Z. Wahrscheinlichkeitstheorie verw. Gebiete 33, 33 – 39 (1975)

Best Approximants in L^1 Space

T. Shintani[1] and T. Ando[2]

[1] Department of Mathematics, Tomakomai Technical College, Tomakomai, Hokkaido, Japan
[2] Division of Applied Mathematics, Research Institute of Applied Electricity, Hokkaido University, Sapporo, Japan

1. Introduction

Consider the Banach space L^1 of real-valued integrable functions on a probability space $(\Omega, \mathscr{A}, P)$. Given a σ-subalgebra $\mathscr{B}$ of the σ-algebra $\mathscr{A}$, let $L^1(\mathscr{B})$ stand for the subspace of all $\mathscr{B}$-measurable functions. This paper concerns the best approximation of a function f in L^1 by functions in $L^1(\mathscr{B})$. A function in $L^1(\mathscr{B})$ that has the minimum distance from f is called a *best approximant* of f. The existence of a best approximant is not trivial.

In the next section best approximants are characterized as $\mathscr{B}$-conditional medians of f. The existence of the maximum $U_{\mathscr{B}} f$ and the minimum $V_{\mathscr{B}} f$ of all best approximants are guaranteed and the distance from f to $L^1(\mathscr{B})$ is provided with a convenient expression. Further every sequence in $L^1(\mathscr{B})$ that minimizes asymptotically the distance to f is shown to be relatively weakly compact.

In the final section convergence of a sequence $\{\mathscr{B}_n\}$ of σ-subalgebras to a σ-subalgebra $\mathscr{B}_\infty$ is taken into consideration. Two kinds of convergence, strong convergence and almost everywhere one, are defined in connection with the corresponding convergences of conditional expectations. If $\mathscr{B}_n$ converges strongly to $\mathscr{B}_\infty$ and if each g_n is a best approximant in $L^1(\mathscr{B}_n)$ of f, then the sequence $\{g_n\}$ is shown to be relatively weakly compact and every weak limiting function of the sequence becomes a best approximant in $L^1(\mathscr{B}_\infty)$. Further the inequalities

$$V_{\mathscr{B}_\infty} f \leqq \liminf_{n\to\infty} g_n \leqq \limsup_{n\to\infty} g_n \leqq U_{\mathscr{B}_\infty} f$$

are proved under the almost everywhere convergence of the sequence $\{\mathscr{B}_n\}$.

2. Best Approximants

Given measurable functions g, h, let us use the following notations:

$$g \vee h = \max(g, h), \qquad g \wedge h = \min(g, h)$$

and

$$g^+ = g \vee 0, \qquad g^- = (-g) \vee 0.$$

If g coincides with h almost everywhere, then they are identified. A constant function with value α is denoted by the same letter α, while the indicator of a set A is denoted by I_A.

In this section $\mathscr{B}$ is a fixed σ-subalgebra and f is a fixed integrable function. The distance from f to the subspace $L^1(\mathscr{B})$ is denoted by $d(f, \mathscr{B})$, that is,

$$d(f, \mathscr{B}) = \inf\{\int |f-g|\, dP;\ g \in L^1(\mathscr{B})\}.$$

Then a function g in $L^1(\mathscr{B})$ is, by definition, a *best approximant* of f if

$$\int |f-g|\, dP = d(f, \mathscr{B}).$$

In the study of best approximants a crucial role will be played by the *$\mathscr{B}$-conditional expectation* operator $E_{\mathscr{B}}$. For convenience, the $\mathscr{B}$-conditional expectation of the indicator of the set $\{f \leqq g\}$ (resp. $\{f < g\}$) will be denoted by $P_{\mathscr{B}}(f \leqq g)$ (resp. $P_{\mathscr{B}}(f < g)$).

