A THIRD-ORDER ITERATIVE FACTORIZATION TECHNIQUE FOR POLYNOMIALS*

H. A. Luther and J. P. McAnally
Texas A&M University
College Station, Texas

Abstract. An iterative technique is displayed whereby factors of arbitrary degree can be found for polynomials in one variable. Convergence is shown to occur always if a certain Jacobian does not vanish and if the initial approximation to a factor is near enough to an actual factor. The process is of third order and uses the second-order partial derivatives of certain associated polynomials. Because of the special nature of these associated polynomials and the technique developed, the computation time of a single iterative step is not excessive.

1. Description of the Process

Let a polynomial of degree n in z be

$$(1) \qquad f(z) = \sum_{i=0}^{n} a_i z^{n-i}, \; a_0 = 1.$$

Let $1 \le m < n$ and let f(z) be factorizable as f(z) = s(z)t(z) where

$$(2) \qquad s(z) = \sum_{i=0}^{m} \rho_i z^{m-i}, \; t(z) = \sum_{i=0}^{n-m} \beta_i z^{n-m-i}, \; \rho_0 = \beta_0 = 1.$$

Let the superscript k refer to the kth step in an iterative process, based on an initial (somewhat arbitrary) choice of the m

* This research was supported by the National Aeronautics and Space Administration under Grant No. NGR-44-001-024.

numbers $p_1^{(1)}, p_2^{(1)}, \cdots, p_m^{(1)}$. For all values of k we consider $p_0^{(k)} = 1$. The process consists in establishing a sequence of polynomials

$$\sum_{i=0}^{m} p_i^{(k)} z^{m-i}$$

which have s(z) as a limit provided a certain Jacobian does not vanish and provided the initial estimate $\sum_{i=0}^{m} p_i^{(1)} z^{m-i}$ is "near enough" to s(z).

The process presented here is a third-order extension of a procedure previously established by Newton-Raphson iteration [1].

Starting with the column vector $P^{(k)} = [p_1^{(k)}\ p_2^{(k)}\ \cdots\ p_m^{(k)}]^t$ a new vector $P^{(k+1)}$ is built in four stages.

In the first stage, numbers $b_s^{(k)}$, $c_s^{(k)}$ and $d_s^{(k)}$ $(0 \le s \le n)$ are built by the recursion relation

$$(3) \qquad \sum_{i=0}^{m} p_i^{(k)} b_{j-i}^{(k)} = a_j,\ b_s^{(k)} = 0 \text{ if } s < 0,\ 0 \le j \le n;$$

$$(4) \qquad \sum_{i=0}^{m} p_i^{(k)} c_{j-i}^{(k)} = -b_j^{(k)},\ c_s^{(k)} = 0 \text{ if } s < 0,\ 0 \le j \le n;$$

$$(5) \qquad \sum_{i=0}^{m} p_i^{(k)} d_{j-i}^{(k)} = -2c_j^{(k)},\ d_s^{(k)} = 0 \text{ if } s < 0,\ 0 \le j \le n.$$

Observe that $c_j^{(k)}$ can be found as soon as $b_j^{(k)}$ is known and $d_j^{(k)}$ can be found as soon as $c_j^{(k)}$ is known. This means that stage one can be carried out in a single loop. (It is, however, true that $c_n^{(k)}$, $d_n^{(k)}$ and $d_{n-1}^{(k)}$ are not needed.)

Next let an mxm matrix be defined as

(6) $$J^{(k)} = [c_{n+1-i-j}^{(k)}],\ 1 \leq i,j \leq m.$$

Here i is the row index and j is the column index. The second stage consists in finding the inverse of (6), which is denoted by

(7) $$K^{(k)} = [f_{ij}^{(k)}],\ 1 \leq i,j \leq m.$$

It is observed that (6) is symmetric, therefore (7) is symmetric. It is a matter of interest that the matrix in (6) is persymmetric.

