TAMKANG JOURNAL OF MATHEMATICS
Volume 23, Number 4, Winter 1992

REMARKS ON THE CONVERGENCE OF NEWTON'S METHOD UNDER HÖLDER CONTINUITY CONDITIONS

IOANNIS K. ARGYROS

Abstract. We use a Newton-like iteration to solve the nonlinear operator equation in a Banach space. The basic assumption is that the Fréchet-derivative of the nonlinear operator is Hölder continuous on some open ball centered at the initial guess. Under natural assumptions, we prove linear convergence of the iteration to a locally unique solution of the nonlinear equation.

Introduction. We introduce the Newton-like iteration

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

to solve the equation

$$F(x) = 0, \tag{2}$$

where F is a nonlinear operator on a Banach space X. Here we assume that the Fréchet derivative $F'(x)$ of F is Hölder continuous on some ball

$$U(x_0, r_0) = \{x \in X \mid \|x - x_0\| < r_0\}, \quad r_0 > 0.$$

The point $x_0 \in x$ and the arbitrary points y_n, $n = 0,1,2,\ldots$ are chosen sufficiently close to the desired solution. Then under natural assumptions we

Received May 19, 1989.
(1980) A.M.S. classification codes: 65J15, 65H10.
Key words and phrases: Newton-like iteration, Banach space, Fréchet-derivative.

show that (1) converges linearly to a unique solution x^* of (2) in $U(x_0, r)$ for some $r > 0$.

Note that for $y_n = x_n$ we obtain Newton's iteration, whereas for $y_n = x_0$ we obtain the modified one. The computer will determine the y_n's as to minimize the effort each time.

Here is an incomplete list of the usual assumptions for the convergence of (1) to a solution of (2):

(a) the existence and boundedness of the second Fréchet derivative of F (see, ex. [2], [3] and [4]).

(b) The assumption of analyticity which eliminates explicit mention of the second derivative [7].

(c) The case when the first Fréchet derivative of F satisfies a Lipschitz condition. (See, ex. [5], [6] and the references there.)

Finally,

(d) various other assumptions mainly based on the possibility of replacing $F'(x_n)^{-1}$ with a sequence of linear operators which are "close" in some sense to $F'(x_n)^{-1}$, $n = 0, 1, 2, \ldots$. (See, ex. [5], [6] and the references there.)

We will need the following definition.

Definition. We assume that F is once Fréchet-differentiable [2] and $F'(x)$ is the first Fréchet-derivative at a point $x \in X$. It is well known that $F'(x) \in L(x)$, the space of bounded linear operators from X to X. We say that the Fréchet-derivative $F'(x)$ is Hölder continuous over a domain $D \subset X$ if for some $c > 0$, $p \in [0, 1]$, and all $x, y \in D$,

$$||F'(x) - F'(y)|| \leq c||x - y||^P, \tag{3}$$

From now on we will find it more convenient to assume that $D = U(x_0, r_0)$ for some fixed $x_0 \in x$ and $r_0 > 0$.

Lemma. *Let $F'(x)$ be Hölder continuous on $U(x_0, r_0)$ for some x_0, r_0, p such that $x_0 \in x$, $r_0 > 0$ and $p \in [0, 1]$. Suppose that $F'(x_0)$ has a bounded inverse. For any $b_0 > ||F'(x_0)^{-1}||$, there is a number $r_3 \leq \min\{1, r_0\}$ such that:*

(a) If $x \in U(x_0, r_3)$, then the linear operator $F'(x)$ has bounded inverse and

$$\|F'(x)^{-1}\| < b_0; \tag{4}$$

(b) if $x_i \in U(x_0, r_3)$, $i = 1, 2, 3$ then

$$\|F(x_1) - F(x_2) - F'(x_3)(x_1 - x_2)\| \leq \frac{1}{2b_0}\|x_1 - x_2\|. \tag{5}$$

