MEASURABLE SUBSPACES AND SUBALGEBRAS[1]

R. R. BAHADUR

1. Introduction. Let (X, S, μ) be a probability measure space. Here X is a set of points x, S is a σ-algebra of subsets of X, and μ is a σ-additive measure on S with $\mu(X)=1$. Let $V=\mathcal{L}_2(X, S, \mu)$ be the real Hilbert space of S-measurable functions $f(x)$ with $\int_X f^2 d\mu < \infty$.

For f and g in V, we write $(f, g)=\int_X f\cdot g d\mu$ where $(f\cdot g)(x)\equiv f(x)\cdot g(x)$, and $\|f\|=(f, f)^{1/2}$. Convergence in V is defined as usual in the norm topology, that is to say, $\lim_{n\to\infty} f_n = f$ means $\lim_{n\to\infty} \|f_n - f\|=0$. If f and g are functions in V such that $\|f-g\|=0$, they are regarded as identical and we write $f=g$. $f\geqq g$ means that $\mu\{x: f(x)<g(x)\}=0$. For any real α, the function which is equal to α for every x is also denoted by α. A function f in V is said to be bounded if there exist α and β such that $\alpha\leqq f\leqq\beta$.

A function on X which takes only the values 0 and 1 is called a characteristic function. For any set $A\subset X$, χ_A denotes the characteristic function which equals 1 on A and vanishes on $X-A$. If S_1 and S_2 are subclasses of S, we write $S_1\subset S_2$ if corresponding to each set $A\in S_1$ there exists a $B\in S_2$ such that $\chi_B=\chi_A$; we write $S_1=S_2$ if $S_1\subset S_2$ and $S_2\subset S_1$.

Let W be a subspace ($\equiv$closed linear manifold) in V. W is *algebraic* if it contains the function 1 (and therefore every constant function) and if, for any two bounded functions f and g in W, the function $f\cdot g$ is also in W. W is *bounded* (so to speak) if the set of bounded functions in W is dense in W. W is *measurable* if there exists a (necessarily unique) σ-subalgebra of S, S^* say, such that $W=\mathcal{L}_2(X, S^*, \mu)$, that is to say, W is the set of all S^*-measurable functions in V.

Let T be the (orthogonal) projection to the subspace W. T is *constant-preserving* if $T\alpha=\alpha$ for every constant function α. T is *positive* if $f\geqq 0$ implies $Tf\geqq 0$.

Let S^0 be the smallest σ-algebra of sets of X such that each f in W is an S^0-measurable function. Let S_0 be the class of all sets $A\subset X$ such that χ_A is in W. While S_0 is not necessarily a σ-algebra, it is a nonempty class of sets, with $S_0\subset S^0\subset S$.

The main result of this note can then be stated as follows.

THEOREM. *The following propositions are equivalent*:

Received by the editors November 18, 1954.

[1] This work was supported in part by the Office of Naval Research under contract N6onr-271, T.O. XI.

(i) *T is constant-preserving and positive.*
(ii) *W is algebraic and bounded.*
(iii) $S_0 = S^0$.
(iv) $W = \mathcal{L}_2(X, S^0, \mu)$.

The proof of the theorem is given in the next section.
By the equivalence of (ii) and (iv) we have

COROLLARY 1. *A subspace is algebraic and bounded if and only if it is measurable.*

This result is essentially about the set $\mathcal{L}_\infty$ of bounded functions in V, and it may be worthwhile to state it entirely in terms of $\mathcal{L}_\infty$ as follows. Regard $\mathcal{L}_\infty$ as a linear algebra under pointwise multiplication of functions, and for the moment suppose closure in $\mathcal{L}_\infty$ to mean closure in the usual $\mathcal{L}_\infty$ topology. Then Corollary 1 can be shown to be equivalent to the proposition that a closed linear subalgebra of $\mathcal{L}_\infty$ contains 1 if and only if it is a measurable algebra, that is to say, it is the closure of the linear algebra generated by a set of characteristic functions including 1.

It is easy to show (cf. the last paragraph of the following section) that if S^* is a σ-subalgebra of S then the projection to $\mathcal{L}_2(X, S^*, \mu)$ is exactly the conditional expectation operator relative to S^*. In other words, a transformation T is a conditional expectation if and only if T is the projection to a measurable subspace.[2] In consequence, the equivalence of (i) and (iv) yields the following characterization of conditional expectation.

