Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Issue 1, June 2001, pp. 22–28.
Submitted: January 3, 2001. Published: June 30, 2001.

LOCAL AND SEMILOCAL CONVERGENCE THEOREMS FOR NEWTON'S METHOD BASED ON CONTINUOUSLY FRÉCHET DIFFERENTIABLE OPERATORS

Ioannis K. Argyros

Abstract. We approximate locally unique solutions of nonlinear equations in a Banach space setting using Newton's method. In contrast with earlier works which are based on the Lipschitz continuity of the operators involved, we only use the continuity of the first Fréchet derivative of the operator involved. Local and semilocal convergence theorems are given and the quadratic convergence of Newton's method under weaker than before conditions is established.

A.M.S. (MOS) Subject Classification Codes:. 65H10, 47H15, 49M17.

Key Words and Phrases:. Banach space, Newton's method, Fréchet-derivative, order of convergence, local-semilocal convergence.

1. Introduction

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

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

where F is a continuously Fréchet-differentiable operator defined on an open convex subset D of a Banach space X with values in a Banach space Y.

The most popular method for approximating x^* is Newton's method

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

Local and semilocal convergence theorems for Newton's method can be found in [1]–[3] and the references there. Most known results on the convergence of iteration (2) are based on the possibility of computing the Lipschitz constant of operator $F'(x_0)^{-1}F'(x)$ $(x \in D)$ (see Remark 2). Computing such a constant is a difficult or impossible task in general despite the fact that sufficient conditions based on this constant are considered to be the weakest known so far.

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

There are many real-life problems whose solution reduces to finding x^* for a suitable operator F and on a proper space. However the computation of the Lipschitz constant is problematic. That is why we need local and semilocal results that avoid the computation of this constant and still guarantee the quadratic convergence of Newton's method. Such results are provided here, based on the continuity of the first Fréchet-derivative $F'(x)$ of F.

2. Convergence Analysis

We can show the following semilocal convergence theorem for Newton's method and continuously Fréchet-differentiable operators.

Theorem 1. *Let $F\colon D \subseteq X \to Y$ be a continuously Fréchet-differentiable operator defined on an open convex subset D of a Banach space X with values in a Banach space Y.*

Assume:

(a) *there exists $x_0 \in D$ such that $F(x_0) \neq 0$;*

(b) *$F'(x)^{-1} \in L(Y, X)$ $(x \in D)$, and there exists $b > 0$ such that*:

$$\|F'(x)^{-1} F'(x_0)\| \leq b \quad (x \in D); \tag{3}$$

(c) *for each fixed $p \in D$ with $F(p) \neq 0$, there exists*

$$\delta = \delta(p, t) \geq t \|F'(p)^{-1} F(p)\|, \quad t \in [0, 1] \tag{4}$$

such that

$$\|F'(x_0)^{-1}[F'(p) - F'(y_t)]\| < t \|F'(p)^{-1} F(p)\| \tag{5}$$

for all

$$y_t \in U(p, \delta) = \{z \in X \mid \|z - p\| < \delta\} \subseteq D, \tag{6}$$

and y_t collinear to p;

(d) *there exists $\eta > 0$ such that*:

$$\|F'(x_0)^{-1} F(x_0)\| \leq \eta, \tag{7}$$

and

$$c = \frac{b\eta}{2} \in (0, 1); \tag{8}$$

and

(e) *$U = U(x_0, \delta^*) \subseteq D$, where*

$$\delta^* \geq \frac{\eta}{1 - c}. \tag{9}$$

Then, Newton's method $\{x_n\}$ $(n \geq 0)$ *generated by (2) is well defined, remains in* U *for all* $n \geq 0$ *and converges to a unique solution* $x^* \in \overline{U}$ *of equation* $F(x) = 0$. *Moreover the following error bounds hold*:

$$\|x_{n+1} - x_n\| < \frac{b}{2}\|x_n - x_{n-1}\|^2 \quad (n \geq 1), \tag{11}$$

and

$$\|x_n - x^*\| \leq a_n\|x_n - x_{n-1}\|^2 \quad (n \geq 1), \tag{12}$$

where,

$$a_n = \frac{b}{2}\sum_{j=0}^{\infty}\left(c^{2^n}\right)^{2^j-1} \leq \frac{b}{2(1-c^{2^n})}. \tag{13}$$

