REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Tome XXVI, N^{os} 1–2, 1997, pp. 103–108

ON THE HERON'S METHOD FOR APPROXIMATING THE CUBIC ROOT OF A REAL NUMBER

DAN LUCA, ION PĂVĂLOIU

1. INTRODUCTION

In this paper we shall specify and go deeply into some problems presented in [1], concerning the Heron's method for approximating the cubic root of a positive number.

The authors of [1] construct a method based on the Heron's algorithm for computing the cubic root of 100.

The method works as follows: Given two real numbers a and b satisfying $a^3 < N < b^3$, the Heron's method for approximating $\sqrt[3]{N}$ is

$$\phi(N,a,b) = a + \frac{bd_1}{bd_1 + ad_2}(b-a), \tag{1.1}$$

where $d_1 = N - a^3$ and $d_2 = b^3 - N$. We shall show that the approximation (1.1) of $\sqrt[3]{N}$ follows from the *regula falsi* applied to the equation $x^2 - \dfrac{N}{x} = 0$ [2]. This will give a rigorous interpretation of (1.1), and the results from [1] will be reached again.

Using results from [4], we shall give other error bounds than those in [1]. On the other hand, the method is generalized to the case $\sqrt[p]{N}$, $p \in \mathbf{N}$, $p \geq 2$, the method also offering bilateral approximations. Some remarks on applying the results in [4] for the error bounds will lead us to the generalization of the Heron's method.

2. HERON'S METHOD AND *REGULA FALSI*

A. In order to approximate the cubic root of $N > 0$ by (1.1), consider the function $f:[a,b] \to \mathbf{R}$ $f(x) = x^3 - N, 0 < a < b$ and the function $g:[a,b] \to \mathbf{R}$, $g(x) = \dfrac{f(x)}{x}$.

1991 AMS Subject Classification: 65G05, 65B05.

It is well known that *regula falsi* applied to the equation $g(x) = 0$ leads to the following approximation of its root

$$c = a - \frac{g(a)}{[a, b; g]}, \tag{2.1}$$

$[u, v; g]$ denoting the first-order divided difference of g on the nodes u and v. It can be easily verified that $c = \phi(N, a, b)$.

Taking into account that $c \in (a, b)$ and denoting by $[u, v, w; g]$ the second-order divided difference of g on the points u, v, w, we get

$$g(x) = g(a) + [a, b; g](x - a) + [a, b, x; g](x - a)(x - b) \tag{2.2}$$

for all $x \in (a, b)$.

For $x = \sqrt[3]{N}$ in (2.2) we obtain

$$g(a) + [a, b; g]\left(\sqrt[3]{N} - a\right) + \left[a, b, \sqrt[3]{N}; g\right]\left(\sqrt[3]{N} - a\right)\left(\sqrt[3]{N} - b\right) = 0,$$

from which, by dividing $[a, b; g]$ it follows

$$c - \sqrt[3]{N} = \frac{\left[a, b, \sqrt[3]{N}; g\right]}{[a, b; g]}\left(\sqrt[3]{N} - a\right)\left(\sqrt[3]{N} - b\right). \tag{2.3}$$

An elementary calculation on $\dfrac{\left[a, b, \sqrt[3]{N}; g\right]}{[a, b; g]}$ shows that

$$\frac{c - \sqrt[3]{N}}{\sqrt[3]{N}} = \frac{\sqrt[3]{N} + \sqrt{ab}}{\sqrt[3]{N}\,[ab(a + b) + N]}\left(\sqrt[3]{N} - a\right)\left(\sqrt[3]{N} - b\right)\left(\sqrt[3]{N} - \sqrt{ab}\right), \tag{2.4}$$

which gives Theorem 3, [1].

Taking into account the above remarks and using the evaluations obtained by T. Popoviciu in [4], (2.3) gives the following error bounds

$$\frac{m_2}{M_1}\left(\sqrt[3]{N} - a\right)\left(b - \sqrt[3]{N}\right) \le \left|c - \sqrt[3]{N}\right| \le \frac{M_2}{m_1}\left(\sqrt[3]{N} - a\right)\left(b - \sqrt[3]{N}\right), \tag{2.5}$$

where

$$m_1 = 3a$$

$$m_2 = \min\left\{\frac{N}{a^3} - 1, 1 - \frac{N}{b^3}\right\}$$

$$M_1 = \max\left\{\frac{2a^3 + N}{a^2}, \frac{2b^3 + N}{b^2}\right\}$$

$$M_2 = \max\left\{\frac{N}{a^3} - 1, 1 - \frac{N}{b^3}\right\}.$$

Note that (2.5) leads to a very good error evaluation; since a and b are close to N, then m_2 and M_2 are close to zero. This is implied by the fact that the function g and its second derivative vanish at the same point $x = \sqrt[3]{N}$.

