

Convergence of the Fraser-Hart Algorithm for Rational Chebyshev Approximation

By Charles B. Dunham

Abstract. The Fraser-Hart variant of the Remez algorithm is used to determine the best rational Chebyshev approximation to a continuous function on an interval. A necessary and sufficient condition for the matrix of the associated linear system to be nonsingular at the solution to the approximation problem is given. It is shown that the Fraser-Hart method may fail even if started arbitrarily close to the solution of the approximation problem. Use of the secant method in place of the Fraser-Hart iteration is also considered.

1. Introduction. Let $[\alpha, \beta]$ be a finite interval. Let w be a positive continuous weight function on $[\alpha, \beta]$. Let $R_m^n[\alpha, \beta]$ be the set of ratios $r = p/q$ of polynomials p of degree at most n to polynomials q of degree at most m , $q(x) > 0$ for $\alpha \leq x \leq \beta$. The approximation problem is: given f continuous on $[\alpha, \beta]$, find $r^* \in R_m^n[\alpha, \beta]$ to minimize

$$e(r) = \max\{|w(x)(f(x) - r(x))|: \alpha \leq x \leq \beta\}.$$

Such an element r^* is called a best approximation to f .

It has been shown that a best approximation exists, has a characteristic number of alternations, and is unique [1]. One of the most popular methods of finding the best approximation is Fraser and Hart's variant of the Remez algorithm [3], [4], hereafter called the FHR algorithm. In the rational Remez algorithm, we attempt to solve the system,

$$(1) \quad f(x_i) - r(x_i) = (-1)^i \lambda / w(x_i), \quad i = 0, \dots, n + m + 1,$$

where $\alpha \leq x_0 < \dots < x_{n+m+1} \leq \beta$. From (1) we obtain

$$p(x_i) + q(x_i)\rho_i(\lambda) = 0, \quad i = 0, \dots, n + m + 1,$$

where $\rho_i(\lambda) = (-1)^i \lambda / w(x_i) - f(x_i)$. Fixing the constant term of q equal to one, we get

$$(2) \quad p(x_i) + \bar{q}(x_i)\rho_i(\lambda) + (-1)^i \lambda / w(x_i) = f(x_i), \quad i = 0, \dots, n + m + 1,$$

where \bar{q} is a polynomial of degree m with constant term zero. If we knew λ , we could evaluate $\rho_i(\lambda)$ and then (2) would be a linear system, which we could easily solve. The approach of Fraser and Hart involves making a guess λ_0 at λ and solving

$$(3) \quad p(x_i) + \bar{q}(x_i)\rho_i(\lambda_0) + (-1)^i \lambda / w(x_i) = f(x_i), \quad i = 0, \dots, n + m + 1.$$

Received June 19, 1974.

AMS (MOS) subject classifications (1970). Primary 41A50.

Key words and phrases. Chebyshev approximation, rational approximation, Fraser-Hart iteration.

Copyright © 1975, American Mathematical Society

We then set $\lambda_0 = \lambda$ and iterate until $|\lambda - \lambda_0|$ is sufficiently small.

2. Singularity of the Matrix. It would be desirable for the coefficient matrix of (3) to be nonsingular whenever $\{x_0, \dots, x_{n+m+1}\}$ is an alternant and $\lambda_0 = (f(x_0) - r^*(x_0))w(x_0)$. Unfortunately, this is not the case when $r^* = 0$, for in this case the factors $[(-1)^i \lambda_0/w(x_i) - f(x_i)]$ drop out and the matrix of (3) is singular. In particular, in the trivial case where $f = 0$ and we set $\lambda_0 = 0$, the factor $\rho_i(\lambda_0) = [(-1)^i \lambda_0/w(x_i) - f(x_i)]$ vanishes for any choice of x_i and the matrix of (3) is singular if $m > 0$. The theorem to follow shows that the matrix may be singular for degenerate r^* and possibly some nondegenerate r^* .

THEOREM 1. *Let $w(f - r^*)$ alternate $n + m + 1$ times and suppose we can write r^* as p_0/q_0 , p_0 of degree at most n , q_0 of degree at most m with constant term zero. Let $\{x_0, \dots, x_{n+m+1}\}$ be an alternant of $(f - r^*)w$ and $\lambda_0 = (f(x_0) - r^*(x_0))w(x_0)$. Then the matrix of the FHR algorithm is singular.*

Proof. We have

$$(4) \quad f(x_i) - p^*(x_i)/q^*(x_i) = (-1)^i \lambda_0/w(x_i), \quad i = 0, \dots, n + m + 1.$$

The FHR matrix corresponding to (3) is singular if there exists p of degree n , \bar{q} of degree m with constant term zero, and λ such that at least one of p, \bar{q}, λ is nonzero, but

$$(5) \quad p(x_i) + \bar{q}(x_i)[(-1)^i \lambda_0/w(x_i) - f(x_i)] + (-1)^i \lambda/w(x_i) = 0, \\ i = 0, \dots, n + m + 1.$$

We use (4) to replace $[(-1)^i \lambda_0/w(x_i) - f(x_i)]$ by $-p^*(x_i)/q^*(x_i)$, and by the hypothesis on r^* , we obtain

$$p_0(x_i) + q_0(x_i)[-p^*(x_i)/q^*(x_i)] = 0, \quad i = 0, \dots, n + m + 1,$$

and (5) is satisfied.

