ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 7, Number 4, Fall 1977

MINIMAL DEGREE RATIONAL APPROXIMATION

F. D. K. ROBERTS*

ABSTRACT. An algorithm is presented for computing a rational approximation to a function defined on a finite set of points. The method requires the accuracy of the approximation to be specified, and computes the least degree rational approximation which achieves this accuracy. Although rational approximating functions are nonlinear, the procedure is based upon determining the feasibility of a sequence of linear inequalities, and this is accomplished by the simplex method of linear programming. Some numerical results obtained with the algorithm are presented.

1. Introduction. In an earlier paper [8], the problem of determining a minimal degree *linear* approximation to a function defined on a finite set of points was discussed, and an algorithm presented for its solution. In this paper, the minimal degree problem is considered for approximation by *rational* functions.

Given a real valued function f defined on a finite point set $X = \{x_1, x_2, \cdots, x_N\}$, and given functions $\phi_1, \phi_2, \cdots, \phi_m, \psi_1, \psi_2, \cdots, \psi_n$ also defined on X, we form a rational approximating function

$$(1) \qquad R = P_m/Q_n = \sum_{i=1}^{m} p_i\phi_i \Big/ \sum_{j=1}^{n} q_j\psi_j.$$

A normalization condition is required on the coefficients p_i, q_j in the rational function. In this paper we impose the normalization

$$(2) \qquad \max\{|p_i|, |q_j|\} = 1.$$

The classical Chebyshev approximation problem is to determine the rational function P_m^*/Q_n^* which minimizes the expression

$$(3) \qquad \max_X |f - R| = \|f - R\|_\infty.$$

subject to the condition

$$Q_n > 0 \text{ on } X.$$

Received by the editors on July 6, 1976, and in revised form on September 3, 1976.

*Work done while visiting the Department of Mathematics, University of Auckland, New Zealand.

The rational approximating function (1) is nonlinear in the parameters q_j, and the Chebyshev problem requires an iterative (non-finite) algorithm for its solution. Many such algorithms have been proposed. A numerical comparison of these methods is given by Lee and Roberts [5]. Traditionally the Remes algorithm (see for example Rice [7], Ralston [6]) has been a popular method, although more recently the differential correction algorithm (see for example Barrodale, Powell and Roberts [1], Cheney and Loeb [2], Kaufman and Taylor [4]) has received considerable attention in the literature.

In Chebyshev approximation, the degrees of numerator and denominator are specified in advance, and the accuracy, i.e., the minimum value of expression (3), is known only after the best approximation has been computed. However, in many practical problems, the precise form of the rational approximating function is unimportant. All that is necessary is that a specified accuracy ϵ be achieved. In this paper we consider the *minimal degree* problem in which the accuracy ϵ required of the approximation is specified in advance, and the least degree rational approximation which achieves this accuracy is to be determined. Specifically, the problem may be stated as follows:

Given a positive tolerance ϵ, determine integers m and n and a rational function $R = P_m/Q_n$ for which the inequalities

$$\max_X |f - R| = \|f - R\|_\infty < \epsilon,$$

$$Q_n > 0 \text{ on } X,$$

are satisfied with $m + n$ as small as possible, and $m - 1 \leqq n \leqq m$. This latter inequality restricts the rational approximation to be of the form P_m/Q_m, or P_{m+1}/Q_m.

Although the rational approximating function is nonlinear, the minimal degree problem can be formulated in terms of systems of linear inequalities, and these can be solved by the finite techniques of linear programming.

A discussion of the formulation is given in the next section, and computational details concerning the application of the simplex method to the solution of the linear programming problem are presented in the 3rd section. Section 4 contains some sample numerical results obtained with the algorithm, and section 5 contains some final comments.

2. **Formulation as a Linear Program.** Perhaps the obvious approach to solving the minimal degree problem is to compute the best Chebyshev approximation to f by the rational functions P_1/Q_1, P_2/Q_1, P_2/Q_2,

$P_3/Q_2, \cdots$, until the desired accuracy ϵ is achieved. The disadvantage of this approach is that each of these problems is nonlinear and requires an iterative (non-finite) algorithm for its solution. However, a solution to the problem can be obtained using a finite linear approach.

