JOURNAL OF APPROXIMATION THEORY **47**, 85–91 (1986)

Best Monotone Approximations in $L_1[0, 1]$

ROBERT HUOTARI

Department of Mathematical Sciences, Indiana University-Purdue University at Fort Wayne, Fort Wayne, Indiana 46805, U.S.A.

AND

AARON D. MEYEROWITZ AND MICHAEL SHEARD

Department of Mathematics, The Ohio State University, Colombus, Ohio 43210, U.S.A.

Communicated by Oved Shisha

Received December 31, 1984

For f in $L_1[0, 1]$, let $\mu_1(f|M)$ be the set of all best L_1-approximations to f by nondecreasing functions, let $\underline{f} = \inf \mu_1(f|M)$ and let $\bar{f} = \sup \mu_1(f|M)$. We show that $\mu_1(f|M)$ is compact in L_1 and provide a description of the elements of $\mu_1(f|M)$ and of its extreme points in terms of $\underline{f}$ and $\bar{f}$.

Let $\Omega = [0, 1]$, $\mu =$ Lebesgue measure on Ω and $\mathfrak{A} =$ the set of Lebesgue measurable subsets of Ω. Let $L_1 = L_1(\Omega, \mathfrak{A}, \mu)$ and let f in L_1 be fixed. If $\mathfrak{B}$ is a sub sigma lattice of $\mathfrak{A}$ (i.e., $\mathfrak{B}$ contains ϕ, Ω, and countable unions and intersections of elements of $\mathfrak{B}$) and B is the set of all $\mathfrak{B}$-measurable functions, a function g in B is a best L_1-approximation to f by elements of B if, for every h in B,

$$\|f-g\|_1 \leqslant \|f-h\|_1.$$

Let $\mu_1(f|B)$ denote the set of all best L_1-approximations to f by elements of B; let $\underline{f} = \inf \mu_1(f|B)$ and $\bar{f} = \sup \mu_1(f|B)$. If $\mathfrak{B}$ is also a sigma algebra, then $\mu_1(f|B)$ can be characterized as follows (see [3, Theorem 2]): g is in $\mu_1(f|B)$ if and only if g is in B and $\underline{f} \leqslant g \leqslant \bar{f}$. A more general problem is to characterize $\mu_1(f|B)$ when $\mathfrak{B}$ is an arbitrary sigma lattice. In attacking this problem, we have considered a particular lattice, which underlies the theory of isotonic approximation.

Let $\mathfrak{M}$ be the collection of all intervals of the form $[a, 1]$ or $(a, 1]$, $0 \leqslant a \leqslant 1$. Then a function g is $\mathfrak{M}$-measurable if and only if g is nondecreasing on Ω. Let M be the set of $\mathfrak{M}$-measurable (nondecreasing) functions. By [2, Lemma 3], $\mu_1(f|M)$ is nonempty. Clearly $\mu_1(f|M)$ is L_1-closed and convex. The purpose of this note is the description of $\mu_1(f|M)$ (Theorem 8) and of its extreme points (Theorem 9).

0021-9045/86 $3.00

For x in Ω and $\delta > 0$, let $B(x, \delta) = (x-\delta, x+\delta) \cap \Omega$. If S is in $\mathfrak{A}$ and Q is a nondegenerate interval in Ω, let $\mu(S; Q)$ denote the relative measure of S in Q, i.e., $\mu(S; Q) = \mu(S \cap Q)/\mu Q$. If S is in $\mathfrak{A}$, then a point p in $(0, 1)$ is said to be a *point of density* of S if

$$\lim_{\delta \downarrow 0} \mu(S; B(p, \delta)) = 1.$$

A function g is said to be *approximately continuous* at p in $(0, 1)$ if, for any open set G containing $g(p)$, p is a point of density of the set $g^{-1}(G)$.

