J. Korea Soc. Math. Educ. Ser. B: Pure Appl. Math. ISSN 1226-0657
Volume 15, Number 2 (May 2008), Pages 111–120

ON THE CONVERGENCE OF NEWTON'S METHOD AND LOCALLY HÖLDERIAN OPERATORS

IOANNIS K. ARGYROS

ABSTRACT. A semilocal convergence analysis is provided for Newton's method in a Banach space setting. The operators involved are only locally Hölderian. We make use of a point-based approximation and center-Hölderian hypotheses. This approach can be used to approximate solutions of equations involving nonsmooth operators.

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 continuous operator defined on a closed subset D of a Banach space X with values in a Banach space Y.

The most popular method for approximating x^* are undoubtedly Newton's method and its variations the so-called Newton-like methods. A survey on locan and semilocal convergence results for such methods can be found in [1]–[3] and the references there. Newton's method of classical analysis is based on its linearization $F(x) + F'(x)(z - x)$, where for given x we compute z. This is possible only if the Fréchet-derivative F' of operator F exists. If this is not the case such a linearization is no longer available.

In [5] a point-based approximation was considered to show that Newton's method converges. The method of nondiscrete mathematical induction was utilized for the semilocal convergence analysis of Newton's method. Under Newton–Kantorovich

Received by the editors March 14, 2007.
2000 *Mathematics Subject Classification.* 65G99, 65H10, 47H10, 49M15.
Key words and phrases. Newton's method, Banach space, locally Hölderian operators, point-based approximation, semilocal convergence, successive substitutions, fixed point.

type hypotheses the quadratic convergence is ensured. Certain assumptions on locally Lipschitz operators are made.

Here we provide a semilocal convergence analysis based on assumptions made on locally p-Hölderian operators for $p \in (0, 1]$. This allows us to consider a wider range of problems than before (see [4], [5], and the references there).

It turns out that even if $p = 1$ our approach provides a larger convergence radius than the corresponding one in [5], and under the same computational cost. This observation is important in computational mathematics [1]–[3].

2. Preliminary Results

We need the following definition of a point-based approximation for p-Hölderian operators:

Definition 1. Let f be an operator from a closed subset D of a metric space (X, d) to a normed linear space Y, and let $x_0 \in D$, $p \in (0, 1]$. We say f has a *point-based approximation* (PBA) on D at $x_0 \in D$ if there exist an operator $A\colon D \times D \to Y$ and scalars ℓ, ℓ_0 such that for each u, and v in D,

$$\|f(v) - A(u, v)\| \leq \ell d(u, v)^p \tag{2}$$

$$\|[A(u, x) - A(v, x)] - [A(u, y) - A(v, y)]\| \leq 2\ell d(u, v)^p \tag{3}$$

and

$$\|[A(u, v) - A(x_0, x)] - [A(u, y) - A(x_0, y)]\| \leq 2\ell_0 d(u, v)^p \tag{4}$$

for all $x, y \in D$.

Assume X is also a normed linear space and D is a convex set. Then this definition is suitable for operators f where Fréchet derivative f' is $(p-1)$-Hölderian with modulus $p\ell$ and $(p-1)$-center Hölderian with modulus $p\ell_0$. Indeed if we set

$$A(u, v) = f(u) + f'(u)(v - u),$$

then (1) says

$$\|F(v) - f(u) - f'(u)(v - u)\| \leq \ell \|v - u\|^p,$$

where as parts (2) and (3) are equivalent to the Hölderian and center Hölderian property of f'.

The following lemma can be regarded as an extension of the Banach lemma on invertible operators [3, Th. 4(2.V)] from linear operators to locally Hölderian operators.

