Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Issue 1, July 2003, pp. 88–95.
Submitted: May 19, 2002. Published: July 1 2003.

SEMILOCAL CONVERGENCE FOR NEWTON'S METHOD ON A BANACH SPACE WITH A CONVERGENCE STRUCTURE AND TWICE FRÉCHET DIFFERENTIABLE OPERATORS

Ioannis K. Argyros

Abstract. We provide a semilocal convergence theorem for Newton's method on a Banach space with a convergence structure. The operators involved are twice Fréchet differentiable. Our result contains earlier ones as special cases, if the Banach space is equipped with a real norm. Moreover it compares favorably with earlier results involving only once Fréchet-differentiable operators, since our sufficient convergence conditions are shown that they can be weaker.

A.M.S. (MOS) Subject Classification Codes. 65G99, 65B05, 65J15, 65N30, 65N35, 47H17, 49M15, CR:1.5.

Key Words and Phrases. Newton's method, Banach space with a convergence structure, majorizing sequence, Fréchet-differentiable operator, Newton–Kantorovich hypothesis.

1. Introduction

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

$$F(x) = 0, \tag{1}$$

where,

$$F(x) = AG(x_0 + x). \tag{2}$$

G is a twice Fréchet differentiable operator on a Banach space X with a convergence structure and values in X, A is meant to be an approximation to $G'(x_0)^{-1}$ and x_0 is the initial guess for Newton's method

$$x_0 = 0, \;\; x_{n+1} = x_n + F'(x_n)^*(-F(x_n)) \quad (n \geq 0). \tag{3}$$

Department of Mathematics, Cameron University, Lawton, OK 73505, U.S.A.
E-mail Address: ioannisa@cameron.edu

By a Banach space X with convergence structure we mean a triple (X, V, E) in the sense of Definition 1 in [3] (see also [1], [2], [6]). Operator $F'(x)^*$ has also been defined in [3].

We provide a semilocal convergence analysis for Newton's method (2) in this general setting. In the case of a Banach space with (real) norm $\|\cdot\|$ our results reduce to the ones in [2], [4], [7].

Moreover our results compare favorably with earlier ones involving only once Fréchet-differentiable operators [2], [6], since our sufficient convergence conditions are shown that they can be weaker.

Finally some numerical examples are provided to show that our results can show convergence of Newton's method to x^*, when earlier ones fail.

2. Semilocal Convergence Analysis for Newton's Method

We provide the main semilocal convergence result for Newton's method (2) on a Banach space with a convergence structure and twice Fréchet-differentiable operators.

Theorem 1. *Let there be given a Banach space X with convergence structure (X, V, E) where $V = (V, C, \|\cdot\|_V)$, and operator $F \in C''(X_0 \to X)$ with $X_0 \subseteq X$, and operator $M \in C'(V_0 \to V)$ with $V_0 \subseteq V$, and a point $p \in C$ satisfying:*

$$U(p) \subseteq X_0, \quad [0, p] \subseteq V_0; \tag{4}$$

M is order-convex on $[0, p]$ and for all $x, y \in U(p)$ with $|x| + |y| \leq p$

$$M'(|x+y|) - M'(x) \in B([F''(x) - F''(x+y)](y)); \tag{5}$$

$$M'(0) \in B(I - F'(0)), \quad (-F(0), M(0)) \in E; \tag{6}$$

$$M(p) \leq p; \tag{7}$$

and

$$M'(p)^n p \to 0 \quad \text{as } n \to \infty. \tag{8}$$

Then, sequence $(x_n, t_n) \in (X \times V)^\pi$, where $\{x_n\}$ is generated by Newton's method (3) and $\{t_n\}$ $(n \geq 0)$ is given by

$$t_0 = 0, \quad t_{n+1} = M(t_n) + M'(|x_n|)(a_n), \quad a_n = |x_{n+1} - x_n| \tag{9}$$

is well defined for all $n \geq 0$, belongs in E^π, and is monotone.

