(x_0)^{-1}F(x_0)\| \leq w\eta < 2$$

and

$$(28) \quad r = \frac{\|F'(x_0)^{-1}F(x_0)\|}{1 - h/2}.$$

Note that

$$(29) \quad 2h < 2 \Rightarrow 2h_1 < 2$$

but not vice versa unless $w_1 = w$. If strict inequality holds in (26) then

$$(30) \quad r_1 < r$$

and our error estimates (11)–(13) are finer than the corresponding ones in [5], [6] obtained by replacing w_1 by w in (11)–(13). Note that all the above advantages are obtained under less computational cost since in practice the computation of w requires that of w_1 .

REMARK 3. Although the computation of the constant w_1 in (6) is in practice easier than the computation of w in (3) one may want to replace (6) by the pair of conditions

$$(31) \quad \|F'(x_0)^{-1}[F'(x + t(y - x)) - F'(x)](y - x)\| \leq w_2 \|y - x\|^2 t$$

and

$$(32) \quad \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| \leq w_0 \|x - x_0\|$$

for all $x, y \in D$, $t \in [0, 1]$.

Using (31) and (32) we showed in [1], [2] that Newton’s sequence converges to a unique solution x^* of equation (1) in $\bar{U}(x_0, 2\eta/(2 - \delta))$ provided that

$$(33) \quad h_2 = (\delta w_0 + w_2)\eta \leq 1, \quad \delta \in [0, 1].$$

Clearly these results will be weaker provided

$$(34) \quad \frac{w_2}{1 - w_0 \|z - x_0\|} < w, \quad z \in \overline{U}\left(x_0, \frac{2\eta}{2 - \delta}\right),$$

since

$$(35) \quad \begin{aligned} \|F'(z)^{-1}[F'(y) - F'(x)](y - x)\| &\leq \|F'(z)^{-1}F'(x_0)\| \|F'(x_0)^{-1}[F'(y) - F'(x)](y - x)\| \\ &\leq \frac{w_2 \|y - x\|^2}{1 - w_0 \|z - x_0\|} \end{aligned}$$

by the Banach lemma on invertible operators [5] and the estimate

$$(36) \quad \|F'(z)^{-1}F'(x_0)\| \leq \frac{1}{1 - w_0 \|z - x_0\|}.$$

Condition (34) certainly holds if

$$\|x - x_0\| < \frac{1}{w_0} \left[1 - \frac{w_2}{w}\right]$$

or

$$(37) \quad \frac{2\eta}{2 - \delta} < \frac{1}{w_0} \left[1 - \frac{w_2}{w}\right],$$

which can happen for sufficiently small η and since w_2/w can be arbitrarily small [1], [2].

We provide an example where (7) and (33) hold but (27) fails.

EXAMPLE 4. Let $X = Y = \mathbb{R}$, $D = [a, 2 - a]$, $a \in [0, 1/2)$, $x_0 = 1$ and define a function F on D by

$$(38) \quad F(x) = x^3 - a.$$

Using (3), (5), (6), (31) and (32) we obtain

$$(39) \quad w = \frac{2(2 - a)}{a^2}, \quad \eta = \frac{1}{3}(1 - a), \quad w_0 = 3 - a, \quad w_1 = \frac{2}{a^2}, \quad w_2 = 2(2 - a).$$

Condition (27) does not hold since

$$(40) \quad 2h = \frac{2}{3a^2}(1 - a)(2 - a) \geq 2 \quad \text{for all } a \in [0, 1/2).$$

That is, there is no guarantee that Newton's method starting at x_0 converges to x^* . However, condition (7) holds, since

$$(41) \quad 2h_1 = \frac{2}{3a^2}(1 - a) < 2 \quad \text{for all } a \in \left(\frac{\sqrt{13} - 1}{6}, \frac{1}{2}\right).$$

Finally, say for $\delta = 1$, condition (33) holds since

$$(42) \quad h_2 = \frac{1}{3}(1 - a)[2(2 - a) + 3 - a] \leq 1 \quad \text{for all } a \in \left[\frac{5 - \sqrt{13}}{3}, \frac{1}{2}\right).$$

In order for us to cover the local convergence of Newton’s method under Newton–Mysovskikh-type conditions, let x^* be a simple solution of the equation $F(x) = 0$. Assume that

$$(43) \quad \|F'(x)^{-1}[F'(x^* + t(x - x^*)) - F'(x)](x - x^*)\| \leq v(1 - t)\|x - x^*\|^2$$

for all $x \in \overline{U}(x^*, r^*)$, $t \in [0, 1]$ where

$$(44) \quad r^* = \frac{2}{v} \quad (v \neq 0).$$

Set also

$$(45) \quad r_1^* = \frac{2}{w} \quad (w \neq 0).$$

We can show the following local convergence theorem for Newton’s method (2):

THEOREM 5. *Let $F: D \subseteq X \rightarrow Y$ be a Fréchet differentiable operator. Assume that*

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

and there exists $x^* \in D$ such that $F(x^*) = 0$.