COROLLARY 2. *A transformation T on V into itself is a conditional expectation if and only if T is linear, idempotent, self-adjoint, constant-preserving, and positive.*

This result is closely related to a characterization of conditional expectation operators on $\mathcal{L}_1$ which was obtained by Moy [1, Theorem 2.2]. An essential part of the proof which follows is based on an argument of Moy.

2. **Proof of the theorem.** Since T is a projection, we have

$$T(\alpha f + \beta g) = \alpha T(f) + \beta T(g), \tag{1}$$

$$T^2 f = Tf, \tag{2}$$

and

$$(Tf, g) = (f, Tg) \tag{3}$$

[2] This fact, which motivated the work presented here, was pointed out to the writer by L. J. Savage.

for all α, β, f and g; moreover, since T is the projection to W,

$$W = \{Tf : f \in V\}, \tag{4}$$

and also

$$W = \{f : Tf = f, f \in V\}. \tag{5}$$

(Cf., e.g., [2].)

We shall prove the theorem by showing that (i)→(ii)→(iii)→(iv)→(i).

Suppose then that (i) holds, that is,

$$T\alpha = \alpha \qquad \text{for every } \alpha \tag{6}$$

and

$$Tf \geqq 0 \qquad \text{whenever } f \geqq 0. \tag{7}$$

It follows from (1) that (6) and (7) are equivalent to

$$\alpha \leqq Tf \leqq \beta \qquad \text{whenever } \alpha \leqq f \leqq \beta \tag{8}$$

and

$$Tf \geqq Tg \qquad \text{whenever } f \geqq g. \tag{9}$$

Consider a fixed g in W. Define $f_n(x) = g(x)$ if $|g(x)| \leqq n$ and $=0$ otherwise, for $n = 1, 2, \cdots$. Clearly, $\{f_n\}$ is a sequence in V such that $\lim_{n\to\infty} f_n = g$. Set $g_n = Tf_n$. Then, for each n, $g_n \in W$ by (4), and $-n \leqq g_n \leqq n$ by (8). Since $\|Tf\| \leqq \|f\|$ for all f (by (3)), and since $Tg = g$ by (5), we have $\|g_n - g\| = \|Tf_n - Tg\| = \|T(f_n - g)\| \leqq \|f_n - g\|$, so that $\lim_{n\to\infty} g_n = g$. Since g is arbitrary, W is thus shown to be bounded.

According to (5) and (6), W does contain every constant function. It remains therefore to show that $f \in W$, $g \in W$ implies $f \cdot g \in W$, provided that f and g are bounded. It will suffice to show that

(10) $\qquad f \in W$ implies $f^2 \in W$ provided that f is bounded.

For, if f and g are bounded functions in W then $f+g$ is such a function also, and it will follow from (10) that f^2, g^2, and $(f+g)^2$ are in W; consequently $f \cdot g = (f+g)^2/2 - f^2/2 - g^2/2$ is in W.

We proceed to establish (10). Choose and fix a bounded $f \in W$, and define $g = Tf^2 - f^2$. In view of (5), we have to show that $g = 0$.

Consider the parabola $v = u^2$ in the uv-plane, and for any real r let $v = a_r u + b_r$ be the tangent to the parabola at the point $u = r$, $v = r^2$ ($a_r = 2r$, $b_r = -r^2$). Let R be a countable everywhere dense set (e.g. the rational points) of the real line. Then, for each u, $-\infty < u < \infty$, $u^2 \geqq a_r u + b_r$ for every r, and $u^2 = \sup_{r \in R}\{a_r u + b_r\}$.

We have $f^2 \geqq a_r f + b_r$ for each fixed r. Hence, for each r, $Tf^2 \geqq T(a_r f + b_r) = a_r f + b_r$, by (9), (1), (5), and (6). It follows that

$$Tf^2 \geqq \sup_{r \in R} \{a_r f + b_r\} = f^2.$$

Thus $g \geqq 0$. We observe next that, for any $h \in V$,

$$\begin{aligned} \int_X h d\mu &= (h, 1) \\ &= (h, T1) \text{ by (6)} \\ &= (Th, 1) \text{ by (3)} \\ &= \int_X Th d\mu. \end{aligned} \tag{11}$$

Since $Tg = T^2 f^2 - Tf^2 = Tf^2 - Tf^2 = 0$ by (1) and (2), it follows from (11) that $\int_X g\, d\mu = 0$, and hence $g = 0$. This completes the proof that (i)→(ii).

