JOURNAL OF APPROXIMATION THEORY 6, 391–403 (1972)

Nonlinear Prediction of Generalized Random Processes

G. Y. H. CHI

Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

Communicated by Oved Shisha

Received August 27, 1970

DEDICATED TO PROFESSOR J. L. WALSH
ON THE OCCASION OF HIS 75TH BIRTHDAY

Let (D) be the Schwartz space of infinitely differentiable scalar functions, φ, on the real line with compact supports, $\sigma(\varphi)$, and (Ω, Σ, P) be a fixed complete probability space. Let $X : (D) \to L^p(\Omega, \Sigma, P)$, with $1 < p < \infty$, be a generalized random process, i.e., a continuous linear transformation, $T = (-\infty, t)$, and $\beta_T = \sigma(X(\varphi) : \sigma(\varphi) \subset T)$, the σ-field generated by the random variables. Let $\{\tilde{\varphi}_i\}_{i=1}^{\infty} \subset (D)$ be such that $\sigma(\tilde{\varphi}_i) \not\subset T, i \geqslant 1$. Then relative to the p-th moment of the error as the criterion of optimality, there exists a weak generalized random process (i.e., a closed linear transformation) $Y : \mathscr{D}(Y) \to L^p(\Omega, \beta_T, P)$, where $\mathscr{D}(Y)$ is dense in (D), such that $\| Y(\tilde{\varphi}_i) - X(\tilde{\varphi}_i)\|_p$ is a minimum for each $i \geqslant 1$.

INTRODUCTION

Let (D) be the Schwartz space of infinitely differentiable scalar functions φ, on the real line with compact supports $\sigma(\varphi)$, and (Ω, Σ, P) be a fixed complete probability space. Let $X : (D) \to L^p(\Omega, \Sigma, P)$, with $1 < p < \infty$, be a generalized random process (g.r.p.), i.e., a continuous linear transformation, $T = (-\infty, t)$, and $\beta_T = \sigma(X(\varphi) : \sigma(\varphi) \subset T)$, the σ-field generated by the random variables. Let $\{\tilde{\varphi}_i\}_{i=1}^{\infty} \subset (D)$ be such that $\sigma(\tilde{\varphi}_i) \not\subset T, i \geqslant 1$. Then the nonlinear prediction problem is: relative to the p-th moment of the error as the criterion of optimality, it is desired to find a weak generalized random process (i.e., a closed linear transformation) $Y : \mathscr{D}(Y) \to L^p(\Omega, \beta_T, P)$, where $\mathscr{D}(Y)$ is dense in (D), such that $\| Y(\tilde{\varphi}_i) - X(\tilde{\varphi}_i)\|_p$ is a minimum for each $i \geqslant 1$.

The solution to the above problem is given in this paper. An approximation theorem is proved. These results generalize some of the results of Rao [8].

1. Preliminaries

In this section, a reformulation and generalization of some known results of Šidak [11] will be stated without proof.

Let (D), (Ω, Σ, P), X, T and β_T be as above, and let

$$S_T = \{\varphi : \varphi \in (D), \sigma(\varphi) \subset T\}.$$

Since (D) is not metrizable, even though it is separable, it is not obvious that S_T is also separable, because separability is not hereditary in arbitrary topological spaces. For a counterexample, see [7, p. 84]. However, in the present case, one has the following

Lemma 1.1. *There exists a countable dense subset, $\mathscr{F}$, of (D) such that* $\overline{\mathscr{F} \cap S_T} = S_T$.

Let $\mathscr{F} = \{\varphi_i\}_{i=1}^{\infty}$ denote the countable dense subset of (D) given in Lemma 1.1, and let $\beta_{\mathscr{F}}$ denote the σ-field generated by $\{X(\varphi_i) : \varphi_i \in \mathscr{F}\}$, then one can consider X as a map from (D) into $L^p(\Omega, \beta_{\mathscr{F}}, P)$. Henceforth, $L^p(\Omega, \beta_{\mathscr{F}}, P)$ will be denoted by $L^p(\Omega, \Sigma, P)$ if no confusion arises. S_T is closed and separable by Lemma 1.1; in fact

$$\mathscr{F}_T = \mathscr{F} \cap S_T = \{\varphi_1{}^T, \varphi_2{}^T, \dots, \varphi_n{}^T, \dots\}$$

is dense in S_T, and $\beta_T = \sigma(X(\varphi_i{}^T) : i = 1, 2, \dots, n, \dots)$. Let

$$\beta_n = \sigma(X(\varphi_1{}^T), X(\varphi_2{}^T), \dots, X(\varphi_n{}^T)).$$

All σ-fields below are completed relative to P and ' $=$ ' and ' $\subset$ ' etc. hold a.e.

Lemma 1.2. $\beta_T = \sigma(\bigcup_{n=1}^{\infty} \beta_n)$.

Definition 1.1 [8]. Let $\mathscr{M} \subset L^p(\Omega, \Sigma, P)$ be a closed linear submanifold. $\mathscr{M}$ is said to be a *measurable* subspace, if there exists a unique (necessarily so, since (Ω, Σ, P) is complete) σ-field $\Sigma_1 \subset \Sigma$ such that $\mathscr{M} = L^p(\Omega, \Sigma_1, P)$, i.e., $\mathscr{M}$ is the set of all Σ_1-measurable functions in $L^p(\Omega, \Sigma, P)$, or equal a.e. to one which is.

The following two results are trivial extensions of Šidak's [11].