Observe first that the distance from f to the space of constant functions is attained by a constant (function) α if and only if α is a *median* of f in the sense that

$$P(f < \alpha) \leqq \tfrac{1}{2} \quad \text{and} \quad P(\alpha < f) \leqq \tfrac{1}{2}.$$

Motivated by this fact, let us consider $\mathscr{B}$-conditional medians (cf. [5, § 29]); a $\mathscr{B}$-measurable function g is called a *$\mathscr{B}$-conditional median* of f if

$$P_{\mathscr{B}}(f < g) \leqq \tfrac{1}{2} \quad \text{and} \quad P_{\mathscr{B}}(g < f) \leqq \tfrac{1}{2}.$$

$\mathscr{B}$-conditional medians are conveniently treated by introducing the *$\mathscr{B}$-conditional distribution function* $F(\omega, \lambda)$ of f (cf. [5, § 27]); $F(\omega, \lambda)$ is defined as the jointly measurable function on $\Omega \times (-\infty, \infty)$ such that for λ fixed, $F(\cdot, \lambda) = P_{\mathscr{B}}(f \leqq \lambda)$ and for ω fixed, $F(\omega, \cdot)$ is increasing and right-continuous with

$$\lim_{\lambda \to -\infty} F(\omega, \lambda) = 0 \quad \text{and} \quad \lim_{\lambda \to +\infty} F(\omega, \lambda) = 1.$$

Define the $\mathscr{B}$-measurable functions $U_{\mathscr{B}} f$ and $V_{\mathscr{B}} f$ by the formulas:

$$U_{\mathscr{B}} f(\omega) = \inf\{\lambda;\ \tfrac{1}{2} < F(\omega, \lambda)\}$$

and

$$V_{\mathscr{B}} f(\omega) = \sup\{\lambda;\ \tfrac{1}{2} > F(\omega, \lambda)\}.$$

Then it is immediate that $U_{\mathscr{B}} f$ (resp. $V_{\mathscr{B}} f$) is the maximum (resp. minimum) of all $\mathscr{B}$-measurable functions g such that

$$P_{\mathscr{B}}(f < g) \leqq \tfrac{1}{2} \quad (\text{resp. } P_{\mathscr{B}}(g < f) \leqq \tfrac{1}{2}).$$

In particular, $U_{\mathscr{B}} f$ and $V_{\mathscr{B}} f$ are the maximum and the minimum $\mathscr{B}$-conditional median of f respectively.

Lemma 1. *The non-linear operators $U_{\mathscr{B}}$ and $V_{\mathscr{B}}$ have the following properties.*

(a) $V_{\mathscr{B}}(-f) = -U_{\mathscr{B}} f$ *and* $V_{\mathscr{B}} f \leqq U_{\mathscr{B}} f$.

(b) $|U_{\mathscr{B}} f| \leqq 2 E_{\mathscr{B}} |f|$ *and* $|V_{\mathscr{B}} f| \leqq 2 E_{\mathscr{B}} |f|$.

Proof. (a) is immediate from definition. To prove (b), suppose first $f \geqq 0$, hence obviously $U_{\mathscr{B}} f \geqq 0$ by definition. The inclusion

$$\{\lambda < U_{\mathscr{B}} f\} \subseteqq \{\tfrac{1}{2} \leqq P_{\mathscr{B}}(\lambda < f)\}$$

implies

$$P(\lambda < U_{\mathscr{B}} f) \leqq 2\, P(\lambda < f),$$

and consequently

$$\int U_{\mathscr{B}} f\, dP = \int_0^\infty P(\lambda < U_{\mathscr{B}} f)\, d\lambda$$

$$\leqq 2 \int_0^\infty P(\lambda < f)\, d\lambda = 2 \int f\, dP.$$

Since the definition of $U_{\mathscr{B}}$ shows

$$U_{\mathscr{B}}(I_A \cdot f) = I_A \cdot U_{\mathscr{B}} f \quad \text{whenever } A \in \mathscr{B}$$

the above consideration yields

$$\int_A U_{\mathscr{B}} f\, dP \leqq 2 \int_A f\, dP \quad \text{for all } A \in \mathscr{B}$$

which is equivalent to

$$U_{\mathscr{B}} f \leqq 2\, E_{\mathscr{B}} f.$$

Now the proof for general f can be derived on the basis of the relation:

$$(U_{\mathscr{B}} f)^+ \leqq U_{\mathscr{B}} f^+.$$

In fact,

$$(U_{\mathscr{B}} f)^+ \leqq U_{\mathscr{B}} f^+ \leqq 2\, E_{\mathscr{B}} f^+ \leqq 2\, E_{\mathscr{B}} |f|$$

and by (a)

$$(U_{\mathscr{B}} f)^- = (V_{\mathscr{B}}(-f))^+ \leqq (U_{\mathscr{B}}(-f))^+ \leqq 2\, E_{\mathscr{B}} |f|.$$

The assertion with $V_{\mathscr{B}}$ instead of $U_{\mathscr{B}}$ follows from (a).