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

$$t_n \leq b, \tag{10}$$

where,

$$b = M^\infty(0), \tag{11}$$

is the smallest fixed point of M in $[0, p]$.

Proof. We first show that $(x_1, b) \in E$. By (6) and (7) we get (for $p = b$)

$$\begin{aligned} M'(0)(b) + |-F(0)| &\leq M'(0)(b-0) + M(0) \\ &\leq M(b) - M(0) + M(0) = M(b) \leq b, \end{aligned} \tag{12}$$

which shows that equation

$$q = (I - F'(0))(q) - F(0) \tag{13}$$

has a solution $q = x_1$, and $(x_1, b) \in E$.

Hence by (9) and the order convexity of M we get

$$|x_1| \leq M'(0)|x_1| + M(0) = t_1 \tag{14}$$

and

$$t_1 = M'(0)|x_1| + M(0) \leq M'(0)b + M(0) \leq M(b) \leq b. \tag{15}$$

Let us assume sequence (x_n, t_n) is well defined and monotone so that

$$0 \leq (x_{k-1}, t_{k-1}) \leq (x_k, t_k) \tag{16}$$

and

$$t_k \leq b \tag{17}$$

for all $k = 1, 2, \dots, n$.

Using (5), (6) and the approximation

$$\begin{aligned} I - F'(x_n) &= I - F'(x_n) - F'(0) + F'(0) \\ &= [I - F'(0)] + [F'(0) - F'(x_n)] \\ &= [I - F'(0)] - \int_0^1 F''(\theta x_n) d\theta (x_n), \end{aligned} \tag{18}$$

we get

$$\begin{aligned} |I - F'(x_n)| &\leq |I - F'(0)| + \int_0^1 |-F''(\theta x_n) d\theta (x_n)| \\ &\leq M(0) + M'(|x_n|) - M'(0) = M'(|x_n|). \end{aligned} \tag{19}$$

That is

$$M'(|x_n|) \in B(I - F'(x_n)). \tag{20}$$

Moreover using (3), (5), (9) and the approximation

$$\begin{aligned} |-F(x_n)| &= |-F(x_n) + F(x_{n-1}) + F'(x_{n-1})(x_n - x_{n-1})| \\ &= \left| -\int_0^1 F''[x_n + \theta(x_{n+1} - x_n)](1-\theta) d\theta (x_{n+1} - x_n)^2 \right| \\ &\leq \int_0^1 [M'(|x_{n-1}| + \theta a_{n-1}) - M'(|x_{n-1}|)] a_{n-1} d\theta \\ &= M(|x_{n-1}| + a_{n-1}) - M(|x_{n-1}|) - M'(|x_{n-1}|) a_{n-1} \\ &\leq M(t_{n-1} + t_n - t_{n-1}) - M(t_{n-1}) - M'(|x_{n-1}|) a_{n-1} \\ &= M(t_n) - t_n. \end{aligned} \tag{21, 22}$$

Let

$$q = b - t_n. \tag{23}$$

Using (20), (22) and (23) we obtain in turn

$$\begin{aligned} M'(|x_n|)(q) + |-F(x_n)| &\le M'(t_n)(b - t_n) + M(t_n) - t_n \\ &\le M(b) - M(t_n) + M(t_n) - t_n \le b - t_n = q. \end{aligned} \tag{24}$$

Hence equation

$$q = (I - F'(x_n))(q) + (-F(x_n)) \tag{25}$$

has a solution x_{n+1} and $a_n \le b - t_n$. We also get

$$t_{n+1} \le M(t_n) + M'(t_n)(b - t_n) \le M(b) \le b, \tag{26}$$

and

$$\begin{aligned} a_n + t_n &\le M'(|x_n|)(a_n) + |-F(x_n)| + t_n \\ &\le M'(|x_n|)(a_n) + M(t_n) = t_{n+1}. \end{aligned} \tag{27}$$

By (9) we obtain using induction on n

$$M^n(0) \le t_n \le b. \tag{28}$$

That is $t_n \to b$ since $M^n(0) \to b$. Hence sequence $\{x_n\}$ converges to $x^* \in U(b)$ which is a solution of F. To show uniqueness, define the modified Newton iteration

$$y_{n+1} = y_n - F(y_n), \quad y_0 = 0. \tag{29}$$

This sequence also converges to x^*, and $(y_n, M^n(0))$ is monotone in $X \times V$. Let $y^* \in U(p)$ be a solution of F. Then it can easily be seen

$$|y^* - y_n| \le M^n(p) - M^n(0) \to 0 \quad \text{as } n \to \infty. \tag{30}$$

Hence we get $x^* = y^*$.

