PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 81, Number 3, March 1981

CONVERGENCE OF L_p APPROXIMATIONS AS $p \to \infty$

RICHARD B. DARST

ABSTRACT. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $\mathcal{B}$ be a subsigma-algebra of $\mathcal{A}$. Let $A = L_\infty(\Omega, \mathcal{A}, \mu)$ and let $B = L_\infty(\Omega, \mathcal{B}, \mu)$. Let $f \in A$, and for $1 < p < \infty$, let f_p denote the best L_p approximation to f by elements of $L_p(\Omega, \mathcal{B}, \mu)$. It is shown that $\lim_{p\to\infty} f_p$ exists a.e. The function f_∞ defined by $f_\infty(x) = \lim_{p\to\infty} f_p(x)$ is a best L_∞ approximation to f by elements of B: $\|f - f_\infty\|_\infty = \inf\{\|f - g\|_\infty; g \in B\}$. Indeed, f_∞ is a best best L_∞ approximation to f by elements of B in the sense that for each $E \in \mathcal{B}$ the restriction, $f_\infty|E$, of f_∞ to E is a best L_∞ approximation to the restriction, $f|E$, of f to E. Since there is at most one best best L_∞ approximation to f, f_∞ is the best best L_∞ approximation to f by elements of B.

The author wishes to thank the referee for his helpful comments and for pointing out several of the references that discuss similar work. In particular, §§12.7 and 12.8 of [5] contain a nice discussion of related work.

We begin with some notation and terminology.

Let $\mathcal{F} = \mathcal{F}(\mathcal{B})$ denote the set of finite partitions of Ω by elements of $\mathcal{B}$ and let $\mathcal{P} = \mathcal{P}(\mathcal{B})$ denote the set of denumerable partitions of Ω by elements of $\mathcal{B}$.

Let $O(f, E)$ denote the essential oscillation of f on $E \in \mathcal{A}$: $O(f, E) = \operatorname{essup}(f, E) - \operatorname{essinf}(f, E)$, where $\operatorname{essup}(f, E) = \operatorname{essinf}(f, E) = 0$ if $\mu(E) = 0$ and for $\mu(E) > 0$,

$$\operatorname{essup}(f, E) = \inf\{\lambda; \mu(\{x \in E; f(x) > \lambda\}) = 0\} \quad \text{and}$$
$$\operatorname{essinf}(f, E) = \sup\{\lambda; \mu(\{x \in E; f(x) < \lambda\}) = 0\}.$$

Let $d(f, B)$ denote the distance from an element f of A to the subspace B of A.

Before plunging into the technical details, we offer a brief outline. Lemma 1 shows that both $\mathcal{F}$ and $\mathcal{P}$ can be used to estimate $d(f, B)$. Lemma 2 asserts that the added flexibility afforded by $\mathcal{P}$ permits us to replace an inf by a min. The partitions corresponding to each min provide an equivalence class of elements of $\mathcal{B}$. These equivalence classes comprise a monotone family parametrized by the positive reals. Lemmas 3 and 4 establish technical relationships between f and the elements of these classes. Lemmas 3 and 4 yield the results.

LEMMA 1. *Let $f \in A$. Then the following inequalities are valid*:

$$d(f, B) \leqslant (1/2) \inf_{\pi \in \mathcal{P}} \sup\{O(f, E); E \in \pi, \mu(E) > 0\}$$
$$\leqslant (1/2) \inf_{\pi \in \mathcal{F}} \sup\{O(f, E); E \in \pi, \mu(E) > 0\} \leqslant d(f, B),$$

so these inequalities are equalities.

Received by the editors November 30, 1979.
1980 *Mathematics Subject Classification*. Primary 41A50, 46E30.
Key words and phrases. Best approximation, L_∞, L_p-space, probability space.

PROOF. The first two inequalities are clear; to verify the third, let $\varepsilon > 0$, let $g \in B$ such that $d(f, g) = \|f - g\|_\infty < d(f, B) + \varepsilon/2$, and let $s = \sum_{i=1}^n c_i I_{E_i}$ be a simple B measurable function (i.e., $\{E_i\}_{i=1}^n \in \mathcal{F}$ and I_E denotes the indicator function of a set E) such that $d(g, s) < \varepsilon/2$. Then $d(f, s) < d(f, B) + \varepsilon$, so $\sup\{O(f, E_i);\ i \leqslant n,\ \mu(E_i) > 0\} < 2(d(f, B) + \varepsilon)$.

