View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Journal of Computational and Applied Mathematics 9 (1983) 65–69
North-Holland Publishing Company

A note on some improvements of the simultaneous methods for determination of polynomial zeros

M.S. PETKOVIĆ and G.V. MILOVANOVIĆ
Department of Mathematics, Faculty of Electronic Engineering, P.O. Box 73, 18000 Niš, Yugoslavia

Received 2 June 1982

Abstract: Applying Gauss–Seidel approach to the improvements of two simultaneous methods for finding polynomial zeros, presented in [9], two iterative methods with faster convergence are obtained. The lower bounds of the R-order of convergence for the accelerated methods are given. The improved methods and their accelerated modifications are discussed in view of the convergence order and the number of numerical operations. The considered methods are illustrated numerically in the example of an algebraic equation.

Keywords: Determination of polynomial zeros, simultaneous iterative methods, converge order.

Consider a monic polynomial of degree $n \geqslant 3$

$$P(z)=\prod_{i=1}^{n}(z-r_i),$$

with simple real or complex zeros $r_1,\dots,r_n$. Let $z_1,\dots,z_n$ be distinct reasonably close approximations of these zeros and let

$$Q(z)=\prod_{i=1}^{n}(z-z_i).$$

Then, for $z=z_k$ $(k=1,\dots,n)$, we have

$$Q'(z_k)=\prod_{\substack{i=1\\ i\neq k}}^{n}(z_k-z_i).$$

Introduce $\Delta_k=-P(z_k)/Q'(z_k)$ for abbreviation. The following iterative method of the second order for the simultaneous finding of polynomial zeros has been the subject of many papers:

$$z_k^*=z_k+\Delta_k \quad (k=1,\dots,n),$$

or, in the form

$$z_k^*=z_k-\frac{P(z_k)}{\prod_{\substack{i=1\\ i\neq k}}^{n}(z_k-z_i)} \quad (k=1,\dots,n) \tag{1}$$

where z_k^* is new approximation to the zero r_k. The iterative formula (1) is classical result introduced by Weierstrass [11, p. 258] in 1891, in connection with a proof of the fundamental theorem of algebra. Different derivations of this formula where given much later by Dočev [4], Kerner [7] and the others.

Using the approximations $z_i^*=z_i+\Delta_i$ instead of z_i $(i\neq k)$, Nourein [9] suggested the following improvement of the methods (1) (called the *improved Durand–Kerner method*):

$$\hat{z}_k=z_k-\frac{P(z_k)}{\prod_{\substack{i=1\\ i\neq k}}^{n}(z_k-z_i-\Delta_i)} \quad (k=1,\dots,n). \tag{2}$$

The convergence order of this method is three (see [9]). The price to be paid in order to achieve faster convergence consists of the increased number of numerical operations because of the additional calculations of $Q'(z_k)$ $(k=1,\dots,n)$. As a result, the iterative process (2) is relatively inefficient in practical application.

Let us put

$$\Delta=\max_{1\leqslant i\leqslant n}|\Delta_i|.$$

Assuming that Δ is small enough (in other words, all starting approximations are taken to be sufficiently close to the zeros), we shall have the development:

$$\prod_{\substack{i=1\\ i\neq k}}^{n}(z_k-z_i-\Delta_i)$$

0377-0427/83/0000–0000/$03.00 © 1983 North-Holland

$$= \prod_{\substack{i=1\\i\neq k}}^{n} (z_k - z_i)\left(1 - \frac{\Delta_i}{z_k - z_i}\right)$$

$$= Q'(z_k)\left[1 - \sum_{\substack{i=1\\i\neq k}}^{n} \frac{\Delta_i}{z_k - z_i} + O(\Delta^2)\right].$$

Taking only the linear terms of Δ_i in the above, it follows from (2) that

$$\hat{z}_k = z_k - \Delta_k\left(1 - \sum_{\substack{i=1\\i\neq k}}^{n} \frac{\Delta_i}{z_k - z_i}\right)^{-1}$$

$$(k = 1,\ldots,n). \tag{3}$$

The convergence of method (3) remains cubic (see [8]).

