REVUE D'ANALISE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION

Tome 23, N° 2, 1994, pp. 167–174

ON A NEWTON TYPE METHOD

IOAN LAZĂR

(Cluj-Napoca)

1. INTRODUCTION

In this paper we shall give a convergence theorem for a Newton type method for solving operator equations in Banach spaces. We shall also give a numerical example.

In order to find the roots of a single nonlinear equation $f(x) = 0$, Moser [2] has proposed the following iterations

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

$$y_{n+1} = y_n - y_n(f'(x_n)y_n - 1), \quad n = 0, 1, \ldots$$

The first iteration is similar to Newton's iteration, in which case y_n is equal to $1/f'(x_n)$. The second iteration is Newton's method applied to $g(y) = 1/y - f'(x_n)$. Thus, if y_n is close to $1/f'(x_n)$, then y_{n+1} is even closer. This scheme can also be interpreted in terms of an approximation to the inverse function of $f(x)$ (in fact the above formulae represent a later interpretation, see [1]). This method was developed as a tool for solving problems with small divisors, for which the application of Newton's method or successive approximation method are dubious. It can be shown that the order of convergence for the above scheme is $(1+\sqrt{5})/2 = 1,62\ldots$

An improved scheme was proposed by Ul'm [4] and Hald [1]. For the equation

$$F(x) = 0, \tag{1.1}$$

where $F : D \subset X \to Y$, X and Y are two Banach spaces and $D \subset X$ an open set, the authors have considered the iterations

$$x_{n+1} = x_n - A_n F(x_n), \tag{1.2}$$

$$A_{n+1} = A_n - A_n(F'(x_{n+1})A_n - I) = A_n(2I - F'(x_{n+1})A_n), \tag{1.3}$$

$n = 0, 1, \ldots$, with the initial guesses $x_0 \in D$ and $A_0 \in L(Y, X)$, where $L(Y, X)$ denotes the Banach space of the bounded linear operators from Y into X.

Iterations (1.2) and (1.3) keep the properties of Moser's iterations and the sequence (x_n) converges quadratically to a solution of (1.1) (see Zehnder [5] and Hald [1]).

Under Kantorovich type assumptions, similarly with those of Hald [1] in the real case, we shall give a convergence theorem for the method (1.2, 1.3). Our a priori error bounds are sharper than those from [1], [3] and [5]. We shall also obtain a posteriori error bounds.

Finally, we shall give a numerical example, related to a Hammerstein integral equation. We shall see that it is advantageous to use the scheme (1.2, 1.3) when we solve a large nonlinear system of equations, because the amount of arithmetic operations for this algorithm is less than in Newton's process.

2. CONVERGENCE THEOREM

Concerning equation (1.1) and the iterations (1.2) and (1.3) we have

THEOREM 2.1. *If F is Fréchet differentiable on $\bar{S}(x_0, r) = \{x \in X : \|x - x_0\| \leqslant r\} \subseteq D$ and for some $k > 0$, $\eta \geqslant 0$ and $q \geqslant 0$ the following conditions hold*

(2.1a) A_0 *is invertible and* $A_0^{-1} \in L(X, Y)$;

(2.1b) $$\|A_0 F(x_0)\| \leqslant \eta;$$

(2.1c) $$\|A_0(F'(x) - F'(y))\| \leqslant k\|x - y\|, \ \forall x, y \in \bar{S}(x_0, r);$$

(2.1d) $$\|I - A_0 F'(x_0)\| \leqslant q;$$

(2.1e) $$d := k\eta + q \leqslant \frac{1}{1 + \sqrt{2}} \ \textit{and} \ r \geqslant \eta + \frac{d^2}{k(1 - 4d^2)}$$

then the sequence (x_n) is well defined by (1.2) and (1.3), remains in $\bar{S}(x_0, r)$ and converges to a solution x^* of equation (1.1). This solution is unique in $D \cap S(x_0, r)$, if $r < (1 - q)/k$.