The method is based upon determining the feasibility (consistency) of the constraints

$$\text{(4)} \qquad \left.\begin{array}{r} |f - P_m/Q_n| < \epsilon \\ Q_n > 0 \end{array}\right\} \quad \text{for all points of } X$$
$$\max\{|p_i|, |q_j|\} = 1,$$

for increasing values of m and n ($m - 1 \leqq n \leqq m$). These constraints may be written in the form

$$\text{(5)} \qquad \left.\begin{array}{r} (f + \epsilon)Q_n - P_m > 0 \\ (-f + \epsilon)Q_n + P_m > 0 \\ Q_n > 0 \end{array}\right\} \text{for all points of } X$$
$$\max\{|p_i|, |q_j|\} = 1.$$

Note that since the accuracy ϵ is specified in advance, these inequalities are *linear* in the coefficients of P_m and Q_n. The feasibility may thus be determined by solving the linear programming problem:

Minimize w subject to

$$\text{(6)} \qquad \left.\begin{array}{r} (f + \epsilon)Q_n - P_m + w \geqq 0 \\ (-f + \epsilon)Q_n + P_m + w \geqq 0 \\ Q_n + w \geqq 0 \end{array}\right\} \text{for all points of } X$$
$$-1 \leqq p_i, q_j \leqq 1.$$

The last inequalities are sufficient to ensure the normalization (2). Clearly $P_m \equiv Q_n \equiv 0$ provides a feasible solution to (6). If the minimum value w^* is zero, then no feasible solution to (4) exists, i.e., an approximation of degree m over n is not sufficiently powerful to achieve the accuracy ϵ. If the minimum value w^* is less than zero, then the solution P_m/Q_n to (6) provides an approximation to f to within accuracy ϵ.

The following result enables us to exclude the nonnegativity constraints on Q_n, and thus reduce the size of the constraint matrix of the linear program.

THEOREM. *If the minimum value w^* of the following linear program is negative, then the nonnegativity constraints $Q_n > 0$ on X are auto-*

matically satisfied:

Minimize w subject to

$$(7)\qquad \left.\begin{aligned} (f+\epsilon)Q_n - P_m + w &\geqq 0\\ (-f+\epsilon)Q_n + P_m + w &\geqq 0 \end{aligned}\right\} \textit{ for all points of } X.$$
$$-1 \leqq p_i, q_j \leqq 1.$$

Proof. The constraints in (7) may be rewritten in the form

$$w + \epsilon Q_n \geqq |fQ_n - P_m| \quad \text{for all points of } X.$$

Since the right hand side is nonnegative, $w^* < 0$ implies $Q_n > 0$.

3. **Computational Details.** A minimal degree approximation can be obtained by solving (7) by the simplex method for increasing values of m and n ($m - 1 \leqq n \leqq m$), until the optimum value w^* is negative. If a solution to (7) does not provide a minimal degree approximation for certain values of m and n, then it is necessary to re-solve the problem with m (or n) increased by 1. This new linear program differs from the previous one only in the addition of a new variable p_{m+1} (or q_{n+1}). However, it is not necessary to solve this new problem completely from the beginning. There is a standard post-optimality technique in linear programming which enables an additional variable to be included into an optimal simplex tableau (see for example Hadley [3]). The simplex iterations resume from this point, and a substantial reduction in computation can be achieved.

In practice it is more efficient to deal with the dual linear program. Letting $\phi_{i,j} \equiv \phi_i(x_j)$, $\psi_{i,j} \equiv \psi_i(x_j)$, this may be stated as:

$$\text{Maximize} \sum_{i=1}^{m+n} - u_i - v_i \text{ subject to}$$

$$\sum_{j=1}^{N} s_j(f_j+\epsilon)\psi_{i,j} + \sum_{j=1}^{N} t_j(-f_j+\epsilon)\psi_{i,j} + u_i - v_i = 0, i = 1, \cdots, n,$$

$$(8)\qquad \sum_{j=1}^{N} - s_j\phi_{i,j} + \sum_{j=1}^{N} t_j\phi_{i,j} + u_{i+n} - v_{i+n} = 0, i = 1, \cdots, m,$$

