JOURNAL OF APPROXIMATION THEORY 32, 271–280 (1981)

Conditions for a Rational Approximation to Be Locally Best in the l_1 Sense

W. FRASER*

Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada

AND

JERRY M. WOLFE

Department of Mathematics, University of Oregon, Eugene, Oregon 97403

Communicated by E. W. Cheney

Received June 27, 1980

Denote by $R_{pq}[\alpha,\beta]$, or R_{pq} when context permits, the class of rational functions $r(x)$ with numerators $p(x)$ of degree not exceeding p and denominators $q(x)$ of degree not exceeding q, and such that $q(x)\neq 0$ for $x\in[\alpha,\beta]$. That is,

$$R_{pq}[\alpha,\beta]=\left\{\frac{p(x)}{q(x)}:p(x)=\sum_{i=0}^{p}a_ix^i,\ q(x)=\sum_{i=0}^{q}b_ix^i,\ q(x)>0 \text{ on } [\alpha,\beta]\right\} \tag{1}$$

Let $\{x_i\}_{i=1}^{m}$ be a set of points belonging to $[\alpha,\beta]$ such that

$$a\leqslant x_1<\cdots<x_m\leqslant\beta. \tag{2}$$

If an approximation $r(x)$ to $f(x)$ provides a local minimum of the sum

$$\sum_{i=1}^{m}|E(x_i)|=\sum_{i=1}^{m}|r(x_i)-f(x_i)| \tag{3}$$

call it a best local l_1 approximation to f over the set $\{x_i\}$.

Problems of existence, uniqueness and degeneracy are discussed in Refs. [2–7] and in the papers to which they refer. It is known that best approximations may not exist, and when they do exist there may be minima which are local rather than global. One point of contrast between l_∞

* Deceased.

0021-9045/81/080271–10$02.00/0

640/32/4-2

approximations and l_1 approximations is that best l_∞ approximations can have degeneracies of high order, whereas although a best l_1 approximation can be degenerate, the degeneracy can be at most of order 1, and can only occur if $r(x_1)=f(x_1)$ and $r(x_m)=f(x_m)$. Accordingly degeneracy is not considered a serious problem and throughout this article attention is paid exclusively to the nondegenerate case.

Discrete l_1 problems present special difficulties when one seeks to characterize (local) best approximations. The error functional will typically not be differentiable at a local best approximation so that standard results from the calculus are not directly applicable. Also, the powerful characterization theorems often available for the l_∞ case do not have counterparts for the l_1 problem. Nevertheless, using one-sided derivative techniques it is possible to develop necessary conditions for a local best approximation and also to develop sufficient conditions for a local best approximation. The conditions developed below are well known in the linear case [8] but have not been used in practice due to the apparent complexity of the conditions themselves and to the efficiency of linear programming methods for linear l_1 problems.

Our principal aim in this paper is to show how the conditions given later can be applied in a reasonably simple way in the case of rational approximation (although other nonlinear families could also be considered). The following elementary lemma is the basis of the analysis.

LEMMA. *Let $\Phi: S\subset R^n \to R$, S open, be such that there is a closed ball $B\subset S$ of radius $r_0>0$ centered at $x_0\in S$ such that* (a) $\lim_{\lambda\downarrow 0} (\Phi(x+\lambda u)-\Phi(x))/\lambda \equiv \Phi'_+(x,u)$ *exists for all $x\in B$, $u\in R^n$.* (b) *The map $D(x,u,\lambda)\equiv(\Phi(x+\lambda u)-\Phi(x))/\lambda$ is bounded on $B\times\mu\times[0,r_0]$ and is jointly continuous in u and λ for each fixed $x\in B$, where $\mu=\{u\in R^n \mid \|u\|=1\}$ and $D(x,u,\lambda)$ is defined by* (a) *if $\lambda=0$.*

Suppose $\Phi'_+(x_0,u)>0$ for each $u\in\mu$. Then x_0 is an isolated local minimum of Φ. Conversely, if x_0 is a local minimum of Φ, then $\Phi'_+(x_0,u)\geqslant 0$ for all $u\in R^n$.

