

A CHARACTERIZATION OF BEST Φ -APPROXIMANTS

BY

D. LANDERS AND L. ROGGE

ABSTRACT. Let T be an operator from an Orlicz space L_Φ into itself. It is shown in this paper that four algebraic conditions and one integration condition assure that T is the best Φ -approximator, given a suitable σ -lattice.

Introduction and notations. Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space and $\Phi: \mathbf{R} \rightarrow [0, \infty)$ a strictly convex symmetric function with $\Phi(0) = 0$. Denote by $L_\Phi(\Omega, \mathcal{A}, \mu)$ respectively $L_\Phi^\infty(\Omega, \mathcal{A}, \mu)$ the system of all μ -equivalence classes of \mathcal{A} -measurable functions f such that $\int \Phi(\alpha f) d\mu < \infty$ for some $\alpha > 0$ respectively for all $\alpha > 0$. If $\Phi(x) = |x|^p$ then $L_\Phi = L_\Phi^\infty$ and we obtain the spaces L_p ($p > 1$).

Let $\mathcal{L} \subset \mathcal{A}$ be a σ -lattice, i.e. a system which is closed under countable unions and intersections and contains \emptyset and Ω . Denote by $L_\Phi(\mathcal{L})$ the system of all equivalence classes in $L_\Phi(\Omega, \mathcal{A}, \mu)$, containing a \mathcal{L} -measurable function, i.e. a function $f: \Omega \rightarrow \mathbf{R}$ with $\{f > a\} \in \mathcal{L}$ for all $a \in \mathbf{R}$. If $L_\Phi = L_\Phi^\infty$ it was shown in [11, Corollary 5], that for each $f \in L_\Phi(\Omega, \mathcal{A}, \mu)$ there exists a unique $\mu_\Phi(f|\mathcal{L}) \in L_\Phi(\mathcal{L})$ such that

$$(*) \quad \int \Phi(f - \mu_\Phi(f|\mathcal{L})) d\mu = \inf \left\{ \int \Phi(f - g) d\mu : g \in L_\Phi(\mathcal{L}) \right\}.$$

We call $\mu_\Phi(f|\mathcal{L})$ the best Φ -approximant of f , given \mathcal{L} . For $p > 1$ and $\Phi(x) = |x|^p$ we write $\mu_p(f|\mathcal{L})$ instead of $\mu_\Phi(f|\mathcal{L})$. The concept of best Φ -approximants, given \mathcal{L} , covers and unifies important concepts of probability theory: the classical concept of conditional expectation ($\Phi(x) = |x|^2$, \mathcal{L} σ -field), the concept of conditional expectation given a σ -lattice ($\Phi(x) = x^2$, \mathcal{L} σ -lattice, see [4], [5]), Ando-Amemiya's [2] concept of p -predictors ($\Phi(x) = |x|^p$, \mathcal{L} σ -algebra) or Brunk's [6] concept of conditional p -means ($\Phi(x) = |x|^p$, \mathcal{L} σ -lattice). Best approximants given σ -lattices play an essential role in several fields of application, e.g. in Bayes-estimation, prediction theory and isotonic regression (see [4]).

It is the aim of this paper to give conditions on an operator $T: L_\Phi \rightarrow L_\Phi$ assuring that T is the best Φ -approximator, given a suitable σ -lattice $\mathcal{L} \subset \mathcal{A}$, i.e. $Tf = \mu_\Phi(f|\mathcal{L})$ for all $f \in L_\Phi$.

Many authors have investigated such characterization problems for special cases. Characterization results for classical conditional expectations were given in [1], [3], [7], [12], [13], [14]; a characterization for a conditional expectation, given a

Received by the editors October 1, 1980.

AMS (MOS) subject classifications (1970). Primary 46E30; Secondary 41A50.

Key words and phrases. Best approximants, σ -lattices, conditional expectations, characterization.