Henceforth f is a fixed element of A. Without loss of generality we suppose that $0 \leqslant f \leqslant 1$.

LEMMA 2. *For $h > 0$ and $\pi \in \mathcal{P}$, let $\delta(h, \pi) = \delta(f, h, \pi) = \{\sum \mu(E);\ E \in \pi,\ O(f, E) \geqslant h\}$ and let $\delta_h = \inf\{\delta(h, \pi);\ \pi \in \mathcal{P}\}$. Then there exists π such that $\delta_h = \delta(h, \pi)$.*

PROOF. For $\pi \in \mathcal{P}$, let $\pi^h = \{E \in \pi;\ O(f, E) \geqslant h\}$, and let $E_\pi^h = \{\bigcup E;\ E \in \pi^h\}$. Let π_k satisfy $\mu(E_{\pi_k}^h) < \delta_h + 2^{-k}$, $k \geqslant 1$. Let

$$\gamma_1 = \left\{F_{1i} = E_{1i};\ E_{1i} \in \pi_1 - \pi_1^h\right\},$$
$$\gamma_2 = \left\{F_{2i} = E_{2i} \cap E_{\pi_1}^h;\ E_{2i} \in \pi_2 - \pi_2^h\right\},$$
$$\gamma_3 = \left\{F_{3i} = E_{3i} \cap E_{\pi_1}^h \cap E_{\pi_2}^h;\ E_{3i} \in \pi_3 - \pi_3^h\right\}, \ldots.$$

Let $\pi = \{F_{ki}\} \cup \{\Omega - \bigcup_{k,i} F_{k,i}\}$.

Lemmas 1 and 2 assure (i) $\delta_h = 0$ if $h > 2d(f, B)$ and (ii) if $h < 2d(f, B)$, then there exists $\pi \in \mathcal{P}$ such that $\delta(h, \pi) = \delta_h > 0$. Notice also that if $\delta(h, \pi) = \delta_h$, $E \in \mathcal{B}$, $E \subset E_\pi^h$ and $\mu(E) > 0$, then $O(f, E) \geqslant h$; thus, E_π^h is uniquely determined up to a set of measure zero by the equation $\delta(h, \pi) = \delta_h$, so we can denote it by E_h. Now observe that if $h_1 < h_2$, then $\mu(E_{h_2} - E_{h_1}) = 0$.

For $E \in \mathcal{A}$, let $m(E) = (1/2)\{\operatorname{essup}(f, E) + \operatorname{essinf}(f, E)\}$.

We have two cases to consider at this point: let $h = 2d(f, B)$; then (1) $\delta_h = 0$ or (2) $\delta_h > 0$. If $\delta_h = 0$, then we can get a best approximation b to f in B by putting $b(x) = m(E)$, $x \in E \in \pi \in \mathcal{P}$, where $\delta(h, \pi) = \delta_h = 0$. On the other hand, suppose that $\delta(h, \pi) = \delta_h > 0$. Then for $E \in \pi$ with $O(f, E) < h$, define b on E by $b|E = m(E)$; it remains to define b on E_π^h. To this end, let $h_n = 2d(f, B) + 2^{-n}$, and let $\pi_n \in \mathcal{P}$ with $\mu(E_{\pi_n}^{h_n}) = 0$.

Let $F = \bigcup \{E_{\pi_n}^{h_n};\ n \geqslant 1\}$, let $H = E_\pi^h - F$, and let $E(\{i_j\}_n) = E_{1,i_1} \cap E_{2,i_2} \cap \cdots \cap E_{n,i_n}$, where $E_{j,i_j} \in \pi_j$. Next define f_n: $H \to R$ by $f_n(x) = m(E(\{i_j\}_n))$ if $x \in H \cap E(\{i_j\}_n)$. Now define b on H by $b(x) = \lim_n f_n(x)$; since $\mu(F) = 0$, $b \in B$ and $d(f, b) = d(f, B)$.

For $1 < p < \infty$, let $A_p = L_p(\Omega, \mathcal{A}, \mu)$, $B_p = L_p(\Omega, \mathcal{B}, \mu)$, $d_p(g, h) = \|g - h\|_p$; let f_p denote the best L_p approximation to f by elements of B_p. We shall show that f_p converges a.e. as $p \to \infty$ to a best L_∞ approximation f_∞ to f by using the following two technical lemmas.