(a) *If condition (43) holds and*

$$(46) \quad \overline{U}(x^*, r^*) \subseteq D$$

then the sequence $\{x_n\}$ generated by Newton’s method (2) is well defined, remains in $U(x^, r^*)$ for all $n \geq 0$ and converges to x^* provided that $x_0 \in U(x^*, r^*)$, with*

$$(47) \quad \|x_{n+1} - x^*\| \leq \frac{v}{2} \|x_n - x^*\|^2 \quad (n \geq 0).$$

(b) *If condition (3) holds and*

$$(48) \quad \overline{U}(x^*, r_1^*) \subseteq D,$$

then the sequence $\{x_n\}$ generated by Newton’s method (2) is well defined, remains in $U(x^, r_1^*)$ for all $n \geq 0$ and converges to x^* provided that $x_0 \in U(x^*, r_1^*)$, with*

$$(49) \quad \|x_{n+1} - x^*\| \leq \frac{w}{2} \|x_n - x^*\|^2 \quad (n \geq 0).$$

Proof. (a) Using (43), (44) and induction on k we get

$$(50) \quad \|x_{k+1} - x^*\| \leq \frac{v}{2} \|x_k - x^*\|^2 < \|x_k - x^*\| < r^*,$$

which shows $x_k \in U(x^*, r^*)$ for all k and $\lim_{k \rightarrow \infty} x_k = x^*$.

(b) Similarly using (3), (45) and induction on k we get

$$(51) \quad \|x_{k+1} - x^*\| \leq \frac{w}{2} \|x_k - x^*\|^2 < \|x_k - x^*\| < r^*.$$

That completes the proof of Theorem 5. ■

Let us further introduce the conditions:

$$(52) \quad \|F'(x^*)^{-1}[F'(x^* + t(x - x^*)) - F'(x)](x - x^*)\| \leq v_1(1 - t)\|x - x^*\|^2$$

and

$$(53) \quad \|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq v_0\|x - x^*\|.$$

Then since

$$(54) \quad \|F'(x)^{-1}z\| \leq \|[F'(x)^{-1}F'(x^*)]\| \|[F'(x^*)^{-1}z]\|,$$

if we set

$$(55) \quad r_2^* = \frac{2}{2v_0 + v_1}$$

then we can show [1], [2] Newton's method converges to x^* provided that $x_0 \in U(x^*, r_2^*)$ and

$$(56) \quad U(x^*, r_2^*) \subseteq D.$$

In general

$$(57) \quad v_1 \leq v \leq w$$

and v/v_1 , w/v_1 , w/v can be arbitrarily large [1], [2]. Therefore by (44) and (45),

$$r_1^* \leq r^*.$$

We can now provide an example where we compare the radii r^* , r_1^* and r_2^* .

EXAMPLE 6. Let $X = Y = \mathbb{R}$, $D = \overline{U}(0, 1)$, $x^* = 0$, and define the function F on D by

$$(58) \quad F(x) = e^x - 1.$$

Using (3), (43), (52), (53) we obtain

$$(59) \quad w = e, \quad v = e - 1, \quad v_0 = e - 1, \quad v_1 = e.$$

By (44), (45), (55) and (59) we get

$$\begin{aligned} r_1^* &= .735758882, \\ r^* &= 1.163953414, \\ r_2^* &= .324947231. \end{aligned}$$

We need to reset $r^* = 1$ so that (46) holds.

It follows that since $r_1^* < r^*$, our approach provides a wider choice of initial guesses x_0 than in the local convergence results using the standard condition (3).

References

- [1] 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.

- [2] I. K. Argyros, *Newton Methods*, Nova Science, New York, 2005.
- [3] J. M. Gutierrez, *A new semilocal convergence theorem for Newton's method*, J. Comput. Appl. Math. 79 (1997), 131–145.
- [4] J. M. Gutierrez, M. A. Hernandez, and M. A. Salanova, *Accessibility of solutions by Newton's method*, Int. J. Comput. Math. 57 (1995), 239–247.
- [5] L. V. Kantorovich and G. P. Akilov, *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.
- [6] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

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

Received on 11.1.2006;
revised version on 7.6.2006

(1800)

