JOURNAL OF APPROXIMATION THEORY 33, 111–118 (1981)

Functions with Many Best L_2-Approximations

M. AKHLAGHI

Department of Mathematics, University of Shiraz, Iran

AND

JERRY M. WOLFE

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

Communicated by E. W. Cheney

Received September 2, 1980

Let X denote a uniformly convex (real) vector space and let M denote a closed subset of M. Then from the work of Efimov and Stechkin [1] it is known that if M is also approximatively compact, then each $x \in X$ will have a unique closest point in M if and only if M is convex. Thus for the usual nonlinear approximating families such as the rational or exponential families, there will exist functions with more than one best approximation (if we are approximating in an L_p space with $1 < p < \infty$).

Specific examples of such functions were given in [2] by Lamprecht for approximation by polynomial rational functions and in [3] by Rice for nonlinear unisolvent families. In these cases and to our knowledge in all the other published examples, a symmetry argument was used to produce a function with two best approximations. Wolfe in [4] showed that for each positive integer k there is an $f \in L_2[0, 1]$ having at least k best *local* approximations from

$$R_m^n[0, 1] = \left\{ p/q \mid p(x) = \sum_{i=0}^{n} a_i x^i,\ q(x) = \sum_{i=0}^{m} b_i x^i,\ q(x) > 0 \text{ for } 0 \leqslant x \leqslant 1 \right\}$$

provided that $m > n$. Braess in [5] (among other things) removed the restriction $m > n$ and asked if it was possible to find a function f having at least three best approximations (not just local best approximations) from one of the standard nonlinear approximating families.

An affirmative answer to this question assuming "three" could be replaced

0021-9045/81/100111–08$02.00/0

640/33/2-3

by a diverging sequence of positive integers would show that there is no uniform upper bound on the number of best approximations that a function can have. This is in contrast to the situation using the uniform norm [6].

In this paper we shall give a straightforward technique that can explicitly produce functions with any specified number (call it N) of global approximations from a class of nonlinear families (with one nonlinear parameter) that includes many of the so called Γ-families of Hobby and Rice [7].

We are not able at this time to prove rigorously that the N approximants formed by our procedure are always best approximations to the function produced, though all our empirical evidence indicates this is so. However, we are able to give sufficient conditions for the approximations produced to be (global) best approximations and we are able to check these conditions numerically for the cases $N=3$ and $N=5$.

Approximating Family

The approximating family we shall use is defined by a continuous real-valued kernel function $K(\cdot,\cdot)$ of two real variables defined on $(-d,d)\times[a,b]$ for some $a>b$ and $d>0$ which satisfies the following conditions:

(i) $(\partial K/\partial\beta)(\beta,x)$ exists and is continuous on $(-d,d)\times[a,b]$,

(ii) $K(\beta_1,\cdot),\ldots,K(\beta_N,\cdot)$, $(\partial K/\partial\beta)(\beta_1,\cdot),\ldots,(\partial K/\partial\beta)(\beta_N,\cdot)$ are linearly independent on $[a,b]$ for any N distinct β's, $N=1,2,\ldots$.

Example 1. (a) $K_1(\beta,x)=e^{\beta x}$ on $(-\infty,\infty)\times[0,1]$

(b) $K_2(\beta,x)=1/(1-\beta x)$ on $(-1,1)\times[-1,1]$.

Given $f\in L_2[a,b]$ we shall consider approximations of the form (*) $\alpha K(\beta,x)$. That is, given f, we seek $\alpha^*\in R$ and $\beta^*\in(-d,d)$ such that

$$\|f-\alpha^*K(\beta^*,\cdot)\|=\inf_{\alpha\in R,\beta\in(-d,d)}\|f-\alpha K(\beta,\cdot)\|,$$

where $\|\ \|$ is the L_2 norm on $[a,b]$ with respect to Lebesgue measure. For notational simplicity let

$$u(\beta,x)=\frac{1}{\|K(\beta,\cdot)\|}K(\beta,x)$$

denote the normalized version of $K(\beta,x)$ and let $u'(\beta,x)$ denote

$(\partial u/\partial\beta)(\beta, x)$. Also, $[\cdot,\cdot]$ will denote the usual inner product in $L_2[a, b]$ defined by

$$[g, h] = \int_a^b g(t)\, h(t)\, dt \qquad \text{for} \quad g, h \in L_2[a, b].$$