Theorem 2. *A function g in $L^1(\mathscr{B})$ is a best approximant of f if and only if*

$$V_{\mathscr{B}} f \leqq g \leqq U_{\mathscr{B}} f.$$

In particular, $U_{\mathscr{B}} f$ and $V_{\mathscr{B}} f$ are the maximum and the minimum best approximant respectively.

Proof. Remark the obvious relation for $h_1, h_2 \in L^1$:

$$|f - h_2| - |f - h_1| = \int_{h_1(\cdot)}^{h_2(\cdot)} \{2 I_{[f(\cdot), \infty)}(\lambda) - 1\}\, d\lambda \quad \text{a.e. on } \{h_1 \leqq h_2\}.$$

If both h_1 and h_2 are $\mathscr{B}$-measurable, the $\mathscr{B}$-conditioning of the above relation yields

$$E_{\mathscr{B}} |f - h_2| - E_{\mathscr{B}} |f - h_1| = \int_{h_1(\cdot)}^{h_2(\cdot)} \{2 F(\cdot, \lambda) - 1\}\, d\lambda \quad \text{a.e. on } \{h_1 \leqq h_2\},$$

where $F(\omega, \lambda)$ is the $\mathscr{B}$-conditional distribution function of f. Since the definition of $U_{\mathscr{B}} f$ and $V_{\mathscr{B}} f$ implies that

$$2F(\cdot, \lambda)-1 \begin{cases} >0 & \text{a.e. on } \{U_{\mathscr{B}} f<\lambda\}, \\ =0 & \text{a.e. on } \{V_{\mathscr{B}} f<\lambda<U_{\mathscr{B}} f\}, \\ <0 & \text{a.e. on } \{\lambda<V_{\mathscr{B}} f\} \end{cases}$$

and since $U_{\mathscr{B}} f$ and $V_{\mathscr{B}} f$ belong to $L^1(\mathscr{B})$ by Lemma 1, it follows that for g in $L^1(\mathscr{B})$

$$E_{\mathscr{B}}|f-g|=E_{\mathscr{B}}|f-U_{\mathscr{B}} f| \quad \text{a.e. on } \{V_{\mathscr{B}} f \leqq g \leqq U_{\mathscr{B}} f\}$$

and

$$E_{\mathscr{B}}|f-g|>E_{\mathscr{B}}|f-U_{\mathscr{B}} f| \quad \text{a.e. on } \{U_{\mathscr{B}} f<g\} \cup\{g<V_{\mathscr{B}} f\},$$

which implies obviously the assertion.

Corollary 3. *A function g in $L^1(\mathscr{B})$ is a best approximant of f if and only if it is a $\mathscr{B}$-conditional median of f.*

In an unpublished paper [4] Kudo constructed a $\mathscr{B}$-conditional median of f which is a best approximant. The following corollary was also established by Kudo [3,4] in its primitive form.

Corollary 4. *The distance from f to $L^1(\mathscr{B})$ is explicitly given by*

$$d(f, \mathscr{B})=\int_{-\infty}^{\infty}\{\tfrac{1}{2}-\int|\tfrac{1}{2}-P_{\mathscr{B}}(f \leqq \lambda)| dP\} d\lambda.$$

Proof. Since $U_{\mathscr{B}} f$ is a best approximant by Theorem 2,

$$d(f, \mathscr{B})=\int|f-U_{\mathscr{B}} f| dP.$$

Now the obvious relation

$$|f-U_{\mathscr{B}} f|=\int_{-\infty}^{\infty}|I_{[f(\cdot), \infty)}(\lambda)-I_{[U_{\mathscr{B}} f(\cdot), \infty)}(\lambda)| d\lambda$$

yields through the $\mathscr{B}$-conditioning and the Fubini theorem

$$\int|f-U_{\mathscr{B}} f| dP=\int E_{\mathscr{B}}|f-U_{\mathscr{B}} f| dP=\int_{\infty}^{\infty} \int|P_{\mathscr{B}}(f \leqq \lambda)-I_{\{U_{\mathscr{B}} f \leqq \lambda\}}| dP\, d\lambda.$$

