PREDICTION THEORY FOR NON-STATIONARY GENERALIZED STOCHASTIC PROCESSES

By José L. Abreu

Let $D(\boldsymbol{R})$ denote the space of infinitely differentiable complex valued functions on the real line $\boldsymbol{R}$ with compact support, equipped with the topology of L. Schwartz [1]. In this paper we give some results concerning the linear least squares prediction theory of generalized stochastic processes which are actually continuous linear maps from $D(\boldsymbol{R})$ into some Hilbert space H (see [2, 3 and 4] for background).

An m-dimensional generalized random field $X = (X_1, \cdots, X_m)$ is a continuous linear map $X: D(\boldsymbol{R}) \to H^m$ where H is a Hilbert space and H^m denotes the cartesian product of m copies of H. $X_1, \cdots, X_m$ are the components of X and $X_k: D(\boldsymbol{R}) \to H$ is a generalized stochastic process for each $k = 1, \cdots, m$. Let $\langle \ , \ \rangle$ denote the inner product and $\| \quad \|$ the norm of H. We define the covariance matrix of X as $B = (B_{ij})$ where $B_{ij}(f \otimes \bar{g}) = \langle X_i f, X_j g\rangle$ for $i, j = 1, \cdots, m; f, g \in D(\boldsymbol{R})$. (Here $\otimes$ denotes tensor product (see [5]) and $\overline{}$ denotes complex conjugation). The existence of the covariance distributions $B_{ij} \in D'(\boldsymbol{R}^2)$ is guaranteed by the theorem of kernels of L. Schwartz [1, 3].

For every $t \in \boldsymbol{R}$, let D_t be the subspace of $D(\boldsymbol{R})$ consisting of those functions with support contained in $(-\infty, t]$. Let H_t = closed linear span $\{X_k f : f \in D_t;\ k = 1, \cdots, m\}$ for each $t \in \boldsymbol{R}$, and also let H_∞ = closed linear span $\{X_k f : f \in D(\boldsymbol{R});\ k = 1, \cdots, m\}$.

For $t \in \boldsymbol{R}$ and also for $t = \infty$, let $P_t : H \to H_t$ denote the orthogonal projections. X is said to be deterministic if $H_t = H_\infty$ for every $t \in \boldsymbol{R}$; X is called linearly free if $\bigcap_{t\in\mathbf{R}} H_t = H_{-\infty} = \{0\}$. (The characterization of deterministic and linearly free generalized random fields on $\boldsymbol{R}$ in terms of their covariance matrices has been studied by Deo [4].).

The prediction problem for X consists in expressing the projections $P_t X_k f$ in terms of linear combinations of $X_j g$ with j = $1, \cdots . m$ and $g \in D_t$.

Theorem 1. *Let $X = (X_1, \cdots, X_m)$ be a generalized random field on $\boldsymbol{R}$ with covariance matrix $B = (B_{ij})$. Then for every t in $\boldsymbol{R}$ there exist functions f_{ij} in D_t, $j = 1, \cdots, m; i = 1, 2, \cdots$* such that for each $n = 1, \cdots, m$,

$$P_t X_n f = \sum_{i=1}^{\infty} \sum_{j=1}^{m} \sum_{k=1}^{m} B_{nj}(f \otimes \bar{f}_{ij}) X_k(f_{ij}) \tag{1}$$

for every $f \in D(\boldsymbol{R})$.