© 1981 American Mathematical Society
0002-9947/81/0000-0417/\$02.50

σ -lattice, can be found in [8] and [12], a characterization for p -predictors is given in [10] and a characterization for conditional p -means is contained in [12]. Hence much is known for $\Phi(x) = x^2$, less for $\Phi(x) = |x|^p$ ($p > 1$) and nothing for general Φ . It is the purpose of this paper to close this gap. In our theorem we prove a characterization result for best Φ -approximants, given a σ -lattice and for best Φ -approximants, given a σ -field. Four of the five describing conditions are purely algebraic; only the fifth condition introduces the measure μ and the function Φ . We consider this an advantage since algebraic conditions are nicer and easier to verify than integration conditions.

The result. Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space, $\Phi: \mathbf{R} \rightarrow [0, \infty)$ strictly convex, symmetric with $\Phi(0) = 0$ and assume that $L_\Phi = L_\Phi^\infty$. Let $\mathcal{L} \subset \mathcal{A}$ be a σ -lattice and put $Tf := \mu_\Phi(f|\mathcal{L})$, $f \in L_\Phi$. Then the operator $T: L_\Phi \rightarrow L_\Phi$ has the following property.

- (i) T is translation invariant, i.e. $T(f + a) = Tf + a$ for all $a \in \mathbf{R}$.
- (ii) T is idempotent, i.e. $TTf = Tf$.
- (iii) T is monotone, i.e. $f \leq g$ implies $Tf \leq Tg$.

Properties (i) and (ii) follow directly from the definition of T , (iii) follows from [11, Theorem 18].

Properties (i)–(iii) were used in most of the cited characterization results. A further property, namely the positive homogeneity of T (i.e. $T(af) = aTf$ for $a \geq 0$), has been used in all characterizations up to now. Though $Tf = \mu_\Phi(f|\mathcal{L})$ is positive homogeneous if $\Phi(x) = |x|^p$, $p > 1$, unfortunately T is not positive homogeneous any more for general Φ . However, even in the general case T has another simple algebraic property, namely

$$(iv) \quad T(f \pm \frac{1}{2} Tf) = (1 \pm \frac{1}{2})Tf.$$

The case $T(f - \frac{1}{2} Tf) = \frac{1}{2}Tf$ follows directly from the definition of $Tf = \mu_\Phi(f|\mathcal{L})$. To see $T(f + \frac{1}{2} Tf) = \frac{3}{2}Tf$ let $h \in L_\Phi(\mathcal{L})$ be given. As $Tf = \mu_\Phi(f|\mathcal{L})$ and Φ is convex we have

$$\begin{aligned} \int \Phi\left(\left(f + \frac{1}{2} Tf\right) - \frac{3}{2} Tf\right) d\mu &= \int \Phi(f - Tf) d\mu < \int \Phi\left(f - \frac{2}{3} h\right) d\mu \\ &= \int \Phi\left(\frac{1}{3}(f - Tf) + \frac{2}{3}\left(f + \frac{1}{2} Tf - h\right)\right) d\mu \\ &\leq \frac{1}{3} \int \Phi(f - Tf) d\mu + \frac{2}{3} \int \Phi\left(f + \frac{1}{2} Tf - h\right) d\mu \end{aligned}$$

which implies $T(f + \frac{1}{2} Tf) = \frac{3}{2}Tf$.

The following theorem shows that if we add to the algebraic conditions (i)–(iv) on T only one integration condition on T (namely that 0 is not a better Φ -approximant of f than Tf), then each operator with these properties is necessarily of the form $Tf = \mu_\Phi(f|\mathcal{L})$ for an appropriate σ -lattice $\mathcal{L} \subset \mathcal{A}$. The added integration condition is a generalization of Dykstra’s condition for $\Phi(x) = x^2$ (see [8]); for $\Phi(x) = |x|^p$ this condition is also used in [10].

We remark that the following characterization result is the first characterization result for σ -lattices which uses only *one* integration condition. For the case

$\Phi(x) = x^2$ Dykstra [8] needs besides three algebraic conditions (namely (ii), (iii) and positive homogeneity) three integration conditions (namely (v), $\int Tf \, d\mu = \int f \, d\mu$ and $\|Tf - Tg\|_2 \leq \|f - g\|_2$). In [12] five algebraic conditions and two integration conditions are used for the case $\Phi(x) = |x|^p, p > 1$.