Proof. By (b), $\inf_{\|u\|=1}\{\Phi'_+(x_0,u)\}\equiv m>0$. Now suppose that x_0 is not an isolated local minimum of Φ. Then there exists a sequence $\{x_\nu\}\to x_0$ such that $\Phi(x_\nu)\leqslant\Phi(x_0)$. Let $u_\nu=(x_\nu-x_0)/\|x_\nu-x_0\|$, $\lambda_\nu=\|x_\nu-x_0\|$. Then $D(x_0,\lambda_\nu,u_\nu)\leqslant 0$ and we can assume without loss of generality that $u_\nu\to u^*\in\mu$. Then

$$0\geqslant\frac{\Phi(x_0+\lambda_\nu u_\nu)-\Phi(x_0+\lambda_\nu u^*)}{\lambda_\nu}+\frac{\Phi(x_0+\lambda_\nu u^*)-\Phi(x_0)}{\lambda_\nu}$$

$$\geqslant\frac{m}{2}+\frac{\Phi(x_0+\lambda_\nu u_\nu)-\Phi(x_0+\lambda_\nu u^*)}{\lambda_\nu}$$

$$= \frac{m}{2} + \frac{\Phi(x_0 + \lambda_\nu u^* + \lambda_\nu(u_\nu - u^*)) - \Phi(x_0 + \lambda_\nu u^*)}{\lambda_\nu}$$

$$= \frac{m}{2} + D(\bar{x}_\nu, v_\nu, \sigma_\nu)\,\|u_\nu - u^*\|$$

for all ν sufficiently large, where

$$\bar{x}_\nu = x_0 + \lambda_\nu u^*, \qquad v_\nu = \frac{u_\nu - u^*}{\|u_\nu - u^*\|}, \qquad \sigma_\nu = \lambda_\nu \|u_\nu - u^*\|.$$

Thus, $-m/2\,\|u_\nu - u^*\| \geqslant D(x_\nu, v_\nu, \sigma_\nu)$ for all ν sufficiently large which contradicts (b). The fact that $\Phi'_+(x_0, u) \geqslant 0$ is necessary for x_0 to be a local minimum of Φ is clear. ∎

To apply the lemma to the rational l_1 approximation problem, let $\Phi(a, b) \equiv \sum_{i=1}^m |(p(a, x_i)/q(b, x_i)) - f(x_i)|$, where

$$p(a, x) = a_0 + a_1 x + \cdots + a_p x^p \qquad a = (a_0, \ldots, a_p),$$
$$q(b, x) = 1 + b_1 x + \cdots + b_q x^q \qquad b = (b_1, \ldots, b_q).$$

Then a simple calculation shows that

$$\Phi'_+(a, b, u)$$
$$= \sum_{i \in N} \operatorname{sgn}(E(a, b, x_i)) \frac{[p(v, x_i)\, q(b, x_i) - (q(w, x_i) - 1)\, p(a, x_i)]}{q^2(b, x_i)}$$
$$+ \sum_{i \in Z} \frac{|p(v, x_i)\, q(b, x_i) - (q(w, x_i) - 1)\, p(b, x_i)|}{q^2(b, x_i)}$$

where $u = (v_0, \ldots, v_p, w_1, \ldots, w_q)$, $v = (v_0, \ldots, v_p)$, $w = (w_1, \ldots, w_q)$, $E(a, b, x) = (p(a, x)/q(b, x)) - f(x)$, $Z = \{j \in \{1, \ldots, m\}: E(a, b, x_j) = 0\}$ and $N = Z^c$.

Note that the normalization $b_0 = 1$ has been made in defining $q(b, x)$. Without this or some other normalization, the condition $\Phi'_+(a, b, u) > 0$ for $u \neq 0$ would be impossible to satisfy, even in the nondegenerate case. We now have the following theorem.

THEOREM. *In the notation above, the point* $(a, b) = (a_0, a_1, \ldots, a_p, b_1 \cdots b_q)$ *is a local minimum of the function* $\Phi(c, d) = \sum_{i=1}^m |(p(c, x_i)/q(d, x_i)) - f(x_i)|$ *if*

$$\left| \sum_{i \in N} \operatorname{sgn} \frac{(E(a,b,x_i))[p(v,x_i)\,q(b,x_i) - (q(w,x_i)-1)\,p(a,x_i)]}{q^2(b,x_i)} \right| < \sum_{i \in Z} \frac{|p(v,x_i)\,q(b,x_i) - (q(w,x_i)-1)\,p(a,x_i)|}{q^2(b,x_i)} \tag{5}$$

holds for all $v \in R^{p+1}$, $w \in R^q$, $(v, w) \neq (0, 0)$.