Proof. First note that for each $t \in \boldsymbol{R}$,

$$H_t = \text{c.l.s.}\ \{\textstyle\sum_{i=1}^{m} X_i f_i : f_i \in D_t,\ i = 1, \cdots, m\}.$$

Now let D^m be the cartesian product of m copies of $D(\boldsymbol{R})$, and for each t in $\boldsymbol{R}$, let D_t^m be the cartesian product of m copies of D_t. We introduce a continuous linear map $\tilde{X} = D^m \to H$ given by $\tilde{X}(f) = \sum_{n=1}^{m} X_n f_n$ for each $f = (f_1, \cdots, f_m) \in D^m$. We also introduce an inner product C on D^m given by $C(f, \mathbf{g}) = \langle \tilde{X} f, \tilde{X} \mathbf{g} \rangle$ for each f, $\mathbf{g}$ in D^m. Define $\| f \|_C = C(f, f)^{1/2}$, and let $\ker C = \{f \in D^m : \| f \|_C = 0\}$. Let K be the Hilbert space obtained by completing the quotient space $D^m/\ker C$ with respect to the norm $\| \quad \|_C$. Finally define $V: D^m \to K$ to be the canonical map and let $K(t)$ be the closure of $V(D_t^m)$ in K.

It is possible to find $f_1, f_2, \cdots$ in D_t^m such that $V(f_1)$, $V(f_2)$, $\cdots$ is an orthonormal basis for $K(t)$. Then it is clear that $\tilde{X}(f_1)$, $\tilde{X}(f_2)$, $\cdots$ is an orthonormal basis for H_t and therefore

$$P_t \tilde{X}(f) = \sum_{i=1}^{\infty} C(f, f_i) \tilde{X}(f_i) \tag{2}$$

for every f in D^m. Now choosing $f = (f_1, \cdots, f_m)$ such that $f_k = 0$ if $k \neq n$ and $f_n = f$, we obtain equation (1) with $f_i = (f_{i1}, \cdots, f_{im})$ for $i = 1, 2, \cdots$ Q.E.D.

Theorem 1 provides an explicit solution to the prediction problem of X in terms of a sequence of values of the past of X.

The generalized random field X is stationary if its covariance matrix $B = (B_{ij})$ satifies: $B_{ij}(f \otimes \bar{g}) = B_{ij}(T_s f \otimes \overline{T_s g})$ for every s in $\mathbf{R}$; $i, j = 1, \cdots, m$ and f, g in $D(\mathbf{R})$. Here T_s denotes the translation operator in $D(\mathbf{R})$, i.e. $T_s f(t) = f(t - s)$ for $f \in D(\mathbf{R})$; $s, t \in \mathbf{R}$.

Corollary 1. *If $X = (X_1, \cdots, X_m)$ is a stationary generalized random field on $\mathbf{R}$ with covariance matrix $B = (B_{ij})$ then there exist functions f_{ij} in D_0 for $i = 1, 2, \cdots; j = 1, \cdots, m$ such that for every $t \in \mathbf{R}$, $f \in D(\mathbf{R})$, $n = 1, \cdots, m$,*

$$P_t X_n(f) = \sum_{i=1}^{\infty} \sum_{j=1}^{m} \sum_{k=1}^{m} B_{nj}(f \otimes \overline{T_t f_{ij}}) X_k(T_t f_{ij}) \tag{3}$$

Proof. It is easy to see that if $\{V f_i\}$ is an orthonormal basis for $K(0)$, then $\{V(T_t f_i)\}$ is an orthonormal basis for $K(t)$. Therefore putting $f_i = (f_{i1}, \cdots, f_{im})$ for every $i = 1, 2, \cdots$ for some sequence $\{f_i\} \subset D_0$ such that $\{V f_i\}$ is an orthonormal basis for $K(0)$, we obtain (3). Q.E.D.

If $m = 1$, theorem 1 and corollary 1 give the solution to the prediction problem of generalized stochastic processes. The stationary case was treated in a different way by Rozanov [6].

Now we give some results which characterize deterministic covariance distributions. (The covariance matrix of a generalized stochastic process consists of a single distribution $B \in D'(\mathbf{R}^2)$ which is called the covariance distribution of the process).