Lemma 1.3. *The correspondence between sub σ-fields $G \subset \Sigma$ and measurable subspaces $L^p(\Omega, G, P)$ is $1-1$ $(1 \leqslant p < \infty)$.*

Lemma 1.4. *Let $\{G_i, i \in I\}$ be sub σ-fields of Σ.*

(a) *If $G_i \subset G_j$, then $L^p(\Omega, G_i, P) \subset L^p(\Omega, G_j, P)$.*

(b) *To every* $\bigcap_{i\in I} G_i$, *there corresponds a* $\bigcap_{i\in I} L^p(\Omega, G_i, P)$.

(c) *To every* $\sigma(\bigcup_{i\in I} G_i)$, *there corresponds a* $\sigma(\bigcup_{i\in I} L^p(\Omega, G_i, P))$, *the smallest measurable subspace containing* $\bigcup_{i\in I} L^p(\Omega, G_i, P)$.

2. Complements to a Theorem of Murray and Mackey

The aim of this section is to prove Theorem 2.1 below, which improves upon a theorem of [6], and furthermore it is used crucially in Sections 3 and 4 in showing the existence of a best predictor as a weak generalized random process.

Definition 2.1. Let $L^p(\Omega, \Sigma, P)$ be the Lebesgue space, $\mathscr{M} = L^p(\Omega, \beta, P)$ a measurable subspace, and $X_0 \in L^p(\Omega, \Sigma, P)$. Then a projection (linear and idempotent) operator, $P_\beta : L^p(\Omega, \Sigma, P) \to \mathscr{M}$ is termed a *prediction operator* of X_0 relative to β provided the following condition (C) holds:

(C) There exists a $Y_0 \in \mathscr{M}$ such that $\| X_0 - Y_0 \|_p \leqslant \| X_0 - Y \|_p$ for all $Y \in \mathscr{M}$ implies X_0 is in $\mathscr{D}(P_\beta)$, the domain of P_β, and $P_\beta(X_0) = Y_0$. (P_β was called a 'closed conditional expectation' in [9]).

Theorem 2.1. *Let* $L^p(\Omega, \Sigma, P)$ $1 < p < \infty$ *be a separable space and* $\mathscr{M} = L^p(\Omega, \beta_T, P)$. Let $\tilde{\mathscr{E}}$ and $\mathscr{G}'$ *be two fixed but arbitrary countable dense subsets of* $L^p(\Omega, \Sigma, P)$ *and* $(L^p(\Omega, \Sigma, P))'$, *respectively. If* $\{\tilde{\varphi}_i\}_{i=1}^\infty \subset (D)$ *are such that* $\sigma(\tilde{\varphi}_i) \not\subset T$, *for every* $i \geqslant 1$, *then there exists a countable dense subset* $\mathscr{E}$ *of* $L^p(\Omega, \Sigma, P)$ *depending on* $\{\tilde{\varphi}_i\}_{i=1}^\infty$, *and a prediction operator* P_{β_T} *with domain* Δ *dense in* $L^p(\Omega, \Sigma, P)$ *such that*

(a) $\Delta = \mathscr{M} \oplus \mathscr{N}$, *where* $(\mathscr{M}, \mathscr{N})$ *is a pair of quasi-complements (i.e.,* $\mathscr{M} \cap \mathscr{N} = \{0\}$ *and the closed subspaces* $\mathscr{M}$ *and* $\mathscr{N}$ *are such that* $\mathscr{M} \oplus \mathscr{N}$ *is dense in* $L^p(\Omega, \Sigma, P)$).

(b) $\Delta \supset \mathscr{E}$.

(c) $P_{\beta_T} : \Delta \to \mathscr{M}$

(d) $P_{\beta_T} X(\tilde{\varphi}_i)$ *is the best nonlinear predictor for* $X(\tilde{\varphi}_i)$ *in* $\mathscr{M}$ *simultaneously for* $i = 1, 2, \ldots, n, \ldots$.

Furthermore, there exists a closed linear operator, P'_{β_T}, *with domain* Δ' *dense in* $(L^p(\Omega, \Sigma, P))'$, *the topological dual, such that*

(a*) $\Delta' = \mathscr{M}' \oplus \mathscr{N}'$, *where* $(\mathscr{M}', \mathscr{N}')$ *is a pair of quasi-complements.*

(b*) $\Delta' \supset \mathscr{G}'$, *a countable dense subset of* $(L^p(\Omega, \Sigma, P))'$.

(c*) $P'_{\beta_T} : \Delta' \to \mathscr{N}'$.

(d*) *For every* $f_i' \in \mathscr{G}'$, $P'_{\beta_T}(f_i')$ *is the element in* $\mathscr{N}'$ closest to f_i', i.e., P'_{β_T} *is also a prediction operator. Moreover,* P'_{β_T} *is the proper adjoint of* P_{β_T}.

In order to prove this theorem, it is necessary now to recall some definitions and facts from the theory of locally convex spaces [4, 5, 10]. In the following discussion, if A, B are subspaces of $L^p(\Omega, \Sigma, P)$, then $A \oplus B$ denotes the smallest subspace of $L^p(\Omega, \Sigma, P)$ containing both A and B with $A \cap B = \{0\}$.

Remark 2.1. If the linearity, of P_{β_T} and P'_{β_T}, is not demanded, then the result, in the present case, follows very simply even with $\Delta = L^p(\Omega, \Sigma, P)$ and $\Delta' = (L^p(\Omega, \Sigma, P))'$. But the linearity will be needed for the main problem of this paper.