In case the linear system (3) uses the power basis for polynomials, degenerate r can always be written as p_0/q_0 , q_0 with constant term zero; we reduce r to lowest terms, then multiply numerator and denominator by x . In the case $0 \notin [\alpha, \beta]$, there exist pole-free nondegenerate rational functions with constant term of denominator equal to zero.

A consequence of the theorem is that if we start the FHR algorithm where it should end, namely on the alternating error extrema and with the optimal value of λ , the matrix is singular and the algorithm fails. Another consequence is that if the algorithm does converge to the best approximation, the matrices become closer and closer to singular, giving numerical problems.

In practice, the best approximation r^* to f is unlikely to be expressible in the form p_0/q_0 , q_0 with constant term zero. However, if r^* is close to such an element, the FHR matrix of (3) at the extrema of $w(f - r^*)$ with $\lambda_0 = (f(x_0) - r^*(x_0))w(x_0)$ will be near singular.

THEOREM 2. *Suppose r^* is best to f and r^* cannot be expressed as p/q , p of degree at most n , q of degree m with constant term equal to zero. Let $\{x_0, \dots, x_{n+m+1}\}$ be an alternant of $(f - r^*)w$ and $\lambda_0 = (f(x_0) - r^*(x_0))w(x_0)$. Then the*

matrix of the FHR algorithm is nonsingular at the solution to the approximation problem.

Proof. Suppose the matrix is singular, then (5) is satisfied for at least one of p, \bar{q}, λ nonzero. Using (4), we can rewrite (5) as

$$(6) \quad p(x_i) + \bar{q}(x_i)[-p^*(x_i)/q^*(x_i)] + (-1)^i \lambda/w(x_i) = 0, \quad i = 0, \dots, n + m + 1.$$

Suppose (6) is satisfied. First suppose that $\lambda = 0$, then $p(x_i)/\bar{q}(x_i) = p^*(x_i)/q^*(x_i)$. This violates our hypothesis on r^* . Next let $\lambda \neq 0$, then $p + \bar{q}(-p^*/q^*)$ has $n + m + 1$ sign changes. But $pq^* - p^*\bar{q}$ is a polynomial of degree at most $n + m$ and so we can have only $n + m$ sign changes. Hence (6) is not satisfied and the matrix is nonsingular.

3. Convergence of the Fraser-Hart-Remez Algorithm. Convergence of the rational Remez algorithm for sufficiently good starting points is guaranteed by the theory of Barrar and Loeb [2] if the best approximation is nondegenerate. Thus if the FHR algorithm solves all systems of the type (1) encountered, it has the same convergence properties. Unfortunately, this is not the case as is shown below.

We assume in this section that $\{x_0, \dots, x_{n+m+1}\}$ is fixed. Let $\bar{\lambda}$ be a number such that (3) has a nonsingular matrix when $\lambda_0 = \bar{\lambda}$. Then the same is true for λ_0 in a neighborhood of $\bar{\lambda}$. We consider only λ_0 in such a neighborhood. The solution λ of (3) depends only on λ_0 and hence is a function of λ_0 . Cramer's rule gives $\lambda(\lambda_0) = N(\lambda_0)/D(\lambda_0)$, where $D(\lambda_0)$ is the determinant of the matrix of (3) and $N(\lambda_0)$ is the determinant of the same matrix with the last column replaced by a column of values of f . Let

$$(7) \quad h(\lambda_0) = (\partial/\partial\lambda_0)\lambda(\lambda_0) = N'(\lambda_0)/D(\lambda_0) - N(\lambda_0)D'(\lambda_0)/D^2(\lambda_0)$$

and it is seen that h is a continuous function of λ_0 in a neighborhood of $\bar{\lambda}$.

As λ depends only on λ_0 for fixed $\{x_0, \dots, x_{n+m+1}\}$, the Fraser-Hart technique of solving (1) via (3) is a linear iteration, as described by Henrici in [5, Chapter 4]. Let us assume that (3) has a nonsingular matrix when $\lambda_0 = \lambda^*$, the solution of (1). Then by the argument of Henrici [5, p. 71], the (local) convergence of the Fraser-Hart iteration depends on the magnitude of $|h(\lambda^*)|$. If it is less than one, local convergence occurs; but if it is more than one, divergence must occur. An example where $|h(\lambda^*)|$ is large is now given.

