SIAM J. NUMER. ANAL.
Vol. 11, No. 2, April 1974

ON THE CONVERGENCE SPEED OF SOME ALGORITHMS FOR THE SIMULTANEOUS APPROXIMATION OF POLYNOMIAL ROOTS*

G. ALEFELD AND J. HERZBERGER†

Abstract. This note gives an analysis of the order of convergence of some modified Newton methods. The modifications we are concerned with are well-known methods—a total-step method and a single-step method—for refining all roots of an nth-degree polynomial simultaneously. It is shown that for the single-step method the R-order of convergence, used by Ortega and Rheinboldt in [6], is at least $2 + \sigma_n > 3$, where $\sigma_n > 1$ is the unique positive root of the polynomial $p_n(\sigma) = \sigma^n - \sigma - 2$.

1. Preliminaries. Suppose $f(x)$ is a polynomial of nth degree given by

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

We assume that the coefficients $a_0, a_1, \cdots, a_n$ are complex numbers and that all the roots $r_1, r_2, \cdots, r_n$ are distinct. Let

$$x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}$$

be approximations for the roots of $f(x)$. To determine the roots of $f(x)$ we consider the following methods.

1. *Total-step method* (TSM).

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f(x_i^{(k)})/f'(x_i^{(k)})}{1 - f(x_i^{(k)})/f'(x_i^{(k)}) \sum\limits_{\substack{j=1 \\ j\neq i}}^{n} \dfrac{1}{x_i^{(k)} - x_j^{(k)}}},$$

$$i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots.$$

2. *Single-step method* (SSM).

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f(x_i^{(k)})/f'(x_i^{(k)})}{1 - f(x_i^{(k)})/f'(x_i^{(k)}) \left[\sum\limits_{j=1}^{i-1} \dfrac{1}{x_i^{(k)} - x_j^{(k+1)}} + \sum\limits_{j=i+1}^{n} \dfrac{1}{x_i^{(k)} - x_j^{(k)}} \right]},$$

$$i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots.$$

(See [1], [2], [3].) For sufficiently good starting values $x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}$, it can be proved that both methods are converging.

Defining

$$\gamma_i^{(k)} = \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{(x_i^{(k)} - r_i)(r_j - x_j^{(k)})}{(x_i^{(k)} - r_j)(x_i^{(k)} - x_j^{(k)})},$$

$$i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots,$$

* Received by the editors September 8, 1972, and in revised form January 24, 1973.

† Institut für Angewandte Mathematik, University of Karlsruhe, D 75 Karlsruhe, Germany.

for (TSM) and

$$\gamma_i^{(k)} = \sum_{j=1}^{i-1} \frac{(x_i^{(k)} - r_i)(r_j - x_j^{(k+1)})}{(x_i^{(k)} - r_j)(x_i^{(k)} - x_j^{(k+1)})} + \sum_{j=i+1}^{n} \frac{(x_i^{(k)} - r_i)(r_j - x_j^{(k)})}{(x_i^{(k)} - r_j)(x_i^{(k)} - x_j^{(k)})}$$

for (SSM), both methods can be written in the form

$$x_i^{(k+1)} = r_i + \frac{\gamma_i^{(k)}}{1 + \gamma_i^{(k)}}(x_i^{(k)} - r_i), \qquad i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots .$$

If we take

$$h_i^{(k)} = x_i^{(k)} - r_i, \qquad i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots ,$$

this may be written as

$$h_i^{(k+1)} = \frac{\gamma_i^{(k)}}{1 + \gamma_i^{(k)}} h_i^{(k)}, \qquad i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots . \tag{1}$$

The following results are given in [3]:

We suppose

$$|h_i^{(k)}| \leqq \tfrac{1}{4} \min_{\substack{1 \leqq j \leqq n \\ j \neq i}} |r_i - r_j| =: d_i/4, \qquad i = 1, 2, \cdots, n,$$

and we further assume, for $i = 1, 2, \cdots, n$,

$$|r_i - x_j^{(k)}| \geqq d_i/2, \qquad j = 1, 2, \cdots, n, \quad j \neq i.$$

Then we get

$$|\gamma_i^{(k)}| \leqq \frac{16}{3d_i^2} |h_i^{(k)}| \sum_{\substack{j=1 \\ j \neq i}}^{n} |h_j^{(k)}|$$

for (TSM) and

$$|\gamma_i^{(k)}| \leqq \frac{16}{3d_i^2} |h_i^{(k)}| \left[\sum_{j=1}^{i-1} |h_j^{(k+1)}| + \sum_{j=i+1}^{n} |h_j^{(k)}| \right] \tag{2}$$

for (SSM).

In addition, if the inequalities $|h_i^{(k)}| \leqq h \leqq d_i/4$ hold, then

$$|h_i^{(k+1)}| \leqq h^3 \delta_i (1 - O(h^2 \delta_i)),$$

where

$$\delta_i = 16(n - 1)/(3d_i^2)$$

for (TSM) and

$$\delta_i = \frac{16}{3d_i^2} \left[h^2 \sum_{j=1}^{i-1} \delta_j + n - i \right]$$

for (SSM).

Hence, both methods are at least cubically convergent.

2. A lower bound for the *R*-order of SSM. The purpose of this paper is to prove that (SSM) converges faster than (TSM). To be precise, we give a lower bound greater than 3 for the order of convergence of the sequence

$$h^{(k)} = \max_{1 \leqq i \leqq n} |h_i^{(k)}|.$$

We use the following definition of the order of convergence (see [6]): Let I be an iterative process with limit point x^*. Then the quantity

$$O_R(I, x^*) = \begin{cases} \infty & \text{if } R_p(I, x^*) = 0 \text{ for all } p \in [1, \infty), \\ \inf\{p \in [1, \infty) | R_p(I, x^*) = 1\} & \text{otherwise}, \end{cases}$$

is called the *R-order of I at* x^*.

$R_p(I, x^*)$ is called the *R-factor of I at* x^* and is defined by

$$R_p(I, x^*) = \sup\{R_p\{x^{(k)}\} | \{x^{(k)}\} \in C(I, x^*)\}, \qquad 1 \leqq p < \infty,$$

where

$$R_p\{x^{(k)}\} = \begin{cases} \limsup\limits_{k \to \infty} |x^{(k)} - x^*|^{1/k} & \text{if } p = 1, \\ \limsup\limits_{k \to \infty} |x^{(k)} - x^*|^{1/p^k} & \text{if } p > 1, \end{cases}$$

and $C(I, x^*)$ is the set of all sequences generated by I and converging to x^*.

THEOREM. *Let $\sigma_n > 1$ be the unique positive root of*

$$\tilde{p}_n(\sigma) = \sigma^n - \sigma - 2 = 0.$$

Then for the R-order of (SSM) *we have*

$$O_R((\mathrm{SSM}), 0) \geqq 2 + \sigma_n.$$

Proof. Using (2), it follows from (1) that

$$\begin{aligned} |h_i^{(k+1)}| &\leqq \frac{16}{3d_i^2} \frac{1}{1 - |\gamma_i^{(k)}|} |h_i^{(k)}|^2 \left[\sum_{j=1}^{i-1} |h_j^{(k+1)}| + \sum_{j=i+1}^{n} |h_j^{(k)}| \right] \\ &\leqq c |h_i^{(k)}|^2 \left[\sum_{j=1}^{i-1} |h_j^{(k+1)}| + \sum_{j=i+1}^{n} |h_j^{(k)}| \right] \end{aligned}$$

as soon as

$$\frac{16}{3d_i^2} \frac{1}{1 - |\gamma_i^{(k)}|} \leqq c, \qquad i = 1, 2, \cdots, n.$$