DEFINITION 2.2. A *pairing* is an ordered pair $\langle Z, L\rangle$ of linear spaces Z and L over the same scalar field together with a fixed bilinear functional on their product $Z \times L$.

The usual inner product notation $\langle \cdot, \cdot \rangle$ will be used to denote the bilinear functional.

DEFINITION 2.3. A pairing is called a *duality* if the bilinear functional satisfies the separation axioms:

(a) $\langle x_0, y\rangle = 0$ for all $y \in L$ implies $x_0 = 0$.

(b) $\langle x, y_0\rangle = 0$ for all $x \in Z$ implies $y_0 = 0$.

If Z is a locally convex space, and Z^* its algebraic dual, then $\langle Z, Z^*\rangle$ is a duality where the bilinear functional is given by $\langle x, x^*\rangle \to x^*(x)$. If Z' denote the topological dual, then $\langle Z, Z^*\rangle$ induces a duality $\langle Z, Z'\rangle$ on the subspace $Z \times Z'$ of $Z \times Z^*$.

If $\langle Z, L\rangle$ is a duality, then let $\sigma(Z, L)$ denote the coarest locally convex (l.c.) topology on Z for which the linear functionals $x \to \langle x, y\rangle$, $y \in L$ are continuous. $\sigma(Z, L)$ is called the *weak topology* on Z relative to the duality $\langle Z, L\rangle$. Clearly the topological dual of $(Z, \sigma(Z, L))$ is L. In like manner, one can consider $\sigma(L, Z)$ on L. Now let $\mathscr{S}$ be the family of all $\sigma(L, Z)$-bounded subsets of L. Then the corresponding $\mathscr{S}$-topology (i.e., the topology of uniform convergence on the sets in $\mathscr{S}$ [10, p. 79]) on Z, denoted by $\beta(Z, L)$, is called the *strong topology* on Z relative to the duality $\langle Z, L\rangle$.

A l.c. topology t on Z is said to be *consistent* with the duality $\langle Z, L\rangle$, if $(Z_t)' = L$. Thus $\sigma(Z, L)$ is the weakest l.c. topology consistent with the duality; but in general $\beta(Z, L)$ is not consistent with the duality.

In the sequel, if Z is a l.c. space, then Z_σ will denote the linear space Z equipped with the topology $\sigma(Z, Z')$, and Z_β the linear space Z equipped with the strong topology $\beta(Z, Z')$.

If Z_1, Z_2 are two l.c. spaces, then a linear transformation $l: Z_1 \to Z_2$ is said to be *weakly continuous* if it is $(\sigma(Z_1, Z_1'), \sigma(Z_2, Z_2'))$-continuous, and *strongly continuous* if it is $(\beta(Z_1, Z_1'), \beta(Z_2, Z_2'))$-continuous.

A l.c. space Z is said to be *reflexive* if $(Z_\beta')_\beta' = Z$. For instance, (D) is reflexive [10, Example 4, p. 147], and $L^p(\Omega, \Sigma, P)$, $1 < p < \infty$ are reflexive.

Let $\langle Z, L\rangle$ be a duality. For any subset $M \subset Z$, the subset

$$M^0 = \{ y \in L : |\langle x, y\rangle| \leqslant 1, x \in M\}$$

of L is called the *polar* of M. If $M \subseteq Z$ is a linear submanifold, then $|\langle x, y\rangle| \leqslant 1$ for all $x \in M$ implies $\langle x, y\rangle = 0$ for all $x \in M$; hence M^0 consists of those elements of L that vanish on M and so is the linear submanifold of L orthogonal to M. Analogously, $(M^0)^0 = M^{00}$ is the polar of M^0, called the *bipolar* of M.

The following lemma holds true even if M_1, M_2 are just subsets of Z. However, since only the linear submanifolds will be of interest here, it will be stated and proved in the present form. For the more general case. see [10, pp. 125, 126].

LEMMA 2.1. *If $M_1 \subseteq M_2$, where M_1, M_2 are linear submanifolds of Z, then*

(a) $M_2^0 \subseteq M_1^0$

(b) $M_1 \subseteq M_1^{00}$

(c) $M_1 = M_1^{00}$ *if and only if* $M_1 = K^0$ *for some linear submanifold* $K \subseteq L$.

Proof. (a) If $l \in M_2^0$, then $l(y) = 0$ for all $y \in M_2$. This implies that $l(y) = 0$ for all $y \in M_1$, and hence $l \in M_1^0$.

(b) If $x \in M_1$, then $l(x) = 0$ for every $l \in M_1^0$. Consequently, $x \in M_1^{00}$.

(c) Since $M_1^{000} = (M_1^{00})^0 \subseteq M_1^0$, and $M_1^{000} = (M_1^0)^{00} \supseteq M_1^0$ follow from (b), hence $M_1^0 = M_1^{000}$. Thus if $M_1 = K^0$ for some $K \subseteq L$, then $K^0 = K_1^{000}$ implies that $M_1 = M_1^{00}$. Conversely, if $M_1 = M_1^{00}$, then simply take $K = M_1^0 \subseteq L$.

LEMMA 2.2. *Let $\langle Z, L\rangle$ be a duality, and $M \subseteq Z$ a linear submanifold. Then $M = M^{00}$ if and only if M is $\sigma(Z, L)$-closed.*

Proof. If M is $\sigma(Z, L)$-closed, then by the Bipolar theorem [10, p. 126] one has $M = M^{00}$. Conversely, if $M = M^{00}$, then by the same theorem just quoted, M is $\sigma(Z, L)$-closed. Q.E.D.

