The Annals of Probability
1978, Vol. 6, No. 5, 891-898

AN ALGORITHM FOR LINEAR PREDICTION OF A BANACH SPACE VALUED STATIONARY STOCHASTIC PROCESS

BY SHASHI PHOHA

The MITRE Corporation and Bucknell University

Wiener and Masani describe a procedure for relating nonlinear prediction of a univariate random process to linear prediction of an infinite-variate process which may not be a Hilbert-space-valued process but may be Banach-space-valued instead. An algorithm for computation of the linear predictor and the generating function of a Banach-space-valued stationary stochastic process is obtained under an extension of the boundedness condition of Wiener and Masani on the spectral density of the process.

1. Introduction and notation. If $\mathscr{X}$ and $\mathscr{Y}$ are two Banach spaces, $\mathscr{B}(\mathscr{X}, \mathscr{Y})$ denotes the Banach space of all bounded linear operators from $\mathscr{X}$ to $\mathscr{Y}$, and $\mathscr{X}^*$ denotes the Banach space of all conjugate linear functionals on $\mathscr{X}$. A bisequence $\{\xi_\eta: -\infty < n < \infty\}$ of elements of $\mathscr{B}(\mathscr{X}, \mathscr{K})$ where $\mathscr{X}$ is a Banach space and $\mathscr{K}$ is a Hilbert space, is called a $\mathscr{B}(\mathscr{X}, \mathscr{K})$-valued weakly stationary stochastic process if the operator $\xi_m{}^*\xi_n$ in $\mathscr{B}(\mathscr{X}, \mathscr{X}^*)$ depends only on $n - m$. And then the operator sequence $R(n) = \xi_0{}^*\xi_n$ defined for $-\infty < n < \infty$ is called the covariance bisequence of the process. Assume that $\mathscr{X}$ is separable. With this stationary stochastic process are associated the following subspaces ([6], page 922):

M_∞, the closed subspace of $\mathscr{K}$ spanned by $\{\xi_k(x): -\infty < k < \infty, x \in \mathscr{X}\}$,

M_n, the closed subspace of $\mathscr{K}$ spanned by $\{\xi_k(x): -\infty < k \leqq n, x \in \mathscr{X}\}$,

and

$$M_{-\infty} = \bigcap_{-\infty<n<\infty} M_n .$$

The process $\{\xi_n: -\infty < n < \infty\}$ is said to be

(i) singular if $M_{-\infty} = M_n$ for $-\infty < n < \infty$;
(ii) nondeterministic if $M_{-\infty} \neq M_n$ for some finite n;
(iii) regular if $M_{-\infty} = \{0\}$.

For $\mathscr{S} \subset \mathscr{B}(\mathscr{X}, \mathscr{K})$ let $\bar{\sigma}(\mathscr{S})$ denote the smallest (strongly) closed subspace of $\mathscr{B}(\mathscr{X}, \mathscr{K})$ containing the set $\{SB: S \in \mathscr{S}, B \in \mathscr{B}(\mathscr{X}, \mathscr{X})\}$ and $\sigma(\mathscr{S})$ denote the smallest closed subspace of $\mathscr{K}$ containing the set $\{Sx: S \in \mathscr{S}, x \in \mathscr{X}\}$. We shall use the same notation also for subsets $\mathscr{S}$ of $\mathscr{B}(\mathscr{K}, \mathscr{K})$. In this notation

Received March 25, 1977; revised October 20, 1977.
AMS 1970 *subject classifications.* Primary 60G25; Secondary 60G20.
Key words and phrases. Spectral density, boundedness condition, factorization, algorithm for the predictor, prediction error matrix.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.*
www.jstor.org

then

$$\bar{\sigma}(\mathscr{S}) = \mathscr{B}(\mathscr{X}, \sigma(\mathscr{S})) \qquad \text{([6], page 922).}$$