Then the assertion follows from the relation

$$|P_{\mathscr{B}}(f \leqq \lambda)-I_{\{U_{\mathscr{B}} f \leqq \lambda\}}|=\tfrac{1}{2}-|\tfrac{1}{2}-P_{\mathscr{B}}(f \leqq \lambda)|$$

which is an immediate consequence of the definition of $U_{\mathscr{B}} f$.

The dual of the Banach space L^1 is canonically realized by the space L^∞ of essentially bounded measurable functions (cf. [6; §4.2]). The *weak* topology is always understood with respect to the pairing (L^1, L^∞).

Since the unit ball of the Banach space L^1 is not weakly compact, the following theorem is not trivial in contrast to L^p-approximation $(1<p<\infty)$ (cf. [1]).

Theorem 5. *Every sequence $\{g_n\}$ in $L^1(\mathscr{B})$ such that*

$$\lim_{n \rightarrow \infty} \int|f-g_n| dP=d(f, \mathscr{B})$$

is relatively weakly compact and every weak limiting function is a best approximant of f.

Proof. Since the subspace $L^1(\mathscr{B})$ is isometric to the Banach space $L^1(\Omega, \mathscr{B}, P)$ in the canonical way and since the sequence $\{g_n\}$ is bounded in norm, the relative weak compactness will follow from its *equi-continuity* in $L^1(\Omega, \mathscr{B}, P)$ (cf. [6, §4.2]):

$$\lim_{n\to\infty} \int_{A_n} |g_n|\, dP = 0 \quad \text{whenever } A_n \in \mathscr{B} \quad \text{and} \quad \lim_{n\to\infty} P(A_n) = 0.$$

The equi-continuity in question is proved as follows. Since

$$\int_{A_n} |g_n|\, dP + \int |f - I_{A_n^c} \cdot g_n|\, dP \leqq \int |f - g_n|\, dP + 2 \int_{A_n} |f|\, dP$$

and

$$\lim_{n\to\infty} \int_{A_n} |f|\, dP = 0,$$

the hypothesis on $\{g_n\}$ implies

$$\limsup_{n\to\infty} \int_{A_n} |g_n|\, dP + \inf_n \int |f - I_{A_n^c} \cdot g_n|\, dP \leqq d(f; \mathscr{B}).$$

On the other hand, since each $I_{A_n^c} \cdot g_n$ is $\mathscr{B}$-measurable, the definition of the distance $d(f, \mathscr{B})$ implies

$$\inf_n \int |f - I_{A_n^c} \cdot g_n|\, dP \geqq d(f, \mathscr{B}).$$

These together yield the convergence of $\int_{A_n} |g_n|\, dP$ to 0, as expected. Finally each weak limiting function g of the sequence belongs to $L^1(\mathscr{B})$ and

$$d(f, \mathscr{B}) \leqq \int |f - g|\, dP \leqq \limsup_{n\to\infty} \int |f - g|\, dP = d(f, \mathscr{B}).$$

Corollary 6. *If* $g_n \in L^1(\mathscr{B})$ *and*

$$\lim_{n\to\infty} \int |f - g_n|\, dP = d(f, \mathscr{B})$$

then the sequence $\{g_n \vee U_{\mathscr{B}} f\}$ *(resp.* $\{g_n \wedge V_{\mathscr{B}} f\}$*) converges to* $U_{\mathscr{B}} f$ *(resp.* $V_{\mathscr{B}} f$*) in* L^1*-norm.*

Proof. For each n the set $A_n = \{g_n > U_{\mathscr{B}} f\}$ belongs to the σ-subalgebra $\mathscr{B}$ and

$$g_n \vee U_{\mathscr{B}} f = I_{A_n} \cdot g_n + I_{A_n^c} \cdot U_{\mathscr{B}} f$$

and

$$g_n \wedge U_{\mathscr{B}} f = I_{A_n^c} \cdot g_n + I_{A_n} \cdot U_{\mathscr{B}_n} f.$$