Our goal is to construct a function $f \in L_2[a, b]$ that has many best approximations of the form (*). To do this we shall consider functions of the form

$$f(x) = \sum_{i=-n}^{n} [a_i u(\beta_i, x) + b_i u'(\beta_i, x)], \tag{1}$$

where $\beta_{-n},..., \beta_{-1}, \beta_0, \beta_1,..., \beta_n$ are $2n+1$ given distinct points in $(-d, d)$ and where we require that f also satisfy the conditions:

$$[f - u(\beta_i, \cdot), u(\beta_i, \cdot)] = 0, \tag{2}$$

$$[f - u(\beta_i, \cdot), u'(\beta_i, \cdot)] = 0. \tag{3}$$

$$i = 0, \pm 1,..., \pm n,$$

The system (2) and (3) is equivalent to the system one obtains by requiring that $(\partial/\partial\beta)\|f - u(\beta, \cdot)\|^2|_{\beta=\beta_i} = 0$, $i = 0, \pm 1,..., \pm n$, which is a necessary condition that each of the functions $u(\beta_i, \cdot) = (1/\|K(\beta_i, \cdot)\|)\, K(\beta_i, \cdot)$ $i = 0, \pm 1,..., \pm n$, be a best approximation to f of the form (*). The use of approximants normalized to unit length is for convenience only. The only essential thing is that they have the same norm.

Thus, we wish to choose β_i's, a_i's, and b_i's so that $u(\beta_i, \cdot)$ is a best approximation to f, $i = 0, \pm 1,..., \pm n$. Before showing that given the β_i's the corresponding a_i's and b_i's defining f are uniquely determined we offer the following example.

EXAMPLE 2. For $K(\beta, x) = e^{\beta x}$ over $(-\infty, \infty) \times [0, 1]$ we have

$$\begin{aligned} u(\beta, x) &= 1, \qquad \beta = 0 \\ &= \left(\frac{2\beta}{e^{2\beta} - 1}\right)^{1/2} e^{\beta x} \end{aligned}$$

and $f(x)$ takes the form $f(x) = \sum_{i=-n}^{n} (a_i + b_i x)\, e^{\beta_i x}$. Also (2) and (3) can be written in the form

$$[f - \alpha(\beta_i)\, e^{\beta_i x}, e^{\beta_i x}] = 0, \tag{2'}$$

$$[f - \alpha(\beta_i) e^{\beta_i x}, x e^{\beta_i x}] = 0, \tag{3'}$$

$$i = 0, \pm 1,..., \pm n,$$

where $\beta_0 = 0$ and $\alpha(\beta) = (1/\|K(\beta, \cdot)\|) = [\int_0^1 e^{2\beta t}\, dt]^{-1/2} = (2\beta/e^{2\beta} - 1)^{1/2}$

LEMMA 1. *Under the above assumptions on the kernel $K(\cdot, \cdot)$ and given $-d < \beta_{-n} < \beta_{-n+1} < \cdots < \beta_{-1} < \beta_0 < \cdots < \beta_n < d$ there exists a unique f of the form* (1) *satisfying* (2) *and* (3).

Proof. The system (2) and (3) can be written as

$$|f, u(\beta_i, \cdot)| = 1, \qquad (4)$$
$$i = 0, \pm 1, \ldots, \pm n,$$
$$|f, u'(\beta_i, \cdot)| = 0, \qquad (5)$$

since $\|u(\beta, \cdot)\| = 1$ for all $\beta \in (-d, d)$. Using the form of f in (1), (4) and (5) take the form

$$\sum_{-n}^{n} (a_j |u(\beta_j, \cdot), u(\beta_i, \cdot)| + b_j |u'(\beta_j, \cdot), u(\beta_i)|) = 1, \qquad (6)$$
$$i = 0, \pm 1, \ldots, \pm n,$$
$$\sum_{-n}^{n} (a_j |u(\beta_j, \cdot), u'(\beta_i, \cdot)| + b_j |u'(\beta_j, \cdot), u'(\beta_i, \cdot)|) = 0. \qquad (7)$$

The system (6), (7) is a linear system of $4n + 2$ equations in the $4n + 2$ unknowns a_j, b_j, $j = 0, \pm 1, \ldots, \pm n$, where the coefficient matrix is a Gram matrix formed from the linearly independent functions

$$u(\beta_{-n}, \cdot), \ldots, u(\beta_n, \cdot), \qquad u'(\beta_{-n}, \cdot), \ldots, u'(\beta_n, \cdot).$$