$$\sum_{j=1}^{N} s_j + \sum_{j=1}^{N} t_j = 1,$$

$$s_j, t_j, u_i, v_i \geqq 0.$$

The addition of a new variable in the primal corresponds to the addition of a new constraint in the dual. Thus the introduction of the new variable p_{m+1} (say) is equivalent to the addition of the new constraint

$$(9) \qquad \sum_{j=1}^{N} - s_j\phi_{m+1,j} + \sum_{j=1}^{N} t_j\phi_{m+1,j} + u_{m+n+1} - v_{m+n+1} = 0$$

in the dual. This constraint may be added to an optimal simplex tableau to (8) by eliminating the coefficients of the basic vectors (see for example Hadley [3]). The variable u_{m+n+1} or v_{m+n+1} provides an additional basic vector, and thus no artificial vector is required. Further simplex iterations are then performed until an optimum solution to the new linear program is obtained. The addition of a new variable q_{n+1} is treated in a similar fashion.

The algorithm proceeds then by first solving (8) for $m = n = 1$. The vectors u_1 and u_2 form part of the initial basis, and so only one artificial variable is required. This is pivoted out of the basis at the first iteration and thereafter normal simplex iterations are performed until an optimal solution is obtained. If the value of the objective function is negative, the algorithm stops and the approximation P_1/Q_1 produced is a minimal degree approximation. If the value of the objective function is zero, a new constraint (9) corresponding to the additional variable p_2 is included, and the simplex iterations resume. If the value of the objective function is still zero, a new constraint corresponding to the variable q_2 is included. The procedure continues alternately adding constraints corresponding to additional variables in numerator and denominator until an optimal solution to (8) is obtained with a negative value of the objective function. At this point the algorithm terminates, and the rational approximation P_m/Q_n so produced solves the minimal degree problem.

4. **Numerical Results.** The algorithm was coded in FORTRAN, and tested on a Burroughs B6700 computer using both single and double precision arithmetic (about 11 and 23 digits respectively). The method was tested on a variety of different problems. Listed below are some typical results obtained with the algorithm.

Table I lists eight standard functions which were approximated by rational functions of the form

$$R(x) = \frac{\sum_{i=1}^{m} p_i x^{i-1}}{\sum_{j=1}^{n} q_j x^{j-1}}$$

In each case 101 equally spaced points in the interval [0, 1] were used, and the accuracy ϵ required of the approximation was chosed as 10^{-6}. The degree of the rational approximation required to achieve this accuracy is given in the table, together with the error achieved by this minimal degree approximation.

The approximation obtained by the algorithm is not the best Chebyshev approximation in the sense of minimizing expression (3). All that is required of the minimal degree approximation is that the accuracy ϵ be achieved. If the minimal degree approximation with least error is required, then this can be obtained for example by the differential correction algorithm (see for example Kaufman and Taylor [4]) using as a starting point the approximation obtained by the minimal degree algorithm. Since the minimal degree approximation is usually close to the best Chebyshev approximation, the convergence of this nonlinear algorithm is typically quite rapid. The error of this best minimal degree approximation is also included in the table.

The same set of functions were approximated by trigonometric rationals of the form

$$R(x) = \frac{p_1 + p_2 \cos\left(\frac{\pi x}{2}\right) + p_3 \sin\left(\frac{\pi x}{2}\right) + p_4 \cos\left(\frac{2\pi x}{2}\right) + \cdots}{q_1 + q_2 \cos\left(\frac{\pi x}{2}\right) + q_3 \sin\left(\frac{\pi x}{2}\right) + q_4 \cos\left(\frac{2\pi x}{2}\right) + \cdots}.$$

The results for this case are also given in table I.