We now set

$$\gamma = \sqrt{(n-1)c}, \quad |h_i^{(k)}| = (1/\gamma)\eta_i^{(k)}, \qquad i = 1, 2, \cdots, n,$$
$$\varepsilon = c/\gamma^2 = 1/(n-1),$$

and get

$$\eta_i^{(k+1)} \leqq \varepsilon(\eta_i^{(k)})^2\left[\sum_{j=1}^{i-1} \eta_j^{(k+1)} + \sum_{j=i+1}^{n} \eta_j^{(k)}\right],$$

$$i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots.$$

Because of $\lim_{k\to\infty} |h_i^{(k)}| = 0$, $i = 1, 2, \cdots, n$, we may assume

$$\eta_i^{(0)} \leqq \eta < 1, \qquad i = 1, 2, \cdots, n.$$

Then we get

$$\eta_i^{(k+1)} \leqq \eta^{m_i^{(k+1)}},$$

$$i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots.$$

Defining the matrix A by

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

the vectors $m^{(k)} = (m_i^{(k)})$ can successively be calculated by

$$m^{(k+1)} = Am^{(k)}, \qquad k = 0, 1, 2, \cdots, \tag{3}$$

with initial values $m_i^{(0)} = 1$, $i = 1, 2, \cdots, n$. The proof is by induction and will be omitted. The matrix A is nonnegative and its directed graph (see [7, p. 20]) is strongly connected, i.e., A is irreducible. By the Perron–Frobenius theorem (see [7, p. 30]) this implies that A has a positive eigenvalue λ_1 equal to its spectral radius. But, by a simple application of Theorem 2.9 of [7, p. 49], we find that A is also primitive. Thus, for the remaining eigenvalues $\lambda_2, \lambda_3, \cdots, \lambda_n$ of A, we get

$$\lambda_1 = \rho(A) > |\lambda_2| \geqq |\lambda_3| \geqq \cdots \geqq |\lambda_n|. \tag{4}$$

Let

$$A^k = (a_{ij}^{(k)}), \qquad k = 1, 2, \cdots,$$

denote the kth power of A. Since A is a primitive matrix, we get

$$A^k > 0 \quad \text{for } k \geqq k_0 \tag{5}$$

(see [7, p. 41]). For an arbitrary matrix with property (4) it can be shown that

$$\lim_{k\to\infty} \frac{a_{ij}^{(k+1)}}{a_{ij}^{(k)}} = \lambda_1.$$

(The proof may be found in [4, p. 179].) If $\varepsilon > 0$ is given, then

$$a_{ij}^{(k+1)}/a_{ij}^{(k)} \geqq \rho(A) - \varepsilon \quad \text{for } k \geqq k(\varepsilon) \geqq k_0,$$

or

$$a_{ij}^{(k+1)} \geqq \alpha[\rho(A) - \varepsilon], \qquad i, j = 1, 2, \cdots, n,$$

where

$$\alpha = \min_{1 \leqq i,j \leqq n} a_{ij}^{(k)} > 0.$$

Therefore,

$$a_{ij}^{(k+2)} \geqq a_{ij}^{(k+1)}[\rho(A) - \varepsilon] \geqq \alpha[\rho(A) - \varepsilon]^2,$$

and, in general,

$$a_{ij}^{(k+r)} \geqq \alpha[\rho(A) - \varepsilon]^r, \qquad i, j = 1, 2, \cdots, n, \quad r = 1, 2, \cdots. \tag{6}$$

Now, combining (3) and (6) into the single inequality

$$m^{(k+r)} = A^{k+r} m^{(0)} = \left(\sum_{j=1}^{n} a_{ij}^{(k+r)} \right) \geqq (n \cdot \alpha[\rho(A) - \varepsilon]^r) e,$$

where $e = (e_i)$, $e_i = 1$, $i = 1, 2, \cdots, n$, we obtain

$$\eta_i^{(k+r)} \leqq \eta^{m_i{}^{(k+r)}} \leqq \eta^{n\alpha[\rho(A)-\varepsilon]^r}, \qquad i = 1, 2, \cdots, n, \quad r = 1, 2, \cdots, \quad k \geqq k(\varepsilon) \geqq k_0,$$