We also have the following a priori error estimates

(2.2) $$\|x_n - x^*\| \leqslant C_1 \frac{(2d)^{2^n}}{2^{n+1}}, \ \|A_n - F'(x^*)^{-1}\| \leqslant C_2 (2d)^{2^n}, \ n = 1, 2, \ldots$$

and the a posteriori error estimates

(2.3) $$\|x_n - x^*\| \leqslant (2d)^{2^{n-1}} \|x_n - x_{n-1}\|, \ n = 2, 3, \ldots$$

where $C_1 = \dfrac{1}{k(1 - 4d^2)}$ *and* $C_2 = \dfrac{3(1 + d)}{1 - 3d^2} \|A_0\|$.

Proof. For x_1 and A_1 given by (1.2) and (1.3) we shall find similar relations to assumptions (2.1).

From (1.3) it follows that

(2.4) $$A_1 w = (2I - A_0 F'(x_1)) A_0 w, \text{ for all } w \in Y,$$

(2.5) $$I - A_1 F'(x_1) = (I - A_0 F'(x_1))^2$$

and

(2.6) $$I - A_1 A_0^{-1} = A_0 F'(x_1) - I.$$

Since, by (2.1)

(2.7) $$\|I - A_0 F'(x_1)\| \leqslant \|I - A_0 F'(x_0)\| + \|A_0(F'(x_0) - F'(x_1))\| \leqslant k\eta + q = d,$$

we obtain, using (2.6), $\|I - A_1 A_0^{-1}\| \leqslant d < 1$, which implies that there exists $(A_1 A_0^{-1}) \in L(X)$ and $\|(A_1 A_0^{-1})^{-1}\| \leqslant 1/(1 - d)$. Thus

(2.8) A_1 is invertible, $A_1^{-1} \in L(X, Y)$ and $\|A_1^{-1}\| \leqslant \|A_0^{-1}\|/(1 - d)$.

From (2.4) and (2.7) we obtain

(2.9) $$\|A_1 F(x_1)\| \leqslant (1 + d)\|A_0 F(x_1)\|.$$

and, by (1.2)

$$A_0 F(x_1) = A_0[F(x_1) - F(x_0) - F'(x_0)(x_1 - x_0)] + A_0[F(x_0) + F'(x_0)(x_1 - x_0)] = A_0[F(x_1) - F(x_0) - F'(x_0)(x_1 - x_0)] + [I - A_0 F'(x_0)] A_0 F(x_0).$$

But, from the assumptions (2.1) it follows that the operator $A_0 F(\cdot)$ is differentiable on $\bar{S}(x_0, r)$, $A_0 F'(\cdot)$ is Lipschitz, so

$$\|A_0[F(x_1) - F(x_0) - F'(x_0)(x_1 - x_0)]\| \leqslant \frac{k}{2}\|x_1 - x_0\|^2,$$

thus, from the above inequality and (2.9)

(2.10) $$\|A_1 F(x_1)\| \leqslant (1 + d)\left[\frac{k}{2}\eta^2 + q\right] = \frac{1 + d}{2k}(d^2 - q^2).$$

We also obtain from (2.4), (2.7) and (2.1c)

(2.11) $$\|A_1(F'(x) - F'(y))\| \leqslant (1 + d)\|A_0(F'(x) - F'(y))\| \leqslant k(1 + d)\|x - y\|,$$

for all $x, y \in \bar{S}(x_0, r)$, and, by (2.5)

(2.12) $$\|I - A_1 F'(x_1)\| \leqslant \|I - A_0 F'(x_1)\|^2 \leqslant d^2.$$

Since (2.8) and (2.10–12) are similar to assumptions (2.1) we can prove by induction that iterations (1.2) and (1.3) are well defined, that is, the sequence (x_n) lies in $\bar{S}(x_0, r)$.