For the stationary stochastic $\mathscr{B}(\mathscr{X}, \mathscr{K})$-valued process $\{\xi_n: -\infty < n < \infty\}$ define

$$\mathscr{M}_n = \bar{\sigma}\{\xi_k: k \leqq n\},$$
$$\mathscr{M}_{-\infty} = \bigcap_n \mathscr{M}_n,$$
$$\mathscr{M}_\infty = \bar{\sigma}\{\xi_k: -\infty < k < \infty\}.$$

In the above notation then, $\mathscr{M}_n = \mathscr{B}(\mathscr{X}, M_n)$, $-\infty \leqq n \leqq \infty$. Furthermore let $\mathscr{B}_n$, B_n for $-\infty \leqq n \leqq \infty$ denote corresponding subspaces for a $\mathscr{B}(\mathscr{K}, \mathscr{K})$-valued stationary process $\{\eta_n: -\infty < n < \infty\}$.

Now for each $\xi_n \in \mathscr{B}(\mathscr{X}, \mathscr{K})$, there exists an operator $(\xi_n | \mathscr{M}_0)$ in $\mathscr{B}(\mathscr{X}, M_0)$ such that $\xi_n - (\xi_n | \mathscr{M}_0)$ is orthogonal to $\mathscr{M}_0$ ([4], Theorem 3.2.5, page 10). $(\xi_n | \mathscr{M}_0)$ is called the projection of ξ_n on $\mathscr{M}_0$, and is denoted by $\hat{\xi}_n$. Similarly define $\hat{\eta}_n$ as $(\eta_n | \mathscr{B}_0)$, $-\infty < n < \infty$. The operator $G = (\xi_0 - \hat{\xi}_0)^*(\xi_0 - \hat{\xi}_0)$ is called the predictor error operator of the process. The process is said to be of full rank if G is boundedly invertible.

Time domain and spectral analysis for such processes, as given below, were obtained by A. G. Miammee [6]. However an extension of the algorithm of Weiner and Masani ([7], 6, pages 123–127) under the boundedness condition ([5], page 1), was obtained only for Hilbert-space-valued random variables using Fourier analysis of infinite matrix valued functions. In dearth of an obvious identity operator in the family of $\mathscr{B}(\mathscr{X}, \mathscr{X}^*)$-valued functions, a suitable boundedness condition on the spectral density of the process must be obtained to work out the corresponding algorithm for prediction.

Time domain analysis. For a $\mathscr{B}(\mathscr{X}, \mathscr{K})$-valued regular stationary stochastic process $\{\xi_n: -\infty < n < \infty\}$ there exist mutually orthogonal isometries S_k and $A_k \in \mathscr{B}(\mathscr{X}, \mathscr{K})$ such that

$$\xi_n = \sum_{k=0}^{\infty} S_{n-k} A_k,$$

convergence being in the strong operator topology ([6] Theorem 3.5, page 924).

Spectral analysis. For the $\mathscr{B}(\mathscr{X}, \mathscr{K})$-valued weakly stationary stochastic process $\{\xi_n: -\infty < n < \infty\}$, the shift operator $\mathscr{U}$ defined on M_∞ as follows

$$\mathscr{U}\xi_n x = \xi_{n+1} x, \qquad x \in \mathscr{X}, \ -\infty < n < \infty$$

has a spectral resolution $\mathscr{U} = 1/2\pi \int_0^{2\pi} e^{-i\theta} E(d\theta)$, where E is a projection valued measure over $([0, 2\pi), \mathscr{B})$, $\mathscr{B}$ being the algebra of Borel sets ([2], pages 359–360). Now define for $B \in \mathscr{B}$, $F(B) = \xi_0^* E(B) \xi_0$. This F is called the spectral distribution function of $\{\xi_n: -\infty < n < \infty\}$. Assume now that there exists a nonnegative $\mathscr{B}(\mathscr{X}, \mathscr{X}^*)$-valued function $f(\theta)$ defined on $[0, 2\pi)$ such that

(i) $f(\theta)$ is strongly measurable;

(ii) $f(\theta)$ is Bochner integrable; and

(iii) for each Borel measurable $B \subset [0, 2\pi)$, $F(B) = \int_B f(\theta)\,d\theta$.