B. It can be easily seen that the method presented at **A** can be generalized. For the approximation of the root of order p of the real number N, $p \in \mathbf{N}$, $p \ge 2$, consider the function $f_1 : [a_1, b_1] \to \mathbf{R}$, $f_1(x) = x^p - N$, $0 < a_1 < b_1$, $a_1^p < N < b_1^p$ and the function $g_1 : [a_1, b_1] \to \mathbf{R}$, $g_1(x) = \dfrac{f_1(x)}{x^q}$, where $q = \dfrac{p-1}{2}$. The function g_1 satisfies $g_1\left(\sqrt[p]{N}\right) = g_1''\left(\sqrt[p]{N}\right)$.

Applying *regula falsi* to the equation $g_1(x) = 0$, we obtain

$$c_1 = a_1 - \frac{g_1(a_1)}{[a_1, b_1; g_1]}. \tag{2.6}$$

Similar to **A**, we obtain

$$c_1 - \sqrt[p]{N} = \frac{\left[a_1, b_1, \sqrt[p]{N}; g_1\right]}{[a_1, b_1; g_1]}\left(\sqrt[p]{N} - a_1\right)\left(\sqrt[p]{N} - b_1\right), \tag{2.7}$$

which gives

$$\frac{t_2}{2T_1}\left(\sqrt[p]{N} - a_1\right)\left(b_1 - \sqrt[p]{N}\right) \le \left|c_1 - \sqrt[p]{N}\right| \le \frac{T_2}{2t_1}\left(\sqrt[p]{N} - a_1\right)\left(b_1 - \sqrt[p]{N}\right), \tag{2.8}$$

where

$$t_1 = p a_1^{\frac{p-1}{2}}$$

$$t_2 = \min\left\{\frac{(p-1)(p+1)\left(N - a_1^p\right)}{4a_1^{\frac{p+3}{2}}}, \frac{(p-1)(p+1)\left(b_1^p - N\right)}{4b_1^{\frac{p+3}{2}}}\right\}$$

$$T_1 = \max\left\{ \frac{(p+1)a_1^p + (p-1)N}{2a_1^{\frac{p+1}{2}}}; \frac{(p+1)b_1^p + (b-1)N}{2b_1^{\frac{p+1}{2}}} \right\}$$

$$T_2 = \max\left\{ \frac{(p-1)(p+1)\left(N - a_1^p\right)}{4a_1^{\frac{p+3}{2}}}; \frac{(p-1)(p+1)\left(b_1^p - N\right)}{4b_1^{\frac{p+3}{2}}} \right\}.$$

3. STEFFENSEN'S METHOD FOR APPROXIMATING THE pTH-ORDER ROOT

Let $I = [\alpha, \beta]$, $\alpha < \beta$ be an interval of the real axis.

Consider the equation

(3.1) $$F(x) = 0,$$

where $F : I \to \mathbf{R}$. Suppose that equation (3.1) has a root $\bar{x} \in (\alpha, \beta)$. Consider also a function $h : I \to \mathbf{R}$ such that the equation

(3.2) $$x - h(x) = 0$$

is equivalent to (3.1).

The Steffensen's method consists in the generation of two sequences (x_n) and $(h(x_n))$ by the relations

(3.3) $$x_{n+1} = x_n - \frac{F(x_n)}{[x_n, h(x_n); F]}, \quad x_0 \in I, \quad n = 0, 1, \dots .$$

As we shall see, this method offers the possibility to obtain better both upper and lower approximations, by starting with a lower approximation of $\sqrt[p]{N}$. Then, by applying only once the *regula falsi* (2.6), the precision can be increased.

As concerns the convergence of (x_n) and $(h(x_n))$ in (3.3), in [3] it is proved the following

THEOREM 3.1 [3]. *If the functions $F : I \to \mathbf{R}$ and $h : I \to \mathbf{R}$ are continuous and satisfy the following conditions:*

i) *the function h is decreasing on I,*
ii) *the function F is increasing and convex on I,*
iii) *there exists $x_0 \in I$ such that $F(x_0) < 0$ and $h(x_0) \in I$,*
iv) *the equations* (3.1) *and* (3.2) *are equivalent,*

then the following properties hold:

j) *the sequence (x_n) is increasing,*
jj) *the sequence $(h(x_n))$ is decreasing,*
jjj) $\lim x_n = \lim h(x_n) = \bar{x}$,
jv) *the relations $x_n \le x_{n+1} \le \bar{x} \le h(x_n)$ hold for all $n = 0, 1, \dots$,*
v) $\bar{x} - x_{n+1} < h(x_n) - x_{n+1}$.

