REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Tome 24, N^os 1–2, 1995, pp. 15–21

REMARKS ON THE CONVERGENCE OF THE NEWTON-RAPHSON METHOD

VASILE BERINDE
(Baia Mare)

1. INTRODUCTION

In this note we are interesed in completing some results concerning the convergence of the Newton-Raphson method for solving a scalar equation f(x) = 0, when conditions involving only f and f' are required ([1] -[2]).

A classical result on the Newton-Raphson (or Newton's) method is given by the following theorem (see, for example, [6], pp. 128).

THEOREM 1. *Let* $f: [a,b] \to \mathbf{R}$, $a<b$, *be a function such that the following conditions are satisfied*

(f_1) $f(a)\cdot f(b)<0$;

(f_2) $f \in C^2[a,b]$ and $f'(x)\cdot f''(x)\neq 0, x\in[a,b]$.

Then the sequence (x_n) *defined by*

(1) $$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)},\ n\geq 0$$

converges to α, *the unique solution of* $f(x) = 0$ *in* $[a, b]$, *for each* $x_o \in [a, b]$ *satisfying* $f(x_0) f''(x_o) > 0$ *and the following estimation*

$$|x_n-\alpha|\leq\frac{M_2}{2m_1}|x_n-x_{n-1}|^2,\ n\geq 1$$

holds, where

$$m_1=\min_{x\in[a,b]}|f'(x)| \text{ and } M_2=\max_{x\in[a,b]}|f''(x)|.$$

Theorem 1 is very convenient for practical purposes but condition (f_2) is too strong, as shown by

Example 1. [5] For $f(x)=\tan x, x\in[a,b]\subset\left(\frac{\pi}{2},\frac{3\pi}{2}\right)$, with $\pi\in[a,b]$, we have $f\in C^2[a,b], f'(x)>0$, but $f''(\pi)=0$ and $\alpha=\pi$ is the unique solution of the equation $f(x)=0$ in $[a,b]$. Theorem 1 does not apply in this case, but the Newton's iteration (1) converges for each $x_0\in[a,b]$, in view of Theorem 2 below. For example, if we take $a=\frac{7\pi}{12}$, $b=\frac{17\pi}{12}$ and $x_0=a$, we obtain the value of π with 14 exact digits after 7 iterations: x_0=1.83; x_1=2.08; x_2=2.51; x_3=2.99; x_4=3.139; x_5=3.141592644; x_6=3.1415926535897944 and x_7=3.141592653589793.

Taking $x_0=\frac{2\pi}{3}$ we obtain the solution after 5 iterations x_0=2.09; x_1=2.527; x_2=2.998; x_3=3.139; x_4=3.141592648 and x_5=3.141592653589793.

So, the following question arises: can we obtain the convergence of (1) under weaker conditions on f? A positive answer - itself a classical result - is given the following theorem due to Ostrowski [9], see [8], pp. 316–318.

THEOREM 2. ([9], Theorem 7.2, pp. 60) *Let $f(x)$ be a real function of the real variable x, $f(x_0)f'(x_0)\neq 0$, and put $h_0=-f(x_0)/f'(x_0)$, $x_1=x_0+h_0$. Consider the interval $I_0=[x_0, x_0+2h_0]$ and assume that $f''(x)$ exists in I_0, that* $\max_{x\in I_0}|f''(x)|=M_2$ and

$$2|h_0|M_2\leq|f'(x_0)|.$$

Then for the sequence (x_n) given by (1) *we have that x_n lie in I_0 and*

$$x_n\to\alpha(n\to\infty),$$

where α is the only zero of f in I_0.

Remark. The assumptions in Theorem 2 are still too strong as shown by

Example 2. [5] Let $f:[-1,1]\to\mathbf{R}$, be given by $f(x)=-x^2+2x$, if $x\in[-1,0)$, and $f(x)=x^2+2x$, if $x\in[0,1]$. The unique solution in $[-1,1]$ of the equation $f(x)=0$ is $\alpha=0$. Since f'' does not exist in $0\in I_0$, Theorem 2 does not apply. However, as will be seen in the next section, the sequence (1) is convergent. For example, if we start with 0.5, we obtain $x_1=0.833333$; $x_2=0.0032051$; $x_3=0.0000129$; $x_4=0.0000001$ and $x_5=0$.

In view of these examples and remarks, in [1] – [5] we have established weaker hypotheses on f which provide the existence and the uniqueness of the solution, the convergence of Newton's method, as well as the errors estimates.

The aim of the present paper is to improve the assumptions and the conclusions from [1] – [5].

2. THE EXTENDED NEWTON'S METHOD

Let $f:[a,b]\to\mathbf{R}$ be a function satisfying (f_1) and

(f'_2) $f\in C^1[a,b]$, $f'(x)\neq 0$, $x\in[a,b]$.