This $f(\theta)$, also denoted by f_θ, will then be called the spectral density of the process $\{\xi_n : -\infty < n < \infty\}$. Let $L_2(\mathscr{K})$ denote the Hilbert space of all $\mathscr{K}$-valued scalarly measurable functions on the unit circle which have square summable norm. The $L_2(\mathscr{K})$ inner product of two functions g_1 and g_2 is

$$\frac{1}{2\pi} \int_0^{2\pi} (g_1(e^{i\theta}), g_2(e^{i\theta}))\,d\theta \,.$$

Then a $\mathscr{B}(\mathscr{H}, \mathscr{K})$-valued function $A(e^{i\theta})$ defined on the unit circle is said to be conjugate analytic if

$$\forall\, x \in \mathscr{H}, \qquad A(e^{i\theta})(x) \in \left\{ g \in L_2(\mathscr{K}) : \frac{1}{2\pi} \int e^{-in\theta} g(e^{i\theta})\,d\theta = 0 \ \text{ for } \ n > 0 \right\} .$$

The spectral density $f(\theta)$ is said to be factorable if there exists a conjugate analytic $\mathscr{B}(\mathscr{H}, \mathscr{K})$-valued function $A(e^{i\theta})$ defined on the unit circle such that $f(e^{i\theta}) = A^*(e^{i\theta})A(e^{i\theta})$, in the sense that $(f(e^{i\theta})x)(y) = (A(e^{i\theta})x, A(e^{i\theta})y)$ for all $x, y \in \mathscr{H}$.

Regarding factorization of the spectral density of a $\mathscr{B}(\mathscr{H}, \mathscr{K})$-valued regular stochastic process, the following has been established in [6], Theorem 4.5, page 930.

1.1 THEOREM. *The spectral distribution F of a regular full rank $\mathscr{B}(\mathscr{H}, \mathscr{K})$-valued stationary stochastic process is absolutely continuous and*

$$\frac{d}{d\theta}(F(e^{i\theta})x)(x) = \|\Phi(e^{i\theta})x\|^2$$

where

$$\Phi(e^{i\theta})(x) = \textstyle\sum_{k=0}^{\infty} e^{-ik\theta} A_k(x) \,, \qquad A_k \in \mathscr{B}(\mathscr{H}, \mathscr{K})$$

and

$$G = A_0{}^* A_0 \,.$$

Φ, as defined in this theorem, is called the generating function of the process $\{\xi_n : -\infty < n < \infty\}$.

2. The boundedness condition on the spectral density. Let the spectral density f_θ, of the $\mathscr{B}(\mathscr{H}, \mathscr{K})$-valued stationary stochastic process $\{\xi_n : -\infty < n < \infty\}$ satisfy

$$0 < m(\theta)A^*A \leqq f_\theta \leqq M(\theta)A^*A \quad \text{a.e.} \quad \theta \in [0, 2\pi)$$

for some $A : \mathscr{H} \to \mathscr{K}$ with $\|A\| = 1$ and $M(\theta)$, $1/m(\theta)$, $M(\theta)/m(\theta)$, summable.

2.1 LEMMA. *Under boundedness condition* 2 *on the spectral density and* $N_\theta = f_\theta/a_\theta - A^*A$ *with* $a_\theta = (m(\theta) + M(\theta))/2$, $\|N_\theta\|_B \leqq (M(\theta) - m(\theta))/(M(\theta) + m(\theta)) < 1$.

