Commentationes Mathematicae Universitatis Carolinae

Hana Petzeltová

Remark on a Newton-Moser type method

Commentationes Mathematicae Universitatis Carolinae, Vol. 21 (1980), No. 4, 719--725

Persistent URL: http://dml.cz/dmlcz/106037

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

21,4 (1980)

REMARK ON A NEWTON-MOSER TYPE METHOD

Hana PETZELTOVÁ

Abstract: The Newton-Moser method for finding the roots of a nonlinear equation $f(x) = 0$ is discussed and the rate of convergence of this method is established.

Key words: Nonlinear equation, rate of convergence, Newton s method.

Classification: 4600 , 65J05

Introduction. In the present paper we will discuss an iterative technique, due to J. Moser, improved by O. Hald [1], for finding the roots of a nonlinear equation $f(x) = 0$. Consider the following method:

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

(2) $$y_{n+1} = y_n - y_n(f'(x_{n+1})y_n - 1)$$

The first equation is similar to the Newton's method, in which case y_n is equal $f'(x_n)^{-1}$. The second equation is the Newton's method applied to $g(y) = y^{-1} - f'(x_{n+1}) = 0$. This method was developed as a technical tool for investigating problems involving small divisors, where the application of the Newton's method is dubious since it is not clear whether $f'(x_n)$ is invertible.

In a series of papers, V. Pták has proposed a new method of estimating the convergence of iterative processes. Instead of defining the rate of convergence as a number, he introduces the following

Definition. Let T be an interval $(0, t_0)$ for some positive t_0. A rate of convergence on T is a function ω defined on T with the following properties

1. ω maps T into itself
2. for each $t \in T$ the series $t + \omega(t) + \omega^2(t) + \dots$ is convergent.

We use the abbreviation ω^n for the n-th iterate of the function ω, so that $\omega^2(t) = \omega(\omega(t))$ and so on. The sum of the above series will be denoted by σ. It is clear that if some sequence $\{x_n\}_{n=0}^{\infty} \subset X$, X being a Banach space, satisfies the implication $|x_n - x_{n-1}| \leq r \Rightarrow |x_{n+1} - x_n| \leq \omega(r)$ for each n and ω is a rate of convergence, then the sequence $\{x_n\}$ converges to some element $x \in X$ and $|x_0 - x| \leq \sigma(r_0)$ where $r_0 = |x_0 - x_1|$.

The Newton-Moser process. Let E and F be two Banach spaces. Let $x_0 \in E$ and $U = U(x_0, q) = \{x \in E, |x - x_0| < q\}$. Let f be a mapping of U into F twice Fréchet differentiable for each $x \in U$ and let

(3) $|f''(x)| \leq M$ for all $x \in U$

We shall use the following notations:

$a(x,x') = f(x') - f(x) - f'(x)(x' - x)$

$b(x,x') = f'(x) - f'(x)$

As a consequence of (3) we get the estimates

(4) $|a(x,x')| \leq \frac{1}{2} M|x-x'|^2$ for $x,x' \in U$
$|b(x,x')| \leq M|x-x'|$

Take $x \in U$, $y \in B(F,E)$ such that

(5) $|yf(x)| \leq r$, $|yf'(x)-I| \leq s$, $|y| \leq t$ and let

(6) $x' = x - yf(x)$

$y' = y-y(f'(x')y-I)$

Let us estimate the norms $|y'f(x')|$, $|y'f'(x')-I|$, $|y'|$ $\cdot$ $f(x') =$
$= f(x) + f'(x)(x'-x) + a(x,x') = f(x) - f'(x)yf(x) + a(x,x')$

$y' = y(f'(x)y-I) - yb(x,x')y$

$|y'f(x')| = |(I-yf'(x))yf(x) + ya(x,x') + (yf'(x)-I)^2 yf(x) -$
$- (yf'(x)-I)ya(x,x') + yb(x,x')(yf'(x)-I)yf(x) - yb(x,x')ya(x,x')| \leq$
$\leq rs + \frac{1}{2} Mr^2t + rs^2 + \frac{1}{2} Mr^2st + Mr^2st + \frac{1}{2} M^2r^3t^2 = r(s + \frac{1}{2} Mrt)\cdot$
$\cdot(1+s+Mrt) = r' = \omega(r)$