Conversely, if (a, b) *is a local minimum of* Φ, *then*

$$\left| \sum_{i \in N} \frac{\operatorname{sgn}(E(a,b,x_i))[p(v,x_i)\,q(b,x_i) - (q(w,x_i)-1)\,p(a,x_i)]}{q^2(b,x_i)} \right| \leqslant \sum_{i \in Z} \frac{|p(v,x_i)\,q(b,x_i) - (q(w,x_i)-1)\,p(a,x_i)|}{q^2(b,x_i)} \tag{6}$$

holds for all $v \in R^{p+1}$, $w \in R^q$.

Proof. Using (4) it is simple to check that the conditions of the lemma are satisfied. Then (5) and (6) represent the application of the lemma taking into account the fact that when (v, w) is replaced by $(-v, -w)$ in (4), the sign of the first sum in (4) changes, while that of the second sum does not.

Remark. It is not difficult to see that (5) cannot hold unless the rational function $r(x) = p(a, x)/q(b, x)$ is nondegenerate. It is also easy to see that the theorem is valid if any of the other coefficients of the denominator are normalized to one. In practice, normalizing b_0 or b_q would occur most frequently.

To simplify notation we shall use the abbreviation h_i for $h(x_i)$, $i = 1 \cdots m$, where h is an arbitrary real valued function on $\{x_1, ..., x_m\}$. Moreover, if $r(x) = p(x)/q(x)$ is an arbitrary member of R_{pq} let $E_i = r(x_i) - f(x_i)$ and $\sigma_i = \operatorname{sgn}(E_i)$, $i = 1, 2, ..., m$.

In order to apply the theorem first consider the matrix

$$A = \begin{pmatrix} q_1 & x_1 q_1 & \cdots & x_1^p q_1 & -p_1 & -x_1 p_1 & \cdots & -x_1^q p_1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ q_m & x_m q_m & \cdots & x_m^p q_m & -p_m & -x_m p_m & \cdots & -x_m^q p_m \end{pmatrix}, \tag{7}$$

where $r(x) = p(x)/q(x)$ is given and let A^l, $l = 0, 1, ..., q$ denote the matrices formed by deleting the $p + l + 1$st column of A. (The matrix A^l thus corresponds to the normalization $b_l = 1$ in the denominator of $r(x)$.) It follows in a straightforward way by use of elementary column transformations that if $r(x)$ is nondegenerate then A and each A^l $(l = 0, ..., q)$ have rank $p + q + 1$ and this is true of any submatrix using at least $p + q + 1$

rows of A or A^l ($l = 0,..., q$). Denote the rows of A by R_i, $i = 1,..., m$ and let R_i^l, $i = 1,..., m$ denote the rows of A^l, $l = 0,..., q$. When context permits, we suppress the superscript l when referring to the rows of A^l. Now assume l is fixed and define $\tilde{c}$ by

$$\tilde{c} = (\tilde{a}_0,..., \tilde{a}_p, \tilde{b}_0,..., \tilde{b}_{l-1}, \tilde{b}_{l+1},..., \tilde{b}_q)^T. \tag{8}$$

This definition sets up a one-to-one relationship between rational functions $\tilde{r}(x)$ (with coefficient $\tilde{b}_l = 1$) and vectors $\tilde{c}$. From (7) it follows that

$$(R_i^l, \tilde{c}) = q_i\tilde{p}_i - p_i(q_i - x_i^l), \qquad i = 1,..., m, \tag{9}$$

where $(.,.)$ is the usual inner product on R^N, $N = p + q + 1$.

