On the Computation of L_1 Approximations by Exponentials, Rationals, and Other Functions

By John R. Rice

1. Introduction. This paper is concerned with the following problem for the closed interval I.

L_1 APPROXIMATION PROBLEM. *Given $f(x)$ and an approximating function $F(A, x)$ depending on n parameters $A = (a_1, a_2, \cdots, a_n)$ in E_n, determine $A^* \in E_n$ so that*

$$\int_I |f(x) - F(A^*, x)| \, dx \leqq \int_I |f(x) - F(A, x)| \, dx = \| f - F(A) \|$$

for all $A \in E_n$.

If A^* is a solution of this problem then $F(A^*)$ is said to be a *best approximation* to f. Recently [3] a characterization theorem for best approximations has been given with mild assumptions on $F(A)$. These assumptions are

(A1) *The set*

$$K = \{(A, d) \mid (A, d) \in E_{n+1}, \| f - F(A) \| \leqq d\} \tag{1}$$

is convex.

(A2) *Set*

$$\nabla F(A) = \left(\frac{\partial F(A)}{\partial a_1}, \cdots, \frac{\partial F(A)}{\partial a_n} \right) \tag{2}$$

then we have, with the dot product notation

$$F(A + tB) = F(A) + tB \cdot \nabla F(A) + to(A, tB) \tag{3}$$

where $o(t) = o(A, tB)$ satisfies $\operatorname{Lim}_{t \to 0} o(t) = 0$.

It appears that the first assumption is essential in the sense that for any nonlinear $F(A)$, there exists an f so that K is not convex. This is in contrast to the situation where F depends linearly on the parameters. It appears, on the other hand, that for "most" $f(x)$ the set K is indeed convex, or at least $\| f - F(A) \|$ has no local minima other than global minima.

We have

THEOREM [3]. *With assumptions* A1 *and* A2, *a necessary and sufficient condition for $F(A^*)$ to be a best approximation to $f(x)$ is that*

$$\left| \int_I A \cdot \nabla F(A^*, x) \operatorname{sgn} [f(x) - F(A^*, x)] \, dx \right| \leqq \int_{Z(A^*)} | A \cdot \nabla F(A^*, x) | \, dx \tag{4}$$

where $Z(A^*) = \{x \mid f(x) = F(A^*, x)\}$.

This result is valid for integrable functions, however in this paper we are concerned only with bounded and continuous functions.

Normally one has that the measure of $Z(A^*)$ is zero and, in this case, condition (4) becomes

Received September 16, 1963.

$$(5) \qquad \int \frac{\partial F(A^*, x)}{\partial a_i} \operatorname{sgn}[f(x) - F(A^*, x)]\, dx = 0 \qquad i = 1, 2, \cdots, n.$$

Now if we knew sgn $[f(x) - F(A^*, x)]$ we could determine $F(A^*, x)$ by interpolating $f(x)$ at the sign changes of this sign function. It is known [4] that for any set $g_i(x)$, $i = 1, 2, \cdots, n$ of integrable functions there is a sign function $s(x)$ *with n or less sign changes* so that

$$(6) \qquad \int g_i(x) s(x)\, dx = 0 \qquad i = 1, 2, \cdots, n.$$

In particular, there is a sign function $s(A, x)$ for each set $\{g_i(x) = \partial F(A, x)/\partial a_i\}$. Thus there is a set of canonical points (the points of sign change of $s(A, x)$) for each A.

In this paper we explore the possibility of using these canonical points to construct best approximations. The idea is to find A_0 so that $F(A_0, x)$ interpolates $f(x)$ at the canonical points associated with A_0. If such an A_0 is found and if, *in addition*, $f(x) - F(A_0, s)$ has no other sign changes *and* if $f(x) - F(A_0, x)$ is not zero on a set of positive measure then the above theorem implies that $F(A_0, x)$ is a best approximation to $f(x)$. This scheme will fail in those cases where $F(A^*, x)$ interpolates $f(x)$ more than n times or $F(A^*, x) = f(x)$ on a set of positive measure. The first class of cases is infrequent and the second class is very rare. Thus one can hope that this idea will lead to the construction of best approximations in most instances. This hope is strengthened by experimental results.