LEMMA 1. *Let $g \in \mu_1(f|M)$. If f is approximately continuous at y and $g(y) \neq f(y)$, then g is constant on a neighborhood of y.*

Proof. Suppose y is in $(0, 1)$, $g(y) \neq f(y)$ and Lemma 1 is false at y. Suppose $f(y) - g(y) = 2\varepsilon > 0$. (If $f(y) < g(y)$, the proof is similar.) Since g is not constant at y, either $x > y \Rightarrow g(x) > g(y)$ or $x < y \Rightarrow g(x) < g(y)$. In either case we obtain a contradiction by constructing a nondecreasing function ϕ with $\|f-\phi\|_1 < \|f-g\|_1$. In the first case, ϕ is obtained by raising g on an interval containing y. Since f is approximately continuous at y, there exists $\sigma > 0$ such that, for $0 < \tau < \sigma$,

$$\mu([f > f(y) - \varepsilon]; B(y, \tau)) > 3/4. \tag{1}$$

Let $z = \min\{y+\sigma, \inf\{x\colon g(x) \geqslant g(y) + \varepsilon\}\}$ and let $\eta = \min\{\varepsilon, g(z+) - g(y-)\}$, where $g(x+) = \lim_{t \downarrow x} g(t)$ and $g(x-)$ is defined similarly. Define ϕ in M by

$$\begin{aligned} \phi(x) &= g(x) + \eta, && x \in (y-\sigma, y), \\ &= g(z+), && x \in [y, z], \\ &= g(x), && x \notin (y-\sigma, z]. \end{aligned}$$

By (1), $\mu([f > f(y) - \varepsilon]; I) > \frac{1}{2}$ for $I = (y-\sigma, y)$ and for $I = (y, z)$. Thus,

$$\int_I |\phi - f| < \int_I |g - f|$$

for $I = (y-\sigma, y)$ and for $I = (y, z)$. Thus ϕ is L_1-closer to f than is g.

In the second case, a better L_1-approximation is produced by lowering g on an interval of the form $(z, y+\sigma]$. When $y = 0$ or 1, a similar argument applies. This concludes the proof of Lemma 1.

COROLLARY 2. *Let $g \in \mu_1(f|M)$. Then Ω is the disjoint union of subsets E and F where, for almost every y in E, $g(y) = f(y)$ and for each y in F, g is constant on a neighborhood of y.*

Proof. Since f is measurable, f is approximately continuous almost everywhere [1, II.5.2]. Let G be the exceptional set, $E=[f=g]\cup G$ and $F=\Omega-E$. Lemma 1 then shows that F has the desired property.

We will say that a function $h\colon \Omega\to\mathbb{R}$ is *constant at* s if there exists $\delta>0$ such that h is constant on $B(s,\delta)$.

LEMMA 3. *Let $g\in\mu_1(f|M)$. If g is not constant at s in Ω, then*

$$\mu([f<g];[s,t])\leqslant\tfrac{1}{2},\qquad s<t\leqslant 1 \tag{2}$$

and

$$\mu([f>g];[t,s])\leqslant\tfrac{1}{2},\qquad 0\leqslant t<s. \tag{3}$$

Furthermore,

$$\mu([f>g];[s,1])\leqslant\tfrac{1}{2}\qquad \textit{for any } s \textit{ in } [0,1) \tag{4}$$

and

$$\mu([f<g];[0,t])\leqslant\tfrac{1}{2}\qquad \textit{for any } t \textit{ in } (0,1]. \tag{5}$$

Proof. Suppose (2) fails. As in the proof of Lemma 1, we will construct a function ϕ in M which is L_1-closer to f than is g. The negation of (2) entails the existence of s and t such that $0\leqslant s<t\leqslant 1$ and g is not constant at s, but $\mu([f<g];[s,t])>\frac{1}{2}$. Since $[f<g]$ is the limit of the increasing sequence of sets $\{[f<g-1/n],\ n\geqslant 1\}$, there exist n' in $\mathbb{N}$ and $\delta>0$ such that

$$\mu([f<g-1/n']\cap[s,t])>(t-s)/2+2\delta.$$

Since g is not constant at s, there exist u, v and n'' such that $u\leqslant s<v$, $v-u<\delta$ and $g(v)-1/n''\geqslant g(u)$. Let $\nu=\min\{1/n',1/n''\}$ and define ϕ: $\Omega\to\mathbb{R}$ by