First, let us consider the sequences (k_n), (η_n), (q_n) and (d_n) defined as follows

(2.13) $$\begin{aligned} k_0 &= k, \ \eta_0 = \eta, \ q_0 = q, \ d_0 = d, \\ d_n &= k_n \eta_n + q_n, \\ k_n &= k_{n-1}(1 + d_{n-1}), \\ q_n &= d_{n-1}^2, \\ \eta_n &= \frac{1 + d_{n-1}}{2k_{n-1}}(d_{n-1}^2 - q_{n-1}^2), \ n = 1, 2, \ldots \end{aligned}$$

For (d_n) we obatin the following recurrent relation

(2.14) $$d_n = \frac{1}{2}(1+d_{n-1})^2(d_{n-1}-q_{n-1}^2)+d_{n-1}^2,$$

and, using (2.1e) we can prove by induction that

(2.15) $$q_n \leqslant d_n \leqslant d_{n-1} \leqslant \frac{1}{1+\sqrt{2}} \text{ and } d_n \leqslant 2d_{n-1}^2, \text{ for all } n \geqslant 1,$$

whence it follows that

(2.16) $$d_n \leqslant \frac{1}{2}(2d)^{2^n},$$

and

(2.17) $$\eta_n = \frac{d_n-q_n}{k_n} \leqslant \frac{d_n-q_n}{k} \leqslant \frac{d_n-d_{n-1}^2}{k} = \frac{d_n-d_{n-1}^2}{k} \leqslant \frac{1}{k}d_{n-1}^2 \leqslant \frac{1}{4k}(2d)^{2^n},$$

for all $n \geqslant 1$.

So, it can be shown by induction that for all $n \geqslant 1$

(2.18a) $x_n \in \bar{S}(x_0, r)$;

(2.18b) there exists $A_n^{-1} \in L(X, Y)$ and $\|A_n^{-1}\| \leqslant \|A_{n-1}^{-1}\|/(1-d_{n-1})$;

(2.18c) $\|A_nF(x_n)\| \leqslant \eta_n$;

(2.18d) $\|A_n(F'(x)-F'(y))\| \leqslant k_n\|x-y\|$, for all $x, y \in \bar{S}(x_0, r)$;

(2.18e) $\|I - A_nF'(x_n)\| \leqslant q_n$.

For $n = 1$, (2.18) hold. We suppose that (2.18) hold for $i = 1, \ldots, n$. Since, for all $j, m \geqslant 1$

(2.19) $$\sum_{i=j}^{j+m-1}\eta_i \leqslant \frac{1}{4k}\sum_{i=j}^{j+m-1}(2d)^{2^i} \leqslant \frac{1}{4k}\sum_{i=0}^{m-1}(2d)^{2^j(1+i)} = \frac{(2d)^{2^j}}{4k}\cdot\frac{1-(2d)^{m\cdot 2^j}}{1-(2d)^{2^j}}$$

it follows that

$$\|x_{n+1}-x_0\| \leqslant \eta + \sum_{i=1}^{n}\eta_i \leqslant \eta + \frac{d^2}{k(1-4d^2)} \leqslant r,$$

hence $x_{n+1} \in \bar{S}(x_0, r)$, that is (2.18a) for $i = n+1$. One can prove (2.18b–e) for $i = n+1$, as we have proved (2.8) and (2.10–2.12).

Now, by (2.18c) and (2.19)

$$\|x_{m+n}-x_n\| \leqslant \sum_{i=n}^{n+m-1}\eta_i \leqslant \frac{(2d)^{2^n}}{4k}\cdot\frac{1-(2d)^{2^n\cdot m}}{1-(2d)^{2^n}}, \text{ for all } n, m \geqslant 1,$$ therefore (x_n) is a Cauchy sequence, so it is convergent.