Efficience of the iterative process (2) can be increased, in certain degree, if calculating new approximations $\hat{z}_k$ we use already calculated approximations $\hat{z}_i$ $(i < k)$ in the same iteration (the so-called Gauss–Seidel approach). In this case we obtain the accelerated iterative process

$$\hat{z}_k = z_k - \frac{P(z_k)}{\prod_{i=1}^{k-1} (z_k - \hat{z}_i) \prod_{i=k+1}^{n} (z_k - z_i^*)}$$

$$(k = 1,\ldots,n). \tag{4}$$

Let $r = [r_1 \ldots r_n]^{\mathrm{T}}$ be the limit point (the vector of the exact zeros) of the iterative process (4). We shall now prove that the R-order of convergence of the method (4), denoted by $O_R((4), r)$ (see [10]), is at least $1 + \sigma_n$, where $\sigma_n \in (2, 3)$ is the unique positive zero of the polynomial

$$f(\sigma) = \sigma^n - \sigma - \sum_{k=0}^{n-1} \sigma^k.$$

The proof is essentially the same as in [2], and some of its steps will be omitted.

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

$$d = \min_{\substack{i,j\\i\neq j}} |r_i - r_j|,$$

$$v_k^{(m)} = z_k^{(m)} - r_k.$$

For simplicity, we shall omit the iteration index always when it cannot cause confusion. We shall write z_k and $\hat{z}_k$ instead of $z_k^{(m)}$ and $z_k^{(m+1)}$ respectively. Let

$$w_k(z) = \prod_{i=1}^{k-1} (z - \hat{z}_i) \prod_{i=k+1}^{n} (z - z_i^*),$$

$$v_k = z_k - r_k, \quad v_k^* = z_k^* - r_k, \quad \hat{v}_k = \hat{z}_k - r_k.$$

It can be proved that

$$\hat{v}_k = v_k\left(\sum_{j=1}^{k-1} \alpha_j^{(k)}\hat{v}_j + \sum_{j=k+1}^{n} \beta_j^{(k)} v_j^*\right)$$

holds, where

$$\alpha_j^{(k)} = \frac{1}{(\hat{z}_j - z_k)w_k'(\hat{z}_j)} \sum_{\substack{i=1\\i\neq k,j}}^{n} (\hat{z}_j - r_i),$$

$$\beta_j^{(k)} = \frac{1}{(z_j^* - z_k)w_k'(z_j^*)} \prod_{\substack{i=1\\i\neq k,j}}^{n} (z_j^* - r_i),$$

$$v_k^* = v_k \sum_{\substack{j=1\\j\neq k}}^{n} a_{jk} v_j,$$

$$a_{jk} = \frac{1}{z_j - z_k} \prod_{\substack{i=1\\i\neq k,j}}^{n} \frac{z_j - r_i}{z_j - z_i}.$$

Suppose that the initial conditions

$$|v_i^{(0)}| < \frac{1}{q} = \frac{d}{2n-1} \quad (i = 1,\ldots,n) \tag{5}$$

are satisfied. Put

$$h_k = q|v_k|, \quad h_k^* = q|v_k^*|, \quad \hat{h}_k = q|\hat{v}_k|.$$

Using the inequality

$$\frac{1}{2n-3}\left(1 + \frac{1}{2n-3}\right)^{n-2} \leqslant \frac{1}{n-1} \quad (n \geqslant 2),$$

we can establish

$$\begin{aligned} \hat{h}_k &\leqslant \frac{1}{n-1} h_k \left\{\sum_{i=1}^{k-1} \hat{h}_i + \sum_{i=k+1}^{n} h_i^*\right\}, \\ h_k^* &\leqslant \frac{1}{n-1} h_k \sum_{\substack{i=1\\i\neq k}}^{n} h_i. \end{aligned} \tag{6}$$

According to (5), the inequalities

$$h_k^{(0)} = q|v_k^{(0)}| < 1$$

hold for each $k = 1,\ldots,n$. If we put

$$h = \max_{1\leqslant k\leqslant n} h_k^{(0)},$$

then

$$h_k^{(0)} \leqslant h < 1 \quad (k = 1,\ldots,n).$$

Besides, we conclude that the iterative process (4) is convergent. Further, we can write

$$h_k^{(m+1)} \leqslant h^{u_k^{(m+1)}} \quad (k = 1,\ldots,n;\ m = 0, 1,\ldots).$$