Proof. Choose $p = x_0$. It follows that $t\|F'(x_0)^{-1}F(x_0)\| > 0$, $t \in (0,1]$ since $F(x_0) \neq 0$. By (5) we obtain

$$\|F'(x_0)^{-1}[F'(x_0) - F'(x_0 + t(x_1 - x_0))]\| < t\|x_1 - x_0\| \tag{14}$$

since $x_1 \in U(x_0, \delta) \subseteq D$. Using the approximation

$$\begin{aligned} x_{k+2} - x_{k+1} = -F'(x_{k+1})^{-1}\int_0^1 \Big[&F'(x_k + t(x_{k+1} - x_k)) \\ &- F'(x_k)\Big](x_{k+1} - x_k)dt \end{aligned} \tag{15}$$

for $k = 0$, (15), and (2) we obtain in turn

$$\|x_2 - x_1\| < \frac{b}{2}\|x_1 - x_0\|^2 \leq c\|x_1 - x_0\|, \tag{16}$$

and

$$\begin{aligned} \|x_2 - x_0\| &\leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq (1+c)\|x_1 - x_0\| \\ &< \frac{1-c^2}{1-c}\|x_1 - x_0\| \leq \delta^*. \end{aligned} \tag{17}$$

Hence (11) holds for $n = 1$ and $x_2 \in U$. Let us assume that (11) and $x_k \in U$ hold for all integer values k smaller or equal to $n-1$. Choose $p = x_{n-1}$. If $x_n = x_{n-1}$ then we found a solution $x^* = x_{n-1}$. If $x_n = x_{n-1}$ then we found a solution $x^* = x_{n-1}$ (by (2)). Otherwise we have $t\|x_n - x_{n-1}\| > 0$, $t \in (0,1]$. We obtain by (5)

$$\|F'(x_0)^{-1}[F'(x_{n-1}) - F'(x_{n-1} + t(x_n - x_{n-1}))]\| < t\|x_n - x_{n-1}\| \tag{18}$$

since $x_n \in U(x_{n-1}, \delta) \subseteq D$. Using (2), (3), (18) for $k = n-1$ we obtain (11). Hence (11) holds for all $n \geq 1$, and

$$\begin{aligned} \|x_{n+1} - x_0\| &\leq \|x_{n+1} - x_n\| + \|x_n - x_{n-1}\| + \cdots + \|x_1 - x_0\| \\ &< \frac{1-c^n}{1-c}\|x_1 - x_0\| \leq \delta^*. \end{aligned} \tag{19}$$

That is $x_n \in U$ for all $n \geq 1$. It follows from (11) and (9)

$$\|x_{n+1} - x_n\| < c^n \|x_1 - x_0\|.$$

Since $c \in (0,1)$ sequence $\{x_n\}$ $(n \geq 0)$ is Cauchy in a Banach space X and as such it converges to some $x^* \in \overline{U}$ (since $\overline{U}$ is a closed set). Letting $n \to \infty$ in (2) we obtain $F(x^*) = 0$. Estimate (12) is obtained by standard majorization techniques [1], [3].

The uniqueness of the solution x^* of $\overline{U}$ can be shown using the identity

$$F(y^*) - F(x^*) = \int_0^1 F'(x^* + t(y^* - x^*))(y^* - x^*)dt \tag{20}$$

where $F(x^*) = F(y^*) = 0$, $y^* \in \overline{U}$. Since

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

it follows that $x^* + t(y^* - x^*) \in \overline{U}$. That is, by (20) $\int_0^1 F'(x^* + t(y^* - x^*))dt$ is an invertible operator. Hence, (20) gives $x^* = y^*$.