Proof. (a) If $x \in U(x_0, r_0)$, then

$$\|F'(x) - F'(x_0)\| \leq c\|x - x_0\|^p.$$

Choose $r_1 > 0$ such that

$$0 < r_1 \leq \min(r_0, (\frac{1}{4b_0 c})^{1/p}),$$

then if $x \in U(x_0, r_1)$

$$\|F'(x) - F'(x_0)\| \leq c\|x - x_0\|^p \leq c{r_1}^p \leq \frac{1}{4b_0}.$$

Since,

$$\|F'(x_0)^{-1}\| \cdot \|F'(x) - F'(x_0)\| < b_0 \cdot \frac{1}{4b_0} < 1,$$

by the Banach lemma $F'(x)$ has a bounded inverse for $x \in U(x_0, r_1)$. Therefore, there exists an $r_2 > 0$, with

$$0 < r_2 \leq r_1$$

such that if $x \in U(x_0, r_2)$, then $\|F'(x)^{-1}\| < b_0$.

That proves (4).

(b) If $r_3 = \min\{1, r_2\}$ and $x_i \in U(x_0, r_3)$, $i = 1, 2, 3$, we first have

$$\begin{aligned} &F(x_1) - F(x_2) - F'(x_3)(x_1 - x_2) \\ &= (F(x_1) - F(x_2) - F'(x_0)(x_1 - x_2)) + (F'(x_0) - F'(x_3))(x_1 - x_2) \\ &= \int_0^1 (F'[tx_1 + (1-t)x_2] - F'(x_0))(x_1 - x_2)dt + (F'(x_0) - F'(x_3))(x_1 - x_2). \end{aligned}$$

But,

$$\|\int_0^1 (F'[tx_1 + (1-t)x_2] - F'(x_0))(x_1 - x_2)dt\| \leq c{r_3}^p\|x_1 - x_2\|$$
$$\leq \frac{1}{4b_0}\|x_1 - x_2\| \qquad (6)$$

and

$$\|(F'(x_0) - F'(x_3))(x_1 - x_2)\| \leq c{r_3}^p\|x_1 - x_2\|$$
$$\leq \frac{1}{4b_0}\|x_1 - x_2\|.$$

Therefore,

$$\|F(x_1) - F(x_2) - F'(x_3)(x_1 - x_2)\| \leq \frac{1}{2b_0}\|x_1 - x_2\|.$$

That proves (5) and completes the proof of the lemma.

We now state and prove the main result.

Theorem. *Let $F'(x)$ be Hölder continuous on $U(x_0, r_3)$, where r_3 is defined in the lemma. Suppose that $x_0 \in X$ is such that $F'(x_0)$ has an inverse satisfying $\|F'(x_0)^{-1}\| < b_0 < \infty$ and*

$$\|F(x_0)\| < \frac{r_3}{2b_0}.$$

Let y_n be arbitrary points such that $y_n \in U(x_0, r_3)$, $n = 0, 1, 2, \ldots$.

Then the iteration $[x_n]$, given by

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

converges to a unique solution x^* of (2) in $U(x_0, r_3)$.

Moreover, the following estimate holds

$$\|x_n - x^*\| < 2^{-n}r_3, \quad n = 0, 1, 2, \ldots. \qquad (7)$$

Proof. Using (1) for $y_0 = x_0$ we obtain

$$\|x_1 - x_0\| = \|F'(x_0)^{-1}F(x_0)\| \leq \|F'(x_0)^{-1}\| \cdot \|F(x_0)\| \leq b_0\|F(x_0)\| < \frac{r_3}{2}.$$

By (5) and the identity

$$F(x_1) = F(x_1) - F(x_0) - F'(y_0)(x_1 - x_0)$$

we get

$$\|F(x_1)\| \leq \frac{1}{2b_0}\|x_1 - x_0\|.$$

Claim. Suppose that x_i, $i = 1, 2, \ldots, n$ have been chosen such that

$$\|x_i - x_0\| < r_3, \tag{8}$$

$$\|x_i - x_{i-1}\| \leq b_0\|F(x_{i-1})\|, \tag{9}$$

and

$$\|F(x_i)\| \leq \frac{1}{2b_0}\|x_i - x_{i-1}\|. \tag{10}$$

Then, (8), (9) and (10) hold for $i = 1, 2, 3, \ldots, n, \ldots$.

