Computing 30, 171 – 178 (1983)

Computing

Short Communications / Kurze Mitteilungen

On the Convergence Order of a Modified Method for Simultaneous Finding Polynomial Zeros

G. V. Milovanović and **M. S. Petković**, Niš

Received April 16, 1982

Abstract — Zusammenfassung

On the Convergence Order of a Modified Method for Simultaneous Finding Polynomial Zeros. Using Newton's corrections and Gauss-Seidel approach, a modification of single-step method [1] for the simultaneous finding all zeros of an n-th degree polynomial is formulated in this paper. It is shown that R-order of convergence of the presented method is at least $2(1+\tau_n)$, where $\tau_n \in (1,2)$ is the unique positive zero of the polynomial $\tilde{f}_n(\tau)=\tau^n-\tau-1$. Faster convergence of the modified method in reference to the similar methods is attained without additional calculations. Comparison is performed in the example of an algebraic equation.

Über die Konvergenz Ordnung der modifizierten Methode zur gleichzeitigen Ermittlung der Polynomwurzeln. In dieser Arbeit wird eine Modifikation einer Einschritt-Methode [1] zur gleichzeitigen Ermittlung aller Nullstellen eines Polynoms n-ter Ordnung unter Verwendung des Gauss-Seidel-Vorgehens und Newtonscher Korrekturen vorgestellt. Es wird gezeigt, daß die R-Ordnung der vorgestellten Methode mindestens $2(1+\tau_n)$ beträgt, wobei $\tau_n \in (1,2)$ die eindeutige positive Wurzel des Polynoms $\tilde{f}_n(\tau)=\tau^n-\tau-1$ darstellt. Es wird eine schnellere Konvergenz der modifizierten Methode im Vergleich zu ähnlichen Methoden erreicht, und zwar ohne zusätzlichen Rechenaufwand. Ein Vergleichsbeispiel mit einer algebraischen Gleichung wird präsentiert.

AMS Subject Classification: 65H05.

Key words: Determination of polynomial zeros, simultaneous iterative methods, accelerated convergence, R-order of convergence.

1. Introduction

Consider a polynomial P of degree $n>1$

$$P(z)=z^n+a_{n-1}z^{n-1}+\ldots+a_1 z+a_0=\prod_{j=1}^{n}(z-r_j) \quad (a_i \in C),$$

with simple real or complex zeros $r_1,\ldots,r_n$. Let $z_1,\ldots,z_n$ be distinct approximations for these zeros.

 0010-485X/83/0030/0171/$01.60

Using the logarithmic derivative of $P(z)$ in the point $z=z_i$, we get

$$\frac{P'(z_i)}{P(z_i)}=\sum_{j=1}^{n}\frac{1}{z_i-r_j}=\frac{1}{z_i-r_i}+\sum_{\substack{j=1\\j\neq i}}^{n}\frac{1}{z_i-r_j},$$

i.e.

$$r_i=z_i-\frac{1}{\dfrac{P'(z_i)}{P(z_i)}-\sum\limits_{\substack{j=1\\j\neq i}}^{n}\dfrac{1}{z_i-r_j}}\qquad(i=1,\ldots,n).\tag{1.1}$$

Putting $r_i\cong\hat{z}_i$ in (1.1), where $\hat{z}_i$ is new approximation for the zero r_i, and taking some approximations for r_j $(j\neq i)$ on the right-hand side of the above identity, some iterative processes for the simultaneous finding zeros of the polynomial P follow from (1.1):

For $r_j:=z_j$ $(j\neq i)$ we obtain the well-known total-step method

$$\hat{z}_i=z_i-\frac{1}{\dfrac{P'(z_i)}{P(z_i)}-\sum\limits_{\substack{j=1\\j\neq i}}^{n}\dfrac{1}{z_i-z_j}}\qquad(i=1,\ldots,n),\tag{1.2}$$

which has been the subject of many papers ([3], [4], [5], [6], [7]). The convergence order of this method is three.