That completes the proof of Theorem 1.

Remark 1. We can replace (3) as follows: if $x_n \in U(x_0, \delta)$ then

$$\|F'(x_0)^{-1}[F'(x_0) - F'(x_n)]\| < \delta^*. \tag{21}$$

Hence, if we choose $\delta^* \in (0,1)$, then by the Banach Lemma on invertible operators [2] $F'(x_n)^{-1} \in L(Y, X)$ and

$$\|F'(x_n)^{-1} F'(x_0)\| \leq \frac{1}{1 - \delta^*} = b.$$

It is simple algebra to show that estimates (8) and (10) hold provided that:

$$\delta^* \in \left[\frac{2 + \eta - \sqrt{\eta^2 - 12\eta + 4}}{4}, \frac{2 + \eta + \sqrt{\eta^2 - 12\eta + 4}}{4}\right] \tag{22}$$

and

$$\eta \in (0, 6 - 4\sqrt{2}]. \tag{23}$$

Hence, we showed:

Theorem 2. *Let F, (a), (c) be as in Theorem 1. Moreover, assume*:

(a) *parameters δ^*, η satisfy (22) and (23) respectively with $b = \frac{1}{1-\delta^*}$*;

(b) $\overline{U} = \overline{U}(x_0, \delta^*) \subseteq D$.

Then, Newton's method $\{x_n\}$ $(n \geq 0)$ *generated by (2) is well defined, remains in* U *for all* $n \geq 0$ *and converges to a unique solution* $x^* \in \overline{U}$ *of equation* $F(x) = 0$, *so that estimates (11), (12) and (13) hold for all* $n \geq 1$.

Condition (5) can be weakened as follows:

Theorem 3. *Let* $F\colon D \subseteq X \to Y$ *be continuously differentiable at* $x_0 \in D$. *Assume*:

(a) *there exists* $x_0 \in D$ *such that* $F(x_0) \neq 0$ *and* $F'(x_0)^{-1} \in L(Y, X)$;

(b) *for all* $\varepsilon > 0$ *there exists* $\delta \geq \varepsilon$ *such that*

$$\|F'(x_0)^{-1}[F'(x_0) - F'(y_t)]\| < \varepsilon, \quad t \in (0, 1]$$

for all $y_t \in U(x_0, \delta) \subseteq D$ *with* y_t *collinear to* x_0;

(c) $\eta \in (0, 5 - 2\sqrt{6})$ (*given by (7)*) *and*

$$\delta^* \in \left[\frac{1 + \eta - \sqrt{\eta^2 - 10\eta + 1}}{6}, \frac{1 + \eta + \sqrt{\eta^2 - 10\eta + 1}}{6}\right];$$

and

(d) $U = U(x_0, \delta^*) \subseteq D$.

Then, Newton's method $\{x_n\}$ $(n \geq 0)$ *generated by (2) is well defined, remains in* U *for all* $n \geq 0$ *and converges to a unique solution* $x^* \in \overline{U}$ *of equation* $F(x) = 0$. *Moreover the following error bounds hold for all* $n \geq 0$

$$\|x_{n+2} - x_{n+1}\| \leq d\|x_{n+1} - x_n\|, \quad d = \frac{2\delta^*}{1 - \delta^*},$$

and

$$\|x_n - x^*\| \leq \frac{d^\eta}{1 - d}\eta.$$

Proof. The proof follows along the lines of Theorem 1 and is based on the approximation

$$\begin{aligned} x_{n+2} - x_{n+1} = -[F'(x_{n+1})^{-1}F'(x_0)]\Big\{ F'(x_0)^{-1} \int_0^1 [F'(x_n + t(x_{n+1} - x_n)) \\ - F'(x_0)](x_{n+1} - x_n)dt + F'(x_0)^{-1}[F'(x_0) - F'(x_n)](x_{n+1} - x_n)\Big\}. \end{aligned}$$

We complete this study with the following local result for Newton's method and continuously Fréchet-differentiable operators:

Theorem 4. *Let* $F\colon D \subseteq X \to Y$ *be a continuously Fréchet-differentiable operator defined on an open convex subset* D *of a Banach space* X *with values in a Banach space* Y.

