

ON NEWTON'S METHOD UNDER MILD DIFFERENTIABILITY CONDITIONS

Ioannis K. Argyros*

*Department of Mathematics
New Mexico State University
Las Cruces, U.S.A.*

الخلاصة :

يُقَدِّم هذا البحث شروطاً تقاربِ طريقة معاودة نيوتن لحل معادلة مؤثر غير خطية في فراغ (بناخ) . والافتراض الوحيد في البحث هو أن مشتقات (فريشيه) للمؤثر غير الخطي تحقق اتصال (هولدر) . كما ويحتوي البحث على بعض أمثلة حيث أن الفرضية العادية لطريقة نيوتن لا تتحقق . ولكن الفرضية المقترحة يمكن تحقيقها .

ABSTRACT

We provide sufficient conditions for the convergence of Newton's iteration to a solution of nonlinear operator equation in Banach space. We assume only that the Fréchet-derivative of the nonlinear operator is Hölder continuous. Some examples are provided where the usual hypotheses for the application of Newton's method are not satisfied but ours are.

*Address for correspondence:
Department of Mathematical Sciences
College of Arts and Science
New Mexico State University
Box 30001, Dept. 3 MB
Las Cruces, New Mexico 88003-0001
U.S.A.

ON NEWTON'S METHOD UNDER MILD DIFFERENTIABILITY CONDITIONS

INTRODUCTION

The Newton–Kantorovich method, namely

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad (1)$$

has been used extensively to solve the nonlinear operator equation

$$F(x) = 0 \quad (2)$$

in a Banach space X [4–6] (and the references there).

Using some ideas of Altman [3], we generalize his results, assuming only that the Fréchet-derivative $F'(x)$ is Hölder (c, p) continuous on X (to be made precise later).

If $p = 1$ and the inverse of $F'(x)$ exists on X then our results reduce to the ones obtained by Kantorovich and others [1], [3], [4].

Some examples are also provided.

Let X and Y be two Banach spaces and let L_1 be a continuous linear operator mapping X onto Y . Denote by e_{L_1} the set of all solutions of the equation $L_1x = 0$. We divide the space X into classes, and we say that x_1 and x_2 belong to the same class \bar{X} , if $x_1 - x_2 \in e_{L_1}$. This quotient space X/e_{L_1} is a Banach space with the norm $\|\bar{X}\| = \inf\|x\|$, $x \in \bar{X}$. The operator L_1 gives rise to an operator $L: X/e_{L_1} \rightarrow Y$ which is bijective and $L\bar{X} = L_1x$ for $x \in \bar{X}$.

We now state the lemmas whose proof can be found in [3].

Lemma 1. Let L_1 and L_2 be two linear operators mapping X onto Y . If

$$\|L_2 - L_1\| < \frac{1}{\|\bar{L}_1^{-1}\|}, \quad (3)$$

then

$$\|\bar{L}_2^{-1}\| \leq \frac{\|\bar{L}_1^{-1}\|}{1 - \|\bar{L}_1^{-1}\|\|L_2 - L_1\|}, \quad (4)$$

where \bar{L}_1 and \bar{L}_2 denote the adjoints of L_1 and L_2 respectively.

Let

$$P = A\bar{X} \quad (5)$$

be the linear transformation of X/e_{L_2} onto Y , induced by the operator L_2 . Then easily

$$\|A^{-1}\| = \|\bar{L}_2^{-1}\|. \quad (6)$$

Now, let $y = F(x)$ be a nonlinear continuous operator on Y . We suppose that $F(x)$ is Fréchet-differentiable in a certain closed sphere $S(x_0, r)$ with center x_0 and of radius $r > 0$. We suppose also that the Fréchet-derivative $F'(x)$ is a linear operator onto Y for every $x \in S(x_0, r)$.

Denote by e_x the set of all solutions z of the equation $F'(x)z = 0$ and consider the quotient space X/e_x . Let us assume that the following is satisfied:

For every $x \in S(x_0, r)$ the norm of $l \in X/e_x$ is reached at a point $z \in l$, i.e., there exists an element $z \in l$ such that

$$\|z\| = \|l\| = \inf\|z'\|, \quad z' \in l. \quad (7)$$

We can now define an iteration for solving (2). We denote by A_n the linear operator defined on X/e_{x_n} and induced by the linear transformation $F'(x_n)$ and put $e_n = e_{x_n}$, $n = 0, 1, 2, \dots$.