Suppose now that (ii) holds. We shall show first that S_0 is a σ-algebra. Since W contains the function 1, $\chi_A \in W$ implies $\chi_{X-A} = 1 - \chi_A \in W$; since W is closed under pointwise multiplication of functions, $\chi_A \in W$, $\chi_B \in W$ implies $\chi_{A \cap B} = \chi_A \cdot \chi_B \in W$; and since W is a closed subset of V, if $f_1, f_2, \cdots$ is a sequence of characteristic functions in W such that $f_r \cdot f_s = 0$ for $r \neq s$, $g = \sum_r f_r$ is a characteristic function in W. By referring to the definition of S_0, we see that the class S_0 is closed under complementation, under intersection, and under countable union of disjoint sets; S_0 is therefore a σ-algebra. We proceed to show that $S^0 \subset S_0$; this will establish (iii), since $S_0 \subset S^0$ in any case.

Let f be a bounded function in W, and suppose that $\alpha \leqq f \leqq \beta$. Let I denote the interval $[\alpha, \beta]$, and let M be the class of Borel measurable sets of I. Define $\nu(C) = \mu(f^{-1}(C))$ for $C \in M$. Then ν is a probability measure on M. It is well known (cf., e.g., [**3**])that the set of polynomial functions on I is dense in $\mathcal{L}_2(I, M, \nu)$. Let us consider a fixed $C \in M$. There exists a sequence $\{p_n\}$ of polynomials on I such that $\lim_{n\to\infty} p_n = \chi_C$. Writing $A = f^{-1}(C)$, we have $\|p_n - \chi_C\| = \|p_n(f) - \chi_A\|$ for every n, so that $\lim_{n\to\infty} p_n(f) = \chi_A$. Since W is algebraic, and f is a bounded function in W, we have $p_n(f) \in W$ for each n; hence, since W is closed, $\chi_A \in W$, that is to say, $A = f^{-1}(C) \in S_0$. Since C is arbitrary, f is S_0-measurable.

We have therefore shown that every bounded $f \in W$ is an S_0-measurable function. Since W is bounded, it follows that every $f \in W$ is S_0-measurable. Consequently $S^0 \subset S_0$, by the definition of S^0. This completes the proof that (ii)→(iii).

Suppose next that S_0 is a σ-algebra, which is less than supposing that (iii) holds. Consider an $f\in\mathcal{L}_2(X, S_0, \mu)$. There exists a sequence $\{f_n\}$ such that each f_n is of the form $\sum_{i=1}^k \alpha_i g_i$ where $g_1, g_2, \cdots, g_k$ are S_0-measurable characteristic functions, and also such that $\lim_{n\to\infty} f_n = f$. It follows from the definition of S_0 that $\{f_n\}$ is a sequence in W; hence $f\in W$, since W is closed. Since f is arbitrary, we have $\mathcal{L}_2(X, S_0, \mu)\subset W$. On the other hand, $f\in W$ implies (by the definition of S^0) that f is S^0-measurable and therefore in $\mathcal{L}_2(X, S^0, \mu)$, so that $W\subset\mathcal{L}_2(X, S^0, \mu)$. Thus $\mathcal{L}_2(X, S_0, \mu)\subset W\subset\mathcal{L}_2(X, S^0, \mu)$, provided only that S_0 is a σ-algebra; if in fact $S_0 = S^0$ the three subspaces must be identical. This proves, in particular, that (iii)→(iv).

Suppose, finally, that (iv) holds. Let U be the conditional expectation operator relative to S^0, that is to say, for each f in $\mathcal{L}_1(X, S, \mu)$ let Uf be the unique function in $\mathcal{L}_1(X, S^0, \mu)$ such that $\int_A f d\mu = \int_A Uf d\mu$ for all $A\in S^0$. The existence of such a function follows easily from the Radon-Nikodym theorem (cf., e.g., [**4**]). Consider an $f\in V$. Since $Tf\in W$ by (4), the hypothesis (iv) implies that

$$Tf\in\mathcal{L}_1(X, S^0, \mu).$$

Also, for any $A\in S^0$ we have

$$\begin{aligned}\int_A f d\mu &= (f, \chi_A)\\ &= (f, T\chi_A) \text{ by (iv) and (5)}\\ &= (Tf, \chi_A) \text{ by (3)}\\ &= \int_A Tf d\mu.\end{aligned}$$

It follows hence by the uniqueness of conditional expectation that Tf differs from Uf on a set of μ-measure zero. Since f is arbitrary, we conclude that $T = U$ on V. (It follows incidentally that $f\in V$ implies $Uf\in V$, $\|Uf\|\leqq\|f\|$). As is well known, U is constant-preserving and positive, so that (i) holds. This completes the proof of the theorem.