Let $n = 0, m = 1$, then we have 3 points $\{x_0, x_1, x_2\}$. Let $x_0 = 0$ and $w = 1$. Then by Cramer's rule we have

$$(8) \quad \lambda(\lambda_0) = \frac{N(\lambda_0)}{D(\lambda_0)} = \frac{\begin{vmatrix} 1 & 0 & f_0 \\ 1 & \rho_1(\lambda_0)x_1 & f_1 \\ 1 & \rho_2(\lambda_0)x_2 & f_2 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 1 \\ 1 & \rho_1(\lambda_0)x_1 & -1 \\ 1 & \rho_2(\lambda_0)x_2 & 1 \end{vmatrix}}$$

if $D(\lambda_0) \neq 0$. $D(\lambda^*)$ is the determinant of the matrix of (3) with $\lambda_0 = \lambda^*$. If it is zero, the Fraser-Hart method of solving (1) will have a singular matrix in solving (3) with $\lambda_0 = \lambda^*$ and will fail.

Let f be a nonzero approximant, then the optimal λ is $\lambda^* = 0$. By the second

theorem, $D(\lambda^*) \neq 0$. As $\lambda(\lambda^*) = \lambda^*$, we must have by Cramer's rule $N(\lambda^*) = 0$. We have by (7)

$$h(\lambda^*) = N'(\lambda^*)/D(\lambda^*).$$

Expanding the determinant of $N(\lambda_0)$, we get

$$N(\lambda_0) = f_2\rho_1(\lambda_0)x_1 - f_1\rho_2(\lambda_0)x_2 + f_0(\rho_2(\lambda_0)x_2 - \rho_1(\lambda_0)x_1);$$

and since $(\partial/\partial\lambda_0)\rho_i(\lambda_0) = (-1)^i$, we have

$$N'(\lambda_0) = x_1f_2 - f_1x_2 + f_0(x_2 + x_1).$$

It is clear that we can make N' as large as we like by choosing f to be a rational large at x_2 and small at x_0 and x_1 . Hence $N'(0)$ can be made as large as we like, and hence $h(0)$ can be made arbitrarily large.

In the case that the best rational approximation is a polynomial p of exact degree n and $\{x_0, \dots, x_{n+m+1}\}$ is an alternant of $w(f-p)$, the optimal λ is obtained after solving (3) *once*. To see this, observe that the right column of the numerator determinant (as in (8)) is of the form

$$f(x_i) = p(x_i) + (-1)^i\lambda^*/w(x_i).$$

By adding multiples of the columns corresponding to the polynomial basis to the right-hand column, we can change this column to $(-1)^i\lambda^*/w(x_i)$ and the value of the determinant is not changed. But this determinant is just λ^* times the denominator determinant and so we get $\lambda = \lambda^*$ regardless of the value of λ_0 .

This suggests that if the denominator of the best rational approximation is near constant, the convergence of the Fraser-Hart iteration for λ^* will be rapid. We recall that the case of nonconvergence of the previous section involved a denominator which was not near constant.

We have seen that for fixed $\{x_0, \dots, x_{n+m+1}\}$, the λ obtained by solving (3) depends only on λ_0 and is, therefore, a function of λ_0 . Define $F(\lambda_0) = \lambda(\lambda_0) - \lambda_0$. For λ^* a solution of (1) we have $F(\lambda^*) = 0$. It seems, therefore, that we can use any method of solving $F(\lambda_0) = 0$ to try to get λ^* . Ralston [7, p. 274] uses the secant method to solve $F(\lambda_0) = 0$. The secant method has superlinear convergence and converges given a sufficiently good starting point if $F'(\lambda_0) = h(\lambda_0) - 1$ is continuous in a neighborhood of λ^* . This happens if the matrix of (3) is nonsingular with $\lambda_0 = \lambda^*$; a sufficient condition for this is that the solution r to (1) satisfies the condition on r^* in Theorem 2.

Whereas the secant method has superlinear convergence, the previously cited analysis of Henrici [5, p. 71] suggests that in general, the Fraser-Hart technique of solving (1) via (3) has only linear convergence with convergence factor $|h(\lambda^*)|$. Experiments with programs using the Fraser-Hart technique appear to confirm this expectation. Theory thus appears to favor the secant method over the Fraser-Hart iteration, as the secant method has a faster rate of convergence and converges if the starting point is close enough.

1. N. I. ACHIESER (AHIEZER), *Lectures On the Theory of Approximation*, OGIZ, Moscow, 1947; English transl., Ungar, New York, 1956. MR 10, 33; 20 #1872.
2. R. BARRAR & H. LOEB, "On the Remez algorithm for nonlinear families," *Numer Math.*, v. 15, 1970, pp. 382-391.
3. W. J. CODY, W. FRASER & J. F. HART, "Rational Chebyshev approximation using linear equations," *Numer. Math.*, v. 12, 1968, pp. 242-251.
4. W. FRASER & J. F. HART, "On the computation of rational approximations to continuous functions," *Comm. ACM*, v. 5, 1962, pp. 401-403.
5. P. HENRICI, *Elements of Numerical Analysis*, Wiley, New York, 1964. MR 29 #4173.
6. A. RALSTON, "Rational Chebyshev approximation by Remes' algorithms," *Numer. Math.*, v. 7, 1965, pp. 322-330. MR 32 #4438.
7. A. RALSTON, "Rational Chebyshev approximation," A. RALSTON & H. WILF, (Editors), *Mathematical Methods for Digital Computers*. Vol. 2, Wiley, New York, 1967, pp. 264-284. MR 35 #2516.