PROOF. For each $x \in \mathscr{H}$

$$m(\theta)(Ax, Ax) \leqq (f(x), x) \leqq M(\theta)(Ax, Ax) \,,$$

i.e.,

$$\left(\frac{m(\theta)}{a_\theta} - 1\right)(A^*Ax, x) \leqq \frac{(f_\theta(x), x)}{a_\theta} - (A^*Ax, x) \leqq \left(\frac{M(\theta)}{a_\theta} - 1\right)(A^*Ax, x),$$

i.e.,

$$\frac{m(\theta) - M(\theta)}{m(\theta) + M(\theta)} \|Ax\|^2 \leqq (N_\theta(x), x) \leqq \frac{M(\theta) - m(\theta)}{M(\theta) + m(\theta)} \|Ax\|^2 .$$

Using the parallelogram law, we have for each $x, y \in \mathscr{X}$

$$\begin{aligned} |(N_\theta(x), y)| &= \tfrac{1}{4}\{(N_\theta(x+y), x+y) - (N_\theta(x-y), x-y)\}| \\ &\leqq \left| \frac{M(\theta) - M(\theta)}{M(\theta) + m(\theta)} \frac{(A^*Ax, y) + (A^*Ay, x)}{2} \right| . \end{aligned}$$

Since $\|A\| = 1$, $\|N_\theta\|_B \leqq (M(\theta) - m(\theta))/(M(\theta) + m(\theta)) < 1$.

2.2 LEMMA. *Let the spectral density f_θ satisfy the boundedness condition 2, and the image $A\mathscr{X}$ be dense in $\mathscr{K}$. If A is one-to-one onto $A\mathscr{X}$ then*

(i) *A^* is one-to-one*;

(ii) *$A^{*-1} = (A^{-1})^*$; and*

(iii) $|(A^{*-1}N_\theta A^{-1}(k), l)| \leqq (M(\theta) - m(\theta))/(M(\theta) + m(\theta))\|k\|\,\|l\|$ a.e. θ *for* $k, l \in A\mathscr{X}$.

PROOF. (i) Since $A\mathscr{X}$ is dense in $\mathscr{K}$, A^* is defined on $\mathscr{K}$ to $\mathscr{X}^*$ as follows: $\forall\, k \in \mathscr{K}$, $A^*(k) = x^*$ where $x^*(y) = (k, Ay)\ \forall\, y \in \mathscr{X}$. A^* is easily seen to be one-to-one.

(ii) In fact

$$\begin{aligned} \mathscr{D}((A^{-1})^*) &= \{x^* \in \mathscr{X}^* : \exists\, k \in \mathscr{K} \text{ with } x^*(A^{-1}l) = (k, l)\ \ \forall\, l \in Ax\} \\ &= \{x^* \in \mathscr{X}^* : \exists\, k \in \mathscr{K} \text{ with } x^*(y) = (k, Ay)\ \ \forall\, y \in x\} \\ &= \text{range of } \ A^* = \mathscr{D}(A^{*-1}) . \end{aligned}$$

Also for each x^* in $\mathscr{D}(A^{*-1})$, $A^{*-1}(x^*) = k \Leftrightarrow (A^{-1})^*(x^*) = k$.

(iii) Now from the proof of Lemma 2.1

$$|(N_\theta(x), y)| \leqq \frac{M(\theta) - m(\theta)}{M(\theta) + m(\theta)} \left| \frac{(Ax, Ay) + (Ay, Ax)}{2} \right| \quad \text{a.e. } \theta .$$

Thus for $k, l \in A\mathscr{X}$

$$\begin{aligned} &|(A^{*-1}N_\theta A^{-1}(k), l)| \\ &\quad = |(N_\theta A^{-1}k, A^{-1}l)| \\ &\quad \leqq \frac{M(\theta) - m(\theta)}{M(\theta) + m(\theta)} \left| \frac{(AA^{-1}k, AA^{-1}l) + (AA^{-1}k, AA^{-1}l)}{2} \right| \quad \text{a.e. } \theta. \end{aligned}$$

$$\therefore \quad \|A^{*-1}N_\theta A^{-1}\|_B \leqq \frac{M(\theta) - m(\theta)}{M(\theta) + m(\theta)} .$$

Hence the result.