Applying this Theorem for $F : [\alpha, \beta] \to \mathbf{R}$, $F(x) = x^p - N$, $h : [\alpha, \beta] \to \mathbf{R}$, $h(x) = x - \frac{F(x)}{p\alpha^{p-1}}$, where $0 < \alpha < \beta$ and $p \in \mathbf{N}$, $p \ge 2$, we obtain

$$x_{n+1} = x_n + \frac{\left(p\alpha^{p-1}\right)^{p-1}\left(x_n^p - N\right)^2}{\left(p\alpha^{p-1}x_n - x_n^p + N\right)^p - \left(p\alpha^{p-1}x_n\right)^p}, \quad x_0 = \alpha, \quad n = 0, 1, \dots; \alpha^p < N.$$

Since F is increasing and convex on $[\alpha, \beta]$, it follows that h is decreasing on $[\alpha, \beta]$, and the equations $F(x) = 0$ and $h(x) - x = 0$ are equivalent. So the conclusion of Theorem 3.1 follows.

The sequences (x_n) and $(h(x_n))$ being convergent, it follows that for all $\varepsilon > 0$ there exists $n_0 \in \mathbf{N}$ such that for $n \ge n_0$ we have

$$h(x_n) - x_n < \varepsilon,$$

which implies $\sqrt[p]{N} - x_n < \varepsilon$ $h(x_n) - \sqrt[p]{N} < \varepsilon$.

If we use (2.6) for $a_1 = x_n$ and $b_1 = h(x_n)$ and $g_1(x) = \frac{F(x)}{x^{\frac{p-1}{2}}}$, and we denote the approximation obtained by c_n, then by (2.8) we have

$$\left|c_n - \sqrt[p]{N}\right| \le \frac{T'_2}{2t'_1}\varepsilon^2,$$

where

$$T'_2 = \max\left\{ \frac{(p-1)(p+1)\left(N - x_n^p\right)}{4x_n^{\frac{p+3}{2}}}; \frac{(p-1)(p+1)\left(h^p(x_n) - N\right)}{4h(x_n)^{\frac{p+3}{2}}} \right\}.$$

$$t'_1 = p x_n^{\frac{p-1}{2}}$$

4. A NUMERICAL EXAMPLE

We intend to apply the method described in Section 3 for the approximation of the number $\sqrt[5]{100}$, i.e., for solving the equation $x^5 - 100 = 0$. In this case we have

$$F(x) = x^5 - 100$$

and, taking $\alpha = 2$, for the function h we have

$$h(x) = x - \frac{1}{80} F(x).$$

Considering $x_0 = \alpha = 2$ and using (3.3), with F and h given above, we obtain for the sequences $(x_n)_{n\geq 0}$ and $(h(x_n))_{n\geq 0}$ the following values:

n	x_n	$h(x_n)$	$\varepsilon_n = h(x_n) - x_n$
0	2.0000000000	2.8500000000	8.5000000000E –01
1	2.3704445072	2.6849117966	3.1446728941E –01
2	2.4927536892	2.5396394928	4.6885803578E –02
3	2.5114651493	2.5125130194	1.0478700715E –03
4	2.5118862213	2.5118867443	5.2291215979E –07
5	2.5118864315	2.5118864315	3.6379788071E –12.

REFERENCES

1. G. Deslauries and S. Dubuc, *Le calcul de la racine cubique selon Héron*, Elemente der Mathematik **51**, *1* (1996), 28–34.
2. M. Ostrowski, *A Solution of Equations and Systems of Equations*, Academic Press, New York-London, 1960.
3. I. Păvăloiu, *On the monotonicity of the sequences of approximations obtained by Steffensen's method*, Mathematica (Cluj) **35 (58)**, *1* (1993), 71–76.
4. T. Popoviciu, *Sur la délimitation de l'erreur dans l'approximation des racines d'une équation par interpolation linéaire ou quadratique*, Rev. Roumaine Math. Pures Appl. **XIII**, *1* (1968), 75–78.

Received May 15, 1996

Institutul de Calcul "Tiberiu Popoviciu"
Str. G. Bilaşcu, nr. 37
C.P. 68, O.P. 1
3400 Cluj-Napoca
România