CS 54

A GENERALIZED BAIRSTOW ALGORITHM

BY

G. H. GOLUB and T. N. ROBERTSON

TECHNICAL REPORT NO. 54
JANUARY 13, 1967

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

A GENERALIZED BAIRSTOW ALGORITHM

Gene H. Golub* and Thomas N. Robertson

I. Introduction: the Bairstow Process

The basic idea for finding the roots of real polynomials by finding a quadratic factor makes use of the following identity

$$(a_0x^n+a_1x^{n-1}+...+a_n) \equiv (x^2-\alpha x-\beta)(b_0x^{n-2}+b_1x^{n-3}+...+b_{n-2}) + Ax + B . \qquad (1)$$

Equating coefficients gives (with $b_{-1} = b_{-2} = 0$):

$$b_k = a_k + \alpha b_{k-1} + \beta b_{k-2} \qquad (2)$$

for k = 0,1,..., n-2 and

$$\left.\begin{aligned} A &= a_{n-1} + \alpha b_{n-2} + \beta b_{n-3} \\ B &= a_n + \beta b_{n-2} \end{aligned}\right\} \qquad (3)$$

Beginning with arbitrary α_0 and β_0, (2) and (3) can be used to define an iterative process for getting a quadratic factor. At the ith step α_i and β_i are used in (2) to provide coefficients b_k; after which, (3) is solved for α_{i+1} and β_{i+1} with A and B set to zero.

This is usually known as Lin's method [2], which was extended and studied by Friedman [3] and Luke and Ufford [4]. The convergence properties have not been fully established, but the method is often slowly convergent.

---

* The work of the first author was in part supported by NSF and ONR.

The Bairstow method [1] consists of solving the system

$$\left.\begin{aligned} A \equiv A(\alpha,\beta) &= 0 \\ B \equiv B(\alpha,\beta) &= 0 \end{aligned}\right\}$$

by Newton's process of successive approximations. The theorem of Kantorovich [5] gives conditions on A and B and on the starting values α_o, β_o which ensure convergence. The verification of these conditions is not computationally feasible, but the method is usually quadratically convergent. More precisely, if we formulate the algorithm as in [6], viz, replacing α_n and β_n by $\alpha_n + \delta$ and $\beta_n + \epsilon$, where δ and ϵ satisfy

$$\left.\begin{aligned} A + \delta A_\alpha + \epsilon A_\beta &= 0 \\ B + \delta B_\alpha + \epsilon B_\beta &= 0 \end{aligned}\right\}$$

(subscripts denote partial differentiation), then the iteration procedure is quadratically convergent if the sequences $\{\alpha_n\}$ and $\{\beta_n\}$ have limits s and t respectively, and further

$$D = \begin{vmatrix} A_\alpha & A_\beta \\ B_\alpha & B_\beta \end{vmatrix} \neq 0 \text{ at } (s,t) .$$

This criterion is applied in the generalization which follows.

II. Generalized Bairstow Algorithm

Consider (cf. identity (1))

$$
\begin{aligned}
&a_oP_n + a_1P_{n-1} + \ldots + a_nP_0 \\
&\equiv (P_2-\alpha P_1-\beta P_0)(b_oP_{n-2}+b_1P_{n-3}+\ldots+b_{n-2}P_o) + AP_1 + BP_0
\end{aligned}
\tag{4}
$$

where $P_n(x)$ are nth degree polynomials satisfying a three term recursion

$$P_{n+1} = (c_nx + d_n)P_n + e_nP_{n-1}, \quad \text{with } P_{-1} = 0,\ P_o = 1 .$$

Thus we can write

$$P_1P_k = \ell_{k+1}P_{k+1} + m_kP_k + r_{k-1}P_{k-1}$$

and

$$P_2P_k = s_{k+2}P_{k+2} + t_{k+1}P_{k+1} + u_kP_k + v_{k-1}P_{k-1} + w_{k-2}P_{k-2}$$

for appropriate ℓ, m, r, s, etc., so that equating coefficients now gives

$$
\begin{aligned}
b_k = \frac{1}{s_{n-k}} \{a_k &+ (\alpha\ell_{n-k} - t_{n-k})b_{k-1} + (\beta + \alpha m_{n-k} - u_{n-k})b_{k-2} \\
&+ (\alpha r_{n-k} - v_{n-k})b_{k-3} - w_{n-k}b_{n-4}\}
\end{aligned}
\tag{4a}
$$