Suppose the rational function $r(x)$ (with $b_l = 1$) interpolates f in k points $x_{i_1},..., x_{i_k}$, where $k \leqslant p + q + 1$. Consider first the case $k < p + q + 1$. Let B be the submatrix of A^l consisting of the rows $R_{i_1},..., R_{i_k}$ and suppose the rows are renumbered if necessary so that they are respectively the rows $i = 1,..., k$. The index set Z will be $\{1, 2,..., k\}$ and N will consist of the indices $\{k + 1,..., m\}$. Determine a set of vectors $u_1,..., u_k$ by the relations

$$(R_j, u_i) = \delta_{ij} \qquad \text{(Kronecker delta with } i, j = 1,..., k). \tag{10}$$

The system $Bx = 0$ has $p + q + 1 - k$ independent solutions, say, $u_{k+1},..., u_{p+q+1}$. The set $\{u_1,..., u_k, u_{k+1},..., u_{p+q+1}\}$ is independent and hence forms a basis for R^{p+q+1}. Thus the polynomial pair $(\tilde{p}, \tilde{q})$ (where $\tilde{r}(x) = \tilde{p}(x)/\tilde{q}(x) \in R_{pq}$ has $\tilde{b}_l = 1$) corresponds to a vector $\tilde{c}$ in R^{p+q+1} which can be expressed uniquely in the form

$$\tilde{c} = \gamma_1 u_1 + \cdots + \gamma_{p+q+1} u_{p+q+1}. \tag{11}$$

Let $S = \{k + 1,..., p + q + 1\}$.

CLAIM. *If $r(x)$ minimizes $\sum_i |E_i|$ then*

$$\left| \sum_{j=k+1}^{m} \frac{\sigma_j(R_j, u_i)}{q_j^2} \right| \leqslant \frac{1}{q_i^2} = \left| \frac{(R_i, u_i)}{q_i^2} \right| \qquad \textit{for each } i \in Z \tag{12}$$

and

$$\left| \sum_{j=k+1}^{m} \frac{\sigma_j(R_j, u_i)}{q_j^2} \right| = 0 \qquad \textit{for each } i \in S. \tag{13}$$

On the other hand, if $r(x)$ satisfies (12) *and* (13) *with strict inequality in* (12), *it is a strict local minimum of $\sum_i |E_i|$.*

Proof. We will prove sufficiency since the same techniques will also give the necessity part of the claim. Thus assume (12) and (13) hold with strict inequality in (12). We will show that (5) of the theorem is satisfied. Thus, from (9) we have

$$\begin{aligned}
&\left|\sum_{j=k+1}^{m} \frac{\sigma_j(\tilde{p}_j q_j - (\tilde{q}_j - x_j^l) p_j}{q_j^2}\right| \\
&\quad = \left|\sum_{j=k+1}^{m} \frac{\sigma_j(R_j, c)}{q_j^2}\right| \\
&\quad = \left|\sum_{j=k+1}^{m} \frac{\sigma_j}{q_j^2} \sum_{i=1}^{p+q+1} \gamma_i(R_j, u_i)\right| = \left|\sum_{i=1}^{p+q+1} \gamma_i \sum_{j=k+1}^{m} \frac{\sigma_j(R_j, u_i)}{q_j^2}\right| \\
&\quad \leqslant \sum_{i=1}^{p+q+1} |\gamma_i| \left|\sum_{j=k+1}^{m} \frac{\sigma_j(R_j, u_i)}{q_j^2}\right| = \sum_{i \in Z} |\gamma_i| \left|\sum_{j=k+1}^{m} \frac{\sigma_j(R_j, u_i)}{q_j^2}\right| \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by (13))} \\
&\quad < \sum_{i \in Z} |\gamma_i| \left|\frac{(R_i, u_i)}{q_i^2}\right| \text{(by (12))} = \sum_{i \in Z} \left|\frac{(R_i, \tilde{c})}{q_i^2}\right| \\
&\quad = \sum_{i \in Z} \left|\frac{\tilde{p}_i q_i - (\tilde{q}_i - x_i^l) p_i}{q_i^2}\right|
\end{aligned}$$

and so (5) holds. Thus by the theorem $r(x) = p(x)/q(x)$ is locally uniquely best and the sufficiency part of the claim is proved. ∎

It is not difficult to see that in fact (12) and (13) with strict inequality in (12) are equivalent to (5). Thus (12) and (13) form a computational approach to (5). It is also easy to check that if (12) fails (strictly) for some i, then r is not a locally best approximation. If $k = p + q + 1$ we can proceed as above, the only difference being that the set S is empty so condition (13) is not used. Since the case $k = p + q + 1$ is the most frequently encountered, we shall examine it in more detail with the aim of simplifying the computational procedure even further.