We first need the following:

Definition 2. Let f be an operator from a metric space (X, d) into a normed linear space Y. We let

$$\delta(f, X) = \inf\left\{\frac{\|F(u) - F(v)\|}{d(u, v)^p},\ u \neq v,\ u, v \in X\right\}. \tag{5}$$

Operator f is one-to-one on X if and only if $\delta(f, X) \neq 0$. Note that by (5) it follows F^{-1} (if it exists) is $\frac{1}{p}$-Hölderian with modulus $\delta(f, X)^{-1/p}$. We also define

$$\delta_0(f, X) = \inf\left\{\frac{\|F(u) - F(x_0)\|}{d(u, x_0)^p},\ u \neq x_0,\ u, x_0 \in D\right\}.$$

Set $\delta = \delta(f, D)$ and $\delta_1 = \delta_0(f, D)$.

Lemma 1. *Let (X, d) be a Banach space, D a closed subset of X, and Y a normed linear space. Let f, y be operators from D into Y, g being p-Hölderian with modulus ℓ and center p-Hölderian with modulus ℓ_0. Let $x_0 \in D$ with $f(x_0) = y_0$. Assume that:*

$$U(y_0, \alpha) = \{y \in Y \mid \|y - y_0\| \leq \alpha\} \subseteq F(D); \tag{6}$$

$$0 \leq \ell < \delta; \tag{7}$$

$$U\big(x_0, (\delta_1^{-1}\alpha)^{1/p}\big) \subseteq D; \tag{8}$$

and

$$\theta = (1 - \ell_0\delta_1^{-1})\alpha - \|g(x_0)\| \geq 0. \tag{9}$$

Then the following hold:

$$(f + g)\big(U(x_0, (\delta_1^{-1}\alpha)^{1/p}\big) \supseteq U(y_0, \theta) \tag{10}$$

and

$$\delta(f + g, D) \geq \delta - \ell > 0. \tag{11}$$

Proof. For $y \in U(y_0, \theta)$ and $x \in U(x_0, (\delta_1^{-1}\alpha)^{1/p})$, define $Ty(x) = f^{-1}(y - g(x))$. Then we have

$$\begin{aligned}\|y - g(x) - y_0\| &\leq \|y - y_0\| + \|y(x) - g(x_0)\| + \|g(x_0)\| \\ &\leq \theta + \ell_0(\delta_1^{-1}\alpha^{1/p})^p + \|g(x_0)\| = \alpha.\end{aligned}$$

That is $Ty(x)$ is a nonempty set. But since $\delta > 0$, Ty is a singleton set. Therefore Ty is an operator on $U(x_0, (\delta_1^{-1}\alpha)^{1/p})$. Moreover by Definition 2

$$d(Ty(x), x_0) = d\big(f^{-1}(y - g(x)), f^{-1}(y_0)\big) \leq (\delta_1^{-1}\alpha)^{1/p}.$$

Hence we deduce operator Ty maps $U(x_0, (\delta_1^{-1}\alpha)^{1/p})$ into itself.

Furthermore, let $x, z \in U(x_0, (\delta_1^{-1}\alpha)^{1/p})$, then again by Definition 2

$$\begin{aligned}d(Ty(x), Ty(z)) &= d\big(f^{-1}(y - g(x)), f^{-1}(y - g(z))\big) \\ &\leq \delta^{-1/p}\ell^{1/p}\|x - z\|.\end{aligned}$$

It follows by hypothesis (10) and the fact operator Ty maps closed set $U(x_0, (\delta_1^{-1}\alpha)^{1/p})$ into itself that Ty is a strong contraction and as such it has unique fixed point $x(y)$ in $U(x_0, (\delta_1^{-1}\alpha)^{1/p})$ (by the contraction mapping principle [3, Th. 1(1.XVI)]).

Clearly $(f+g)(x(y)) = y$, and $x(y)$ is the only point in D satisfying this equation since

$$\begin{aligned}\delta(f+g, D) &= \inf\left\{\frac{\|[f(u) - f(v)]\| + \|[g(u) - g(v)]\|}{d(u,v)^p},\ u \neq v,\ u, v \in D\right\} \\ &\geq \delta(f, D) - \sup\left\{\frac{\|g(u) - g(v)\|}{d(u,v)^p},\ u \neq v,\ u, v \in D\right\} \\ &\geq \delta - \ell > 0 \quad \text{(by hypothesis (2))}.\end{aligned}$$

In particular $f + g$ is one-to-one on D. That completes the proof of Lemma 1. □

We need the following result in order to study the uniqueness of the solution x^*.