Defining the matrix A by

$$A = \begin{bmatrix} 2 & 1 & & & 0 \\ 1 & 1 & 1 & & \\ \vdots & & \ddots & \ddots & \\ 1 & 0 & & 1 & 1 \\ 2 & 1 & & 0 & 1 \end{bmatrix},$$

the vectors

$$\boldsymbol{u}^{(m)} = \left[u_1^{(m)} \dots u_n^{(m)} \right]^{\mathrm{T}}$$

can be successively calculated by

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

starting with $\boldsymbol{u}^{(0)} = [1 \cdots 1]^{\mathrm{T}}$. The characteristic polynomial of A is

$$f_n(\lambda) = (\lambda - 1)^n - (\lambda - 1) - \sum_{k=0}^{n-1} (\lambda - 1)^k.$$

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

$$\tilde{f}_n(\sigma) = f_n(\sigma + 1) = \sigma^n - \sigma - \sum_{k=0}^{n-1} \sigma^k.$$

Since the spectral radius of matrix A is $\rho(A) = 1 + \sigma_n$, where $\sigma_n \in (2, 3)$ is the unique positive zero of $\tilde{f}_n(\sigma)$, we can prove, similarly as in [2], that the lower bound of the R-order of the iterative method (4) is given by

$$O_R((4), \boldsymbol{r}) \geqslant \rho(A) = 1 + \sigma_n.$$

To compare the efficience of modifications (2), (3) and (4) of Weierstrass' formula (1), with regard to the number of numerical operations, we present Table 1 for an n-th degree polynomial P with all real zeros. The number of calculations of the polynomial values $P(z_1), \dots, P(z_n)$ is the same for all methods.

From Table 1 we can infer that the improved method (2) requires two times more numerical operations than the basic method (1). Thus, the method (2) is not quite applicable. The following note confirms this conclusion: Using the same number of operations (plus the additional calculations of the values $P(\hat{z}_1), \dots, P(\hat{z}_n)$), the basic method (1) can be applied successively two times giving, in a certain sence, the process of the fourth order. From Table 1 we also conclude that the modified method (3) uses a lot of divisions. Further, many numerical examples show the similar behaviour of the methods (2) and (3) under the same conditions. Finally, besides the faster convergence, the iterative process (4) requires less numerical operations and occupies less storage space at digital computer (because of the use of the previous calculated approximations in the same iteration) in compared to the procedures (2) and (3).

Let $\delta_k = -P(z_k)/P'(z_k)$ be Newton's correction. For the simultaneous determination of all zeros of the polynomial P, the following modified Newton method is well known (see Börsch-Supan [3], Dočev and Byrnev [5], Ehrlich [6], Aberth [1]):

$$\hat{z}_k = z_k + \delta_k \left(1 + \delta_k \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{1}{z_k - z_i} \right)^{-1} \quad (k = 1, \dots, n). \tag{7}$$

For sufficiently good starting values $z_1, \dots, z_n$, the method (7) converges cubically.

Taking Newton's approximation $z_i + \delta_i$ instead of z_i $(i \neq k)$ in (7), Nourein [9] obtained the following modified method (called the *improved Ehrlich method*):

$$\hat{z}_k = z_k + \delta_k \left(1 + \delta_k \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{1}{z_k - z_i - \delta_i} \right)^{-1} \quad (k = 1, \dots, n). \tag{8}$$

The convergence order of this method is four. Contrary to the improved method (2), the acceleration of convergence of the method (8) is attained without the additional calculations since the al-

Table 1

	(1)	(2)	(3)	(4)
addition and substraction	n^2	$2n^2$	$n(3n-2)$	$2n^2 - n$
multiplication	$n(n-2)$	$2n(n\text{-}2)$	$n(n-2)$	$2n^2 - 5n + 2$
division	n	$2n$	$n(n+1)$	$2n - 1$

Table 2

	(7)	(8)	(9)
addition and substraction	$n(2n-1)$	$n(3n-2)$	$n(5n-3)/2$
multiplication	n^2	n^2	n^2
division	n	n	n

Table 3

	(2)	(4)	(8)	(9)
$z_1^{(1)}$	**0.3226**164	**0.3226**1642	**0.322546**633902	**0.322546**633902
$z_2^{(1)}$	**1.745**8725	**1.7457**8171	**1.745765**055593	**1.745760**941476
$z_3^{(1)}$	**4.5366**340	**4.53662**649	**4.536620**461362	**4.536620**265610
$z_4^{(1)}$	**9.3950**698	**9.39507**132	**9.395070**928986	**9.395070912**322

ready calculated Newton corrections δ_i (which appear in (7) too) are used. Therefore, the iterative method (8) is more suitable in practical realization relative to the simultaneous methods (2), (3) and (4).