That completes the proof of Theorem 1. □

Remark. In order for us to obtain a posteriori estimates we define as in [3], [6]

$$R_n(q) = (I - M'(|x_n|)^* S_n(q) + a_n, \tag{31}$$

$$S_n(q) = M(|x_n| + q) - M(|x_n|) - M'(|x_n|)(q) \tag{32}$$

on interval

$$I_n = [0, p - |x_n|]. \tag{33}$$

As in [3], [6] we can show the following Propositions but for twice Fréchet-differentiable operators F.

Proposition 1. *If $q \in I_n$ satisfies*

$$R_n(q) \le q, \tag{34}$$

then the following error bounds hold for all $n \ge 0$:

$$a_n \le R_n(q) = d \le q \tag{35}$$

and

$$R_{n+1}(d - a_n) \le d - a_n. \tag{36}$$

Proposition 2. *Assume hypotheses of Theorem 1 hold and there exists $q_n \in I_n$ satisfying (34). Then the following hold:*

$$|x^* - x_m| \le \overline{t}_m, \tag{37}$$

where sequence $\{\overline{t}_n\}$ $(n \ge 0)$ is defined by

$$\overline{t}_n = q_n, \quad \overline{t}_{m+1} = R_m(\overline{t}_m) - a_m \quad (m \ge n). \tag{38}$$

Proposition 3. *Let $q \in I_n$ be any solution of (34), and assume hypotheses of Theorem 1 hold. Then the following error bounds hold for all $n \ge 0$:*

$$|x^* - x_n| \le R_n^\infty(0) \le q. \tag{39}$$

Remark 2. In order for us to consider monotone convergence for method (3) as in [3], [6] we assume X is as partially ordered. Set

$$X = V, \quad D = C^2 \quad \text{and} \quad |\cdot| = I. \tag{40}$$

Then we obtain the result:

Proposition 4. *Let Y be a Banach space, $G \in C''(V_0 \to Y)$, $A \in L(Y \to X)$, and $x, y \in V$ be such that*

$$[x, y] \in V_0; \tag{41}$$

$$I - AG'(x) \in L_+(V); \tag{42}$$

for all

$$z_1, z_2 \in [x, y], \quad z_1 \le z_2 \Rightarrow AG'(z_1) \ge AG'(z_2); \tag{43}$$

$$AG(x) \le 0, \quad AG(y) \ge 0 \tag{44}$$

and

$$(I - AG'(y))^n (y - x) \to 0 \quad \text{as } n \to \infty. \tag{45}$$

Then Newton's method $\{z_n\}$ $(n \ge 0)$ given by

$$z_0 = x, \quad z_{n+1} = z_n + [AG'(z_n)]^*[-AG(z_n)] \tag{46}$$

is well defined for all $n \ge 0$, and monotone. Moreover it converges to the unique solution x^* of AG in $[x, y]$.

Proof. Define

$$F(z) = AG(x + z), \quad M(q) = q - F(q), \quad p = y - x. \tag{47}$$

Then the results follow as special cases of Theorem 1.

That completes the proof of Proposition 4. □

Remark 3. It turns out that the error bounds obtained here can be improved under the same hypotheses of Theorem 1.