Let $r_j:=\hat{z}_j$ $(j<i)$ and $r_j:=z_j$ $(j>i)$, then single-step method follows from (1.1)

$$\hat{z}_i=z_i-\frac{1}{\dfrac{P'(z_i)}{P(z_i)}-\sum\limits_{j=1}^{i-1}\dfrac{1}{z_i-\hat{z}_j}-\sum\limits_{j=i+1}^{n}\dfrac{1}{z_i-z_j}}\qquad(i=1,\ldots,n)\tag{1.3}$$

(see [1], [5]). Alefeld and Herzberger have proved in [1] that the R-order of convergence of this method is at least $2+\sigma_n\,(>3)$, where $\sigma_n>1$ is the unique positive zero of the polynomial $p_n(\sigma)=\sigma^n-\sigma-2$.

Taking

$$r_j:=z_j+\Delta z_j=z_j-\frac{P(z_j)}{P'(z_j)}\qquad(j\neq i)$$

in (1.1) we obtain

$$\hat{z}_i=z_i-\frac{1}{\dfrac{P'(z_i)}{P(z_i)}-\sum\limits_{\substack{j=1\\j\neq i}}^{n}\dfrac{1}{z_i-z_j-\Delta z_j}}\qquad(i=1,\ldots,n),\tag{1.4}$$

which is considered in [8]. It is shown in this paper that the iterative process (1.4) has the convergence order equal to four.

Finally, for

$$r_j:=\hat{z}_j\ (j<i)\quad\text{and}\quad r_j:=z_j+\Delta z_j=z_j-\frac{P(z_j)}{P'(z_j)}\quad(j>i),$$

we have

$$\hat{z}_i = z_i - \cfrac{1}{\cfrac{P'(z_i)}{P(z_i)} - \sum_{j=1}^{i-1} \cfrac{1}{z_i - \hat{z}_j} - \sum_{j=i+1}^{n} \cfrac{1}{z_i - z_j - \Delta z_j}} \quad (i=1, \ldots, n). \tag{1.5}$$

The last iterative formula will be considered in this paper. Note that (1.5) is combination of the formulas (1.3) and (1.4).

Assume that $z_1^{(0)}, \ldots, z_n^{(0)}$ are distinct and sufficiently good approximations for the zeros $r_1, \ldots, r_n$. Taking $z_i = z_i^{(m)}$, $\hat{z}_i = \hat{z}_i^{(m+1)}$ and

$$\Delta z_i = \Delta z_i^{(m)} = -P(z_i^{(m)})/P'(z_i^{(m)})$$

(Newton's correction), starting from (1.5) we obtain the following iterative process for the simultaneous determination of polynomial complex zeros:

$$z_i^{(m+1)} = z_i^{(m)} - \cfrac{1}{\cfrac{P'(z_i^{(m)})}{P(z_i^{(m)})} - \sum_{j=1}^{i-1} \cfrac{1}{z_i^{(m)} - z_j^{(m+1)}} - \sum_{j=i+1}^{n} \cfrac{1}{z_i^{(m)} - z_j^{(m)} - \Delta z_j^{(m)}}} \quad (i=1, \ldots, n). \tag{1.6}$$

In the next section, we shall prove that the method (1.6) converges with the R-order of convergence higher than 4 for sufficiently good starting values $z_1^{(0)}, \ldots, z_n^{(0)}$.

2. Convergence Analysis

Ortega and Rheinboldt have introduced in [9] the following definition of the convergence order:

Definition: Let IP be an iterative process with the limit point $\boldsymbol{r}$. Then the quantity

$$O_R(IP, \boldsymbol{r}) = \begin{cases} +\infty \text{ if } R_p(IP, \boldsymbol{r}) = 0 \text{ for all } p \in [1, +\infty), \\ \inf\{p \in [1, +\infty) \mid R_p(IP, \boldsymbol{r}) = 1\} \text{ otherwise} \end{cases}$$

is called the R-order of IP at $\boldsymbol{r}$.