Lemma 2. *Let X and Y be normed linear spaces, and let D be a closed subset of X. Let $f\colon D \to Y$, and let A be a PBA for f on D at $x_0 \in D$. Denote by d the quantity $\delta(A(x_0, \cdot), D)$. If $U(x_0, \rho) \subseteq D$, then*

$$\delta(f, U(x_0, \rho)) \geq \delta_0 - 2\ell_0 - \frac{\ell}{2^{p-1}}$$

In particular, if $\delta_0 > 2\ell_0 + \frac{\ell}{2^{p-1}}$, then f is one-to-one on $U(x_0, \rho)$.

Proof. Let x_1 and x_2 belong to $U(x_0, \rho)$. Set $x = \frac{x_1+x_2}{2}$. Then we can write

$$f(x_1) - f(x_2) = [f(x_1) - A(x, x_1)] + [A(x, x_1) - A(x, x_2)] + [A(x, x_2) - f(x_2)].$$

By Definition 1 we have for $i = 1, 2$,

$$\|f(x_i) - A(x, x_i)\| \le \ell \|x - x_i\|^p \le \frac{\ell}{2^p}\|x_1 - x_2\|^p.$$

By the triangle inequality we get

$$\|A(x, u) - A(x, v)\| \ge \|A(x_0, u) - A(x_0, v)\| - \|[A(x, u) - A(x_0, u)] - [A(x, v) - A(x_0, v)]\|,$$

and

$$\begin{aligned}\delta(A(x, \cdot), D) &= \inf\left\{\frac{\|A(x, u) - A(x, v)\|}{\|u - v\|^p},\ u \ne v,\ u, v \in D\right\} \\ &\ge \delta(A(x_0, \cdot), D) - \sup\left\{\frac{\|[A(x, u) - A(x_0, u)] - [A(x, v) - A(x_0, v)]\|}{\|u - v\|^p},\right. \\ &\qquad \left. u \ne v,\ u, v \in D\right\} \\ &\ge \delta_0 - 2\ell_0.\end{aligned}$$

Hence, we get

$$\begin{aligned}\|f(x_1) - f(x_2)\| &\ge (\delta_0 - 2\ell_0)\|x_1 - x_2\|^p - \frac{\ell}{2^{p-1}}\|x_1 - x_2\|^p \\ &\ge \left(\delta_0 - 2\ell_0 - \frac{\ell}{2^{p-1}}\right)\|x_1 - x_2\|^p\end{aligned}$$

and for $x_1 \ne x_2$,

$$\frac{\|f(x_1) - f(x_2)\|}{\|x_1 - x_2\|^p} \ge \delta_0 - 2\ell_0 - \frac{\ell}{2^{p-1}}.$$

That completes the proof of Lemma 2. □

Remark 1. If $p = 1$, then for $\ell_0 < \ell$ or $\delta_0 > \delta$ Lemma 1 improves (enlarges) the range of θ, and Lemma 2 enlarges the radius of convergence for Newton's method given in corresponding Lemma 3.1 in [4, p. 298] and Lemma 2.3 in [5, p. 294] respectively. These observations are important in computational mathematics [1]. Note also that in general $\ell_0 \le \ell$, $\delta_0 \ge \delta$ hold and $\frac{\ell}{\ell_0}$, $\frac{\delta_0}{\delta}$ can be arbitrarily large [2].