The convergence of Nourein's modification (8) can be accelerated applying the Gauss–Seidel approach. If in determination of $\hat{z}_k$ we use the already calculated approximations $\hat{z}_i$ $(i < k)$, which are better than Newton's approximations $z_j + \delta_i$, we obtain

$$\hat{z}_k = z_k + \delta_k \gamma_k$$

where

$$\gamma_k = \left(1 + \delta_k \left(\sum_{i=1}^{k-1} \frac{1}{z_k - \hat{z}_i} + \sum_{i=k+1}^{n} \frac{1}{z_k - z_i - \delta_i} \right)\right)^{-1}$$

$$(k = 1, \ldots, n). \tag{9}$$

For this method we can derive the following relation:

$$h_k^{(m+1)} \leqslant \frac{1}{n-1} h_k^{(m)^2} \left(\sum_{i=1}^{k-1} h_i^{(m+1)} + \sum_{i=k+1}^{n} h_i^{(m)^2} \right).$$

The derivation is similar to that of the method (2) (see also [2] and will be omitted. The corresponding matrix is

$$B = \begin{bmatrix} 2 & 2 & & & 0 \\ & 2 & 2 & & \\ & & \ddots & \ddots & \\ & 0 & & 2 & 2 \\ 2 & 2 & \ldots & 0 & 2 \end{bmatrix}.$$

The spectral radius $\rho(B)$ of this matrix determines the lower bound of the R-order of the method (9). Since $\rho(B) = 2(1 + \tau_n)$, where $\tau_n \in (1, 2)$ is the unique positive root of the equation $\tau^n - \tau - 1 = 0$, we have

$$O_R((9), r) \geqslant \rho(B) = 2(1 + \tau_n) > 4.$$

The increase of the convergence order of the method (9) is obtained without the additional calculations. Moreover, this method occupies less storage space and uses less numerical operations than the method (8) (see Table 2).

By virtue of the previous, we conclude that the accelerated simultaneous method (9) is more effective comparing to the other methods considered in this paper. The following example illustrates this conclusion.

Example. Consider Laguerre's polynomial of the fourth degree

$$x^4 - 16x^3 + 72x^2 - 96x + 24.$$

Beginning with the initial approximations

$$z_1^{(0)} = 0.26, \qquad z_2^{(0)} = 1.7,$$

$$z_3^{(0)} = 4.5, \qquad z_4^{(0)} = 9.36,$$

after the first iteration we obtained the values presented in Table 3. The correct digits in the table are printed boldface.

References

[1] O. Aberth, Iteration methods for finding all zeros of a polynomial simultaneously, *Math. Computation* **27** (1973) 339–344.

[2] G. Alefeld and J. Herzberger, On the convergence speed of some algorithms for the simultaneous approximation of polynomial roots, *SIAM J. Numer. Anal.* **2** (1974) 237–243.

[3] W. Börsch-Supan, A posteriori error bounds for the zeros of polynomials, *Numer. Math.* **6** (1963) 380–398.

[4] K. Dočev, A modified Newton method for the simultaneous approximation of all roots of the given algebraic equation *Fiz.-Mat. Spis Bulgar. Akad. Nauk* **5** (1962) 136–139 (in Bulgarian).

[5] K. Dočev and P. Byrnev, Certain modifications of Newton's method for the approximate solution of algebraic equations, *Z. Vyčisl. Mat. i Fiz.* **4** (1964) 915–920.

[6] L.W. Ehrlich, A modified Newton method for polynomials, *Comm. ACM* **10** (1967) 107–108.

[7] I.O. Kerner, Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen, *Numer. Math.* **8** (1966) 290–294.

[8] A.W. Nourein, An iteration formula for simultaneous determination of the zeroes of a polynomial, *J. Comput. Appl. Math.* **4** (1975) 251–254.

[9] A.W. Nourein, An improvement on two iteration methods for simultaneous determination of the zeroes of a polynomial, *Int. Comput. Math.* **3** (1977) 241–252.

[10] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).

[11] K. Weierstrass, Neuer Beweis des Satzes, dass jede Ganze Rationale Function einer Veräderlichen dargestellt werden kann als ein Product aus Linearen Functionen derselben Veränderlichen. *Ges. Werke* Vol. 3 (1903) 251–269.