$R_p(IP, \boldsymbol{r})$ is called the R-factor of IP at $\boldsymbol{r}$ and is defined by

$$R_p(IP, \boldsymbol{r}) = \sup\{R_p\{\boldsymbol{z}^{(m)}\} \mid \{\boldsymbol{z}^{(m)}\} \in C(IP, \boldsymbol{r})\} \quad (1 \leq p < +\infty),$$

where

$$R_p\{\boldsymbol{z}^{(m)}\} = \begin{cases} \limsup_{m \to +\infty} \|\boldsymbol{z}^{(m)} - \boldsymbol{r}\|^{1/m} & \text{if } p = 1, \\ \limsup_{m \to +\infty} \|\boldsymbol{z}^{(m)} - \boldsymbol{r}\|^{1/p^m} & \text{if } p > 1, \end{cases}$$

and $C(IP, \boldsymbol{r})$ is the set of all sequences produced by IP and converging to $\boldsymbol{r}$.

We shall use this definition for analysis of the convergence order of the iterative method (1.6).

Let $m=0,1,\ldots$ be the index of iteration and let

$$d=\min_{i\neq j}|r_i-r_j|,\quad v_j^{(m)}=z_j^{(m)}-r_j,$$

$$w_j^{(m)}=z_j^{(m)}+\Delta z_j^{(m)}\quad\text{(Newton approximation)},$$

$$\hat{g}_i^{(m)}=\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{z_i^{(m)}-r_i}{z_i^{(m)}-r_j},$$

$$g_i^{(m)}=\sum_{j=1}^{i-1}\frac{(z_i^{(m)}-r_i)(r_j-z_j^{(m+1)})}{(z_i^{(m)}-r_j)(z_i^{(m)}-z_j^{(m+1)})}+\sum_{j=i+1}^{n}\frac{(z_i^{(m)}-r_i)(r_j-w_j^{(m)})}{(z_i^{(m)}-r_j)(z_i^{(m)}-w_j^{(m)})}.$$

It is easy to show that Newton's method can be written in the form

$$w_j^{(m)}-r_j=\frac{\hat{g}_j^{(m)}}{1+\hat{g}_j^{(m)}}(z_j^{(m)}-r_j)=\frac{\hat{g}_j^{(m)}}{1+\hat{g}_j^{(m)}}v_j^{(m)}. \tag{2.1}$$

Similarly, for the iterative process (1.6) we have

$$z_i^{(m+1)}=r_i+\frac{g_i^{(m)}}{1+g_i^{(m)}}(z_i^{(m)}-r_i)\qquad(i=1,\ldots,n;\ m=0,1,\ldots),$$

wherefrom

$$v_i^{(m+1)}=\frac{g_i^{(m)}}{1+g_i^{(m)}}v_i^{(m)}\qquad(i=1,\ldots,n;\ m=0,1,\ldots). \tag{2.2}$$

Suppose that the initial conditions

$$|v_i^{(0)}|<\frac{1}{q}=\frac{d}{2n-1}\quad(i=1,\ldots,n) \tag{2.3}$$

are satisfied. Then, for $i\neq j$, we have

$$|z_i^{(0)}-r_j|\geqq|r_i-r_j|-|z_i^{(0)}-r_i|>d-\frac{d}{2n-1},$$

$$|z_i^{(0)}-z_j^{(0)}|\geqq|z_i^{(0)}-r_j|-|z_j^{(0)}-r_j|>\left(d-\frac{d}{2n-1}\right)-\frac{d}{2n-1},$$

wherefrom

$$|z_i^{(0)}-r_j|>\frac{2(n-1)}{q}\quad\text{and}\quad|z_i^{(0)}-z_j^{(0)}|>\frac{2n-3}{q}\geqq\frac{1}{q}. \tag{2.4}$$

We shall now determine the estimate for $|\hat{g}_i^{(0)}|$. On the basis of (2.3) and (2.4) we have

$$|\hat{g}_i^{(0)}|\leqq|v_i^{(0)}|\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{|z_i^{(0)}-r_j|}<\frac{q}{2}|v_i^{(0)}|<\frac{1}{2}. \tag{2.5}$$

Then, from (2.1), for $m=0$ it follows

$$|r_j-w_j^{(0)}|\leqq\frac{|\hat{g}_j^{(0)}|}{1-|\hat{g}_j^{(0)}|}|v_j^{(0)}|<q|v_i^{(0)}|^2<\frac{1}{q}. \tag{2.6}$$