Then f has a unique solution $\alpha\in(a,b)$. Let $F:[a,b]\to\mathbf{R}$ be given by

(2) $$F(x)=-\frac{f(x)}{f'(x)},\quad x\in[a,b].$$

Then the equation

$$f(x)=0$$

is equivalent to the fixed point problem

$$x=F(x).$$

If f satisfies the conditions in Theorem 1, then for $x\in[a,b]$ satisfying $f(x)f''(x)>0$ we have

$$F(x)\in[a,b],$$

and the sequence of succesive approximations, $(F^n(x))$ converges to α as $n\to\infty$.

Let us now assume that f satisfies (f_1), (f'_2) and that Newton's iteration

(3) $$x_{n+1}=F(x_n)$$

converges to the unique solution of the equation $f(x)=0$, for each $x_0\in[a,b]$. This means

1) $F:[a,b]\to\mathbf{R}$ has a unique fixed point α:

2) For each $x_0\in[a,b]$, the sequence of successive approximations converges to α,

$$F^n(x_0)\to\alpha,\ \text{as } n\to\infty.$$

These two properties suggest to consider Bessaga's theorem, one of the crucial results in the fixed point theory (see, for example [10]) .

THEOREM 3. *Let X be a nonempty set and $F:X\to X$ be a mapping such that*

$$F_{F^k}=\{x^*\},\ \forall k\in\mathbf{N}^*\left(\text{where } F_g=\{x\in X\,/\,g(x)=x\}\right).$$

Let $\gamma\in(0,1)$ be a constant. Then there exists a metric d on X such that

a) *(X,d) is a complete metric space;*

b) *F is a contraction with respect to d, i.e.*

$$d(F(x),\ F(y)) \le \gamma \cdot d(x,y)\ ,\ \forall x,y \in X\ .$$

If our previous assumptions are satisfied, Bessaga's theorem considered for $X=[a,b]$ and F given by (2) shows that there exists a metric on X such that F is a contraction.

This was the basic idea in proving the convergence of Newton's method in [1] - [5], where, in order to assure the invariance of the domain with respect to F, f is prolonged to the whole real axis and the prolongation is denoted by f too:

$$f(x)=\begin{cases} f'(a)\cdot(x-a)+f(a), & \text{if } x<a. \\ f(x), & x\in[a,b] \\ f'(b)\cdot(x-b)+f(b), & \text{if } x>b\,. \end{cases}$$

The main result of this paper is given by

THEOREM 4. *Let* $f\colon [a,\ b]\to\mathbf{R}$ $a<b$, *be a function satisfying* (f_1), (f'_2) and (f_3)

$$2m>M,$$

where

$$m=\min_{x\in[a,b]}|f'(x)|\ ,\ M=\max_{x\in[a,b]}|f'(x)|\ .$$

Then the extended Newton iteration (1) converges to α, *the unique solution of* $f(x)=0$ *in* $[a,\ b]$ *and the following estimation*

$$(4)\qquad |x_n-\alpha| \le \frac{|f'(x_n)|}{m}|x_n-x_{n+1}|\ ,\ n\ge 0\ ,$$

holds.

Proof. The basic idea of the proof is similar to the one in [1]–[5] but there exist diferent arguments.

Obviously, from (f_1) and (f'_2) it results that $f(x)=0$ has a unique solution $\alpha\in(a,b)$

Let $F\colon\mathbf{R}\to\mathbf{R}$ be given by

$$(5)\qquad F(x)=x-\frac{f(x)}{f'(x)}\ ,\ x\in\mathbf{R}\ .$$

Then, α is a solution of $f(x)=0$ if and only if α is a fixed point of F, that is

$$F(\alpha)=\alpha$$

and

$$F(x)-\alpha=x-\frac{f(x)}{f'(x)}-\alpha=x-\alpha-\frac{f(x)}{f'(x)}$$

But

$$f(x)=f(x)-0=f(x)-f(\alpha)\,,$$

hence, from (f_2) and the mean value theorem we deduce

$$f(x)=f'(y)\cdot(x-\alpha)\,,$$

where $y=\alpha+\lambda(x-\alpha)$, $0<\lambda<1$.

Then

$$F(x)-\alpha=(x-\alpha)\cdot\left(1-\frac{f'(y)}{f'(x)}\right),\ \ \forall x\in\mathbf{R}\,.$$

Using (f'_2) it results that f' preserves a sign on $[a,\ b]$, hence $f'(y)/f'(x)>0$ on $\mathbf{R}$ (since $f'(x)=f'(a)$, *if* $x<a$ *and* $f'(x)=f'(b)$, *if* $x>b$).

This means

$$1-\frac{f'(y)}{f'(x)}<1,\ \forall x,y\in\mathbf{R}\,.$$

On the other hand, from (f_3) we obtain

$$\frac{f'(y)}{f'(x)}=\left|\frac{f'(y)}{f'(x)}\right|=\frac{|f'(y)|}{|f'(x)|}\le\frac{M}{m}<2\,,$$

which shows that

$$1-\frac{f'(y)}{f'(x)}>-1,\ \forall x,y\in[a,b]\,.$$

Now from the continuity of f and the fact that f is actually defined on the compact interval $[a,\ b]$ it results

$$(6)\qquad k=\max_{x,\,y\in\mathbf{R}}\left|1-\frac{f'(y)}{f'(x)}\right|<1\,,$$

which together with the previous relations, yields

$$|F(x)-\alpha|\le k\cdot|x-\alpha|\,,\ \forall x\in\mathbf{R}\,,$$

and $0\le k<1$. By induction we then obtain

$$|F^n(x)-\alpha|\le k^n\cdot|x-\alpha|\,,$$

which shows that

$$F^n(x_0)\to\alpha\ ,\ \text{as}\ n\to\infty\,,$$

for each $x_0\in\mathbf{R}$.