Let $x^* = \lim_{n\to\infty} x_n$. From the above inequality it follows that

(2.20) $$\|x_n - x^*\| \leqslant \frac{1}{4k}\cdot\frac{(2d)^{2^n}}{1-(2d)^{2^n}}, \text{ for all } n \geqslant 1,$$

and $\|x^*-x_0\| \leqslant \|x_1-x_0\| + \|x_1-x^*\| \leqslant \eta + \frac{d^2}{k(1-4d^2)} \leqslant r$, that is $x^* \in \bar{S}(x_0, r)$.

Further, we prove that x^* is a solution of (1.1). Since (1.2) can be written

$$A_n^{-1}(x_{n+1}-x_n) = -F(x_n),$$

it is sufficient to show that the sequence $(\|A_n^{-1}\|)$ is bounded.

First, by (2.18)

(2.21) $$\|I-A_nF'(x^*)\| \leqslant \|I-A_nF'(x_n)\| + \|A_n(F'(x_n)-F'(x^*))\| \leqslant q_n + k_n\|x_n-x^*\|, \text{ for all } n \geqslant 1,$$

and, by (2.13) and (2.15)

$$\eta_{n+1}/\eta_n = \frac{1}{2}(1+d_{n-1})(d_n+d_{n-1}^2) < \frac{1}{2},$$

whence, $\eta_{n+m} < \frac{1}{2^m}\eta_n$, for all $n, m \geqslant 1$,

which implies

(2.22) $$\|x_n-x^*\| \leqslant \sum_{m=0}^{\infty}\eta_{n+m} \leqslant \eta_n\sum_{m=0}^{\infty}\frac{1}{2^m} = 2\eta_n, \text{ for all } n \geqslant 1.$$

So, the estimates (2.21) become

(2.23) $$\|I-A_nF'(x^*)\| \leqslant q_n + 2k_n\eta_n = 2d_n - q_n \leqslant 3d_{n-1}^2 \leqslant \frac{3}{4}(2d)^{2^n}.$$

By (2.18b) and (2.15) we have $\|A_n^{-1}\| \leqslant \|A_0^{-1}\|/(1-d)^n$ and using (2.23)

$$\|A_n^{-1} - F'(x^*)\| \leqslant \|A_n^{-1}\|\cdot\|I-A_nF'(x^*)\| \leqslant \frac{3}{4}\|A_0^{-1}\|\frac{(2d)^{2^n}}{(1-d)^n},$$

whence $\|A_n^{-1}-F'(x^*)\| \to 0$, when $n \to \infty$, so $(\|A_n^{-1}\|)$ is bounded.

For the estimates (2.2), let $e_n := \|x_n - x^*\|$, $n \geqslant 0$. Since

$$x_{n+1}-x^* = x_n - A_nF(x_n) - x^* =$$
$$= [I-A_nF'(x_n)](x_n-x^*) + A_n[F(x^*) - F(x_n) - F'(x_n)(x^*-x_n)],$$

by (2.18) we get

$$e_{n+1} \leqslant \left(\frac{k_n e_n}{2} + q_n\right) e_n,$$

and, by (2.22)

(2.24) $e_{n+1} \leqslant (k_n \eta_n + q_n) e_n = d_n e_n$, for all $n \geqslant 1$, thus, using (2.16)

$$e_n \leqslant \left(\prod_{i=1}^{n-1} d_i\right) e_1 \leqslant \frac{1}{2^{n-1}} (2d)^{\sum_{i=1}^{n-1} 2^i} e_1 = \frac{(2d)^{2^n-2}}{2^{n-1}} e_1,$$

and estimating e_1 from (2.20) there follows the first part of (2.2). For the second part, by (2.23) for $n = 1$ there follows $\| I - A_1 F'(x^*) \| \leqslant 3d^2 < 1$, hence there exists $[A_1 F'(x^*)]^{-1} \in L(X)$ and $\| A_1 F'(x^*)^{-1} \| \leqslant 1/(1-3d^2)$, therefore using (2.4) and (2.7)

$$\| F'(x^*)^{-1} \| \leqslant 1/(1-3d^2) \| A_1 \| \leqslant \frac{1+d}{1-3d^2} \| A_0 \|.$$

Finally, because $A_n - F'(x^*)^{-1} = -[I - A_n F'(x^*)] F'(x^*)^{-1}$, using (2.23) we get $\| A_n - F'(x^*)^{-1} \| \leqslant C_2 (2d)^{2^n}$.