Remark 2.2. The operation '0' of taking polar sets up a $1 - 1$ inclusion inverting correspondence between the $\sigma(Z, L)$-closed linear submanifolds and the $\sigma(L, Z)$-closed linear submanifolds.

LEMMA 2.3. *Let $\langle Z, L\rangle$ be a duality such that Z and L are $\aleph_0$-dimensional, i.e., have $\aleph_0$ independent generators. Then if M is any $\sigma(Z, L)$-closed linear submanifold, there exists a second $\sigma(Z, L)$-closed linear submanifold N such that $M \oplus N = Z$ and $M^0 \oplus N^0 = L$.*

Proof. For proof of this lemma, see [5, Lemma 1, pp. 322–323]. Consider now the proof of Theorem 2.1.

Proof of Theorem 2.1. Let $L^p(\Omega, \Sigma, P)$ be the given separable space with $1 < p < \infty$, and $\mathscr{M} = L^p(\Omega, \beta_T, P)$. Since $(L^p(\Omega, \Sigma, P))'$ is total for $L^p(\Omega, \Sigma, P)$ and if $\mathscr{G}'$ is a countable dense subset of $(L^p(\Omega, \Sigma, P))'$, then $\mathscr{G}'$ is also total for $L^p(\Omega, \Sigma, P)$.

Since $\mathscr{M}$ is closed, and $L^p(\Omega, \Sigma, P)$ is separable, then $L^p(\Omega, \Sigma, P)/\mathscr{M}$ is also separable. Hence there exists a countable total set $\mathscr{H}'$ in $(L^p(\Omega, \Sigma, P)/\mathscr{M})'$. Now every member F' in $(L^p(\Omega, \Sigma, P)/\mathscr{M})'$ has associated with it in the following manner a member f' in $(L^p(\Omega, \Sigma, P))'$ which vanishes throughout $\mathscr{M}$.

$$f'(x) = F'(x + \mathscr{M}), \text{ for every } x \in L^p(\Omega, \Sigma, P).$$

Therefore, corresponding to $\mathscr{H}'$, there is a countable but not necessarily total subset $\mathscr{G}_1'$ of $(L^p(\Omega, \Sigma, P))'$. Note that each $f' \in \mathscr{G}_1'$ vanishes on $\mathscr{M}$ only. For, if $x \in \mathscr{M}$ such that $f'(x) = 0$, then the corresponding F' in $(L^p(\Omega, \Sigma, P)/\mathscr{M})'$ would be such that $F'(x + \mathscr{M}) = f'(x) = 0$, i.e., this implies that $x + \mathscr{M} = \mathscr{M}$, i.e., $x \in \mathscr{M}$, which is a contradiction. Hence the intersection of the null spaces of f' in $\mathscr{G}_1'$ is $\mathscr{M}$.

Let $d_i = \inf_{Y \in \mathscr{M}}\{\| X(\tilde{\varphi}_i) - Y \|_p\}$, $i \geq 1$. Then $d_i \geq 0$. Without loss of generality, one can assume that for some i, $d_i > 0$, otherwise the problem is trivial. Since $1 < p < \infty$, $L^p(\Omega, \Sigma, P)$ is uniformly convex, and $\mathscr{M}$ is a closed linear submanifold, therefore there is a unique element $Y_i \in \mathscr{M}$ such that $d_i = \| X(\tilde{\varphi}_i) - Y_i \|_p$, for every $i \geq 1$ [2, *II*. 4.29].

Now consider the subset $\mathscr{V}$ of $L^p(\Omega, \Sigma, P)$ defined by

$$\mathscr{V} = \{h : \| h \|_p \leq \| g + h \|_p \,,\, g \in \mathscr{M}\},$$

i.e., $\mathscr{V}$ contains elements h of $L^p(\Omega, \Sigma, P)$ that are minimal relative to $\mathscr{M}$. Note that $\mathscr{V}$ is not necessarily linear but is closed [6]. Now by [6, p. 80], every element of $L^p(\Omega, \Sigma, P)$ can be written as $f = g + h$, where $g \in \mathscr{M}$, $h \in \mathscr{V}$. Since $\| h \|_p = \| f - g \|_p$, g is nearest to f in $\mathscr{M}$. Thus let $X(\tilde{\varphi}_i) = Y_i + Z_i$, where $Y_i \in \mathscr{M}$, $Z_i \in \mathscr{V}$, $i \geq 1$ be such decompositions.

Since $L^p(\Omega, \Sigma, P)$ is separable, there exists a countable dense subset $\tilde{\mathscr{E}}$.

$\tilde{\mathscr{E}} \cap \mathscr{M}$ is necessarily dense in $\mathscr{M}$. Let $\mathfrak{A} = \{X(\tilde{\varphi}_i), Z_i, i \geq 1\}$, $\mathscr{C} = \mathfrak{A} - \tilde{\mathscr{E}}$, and $\mathscr{E} = \mathscr{C} \cup \tilde{\mathscr{E}}$. Then $\mathscr{E}$ is still countable, dense in $L^p(\Omega, \Sigma, P)$, and in fact is total for $(L^p(\Omega, \Sigma, P))'$.