Set

$$\bar{x}_1 = \bar{x}_0 - A_0^{-1}F(x_0),$$

where $x_0 \in \bar{x}_0$ and $\bar{x}_0, \bar{x}_1 \in X/e_0$, and choose an element x_1 in \bar{x}_1 such that

$$\|x_1 - x_0\| = \|\bar{x}_1 - \bar{x}_0\|.$$

If the approximate solutions x_1, \dots, x_n are already defined, then we put

$$\bar{x}_{n+1} = \bar{x}_n - A_n^{-1}F(x_n), \quad (8)$$

where $\bar{x}_n \in X/e_n$ and $x_n \in \bar{x}_n$. Further, we choose elements x_{n+1} and $\bar{x}_{n+1} \in X/e_n$ such that

$$\|x_{n+1} - x_n\| = \|\bar{x}_{n+1} - \bar{x}_n\|. \quad (9)$$

It is well known that at branch and limit points of nonlinear functional equations the first Fréchet-derivative is singular and an interest in the computation of such solution points (see, for example [7] and the references therein) has provided some of the motivation for the construction of operators A_n in this paper.

Conditions for the convergence of iterations (8) and (9), which are also sufficient conditions for the existence of a solution of (2) will be given in the main theorem that follows. But first we need the following:

Definition 1. Assume that F is Fréchet-differentiable and $F'(x)$ is the first Fréchet-derivative at a point x . We say that the Fréchet-derivative is Hölder continuous over a domain R if for some $c > 0$, $p \in [0, 1]$, and all $x, y \in R$

$$\|F'(x) - F'(y)\| \leq c \|x - y\|^p. \quad (10)$$

In this case, we say that $F'(x) \in H_R(c, p)$.

We will need the following lemma ([4], p. 142).

Lemma 2. Let $F: X \rightarrow Y$ and $\tilde{D} \subseteq X$. Assume \tilde{D} is open and that $F'(x) \in H_{\tilde{D}_0}(c, p)$ for some convex $\tilde{D}_0 \subseteq \tilde{D}$. Then for all $x, y \in \tilde{D}_0$

$$\|F(x) - F(y) - F'(x)(x - y)\| \leq \frac{c}{p+1} \|x - y\|^{p+1}. \quad (11)$$

MAIN RESULTS

Definition 2. Define the functions g_1 , g_2 , and g_3 by,

$$g_1(d) = g_1 z(d) = v z^{p+1} + z^p - v, \\ \text{where } z = d^{1/(p+1)}, v = (p+1)^{p/(p+1)},$$

$$g_2(d) = d^p + v d - v,$$

$$g_3(p) = 1 - \sqrt{\left(\frac{1}{(p+1)v}\right)}.$$

Claim. There exists d , with $0 < d < 1$ such that

$$g_1(d) \leq 0 \quad (12)$$

$$g_2(d) \leq 0 \quad (13)$$

and

$$d \leq g_3(p). \quad (14)$$

Note:

$$g_1(0) = -v < 0, \\ g_1'(z) = (p+1)vz^p + pz^{p-1} > 0, z \in (0, +\infty).$$

Therefore, g_1 is increasing on $(0, +\infty)$. But,

$$g_1(1) = 1 > 0.$$

That is, there exists, $0 < d_1 < 1$ such that

$$g_1(d) \leq 0 \text{ for all } d \in (0, d_2).$$

Similarly there exists $0 < d_2 < 1$ such that

$$g_2(d) \leq 0 \text{ for all } d \in (0, d_1).$$

Set,

$$d_3 = \min(d_0, d_1, d_2, g_3(p)), d_0 = \frac{v-1}{v}. \quad (15)$$

Then

$$g_1(d) \leq 0,$$

$$g_2(d) \leq 0,$$

and

$$d \leq g_3(p) < 1 \text{ for all } d \in (0, d_3).$$

By the choice of d above, it is possible to choose w, y such that

$$\frac{1}{(1-d)v} \leq w \leq \min(1, (p+1)(1-d)), \quad (16)$$

$$\frac{1}{(p+1)(1-d)} \leq y < \frac{1}{w}. \quad (17)$$

Theorem. Let us assume the following conditions are satisfied:

(a) For the Fréchet-derivative $F'(x_0)^{-1}$ there exists $B_0 > 0$ such that

$$\|F'(x_0)^{-1}\| \leq B_0; \quad (18)$$

(b) the first approximate solution x_1 satisfies

$$\|x_1 - x_0\| \leq k_0; \quad (19)$$

(c) for $R = S(x_0, r)$ the Fréchet-derivative $F'(x) \in H_R(c, p)$ with $x \in R$, for some $p \in [0, 1]$;

(d) the constants B_0, k_0, c satisfy the inequality

$$h_0 = cB_0k_0^p \leq d \leq d_0 = \frac{v-1}{v} < 1, \quad (20)$$

and

$$\frac{k_0}{1-ywD^p} \leq r, \quad (21)$$

where $d \in (0, d_3)$ and d_3, y, w, v , are as in definition 2 with $D = d^{1/p}$.

Then (2) has a solution $x^* \in S(x_0, r)$. The sequence $\{x_n\}$, $n = 0, 1, 2, \dots$ defined by (8) and (9), remains in $S(x_0, r)$ and converges to x^* , with

$$\|x^* - x_n\| \leq \frac{k_0(yw)^n D^{(p+1)^n}}{1-ywD^p}, n = 0, 1, 2, \dots \quad (22)$$

Moreover if

$$cB_0(3r)^p < 1 \quad (23)$$

then

$$\|x_n - x^*\| \leq \frac{cB_0}{(p+1)[1 - cB_0(2r + \|x_n - x_0\|)^p]} \|x_n - x_{n-1}\|^{p+1},$$

$$n = 1, 2, \dots \quad (24)$$

Proof. We will follow the standard inductive Newton-Kantorovich proof used also in ([3], p. 792) and ([4], p. 143) with some modifications. By (8), we have

$$A_n(\bar{x}_{n+1} - \bar{x}_n) = -F(x_n), \quad n = 0, 1, 2, \dots \quad (25)$$

Hence, applying the definition of the transformation A_n , we obtain

$$F'(x_n)(x_{n+1} - x_n) = -F(x_n). \quad (26)$$

Using (11) and the approximation

$$F(x_n) = F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1}) \quad (27)$$

we derive

$$\|F(x_n)\| \leq \frac{c}{p+1} \|x_n - x_{n-1}\|^{p+1}. \quad (28)$$

By (8), (9), and (28), we obtain

$$\|x_{n+1} - x_n\| \leq \frac{c}{p+1} \|A_n^{-1}\| \cdot \|x_n - x_{n-1}\|^{p+1}.$$

We shall now estimate the norm $\|A_n^{-1}\|$ using lemma 1 (5) and (6) to obtain

$$\|A_n^{-1}\| \leq \frac{\|\bar{F}'(x_{n-1})^{-1}\|}{1 - \|\bar{F}'(x_{n-1})^{-1}\| \|F'(x_n) - F'(x_{n-1})\|}$$

$$\leq \frac{B_{n-1}}{1 - B_{n-1}ck_{n-1}^p}$$

where

$$\|A_{n-1}^{-1}\| = \|\bar{F}'(x_{n-1})^{-1}\| \leq B_{n-1},$$

$$\|x_n - x_{n-1}\| \leq k_{n-1},$$

and

$$h_{n-1} = B_{n-1}ck_{n-1}^p.$$

Using

$$\|x_{n+1} - x_n\| \leq k_n,$$

$$B_n = \frac{B_{n-1}}{1 - h_{n-1}},$$

and

$$k_n = \frac{1}{p+1} \cdot \frac{h_{n-1}}{1 - h_{n-1}} \cdot k_{n-1},$$

we obtain

$$h_n = B_nck_n^p = \left(\frac{1}{p+1}\right)^p \cdot \left(\frac{h_{n-1}}{1 - h_{n-1}}\right)^{p+1},$$

$$n = 0, 1, 2, \dots$$

We now note the following:

Claim 1. For, $h_0 \leq d < 1$, assuming $h_{n-1} \leq d$ we can show

$$h_n \leq d.$$

To show,

$$\left(\frac{1}{p+1}\right)^p \left(\frac{h_{n-1}}{1 - h_{n-1}}\right)^{p+1} \leq d$$

it suffices to show

$$g_1(d) = g_1(z) \leq 0,$$

$$\text{with } z = d^{1/(p+1)}, \quad v = (p+1)^{p/(p+1)},$$

which is true for $d \in (0, d_3)$.