3. Relationship with the case of Hilbert-space-valued random variables.

MAIN THEOREM 1. *If the spectral density f_θ satisfies the boundedness condition* 2 *and $A: \mathscr{X} \to \mathscr{K}$ is one-to-one and $A\mathscr{X}$ dense in $\mathscr{K}$ then there is a unique stationary stochastic process $\{\eta_n: -\infty < n < \infty\}$ which is $\mathscr{B}(\mathscr{K}, \mathscr{K})$-valued and is such that*

(i) $R_\eta(n) = A^{*-1}R_\xi(n)A^{-1}$ *on* $A\mathscr{X}$.

(ii) $f_\eta(\theta) = 2/(M(\theta) + m(\theta))[I_{\mathscr{K}} + A^{*-1}N_\theta A^{-1}] = A^{*-1}f_\xi(\theta)A^{-1}$ *on* $A\mathscr{X}$ *where* $I_{\mathscr{K}}$ *denotes the identity operator on* $\mathscr{K}$ *and* $f_\xi(\theta)$ *is* f_θ *in our previous notation.*

PROOF. By Lemma 2.2 $A^{*-1}f_\xi(\theta)A^{-1} = 2/(M(\theta) + m(\theta))[I_{\mathscr{K}} + A^{*-1}N_\theta A^{-1}]$ is a bounded operator defined on $A\mathscr{X}$. Let g_θ denote its unique continuous extension to $\mathscr{K}$. g_θ is then a nonnegative $\mathscr{B}(\mathscr{K}, \mathscr{K})$-valued continuous function defined on the unit circle. Further g_θ is strongly measurable since $f(\theta)$ is assumed to be so. Also by 2.2 (iii) $\|g_\theta\| \in L_1[0, 2\pi)$. Hence g_θ is Bochner integrable. Thus for any n, $\xi_n A^{-1}$ defined on $A\mathscr{X}$ is such that $\forall\, k \in A\mathscr{X}$

$$\begin{aligned}\|\xi_n A^{-1}(k)\|^2 &= \frac{1}{2\pi}\int_0^{2\pi} (A^{*-1}f_\theta A^{-1}(k), k)\, d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} (g_\theta(k), k)\, d\theta \\ &\leqq \|g\|_{L_1[0,2\pi)}\|k\|^2 < \infty\,.\end{aligned}$$

Hence $\xi_n A^{-1}$ admits a unique continuous extension to $\mathscr{K}$, say η_n. $\{\eta_n: -\infty < n < \infty\}$ is then a $\mathscr{B}(\mathscr{K}, \mathscr{K})$-valued stationary stochastic process. In fact $\{\eta_n: -\infty < n < \infty\}$ is stationary as shown by the following reasoning.

For $k, l \in \mathscr{K}$ we must show that $(\eta_n k, \eta_m l)$ depends only on $n - m$. Since $A\mathscr{X}$ is dense in $\mathscr{K}$, there exist sequences $\{x_p\}$, $\{y_q\}$ in $\mathscr{X}$ such that

$$A(x_p) \to k \quad \text{in} \quad \mathscr{K} \qquad \text{and} \qquad A(y_q) \to l \quad \text{in} \quad \mathscr{K}\,.$$

Then, since η_n and η_m are bounded

$$\begin{aligned}(\eta_n k, \eta_m l) &= \lim_{p\to\infty} (\eta_n(Ax_p), \eta_m(Ay_p)) \\ &= \lim_{p\to\infty} (\xi_n x_p, \xi_m y_p) \\ &= \lim_{p\to\infty} \frac{1}{2\pi}\int_0^{2\pi} \bar{e}^{i(n-m)\theta}(f_\theta x_p, y_p)\, d\theta\end{aligned}$$

which depends on m and n only through $n - m$.