Since, for $i \neq j$

$$|z_i^{(0)} - w_j^{(0)}| \geqq |z_i^{(0)} - r_j| - |r_j - w_j^{(0)}| > \frac{2n-3}{q} \geqq \frac{1}{q},$$

according to the previous estimates we find

$$\frac{|r_j - w_j^{(0)}|}{|z_i^{(0)} - r_j|\,|z_i^{(0)} - w_j^{(0)}|} < \frac{q^3}{2(n-1)}\,|v_j^{(0)}|^2. \tag{2.7}$$

Estimate the values $|v_i^{(1)}|$ and $|g_i^{(0)}|$. In regard to (2.2) we obtain

$$|v_i^{(1)}| \leqq \frac{|g_i^{(0)}|}{1 - |g_i^{(0)}|}\,|v_i^{(0)}| \qquad (i = 1, \ldots, n). \tag{2.8}$$

Since

$$|g_1^{(0)}| = \left| \sum_{j=2}^{n} \frac{(z_1^{(0)} - r_1)(r_j - w_j^{(0)})}{(z_1^{(0)} - r_j)(z_1^{(0)} - w_j^{(0)})} \right|,$$

applying (2.3) and (2.7) we find

$$|g_1^{(0)}| < \frac{q^3}{2(n-1)} \cdot \frac{n-1}{q^2}\,|v_1^{(0)}| < \frac{1}{2},$$

wherefrom, according to (2.8), it follows

$$|v_1^{(1)}| < |v_1^{(0)}| < \frac{1}{q}.$$

Using the above consideration and the inequality

$$|z_2^{(0)} - z_1^{(1)}| \geqq |z_2^{(0)} - r_1| - |v_1^{(1)}| > \frac{1}{q},$$

we successively obtain for $i = 2, \ldots, n$ the following estimates:

$$|g_i^{(0)}| < |v_i^{(0)}| \left(\frac{q^2}{2(n-1)} \sum_{j=1}^{i-1} |v_j^{(1)}| + \frac{q^3}{2(n-1)} \sum_{j=i+1}^{n} |v_j^{(0)}|^2 \right) < \frac{1}{2},$$

$$|v_i^{(1)}| < \frac{q^2}{n-1}\,|v_i^{(0)}|^2 \left(\sum_{j=1}^{i-1} |v_j^{(1)}| + q \sum_{j=i+1}^{n} |v_j^{(0)}|^2 \right) < \frac{1}{q}.$$

Applying mathematical induction we prove that the inequality

$$|v_i^{(m+1)}| < \frac{q^2}{n-1}\,|v_i^{(m)}|^2 \left(\sum_{j=1}^{i-1} |v_j^{(m+1)}| + q \sum_{j=i+1}^{n} |v_j^{(m)}|^2 \right) < \frac{1}{q} \tag{2.9}$$

hold for every $m = 0, 1, \ldots$.

Theorem: *Under the conditions* (2.3), *the iterative process* (1.6) *is convergent with the R-order*

$$O_R((1.6), r) \geqq 2(1 + \tau_n),$$

where $\tau_n \in (1, 2)$ *is the unique positive root of the equation*

$$\tilde{f}_n(\tau) = \tau^n - \tau - 1$$

and $\boldsymbol{r} = [r_1\, r_2 \ldots r_n]^T$ *is the limit point.*

Proof: Substituting $q\,|v_i^{(m)}|=h_i^{(m)}$ in (2.9), we get

$$h_i^{(m+1)}<\frac{1}{n-1}(h_i^{(m)})^2\left(\sum_{j=1}^{i-1}h_i^{(m+1)}+\sum_{j=i+1}^{n}(h_i^{(m)})^2\right), \tag{2.10}$$

where $i=1,\dots,n$ and $m=0,1,\dots$. On the basis of (2.3), $h_i^{(0)}<1$ holds for each $i=1,\dots,n$. If we put $h=\max_{1\le i\le n}h_i^{(0)}$, then

$$h_i^{(0)}\leqq h<1\quad (i=1,\dots,n).$$

According to (2.10) we conclude that the iterative process is convergent. Further, we can write