Claim 2. For w , given by (16),

$$h_n \leq wh_{n-1}^{p+1}.$$

$$\left(\frac{1}{p+1}\right)^p \left(\frac{h_{n-1}}{1 - h_{n-1}}\right)^{p+1} \leq wh_{n-1}^{p+1},$$

assuming

$$h_{n-1} \leq d,$$

it suffices to show

$$w \geq \frac{1}{(1-d)v},$$

which is true by the choice of w .

For consistency,

$$h_n \leq wh_{n-1}^{p+1} \leq wd^{p+1} \leq d,$$

or

$$w \leq \frac{1}{d^p}.$$

That is, we must have

$$\frac{1}{(1-d)v} \leq w \leq \frac{1}{d^p},$$

which is true by the choice of w and the fact that

$$g_2(d) \leq 0 \text{ for } d \in (0, d_3).$$

Claim 3. For y , given by (17)

$$k_n = \frac{1}{p+1} \frac{h_{n-1}k_{n-1}}{1-h_{n-1}} \leq yh_{n-1}k_{n-1}. \quad (29)$$

To show (29) it is enough to choose

$$(1-d)(p+1)y \geq 1,$$

which is true by the choice of y .

Now,

$$h_n \leq wh_{n-1}^{p+1} \leq w(h_{n-1}^{p+1})^{p+1} \leq \dots \leq w(h_0)^{(p+1)^n}$$

and

$$\begin{aligned} k_n &\leq yh_{n-1}k_{n-1} \leq y^2h_{n-1}h_{n-2}k_{n-2} \\ &\leq y^n h_{n-1}h_{n-2} \dots h_0 k_0 \\ &\leq (yw)^n d^{[(p+1)^n - 1]/p} k_0. \end{aligned}$$

Now, using a standard argument as in ([4], p.143), we can show by induction that the approximate solution x_n and the corresponding numbers B_n , k_n , and h_n can be defined for every n , and, moreover conditions (a)-(d) are satisfied.

Moreover,

$$\begin{aligned} \|x_{n+q} - x_n\| &\leq k_{n+q-1} + k_{n+q-2} + \dots + k_n \\ &\leq (yw)^{n+q-1} d^{[(p+1)^{n+q-1} - 1]/p} k_0 \\ &\quad + (yw)^{n+q-2} d^{[(p+1)^{n+q-2} - 1]/p} k_0 \\ &\quad + \dots + (yw)^n d^{[(p+1)^n - 1]/p} k_0. \\ &\leq D^{(p+1)^n - 1} k_0 (yw)^n [(yw)^{q-1} D^{(p+1)^{q-1}} \\ &\quad + (yw)^{q-2} D^{(p+1)^{q-2}} + \dots + 1] \\ &\leq D^{(p+1)^n} k_0 (yw)^n [(yw)^{q-1} D^{p(q-1)} \\ &\quad + (yw)^{q-2} D^{p(q-2)} + \dots + 1] \end{aligned}$$

(since for $D < 1$, $p \in [0, 1]$, $D^{(p+1)^{q-1}} \leq D^{p(q-1)}$)

$$\leq D^{(p+1)^n} k_0 (yw)^n \left[\frac{1 - (ywD^p)^n}{1 - ywD^p} \right]. \quad (30)$$

The right hand side of (30) and the choice of y , w , d show that the sequence $\{x_n\}$, $n = 0, 1, 2, \dots$ is a Cauchy sequence and as such it converges to some $x^* \in X$. Letting $q \rightarrow \infty$ in (30), we obtain (22).

By (30), setting $n = 0$ and $q = n$, we can easily obtain that

$$\|x_n - x_0\| \leq r,$$

that is $x_n \in S(x_0, r)$, $n = 0, 1, 2, \dots$ if (21) holds.

By the continuity of $F(x)$ and since the norms of the operators $F'(x_n)$ are bounded, it follows that x^* is a solution of (2).

Moreover, $x^* \in S(x_0, r)$ since $S(x_0, r)$ is closed and (22) holds.

Let us now observe that

$$F(x_n) = F(x_n) = F(x^*) = Q_n(x_n - x^*), \quad (31)$$

where

$$Q_n = \int_0^1 F'(x^* + t(x_n - x^*)) dt, \quad n = 0, 1, 2, \dots$$

We want to prove that the linear operator Q_n is invertible for all n .