Let $Z = \mathscr{L}(\mathscr{E})$ denote the linear span of $\mathscr{E}$, $M = \mathscr{M} \cap Z$, and $L = \mathscr{L}(\mathscr{G}' \cup \mathscr{G}_1')$, then it is claimed that M, Z, and L satisfy the hypothesis of Lemma 2.3. For,

(a) (Z, L) is clearly a duality. For it is a pairing with the induced bilinear functional from $(L^p(\Omega, \Sigma, P), (L^p(\Omega, \Sigma, P))')$. The induced bilinear functional satisfies the separation axioms, since both Z and L are total in $(L^p(\Omega, \Sigma, P))$ and $L^p(\Omega, \Sigma, P)'$ respectively.

(b) It is clear that both Z and L are $\aleph_0$-dimensional.

(c) M is $\sigma(Z, L)$-closed, for let

$$K = \{F' : F' \in L, F'(x) = 0, \text{ for every } x \in \mathscr{M}\}. \tag{1}$$

Then K is a proper linear submanifold of L. This is seen as follows: Clearly $K \neq \varnothing$, since $K \supseteq \mathscr{G}_1'$, and $K \neq L$, for if $K = L$, then $\mathscr{M} = \{0\}$, since L is total for $L^p(\Omega, \Sigma, P)$. Hence a contradiction. The linearity of K is trivial, thus K is a proper submanifold of the linear manifold L. Now it will be shown that $M = K^0$. Since it follows from (1) and definitions of K^0 that $M \subset K^0$, only the opposite inclusion has to be shown. However, it is clear from (1) that the null space of each $F' \in K$ contains $\mathscr{M}$ and that $K \supseteq \mathscr{G}_1'$, therefore the intersection of null spaces of $F' \in K$ equals $\mathscr{M}$, since the intersection of null spaces of $F' \in \mathscr{G}_1'$ equals $\mathscr{M}$. Hence if $z \in K^0$, then by definition $z \in Z$, and $F'(z) = 0$, for every $F' \in K$. But the latter implies that $z \in \mathscr{M}$. Thus $z \in Z \cap \mathscr{M} = M$, i.e., $K^0 \subseteq M$. Thus $M = K^0$ as desired. Now by Lemma 2.1 (c), M is $\sigma(Z, L)$-closed.

From (a), (b), and (c), thus indeed M, Z, and L satisfy the hypothesis of Lemma 2.3. Hence there exists a second $\sigma(Z, L)$-closed linear submanifold N of Z such that

(i) $N = N^{00}$,

(ii) $M \oplus N = Z$,

(iii) $M^0 \oplus N^0 = L$.

Now let $\mathscr{N}$ be the closure of N in $L^p(\Omega, \Sigma, P)$, then $\Delta = \mathscr{M} \oplus \mathscr{N}$ is dense in $L^p(\Omega, \Sigma, P)$, where $(\mathscr{M}, \mathscr{N})$ is a pair of quasi-complements, in the sense of Murray [6]. For since $\mathscr{M} \supseteq M$, $\mathscr{N} \supseteq N$, and Z is dense in $L^p(\Omega, \Sigma, P)$, so $\Delta (= \mathscr{M} \oplus \mathscr{N} \supseteq M \oplus N = Z)$ is dense in $L^p(\Omega, \Sigma, P)$. Moreover, $\mathscr{M} \cap \mathscr{N} = \{0\}$, since $\mathscr{M} = \overline{M}$, hence every element in M^0 vanishes throughout $\mathscr{M}$. Similarly, $\mathscr{N} = \overline{N}$, and $N = N^{00}$ by (i) implies that every element in N^0 vanishes throughout $\mathcal{N}$. Hence from (iii) and the fact that L is total, it follows that $\mathcal{M} \cap \mathcal{N} = \{0\}$.

Observe the additional fact that $\Delta \supseteq \mathcal{E}$.

Thus by [6, Lemma 10b], there exists a closed projection $P_{\beta_T} : \Delta \to \mathcal{M}$ such that if $h \in \Delta$, and $h = f + g$, where $f \in \mathcal{M}$, $g \in \mathcal{N}$ is the unique decomposition of h relative to $\mathcal{M}$ and $\mathcal{N}$ as the sum of an element from $\mathcal{M}$ and one from $\mathcal{N}$, then one has

$$P_{\beta_T}(h) = f.$$

In particular, since $X(\tilde{\varphi}_i) \in \Delta$, and $X(\tilde{\varphi}_i) = Y_i + Z_i$, where $Y_i \in \mathcal{M}$, $Z_i \in \mathcal{N}$, for $i \geqslant 1$, one has

$$P_{\beta_T} X(\tilde{\varphi}_i) = Y_i .$$

In other words, P_{β_T} is the prediction operator for the problem.

To prove the second half of this theorem, one proceeds as follows:

Denote $\mathcal{N}' = \bar{N}^0$. Thus $\mathcal{N}'$ is a subspace of $(L^p(\Omega, \Sigma, P))'$. Let

$$\mathcal{V}' = \{h' : \| h' \|_q \leqslant \| h' + g' \|_q \text{, for every } g' \in \mathcal{N}'\},$$

i.e., $\mathcal{V}'$ is the set of elements of $(L^p(\Omega, \Sigma, P))'$ minimal relative to $\mathcal{N}'$. Now for every $f' \in (L^p(\Omega, \Sigma, P))'$, $f' = g' + h'$, where $g' \in \mathcal{N}'$, $h' \in \mathcal{V}'$ is the unique decomposition [6]. Now this decomposition holds in particular for every $f_i' \in \mathcal{G}' \subseteq L$, i.e., $f_i' = g_i' + h_i'$, $g_i' \in \mathcal{N}'$, $h_i' \in \mathcal{V}'$, $i \geqslant 1$.