In order to obtain (4), from (5) we deduce, by using the mean value theorem,

$$F^{n+1}(x_0) - F^n(x_0) = x_{n+1} - x_n = \frac{f'(c_n)}{f'(x_n)} \cdot (x_n - \alpha)$$

where

$$c_n = \alpha + \mu(x_n - \alpha)\,,\; 0 < \mu < 1\,,$$

which yields

$$|x_n - \alpha| \le \frac{|f'(x_n)|}{m} |x_n - x_{n+1}|\,,\; n \ge 0\,.$$

The proof is now complete.

Remarks. 1) Let us observe that if some x_p (say x_0) does not belong to $[a, b]$, this fact is unimportant for the convergence of the method. If, for example, we have $x_p < a$, then

$$x_{p+1} = x_p - \frac{f(x_p)}{f'(x_p)} = x_p - \frac{f'(a)\cdot(x_p - a) + f(a)}{f'(a)}\,,$$

hence

$$x_{p+1} = a - \frac{f(a)}{f'(a)} > a\,,$$

because, from (f_1) and (f'_2) we have $f(a) \cdot f'(a) < 0$.
In a similar manner we obtain that $x_{p+1} < b$, if $x_p > b$.
This means that the extended Newton's method consists in applying Newton's method on $[a, b]$ and the modified Newton's method for $x > b$ and $x < a$.

2) The estimate (4) is an improvment of the corresponding estimate in [1]–[5]. It shows a linear convergence for Newton's method.

3) The convergence of Newton's method for the function $f(x)$ in Example 1 is now an easy consequence of Theorem 4.

4) In view of Edelstein's fixed point theorem, see [10], condition (f_3) in Theorem 4 may be replaced by a weaker condition. We thus obtain a more general result given by

THEOREM 5. *If f satisfies (f_1), (f'_2) and*

(f'_3) $2m \ge M$,

then the conclusion of Theorem 4 remains true.

Example 3. If f is as in Example 2, we have M=4 and m=2, Theorem 4 does not apply but we can apply Theorem 5. Thus Newton's iteration is convergent.

Remarks. 1) Theorem 4 or 5 may be extended to $\mathbf{R}^n$, as in [1];

2) For f as in Example 1, if we start with any $x_0 < a$ we obtain x_1=2.08, x_3=2.51 and so on, the same iterations as in the case x_0=a;

3) If $f \in C^1[a,b] \setminus C^2[a,b]$ then Newton's method generaly converges linearly, but if however f' is lipschitzian or there exists f'', then the convergence is quadratic;

4) As shown by some recent numerical tests performed on an IBM PC compatible computer, under MATCHAD, condition (f_3) and respectively (f'_3) seems not to be necessary fot the convergence of Newton's iteration.

REFERENCES

1. Berinde, V., *Generalized contractions and applications (Romanian)*, Ph. D. Thesis, Univ. "Babeş-Bolyai" Cluj-Napoca, 1993.
2. Berinde, V., *An extension of the Newton-Raphson method for solving nonlinear equations*, Proceed. Int. Conference MICRO CAD '93, 9–11 nov. 1993, Technical University Košice, pp. 63–64.
3. Berinde, V., *A fixed point proof of the convergence of the Newton method*, Proceed. Int. Conference Micro CAD '94, Univ. of Miskolc, 2–4 March 1994, pp. 14–21.
4. Berinde, V., *On some exit criteria for the Newton method*, Rev. – Res. Fac. Sc. – Univ. Novi Sad (submmited).
5. Berinde, V., *On the convergence of the Newton method*, Transactions of the Technical Univ. of Košice (to appear).
6. Demidovici, B. P., Maron, A. I., *Computational mathematics*, MIR Publishers, Moscow, 1987.
7. Kantorovici, L. V., Akilov, G. P., *Functional analysis* (Romanian), Editura Ştiinţifică şi Enciclopedică Bucureşti, 1986.
8. Ortega, J., Rheinboldt, W. C., *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, 1970.
9. Ostrowski, A., *Solution of equations and systems of equations*, Academic Press, New York, 1966.
10. Rus, A. I., *Principles and applications of the fixed point theory* (Romanian), Editura Dacia, Cluj-Napoca, 1979.
11. Măruşter, Şt., *Numerical methods in solving nonlinear equations* (Romanian), Editura Tehnică, Bucureşti, 1981.

Received 15 XII 1994

Department of Mathematics
University of Baia Mare
Victoriei 76, 4800, Baia Mare
Romania