Thus, assume that $r(x_i) = f_i$, $i = 1,..., p + q + 1$ and that $r(x)$ is nondegenerate. Since the matrix A has rank $p + q + 1$ (as does any submatrix consisting of $p + q + 1$ rows) it follows that all rows $R_{p+q+2},..., R_m$ of A can be uniquely expressed as linear combinations of $R_1,..., R_{p+q+1}$. (Since r is nondegenerate, the same statement is valid for any appropriate A^l with rows $R_1^l,..., R_m^l$.) In particular, there are uniquely determined constants $\lambda_1,..., \lambda_{p+q+1}$ such that

$$\sum_{j=p+q+2}^{m} \left(\frac{\sigma_j}{q_j^2}\right) R_j = \left(\frac{\lambda_1}{q_1^2}\right) R_1 + \cdots + \left(\frac{\lambda_{p+q+1}}{q_{p+q+1}^2}\right) R_{p+q+1}. \tag{14}$$

CLAIM. *If $|\lambda_i| < 1$, $i = 1,\dots, p+q+1$ then $r(x)$ is a local best l_1 approximation to f. If $|\lambda_i| > 1$ for some i, then $r(x)$ is not a local best approximation to f.*

Proof. Pick an $l \in \{0, 1,\dots, q\}$ such that the coefficient b_l of $q(x)$ is nonzero. Then, without loss of generality, $b_l = 1$. Consider the corresponding matrix A^l. Then (14) holds with each R_i replaced by R_i^l and hence for an arbitrary nonzero vector $\tilde{c}$ in R^{p+q+1} we have

$$\sum_{i=1}^{p+q+1} \left| \frac{\tilde{p}_i q_i - (\tilde{q}_i - x_i^l) p_i}{q_i^2} \right|$$

$$= \sum_{i=1}^{p+q+1} \left| \frac{(R_i^l, \tilde{c}}{q_i^2} \right| > \sum_{i=1}^{p+q+1} |\lambda_i| \left| \frac{(R_i^l, \tilde{c})}{q_i^2} \right|$$

$$\geqslant \left| \left(\sum_{i=1}^{p+q+1} \frac{\lambda_i R_i^l}{q_i^2}, \tilde{c} \right) \right| = \left| \left(\sum_{i=p+q+2}^{m} \frac{\sigma_j R_j^l}{q_j^2}, \tilde{c} \right) \right|$$

$$= \left| \sum_{i=p+q+2}^{m} \left(\frac{\sigma_j R_j^l}{q_j^2}, \tilde{c} \right) \right| = \left| \sum_{i=p+q+2}^{m} \sigma_j \left[\frac{\tilde{p}_i q_i - (\tilde{q}_i - x_i^l) p_i}{q_i^2} \right] \right|$$

and since $\tilde{c}$ is arbitrary we see that (5) of the theorem holds and $r(x)$ is locally best. If $|\lambda_i| > 1$ for some i, then it follows immediately by letting $\tilde{c} = u_i$ that (12) fails and so r is not locally best. ∎

Remark. If $|\lambda_i| \leqslant 1$, $i = 1,\dots, p+q+1$ but $\lambda_j = 1$ for some j, then terms of higher order in the expansion of $\sum |E_i|$ could be used to help decide if $r(x)$ is locally best.