We need the following result on fixed points:

Theorem 1. *Let $Q\colon X \subset X \to Y$ be a continuous operator, let $p \in [0,1)$, $q \geq 0$ and $x_0 \in D$ be such that*

$$\|Q(x) - Q(y)\| \leq q\|x - y\|^p \quad \text{for all } x, y \in D \tag{12}$$

$$d \in [0,1), \tag{13}$$

and

$$U(x_0, r) \subseteq D \tag{14}$$

where,

$$b = q^{\frac{1}{1-p}}, \tag{15}$$

$$d = b^{-1}\|x_0 - Q(x_0)\|, \quad \text{and} \quad r = \frac{b}{1-d^p}. \tag{16}$$

Then sequence $\{x_n\}$ $(n \geq 0)$ generated by successive substitutions:

$$x_{n+1} = Q(x_n) \quad (n \geq 0) \tag{17}$$

converges to a fixed point $x^* \in U(x_0, r)$ of operator Q, so that for all $n \geq 0$:

$$\|x_{n+1} - x_n\| \leq d^{p^n} b, \tag{18}$$

and

$$\|x_n - x^*\| \leq \frac{d^{np}}{1-d^p} b. \tag{19}$$

Moreover if

$$d_1 = b^{-1}(\|x_0 - Q(x_0)\| + qr) < 1, \tag{20}$$

or

$$q^{1/p} r < 1 \tag{21}$$

then x^* is the unique fixed point of Q in $U(x_0, r)$.

Proof. By hypothesis (14) $x_1 \in U(x_0, r)$. Assume $x_k \in U(x_0, r)$, $k = 0, 1, \ldots, n$. Then x_{n+1} is defined by (18), and using (12) we can have in turn:

$$\begin{aligned}\|x_{n+1} - x_n\| &= \|Q(x_n) - Q(x_{n-1})\| \leq q\|x_n - x_{n-1}\|^p \\ &\leq q\big[q\|x_{n-1} - x_{n-2}\|^p\big]^p = q^{\frac{1-p^2}{1-p}} \|x_{n-1} - x_{n-2}\|^{p^2} \\ &\leq \cdots \leq q^{\frac{1-p^n}{1-p}} \|x_1 - x_0\|^{p^n} \leq d^{p^n} b, \end{aligned} \tag{22}$$

which shows (18).

Moreover for all $m = 0, 1, 2, \ldots$ we have

$$
\begin{aligned}
\|x_{n+m} - x_n\| &\le \|x_{n+m} - x_{n+m-1}\| + \|x_{n+m-1} - x_{n+m-2}\| + \cdots + \|x_{n+1} - x_n\| \\
&\le \left[d^{p^{n+m-1}} + \cdots + d^{p^n}\right] b \\
&\le \left[d^{(n+m-1)p} + \cdots + d^{np}\right] b = \frac{1-d^{mp}}{1-d^p} d^{np} b. \qquad (23)
\end{aligned}
$$

It follows from (13) and (23) that sequence $\{x_n\}$ is Cauchy in a Banach space X, and as such it converges to some $x^* \in U(x_0, r)$ (since $U(x_0, r)$ is a closed set). By letting $m \to \infty$ in (23) we get (19). In particular for $n = 0$, and $m = n + 1$ (22) gives $x_{n+1} \in U(x_0, r)$. That is $x_n \in U(x_0, r)$ for all $n \ge 0$. By letting $n \to \infty$ in (17) we get $x^* = Q(x^*)$.

To show uniqueness let $y^* \in U(x_0, r)$ be a fixed point of Q. Then as in (22) we get

$$\|x_{n+1} - y^*\| = \|Q(x_n) - Q(y^*)\| \le d_1^{p^n} b. \qquad (24)$$

By letting $n \to \infty$ in (24), and using hypothesis (20) we get $\lim_{n\to\infty} x_n = y^*$. But we showed $\lim_{n\to\infty} x_n = x^*$. Hence, we deduce $x^* = y^*$. Similarly using (21) instead of (20) we deduce again $x^* = y^*$. That completes the proof of Theorem 1. □

Remark 2. (a) The case $p = 1$ is not covered above since it is already well known by the contraction mapping principle [3]. Simply assume $q \in [0, 1)$ and $U(x_0, r_1) \subseteq D$ where $r_1 = \frac{\|x_0 - Q(x_0)\|}{1-q}$. Results corresponding to (19) and (20) will be

$$\|x_{n+1} - x_n\| \le q^n \|x_1 - x_0\| \qquad (25)$$

and

$$\|x_n - x^*\| \le \frac{q^n}{1-q} \|x_0 - Q(x_0)\|, \qquad (26)$$

respectively.