To this effect we note that according to (10) and (23) we have:

$$\begin{aligned} \|F'(x_0)^{-1} \left(\int_0^1 [F'(x^* + t(x_n - x^*)) - F'(x_0)] dt \right)\| \\ \leq cB_0 \int_0^1 \|x^* + t(x_n - x^*) - x_0\|^p dt \\ \leq cB_0 (2\|x_n - x^*\| + \|x_n - x_0\|)^p \\ \leq cB_0 (3r)^p < 1. \end{aligned}$$

By virtue of Banach's lemma, B_n is invertible and the following norm estimation holds:

$$\|Q_n^{-1}\| \leq \frac{B_0}{1 - cB_0(2r + \|x_n - x_0\|)^p}. \quad (32)$$

Finally from (31), (28), and (32) we deduce (24), since

$$\|x_n - x^*\| \leq \|Q_n^{-1}\| \cdot \|F(x_n)\|.$$

That completes the proof of the theorem.

A theorem similar to the above can easily be stated here for the modified Newton's iteration

$$x_{n+1} = x_n - F'(x_0)^{-1} F(x_n), \quad n = 0, 1, 2, \dots$$

However, we leave that to the motivated reader. We now provide some examples.

APPLICATIONS

Example 1.

Consider the function G defined on $[0, b]$ by

$$G(t) = \frac{2}{3}t^{3/2} + t - 3$$

for some $b > 0$.

After four iterations we get a vector

$$z_4 = \begin{bmatrix} 3.35740 \text{ E} + 01 \\ 6.52027 \text{ E} + 01 \\ 9.15664 \text{ E} + 01 \\ 1.09168 \text{ E} + 02 \\ 1.15363 \text{ E} + 02 \\ 1.09168 \text{ E} + 02 \\ 9.15664 \text{ E} + 01 \\ 6.52027 \text{ E} + 01 \\ 3.35740 \text{ E} + 01 \end{bmatrix} .$$

We choose z_4 as our x_0 for the theorem. We get the following results:

$$B_0 = \|F'(x_0)^{-1}\| = 2.55882 \text{ E} + 01 ,$$

$$K_0 = \|x_1 - x_0\| = 9.15311 \text{ E} - 05 ,$$

$$c = \frac{3}{2} h^2 = 0.0015 ,$$

$$p = \frac{1}{2} ,$$

$$v = 1.44714243 ,$$

$$d_0 = 0.126419535 .$$

Choose $d = 4\text{E}-03$ to get from (20), thus

$$h_0 = 3.672107076 \text{ E} - 03 < 4\text{E} - 03 d_0 < 1 .$$

Finally choose $y = 1$ and $w = 0.9$ to obtain from (21)

$$r \geq 9.18618024 \text{ E} - 05 = r_0 .$$

With the above values it can easily be seen that all the hypotheses of the theorem are satisfied. Therefore the sequence of iterates remains in $S(x_0, r_0)$ and converges to a nontrivial solution x^* of Equation (36).

REFERENCES

- [1] M. Altman, "A Generalization of Newton's Method", *Academie Polonaise des Sciences Bulletin, CI III, 2* (1955), p. 189.
- [2] M. Altman, "On the Approximate Solution of Non-linear Functional Equations", *Academie Polonaise des Sciences Bulletin, CI III, 5* (1957), p. 457.
- [3] M. Altman, "On the Generalization of Newton's Method", *Academie Polonaise des Sciences Bulletin, CI III, 5(8)* (1957), p. 789.
- [4] M. A. Krasnoleskii, G. M. Vainikko *et al.*, *Approximate Solution of Operator Equations*. Groningen: Wolters-Noordhoff Publishing, 1972.
- [5] W. C. Rheinboldt and J. M. Ortega, *Iterative Solution of Nonlinear Equations in Several Variables*. New York: Academic Press, 1970.
- [6] W. C. Rheinboldt, *Numerical Analysis of Parametrized Nonlinear Equations*. New York: Wiley, 1986.
- [7] R. Seydel, "Numerical Computation of Branch Points in Ordinary Differential Equations", *Numerische Mathematik, 32* (1979), p. 439.

Paper Received 15 March 1988; Revised 28 November 1988.