Applications of this last test were made in the case of two approximations discussed in the article by Barrodale and Mason [1]. The first of these to be discussed here is their Example B3(b) which can be stated:

Given the function

$$\begin{aligned} f(x) &= 1, && x = 0, 0.2, 0.4 \\ &= 0, && x = 0.5 \\ &= -1, && x = 0.6, 0.8, 0.10, \end{aligned}$$

find a best l_1 approximation in the class R_{22}. The algorithm being tested by Barrodale and Mason produced an approximation which could be written in the form

$$r(x) = \frac{p_2(x)}{q_2(x)} = \frac{12 - 29x + 10x^2}{12 - 41x + 40x^2}.$$

First construct the table:

x	$q(x)$	$p(x)$	$r(x)$	$f(x)$	$r(x)-f(x)$
0	12	12	1	1	0
0.2	27/5	33/5	33/27	1	2/9 $\operatorname{sgn}(r-f)=+1$
0.4	2	2	1	1	0
0.5	3/2	0	0	0	0
0.6	9/5	−9/5	−1	−1	0
0.8	24/5	−24/5	−1	−1	0
1.0	11	−7	−7/11	−1	4/11 $\operatorname{sgn}(r-f)=+1$

The matrix A of this article is given by

$$A = \begin{bmatrix} 12 & 0 & 0 & -12 & 0 & 0 \\ 27/5 & 27/25 & 27/125 & -33/5 & -33/25 & -33/125 \\ 2 & 4/5 & 8/25 & -2 & -4/5 & -8/25 \\ 3/2 & 3/4 & 3/8 & 0 & 0 & 0 \\ 9/5 & 27/25 & 81/125 & 9/5 & 27/25 & 81/125 \\ 24/5 & 96/25 & 384/125 & 24/5 & 96/25 & 384/125 \\ 11 & 11 & 11 & 7 & 7 & 7 \end{bmatrix}$$

Multiplying row A_j by $1/q_j^2$, and rearranging rows so that those which correspond to points of interpolation occur first, the resulting matrix is

$$\begin{bmatrix} 1/12 & 0 & 0 & -1/12 & 0 & 0 \\ 1/2 & 1/5 & 2/25 & -1/2 & -1/5 & -2/25 \\ 2/3 & 1/3 & 1/6 & 0 & 0 & 0 \\ 5/9 & 1/3 & 1/5 & 5/9 & 1/3 & 1/5 \\ 5/24 & 1/6 & 2/15 & 5/24 & 1/6 & 2/15 \\ 5/27 & 1/27 & 1/135 & -55/243 & -11/243 & -11/1215 \\ 1/11 & 1/11 & 1/11 & 7/121 & 7/121 & 7/121 \end{bmatrix}$$

When cleared of fractions on the left side, the equations to be solved for $\lambda_1,\ldots,\lambda_5$ become

$$\begin{aligned} 6\lambda_1 + 36\lambda_2 + 48\lambda_3 + 40\lambda_4 + 15\lambda_5 &= 19.878788, \\ 6\lambda_2 + 10\lambda_3 + 10\lambda_4 + 5\lambda_5 &= 3.838384, \\ 12\lambda_2 + 25\lambda_3 + 30\lambda_4 + 20\lambda_5 &= 14.747475, \\ -6\lambda_1 - 36\lambda_2 \qquad + 40\lambda_4 + 15\lambda_5 &= -12.131007, \\ -\ 6\lambda_2 \qquad + 10\lambda_4 + 5\lambda_5 &= 0.377512, \\ -\ 6\lambda_2 \qquad + 15\lambda_4 + 10\lambda_5 &= 3.659831. \end{aligned}$$

The solution of this system is

$$\lambda_1 = 1.038262,$$
$$\lambda_2 = 0.204062,$$
$$\lambda_3 = 0.101213,$$
$$\lambda_4 = -0.336088,$$
$$\lambda_5 = 0.992552.$$

Since $|\lambda_1| > 1$, the function $r(x)$ being tested is not a local minimum of $\sum |r(x_i) - f(x_i)|$. This conclusion is in agreement with the findings of Barrodale and Mason.

The second approximation to be considered is Example A1 which is:

Given the function $f(x) = e^x$ at 21 equally spaced points on the interval $[-1, 1]$, i.e., the points $-1, -0.9, -0.8, \ldots, 0.8, 0.9, 1$, find a best l_1 approximation in the class R_{22}.