Now denote $\mathcal{M}' = \bar{M}^0$, then $\Delta'(= \mathcal{M}' \oplus \mathcal{N}' \supseteq M^0 \oplus N^0 \supseteq L)$. Since L is dense in $(L^p(\Omega, \Sigma, P))'$, so Δ' is also dense in $(L^p(\Omega, \Sigma, P))'$. Moreover, by a similar argument, one can show that $\mathcal{M}' \cap \mathcal{N}' = \{0\}$, i.e., $(\mathcal{M}', \mathcal{N}')$ is a pair of quasi-complements. Note however that h_i' need not belong to M^0, but do belong to $\mathcal{M}'$.

One also has the additional fact that $\Delta' \supseteq \mathcal{G}'$.

Thus by [6, Lemma 10b], there exists a closed projection $\tilde{P}_{\beta_T} : \Delta' \to \mathcal{N}'$, such that if $f' \in \Delta'$, and $f' = g' + h'$, where $g' \in \mathcal{N}'$, $h' \in \mathcal{M}'$ is the unique decomposition of f' as an element from $\mathcal{N}'$ and one from $\mathcal{M}'$, then one has

$$f_i' = g_i' + h_i', \; g_i' \in \mathcal{N}', \; h_i' \in \mathcal{M}', \text{ for } i \geqslant 1,$$

and, moreover,

$$\tilde{P}_{\beta_T}(f_i') = g_i'.$$

However, from the definition of M, and the fact that L is dense in $L^p(\Omega, \Sigma, P)$, M in $\mathcal{M}$, one has that $\mathcal{M}'$ is simply the annihilator of $\mathcal{M}$, and $\mathcal{N}'$ that of $\mathcal{N}$. Thus by [6, Lemma 11], $\tilde{P}_{\beta_T}$ is also the proper adjoint of P_{β_T}. This concludes the proof of Theorem 2.1. Q.E.D.

Remark 2.3. The (linear operator) P_{β_T} need not be defined for all $X(\varphi)$, since Δ, the domain of P_{β_T}, is only dense in $L^p(\Omega, \Sigma, P)$.

Remark 2.4. In general, P_{β_T} is not a linear predictor, i.e., if $X(\tilde{\varphi}_1)$, $X(\tilde{\varphi}_2)$, and $\alpha X(\tilde{\varphi}_1) + \beta X(\tilde{\varphi}_2) \in \Delta$, for some scalars α, β, then

$$P_{\beta_T}(\alpha X(\tilde{\varphi}_1) + \beta X(\tilde{\varphi}_2)) = \alpha P_{\beta_T} X(\tilde{\varphi}_1) + \beta P_{\beta_T} X(\tilde{\varphi}_2)$$

need not be the best predictor for $\alpha X(\tilde{\varphi}_1) + \beta X(\tilde{\varphi}_2)$. (In fact, if $p \neq 2$, then P_{β_T} is not linear as shown in [8] for the ordinary process.)

Remark 2.5. Theorem 2.1 can be generalized to the case where the range space of $X(\cdot)$ is any separable reflexive rotund Banach space.

3. An Existence Theorem

In this section, the existence of a weak generalized random process, $Y : \mathscr{D}(Y) \to L^p(\Omega, \beta_T, P)$ such that $Y(\tilde{\varphi}_i) = Y_i$ for $i \geqslant 1$, where Y_i are the best predictors of $X(\tilde{\varphi}_i)$, and $\mathscr{D}(Y)$ is dense in (D), will be proved via the first half of Theorem 2.1; the second half of the theorem will not be needed.

Theorem 2.1 shows that $P_{\beta_T} X(\tilde{\varphi}_i)$ is the best nonlinear predictor for $X(\tilde{\varphi}_i)$ for $i \geqslant 1$. In this section, this will be taken for granted, and held fixed throughout the discussion.

Consider the following diagram:

$$(D) \xrightarrow{X} L^p(\Omega, \Sigma, P) \supset \Delta \xrightarrow{P_{\beta_T}} L^p(\Omega, \beta_T, P)$$

$$(D)' \xleftarrow{X'} (L^p(\Omega, \Sigma, P))' \xleftarrow{P'_{\beta_T}} \Delta' \subset (L^p(\Omega, \beta_T, P))',$$

where P'_{β_T} is the adjoint of P_{β_T}, a closed linear transformation [10, Theorem 7.1, p. 155; Corollary, p. 156; Theorem 3.1, p.130] with domain

$$\Delta' = \{y' \in (L^p(\Omega, \beta_T, P))' : x \to \langle P_{\beta_T}(x), y' \rangle \text{ is continuous on } \Delta\}$$

dense in $(L^p(\Omega, \beta_T, P))'$, and satisfying the relation

$$\langle x, P'_{\beta_T}(y') \rangle = \langle P_{\beta_T}(x), y' \rangle, \qquad \text{for every} \quad y' \in \Delta', \quad x \in \Delta,$$

and X' is the adjoint of X, a continuous linear transformation [10, Theorem 7.4, p. 158] defined by

$$\langle \varphi, X'f \rangle = \langle X(\varphi), f \rangle, \text{ for every } f \in (L^p(\Omega, \Sigma, P))', \ \varphi \in (D).$$

Consider now the operator,

$$Y' = X' \circ P'_{\beta_T} : \Delta' \to (D)'.$$

Since P'_{β_T} is a closed linear transformation and X' is a continuous linear transformation, Y' is also a closed linear transformation with domain Δ' dense in $(L^p(\Omega, \beta_T, P))'$, and range in $(D)'$.