Thus (6), (7) has a unique solution as claimed. ∎

In view of Lemma 1, our problem is to make a proper choice of β_i, $i = 0, \pm 1, \ldots, \pm n$, to insure that the corresponding function f of the required form actually has $u(\beta_i, \cdot)$, $i = 0, \pm, \ldots, \pm n$, as best approximations. A natural choice is to place the nodes symmetrically about the origin and this is what we shall do. (Empirically we found that nonsymmetric choices often produced only local best approximations or even saddle points.) Thus for symmetry we require:

$$\beta_0 = 0, \qquad (8)$$
$$0 < \beta_i < \beta_{i+1} < d, \qquad i = 0, \ldots, n-1, \qquad (9)$$
$$\beta_i = -\beta_{-i}. \qquad (10)$$

We shall also assume that the normalized kernel function $u(\beta, \cdot)$ satisfies the midpoint symmetry condition.

$$u(-\beta, x) = u(\beta, 2p - x), \qquad \text{where} \quad p = (a + b)/2. \qquad (11)$$

Remark. For both $u_1(\beta, x) = (2\beta/e^{2\beta} - 1)^{1/2} e^{\beta x}$ and $u_2(\beta, x) = ((1 - \beta^2)/2)^{1/2} 1/(1 - \beta x)$ (the normalized kernels for the functions of example 1) it is simple to check that (11) holds on $(-\infty, \infty) \times [0, 1]$ and $(-1, 1) \times [-1, 1]$, respectively.

LEMMA 2. *Let f be a function of the form* (1) *with the β's satisfying* (8), (9), *and* (10) *and such that u satisfies* (11). *Then*

(a) $f(x) = f(2p - x)$, $x \in [a, b]$, $p = (a + b)/2$,

$$\left.\begin{aligned} a_{-i} &= a_i, \\ b_{-i} &= -b_i, \end{aligned}\right. \quad i = 0, 1, \ldots, n \text{ and in particular } b_0 = 0. \tag{b}$$

Proof. Let $h(x) \equiv f(x) - f(2p - x)$. By differentiating both sides of (11) with respect to β we obtain the identity.

$$u'(-\beta, x) = u'(\beta, 2p - x). \tag{12}$$

Using (11) and (12) and the definition of h, we can write h is the form

$$h(x) = \sum_{i=-n}^{n} \{(a_i - a_{-i})\, u(\beta_i, x) + (b_i + b_{-i})\, u'(\beta_i, x)\} \tag{13}$$

which shows that h is a linear combination of $u(\beta_i, \cdot)$ and $u'(\beta_i, \cdot)$, $i = 0, \pm 1, \ldots, \pm n$.

CLAIM 1. $\int_a^b h(x)\, u(\beta_i, x)\, dx = 0$, $i = 0, \pm 1, \ldots, \pm n$.

CLAIM 2. $\int_a^b h(x)\, u'(\beta_i, x)\, dx = 0$, $i = 0, \pm 1, \ldots, \pm n$.

If these claims are proved then by considering (13) the four conclusions (a)–(d) would obviously follow immediately.

Proof of Claim 1. $\int_a^b h(x)\, u(\beta_i, x)\, dx = \int_a^b f(x)\, u(\beta_i, x) - \int_a^b f(2p - x)\, u(\beta_i, x)\, dx = 1 - \int_a^b f(t)\, u(\beta_i, 2p - t)\, dt$ by taking $t = 2p - x$ and using (4).

But by (11), $\int_a^b f(t)\, u(\beta_i, 2p - t)\, dt = \int_a^b f(t)\, u(-\beta_i, t)\, dt = 1$ and so $\int_a^b h(x)\, u(\beta_i, x)\, dx = 0$, $i = 0, \pm 1, \ldots, \pm n$, proving Claim 1.

Ptoof of Claim 2. We have $\int_a^b h(x)\, u'(\beta_i, x)\, dx = 0 - \int_a^b f(t)\, u'(\beta_i, 2p - t)\, dt$ again letting $t = 2p - x$ and using (5). By (12), $\int_a^b f(t)\, u'(\beta_i, 2p - t)\, dt = -\int_a^b f(t)\, u'(-\beta_i, t)\, dt = 0$ (by (5)) so $\int_a^b h(x)\, u'(\beta_i, x)\, dx = 0$, $i = 0, \pm 1, \ldots, \pm n$. ∎