The result found by Barrodale and Mason was the rational function

$$r(x) = \frac{p_2(x)}{q_2(x)} = \frac{p_0 + p_1 x + p_2 x^2}{1 + q_1 x + q_2 x^2} = \frac{1.00006 + 0.50876x + 0.08603x^2}{1 - 0.49103x + 0.07780x^2}.$$

The results of calculations stated below were obtained by use of the University of Guelph computing facility. By making the assumption that the function to be tested was an interpolant of e^x at $x_1 = -0.9$, $x_2 = -0.4$, $x_3 = 0.2$, $x_4 = 0.7$, and $x_5 = 0.9$, the rational function obtained was

$$r(x) = \frac{1.000061 + 0.5087647x + 0.08603436x^2}{1 - 0.4910249x + 0.07779443x^2}.$$

The matrix whose rows are respectively

$$\frac{1}{[q_2(x_1)]^2} A_1, \ldots, \quad \frac{1}{[q_2(x_5)]^2} A_5, \quad \sum_{i=6}^{21} \frac{\sigma_i}{[q_2(x_i)]^2} A_i$$

was calculated to be

0.6644805	−0.5980324	0.5382290	−0.2701575	0.2431417	−0.2188275
0.8272289	−0.3308915	0.1323566	−0.5545085	0.2218033	−0.0887213
1.1050850	0.2210172	0.0442034	−1.3497530	−0.2699507	−0.0539901
1.4400880	1.0080610	0.7056431	−2.8999770	−2.0299840	−1.4209890
1.6100690	1.4490620	1.3041550	−3.9601260	−3.5641120	−3.2077000
−2.7748270	−1.9488850	−2.3410450	6.1793010	5.2366760	4.6271480

Solution of the first five equations of the system

$$\begin{pmatrix} 0.6644805 & 0.8272289 & 1.1050850 & 1.4400880 & 1.6100690 \\ -0.5980324 & -0.3308915 & 0.2210172 & 1.0080610 & 1.4490620 \\ 0.5382290 & 0.1323566 & 0.0442034 & 0.7056431 & 1.3041550 \\ -0.2701575 & -0.5545085 & -1.3497530 & -2.8999770 & -3.9601260 \\ 0.2431417 & 0.2218033 & -0.2699507 & -2.0299840 & -3.5641120 \\ -0.2188275 & -0.0887213 & -0.0539901 & -1.4209890 & -3.2077000 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \end{pmatrix} \begin{matrix} = -2.7748270 \\ = -1.9488850 \\ = -2.3410450 \\ = 6.1793010 \\ = 5.2366760 \\ = 4.6271480 \end{matrix}$$

gave

$$\begin{aligned} \lambda_1 &= -0.85476, \\ \lambda_2 &= 0.44233, \\ \lambda_3 &= 0.36591, \\ \lambda_4 &= -0.98901, \\ \lambda_5 &= -0.96447. \end{aligned}$$

According to the test the approximation found is a local minimum since each $|\lambda_i| < 1$, $i = 1,..., 5$, and this agrees with the surmise of Barrodale and Mason.

The matrix which enters the calculation can in some cases be poorly conditioned, and in order to determine the numbers $\{\lambda_i\}$ as accurately as possibly it may be necessary to use double or higher order precision, and also to use a polynomial base other than the power polynomials, for example, the Chebyshev polynomials.

References

1. I. Barrodale and J. C. Mason, Two simple algorithms for discrete rational approximation, *Math. Comp.* **24** (1970), 877–891.
2. D. Braess, On a paper of C. B. Dunham concerning degeneracy in mean nonlinear approximation, *J. Approx. Theory* **9** (1973), 313–317.
3. E. W. Cheney and A. A. Goldstein, Mean square approximation by generalized rational functions, *Math. Z.* **95** (1967), 232–241.
4. C. B. Dunham, Degeneracy in mean rational approximation, *J. Approx. Theory* **4** (1971), 225–229.
5. C. B. Dunham, Best mean rational approximation, *Computing* **9** (1972), 87–93.
6. G. Lamprecht, Zur Mehrdeutigkeit bei der Approximation in der L_p-Norm mit Hilfe Rationaler Funktionen, *Computing* **5** (1970), 349–355.
7. J. M. Wolfe, Nonlinear L_1 approximation of smooth functions, *J. Approx. Theory* **17** (1976), 166–176.
8. J. R. Rice, "The Approximation of Functions: Vol. I," Addison–Wesley, Reading, Mass., 1964.