It is not difficult to see that every covariance distribution B has representations of the form $B = \sum_{n=1}^{\infty} \bar{T}_n \otimes T_n$ with $T_n \in D'(\mathbf{R})$ for $n = 1, 2, \cdots$; $\bar{T}(f) = \overline{T(\bar{f})}$ for every $f \in D(\mathbf{R})$, $T \in D'(\mathbf{R})$; and $T_n(f) = C(f, g_n)$ for every $f \in D(\mathbf{R})$, where $\{g_n\}$ is an orthonormal basis for the Hilbert space K obtained by completion of $D(\mathbf{R})/\ker C$ with respect to the norm $\| \quad \|_C$. (C and $\| \quad \|_C$ were defined in the proof of Theorem 1). Such a representation is called an orthonormal representation of B.

THEOREM 2. *If* $B = \sum_{n=1}^{\infty} \bar{T}_n \otimes T_n$ *is an orthonormal representation of the covariance distribution* B, *then* B *is deterministic if and only if whenever* c_n *are complex numbers such that* $0 < \sum_{n=1}^{\infty} |c_n|^2 < \infty$, *the support of the distribution* $\sum_{n=1}^{\infty} c_n T_n$ *extends to* $-\infty$, *i.e. it is not contained in* (t, ∞) *for any real number* t.

Proof. For each $t \in \boldsymbol{R}$ let $K(t)$ be the closure in K of $V(D_t)$ where $V:D(\boldsymbol{R}) \to K$ is the natural map. Let $W:K \to \ell^2$ be the map given by $W(f) = \{\bar{T}_n f\}$. W is a unitary isomorphism. The map $W \circ V:D(\boldsymbol{R}) \to \ell^2$ is a generalized stochastic process with covariance distribution B. Hence B is deterministic if and only if $W(K(t)) = \ell^2$ for every real t; i.e. B is deterministic if and only if for every non-zero sequence $\{c_n\} \in \ell^2$, $\{c_n\}$ is not orthogonal to $W(K(t))$. Since $V(D_t)$ is dense in $K(t)$, we conclude that B is deterministic if and only if whenever $\{c_n\}$ is a non-zero element of ℓ^2, and t is a real number, there exists $f \in D_t$ such that $\sum_{n=1}^{\infty} c_n T_n f \neq 0$. Q.E.D.

PROPOSITION 1. *A necessary and sufficient condition for a representation* $B = \sum_{n=1}^{\infty} \bar{T}_n \otimes T_n$ *of a covariance distribution* B *to be an orthonormal representation is that for every sequence* $\{c_n\} \in \ell^2$, $\sum_{n=1}^{\infty} c_n T_n = 0$ *necessarily implies* $c_n = 0$ for $n = 1, 2, \cdots$ (*i.e. the* T_n*'s are Hilbert-free*).

Proof. The necessity is obvious. To prove the sufficiency suppose the T_n's are Hilbert-free. There exists a sequence $\{g_n\} \subset K$ such that $\bar{T}_n f = C(Vf, g_n)$ for $n = 1, 2, \cdots; f \in \boldsymbol{D}(\boldsymbol{R})$. Furthermore $\{\sum_{i=1}^{m} x_i g_i : x_i \in \boldsymbol{C}, i = 1, \cdots, m; m = 1, 2, \cdots\}$ is dense in K, because the g_n's are Hilbert-free. Let ℓ_0^2 be the subspace of ℓ^2 consisting of those sequences which have only finitely many non-zero elements. The natural map $U:\ell_0^2 \to K$ given by $U\{x_i\} = \sum_{i=1}^{\infty} x_i g_i$ is weakly bounded. Indeed, if $\{x_i\} \in \ell_0^2$ and $\sum_{i=1}^{\infty} |x_i|^2 \leq 1$, then $|\sum_{i=1}^{\infty} x_i T_i f|^2 \leq \sum_{i=1}^{\infty} |x_i|^2 \sum_{i=1}^{\infty} |T_i f|^2 \leq B(f \otimes \bar{f})$ for every $f \in D(\boldsymbol{R})$. Thus the image of the unit ball of ℓ_0^2 under U is weakly bounded, and by Mackey's theorem (see theorem 3.5.3 of [5]) it is also strongly bounded. Therefore U may be continuously extended to ℓ^2. Since the g_n's are Hilbert-free, U is injective; and since the closed linear subspace generated by $\{g_n\}$ coincides with K, it follows that U is onto. Now given $f \in D(\boldsymbol{R})$, $\{\bar{T}_i f\}$ belongs to ℓ^2 and for every $g \in D(\boldsymbol{R})$ we have