3. **Applications to a finite-dimensional function space.** By way of an example, let $X = \{1, 2, \cdots, n\}$, S = the class of all sets of X, and μ = the uniform distribution on X, that is to say, $\mu(\{x\}) = 1/n$ for $x = 1, 2, \cdots, n$. In this case V is the real vector space of all real-valued functions on X, every $f\in V$ is bounded, and every linear manifold in V is closed.

Let W be a linear manifold in V. By definition, W is algebraic if it contains every constant function, and if W is closed under pointwise multiplication of functions. It follows from Corollary 1 that a k-dimensional linear manifold W is algebraic if and only if there exists a partition of X into k mutually exclusive nonempty sets, $X = \bigcup_{r=1}^{k} X_r$ say, such that W is the set of *all* functions of the form $f(x) = \alpha_r$ for $x \in X_r$ $(r = 1, 2, \cdots, k)$. The verification is omitted.

For each $i = 1, 2, \cdots, n$ let $f_i(x) = 1$ for $x = i$ and $f_i(x) = 0$ for $x \neq i$. Then $\{f_1, f_2, \cdots, f_n\}$ is a basis in V. Let $\{t_{ij}\}$ be an $n \times n$ symmetric probability matrix (i.e. $t_{ij} = t_{ji}$, $t_{ij} \geqq 0$, $\sum_{j=1}^{n} t_{ij} = 1$), and for $f = \sum_{i=1}^{n} \alpha_i f_i$ let $Tf = \sum_{i=1}^{n} \alpha_i Tf_i$, where $Tf_i = \sum_{j=1}^{n} t_{ij} f_j$. Then the transformation T is linear, self-adjoint, constant-preserving, and positive.

Clearly, T is idempotent if and only if the matrix $\{t_{ij}\}$ is idempotent (i.e. $\sum_{r=1}^{n} t_{ir} t_{rj} = t_{ij}$). Moreover, it follows easily from the definition of a conditional probability that T is a conditional expectation, with a k-dimensional subspace as its range, if and only if there exists a partition of n into k positive integers, $n = n_1 + n_2 + \cdots + n_k$ say, such that (except for a permutation of rows and columns) $\{t_{ij}\}$ is the matrix

$$M_k(n_1, n_2, \cdots, n_k) = \begin{pmatrix} N_1 & & & 0 \\ & N_2 & & \\ & & \ddots & \\ 0 & & & N_k \end{pmatrix}.$$

where N_r is the $n_r \times n_r$ matrix with each element equal to $1/n_r$.

It now follows from Corollary 2 that *a symmetric $n \times n$ probability matrix of rank k is idempotent if and only if there exists a partition of n into k positive integers $n_1, n_2, \cdots, n_k$ such that the matrix is a permutation of the rows and columns of $M_k(n_1, n_2, \cdots, n_k)$.*

References

1. Shu-Teh Chen Moy, *Characterizations of conditional expectation as a transformation on function spaces*, Pacific Journal of Mathematics vol. 4 (1954) p. 47.

2. P. R. Halmos, *Introduction to Hilbert space and the theory of spectral multiplicity*, New York, Chelsea, 1951, pp. 1–48.

3. G. Szegö, *Orthogonal polynomials*, Amer. Math. Soc. Colloquium Publications, vol. 23, 1939, pp. 10–11.

4. J. L. Doob, *Stochastic processes*, New York, Wiley, 1953, pp. 15–18.

5. P. R. Halmos, *Measure theory*, New York, Van Nostrand, 1951.

6. G. Birkhoff, *Moyennes des fonctions bornées*, Congrès d'Algèbre et Théorie des Nombres, Paris, 1949, pp. 143–156.

The University of Chicago