(with $b_{-4} = b_{-3} = b_{-2} = b_{-1} = 0$) for $k = 0,1,\ldots, n-2$, and

$$A = a_{n-1} + \alpha b_{n-2} + (\beta + \alpha m_1 - r_1)b_{n-3} + (\alpha r_1 - v_1)b_{n-4} - w_1b_{n-5} \tag{4b}$$

$$B = a_n + \beta b_{n-2} + \alpha r_0b_{n-3} - w_0b_{n-4} .$$

Equations (4a) and (4b) can be used to compute the factor $(P_2 - \alpha P_1 - \beta P_o)$ by a natural extension of Bairstow's process. Begin by choosing starting values α_o, β_o . Having computed α_i and β_i, (4a) will provide values for the quantities b_k and their partial derivatives with respect to α and β. These in turn are used in connection with (4b) to provide values for $A, B, A_\alpha, A_\beta, B_\alpha, B_\beta$. Newton's process will yield values for δ and ϵ, from which $\alpha_{i+1}, \beta_{i+1}$ follow,

To establish the convergence properties, write (4) as follows:

$$P(x) \equiv (P_2 - \alpha P_1 - \beta P_o)Q(x) + AP_1 + BP_o \ .$$

<u>Theorem</u>. If $P_2 - sP_1 - tP_o$ is an exact factor with roots r_1 and r_2, then the convergence $\alpha \to s$ and $\beta \to t$ is quadratic if $Q(r_1)$ and $Q(r_2)$ are non-zero,

<u>Proof</u>. Differentiation of (4) w.r.t. α gives

$$0 = -P_1Q(x) + (P_2 - \alpha P_1 \beta P_o)Q_\alpha(x) + A_\alpha P_1 + B_\alpha P_o \ , \qquad (5)$$

so that evaluation at $\alpha = s$, $\beta = t$, and $x = r_1$ gives

$$A_\alpha P_1(r_1) + B_\alpha = P_1(r_1)Q(r_1) \ .$$

Similarly,

$$A_\beta P_1(r_1) + B_\beta = Q(r_1)$$

Then

$$D = \begin{vmatrix} A_\alpha & A_\beta \\ B_\alpha & B_\beta \end{vmatrix} = Q(r_1)\left(A_\alpha - A_\beta P_1(r_1)\right) \ .$$

So (i) if $r_1 \neq r_2$, evaluation of **(5)** at $x = r_2$ gives two more equations which solve to yield

$$A_\alpha = \frac{P(r_1)Q(r_1) - P(r_2)Q(r_2)}{P_1(r_1) - P_1(r_2)} ,$$

$$A_\beta = \frac{Q(r_1) - Q(r_2)}{P_1(r_1) - P_1(r_2)}$$

and thus $D = Q(r_1)Q(r_2)$;

(ii) if $r_1 = r_2 = r$, differentiation of (5) w.r.t. x gives

$$o = -P_1'Q(x) - P_1Q'(x) + (P_2 - \alpha P_1 - \beta P_o)'Q_\alpha(x)$$
$$+ (P_2 - \alpha P_1 - \beta P_o)Q_\alpha'(x) + A_\alpha P_1' ,$$

and evaluation at $\alpha = s$, $\beta = t$ and $x = r$ gives

$$A_\alpha = Q(r) + P_1(r)Q'(r)/a_1 .$$

Similarly,

$$A_\beta = Q'(r)/a_1 ,$$

so

$$D = Q^*(r) .$$

Thus in either case $D \neq 0$ if $Q(r_1)$ and $Q(r_2)$ are non-zero.

The proof (i) above is the generalization of the result given by Henrici [6] which ensures convergence of the Bairstow algorithm to a quadratic factor of a polynomial if its roots have multiplicity one. We have shown in (ii) that a root of multiplicity two can be extracted, and the procedure remains quadratically convergent.

It is interesting to experiment with the classical Bairstow method upon polynomials having repeated roots. It has been observed, for example, that if r_1 has multiplicity two and r_2 is any other root, even if an initial approximation closer to $(x - r_1)(x - r_2)$ than to $(x - r_1)^2$ is taken, the scheme "prefers" to converge to $(x - r_1)^2$, avoiding $Q(r_1) = 0$. Similarly in the generalization we can say that the extraction of "quadratic factors" $P_2 - sP_1 - t$ can be accompanied with quadratic convergence if the roots of the linear combination $\sum_{k=0}^{n} a_k P_k$ have multiplicity two or less.

III. Applications

(i) Orthogonal polynomials

Orthogonal polynomials are important in curve fitting, cf. [7]; they also play an important rôle in Gaussian quadrature. The method we have presented is applicable to finding the zeros of linear combinations of orthogonal polynomials, since such polynomials satisfy a three term recurrence relationship.