$$C(U\{\bar{T}_i f\}, Vg) = C(\textstyle\sum_{i=1}^{\infty} (\bar{T}_i f) g_i, Vg) = \sum_{i=1}^{\infty} \bar{T}_i f C(g_i, Vg) = \sum_{i=1}^{\infty} \bar{T}_i f T_i \bar{g}$$
$$= B(f \otimes \bar{g}) = C(Vf, Vg)$$

Hence $U\{\bar{T}_i f\} = Vf$ for every $f \in D(\mathrm{R})$. Therefore for $f, g \in D(\boldsymbol{R})$

$$C(U\{\bar{T}_i f\}, U\{\bar{T}_j g\}) = C(Vf, Vg) = \sum_{i=1}^{\infty} \bar{T}_i f T_i \bar{g} = \sum_{i=1}^{\infty} \bar{T}_i f (\bar{T}_i g)$$

Since $V(D(\boldsymbol{R}))$ is dense in K we conclude that for every $\{x_i\}, \{y_i\} \in \ell^2$, $C(U\{x_i\}, U\{y_j\}) = \sum_{i=1}^{\infty} x_i \bar{y}_i$.

We just proved that U is a unitary isomorphism from ℓ^2 onto K, and $\{g_n\}$ is the image under U of the canonical orthonormal basis in ℓ^2. Hence $\{g_n\}$ is an orthonormal basis for K. Q.E.D.

We say that a covariance distribution B is degenerate if the corresponding space K defined above is finite dimensional. It follows that B is degenerate if and only if $B = \sum_{n=1}^{N} \bar{T}_n \otimes T_n$ for some integer N and $T_n \in D'(\mathrm{R})$ for $n = 1, 2, \cdots, N$.

COROLLARY 2. *$B = \sum_{n=1}^{N} \bar{T}_n \otimes T_n$ is an orthonormal representation of the degenerate covariance distribution B if and only if the T_n's are linearly independent. If $B = \sum_{n=1}^{N} \bar{T}_n \otimes T_n$ is an orthonormal representation of B, then B is deterministic if and only if the restrictions of the T_n's* to *D_t are linearly independent for any real number t.*

The proof follows easily from theorem 2 and proposition 1.

This paper is part of the Ph.D. Thesis which was presented by the author to the Department of Mathematics of the Massachusetts Institute of Technology on June, 1970. The author wishes to express his gratitude to Dr. Richard M. Dudley for his guidance in the preparation of the original manuscript.

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

REFERENCES

[1] L. SCHWARTZ, Théorie des Distributions, Hermann, Paris, 1966.
[2] K. ITO, *Stationary random distributions,* Mem. of the College of Science, University of KYOTO, Ser. A. 28, No. **3** (1953), 209–23.
[3] I. M. GELFAND AND N. YA. VILENKIN, Generalized Functions, Vol. **4**, Applications of Harmonic Analysis, Academic Press, New York and London, 1964.
[4] CH. M. DEO, *Prediction theory for non-stationary processes,* Sankyā: The Indian Journal of Statistics, **27** (1965) 113–32.
[5] J. HORVÁTH, Topological Vector Spaces and Distributions, Vol. 1, Addison-Wesley, Reading, Mass., 1966.
[6] YU. A. ROZANOV, *On the extrapolation of generalized stationary random processes.* Theor. Probability Appl., **4** (1959)—(Translated from Russian), 426–31.