$$h_i^{(m+1)}\leqq h^{u_i^{(m-1)}}\quad (i=1,\dots,n;\ m=0,1,\dots).$$

Defining the matrix B by $B=2A$, where

$$A=\begin{bmatrix}1 & 1 & & & & \\ & 1 & 1 & & 0 & \\ & & & \ddots & & \\ & 0 & & \ddots & & \\ & & & & \ddots & \\ & & & & 1 & 1\\ 1 & 1 & 0 & \dots & 0 & 1\end{bmatrix},$$

the vectors $\boldsymbol{u}^{(m)}=[u_1^{(m)}\dots u_n^{(m)}]^T$ can successively be calculated by

$$\boldsymbol{u}^{(m+1)}=B\,\boldsymbol{u}^{(m)}\quad (m=0,1,\dots),$$

with $\boldsymbol{u}^{(0)}=[1\dots 1]^T$. The proof is by induction and will be omitted. The matrix A is the same as the corresponding matrix in [2, Ch. 8] and consequently, following the proof given in [2] (see, also [1]), we conclude that

$$O_R((1.6),r)\geqq\rho(B)=2\rho(A), \tag{2.11}$$

where $\rho(A)$ and $\rho(B)$ are the spectral radii of A and B.

The characteristic polynomial f_n of the matrix A is

$$f_n(\lambda)=(\lambda-1)^n-(\lambda-1)-1.$$

Putting $\tau=\lambda-1$, we obtain

$$\tilde{f}_n(\tau)=f_n(1+\tau)=\tau^n-\tau-1.$$

Since $\tilde{f}_n(1)=-1<0$ and $\tilde{f}_n(2)=2^n-3>0$, there is a zero $\tau_n\in(1,2)$. Besides, by Descartes' rule of signs, there can be no other positive zero of $\tilde{f}_n(\tau)$. Hence, we conclude that $\rho(A)=1+\tau_n$. According to this, from (2.11) we find a lower bound for the R-order of the iterative method (1.6)

$$O_R((1.6),r)\geqq 2(1+\tau_n).\quad\blacksquare$$

The lower bound for $O_R((1.6),r)$ given by $\rho(B)=2(1+\tau_n)$, is tabulated for $n=2(1)10$.

Table 1

n	2	3	4	5	6	7	8	9	10
$\rho(B)$	5.236	4.649	4.441	4.335	4.269	4.226	4.194	4.170	4.152

3. Numerical Results

In order to test the modified method (1.6), the routine on FORTRAN IV was written for PDP 11/40 system in double precision arithmetic. First, before calculating new approximations, the values

$$z_i^{(m)} = -P(z_i^{(m)})/P'(z_i^{(m)}) \qquad (i=1, \ldots, n),$$

which appear in front of the sums in (1.6), were calculated. The same values were used as the correction terms in the second sum too. Thus, the iterative process (1.6) does not require the additional calculation of the values of P and P' in reference to the methods (1.2), (1.3) and (1.4). Comparing with Nourein's modification (1.4), the method (1.6): (i) requires less numerical operations and (ii) occupies less storage space at digital computer because of the use of the previous calculated approximations in the same iteration.

According to the previous remarks and the fact that the convergence order is higher than 4, we conclude that the modified method (1.6) has the advantage in reference to the methods (1.2), (1.3) and (1.4), obtained from (1.1). This conclusion is illustrated numerically in the example of the algebraic equation

$$z^7 + z^5 - 10\,z^4 - z^3 - z + 10 = 0$$

which the exact zeros are $r_1 = 2$, $r_{2,3} = \pm 1$, $r_{4,5} = \pm i$, $r_{6,7} = -1 \pm 2\,i$. As the initial approximations the following complex numbers were taken:

$$z_1^{(0)} = 2.2,\ z_2^{(0)} = 1.2 + 0.1\,i,\ z_3^{(0)} = -0.8 - 0.1\,i,\ z_4^{(0)} = 0.1 + 1.2\,i,$$
$$z_5^{(0)} = -0.1 - 0.8\,i,\ z_6^{(0)} = -1.1 + 2.2\,i,\ z_7^{(0)} = -1.1 - 1.8\,i.$$

These starting values were chosen under weaker conditions than (2.3) ((2.3) requires $|z_i^{(0)} - r_i| < q = 1/13 \cong 0.077$, in the example we have $|z_i^{(0)} - r_i| < 0.224 \cong 2.9\,q$). Many examples of the algebraic equations show that the initial distances $|z_i^{(0)} - r_i|$ can be chosen to be considerably greater than q, defined by (2.3). The condition (2.3) is required to be strong because the convergence analysis is done using crude estimates and weak inequalities.