THEOREM 1. *Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space and $\Phi: \mathbf{R} \rightarrow \mathbf{R}_+$ be a symmetric, strictly convex, differentiable function with $\Phi(0) = 0$ such that $L_\Phi = L_\Phi^\infty$. Let $T: L_\Phi(\Omega, \mathcal{A}, \mu) \rightarrow L_\Phi(\Omega, \mathcal{A}, \mu)$ be an operator with the following properties.*

- (i) *T is translation invariant.*
- (ii) *T is idempotent.*
- (iii) *T is monotone.*
- (iv) $T(f \pm \frac{1}{2}Tf) = (1 \pm \frac{1}{2})Tf$.
- (v) $\int \Phi(f - Tf) \, d\mu \leq \int \Phi(f) \, d\mu$.

Then there exists a σ -lattice $\mathcal{L} \subset \mathcal{A}$ such that $Tf = \mu_\Phi(f|\mathcal{L})$ for all $f \in L_\Phi$. If, moreover, $T(-Tf) = -Tf$, then there exists a σ -algebra $\mathcal{A}_0 \subset \mathcal{A}$ such that $Tf = \mu_\Phi(f|\mathcal{A}_0)$ for all $f \in L_\Phi$.

PROOF. From the properties of Φ we obtain (see e.g. [9])

(1) Φ' is strictly increasing, continuous, $\Phi'(0) = 0$ and $\Phi'(-x) = -\Phi'(x)$.
Hence $0 \leq x\Phi'(x) \leq \int_x^{2x} \Phi'(t) \, dt \leq \Phi(2x), x \geq 0$. As $L_\Phi = L_\Phi^\infty$ this implies $h\Phi'(h) \in L_1$ if $0 \leq h \in L_\Phi$. For $h = |f| + |g|$ this yields together with (1)

(2)
$$\Phi'(f)g \in L_1 \quad \text{for all } f, g \in L_\Phi.$$

Applying $T(g - \frac{1}{2}Tg) = \frac{1}{2}Tg$ to $g = f - (1 - 1/2^n)Tf$ one obtains inductively

(3)
$$T(f - (1 - 1/2^n)Tf) = (1/2^n)Tf, \quad n \in \mathbf{N}.$$

As by (i) and (3) for all $a \in \mathbf{R}$

$$f - Tf = f - (1 - 1/2^n)Tf + a - T(f - (1 - 1/2^n)Tf + a)$$

we obtain from (v) that

(4)
$$\int \Phi(f - Tf) \, d\mu \leq \int \Phi(f - (1 - 1/2^n)Tf + a) \, d\mu, \quad n \in \mathbf{N}.$$

Hence $n \rightarrow \infty$ implies by the Theorem of Lebesgue

$$\int \Phi(f - Tf) \, d\mu \leq \int \Phi(f - Tf + a) \, d\mu, \quad \text{for all } a \in \mathbf{R}.$$

Therefore the function $a \rightarrow \int \Phi(f - Tf + a) \, d\mu, a \in \mathbf{R}$, attains its minimum at $a = 0$. Since $\sup_{|a| \leq 1} |\Phi'(f - Tf + a)| \leq \Phi'(|f - Tf| + 1) \in L_1$ by (1) and (2), we obtain from the Theorem of Lebesgue that

(5)
$$\int \Phi'(f - Tf) \, d\mu = 0.$$

In the next step we show

(6) $f_n \in L_\Phi, f_n \uparrow f \in L_\Phi \quad \text{or} \quad f_n \downarrow f \in L_\Phi \text{ implies } Tf_n \rightarrow Tf.$

We prove only the increasing case, the decreasing case runs analogously. Since $Tf_n \leq Tf_{n+1} \leq Tf$ by (iii), we obtain from (5) and the Theorem of Lebesgue (use (1)

and (2)) that

$$\int \Phi'(f - Tf) d\mu = 0 = \lim_{n \in \mathbb{N}} \int \Phi'(f_n - Tf_n) d\mu = \int \Phi'(f - \lim_{n \in \mathbb{N}} Tf_n) d\mu.$$

Since $\lim_{n \in \mathbb{N}} Tf_n \leq Tf$ and Φ' is strictly increasing, this implies $Tf = \lim_{n \in \mathbb{N}} Tf_n$, i.e. (6). Put $H = \{f \in L_\Phi : Tf = f\}$. From (i), (iv) and (6) we obtain

(7) $H + b \subset H$ for all $b \in \mathbb{R}$.