(b) If instead of (13) and (20) we assume

$$h = 2d^p < 1, \qquad (27)$$

and

$$2b \le q^{-1/p} \qquad (28)$$

then the conclusions of Theorem 2 hold in the ball $U(x_0, r_0)$ where

$$r_0 = 2b. \qquad (29)$$

We set

$$\|x_0 - Q(x_0)\| \leq \eta. \tag{30}$$

3. SEMILOCAL CONVERGENCE ANALYSIS

We can state and prove the main semilocal convergence result for Newton's method involving a p-(PBA) approximation for f.

Theorem 2. *Let X and Y be Banach spaces, D a closed convex subset of X, $x_0 \in D$, and F a continuous operator from D into Y. Suppose that F has a p-(PBA) approximation at x_0. Moreover assume:*

(a)

$$\delta(A(x_0,\cdot),D) \geq \delta_0 > 0, \tag{31}$$

$$\eta \leq b, \quad 2\ell_0 < \delta_0, \quad (1 - 2\ell_0\delta_1^{-1}r_0)\delta_0\eta - \ell\eta^p \geq 0 \tag{32}$$

and conditions (9), (27), (28) *are satisfied for $\alpha = \delta_0\eta$ and $q = \ell(\delta_0 - 2\ell_0)^{-1}$, respectively where,*

$$\delta_1 = \delta_0(A(x_0,\cdot),D).$$

(b) *for each $y \in U(0,\delta_0\eta)$ the equation $A(x_0,x) = y$ has a solution x;*

(c) *the solution $T(x_0)$ of $A(x_0,T(x_0)) = 0$ satisfies $\|x_0 - T(x_0)\| \leq \eta$; and*

(d) *$U(x_0,r_0) \subseteq D$, where r_0 is given by* (29).

Then the Newton iteration defining x_{n+1} by

$$A(x_n,x_{n+1}) = 0 \tag{33}$$

remains in $U(x_0,r_0)$, and converges to a solution $x^* \in U(x_0,r_0)$ of equation $F(x) = 0$, so that estimates (18) *and* (19) *hold.*

Proof. We use Lemma 1 with quantities f, g, x_0 and y_0 replaced by $A(x,\cdot)$, $A(x_1,\cdot) - A(x,\cdot)$, $x_1 = T(x_0)$, and 0 respectively. Hypothesis (6) of Lemma 1 follows from the fact that $A(x_0,x) = y$ has a unique solution x_1 (since $\delta_0 > 0$). For hypothesis (7) we have

$$\delta(A(x_1,\cdot),D) \geq \delta(A(x_0,\cdot),D) - 2\ell_0 \geq \delta_0 - 2\ell_0 > 0. \tag{34}$$

To show (8) we must have:

$$U(x_1,\eta) \subset U(x_0,r_0),$$

which is true since

$$\|x_1 - x_0\| + \eta \le 2\eta \le 2b = r_0.$$

We also get by (2)

$$\|A(x_0, x_1) - A(x_1, x_1)\| \le \ell \|x_1 - x_0\|^p \le \ell \eta^p. \tag{35}$$

We need to find an upper bound on θ which is given by

$$\theta \ge \left[1 - 2\ell_0 \delta_1^{-1} r_0\right] \delta_0 \eta - \|A(x_0, x_1) - A(x_1, x_1)\|. \tag{36}$$

However the upper bound on θ given by (36) is nonnegative by (32) and (35), which shows hypothesis (9) of Lemma 1. Condition (10) holds by the last hypothesis in (a) above. Hence all hypotheses of Lemma 1 are satisfied. It follows that for each $y \in U(0, r_0 - \|x_1 - x_0\|)$, the equation $A(x, z) = y$ has a unique solution, since $\delta(A(x_1, \cdot), D) > 0$. We also have $A(x_0, x_1) = A(x_1, x_2) = 0$ and $A(x_1, x_1) = F(x_1)$. By (5) we have

$$\begin{aligned} \|x_2 - x_1\| &\le \delta(A(x_1, \cdot), D)^{-1} \|A(x_0, x_1) - F(x_1)\| \\ &\le q \|x_1 - x_0\|^p. \end{aligned} \tag{37}$$