One of the advantages of a discrete approximation algorithm is that it can be used to compute multi-dimensional approximations. Since the approximation is obtained on only a finite set of points, the dimension of the problem presents no difficulty. As an example of a two-dimensional approximation, the functions $e^{xy/4}$, $\sin(x + y/2)$, $e^{-(x^2+y^2)/2}$ were approximated on the 11×11 grid $x = 0(.1)1$, $y = 0(.1)1$ by the function

$$R(x, y) = \frac{p_1 + p_2 x + p_3 y + p_4 x^2 + p_5 xy + \cdots}{q_1 + q_2 x + q_3 y + q_4 x^2 + q_5 xy + \cdots}.$$

The accuracy of the approximation was specified as 10^{-3}. The algorithm produced approximation of degrees (5,4), (6,5), (7,6) and errors .665150(−3), .746464(−3), .655787(−3), respectively. The errors of the best minimal degree approximations computed by the differential correction algorithm are .640776(−3), .712340(−3), .514046(−3).

Table I. Sample results obtained with the algorithm

Function $F(x)$	Polynomial rational approximation			Trigonometric rational approximation		
	Minimal degree (m, n)	Error of minimal degree approximation	Error of best minimal degree approximation	Minimal degree (m, n)	Error of minimal degree approxmiamtion	Error of best minimal degree approximation
$(1 + x)^{1/2}$	(3, 3)	.788257(−6)	.716218(−6)	(4, 4)	.447982(−6)	.400783(−6)
$\sin(\pi x/2)$	(5, 4)	.176377(−6)	.415115(−7)	(3, 2)	0	0
e^x	(4, 3)	.329174(−6)	.112018(−6)	(5, 4)	.528322(−6)	.109934(−6)
$\log(1 + x)$	(4,3)	.565583(−6)	.888585(−7)	(4,4)	.715657(−6)	.680203(−6)
$\sinh(x)$	(4, 3)	.379963(−6)	.364797(−6)	(4, 4)	.686327(−6)	.646034(−6)
$\Gamma(2 + x)$	(4, 4)	.175174(−6)	.102236(−6)	(4,4)	.749728(−6)	.677444(−6)
$\mathrm{erf}(x)$	(5, 4)	.207914(−6)	.721495(−7)	(5, 4)	.385888(−6)	.252509(−6)
$e^{-x^2/2}$	(4, 4)	.422584(−6)	.387183(−6)	(4, 4)	.405673(−6)	.102532(−6)

5. **Comments.** We conclude by pointing out four advantages of minimal degree rational approximations over best Chebyshev approximations:

1. In many practical problems, the accuracy of the approximation is specified in advance rather than the precise degree of the rational approximating function. Thus minimal degree techniques are more appropriate in these cases.
2. A solution to the minimal degree problem is guaranteed to exist. The Chebyshev problem may have no solution.
3. Degeneracy and near degeneracy can cause serious problems when computing Chebyshev approximations. These are not problems with minimal degree approximations.
4. The Chebyshev approximation problem is *nonlinear* and requires an iterative (non-finite) algorithm for its solution. The minimal degree rational problem is *linear* and can be solved by the finite techniques of linear programming.

ACKNOWLEDGEMENT. The author wishes to thank the National Research Council of Canada for financial support provided by N.R.C. grant A7143.

REFERENCES

1. I. Barrodale, M. J. D. Powell, and F. D. K. Roberts, *The Differential Correction Algorithm for rational ℓ_∞ approximation,* SIAM J. Numer. Anal. **9** (1972), 493–504.

2. E. W. Cheney and H. L. Loeb, *Two new algorithms for rational approximation,* Numer. Math. 3 (1961), 72–75.

3. G. Hadley, *Linear Programming,* Addison-Wesley, Reading, Mass.: 1962.

4. E. H. Kaufman, Jr. and G. D. Taylor, *Uniform rational approximation of functions of several variables,* Int. J. Num. Meth. Engng. **9** (1975), 297–323.

5. C. M. Lee and F. D. K. Roberts, *A comparison of algorithms for rational ℓ_∞ approximation,* Math. Comp. **27** (1973), 111–121.

6. A. Ralston, *A First Course in Numerical Analysis,* McGraw-Hill, New York: 1965.

7. J. R. Rice, *The Approximation of Functions. Vol. 2: Nonlinear and Multivariate Theory,* Addison-Wesley, Reading, Mass.: 1969.

8. F. D. K. Roberts, *An algorithm for minimal degree linear Chebyshev approximation on a discrete set,* Int. J. Num. Meth. Engng. **10** (1976), 619–635.

UNIVERSITY OF VICTORIA, VICTORIA, CANADA