In the remainder of the paper are presented the canonical points for several nonlinear approximating functions. Since these points depend on the parameters A, these results are presented in the form of curves. (All of the cases considered but two are nonlinear in essentially only one parameter and hence curves may be presented. In the other two cases no systematic presentation is made of these points.) The nonlinear approximating functions considered are

$ae^{tx} + b$	$\log(a + bx)$
Exponential sums	$(a + bx)^3$(a unisolvent function)
$\dfrac{a + bx}{1 + cx}$	$(a + bx^2)^2$
$\dfrac{\text{Polynomial}}{1 + cx}$	$a^2 + b^2x + 2abx^2$.

Some of these approximating functions are of considerable interest and others are included simply to explore the nature of the problem.

In addition to the canonical points, some best approximations have been constructed for three of these approximating functions. They are $ae^{tx} + b$, $(a + bx)/(1 + cx)$ and $a^2 + b^2 + 2abx^2$. These approximations were determined by the following simple scheme:

(i) guess at the parameters of a best approximation,

(ii) determine the canonical points for these parameters,

(iii) interpolate $f(x)$ at these canonical points,

(iv) use the parameters obtained by interpolation as the next guess at the parameters of a best approximation and return to (ii).

This scheme worked in almost all cases and in those where it failed it is probably due to the lack of attention to details and exceptional situations in the computer program. No special effort was made in the case of a failure. There were no failures due to the lack of applicability of condition (5).

2. The Approximating Functions. The information for each approximating function is given in the following format:

definition
equations for canonical points
list of functions approximated (if applicable)
comments
graph of canonical points.

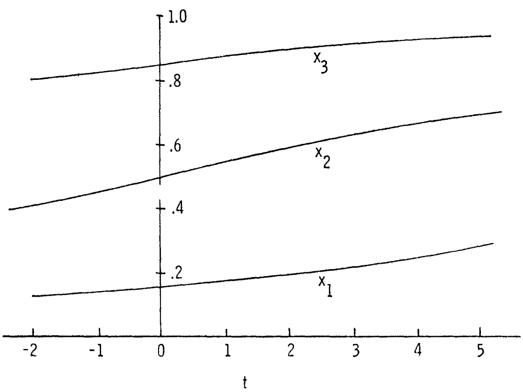
1. $F(A, x) = ae^{tx} + b, x \in [0, 1]$.

There are three canonical points, x_1, x_2 and x_3. Set $y_1 = tx_1 - t/2, y_2 = tx_2 - t/2$ and $y_3 = tx_3 - t/2$. Then we have

$$y_1 - y_2 + y_3 = 0, \qquad e^{y_1} - e^{y_2} + e^{y_3} = \cosh\left(\frac{t}{2}\right),$$
$$y_1 e^{y_1} - y_2 e^{y_2} + y_3 e^{y_3} = \frac{t}{2}\sinh\left(\frac{t}{2}\right).$$

Some functions approximated on [0, 1] are $\sin x$, $\tan x$, x^2, e^{-x^2}, $1/(1 + 25x^2)$, $\ln(1 + x)$, xe^{-x} and $x \sin x + 1 + 2x^2$.

Comments. The equations for the canonical points are readily solved by Newton's method. The interpolation problem for this function is non-trivial and a good method is given in [1]. The curves for the canonical points are symmetric about the point $t = 0, x = .5$.



2. $F(A, x) = \sum_{i=1}^{n} a_i e^{t_i x}, x \in [0, 1]$.

There are $2n$ canonical points x_i, $i = 1, 2, \cdots, 2n$ which satisfy the equations

$$\sum_{j=1}^{2n} (-1)^{j+1} e^{t_i x_j} = \frac{1 - e^{t_i}}{2} \qquad i = 1, 2, \cdots, n$$

$$\sum_{j=1}^{2n} (-1)^{j+1} t_i x_j e^{t_i x_j} = \frac{t_i e^{t_i}}{2} \qquad i = 1, 2, \cdots, n.$$

Comments. Several methods were explored for solving this system of equations and Newton's method was the most successful. Even with Newton's method there are three distinct sources of difficulty. They are i) two or more values of t_i close together, ii) very large values of t_i, and iii) many equations. Newton's method does not converge, for example, when $n = 1$, $t_1 = -20$; $n = 3$, $t_1 = 1.0$, $t_2 = 1.1$ and $t_3 = 1.3$.