No further matrix inversions are needed; matrix multiplication suffices to finish the process. It is observed that the computations thus far (except for (5)) are those needed for the Newton-Raphson approach [1, p. 109].

For the third stage, first define the mxm persymmetric matrices (see (5))

(8) $$D_i^{(k)} = [d_{n+1-i-r-s}^{(k)}],\ 1 \leq i,r,s \leq m.$$

Here r is the row index and s is the column index.

Next define (see (3)) the column matrix

(9) $$B^{(k)} = [b_n^{(k)}\ b_{n-1}^{(k)}\ \cdots\ b_{n-m+1}^{(k)}]^t$$

and let

$$\Phi_i^{(k)} = (B^{(k)})^t K^{(k)} D_i^{(k)} K^{(k)} B^{(k)}, \quad 1 \leq i \leq m, \tag{10}$$

while

$$\Psi^{(k)} = [\Phi_1^{(k)} \Phi_2^{(k)} \cdots \Phi_m^{(k)}]^t. \tag{11}$$

This completes the third stage.

For the fourth and final stage let

$$P^{(k+1)} = P^{(k)} - K^{(k)} B^{(k)} - \frac{1}{2} K^{(k)} \Psi^{(k)}. \tag{12}$$

(Were we to use $P^{(k)} - K^{(k)}B^{(k)}$ for the right member of (12), the result would be a Newton-Raphson iteration. This feature is useful in application, since apparently the second-order technique locates <u>some</u> factor more readily.)

By the parametric use of functions, techniques can be found which influence both the rate of convergence and the character of the region from which convergence to a given factor occurs. Let $g_1(P)$ and $g_2(P)$ have derivatives of the necessary orders in the neighborhood of a solution point $\boldsymbol{\rho} = [\rho_1 \rho_2 \cdots \rho_m]^t$ $\left(\text{see } (2)\right)$. In place of (12) use

$$P^{(k+1)} = P^{(k)} - g_1(P^{(k)}) K^{(k)} B^{(k)} - \frac{1}{2} g_2(P^{(k)}) K^{(k)} \Psi^{(k)}. \tag{13}$$

Then the situation is as follows.

If $g_2(P)$ is arbitrary (probably zero is a good choice) and $g_1(\boldsymbol{\rho})$ falls between zero and two, then (13) converges linearly.

If $g_2(P)$ is arbitrary and $g_1(\mathcal{P})$ is one, then (13) has quadratic convergence.

If $g_2(\mathcal{P})$ is one while $g_1(\mathcal{P})$ is one and $\dfrac{\partial g_1(\mathcal{P})}{\partial p_i} = 0$ for $1 \le i \le m$ then (13) is a cubic iterative process.

2. Related Processes and Proof of Convergence

Let a polynomial of degree n in z be described by (1). Let

$$g(z) = \sum_{i=0}^{m} p_i z^{m-i},\ p_0 = 1;\ h(z) = \sum_{i=0}^{n-m} b_i z^{n-m-i},\ b_0 = 1.$$

For $0 \le j \le n$ let

$$\sum_{i=0}^{m} p_i b_{j-i} = a_j,\ b_s = 0 \text{ for } s < 0. \tag{14}$$

Let $r_j = \sum_{i=0}^{m-j-1} p_i b_{n-j-i}$, $0 \le j \le m - 1$. Then it is known and easily verified that

$$f(z) = g(z)h(z) + \sum_{i=0}^{m-1} r_i z^i.$$

Let $P = [p_1\ p_2\ \cdots\ p_m]^t$. Then it is clear that $b_j(P)$, $0 \le j \le m$, is a polynomial in P. It is also clear that $r_j(P)$, $0 \le j \le m - 1$, is a polynomial in P. Moreover, factorization occurs if and only if $r_j(P) = 0$ for $0 \le j \le m - 1$, or equivalently, if and only if $b_j(P) = 0$, $n - m + 1 \le j \le n$.