Then the relation

$$\begin{aligned} \int_{A_n} |f - U_{\mathscr{B}} f|\, dP + \int_{A_n^c} |f - U_{\mathscr{B}} f|\, dP &= \int |f - U_{\mathscr{B}} f|\, dP = d(f, \mathscr{B}) \\ &\leqq \int |f - g_n \wedge U_{\mathscr{B}} f|\, dP = \int_{A_n} |f - U_{\mathscr{B}} f|\, dP + \int_{A_n^c} |f - g_n|\, dP, \end{aligned}$$

implies

$$\int_{A_n^c} |f - U_{\mathscr{B}} f| \, dP \leqq \int_{A_n^c} |f - g_n| \, dP,$$

and consequently

$$d(f, \mathscr{B}) \leqq \int |f - g_n \vee U_{\mathscr{B}} f| \, dP \leqq \int |f - g_n| \, dP.$$

Therefore the hypothesis on $\{g_n\}$ implies

$$\lim_{n \to \infty} \int |f - g_n \vee U_{\mathscr{B}} f| \, dP = d(f, \mathscr{B}).$$

Now in view of Theorem 5 and Theorem 2 the sequence $\{g_n \vee U_{\mathscr{B}} f\}$ is relatively weakly compact and every weak limiting function is majorated by $U_{\mathscr{B}} f$. This leads to the conclusion:

$$0 \leqq \limsup_{n \to \infty} \int |g_n \vee U_{\mathscr{B}} f - U_{\mathscr{B}} f| \, dP = \limsup_{n \to \infty} \int (g_n \vee U_{\mathscr{B}} f) \, dP - \int U_{\mathscr{B}} f \, dP \leqq 0.$$

The assertion on $\{g_n \wedge V_{\mathscr{B}} f\}$ is proved similarly.

3. Convergence

In this section $\mathscr{B}_1, \mathscr{B}_2, \ldots, \mathscr{B}_\infty$ are σ-subalgebras. The sequence $\{\mathscr{B}_n\}$ is said to converge *strongly* (resp. *almost everywhere*) to $\mathscr{B}_\infty$ if for every function f in L^1 the sequence $\{E_{\mathscr{B}_n} f\}$ converges in L^1-norm (resp. almost everywhere) to $E_{\mathscr{B}_\infty} f$. Strong convergence results from less restrictive conditions. Indeed, Kudo [3] proved that $\{\mathscr{B}_n\}$ converges strongly to $\mathscr{B}_\infty$ if for every f in L^1

$$\lim_{n \to \infty} \int |E_{\mathscr{B}_n} f| \, dP = \int |E_{\mathscr{B}_\infty} f| \, dP$$

while Becker [2] pointed out that the strong convergence is a consequence of the condition that for every f in L^1 the sequence $\{E_{\mathscr{B}_n} f\}$ converges weakly to $E_{\mathscr{B}_\infty} f$.

Almost everywhere convergence implies strong convergence. In view of the Martingale theorem (cf. [6; §4.5]) every monotone sequence of σ-subalgebras is almost everywhere convergent.

Convergence problems of best approximants in L^p-version $(1 < p < \infty)$ were discussed in a previous paper [1].

Theorem 7. *Suppose that the sequence $\{\mathscr{B}_n\}$ converges strongly to $\mathscr{B}_\infty$. If each g_n is a best approximant in $L^1(\mathscr{B}_n)$ of one and the same function f, then the sequence $\{g_n\}$ is relatively weakly compact and the sequence $\{g_n \vee U_{\mathscr{B}_\infty} f\}$ (resp. $\{g_n \wedge V_{\mathscr{B}_\infty} f\}$) converges to $U_{\mathscr{B}_\infty} f$ (resp. $V_{\mathscr{B}_\infty} f$) in L^1-norm. Further every weak limiting function of the sequence $\{g_n\}$ is a best approximant in $L^1(\mathscr{B}_\infty)$ of f.*

Proof. Since in view of Theorem 2 and Lemma 1

$$|g_n| \leqq U_{\mathscr{B}_n} f \vee (-V_{\mathscr{B}_n} f) \leqq 2 E_{\mathscr{B}_n} |f|,$$

the boundedness and the equi-continuity of the sequence $\{g_n\}$ follow from the convergence of $E_{\mathscr{B}_n} |f|$ in L^1-norm. Therefore the sequence $\{g_n\}$ is relatively weakly compact (cf. [6, §4.2]).