The a posteriori error bound (2.3) follows from

$$e_n \leqslant d_{n-1}\, e_{n-1},\ e_{n-1} \leqslant 2\eta_{n-1} \text{ and } (2.16).$$

The uniqueness of the solution x^*, if $r < (1-q)/k$, follows from the fact that the operator $P(x) = x - A_0 F(x)$ is contractive on $\bar{S}(x_0, r)$. Indeed, P is differentiable, $\| P'(x) \| = \| I - A_0 F'(x) \| \leqslant \| I - A_0 F'(x_0) \| + \| A_0 (F'(x_0) - F'(x)) \| \leqslant kr + q < 1$, if $x \in \bar{S}(x_0, r)$.

3. NUMERICAL EXAMPLE

Let us consider the Hammerstein integral equation

(3.1) $$x(s) - \int_0^1 s t^2 (x(t))^2 \, dt = \frac{9}{20} s,\ s \in [0, 1],$$

in $C\,[0, 1]$, and let $H(s, t, u) = stu^2$, $b(s) = 9/20\ s$.

Using the repeated trapezoidal rule with $h_N = 1/N$, $s_i = i/N$, $i = 0, 1, \ldots, N$, $w_i = h_N$, $i = 1, \ldots, N-1$ and $w_0 = w_N = h_N/2$, we approximate the exact solution $x(s) = s/2$ of (3.1) by the solution of the equation

(3.2) $$\bar{x}(s) - \sum_{j=0}^{N} H(s, s_j, \bar{x}(s_j)) w_j = b(s),\ s \in [0, 1].$$

We are not concerned with the existence of the solutions of (3.1) and (3.2) or with the approximation error but we shall, rather estimate the

amount of arithmetic operations for the method (1.2,3) applied to solve (3.2).

Writing $s = s_i$ in (3.2) we obtain the nonlinear system

(3.3) $$\bar{x}_i - \sum_{j=0}^{N} H(s_i, s_j, \bar{x}_i) w_j = b(s_i),\ i = 0, 1, \ldots, N.$$

Let $\bar{x} = (\bar{x}_i)_{i=0}^N$, $\bar{b} = b(s_i)_{i=0}^N$. So the system (3.3) can be written

(3.4) $$F(\bar{x}) = (I - \bar{H})\bar{x} - \bar{b} = 0, \text{ where } F : R^{N+1} \to R^{N+1}.$$

For $\bar{x} \in R^{N+1}$ we set $\bar{x} = \max_{c \leqslant i \leqslant N} |\bar{x}_i|$. We see that F is differentiable on R^{N+1}, $F'(\bar{x}) = I - \bar{H}'(\bar{x})$ and $\| F'(y) - F'(z) \| = \| \bar{H}'(y) - \bar{H}'(z) \| = 2 \max_{0 \leqslant i \leqslant N} \sum_{j=0}^{N} s_i s_j^2 w_j | y_j - z_j | \leqslant (2N^2 + 1)/(3N^2) \| y - z \|$.

In order to get a solution of (3.4) we shall use the iterations (1.2) and (1.3) with $\bar{x}^{(0)} = \left(\frac{1}{4} s_i\right)_{i=0}^N$ and $A_0 = I$ (the identity matrix). In the method (1.2) and (1.3) the approximate inverse of F' is available and this facilitates the estimates of the errors (2.2) and (2.3). But this advantage disappears because of the need of matrix multiplications in (1.3).