LEMMA 3. *Suppose $0 < h_1$, $F \in \mathcal{B}$, $F \subset E_{h_1}$, $\mu(F) > 0$, $h = O(f, F)$. (Then $h_1 \leqslant h$.) Let $l_0 = \operatorname{essinf}(f, F)$, $u_0 = \operatorname{essup}(f, F)$, $\lambda = (h - h_1) + \varepsilon$, where $\varepsilon > 0$; let $l = l_0 + \lambda$, $u = u_0 - \lambda$, $L = \{x \in F;\ f(x) \leqslant l\}$ and $U = \{x \in F;\ f(x) \geqslant u\}$. Let $\alpha > 0$. There exists $\beta > 0$ such that if $H \in \mathcal{B}$, $H \subset F$ and $\mu(H) \geqslant \alpha$, then $\mu(H \cap L) \geqslant \beta\mu(H)$ and $\mu(H \cap U) \geqslant \beta\mu(H)$.*

PROOF. If we establish the first inequality, then the second follows by symmetry. To establish the first, suppose, on the contrary, that we can find sets $E_k \in \mathfrak{B}$ with $E_k \subset F$, $\mu(E_k) \geqslant \alpha$ and $\mu(E_k \cap L) \leqslant 2^{-k}\mu(E_k)$. Let $E = \limsup_k E_k$. Then $\mu(E) \geqslant \alpha$. But $\sum \mu(E_k \cap L) \leqslant 1$, so $\mu(E \cap L) = 0$ which implies the contradiction $O(f, E) < h_1$.

Define $\bar{f}$ by $\bar{f}(x) = \limsup_{p\to\infty} f_p(x)$, $x \in \Omega$, and let $\underline{f}$ denote the corresponding lim inf.

Let $0 < 2\gamma < h_1$, and let $h_2 = h_1 + \gamma$. Let $\delta(h_2, \pi) = \delta_{h_2}$, where $\pi = \{E_i\} \cup E_{h_2}$ and $O(f, E_i) < h_2$. Let $F_i = E_i \cap E_{h_1}$.

LEMMA 4. *$\bar{f} - \underline{f} \leqslant 2\gamma$ a.e. on $E_{h_1} - E_{h_2}$.*

PROOF. Notice that it suffices to show that if $\mu(F_i) > 0$ and $\varepsilon > 0$, then $\mu(\{x \in F_i; \bar{f}(x) - \underline{f}(x) \geqslant 2\gamma + 6\varepsilon\}) < \varepsilon$ as follows. Without loss of generality, we simplify the notation by fixing i, letting F denote F_i, supposing that $\mu(F) = 1$ and letting $\varepsilon > 0$ satisfy $2\gamma + 8\varepsilon < h = u_0 - l_0$ (cf. Lemma 3). Let m be a positive integer with $2^{-m} < \varepsilon$. For $1 < p < \infty$, let $G_{p,j} = [f_p \geqslant (j-1)/2^m] - [f_p \geqslant j/2^m]$, $1 \leqslant j < 2^m$, $G_{p,2^m} = [f_p \geqslant 1 - 2^{-m}]$ and $H_{p,j} = F \cap G_{p,j}$. Notice that $\sum\{\mu(H_{p,j}); \mu(H_{p,j}) < 4^{-m}\} < 2^{-m}$. Referring to Lemma 3, let β correspond to $\alpha = 4^{-m}$. Consider j with $\mu(H_{p,j}) \geqslant \alpha$. Simplifying the notation again, let H denote $H_{p,j}$. Then $O(f_p, H) < \varepsilon$ and according to Lemma 3, the numbers $\mu(H)$, $\mu(\{x \in H; f(x) \leqslant l_0 + \gamma + \varepsilon\})$ and $\mu(\{x \in H; f(x) \geqslant u_0 - \gamma - \varepsilon\})$ are balanced, independent of p. Let $m_0 = (1/2)(u_0 + l_0)$ and observe that for large p if $f_p(H)$ were not a subset of $[m_0 - \gamma - 3\varepsilon, m_0 + \gamma + 3\varepsilon]$, then there would be a better L_p approximation to f on H. Thus, $\bar{f} - \underline{f} \leqslant 2\gamma$ a.e. on $E_{h_1} - E_{h_2}$.