Hence we showed (18) and

$$U(x_{n+1}, r_0 - \|x_{n+1} - x_0\|) \subseteq U(x_n, r_0 - \|x_n - x_0\|) \tag{38}$$

hold for $n = 0, 1$. Moreover for every $v \in U(x_1, r_0 - \|x_1 - x_0\|)$ it follows

$$\|v - x_0\| \le \|v - x_1\| + \|x_1 - x_0\| \le r_0 - \|x_1 - x_0\| + \|x_1 - x_0\| = r_0,$$

implies $v \in U(x_0, r_0)$. Given they hold for all $n = 0, 1, \ldots, j$, then

$$\|x_{j+1} - x_0\| \le \sum_{i=1}^{j+1} \|x_i - x_{i-1}\| \le r_0.$$

Then the induction can easily be completed by simply replacing above x_0, x_1 by x_n, x_{n+1} respectively. Indeed the crucial upper bound on θ is

$$\left[1 - 2\ell_0 \delta_1^{-1} r_0\right] \delta_0 \eta - \|A(x_n, x_{n+1}) - A(x_{n+1}, x_{n+1})\|$$

which is bounded above by (since $\delta_1 \ge \delta_0$)

$$[1 - 2\ell_0 \delta_1^{-1} r_0] \delta_0 \eta - \ell \|x_{n+1} - x_n\|^p$$

and the latter is again bounded above by

$$[1 - 2\ell_0 \delta_1^{-1} r_0] \delta_0 \eta - \ell \eta^p$$

which is nonnegative by (32).

It follows by (18) and (38) that $\{x_n\}$ is a Cauchy in a Banach space X and as such it converges to some $x^* \in U(x_0, r_0)$ (since $U(x_0, r_0)$ is a closed set). By (2) and Theorem 1 we get

$$\begin{aligned} \|F(x_{n+1})\| = \|F(x_{n+1}) - A(x_n, x_{n+1})\| &\leq \ell \|x_{n+1} - x_n\|^p \\ &\leq q\|x_{n+1} - x_n\|^p \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

Therefore $\|F(x_{n+1})\|$ converges to zero as $n \to \infty$. By the continuity of F we deduce $F(x^*) = 0$. That completes the proof of Theorem 2. □

Remark 3. The uniqueness of the solution x^* was not considered in Theorem 2. Using Lemma 2 we can obtain a uniqueness result so that if

$$r_0 < \rho \quad \text{and} \quad \delta_0 > 2\ell_0 + \frac{\ell}{2^{p-1}},$$

then operator F is one-to-one in a neighborhood of x^*, since $x^* \in U(x_0, r_0)$. That is x^* is an isolated zero of F in this case.

REFERENCES

1. Argyros, I.K.: *Advances in the Efficiency of Computational Methods and Applications.* World Scientific Publ. Co., River Edge, NJ, 2000.
2. Argyros, I.K.: A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space. *J. Math. Anal. and Applic.* **298** (2004), no. 2, 374–397.
3. Kantorovich, L.V. & Akilov, G.P.: *Functional Analysis in Normed Spaces.* Pergamon Press, Oxford, 1982.
4. Robinson, S.M.: An implicit function theorem for a class of nonsmooth functions. *Mathematics of Operations Research.* **16** (1991), no. 2, 292–309.
5. Robinson, S.M.: Newton's method for a class of nonsmooth functions. *Set-Valued Analysis* **2** (1994), 291–305.

DEPARTMENT OF MATHEMATICAL SCIENCES, CAMERON UNIVERSITY, LAWTON, OK 73505, U.S.A.
Email address: iargyros@cameron.edu