$$\begin{aligned}\phi(x)&=\min\{g(x),g(v)-\nu\}, &&x\in[u,v],\\ &=g(x)-\nu, &&x\in(v,t],\end{aligned}$$

and $\phi=g$ on $\Omega-[u,t]$. Then ϕ is nondecreasing and

$$\begin{aligned}&\int_u^t|\phi-f|\\ &\quad\leqslant\int_u^v|g-f|+\nu\delta\\ &\qquad+\int_v^t|g-f|-\nu[(t-s)/2+\delta]+\nu[(t-s)/2-2\delta]\\ &\quad\leqslant\int_u^t|g-f|-2\nu\delta.\end{aligned}$$

Thus ϕ is a better L_1-approximation to f than is g, a contradiction. That (3), (4), and (5) hold is proven similarly. This concludes the proof of Lemma 3.

Let $\underline{f}=\inf \mu_1(f|M)$ and $\bar{f}=\sup \mu_1(f|M)$. By [2, Lemma 3], $\underline{f}$ and $\bar{f}$ are in $\mu_1(f|M)$. Items (4) and (5) of Lemma 3 (with $s=0$ and $t=1$) imply the following:

COROLLARY 4. *If* $g \in \mu_1(f|M)$, *then* $\mu[f>g] \leqslant \frac{1}{2}$ *and* $\mu[f<g] \leqslant \frac{1}{2}$. *Consequently* $\mu[f \geqslant \bar{f}] \geqslant \frac{1}{2}$, $\mu[f \leqslant \underline{f}] \geqslant \frac{1}{2}$ *and* $\mu[\underline{f}<f<\bar{f}]=0$.

THEOREM 5. *If* $\bar{f}$ *(resp.* $\underline{f}$*) is constant on* $(a, b) \subset \Omega$, *then there exists a disjoint sequence* $\{(\alpha_i, b_i)\colon i \geqslant 1\}$ *of subintervals of* (a, b) *such that* $\underline{f}$ *(resp.* $\bar{f}$*) is constant on each* (a_i, b_i) *and* $\mu[\bigcup_{i \geqslant 1} (a_i, b_i)] = b-a$.

Proof. Suppose without loss of generality that $\bar{f}$ is constant on (a, b) and not constant at a. Let E be the set of all points in (a, b) at which $\underline{f}$ is not constant. If $\mu E>0$, then, by the Lebesgue Density Theorem [1, II.5.1], there exists y in E such that y is a point of density of E. If $\underline{f}(y+)=\bar{f}(y)$, then there exists $\delta>0$ such that $\underline{f}$ is constant on $(y, y+\delta)$, which contradicts the choice of y. Thus $\underline{f}(y+)<\bar{f}(y)$, so there exist c and d in E such that $c<y<d$, $\underline{f}(d)<\bar{f}(d)$ and $\mu(E;\, [y, d])>\frac{1}{2}$.

Let $I_1=(a,\ y)$ and $I_2=(a,\ d)$. Lemma 3 implies that, for $i=1, 2$, $\mu([f \geqslant \bar{f}];\ I_i) \geqslant \frac{1}{2}$ and $\mu([f \leqslant \underline{f}];\ I_i) \geqslant \frac{1}{2}$. Since $\underline{f}<\bar{f}$ on (a, d) and $\mu[\underline{f}<f<\bar{f}]=0$, $\mu([f \leqslant \underline{f}];\ I_i)=\frac{1}{2}$, $i=1, 2$, so $\mu([f \leqslant \underline{f}];\ [y, d])=\frac{1}{2}$.