As in [1], in the numerical computations below we use the following version of (1.2) and (1.3)

(3.5) $$\bar{x}^{(k+1)} = \bar{x}^{(k)} - A_k F(\bar{x}^{(k)}),$$

(3.6) $$A_i u = A_{i-1} (2I - F'(\bar{x}^{(i)}) A_{i-1})\, u,\ i = 1, 2, \ldots, k.$$

Thus to compute $A_k F(\bar{x}^{(k)})$, instead of computing the matrix A_k at each step k, we can use (3.6) recursively. So we must save the Fréchet derivative from all the previous steps.

To estimate the amount of arithmetic operations for this algorithm, let c_i and c be the number of multiplicative operations needed to calculate $v = A_i u$ and $F'(\bar{x}^{(i)}) v$, respectively. It follows from (3.6) that $c_i = 2c_{i-1} + c$, and since $c_0 = 0$ we see that $c_k = (2^k - 1)c$. If the cost of computing $F(\bar{x}^{(i)})$ is less than or equal to c, then the total amount of work for the computation of $\bar{x}^{(i)}, \ldots, \bar{x}^{(k)}$ is less than $2^{k+1} c$.

We observe that the cost per step of the method (3.5) and (3.6) grows exponentially, but it is important that the method converges quadratically. So, if c is small compared to the number of operations needed to solve the linear system involved in Newton's method at each step (especially for large systems), then we can make several steps of (1.2) and (1.3) in the time needed for one step of Newton's method (for a comparison of these methods, see [1]).

The table below contains the results of application of (3.5) and (3.6) after four iterations: the initial estimates for q_0, d_0; the a priori estimates (2.2) for $\| \bar{x}^{(4)} - x^* \|$, the a posteriori estimates (2.3) for $\| \bar{x}^{(4)} - x^* \|$, and $\| \bar{x}^{(4)} - \tilde{x} \|$, where $\tilde{x} = \left(\frac{1}{2} s_i\right)_{i=0}^N$.

N	q_0	d_0	a priori estimates	a posteriori estimates	$\|\bar{x}^{(4)} - \widetilde{x}\|$
4	$1.32 \cdot 10^{-1}$	$2.96 \cdot 10^{-1}$	$2.10 \cdot 10^{-3}$	$1.03 \cdot 10^{-5}$	$6.62 \cdot 10^{-3}$
16	$1.25 \cdot 10^{-1}$	$2.81 \cdot 10^{-1}$	$1.32 \cdot 10^{-3}$	$3.31 \cdot 10^{-6}$	$4.03 \cdot 10^{-4}$
64	$1.24 \cdot 10^{-1}$	$2.78 \cdot 10^{-1}$	$1.29 \cdot 10^{-3}$	$3.12 \cdot 10^{-6}$	$2.51 \cdot 10^{-5}$

REFERENCES

1. Hald, O. H. *On a Newton Type Method.* Numer. Math. **23**, 411–426 (1975)
2. Moser, J. *Stable and Random Motions in Dynamical Systems with Special Emphasis on Celestial Mechanics.* Herman Wely Lectures, Annals of Mathematics Studies, no. 77. Princeton, New Jersey : Princeton University Press 1973.
3. Păvăloiu, I. Diaconu, A. *Sur quelque méthodes itératives pour la résolution des équations opérationnelles.* Revue d'analyse numérique et de la Théorie de l'Approximation, Tome 1, pp. 45–61 (1973).
4. Ul'm, S. *On the Iterative Methods with Simultaneous Approximation of the Inverse of the Operator.* Izv. Acad. Nauk Estonskoi S.S.R., **16**, 4, 403–411 (1967).
5. Zehnder, E. J. *A Remark about Newton's Method.* Comm. Pure Appl. Math., **37**, 361–366 (1974).

Received 15 XII 1993

Institute of Calculus
Republicii 37
PO Box 68
3400 Cluj-Napoca
România