Lin's method and related methods [2], [3], are first-order iteration techniques based on equations $r_j(P)$. Convergence occurs only for special polynomials $f(z)$.

Bairstow's method finds quadratic factors only, and is a second-order process based on the application of Newton-Raphson to $r_0(P)$ and $r_1(P)$ [4, p. 472]. One of the present authors has extended this to factors of arbitrary order, and it is hoped to serve in conjunction with a future article.

A second-order technique for factors of arbitrary degree, based on the equations $b_j(P)$ and Newton-Raphson, is found in [1].

It is known, and from (14) it is easy to show that if, for $0 \le j \le n$,

$$\sum_{i=0}^{m} p_i c_{j-i} = -b_j, \quad c_s = 0 \text{ for } s < 0, \tag{15}$$

then $c_{j-\ell}(P) = \dfrac{\partial b_j(P)}{\partial p_\ell}$, $1 \le \ell \le m$, $0 \le j \le n$.

It may have passed unobserved, but from (15) it is easy to show that if, for $0 \le j \le n$,

$$\sum_{i=0}^{m} p_i d_{j-i} = -\,2c_j, \quad d_s = 0 \text{ if } s < 0, \tag{16}$$

then $d_{j-r-s}(P) = \dfrac{\partial^2 b_j(P)}{\partial p_r \partial p_s}$, $1 \le r,s \le m$, $0 \le j \le n$.

Indeed, still higher derivatives can be similarly found. Moreover, their computational evaluation is simple, as indicated in conjunction with relations (3), (4) and (5). Thus in this instance, the establishment of high-order iteration techniques seems more easily accomplished by use of high-order derivatives than by multipoint iteration. (In this connection see [5, p. 215].)

To establish (12) and (13) proceed as follows. Let

$$\frac{\partial B(P)}{\partial p_i} = \left[\frac{\partial}{\partial p_i} b_n(P) \; \frac{\partial}{\partial p_i} b_{n-1}(P) \; \cdots \; \frac{\partial}{\partial p_i} b_{n-m+1}(P) \right]^t$$

$$\frac{\partial^2 B(P)}{\partial p_i \partial p_j} = \left[\frac{\partial^2}{\partial p_i \partial p_j} b_n(P) \; \frac{\partial^2}{\partial p_i \partial p_j} b_{n-1}(P) \; \cdots \; \frac{\partial^2}{\partial p_i \partial p_j} b_{n-m+1}(P) \right]^t$$

$$\frac{\partial^3 B^*(P,\rho)}{\partial p_i \partial p_j \partial p_k} = \left[\frac{\partial^3}{\partial p_i \partial p_j \partial p_k} b_n(P + \theta_1 \overline{\rho - P}) \frac{\partial^3}{\partial p_i \partial p_j \partial p_k} b_{n-1}(P + \theta_2 \overline{\rho - P}) \right.$$

$$\left. \cdots \; \frac{\partial^3}{\partial p_i \partial p_j \partial p_k} b_{n-m+1}(P + \theta_m \overline{\rho - P}) \right]^t$$

Because $B(\rho) = 0$ (see 12) , by Taylor's Theorem

$$(17) \qquad -B(P) = \sum_{r=1}^{m} (\rho_r - p_r) \frac{\partial}{\partial p_r} B(P)$$

$$+ \frac{1}{2!} \sum_{r,s=1}^{m} (\rho_r - p_r)(\rho_s - p_s) \frac{\partial^2}{\partial p_r \partial p_s} B(P)$$

$$+ \frac{1}{3!} \sum_{r,s,t=1}^{m} (\rho_r - p_r)(\rho_s - p_s)(\rho_t - p_t) \frac{\partial^3}{\partial p_r \partial p_s p_t} B^*(P,\rho).$$

If we write $J(P) = \left[\frac{\partial}{\partial p_j} b_{n+1-i}(P) \right]$ where i is the row index and j is the column index (see (6)) then the first term on the right side of (17) is

(18) $J(P)(\rho - P)$.