Since (D) and $L^p(\Omega, \beta_T, P)$ are reflexive, and moreover the closure of a convex subset $C \subset (D)$ is the same for all l.c. topologies on (D) consistent with the duality $\langle (D), (D)' \rangle$ [10, Theorem 3.1, p. 130], one can conclude from [10, Corollary, p. 156] that

$$Y : \mathscr{D}(Y) \subset (D) \to L^p(\Omega, \beta_T, P),$$

the adjoint of Y', exists and is a closed linear transformation with domain

$$\mathscr{D}(Y) = \{\varphi \in (D) : y' \to \langle \varphi, Y'y' \rangle \text{ is continuous on } \Delta'\}$$

dense in (D) and range in $L^p(\Omega, \beta_T, P)$ such that

$$\langle Y(\varphi), y' \rangle = \langle \varphi, Y'y' \rangle, \text{ for every } \varphi \in \mathscr{D}(Y), \text{ and } y' \in \Delta'. \tag{2}$$

It is clear that $\mathscr{D}(Y) \supset \{\tilde{\varphi}_i\}_{i=1}^{\infty}$. For, for every $\tilde{\varphi}_i \in \{\tilde{\varphi}_i\}_{i=1}^{\infty}$, the linear functional

$$\begin{aligned} y' \to \langle \tilde{\varphi}_i, Y'y' \rangle &= \langle \tilde{\varphi}_i, X'P'_{\beta_T}y' \rangle \\ &= \langle X(\tilde{\varphi}_i), P'_{\beta_T}y' \rangle \\ &= \langle P_{\beta_T}X(\tilde{\varphi}_i), y' \rangle \end{aligned}$$

is certainly continuous on Δ', since $X(\tilde{\varphi}_i) \in \Delta$, for every $i \geqslant 1$. Hereafter, Y will be termed a *weak generalized random process*.

Remark 3.1. In general, Y is not unique since P_{β_T} is not unique [6, Theorem, p. 93].

Remark 3.2. In (2), it is tempting to say that

$$\langle Y(\varphi), y' \rangle = \langle \varphi, X'P'_{\beta_T}y' \rangle = \langle P_{\beta_T}X(\varphi), y' \rangle.$$

However, it is not *a priori* obvious that the range of X is contained in the domain Δ of P_{β_T}.

LEMMA 3.1. *Let* X, β_T, $L^p(\Omega, \Sigma, P)$, *etc., be as defined in Theorem* 2.1. *Let*

$Y : \mathscr{D}(Y) \to L^p(\Omega, \beta_T, P)$ be defined by Eq. (2). Then Y (a closed linear transformation) satisfies

$$Y(\tilde{\varphi}_i) = P_{\beta_T} X(\tilde{\varphi}_i),$$

for every $\tilde{\varphi}_i \in \{\tilde{\varphi}_i\}_{i=1}^{\infty}$.

Proof. The conclusion of the lemma follows from the fact that for every $y' \in \Delta'$,

$$\begin{aligned} \langle Y(\tilde{\varphi}_i), y' \rangle &= \langle \tilde{\varphi}_i, X' P'_{\beta_T} y' \rangle \\ &= \langle X(\tilde{\varphi}_i), P'_{\beta_T} y' \rangle \\ &= \langle P_{\beta_T} X(\tilde{\varphi}_i), y' \rangle. \end{aligned} \qquad \text{Q.E.D.}$$

This establishes the main result of this section:

THEOREM 3.1. *Let $X : (D) \to L^p(\Omega, \Sigma, P)$ be a generalized random process, $\{\tilde{\varphi}_i\}_{i=1}^{\infty} \subset (D)$ be such that $\sigma(\tilde{\varphi}_i) \not\subset T$. Then there exists a weak generalized random process, $Y : \mathscr{D}(Y) \to L^p(\Omega, \beta_T, P)$ such that $\mathscr{D}(Y)$ is dense in (D) and $Y(\tilde{\varphi}_i)$ $(= P_{\beta_T} X(\tilde{\varphi}_i))$ gives the best nonlinear predictor for $X(\tilde{\varphi}_i)$, for $i \geqslant 1$.*

The existence problem of the prediction theory proposed in introduction is thus solved. In the next section, an approximation theorem for Y will be given.

4. An Approximation Theorem

Let $\{\beta_n\}_{n=1}^{\infty}$ be the sequence of increasing sub σ-fields as defined in Section 1 such that $\beta_n \uparrow \beta_T$ [cf. Lemma 1.2], and Y_n, Y be the corresponding weak g.r.p.'s as given by

$$\langle Y_n(\varphi), y' \rangle = \langle \varphi, X' P'_{\beta_n} y' \rangle, \qquad \text{for every} \quad y' \in \Delta_n', \quad \varphi \in \mathscr{D}(Y_n),$$

where Δ_n' is the domain of P'_{β_n}, and $\mathscr{D}(Y_n)$ is the domain of Y_n dense in (D). Note that, since $\Delta_n \supset \{X(\tilde{\varphi}_i)\}_{i=1}^{\infty}$, hence $\mathscr{D}(Y_n) \supset \{\tilde{\varphi}_i\}_{i=1}^{\infty}$, for every $n \geqslant 1$. So for every $\tilde{\varphi}_i \in \{\tilde{\varphi}_i\}_{i=1}^{\infty}$ Theorem 3.1 implies that $Y_n(\tilde{\varphi}_i) = P_{\beta_n} X(\tilde{\varphi}_i)$ a.e. for every $n \geqslant 1$. Hence it follows from Lemma 1.3, Lemma 1.4, and the ordinary theory [8] that $Y_n(\tilde{\varphi}_i) \to Y(\tilde{\varphi}_i)$ in L^p-norm and also $Y_n(\tilde{\varphi}_i) \to Y(\tilde{\varphi}_i)$ a.e. for every $\tilde{\varphi}_i \in \{\tilde{\varphi}_i\}_{i=1}^{\infty}$. Thus one has proved