Furthermore

$$(\eta_n k, \eta_m l) = \lim_{p\to\infty} \frac{1}{2\pi}\int_0^{2\pi} e^{-i(n-m)\theta}(f_\theta A^{-1}(Ax_p), A^{-1}(Ay_p))\, d\theta$$

$$\begin{aligned}(\eta_n k, \eta_m l) &= \lim_{p\to\infty} \frac{1}{2\pi}\int_0^{2\pi} e^{-i(n-m)\theta}(A^{*-1}f_\theta A^{-1}(Ax_p), Ay_p)\, d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{-i(n-m)}\{\lim_{p\to\infty} (A^{*-1}f_\theta A^{-1}(Ax_p), Ay_p)\}\, d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{-i(n-m)\theta}(g_\theta(k), (l))\, d\theta\,,\end{aligned}$$

the last two steps being true since g_θ is a bounded operator. Thus on $A(\mathscr{X})$

$$R_{\eta_n} = (\eta_n, \eta_0) = (\xi_n A^{-1}, \xi_0 A^{-1}) = A^{*-1} R_{\xi_n} A^{-1}$$
$$f_{\eta_n}(\theta) = A^{*-1} f_{\xi_n}(\theta) A^{-1},$$

and due to continuity of all functions involved, R_{η_n} and $f_{\eta_n}(\theta)$ are the unique continuous extensions of $A^{*-1}R_{\xi_n}A^{-1}$ and $A^{*-1}f_{\xi_n}(\theta)A^{-1}$ respectively to $\mathscr{K}$.

3.1 *Factorization of the spectral density.*

COROLLARY. *If $\Phi_\theta: \mathscr{K} \to \mathscr{K}$ is the generating function for the $\mathscr{B}(\mathscr{K}, \mathscr{K})$-valued stationary stochastic process $\{\eta_n: -\infty < n < \infty\}$ and if f_θ satisfies the boundedness condition 2, and $A: \mathscr{X} \to \mathscr{K}$ is one-to-one and $A\mathscr{X}$ dense in $\mathscr{K}$ then $\Phi_\theta A: \mathscr{X} \to \mathscr{K}$ is such that $f_\theta = (\Phi_\theta A)^*(\Phi_\theta A)$.*

PROOF. By Main Theorem I, g_θ is the unique continuous extension of $A^{*-1}f_\theta A^{-1}$. So that

$$\begin{aligned} f_\theta &= A^* g_\theta A \\ &= A^* \Phi_\theta^* \Phi_\theta A = (\Phi_\theta A)^*(\Phi_\theta A) . \end{aligned}$$

3.2 *The prediction error matrix and the predictor for a Banach-space-valued process.* For the $\mathscr{B}(\mathscr{K}, \mathscr{K})$-valued stationary stochastic process $\{\eta_n: -\infty < n < \infty\}$ a schematic algorithm to obtain the prediction error matrix G_η and the linear predictor $\hat{\eta}_\nu$ of η_ν for $\nu > 0$ based on the past $\{\eta_n: n \leqq 0\}$, is given in [5]. We shall now find the same for the $\mathscr{B}(\mathscr{X}, \mathscr{K})$-valued process $\{\xi_n: -\infty < n < \infty\}$.

MAIN THEOREM II. *The two stationary stochastic processes $\{\xi_n: -\infty < n < \infty\}$ and $\{\eta_n: -\infty < n < \infty\}$ are further related as follows:*

(i) *For each integer $\nu > 0$,*

$$\hat{\xi}_\nu = \hat{\eta}_\nu A .$$

(ii) $G_\xi = A^* G_\eta A$. *Note that $G_\eta = A_0^* A_0$ where $\Phi(e^{i\theta})(x) = \sum_{k=0}^\infty e^{-ik\theta} A_k(x)$ is the generating function of the process $\{\eta_n: -\infty < n < \infty\}$.*

(iii) $\hat{\xi}_\nu = \lim_{n\to\infty} \sum_{k=0}^n E_{\nu k} \xi_{-k}$ *where $E_{\nu k}$ is the* kth *Fourier coefficient of* $[e^{-i\nu\theta}\Phi(e^{i\theta})_{0+}\Phi^{-1}]$.