In view of Corollary 2 and Theorem 5, we can state an initial characterization of $\mu_1(f|M)$: Let U be the set of points at which $\underline{f}$ and $\bar{f}$ are both constant and unequal. Then U is an open set and $\underline{f}(y)=\bar{f}(y)$ for almost every y in $\Omega-U$. Thus we will have characterized $\mu_1(f|M)$ if we can describe the behavior of an arbitrary $g \in \mu_1(f|M)$ on an arbitrary component of U. To this end, define $h\colon \Omega \rightarrow \mathbb{R}$ by

$$
\begin{aligned}
h(x) &= 1, && \underline{f}(x)<\bar{f}(x) \leqslant f(x), \\
&= -1, && f(x) \leqslant \underline{f}(x)<\bar{f}(x), \\
&= 0, && \text{otherwise},
\end{aligned}
$$

and let $k(x)=\int_0^x h(t)\, dt$. Then $h(x)=0$ almost everywhere on $\Omega-U$ and for any component (u, v) of U, $\int_u^v h(t)\, dt=0$. Hence $[h=0] \supset \Omega-U$. Furthermore, we have

LEMMA 6. *The set* $[k=0] \cap U$ *has measure zero.*

Proof. Suppose $y \in U$ is a point of density of $[f \geqslant \bar{f}]$. Then there exists $\delta>0$ such that

$$\mu([f \geqslant \bar{f}]; I)>\tfrac{1}{2} \tag{6}$$

for every interval I which contains y and is contained in $B(y, \delta)$. If y is also a point of density of $[k=0]$, then there exist z and w such that $y-\delta<z<y<w<y+\delta$ and $k(z)=k(w)=0$. Thus $\int_z^w h(t)\,dt=0$. But (6) implies that $\int_z^w h(t)\,dt>0$. Thus $[f \geqslant \bar{f}]$ and $[k=0]$ have no common density points. Similarly, $[f \leqslant \underline{f}]$ and $[k=0]$ have no common density points. Since almost every x in Ω is a density point of $[f \geqslant \bar{f}]$ or of $[f \leqslant \underline{f}]$, the measure of the set of density points of $[k=0]$ is zero, i.e., $\mu[k=0]=0$.

Since k is continuous, $[k \neq 0]$ is an open set. Let $V=U \cap [k \neq 0]$. Our characterization of $\mu_1(f|M)$ focuses on the components of V.

THEOREM 7. *If $g \in \mu_1(f|M)$ and k is defined as above, then g is constant on each component of $[k \neq 0]$.*

Proof. Suppose g is not constant at y, where $z<y<w$ and (z, w) is a component of $[k \neq 0]$ such that (z, w) is contained in the component (u, v) of U. Then (3) in Lemma 3 implies that

$$\mu([f \leqslant \underline{f}];\ [z, y]) \geqslant \tfrac{1}{2}. \tag{7}$$

If equality holds in (7), then $\int_z^y h(t)\,dt=0$ so $k(y)=0$, a contradiction. Since $k(z)=0$,

$$\mu([f \leqslant \underline{f}];\ [u, z])=\tfrac{1}{2}$$

so

$$\mu([f \leqslant \underline{f}];\ [u, y])>\tfrac{1}{2},$$

which contradicts (2) in Lemma 3. This establishes Theorem 7.

The following theorem is a summary of our characterization of $\mu_1(f|M)$.

THEOREM 8. *If $f \in L_1$, then $g \in \mu_1(f|M)$ if and only if $g \in M$, $\underline{f} \leqslant g \leqslant \bar{f}$, g is constant on each component of V and, for almost every x in $\Omega - V$, $g(x)=f(x)$.*

Another way to characterize $\mu_1(f|M)$ is by means of its extreme points. We have noted above that $\mu_1(f|M)$ is convex. Thus, if we can also show that $\mu_1(f|M)$ is compact, the Krein–Milman Theorem will assert that $\mu_1(f|M)$ is the closed convex hull of its extreme points. To see that it is compact, let $\{(\alpha_n, \beta_n): n \geqslant 1\}$ be the set of components of V and, for any g in $\mu_1(f|M)$ and $n \geqslant 1$, let g_n be the value of g on (α_n, β_n). Suppose $\{g^m: m \geqslant 1\}$ is a sequence in $\mu_1(f|M)$. Since $\bar{f}$ and $\underline{f}$ are integrable, it is necessary that $\{g_n^m: m \geqslant 1\}$ is bounded for each $n \geqslant 1$. Thus, there exists a subsequence $\{g^{m,1}\}$ of $\{g^m\}$ such that $\{g_1^{m,1}=(g^{m,1})_1\}$ converges and, for each $k \geqslant 1$, there exists a subsequence $\{g^{m,k+1}\}$ of $\{g^{m,k}\}$ such that $\{g_{k+1}^{m,k+1}\}$