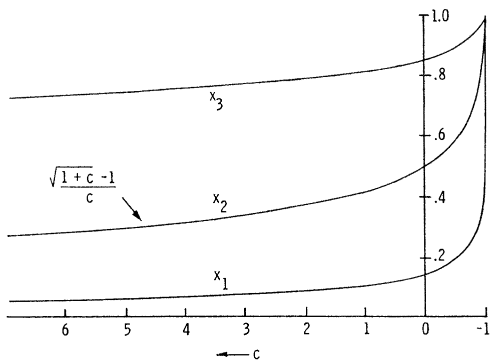
3. $F(A, x) = (a + bx)/(1 + cx), x \in [0, 1], 1 + cx > 0$.

There are three canonical points x_1, x_2 and x_3. Set $y_1 = 1 + cx_1$, $y_2 = 1 + cx_2$ and $y_3 = 1 + cx_3$. The equations for the canonical points then take the form

$$x_2 = \frac{\sqrt{1+c} - 1}{c}, \qquad y_1 y_3 = y_2 \sqrt{1+c}, \qquad y_1 + y_3 = y_2 + 1 + \frac{c}{2}.$$

Some functions approximated on [0, 1] are $\sin x$, $\tan x$, x^2, e^{-x^2}, $1/(1 + 25x^2)$, $\ln (1 + x)$, xe^{-x} and $x \sin x + 1 + 2x^2$.

Comments. The equations for the canonical points may be solved directly.



4. $F(A, x) = \sum_{j=0}^{n} a_j x^j/(1 + cx), x \in [0, 1], 1 + cx > 0$.

We only write the equations which determine the $n + 2$ canonical points. Set $y_i = 1 + cx_i$, where x_i is the ith canonical point. The equations are:

$$\sum_{j=1}^{n} (-1)^{j+1} y_j^p = \frac{1}{2}[1 + (-1)^{n+1}(1 + c)^p] \qquad p = -1, 1, 2, 3, \cdots, n,$$

$$\prod_{j \text{ odd}} y_j = [\sqrt{1+c}]^{(-1)^{n+1}} \prod_{j \text{ even}} y_j.$$

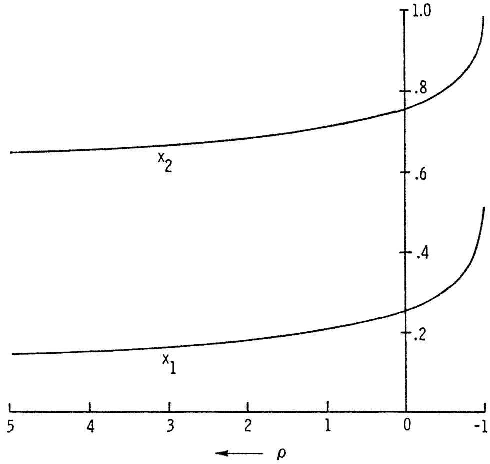
5. $F(A, x) = \ln (a + bx) = \ln \alpha(1 + \rho x), x \in [0, 1], 1 + \rho x > 0$.

The canonical points are given by

$$x_1 = x_2 - \frac{1}{2}, \qquad x_2 = \frac{\sqrt{1+\rho}}{2(\sqrt{1+\rho}-1)} - \frac{1}{\rho}.$$

For small ρ, we have x_2 given approximately by

$$x_2 = \frac{3 - 2.5\rho}{4 - 3\rho}.$$

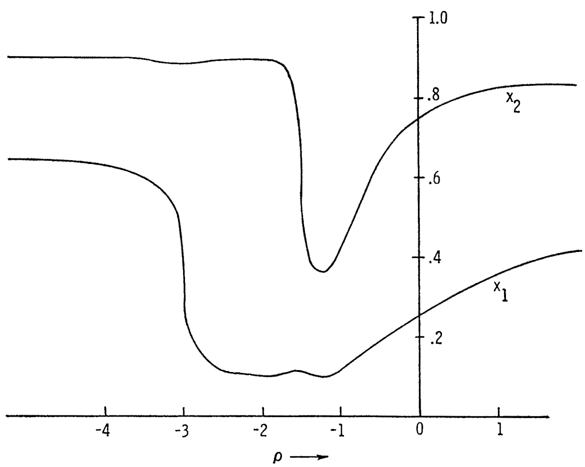


6. $F(A, x) = (a + bx)^3 = \alpha(1 + \rho x)^3, x \in [0, 1]$.