$|y'f'(x')-I| = |yf'(x')-y(f'(x')y-I)f'(x')-I| = |(yf'(x')-I)^2| =$
$= |(yf'(x)-I + yb(x,x'))^2| \leq (s+Mrt)^2 = s'$

$|y'| \leq t(1+s+Mrt) = t'$

It is easy to verify that i

(7) $tr = \frac{(1-s)^2}{2M} - t^2d$ for some $d \geqq 0$, i.e.

$$t = \begin{cases} \frac{1}{2d}(-r + (r^2+2M^{-1}d(1-s)^2)^{1/2}) & \text{for } d \geqq 0 \\ \frac{(1-s)^2}{2Mr} & \text{for } d=0 \end{cases}$$

then the same is true for r', s', t'.

For such t we get the inequality

$Mrt \leq \frac{(1-s)^2}{2}$ which yields

$\omega(r) \leq r \left(\frac{1+s}{2}\right)^2 \quad \frac{3+s^2}{2}$

$s' \leq \left(\frac{1+s^2}{2}\right)^2$

If $s_0 = |y_0 f(x_0)-I| < 1$ and $s^* < 1$ satisfies $s^* = \left(\frac{1+s^{*2}}{2}\right)^2$

then the last inequality shows that during all the process $s \leq \hat{s} = \max(s_0, s^*)$. Now, it suffices to choose s_0 such that $(\frac{1+s_0}{2})^2 \ \frac{3+s_0}{2} < 1$ for ω to be a rate of convergence. (For s^* this inequality is satisfied.)

If x,y satisfy (5), then

$$(8) \quad |f(x)| \leq |y^{-1}| r \leq \frac{|f'(x)|}{1-s} r \leq \frac{f'(x_0) + 2Mq}{1 - \hat{s}} r = Cr$$

The corresponding function σ can be obtained with the help of the function $g: R \longrightarrow R$, $g(z) = \frac{1}{2}Mz^2 - d$, $d \geqq 0$ for which all the above inequalities become equalities, in the following way: If r_0, s_0 are given, we find y_0, $z_0 > (\frac{2d}{M})^{1/2}$ such that $y_0 g(z_0) = r_0$, $y_0 g'(z_0) - 1 = s_0$. Then $\sigma(r_0) = z_0 - (\frac{2d}{M})^{1/2}$.
We get the quadratic equation for z_0:

$$\frac{1+s_0}{Mz_0}(\frac{1}{2} Mz_0^2 - d) = r_0$$

Hence $z_0 = (1+s_0)^{-1} (r_0 + (r_0^2 + 2M^{-1}d(1+s_0)^2)^{1/2})$ and

$$\sigma(r_0) = (1+s_0)^{-1} (r_0 + (r_0^2 + 2M^{-1}d(1+s_0)^2)^{1/2}) - (\frac{2d}{M})^{1/2}$$

The relation (7) was derived with help of the function g as well.

<u>Theorem</u>. Let E and F be two Banach spaces. Let $x_0 \in E$. Let f be a mapping of $U(x_0,q)$ into F satisfying (3) and the following assumptions:

1^o there exists a linear invertible operator $y_0 \in B(F,E)$, non-negative numbers r_0, s_0, d such that

$$|y_0 f(x_0)| = r_0, \quad |y_0 g'(x_0) - I| = s_0$$

$$2^o \quad |f(x_0)| \leq \frac{1}{2M} (\frac{1-s_0}{|y_0|})^2 - d$$

3^o $q \geq (1+s_o)^{-1} (r_o+(r_o^2 + 2M^{-1} d(1+s_o)^2)^{1/2}) -$

$- (2M^{-1} d)^{1/2} = \sigma(r_o)$

4^o $\frac{3 + s_o^2}{2} (\frac{1+s_o}{2})^2 < 1$

Then the process (1),(2) starting at x_o is meaningful and converges to a point x such that $f(x) = 0$. The rate of convergence $\omega(r) = r(s(r) + \frac{1}{2} M r t(r)) (1+s(r) + M r t(r))$, $r \in (0,r_o\rangle$ where $s(r_o) = s_o$ and $s(\omega(r)) = (s(r) + M r t(r))^2$

$$(9) \quad t(r) = \begin{cases} \frac{1}{2d}(-r+(r^2 + 2M^{-1} d(1-s(r))^2)^{1/2}) & \text{for } d>0 \\ \frac{(1-s(r))^2}{2Mr} & \text{for } d = 0 \end{cases}$$

yields the following estimates

$$|x_{n+1} - x_n| \leq \omega^n(r_o), \quad |x - x_o| < \sigma(r_o)$$

Proof. According to what has been said above, it suffices to prove that $|y_o| \leq t(r_o)$. From 2^o we get $|y_o| r_o \leq |y_o|^2 |f(x_o)| \leq$ $\leq \frac{(1-s_o)^2}{2M} - d|y_o|^2$ which is equivalent to $|y_o| \leq t(r_o)$.