Indeed let us assume that there exists $M_0 \in C'(V_0 \to V)$ which is order convex on $[0, p]$ and for all $x \in U(p)$ with $|x| \leq p$

$$M_0'(|x|) - M_0'(0) \in B([F''(0) - F''(x)](x)), \tag{48}$$

$$M_0'(0) \in B(I - F'(0)), \quad (-F(0), M_0(0)) \in E, \tag{49}$$

and define sequence $\{r_n\}$ by

$$r_0 = 0, \quad r_{n+1} = M(r_n) + M_0'(|x_n|)(a_n). \tag{50}$$

Furthermore assume:

$$M_0(p_0) \leq M(p_0) \tag{51}$$

and

$$M_0'(p_0) \leq M'(p_0) \tag{52}$$

for all $p_0 \in [0, p]$.

It can then easily be seen from the proof of Theorem 1 that sequence $(x_n, r_n) \in (X \times V)^\pi$ is well defined for all $n \geq 0$, monotone and satisfies

$$r_n \leq t_n, \tag{53}$$

$$r^* = \lim_{n \to \infty} r_n \leq b. \tag{54}$$

Note also that M' must be replaced by M_0' in the definition of operator R_n and S_n. Conditions (51) and (52) hold in general. Moreover if strict inequality holds in (51) or (52) then so does in (53) say. That is under the same hypotheses as in Theorem 1 and Propositions 1–3 we can improve on the error bounds. Moreover one hopes that sequence $\{r_n\}$ must converge under conditions weaker than (7), (8) corresponding to $\{t_n\}$. If we can show the convergence of $\{r_n\}$ without using (7) and (8), then one hopes that the new sufficient convergent conditions must be indeed weaker than (7) and (8) because of (51) and (52).

In the next section we justify these claims.

3. Applications

As an application let us discuss the case of a Banach space with a (real) norm $\|\cdot\|$. To check the conditions of Theorem 1 and also compare it with earlier ones assume there exist parameters ℓ, ℓ_0, ℓ_1 and ℓ_2 such that for $F'(0) = I$

$$\|F''(x) - F''(y)\| \leq \ell \|x - y\|, \tag{55}$$

$$\|F''(x) - F''(x_0)\| \leq \ell_0 \|x - x_0\|, \tag{56}$$

$$\|F''(x_0)\| \leq \ell_1, \tag{57}$$

$$\|F(x_0)\| \leq \ell_2, \tag{58}$$

and

$$\|F(x) - F'(y)\| \leq \ell_3 \|x - y\|. \tag{59}$$

Define sequences $\{\alpha_n\}$, $\{\beta_n\}$, $\{\gamma_n\}$, $(n \geq 0)$ by

$$\begin{aligned} &\alpha_0 = 0, \quad \alpha_1 = \ell_2, \\ &\alpha_{n+2} = \alpha_{n+1} + \frac{\frac{\ell}{2}[\alpha_n^2 + \frac{1}{3}(\alpha_{n+1} - \alpha_n)] + \ell_1}{1 - \ell_1 \alpha_{n+1} - \frac{1}{2}\ell t_{n+1}^2}(\alpha_{n+1} - \alpha_n)^2 \end{aligned} \tag{60}$$

$$\begin{aligned} &\beta_0 = 0, \quad \beta_1 = \ell_2, \\ &\beta_{n+2} = \beta_{n+1} + \frac{\frac{\ell}{2}[\beta_n^2 + \frac{1}{3}(\beta_{n+1} - \beta_n)] + \ell_1}{1 - \ell_1 \beta_{n+1} - \frac{1}{2}\ell_0 \beta_{n+1}^2}(\beta_{n+1} - \beta_n)^2 \end{aligned} \tag{61}$$

and

$$\gamma_0 = 0, \quad \gamma_1 = \ell_2, \quad \gamma_{n+2} = \gamma_{n+1} + \frac{\ell_3(\gamma_{n+1} - \gamma_n)^2}{2(1 - \ell_3 \gamma_{n+1})}. \tag{62}$$

In general we have:

$$\ell_0 \leq \ell. \tag{63}$$

Note that if equality holds in (62) then $\alpha_n = \beta_n$ for all $n \geq 0$.