The convergence of $g_n \vee U_{\mathscr{B}_\infty} f$ to $U_{\mathscr{B}_\infty} f$ in L^1-norm will follow from the convergence in measure and the above proved equi-continuity. To prove the convergence in measure, it suffices to show that each sequence of integers contains a subsequence $n(1) < n(2) < \cdots$ such that

$$\limsup_{m\to\infty} U_{\mathscr{B}_{n(m)}} f \leqq U_{\mathscr{B}_\infty} f.$$

To this end, consider the $\mathscr{B}_n$-conditional distribution function $F_n(\omega, \lambda)$ of f. Since

$$F_n(\cdot, \lambda) = P_{\mathscr{B}_n}(f \leqq \lambda),$$

the strong convergence of $\{\mathscr{B}_n\}$ implies that for every λ the sequence $\{F_n(\cdot, \lambda)\}$ converges to $F_\infty(\cdot, \lambda)$ in L^1-norm. Therefore it is possible to choose a subsequence $n(1) < n(2) < \cdots$ such that

$$\lim_{m\to\infty} F_{n(m)}(\cdot, \lambda) = F_\infty(\cdot, \lambda)$$

for all rational numbers λ. The inspection of the definition of $U_{\mathscr{B}_n} f$ shows that this subsequence meets the requirement. The convergence of $g_n \wedge V_{\mathscr{B}_\infty} f$ to $V_{\mathscr{B}_\infty} f$ in L^1-norm is similarly established.

Finally suppose that a subsequence $\{g_{n(m)}\}$ converges weakly to g_∞, say. The above argument yields

$$V_{\mathscr{B}_\infty} f \leqq g_\infty \leqq U_{\mathscr{B}_\infty} f.$$

Since for each function h in L^∞ the sequence $\{E_{\mathscr{B}_n} h\}$ converges in measure to $E_{\mathscr{B}_\infty} h$ by hypothesis, the equi-continuity of $\{g_n\}$ implies

$$\int h \cdot g_\infty \, dP = \lim_{m\to\infty} \int h \cdot g_{n(m)} \, dP = \lim_{m\to\infty} \int E_{\mathscr{B}_{n(m)}} h \cdot g_{n(m)} \, dP = \int E_{\mathscr{B}_\infty} h \cdot g_\infty \, dP$$
$$= \int h \cdot E_{\mathscr{B}_\infty} g_\infty \, dP,$$

hence $g_\infty = E_{\mathscr{B}_\infty} g_\infty$, that is, g_∞ is $\mathscr{B}_\infty$-measurable. Now g_∞ is a best approximant in $L^1(\mathscr{B}_\infty)$ by Theorem 2. This completes the proof.

The inspection of the above proof will show

Corollary 8. *If the sequence $\{\mathscr{B}_n\}$ converges almost everywhere to $\mathscr{B}_\infty$, then for every function f in L^1*

$$V_{\mathscr{B}_\infty} f \leqq \liminf_{n\to\infty} V_{\mathscr{B}_n} f \leqq \limsup_{n\to\infty} U_{\mathscr{B}_n} f \leqq U_{\mathscr{B}_\infty} f.$$

The authors thank the referee for valuable comments.

References

1. Ando, T., Amemiya, I.: Almost everywhere convergence of prediction sequences in L_p. Z. Wahrscheinlichkeitstheorie verw. Gebiete **4**, 113–120 (1965)
2. Becker, R.: Ensembles compacts de tribus. Z. Wahrscheinlichkeitstheorie verw. Gebiete **29**, 229–234 (1974)
3. Kudo, H.: A note on the strong convergence of σ-algebras. Ann. Probability **2**, 76–83 (1974)
4. Kudo, H.: An appendix to [3] (unpublished)
5. Loève, M.: Probability Theory, 2nd ed. Princeton: Van Nostrand 1961
6. Neveu, J.: Bases Mathematiques du Calcul des Probabilités, 2ième ed. Paris: Masson et C^ie 1970

Received April 4, 1975; in revised form June 15, 1975