Set $u_i = 1 + \rho x_i, i = 1, 2$, and then the equations for the canonical points x_1 and x_2 appear as

$$u_1^4 - u_2^4 = \frac{1 - (1+\rho)^4}{2}, \qquad u_1^3 - u_2^3 = \frac{1 - (1+\rho)^3}{2}.$$

Comments. This approximating function is a *unisolvent* function [2].



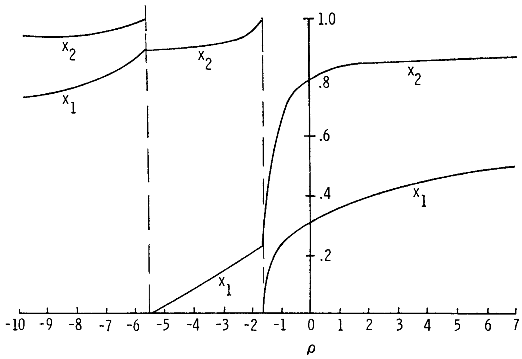
7. $F(A, x) = (a + bx^2)^2 = \alpha^2(1 + \rho x^2)^2, x \in [0, 1]$.

The equations for the canonical points x_1 and x_2 are

$$3(x_1 - x_2) + \rho(x_1^3 - x_2^3) = -\tfrac{1}{2}(3 + \rho),$$

$$5(x_1^3 - x_2^3) + 3\rho(x_1^5 - x_2^5) = -\tfrac{1}{2}(5 + 3\rho).$$

Comments. The points of derivative discontinuity in the curves for the canonical points are due to change of the solution from one branch to another. There are actually three pairs of solutions to these equations, but only one pair lies in the interval [0, 1] for any particular value of ρ.



8. $F(A, x) = a^2 + b^2 x + 2\, abx^2 = \alpha^2(\rho^2 + x + 2\rho x^2)$, $x \in [0, 1]$.

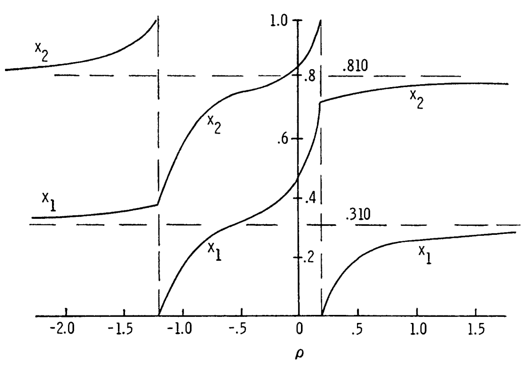
Set $u = x_1 - x_2$, $v = x_1 + x_2$, then we have

$$u^4 + u^2[12\rho + 12\rho^4] + u[12\rho^2(\rho^2 - \tfrac{1}{2}) + 2(3\rho + 1)] + 3(\rho^2 - \tfrac{1}{2}) = 0,$$

$$v = \frac{2\rho^2 u + \rho^2 - \frac{1}{2}}{u}$$

from which we may determine the two canonical points, x_1 and x_2. Some functions approximated on [0, 1] are $\sin x$, $|x - .6|$, e^{-x^2}, xe^{-x} and $x \sin x + 1 + 2x^2$.

Comments. This approximating function is not unisolvent. No curve of this form passes through (0, 4) and $(\frac{1}{2}, 1)$ and both $1 + 9x - 6x^2$ and $1 + x + 2x^2$ pass through (0, 1) and (1, 4). The points of discontinuous derivative in the canonical point curves are due to the same source as in the preceding example.



Mathematics Department
General Motors Research Laboratories
Warren, Michigan

1. H. J. MAEHLY, "Numerical solution of the equation $f(b) = (b^{x_4} - b^{x_2})/(b^{x_3} - b^{x_1}) = u$, $u > 0$. (A remark on a paper of J. R. Rice)," *J. Soc. Indust. Appl. Math.*, v. 10, 1962, p. 30–34.
2. T. S. MOTZKIN, "Approximation by curves of a unisolvent family," *Bull. Amer. Math. Soc.*, v. 55, 1949, p. 789–793.
3. J. R. RICE, "On nonlinear L_1 approximation," *Arch. Rational Mech. Anal.* (To appear.)
4. J. R. RICE, "A moment problem," Abstract 63T-206, *Notices Amer. Math. Soc.*, v. 10, 1963, p. 368.