PROOF. (i) Note that $\xi_k = \eta_k A$ for each k. Also $A\mathscr{X}$ is dense in $\mathscr{K}$ and η_k is bounded. Therefore for each k

$$\sigma\{\xi_k x: x \in \mathscr{X}\} = \sigma\{\eta_k l: l \in \mathscr{K}\} .$$

Now for each $x \in \mathscr{X}$

$$\begin{aligned} \hat{\eta}_\nu A(x) &= (\eta_\nu \mid \mathscr{B}_0)(Ax) = (\eta_\nu Ax \mid B_0) = (\xi_\nu A^{-1}Ax \mid B_0) = (\xi_\nu x \mid M_0) \\ &= (\xi_\nu \mid \mathscr{M}_0)(x) = \hat{\xi}_\nu(x) . \end{aligned}$$

(ii) Now for each $x, y \in \mathscr{X}$

$$\begin{aligned}(A^*G_\eta A(x))(y) &= ((\eta_1 - \hat{\eta}_1)(Ax), (\eta_1 - \hat{\eta}_1)(Ay)) \qquad \text{(by definition of } G_\eta) \\ &= (\eta_1 Ax - \hat{\eta}_1 Ax, \eta_1 Ay - \hat{\eta}_1 Ay) \\ &= (\xi_1 A^{-1}(Ax) - \hat{\xi}_1(x), \xi_1 A^{-1}(Ay) - \hat{\xi}_1(y)) \\ &= (\xi_1 x - \hat{\xi}_1 x, \xi_1 y - \hat{\xi}_1 y) \\ &= (G_\xi(x))(y) .\end{aligned}$$

(iii) For each integer $\nu > 0$

$$\hat{\eta}_\nu(x) = \lim_{n\to\infty} (\textstyle\sum_{k=0}^n E_{\nu k}\eta_{-k})(x) \qquad \text{([4], Theorem 7.4.11, page 108).}$$

Therefore

$$\begin{aligned}\hat{\xi}_\nu(x) &= \hat{\eta}_\nu A(x) = \hat{\eta}_\nu(Ax) \\ &= \lim_{n\to\infty} (\textstyle\sum_{k=0}^n E_{\nu k}\eta_{-k})(Ax) \\ &= \lim_{n\to\infty} (\textstyle\sum_{k=0}^n E_{\nu k}\xi_{-k}A^{-1})(Ax) \\ &= \lim_{n\to\infty} \textstyle\sum_{k=0}^n E_{\nu k}\xi_{-k}A^{-1}(Ax) \\ &= \lim_{n\to\infty} \textstyle\sum_{k=0}^n E_{\nu k}\xi_{-k}(x) .\end{aligned}$$

4. Note. For results in this chapter the boundedness assumption 2 was made on the spectral density of the process and it was further assumed that the map $A: \mathscr{X} \to \mathscr{K}$ be one-to-one with the image of $\mathscr{X}$ dense in $\mathscr{K}$. The restriction of $A\mathscr{X}$ being dense in $\mathscr{K}$ is easily deleted by replacing $\mathscr{K}$ by the Hilbert space $\mathscr{H}$ generated by $A\mathscr{X}$ in defining the process $\{\eta_n: -\infty < n < \infty\}$. Generalization when A is not one-to-one calls for a closer look and may be handled as follows: let $K(P)$ denote the kernel of any operator P. Then due to boundedness assumption 2

$$K(A) = K(\hat{f}_\theta) \quad \text{a.e.} \quad \theta \tag{4.1}$$

where $\hat{f}_\theta$ denotes the quadratic form of f_θ. Let the quotient space, denoted by $\tilde{\mathscr{X}}$, be such that

$$\forall\, x \in \mathscr{X}, \quad \|\tilde{x}\| = \inf_{\delta\in K(A)} \|x - \delta\| = d(x, K(A))$$

where $\tilde{x}$ is the equivalence class $x + K(A)$ of elements of x. Now $(\tilde{x}, \|\ \|)$ is a Banach space ([1], page 140). The linear map A_Q defined on it as follows

$$A_Q(\tilde{x}) = Ax \qquad \text{for} \quad x \in \tilde{\mathscr{X}}$$

is continuous in the norm of $\tilde{\mathscr{X}}$. This is shown as follows:

$$\|A_Q\| = \sup_{\|\tilde{x}\|=1} \|A(\tilde{x})\| = \sup_{x:d(x,K(A))=1} \|Ax\| .$$