Programming experiments using an IBM 1620 tested the method on combinations of the form $\sum_{k=0}^{n} a_k T_k$, where $T_k(x)$ is the kth degree Chebyshev polynomial, and confirmed the properties of convergence to the "quadratic factors" $T_2 - sT_1 - t$. The recursion formulas

$$T_1 T_k = \tfrac{1}{2}(T_{k+1} + T_{k-1}), \quad T_2 T_k = \tfrac{1}{2}(T_{k+2} + T_{k-2})$$

for Chebyshev polynomials yield

$$b_k = 2a_k + \alpha b_{k-1} + 2\beta b_{k-2} + \alpha b_{k-3} - b_{k-4} \qquad (6)$$

with $b_{-1} = b_{-2} = b_{-3} = b_{-4} = 0$

$$A = a_{n-1} + \alpha b_{n-2} + (\beta - \tfrac{1}{2}) b_{n-3} + \frac{\alpha}{2} b_{n-4} - \tfrac{1}{2} b_{n-5}$$

$$B = a_n + \beta b_{n-2} + \frac{\alpha}{2} b_{n-3} - \tfrac{1}{2} b_{n-4} .$$

The formulas for $\frac{\partial b_k}{\partial \alpha}$, $\frac{\partial b_k}{\partial \beta}$ follow easily from (6) and apply to provide A_α, A_β, B_α, B_β .

(ii) Eigenvalue problems.

For the tridiagonal matrix

$$A = \begin{pmatrix} d_1 & u_1 & & & \mathbf{0} \\ \ell_2 & d_2 & \cdot & & \\ & \cdot & \cdot & \cdot & \\ & & & \cdot & u_{n-1} \\ \mathbf{0} & & & \ell_n & d_n \end{pmatrix} .$$

the characteristic polynomial

$$\det(xI - A) = P_n(x)$$

satisfies

$$P_0 = 1,\ P_1 = x - d_1 \text{ and } P_{k+1} = (x - d_{k+1})P_k - u_k \ell_{k+1} P_{k-1} ,$$

so that the method applies to the eigenvalue problem for arbitrary tridiagonal matrices.

Programming tests on symmetric tridiagonal matrices with known eigenvalues gave good convergence and accuracy. The eigenvalues were found in pairs, each pair being deflated out before the subsequent pair was obtained.

There have been a number of algorithms proposed to reduce an arbitrary matrix to tridiagonal form; references to these algorithms are given in [8]. The method presented, used in conjunction with such a routine, offers a contribution to the solution of the complete eigenvalue problem. In particular, when approximations to the eigenvalues are known this generalization of the Bairstow process is an efficient means of obtaining final values.

(iii) Symmetric polynomials

Consider a 2n*th* degree polynomial of the form

$$P_{2n}(z) = a_o z^{2n} + a_1 z^{2n-1} + \ldots + a_n z^n + \ldots + a_o$$
$$= a_o(z^{2n} + 1) + a_1(z^{2n-1} + z) + \ldots + a_n z^n .$$

It is easy to see that if $P_{2n}(z^*) = 0$, then $P_{2n}(1/z^*) = 0$. Now

$P_{2n}(z) = 0$ when $W(z) = \dfrac{P_{2n}(z)}{z^n} = a_o\left(\dfrac{z^n+z^{-n}}{2}\right) + a_1\left(\dfrac{z^{n-1}+z^{-(n-1)}}{2}\right) + \ldots + \dfrac{a_n}{2} = 0$.

Let us write

$$R_k(z) = \frac{z^k+z^{-k}}{2} \text{ , so that}$$

$$w(z) = a_o R_n(z) + a_1 R_{n-1}(z) + \ldots + \frac{a_n}{2} R_o(z) .$$

Note that

$$R_{k+1}(z) = 2R_1(z)R_k(z) - R_{k-1}(z) .$$

It is easy to see that the method presented here is applicable even though $R_k(z)$ is not a k*th* degree polynomial.

Computer Science Department, Stanford, California 94305

Mathematics Department, Occidental College

IV. References

1. Bairstow, L., "Investigations Relating to the Stability of the Aeroplane", R and M 154 of Advisory Committee for Aeronautics (1914).

2. Lin, Shih-nge, "A Method for Finding Roots of Algebraic Equations", J. Math. and Phys. 22:60-77 (1943).

3. Friedman, B., "Note on Approximating Complex Zeros of a Polynomial", Comm. Pure Appl. Math Z:195-208 (1949).

4. Luke, Y. L., and Ufford, D., "On the Roots of Algebraic Equations", J. Math. and Phys. 30:94-101 (1951).

5. Kantorovich, L. V., "Functional Analysis and Applied Mathematics", N.B.S. Report 1509 (translation by C. D. Benster and edited by G. E. Forsythe).

6. Henrici, Peter, "Elements of Numerical Analysis", Wiley and Sons, New York (1964).

7. Forsythe, G. E., "Generation and the use of Orthogonal Polynomials for Data Fitting with a Digital Computer" J. Soc. Indust. Appl. Math. 5:74-88 (1957).

8. Waltmann, W. L., and Lambert, R. J., "T-Algorithm for Tridiagonalization", J. Indust. Appl. Math. 13:1069-78 (1965).