Our original approximation problem involves approximating f by functions of the form $\alpha K(\beta, \cdot)$ and hence appears to have two parameters α and β. However, we may eliminate α by noting that if $\alpha^* K(\beta^*, \cdot)$ is a (local) best approximation to f it must satisfy the condition

$$\frac{\partial}{\partial \alpha} [f - \alpha K(\beta, \cdot), f - \alpha K(\beta, \cdot)]\Big|_{\substack{\alpha = \alpha^* \\ \beta = \beta^*}} = 0 \tag{14}$$

and this yields

$$[f - \alpha^* K(\beta^*, \cdot), K(\beta^*, \cdot)] = 0. \tag{15}$$

which implies that

$$\alpha^* = \frac{[f, K(\beta^*, \cdot)]}{[K(\beta^*, \cdot), K(\beta^*, \cdot)]}.$$

Thus if we let

$$\psi(\beta) = \| f - r(\beta, \cdot)\|^2, \qquad \text{where} \quad r(\beta, x) = s(\beta)\, K(\beta, x) \tag{16}$$

with $s(\beta) = [f, K(\beta, \cdot)]/\|K(\beta, \cdot)\|^2$ for $\beta \in (-d, d)$, then the problem of finding a best approximation to f is equivalent to finding a $\beta \in (-d, d)$ that minimizes ψ.

Lemma 3. *For each* $\beta \in (-d, d)$, $\psi(\beta) = \| f \|^2 - [f, u(\beta, \cdot)]^2$.

Proof. $\psi(\beta) = \| f \|^2 - 2[f, r(\beta, \cdot)] + \| r(\beta, \cdot)\|^2$. But from (16),

$$\| r(\beta, \cdot)\|^2 = \| s(\beta)\, K(\beta, \cdot)\|^2 = \frac{[f, K(\beta, \cdot)]^2}{\|K(\beta, \cdot)\|^4} \| K(\beta, \cdot)\|^2$$

$$= \left[f, \frac{K(\beta, \cdot)}{\| K(\beta, \cdot)\|} \right]^2 = [f, u(\beta, \cdot)]^2$$

and also

$$[f, r(\beta, \cdot)] = s(\beta)\, [f, K(\beta, \cdot)] = \frac{[f, K(\beta, \cdot)]^2}{\| K(\beta, \cdot)\|^2}$$

$$= [f, u(\beta, \cdot)]^2.$$

Thus, $\psi(\beta) = \| f \|^2 - [f, u(\beta, \cdot)]^2$. ∎

Lemma 4. *For each* $\beta \in (-d, d)$, $[f, u(-\beta, \cdot)] = [f, u(\beta, \cdot)]$.

Proof. $[f, u(-\beta, \cdot)] = \int_a^b f(x)\, u(-\beta, x)\, dx = \int_a^b f(x)\, u(\beta, 2p - x)\, dx = \int_a^b f(2p - t)\, u(\beta, t)\, dt = \int_a^b f(t)\, (\beta, t)\, dt = [f, u(\beta, \cdot)]$, where $t = 2p - x$ and where we have used Lemma 2. ∎

From Lemma 4, the following two corollaries are immediate.

Corollary 1. $\psi(\beta) = \psi(-\beta)$ *for every* $\beta \in (-d, d)$.

Corollary 2. *For* $i = 0, \pm 1, \ldots, \pm n$, $\psi(\beta_i) = \| f \|^2 - 1$ *and hence* $\| f \| > 1$.

From the definition of β_i we have that $u(\beta_i, \cdot) = r(\beta_i, \cdot)$, $i = 0, \pm 1, \pm 2, \ldots, \pm n$, but it remains to be seen whether or not the value $\psi(0) = \|f\|^2 - 1$ is a global minimum for ψ. The following lemma demonstrates that we may confine ourselves to a bounded interval in checking this.

LEMMA 5. *For β satisfying $\|u(\beta, \cdot)\|_1 < 1/\|f\|_\infty$, we have $\psi(\beta) > \psi(0)$, where $\| \ \|_1$ denotes the L_1-norm.*

Proof. $\psi(\beta) = \|f\|^2 - [f, u(\beta, \cdot)]^2 \geqslant \|f\|^2 - \|f\|_\infty^2 \|u(\beta, \cdot)\|_1^2$ and therefore $\psi(\beta) - \psi(0) = 1 - [f, u(\beta, \cdot)]^2 \geqslant 1 - \|f\|_\infty^2 \|u(\beta, \cdot)\|_1^2$. Hence if $1 - \|f\|_\infty^2 \|u(\beta, \cdot)\|_1^2 > 0$, then $\psi(\beta) > \psi(0)$ and this condition is equivalent to $\|u(\beta, \cdot)\|_1 < 1/\|f\|_\infty$. ∎