Assume:

(a) *there exists a simple zero* $x^* \in D$ *of* F *in the sense that* $F'(x^*)^{-1} \in L(Y, X)$;

(b) $F'(x)^{-1} \in L(Y, X)$ $(x \in D)$ *and there exists* $q > 0$ *such that*:

$$\|F'(x)^{-1}F'(x^*)\| \leq q \quad (x \in D);$$

(c) *for each fixed* $p \in D$ *with* $F(p) \neq 0$ *there exists* $\delta \geq t\|p - x^*\|$, $t \in (0,1]$ *such that*:

$$\|F'(x^*)^{-1}[F'(p) - F'(p_t)]\| < t\|p - x^*\|$$

for all $p_t = x + t(x^* - x) \in U(p, \delta) \subseteq D$;

(d) $U^* = U(x^*, r^*) \subseteq D$, *where*

$$r^* > \frac{2}{q}.$$

Then, Newton's method $\{x_n\}$ $(n \geq 0)$ *generated by (2) is well defined, remains in* U^* *for all* $n \geq 0$ *and converges to* x^* *provided that* $x_0 \in U^*$. *Moreover the following error bounds hold for all* $n \geq 0$:

$$\|x_{n+1} - x^*\| \leq \frac{q}{2}\|x_n - x^*\|^2.$$

Proof. The proof follows along the lines of Theorem 1 and is based on the approximation:

$$\begin{aligned} x_{n+1} - x^* = [F'(x_n)^{-1}F(x^*)]\Big\{F'(x^*)^{-1}\int_0^1 \Big[F'(x_n + t(x^* - x_n)) \\ - F'(x_n)](x^* - x_n)dt\Big\}. \end{aligned}$$

Remark 2. The results obtained here do not make use of the Lipschitz constant appearing in the famous Newton–Kantorovich Theorem [1], [3]. It is stated there that if

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

and

$$2\ell\eta \leq 1, \tag{24}$$

then equation (1) has a unique solution x^* inside a ball centered at x_0 and included in D. Our results are weaker than the above and promising especially if one compares condition (8) with (24) (if, e.g., $b < 2\ell$ and (c) is violated). Moreover condition (5) is a consequence of the continuity of the operator $F'(x_0)^{-1}F'(x)$ on D. Condition (4) may not hold in general. However since

$$F'(x_0)^{-1}[F'(p) - F'(y_t)] = F'(x_0)^{-1}\int_0^1 F''[y_t + t(p - y_t)](p - y_t)dt \tag{25}$$

it follows from (25) that (4) and (5) hold if, e.g.,

$$\overline{\ell} = \left\|F'(x_0)^{-1}\int_0^1 F''[y_t + t(p - y_t)]dt\right\| \leq 1, \quad t \in [0,1]$$

but not only then. Let us provide an example. For simplicity, let $X = Y = \mathbf{R}$, $D = [1.5, 2.5]$, $x_0 = 1.9$ and $F(x) = x^2 - 4$. Then we get $b = 1.2\overline{6}$, $\eta = .102631579$, $c = .065$ and $\delta^* = .109766395$. That is, all hypotheses of Theorem 1 are satisfied. Moreover note that $\overline{\ell} = \frac{2}{2(1.9)} = .526315789 < 1$.

REFERENCES

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

[2] Kantorovich, L.V. and Akilov, G.P. *Functional Analysis*, Pergamon Press, Oxford, 1982.

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