By (4), we have

$$\|x_{n+1} - x_n\| \leq b_0\|F(x_n)\| < \frac{1}{2}\|x_n - x_{n-1}\| \tag{11}$$

that proves (9). Also, by (11)

$$\begin{aligned}\|x_{n+1} - x_0\| &\leq \{\sum_{i=0}^{n} 2^{-i}\}\|x_1 - x_0\| \\ &< \{1 - 2^{-(n+1)}\}r_3 < r_3\end{aligned}$$

which proves (8).

Moreover, using (5) and

$$F(x_{n+1}) = F(x_{n+1}) - F(x_n) - F'(y_n)(x_{n+1} - x_n),$$

we obtain

$$\|F(x_{n+1})\| \leq \frac{1}{2b_0}\|x_{n+1} - x_n\|,$$

which proves (10). The claim is now proved.

Let n, q be two integers, then

$$\begin{aligned}||x_{n+p} - x_n|| &\le \sum_{j=1}^{q} ||x_{n+j} - x_{n+j-1}|| \\ &\le b_0||F(x_0)|| \cdot 2^{-n}\{\sum_{j=0}^{q-1} 2^{-j}\} < 2^{-n} r_3.\end{aligned}$$

Therefore, $\{x_n\}$, $n = 0, 1, 2, \ldots$ constitutes a Cauchy sequence in a Banach space and as such it converges to some $x^* \in X$. By (8) and (10) respectively, we get

$$||x^* - x_0|| \le r_3$$

and

$$F(x^*) = 0.$$

Finally, to show that x^* is the unique solution of (2) in $U(x_0, r_3)$ let us assume that x_1^* is another solution of (2) in $U(x_0, r_3)$, with $x^* \neq x_1^*$. Then

$$\begin{aligned}||x^* - x_1^*|| = ||F'(x_0)^{-1}F'(x_0)(x^* - x_1^*)|| &\le b_0||F'(x_0)(x^* - x_1^*)|| \\ &\le \frac{1}{2}||x^* - x_1^*||,\end{aligned}$$

which contradicts the assumption $x^* \neq x_1^*$.

Therefore, $x^* = x_1^*$. Letting $q \to \infty$ in (12), we obtain (7) and that completes the proof of the theorem.

If $y_n = x_0$, $n = 0, 1, 2, \ldots$, (1) reduces to the modified Newton's iteration which requires the evaluation of the same inverse $F'(x_0)^{-1}$ at each step of the iteration.

However, if $y_n \neq x_0$, (1) requires the evaluation of the inverse operators $F'(y_n)^{-1}$, $n = 0, 1, 2, \ldots$ at each step, which constitutes a difficult task in general.

A usual alternative is then to find a sequence of bounded linear operators L_n, $n = 0, 1, 2, \ldots$, such that

$$||L_n - F'(x_0)|| < \frac{1}{4b_0}$$

and

$$\|L_n^{-1}\| \;<\; b_0.$$

Following the proof of the above theorem, we can then easily show that the iteration

$$x_{n+1} \;=\; x_n - L_n^{-1}F(x_n), \quad n = 0,1,2,\ldots \tag{13}$$

converges to a solution x^* of (2).

One can refer to [5], [6] and the references there for an extensive analysis of iterations like (13).

Some of the results in [1] (especially Theorems 1 and 2) are similar to ours for $p = 1$ only. However, the results there cannot be applied here for $p \neq 1$.

The motivation for the introduction of an interation like (1) when $F'(x)$ is Hölder continuous on some open ball is due to the existence of problems like the one illustrated in the example.