The numerical results, obtained after the first and the second iteration applying the method (1.6), are shown in Table 2.

Table 2

	$m=1$	$m=2$
$z_1^{(m)}$	$1.99936 - 4.46 \times 10^{-4}\,i$	$2.0000000000003951 + 3.03 \times 10^{-13}\,i$
$z_2^{(m)}$	$1.00112 + 2.02 \times 10^{-3}\,i$	$0.9999999999401543 - 6.69 \times 10^{-11}\,i$
$z_3^{(m)}$	$-1.00054 + 7.35 \times 10^{-4}\,i$	$-1.0000000000141856 + 4.12 \times 10^{-13}\,i$
$z_4^{(m)}$	$-2.06 \times 10^{-3} + 1.00226\,i$	$1.43 \times 10^{-11} + 0.9999999999529638\,i$
$z_5^{(m)}$	$3.26 \times 10^{-3} - 1.00179\,i$	$2.94 \times 10^{-13} - 1.0000000000003203\,i$
$z_6^{(m)}$	$-1.00010 + 1.99957\,i$	$-1.0000000000000003 + 1.9999999999999997\,i$
$z_7^{(m)}$	$-0.99990 - 2.00005\,i$	$-1.0000000000000000 - 2.0000000000000000\,i$

Let $z^{(m)}=[z_1^{(m)} \dots z_n^{(m)}]^T$ be the vector of approximations to the zeros in m-th iteration. Take Euklid's norm

$$e^{(m)}:=\| z^{(m)}-r \|_E=\left(\sum_{i=1}^{n} | z_i^{(m)}-r_i|^2\right)^{1/2}$$

as a measure of closeness of approximations with regard to the exact zeros. In order to compare the method (1.6) with the methods (1.2), (1.3) and (1.4), the above equation was solved with the same initial values applying the mentioned methods. The corresponding values $e^{(m)}$ for every process are given in Table 3.

Table 3

m	(1.2)	(1.3)	(1.4)	(1.6)
1	2.80×10^{-2}	1.78×10^{-2}	9.96×10^{-3}	5.49×10^{-3}
2	4.01×10^{-6}	8.47×10^{-7}	2.19×10^{-9}	1.03×10^{-10}

From the above table, faster convergence of the method (1.6) is evident.

References

[1] Alefeld, G., Herzberger, J.: On the convergence speed of some algorithms for the simultaneous approximation of polynomial roots. SIAM J. Numer. Anal. *2*, 237–243 (1974).
[2] Alefeld, G., Herzberger, J.: Einführung in die Intervallrechnung. Zürich 1974.
[3] Börsch-Supan, W.: A posteriori error bounds for the zeros of polynomials. Numer. Math. *5*, 380–398 (1963).
[4] Dočev, K., Byrnev, P.: Certain modifications of Newton's method for the approximate solution of algebraic equations. Ž. Vyčisl. Mat. i Fiz. *4*, 915–920 (1964).
[5] Ehrlich, L. W.: A modified Newton method for polynomials. Comm. ACM *10*, 107–108 (1967).
[6] Farmer, M. R., Loizou, G.: A class of iteration functions for improving simultaneously approximations to the zeros of a polynomial. BIT *15*, 250–258 (1975).
[7] Maehly, H. J.: Zur iterativen Auflösung algebraischer Gleichungen. Z. Angew. Math. Phys. *5*, 260–263 (1954).
[8] Nourein, A. W.: An improvement on two iteration methods for simultaneous determination of the zeros of a polynomial. Int. J. Comput. Math. *3*, 241–252 (1977).
[9] Ortega, J. M., Rheinboldt, W. C.: Iterative solution of nonlinear equations in several variables. New York: Academic Press 1970.

Dr. G. V. Milovanović
Dr. M. S. Petković
Faculty of Electronic Engineering
Department of Mathematics
Beogradska 14
18000 Niš
Yugoslavia