(8) $(\frac{3}{2})^n H \subset H, \quad n \in \mathbb{N}$.

(9) $f_n \in H, \quad f_n \uparrow f \in L_\Phi$ or $f_n \downarrow f \in L_\Phi \Rightarrow f \in H$.

We show that furthermore

(10) $f, g \in H \Rightarrow f \wedge g, \quad f \vee g \in H$.

We prove only $f \vee g \in H$, the case $f \wedge g \in H$ runs similarly. As $T(f \vee g) \geq Tf \vee Tg = f \vee g$ by (iii) and $\int \Phi'(f \vee g - T(f \vee g)) d\mu = 0$ by (5) we obtain, using $\Phi'(0) = 0$ and $\Phi'(x) < 0$ for $x < 0$, that $f \vee g = T(f \vee g)$, i.e. $f \vee g \in H$.

Let $\mathcal{L} = \{A \in \mathcal{Q} : 1_A \in H\}$. As $T0 = 0, T(1) = 1$ by (v) and (i), \mathcal{L} is a σ -lattice according to (9) and (10). If additionally $T(-Tf) = -Tf$, then \mathcal{L} is a σ -field, since $T1_A = 1_A$ implies by (i) that

$$T(1_{\bar{A}}) = T(1 - 1_A) = 1 + T(-1_A) = 1 - T1_A = 1_{\bar{A}}.$$

Now we show that

(11) $H = L_\Phi(\mathcal{L})$.

Since $((\frac{3}{2})^n(f - b) \vee 0) \wedge 1 \uparrow 1_{\{f > b\}}, f \in H$ implies by (7)–(10) that $1_{\{f > b\}} \in H$. Hence $\{f > b\} \in \mathcal{L}$, whence $f \in L_\Phi(\mathcal{L})$. It remains to prove $L_\Phi(\mathcal{L}) \subset H$. Let $f \in L_\Phi(\mathcal{L})$. According to (7) and (9) we may assume $f \geq 0$. Since

$$\sup \left\{ \frac{\nu}{2^n} 1_{\{f > \nu/2^n\}} : \nu = 0, \dots, n2^n \right\} \uparrow f$$

it suffices to prove according to (9) and (10) that $C \in \mathcal{L}$ implies $a1_C \in H$ for all $a \geq 0$. According to (8) we may assume $0 \leq a \leq 1$. Thus we have to prove $T1_C = 1_C$ implies $T(a1_C) = a1_C, 0 \leq a \leq 1$. As T is monotone by (iii) we have $T(a1_C) \leq T(a) \wedge T(1_C) = a1_C$. Hence (5) implies $T(a1_C) = a1_C$, i.e. $a1_C \in H$. Thus (11) is shown.

In the next step we prove that

(12) $\int \Phi'(f - Tf)g d\mu \leq 0$ if $f \in L_\Phi, g \in L_\Phi(\mathcal{L})$.

Using (i), (6) and (2) we may w.l.g. assume that $0 \leq f \leq b$. Using (5) and (2) we may w.l.g. assume that $0 \leq g \in L_\Phi(\mathcal{L})$. As each \mathcal{L} -measurable function $g \geq 0$ is increasing limit of \mathcal{L} -measurable simple functions (see e.g. [4, p. 313]), it suffices to prove (12) for $g = 1_C$ with $C \in \mathcal{L}$. Since $b1_C \in L_\Phi(\mathcal{L}) = H$ by (11) we have $T(b1_C) = b1_C$. As T is monotone and $0 \leq f \leq b$ we obtain $T(f1_C) \leq Tf$ and $T(f1_C) \leq T(b1_C) = b1_C$. Hence $T(f1_C) \leq (Tf)1_C$ and (1) and (5) imply that

$$\int \Phi'(f - Tf)1_C d\mu = \int \Phi'(f1_C - (Tf)1_C) d\mu \leq \int \Phi'(f1_C - T(f1_C)) d\mu = 0.$$

Thus (12) is shown.