In similar fashion, using the notation of (8), the second term of the right side of (17) is

(19) $$\frac{1}{2!} Q(P) = \frac{1}{2!} [Q_1(P) Q_2(P) \cdots Q_m(P)]^t$$

where

$$Q_i(P) = (\rho - P)^t D_i(P)(\rho - P)$$

and

$$D_i(P) = [d_{n+1-i-r-s}(P)]$$

since $d_{n+1-i-r-s}(P) = \dfrac{\partial^2 b_{n+1-i}(P)}{\partial p_r \partial p_s}$ $\Big($(see (16)$\Big)$.

Thus from (17)

(20) $$-B(P) = J(P)(\rho - P) + \frac{1}{2!} Q(P) + R_1(P)$$

where $R_1(P)$ may be viewed as a vector of cubic forms in the $\rho_r - p_r$.

Next consider $\Big($see (8), (9), (10) and (11)$\Big)$

(21) $$\Psi(P) = [\Phi_1(P) \Phi_2(P) \cdots \Phi_m(P)]^t$$

Since $B(\rho) = 0$, from Taylor's expansion

(22) $$-B(P) = J(P)(\rho - P) + R_2(P)$$

where $R_2(P)$ may be viewed as a vector of quadratic forms in the $\rho_r - p_r$.

Then since $K(P) = J^{-1}(P)$ and is symmetric,

$$(23) \qquad \Phi_i(P) = (\rho - P)^t D_i(P)(\rho - P) + R_{3i}(P)$$

where $R_{3i}(P)$ contains only third and fourth order terms in the $\rho_r - p_r$.

Thus $\left[\text{see (12)}\right]$, from (20) and (23),

$$(24) \qquad P - K(P)B(P) - \frac{1}{2!}K(P)\Psi(P) = \rho + R(P)$$

where $R(P)$ is a column vector each of whose entries contains only terms of third and fourth degree in the $\rho_r - p_r$.

If $J(\rho)$ is not singular, there exists a closed region $\mathcal{R}$ having ρ as an interior point and within which the coefficients of the terms $(\rho_r - p_r)(\rho_s - p_s)(\rho_t - p_t)$ in $R(P)$ are bounded. If, for a column vector Z with elements z_i, by $\|Z\|$ we mean $\max_i |z_i|$, it follows that there is a constant M such that for $P \varepsilon \mathcal{R}$

$$(25) \qquad \|R(P)\| \leq M\|\rho - P\|^3$$

Now let $P^{(1)}$ be so chosen that $\|\rho - P^{(1)}\| < \theta/\sqrt{M}$ where $0 < \theta < 1$ and $P^{(1)}$ lies in $\mathcal{R}$. Then inductively see $\left[(12), (24) \text{ and } (25)\right]$

$$\|P^{(k+1)} - \rho\| \leq M\|P^{(k)} - \rho\|^3 \leq \theta^{3k}/\sqrt{M}.$$

Thus $\lim_{k=\infty} P^{(k)} = \rho$ and the process is a third-order process.

The truth of the statements involving (13) follows from similar considerations.

References

1. Luther, H. A. An iterative factorization technique for polynomials. *Comm. ACM 6*, 3(March 1963), 108-110.

2. Luther, H. A. A class of iterative techniques for the factorization of polynomials. *Comm. ACM 7*, 3(March 1964), 177-179.

3. Lin, Shih-nge. A method for finding roots of algebraic equations. *J. Math. Phys. 22* (1943), 60-77.

4. Hildebrand, F. B. *Introduction to Numerical Analysis*. McGraw-Hill Book Co., Inc. 1956.

5. Traub, J. F. *Iterative Methods for the Solution of Equations.* Prentice-Hall, Inc. 1964.