converges. Then the subsequence $\{g^{1,1}, g^{2,2}, g^{3,3},...\}$ of $\{g^m\}$ converges at every point of V. Since $\underline{f}=f=\bar{f}$ almost everywhere on $\Omega-V$, $\{g^m\}$ converges almost everywhere. Clearly $g^0=\lim g^m$ is nondecreasing and, by the Lebesgue Convergence Theorem, g^0 is in $\mu_1(f|M)$ and $\|g^m-g^0\|_1\to 0$.

THEOREM 9. *Let $f\in L_1$. A function g in $\mu_1(f|M)$ is an extreme point of $\mu_1(f|M)$ if and only if, for any $n\geqslant 1$,*

$$g_n\in \mathrm{Cl}\{g_m: g_m=\bar{f}_m \text{ or } g_m=\underline{f}_m\}. \tag{8}$$

Proof. The condition is sufficient. Indeed, suppose that h and k are in $\mu_1(f|M)$ and $g=\rho h+(1-\rho)k$, $0<\rho<1$. If $g_m=\bar{f}_m$ or $g_m=\underline{f}_m$, and h_m and k_m are distinct from g_m, then either h_m or k_m is excluded from $[\underline{f}_m, \bar{f}_m]$, a contradiction. Thus $h_m=k_m=g_m$, so $h_n=k_n=g_n$ whenever g_n is a limit point of $\{g_m: g_m=\bar{f}_m \text{ or } \underline{f}_m\}$. Since $h=k=f$ almost everywhere on $\Omega-V$, the three functions are equivalent, i.e., g is an extreme point.

Conversely, suppose that there exists n such that g_n does not satisfy (8). Let (α, β) be the largest open interval on which $g=g_n$.

Suppose first that g is continuous at α and at β. By the negation of (8), there exists $\eta>0$ such that for any $m\geqslant 1$, $\bar{f}_m$ and $\underline{f}_m$ are excluded from $B(g_n, 2\eta)$. Choose $z>\beta$ and $w<\alpha$ so that

$$0<g(z)-g(\beta)=\varepsilon_2\leqslant\eta$$

and

$$0<g(\alpha)-g(w)=\varepsilon_1\leqslant\eta.$$

Let $\varepsilon=\min\{\varepsilon_1, \varepsilon_2\}$ and define g^u by

$$\begin{aligned}
g^u(x)&=g(x)+\varepsilon(g(x)-g(w))/(g(\alpha)-g(w)), && x\in(w,\alpha),\\
&=g(x)+\varepsilon, && x\in[\alpha,\beta],\\
&=g(x)+\varepsilon(g(z)-g(x))/(g(z)-g(\beta)), && x\in(\beta,z),\\
&=g(x), && x\notin(w,z).
\end{aligned}$$

Define g^l similarly, with each addition replaced by a subtraction. Since g^u is constant exactly where g is constant and $\mu[g<f<g^u]=0$, $\|g^u-f\|_1=\|g-f\|_1$, so g^u is in $\mu_1(f|M)$. Similarly, $g^l\in\mu_1(f|M)$. Since $g=(g^u+g^l)/2$, g is not an extreme point of $\mu_1(f|M)$.

If g is discontinuous at α, choose $\varepsilon<\min\{g(\alpha+)-g(\alpha-), \eta\}$, let $w=\alpha$ and define g^u and g^l as before. (The first line in the above definition of g^u is now vacuous.) If g is discontinuous at β, similar changes are made.

References

1. A. M. Bruckner, "Differentiation of Real Functions," Lecture Notes in Math. No. 659, Springer–Verlag, Berlin/Heidelberg/New York, 1978.
2. D. Landers and L. Rogge, Natural choice of L_1-approximants, *J. Approx. Theory* 33 (1981), 268–280.
3. T. Shintani and T. Ando, Best approximants in L^1-space, *Z. Warsch. Verw. Gebiete* 33 (1975), 33–39.

Printed in Belgium