Now we prove that

$$(13) \quad \int \Phi'(f - Tf)Tf \, d\mu \geq 0 \quad \text{if } f \in L_\Phi.$$

Using (1) and (2) we obtain from the Theorem of Lebesgue that the function $\varphi(t) := \int \Phi(f - tTf) \, d\mu$, $t \in \mathbf{R}$, is differentiable. Let $t_n = 1 - 1/2^n$; we obtain from (4), applied to $a = 0$ that

$$\varphi'(1) = \lim_{n \in \mathbf{N}} \frac{\varphi(1) - \varphi(t_n)}{1 - t_n} < 0.$$

As $\varphi'(1) = \int \Phi'(f - Tf)(-Tf) \, d\mu$ by (2) and the Theorem of Lebesgue, this implies (13).

By (12) and (13) we have

$$(14) \quad \int \Phi'(f - Tf)(g - Tf) \, d\mu \leq 0 \quad \text{if } f \in L_\Phi, g \in L_\Phi(\mathcal{L}).$$

Since $T(Tf) = Tf$ by (ii), we have $Tf \in H$. Hence $Tf \in L_\Phi(\mathcal{L})$ by (11). To prove $Tf = \mu_\Phi(f|\mathcal{L})$ we have therefore to show that

$$(15) \quad \int \Phi(f - Tf) \, d\mu \leq \int \Phi(f - g) \, d\mu \quad \text{for all } g \in L_\Phi(\mathcal{L}).$$

Let $f \in L_\Phi$, $g \in L_\Phi(\mathcal{L})$ be given and put

$$\Psi(\alpha) := \int \Phi(f - (\alpha Tf + (1 - \alpha)g)) \, d\mu, \quad \alpha \in \mathbf{R}.$$

Then ψ is a convex function as Φ is a convex function. By (2), the Theorem of Lebesgue and (14) we obtain that

$$(16) \quad \psi'(1) = \int \Phi'(f - Tf)(g - Tf) \, d\mu \leq 0.$$

As ψ is convex, $\psi'(\alpha)$, $\alpha \in \mathbf{R}$, is a nondecreasing function. Hence $\psi(0) \geq \psi(1)$ by (16), i.e. (15) holds. This finishes the proof.

The following lemma clarifies the connection of our integration condition (v) and other integration conditions.

LEMMA 2. Let $(\Omega, \mathcal{A}, \mu)$ and Φ be as in the theorem above. Let $T: L_\Phi \rightarrow L_\Phi$ be an operator with $T(f - \frac{1}{2}Tf) = \frac{1}{2}Tf$. Then the following three conditions are equivalent.

- (v) $\int \Phi(f - Tf) \, d\mu \leq \int \Phi(f) \, d\mu$ for $f \in L_\Phi$,
- (v*) $\int \Phi(f - Tf) \, d\mu \leq \int \Phi(f - aTf) \, d\mu$ for $f \in L_\Phi$ and some $a = a(f) \in [0, 1)$,
- (v**) $\int \Phi'(f)Tf \, d\mu \geq 0$ for $f \in L_\Phi$.

PROOF. (v) obviously implies (v*). Let $f \in L_\Phi$ be given and put $\varphi(a) = \int \Phi(f - aTf) \, d\mu$, $a \in \mathbf{R}$. Since Φ is a convex function, φ is a convex function, too.