THEOREM 4.1. *Let $X : (D) \to L^p(\Omega, \Sigma, P)$ be a generalized random process,* *and Y_n, Y be the weak* g.r.p.'s *as given by* (2), *corresponding to β_n, β_T where $\beta_n \uparrow \beta_T$. Then for every $\tilde{\varphi}_i \in \{\tilde{\varphi}_i\}_{i=1}^{\infty}$, $Y_n(\tilde{\varphi}_i) \to Y(\tilde{\varphi}_i)$ in L^p-norm, and also* $Y_n(\tilde{\varphi}_i) \to Y(\tilde{\varphi}_i)$ a.e.

5. The Special Case of $p = 2$

In the case when $p = 2$, $L^2(\Omega, \Sigma, P)$ is a Hilbert space. Thus if $\mathscr{M} = L^2(\Omega, \beta_T, P)$ and $X(\tilde{\varphi}_i) \in L^2(\Omega, \Sigma, P)$, $d_i = \inf\{\| X(\tilde{\varphi}_i) - Y \|_2 : Y \in \mathscr{M}\}$, then again there exists $Y_i^0 \in \mathscr{M}$ such that d_i is attained. In fact, Y_i^0 is obtained by taking it to be the orthogonal projection, π_{β_T}, of $X(\tilde{\varphi}_i)$ onto $\mathscr{M}$. π_{β_T} is independent of $\tilde{\varphi}_i$, since it only depends on $\mathscr{M}$. If

$$\mathscr{V} = \{h : \| h \|_2 \leqslant \| g + h \|_2 , g \in \mathscr{M}\},$$

then one can show that $\mathscr{V} = \mathscr{M}^{\perp}$. In fact, this is just the content of the projection theorem in Hilbert Space, and then the quasi-complements reduce to orthogonal complements.

Thus for every $f \in L^2(\Omega, \Sigma, P)$, $f = g + h$, where $g \in \mathscr{M}, h \in \mathscr{M}^{\perp}$ and $P_{\beta_T} f = g$, and P_{β_T} coincides with the orthogonal projection which is independent of $\varphi \in (D)$. Thus as a consequence of Theorem 3.1, Y is a g.r.p. such that for every $\varphi \in (D)$, $Y(\varphi) = \pi_{\beta_T} X(\varphi)$ is the best nonlinear predictor relative to $\mathscr{M}$ of $X(\varphi)$. In this case, the g.r.p., Y, is the best nonlinear predictor for the problem and the prediction operator $P_{\beta_T} = \pi_{\beta_T}$ is actually linear.

One should observe that when the g.r.p. X is real Gaussian with zero mean functional, then the linear least-square predictor coincides with the nonlinear predictor as in the scalar case [1, p. 561].

Remark 5.1. If $\{\beta_n\}_{n=1}^{\infty}$ is an increasing sequence of sub σ-fields of Σ such that $\beta_n \uparrow \beta_T$, then it is obvious from the above discussion that $Y_n \to Y$ in the topology of simple convergence, i.e., $Y_n(\varphi) \to Y(\varphi)$ in L^2-norm for every $\varphi \in (D)$, and also $Y_n(\varphi) \to Y(\varphi)$ a.e.. This result can also be deduced from the generalized martingale convergence theorem (which will be published shortly).

Remark 5.2. The results proved here can easily be generalized to multidimensional generalized random fields.

Acknowledgment

This is part of the author's thesis completed at Carnegie-Mellon University, Pittsburgh, Pennsylvania. The author wishes to thank his thesis advisor Professor M. M. Rao for many hours of stimulating discussion.

REFERENCES

1. J. L. DOOB, "Stochastic Processes," Wiley & Sons, New York, 1965.
2. N. DUNFORD AND J. T. SCHWARTZ, "Linear Operators, Part I: General Theory," Interscience, New York, 1958.
3. T. HIDA, Canonical representation of Gaussian processes and their applications, *Mem. College Sci., Univ. Kyoto, Ser. A* **33** (1960), 109–155.
4. J. L. KELLEY AND I. NAMIOKA, "Linear Topological Spaces," Van Nostrand, Princeton, 1963.
5. G. W. MACKEY, Note on a theorem of Murray, *Bull. Amer. Math. Soc.* **52** (1946), 322–325.
6. F. J. MURRAY, Quasi-complements and closed projections in reflexive Banach spaces, *Trans. Amer. Math. Soc.* **58** (1945), 77–95.
7. W. J. PERVIN, "Foundations of General Topology," Academic Press, New York, 1964.
8. M. M. RAO, Inference in stochastic processes—III, *Z. Wahrscheinlichtkeitstheorie und Verw. Gebiete* **8** (1967), 49–72.
9. M. M. RAO, Conditional expectations and closed projections, *Indag. Math.* **27** (1965), 100–112.
10. H. H. SCHAEFER, "Topological Vector Spaces," MacMillan, New York, 1966.
11. Z. ŠIDAK, Relations between strict sense and wide sense conditional expectations, *Theor. Probability Appl.* **2** (1957), 283–288.