Example. Consider the differential equation

$$x'' + x^{1+p} \;=\; 0, \quad p \in [0,1] \tag{14}$$

$$x(0) \;=\; x(1) \;=\; 0.$$

We divide the interval [0,1] into n subintervals and we set $h = \frac{1}{n}$. Let $\{v_k\}$ be the points of subdivision with

$$0 \;=\; v_0 \;<\; v_1 \;<\; \ldots \;<\; v_n \;=\; 1.$$

A standard approximation for the second derivative is given by

$$x_i'' \;=\; \frac{x_{i-1} - 2x_i + x_{i+1}}{h^2}, \; x_i \;=\; x(v_i), \quad i = 1,2,\ldots,n-1.$$

Take $x_0 = x_n = 0$ and define the operator $F : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ by

$$F(x) \;=\; H(x) + h^2\varphi(x) \tag{15}$$

$$H = \begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & -1 & \\ 0 & & \ddots & \ddots & \\ & & -1 & 2 & \end{bmatrix},$$

$$\varphi(x) = \begin{bmatrix} x_1^{1+p} \\ x_2^{1+p} \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1}^{1+p} \end{bmatrix},$$

and

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \end{bmatrix}.$$

Then

$$F'(x) = H + h^2(p+1) \begin{bmatrix} x_1^p & & & & 0 \\ & x_2^p & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & x_{n-1}^p \end{bmatrix}.$$

Newton's method cannot be applied to the equation

$$F(x) = 0 \tag{16}$$

We may not be able to evaluate the second Fréchet-derivative since it would involve the evaluation of quantities of the form x_i^{-p} and they may not exist.

Let $x \in \mathbb{R}^{n-1}$, $H \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ and define the norms of x and H by

$$\|x\| = \max_{1 \leq j \leq n-1} |x_j|$$

$$\|H\| = \max_{1 \leq j \leq n-1} \sum_{k=1}^{n-1} |h_{jk}|.$$

For all $x, z \in \mathbb{R}^{n-1}$ for which $|x_i| > 0$, $|z_i| > 0$, $i = 1, 2, \ldots, n-1$ we obtain, for $p = \frac{1}{2}$ say,

$$\begin{aligned} \|F'(x) - F'(z)\| &= \|diag\{(1 + \frac{1}{2})h^2(x_j^{1/2} - z_j^{1/2})\}\| \\ &= \frac{3}{2}h^2 \max_{1 \le j \le n-1} |x_j^{1/2} - z_j^{1/2}| \le \frac{3}{2}h^2[\max |x_j - z_j|]^{1/2} \\ &= \frac{3}{2}h^2 \|x - z\|^{1/2}. \end{aligned}$$

Therefore, under the assumptions of the theorem, iteration (1) will converge to the solution x^* of (16).

References

[1] L. M. Graves, "Some mapping theorems", *Duke Math. J.*, Vol 17 (1950), pp. 111-114.
[2] L. V. Kantorovich, "On Newton's method", *Math. Reviews*, Vol. 12 (1951), p. 419.
[3] L. V. Kantorovich and G. P. Akilov, "Functional analysis in normed spaces", *Oxford Publ.*, Pergamon Press, 1964.
[4] I. P. Mysovkih, "On the convergence of Newton's method", *Math. Reviews*, Vol. 12 (1951), p.419.
[5] L. B. Rall, "Nonlinear functional analysis and applications", *(Article by J. Dennis.)* Academic Press, 1971.
[6] W. C. Rheinboldt, "Numerical analysis of parametrized nonlinear equations", *John Wiley*, Publ. 1986.
[7] L. M. Stein, "Sufficient conditions for the convergence of Newton's method in complex Banach spaces", *Proc. Amer. Math. Soc.* Vol. 3 (1952), pp. 858-863.

Department of Mathematics, Cameron University, Lawton, OK, 73505, U. S. A.