COROLLARY 3. *For $K(\beta, x) = e^{\beta x}$ over $(-\infty, \infty) \times [0, 1]$ if $\beta > 2\|f\|_\infty^2$ then $\psi(\beta) > \psi(0)$.*

Proof. We simply calculate

$$\|u(\beta, \cdot)\|_1 = \int_0^1 \left(\frac{2\beta}{e^{2\beta} - 1}\right)^{1/2} e^{\beta x}\, dx = \left(\frac{2\beta}{e^{2\beta} - 1}\right)^{1/2} \frac{e^\beta - 1}{\beta}$$
$$= \left(\frac{2}{\beta}\right)^{1/2} \left(\frac{e^\beta - 1}{e^\beta + 1}\right)^{1/2}.$$

For $\beta > 0$, $(e^\beta - 1)/(e^\beta + 1) < 1$ and hence $\|u(\beta, \cdot)\|_1 < (2/\beta)^{1/2}$. By Lemma 5, if $(2/\beta)^{1/2} < 1/\|f\|_\infty$ then $\psi(\beta) > \psi(0)$. But this is equivalent to $\beta > 2\|f\|_\infty^2$. ∎

As mentioned in the introduction we do not have a proof of the following theorem in the strict mathematical sense. We are able to give a computational "proof" in the manner described below which shows that Braess's question has been answered in the affirmative up to the accuracy of our numerical procedure.

THEOREM. *For $n = 1$, $\beta_1 = 2$, and $K(\beta, x) = e^{\beta x}$ defined over $(-\infty, \infty) \times [0, 1]$, there exists a unique function $f(x)$ of the form* (1) *have exactly $2n + 1 = 3$ global best approximations (in the L_2 sense), namely, $u(-2, x)$, $u(0, x) = 1$, and $u(2, x)$, where $u(\beta, x) = (2\beta/e^{2\beta} - 1)^{1/2} e^{\beta x}$.*

"*Proof.*" The "proof" of this result was accomplished as follows. The values a_0, a_1, and b_1 that determine f were found by solving the linear system (6), (7) numerically. Then by computing the values of f on $[0, 1]$ on a grid of equally spaced points whose common spacing was sufficiently "small" (a spacing of 0.05 was found to be sufficient), the inequality

$\|f\|_\infty < 3.76$ was obtained. Finally, a similar search verified that $\beta = \pm 2, 0$ were the only minima of $\psi(\beta)$ in $(-32, 32)$ and since $2\|f\|_2^2 < 32$ we concluded from Corollary 3 that f has $u(0, x)$, $u(2, x)$, and $u(-2, x)$ as its only global best approximations. “∎”

This same technique was successfully applied to the case $n = 2$ (i.e., $N = 5$). For $n \geqslant 3$ (i.e., $N \geqslant 7$) the interval obtained from Corollary 3 was so large that overflow occurred in the computations and so the results were unreliable. However, all the evidence at our disposal indicates that the following conjecture is correct.

Conjecture. Given $n \geqslant 1$ and any $2n + 1$ distinct values β_i, $i = 0 \pm 1, \ldots, \pm n$, symmetrically placed about the origin in $(-d, d)$ there is a unique f of the form (1) having each function $u(\beta_i, \cdot)$, $i = 0, \pm 1, \ldots, \pm n$, as its set of global best L_2-approximations from the approximating family $f = \{\alpha K(\beta, \cdot) \mid \alpha \in R, \beta \in (-d, d)\}$.

References

1. N. V. Efimov and S. B. Stechkin, Approximative compactness and Chebychev sets, *Dokl. Akad. Nauk. SSSR* **140** (1961), 522–524.
2. G. Lamprecht, Zur Mehrdentigkeit bei der Approximation in der L_p-norm, *Computing* **5** (1970), 349–355.
3. J. R. Rice, "The Approximation of Functions," Vol. II, Addison-Wesley, Reading, Mass., 1969.
4. J. M. Wolfe, On the unicity of nonlinear approximation in smooth spaces, *J. Approx. Theory* **12** (1974), 165–181.
5. D. Braess, On rational L_2-approximation, *J. Approx. Theory* **18** (1976), 136–151.
6. D. Braess, On the number of best approximations in certain nonlinear families of function, *Aequationes Math.* **12** (No. 2) (1975), 184–199.
7. C. R. Hobby and J. R. Rice, Approximation from a curve of functions, *Arch. Rational Mech. Anal.* **24** (1967), 91–106.