or

$$|h_i^{(k+r)}| \leqq \frac{1}{\gamma} \eta^{n\alpha[\rho(A)-\varepsilon]^r}.$$

For

$$h^{(k)} = \max_{1 \leqq i \leqq n} |h_i^{(k)}|$$

we also have

$$h^{(k+r)} \leqq \frac{1}{\gamma} \eta^{\alpha n[\rho(A)-\varepsilon]^r}.$$

Thus, it follows that

$$R_{\rho(A)-\varepsilon}\{h^{(k)}\} = \limsup_{r \to \infty} [h^{(k+r)}]^{1/[\rho(A)-\varepsilon]^r} \leqq \limsup_{r \to \infty} \left[\frac{1}{\gamma} \eta^{\alpha n[\rho(A)-\varepsilon]^r} \right]^{1/[\rho(A)-\varepsilon]^r} = \eta^{\alpha n} < 1,$$

and therefore

$$O_R((\mathrm{SSM}), 0) \geqq \rho(A) - \varepsilon.$$

This inequality holds for all $\varepsilon > 0$ and we immediately have

$$O_R((\mathrm{SSM}), 0) \geqq \rho(A). \tag{7}$$

We now consider the characteristic polynomial $p_n(\lambda)$ of A:

$$p_n(\lambda) = (\lambda - 2)^n - (\lambda - 2) - 2.$$

If we set $\sigma = \lambda - 2$, then simply substituting this in the polynomial above yields

$$\tilde{p}_n(\sigma) = \sigma^n - \sigma - 2.$$

Since $\tilde{p}_n(1) = -2$ and $\tilde{p}_n(2) \geqq 0$ for $n \geqq 2$, there is a root σ_n with $1 < \sigma_n \leqq 2$, and, by Descartes' rule of signs, there can be no other positive root of $\tilde{p}_n(\sigma)$. Thus, for the spectral radius $\rho(A)$ of A we have

$$\rho(A) = 2 + \sigma_n,$$

and the combination with (7) gives

$$O_R((\mathrm{SSM}), 0) \geqq 2 + \sigma_n,$$

which completes the proof.

3. Remark. Two other methods for calculating the roots of a polynomial simultaneously are the following.

1′. *Total step-method* (TSM)′.

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f(x_i^{(k)})}{\prod_{\substack{j=1\\ j\neq i}}^{n} (x_i^{(k)} - x_j^{(k)})}, \qquad i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots.$$

2′. *Single-step method* (SSM)′.

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f(x_i^{(k)})}{\prod_{j=1}^{i-1} (x_i^{(k)} - x_j^{(k+1)}) \prod_{j=i+1}^{n} (x_i^{(k)} - x_j^{(k)})}, \qquad i = 1, 2, \cdots, n, \quad k = 0, 1, 2, \cdots.$$

Here we assume $a_n = 1$.

It is well known that the order of convergence of (TSM)′ is at least 2 (see [5]). For (SSM)′ we can, as in the proof of the above theorem, deduce the relation

$$|h_i^{(k+1)}| \leqq c|h_i^{(k)}| \left[\sum_{j=1}^{i-1} |h_j^{(k+1)}| + \sum_{j=i+1}^{n} |h_j^{(k)}| \right].$$

Thus, it follows that the R-order of convergence is at least $1 + \tau_n > 2$, where τ_n denotes the unique positive root of

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

The proof of this statement is similar to the proof given above. In this case, we only have to replace all elements in the main diagonal of A and the element in the first column and last row by 1.

REFERENCES

[1] W. BÖRSCH-SUPAN, *A posteriori error bounds for the zeros of polynomials*, Numer. Math., 5 (1963), pp. 380–398.

[2] K. DOCHEV AND P. BYRNEV, *Certain modifications of Newton's method for the approximate solution of algebraic equations*, Zh. Vychisl. Mat. i Mat. Fiz., 4 (1964), pp. 915–920; U.S.S.R. Comput. Math. and Math. Phys., 4 (1964), pp. 915–920.

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

[4] W. GRÖBNER, *Matrizenrechnung*, Bibliographisches Institut, Mannheim-Wien-Zürich, 1966.

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

[6] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[7] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.