It was shown in [2], [4], [7] that if

$$3\ell_2 \ell^2 + 3\ell_1 \ell + \ell_1^3 \leq [\ell_1^2 + 2\ell]^{3/2} \tag{64}$$

then $\{\alpha_n\}$ $(n \geq 0)$ is monotonically convergent to some α^* and majorizes sequence $\{x_n\}$ $(n \geq 0)$. Moreover sequence $\{x_n\}$ $(n \geq 0)$ has a solution x^*. If strict inequality holds in (63) then under (64) sequence $\{\beta_n\}$ $(n \geq 0)$ is also monotonically convergent to some β^* and majorizes sequence $\{x_n\}$, with

$$\beta_n < \alpha_n, \tag{65}$$

$$\beta_{n+1} - \beta_n < \alpha_{n+1} - \alpha_n \tag{66}$$

and

$$\beta^* \leq \alpha^* \tag{67}$$

for all $n > 1$.

That is the error bounds are finer and the information on the location of the solution at least as precise. It can also easily be seen that in the case when strict inequality holds in (63) then sequence $\{\beta_n\}$ must hold, be monotonically convergent under sufficient conditions at least as weak as (64).

Furthermore sequences $\{\alpha_n\}$, $\{\beta_n\}$ correspond to $\{t_n\}$ and $\{r_n\}$ respectively in the (real) norm case $\|\cdot\|$. All the above justify the claims made in Remark 3.

In order for us to compare our Theorem 1 with earlier ones [2], [6] we note that $\{\gamma_n\}$ is monotonically convergent to some γ^* and majorizes sequence $\{x_n\}$ if the famous Newton–Kantorovich hypothesis

$$2\ell_3 \ell_2 \leq 1 \tag{68}$$

holds [2], [5]. However condition (64) can be weaker than (68), which implies that our Theorem 1 can guarantee convergence of sequence x_n to x^* but the corresponding Theorem 5 in [6] cannot (note that γ_n corresponds to d_n there).

We complete this study by providing such a case considered also in [2], [4], [7]:

Example 1. Let $X = [-1,1]$, $x_0 = 0$ and define G on X by

$$G(x) = \frac{1}{6}x^3 + \frac{1}{6}x^2 - \frac{5}{6}x + \frac{1}{3}. \tag{69}$$

Choose $A = G'(x_0)^{-1}$. Then using (55)–(59) and (60) we get

$$\ell = \ell_0 = \frac{6}{5}, \quad \ell_1 = \frac{2}{5}, \quad \ell_2 = \frac{2}{5} \quad \text{and} \quad \ell_3 = \frac{8}{5}. \tag{70}$$

Then we get

$$2\ell_2\ell_3 = \frac{32}{25} > 1. \tag{71}$$

Therefore, the Newton–Kantorovich hypothesis fails. Hence Theorem 5 in [6] cannot guarantee the convergence of Newton's method (3) starting at $x_0 = 0$. However condition (64) becomes

$$\frac{404}{125} < \frac{512}{125}\,.$$

That is (64) holds in this case, which guarantees the convergence of sequence x_n to x^*.

REFERENCES

[1] Argyros, I.K., Newton methods on Banach spaces with a convergence structure and applications, *Computers and Math. with Applic. Intern. J.* **40**, 1 (2000), 37–48.

[2] Argyros, I.K., *Advances in the Efficiency of Computational Methods and Applications*, World Scientific Publ. Co., River Edge, NJ, 2000.

[3] Argyros, I.K., A unifying theorem on Newton's method in a Banach space with a convergence structure under weak assumptions, *Southwest J. Pure Appl. Math.* **2** (2003).

[4] Kantorovich, L.V. and Akilov, G.P., *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.

[5] Gutierrez, J.M., A new semilocal convergence theorem for Newton's method, *J. Comp. Appl. Math.* **79** (1997), 131–145.

[6] Meyer, P.W., Newton's method in generalized Banach spaces, *Numer. Funct. Anal. and Optimization* **9** (1987), 249–259.

[7] Zhengda, Huang, A note on the Kantorivich theorem for Newton iteration, *J. Comp. Appl. Math.* **47** (1993), 211–217.