THEOREM 1. *Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $\mathfrak{B}$ be a subsigma-algebra of $\mathcal{A}$. Let $A = L_\infty(\Omega, \mathcal{A}, \mu)$ and let $B = L_\infty(\Omega, \mathfrak{B}, \mu)$. Let $f \in A$, and for $1 < p < \infty$, let f_p denote the best L_p approximation to f by elements of $L_p(\Omega, \mathfrak{B}, \mu)$. Then $\lim_{p\to\infty} f_p$ exists a.e.; moreover, the function f_∞ is a best L_∞ approximation to f by elements of B.*

PROOF. Let $0 < \varepsilon < 4^{-1}$; let $h_1 = \varepsilon$ and let $h_{n+1} = h_n + \varepsilon/4$, $n \geqslant 1$. Then Lemma 4 asserts that $\bar{f} - \underline{f} \leqslant \varepsilon/2$ a.e. on E_ε. Let $\varepsilon \to 0$ and recalling that $f_p = f$ on $\Omega - U_n E_{1/n}$, we have that f_p converges a.e. to a function f_∞. It is clear that f_∞ is a best L_∞ approximation to f.

Not only is f_∞ a best L_∞ approximation to f by elements of B on Ω, but for each set E in $\mathfrak{B}$, the restrictions to E of the functions considered above maintain their relationships: $f_p|E \to f_\infty|E$ a.e. and $f_\infty|E$ is a best L_∞ approximation to $f|E$ by elements of $B|E$. We verify that this latter relationship characterizes f_∞ below.

LEMMA 5. *If each of g and u is a best L_∞ approximation to f by elements of B and for each $E \in \mathfrak{B}$ each of $g|E$ and $u|E$ is a best L_∞ approximation to $f|E$ by elements of $B|E$, then $d(g, u) = 0$.*

PROOF. Let $h > 0$ and $0 < \varepsilon < h/8$. It suffices to show that $|g - u| \leqslant 4\varepsilon$ a.e. on $E_h - E_{h+\varepsilon}$ as follows. Suppose $E \subset (E_h - E_{h+\varepsilon})$ with $\mu(E) > 0$, $O(g, E) < \varepsilon$,

$O(u, E) < \varepsilon$ and $O(f, E) \leqslant h + \varepsilon$. Since each of $g|E$ and $u|E$ is a best L_∞ approximation of $f|E$, refer to Lemma 3 and find out that $|g - u| \leqslant 4\varepsilon$ a.e. on E. Then, refer to the proof of Lemma 4 and infer the desideratum.

Lemma 5 and the paragraph that precedes it combine with the following definition to characterize f_∞.

DEFINITION. Let $f \in A$. Then $g \in B$ is said to be a best best L_∞ approximation to f by elements of B if for each $E \in \mathfrak{B}$, the restriction, $g|E$, of g to E is a best L_∞ approximation to the restriction, $f|E$, of f to E.

THEOREM 2. *Let $f \in A$. Then f_∞ is the unique (up to a set of measure zero) best best L_∞ approximation to f by elements of B.*

REFERENCES

1. C. K. Chui, P. W. Smith and J. D. Ward, *Favard's solution is the limit of $W^{k,p}$-splines*, Trans. Amer. Math. Soc. **220** (1976), 299–305.

2. Jean Descloux, *Approximations in L^p and Chebyshev approximations*, J. Soc. Indust. Appl. Math. **11** (1963), 1017–1026.

3. L. A. Karlovitz, *Construction of nearest points in the L^p, p even and L^∞ norms*. I, J. Approximation Theory **3** (1970), 123–127.

4. G. Polya, *Sur un algorithme toujours convergent pour obtenir les polynomes de meilleure approximation de Tchebycheff pour une fonction continue quelconque*, Compt. Rend. **157** (1913), 480–483.

5. J. R. Rice, *The approximation of functions*. II, Addison-Wesley, Reading, Mass., 1969.

6. V. A. Ubhaya, *Isotone optimization*. II, J. Approximation Theory **12** (1974), 315–331.

DEPARTMENT OF MATHEMATICS, COLORADO STATE UNIVERSITY, FORT COLLINS, COLORADO 80523