Using (1) and (2) of the proof of the preceding theorem we obtain from the Theorem of Lebesgue that the function φ is differentiable with

$$(1) \quad \varphi'(a) = - \int \Phi'(f - aTf)Tf \, d\mu, \quad a \in \mathbf{R}.$$

If (v*) is fulfilled, then $\varphi(a) \geq \varphi(1)$ for some $0 \leq a < 1$. As φ is convex and differentiable, φ' is nondecreasing; hence $\varphi'(0) \leq 0$. Together with (1) this implies

(v**). Let (v**) be fulfilled. Since $T(f - \frac{1}{2}Tf) = \frac{1}{2}Tf$ we have $T(f - (1 - 1/2^n)Tf) = (1/2^n)Tf$ (see (3) in the proof of Theorem 1). Hence (1) and (v**) imply

$$\begin{aligned}\varphi'(1 - 1/2^n) &= -\int \Phi'(f - (1 - 1/2^n)Tf)Tf \, d\mu \\ &= -2^n \int \Phi'(f - (1 - 1/2^n)Tf)T(f - (1 - 1/2^n)Tf) \, d\mu \leq 0.\end{aligned}$$

Since φ is convex this directly implies $\varphi(0) \geq \varphi(1)$, i.e. (v).

Lemma 2 shows that the assertion of Theorem 1 holds also true if we replace condition (v) by

$$(v^{**}) \int \Phi'(f)Tf \, d\mu \geq 0 \text{ for } f \in L_\Phi.$$

If $\Phi(x) = x^2$ condition (v**) means $\int fTf \, d\mu \geq 0$ for $f \in L_2$. The stronger integration condition $\int fTf \, d\mu > 0$ for $f \in L_2$ with $Tf \neq 0$ was used in [4] and [12] together with other integration conditions to characterize conditional expectations, given a σ -lattice.

If $\Phi(x) = |x|^p$, $p > 1$, condition (v**) means $\int \operatorname{sgn} f |f|^{p-1}Tf \, d\mu \geq 0$ for $f \in L_p$. A weaker form of this integration condition was used together with p -expectation invariance (i.e. $\int |f - Tf|^{p-1} \operatorname{sgn}(f - Tf) \, d\mu = 0$) and five algebraic conditions to characterize conditional p -means, given a σ -lattice.

REFERENCES

1. T. Ando, *Contractive projections in L_p -spaces*, Pacific J. Math. **17** (1966), 391–405.
2. T. Ando and I. Amemiya, *Almost everywhere convergence of prediction sequence in L_p ($1 < p < \infty$)*, Z. Wahrsch. Verw. Gebiete **4** (1965), 113–120.
3. R. R. Bahadur, *Measurable subspaces and subalgebras*, Proc. Amer. Math. Soc. **6** (1955), 565–570.
4. R. E. Barlow, D. J. Bartholomew, J. M. Bremner and H. D. Brunk, *Statistical inference under order restrictions*, Wiley, New York, 1972.
5. H. D. Brunk, *On an extension of the concept conditional expectation*, Proc. Amer. Math. Soc. **14** (1963), 298–304.
6. ———, *Uniform inequalities for conditional p -means given σ -lattices*, Ann. Probab. **3** (1975), 1025–1030.
7. R. G. Douglas, *Contractive projections on an L_1 -space*, Pacific J. Math. **15** (1965), 443–462.
8. R. L. Dykstra, *A characterization of a conditional expectation with respect to a σ -lattice*, Ann. Math. Statist. **41** (1970), 698–701.
9. M. A. Krasnoselskii and Y. B. Rutickii, *Convex functions and Orlicz spaces*, Noordhoff, Groningen, 1961.
10. D. Landers and L. Rogge, *Characterization of p -predictors*, Proc. Amer. Math. Soc. **76** (1979), 307–309.
11. ———, *Best approximants in L_Φ -spaces*, Z. Wahrsch. Verw. Gebiete **51** (1980), 215–237.
12. ———, *Isotonic approximation in L_p* , J. Approx. Theory (to appear).
13. S. C. Moy, *Characterization of conditional expectation as transformation on function spaces*, Pacific J. Math. **4** (1954), 47–63.
14. J. Pfanzagl, *Characterization of conditional expectations*, Ann. Math. Statist. **38** (1967), 415–421.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT ZU KÖLN, WEYERTAL 86-90, D-5000 KÖLN 41, FEDERAL REPUBLIC OF GERMANY

FACHBEREICH WIRTSCHAFTSWISSENSCHAFTEN/STATISTIK, UNIVERSITÄT KONSTANZ, POSTFACH 7733, D-7750 KONSTANZ, FEDERAL REPUBLIC OF GERMANY