<u>Corollary</u>. In the case that $d>0$, the sequence $\{y_n\}$ is bounded, $f'(x)$ is invertible and $y_n \to f'(x)^{-1}$. The rate of convergence is almost quadratic, more precisely $\omega(r) \leq c\, r^{1+\alpha}$ for r sufficiently small and $\alpha < 1$.

Proof. It follows immediately from (9) that $|y_n| \leq$ $\leq t(\omega^n(r_o)) \leq (2M\, d)^{-1/2} = K$, $s_n = s(\omega^n(r_o)) \to 0$. Let y, y' satisfy (5),(6). Then

$$|y' - y| \leq |(yf'(x) - I)y| + |yb(x,x')y| \leq K(s(r) + MKr) =$$
$$= Kh(r)$$

$$(10) \quad h(\omega(r)) \leq (s(r) + MKr)^2 + MKr(s(r) + \frac{1}{2} MKr)\cdot(1+s(r)+ MKr) \leq$$
$$\leq h(r)^2 (1 + \frac{1}{2} (1+s(r) + MKr))$$

This assures the convergence of the sequence $\{y_n\}$. Since $y_n \to y$ and $s_n \to 0$, we get $y = f'(x)^{-1}$.

Let $\alpha \in (0,1)$ be given, r, s, t satisfy (5) and $s^{1-\alpha}(1 + \frac{Mrt}{s})^2 < 1$ (This can be achieved, as $s(\omega^n(r)) \to 0$ and $s(\omega^n(r)) \geq \text{const.}\,\omega^n(r)$ for $n \geq 1$). Let $s = kr^{\alpha}$. Then $\omega(r) + r(s + \frac{1}{2} Mrt)(1+s+Mrt) = rs\,(1 + \frac{1}{2} M \frac{rt}{s})(1+s+Mrt) \leq$ $\leq Cr^{1+\alpha}$, $s' = (s+Mrt)^2 = s^2(1+ \frac{Mrt}{s})^2 = k\omega(r)^{\alpha} \cdot s^{1-\alpha}(1+\frac{Mrt}{s})^2 \cdot$ $\cdot (\frac{\omega(r)}{rs})^{-\alpha} \leq k\,\omega(r)^{\alpha}$.

Remark. In the case that $f'(x)$ is invertible in some neighbourhood U of the solution and $d\,f'(x) = \inf\{|f'(x)|, |x| \geq 1\} \geq m$ for $x \in U$, we get $|y| \leq \frac{1+s(r)}{m}$ if $|yf'(x) - I| \leq$ $\leq s(r)$, $x \in U$. Moreover, if $s(r) + \frac{3M}{2m} r \leq \frac{1}{2}$, then (10) gives $t(\omega(r)) + \frac{3M}{2m}\omega(r) \leq \frac{7}{4}(t(r) + \frac{3M}{2m} r)^2$, see [1].

<u>Corollary</u>. Let d = 0 in 2^0 for all x. Then the sequence y_n does not converge, but even in this case (8) assures that $f(x_n) \to 0$. The rate of convergence becomes linear, $\omega(r) \leq$ $\leq r\,\frac{3+s(r)^2}{2}(\frac{1+s(r)}{2})^2$ and $s(r) \to s^*$, s^* being the solution of the equation $(\frac{1+s^2}{2})^2 = s$.

References

[1] O. HALD: On a Newton-Moser type method, Num. Math. 23, 411-426(1975).

[2] J. MOSER: Stable and Random Motions in Dynamical Systems, Princeton University Press 1973.

[3] V. PTÁK: Nondiscrete mathematical induction and iterative existence proofs, Lin. Algebra and its Appl. 13, 223-238(1976).

[4] V. PTÁK: The rate of convergence of Newton s process,

Numer. Math. 25, 279-285(1976).

Matematický ústav ČSAV
Žitná 25, 11567 Praha 1
Československo

(Oblatum 6.6. 1980)