Also for each $x \in \tilde{\mathscr{X}}$ with $d(x, K(A)) = 1$, $x = x + \delta - \delta$ whatever $\delta \in K(A)$. So

$$\|Ax\| \leqq \inf_{\delta\in K(A)} \{\|A(x + \delta)\| + \|A\delta\|\} \leqq \inf_{\delta\in K(A)} \{\|x + \delta\|\} \leqq 1$$

since $A\delta = 0$ for $\delta \in K(A)$ and $\|A\| = 1$. Hence A_Q is continuous. Furthermore

A_Q is such that

$$(A_Q^* A_Q(\tilde{x}), \tilde{y}) = (Ax, Ay) = (A^*Ax, y) \quad \text{for} \quad x, y \in \mathscr{X}.$$

Thus

$$m(\theta)A_Q^*A_Q \leqq f_\theta \leqq M(\theta)A_Q^*A_Q \quad \text{a.e.} \quad \theta$$

and $A_Q : \tilde{\mathscr{X}} \to \mathscr{K}$ is one-to-one, and without loss of generality $A_Q(\tilde{\mathscr{X}})$ is dense in $\mathscr{K}$.

To make sense of the definition of $\xi_n A_Q^{-1}$ on the image of $\tilde{\mathscr{X}}$ under A_Q, we must have ξ_n uniquely defined on $\tilde{\mathscr{X}}$. It is here that we would need the assumption of linearity of ξ_n. Let, for x and y in $\mathscr{X}$, $Ax = Ay$. Then

$$\begin{aligned} \|\xi_n(x) - \xi_n(y)\|^2 &= \|\xi_n(x-y)\|^2 \\ &= (\xi_n(x-y), \xi_n(x-y)) \\ &= (\xi_0(x-y), \xi_0(x-y)) \\ &= \frac{1}{2\pi} \int_0^{2\pi} \hat{f}_\theta(x-y)\, d\theta \\ &= 0 \quad (\text{due to } (4.1)). \end{aligned}$$

So if $x = y \bmod K(A)$ then $\xi_n(x) = \xi_n(y)$. Hence $\forall\, x \in \mathscr{X}$ we may define $\xi_n(\tilde{x}) = \xi_n(x)$. And the preceding procedures now apply to $\xi_n A^{-1}$ to ultimately yield the prediction error matrix and the predictor for the process $\{\xi_n : -\infty < n < \infty\}$.

REFERENCES

[1] Kato, T. (1960). *Perturbation Theory for Linear Operators.* Springer-Verlag, New York.

[2] Masani, P. (1966). *Recent Trends in Multivariate Prediction Theory*, *Multivariate Analysis.* Academic Press, New York, 351–382.

[3] Masani, P. and Wiener, N. (1959). Nonlinear prediction. In *Probability and Statistics*, *The Harold Cramér Volume* (Ulf Grenander, ed.) 190–212. Stockholm.

[4] Miammee, A. G. (1973). Banach space valued stationary stochastic processes and factorization of nonnegative operator valued functions on a Banach space. Ph. D. Thesis, Michigan State Univ.

[5] Miammee, A. G. (1974). An algorithm for determining the generating function and the linear predictor of an infinite dimensional stationary stochastic process. *Sankhya* **A36** 000–014.

[6] Miammee, A. G. (1976). On $\mathscr{B}(\mathscr{X}, \mathscr{K})$-valued stationary stochastic processes. *Indiana Univ. Math. J.* **25** 921–932.

[7] Wiener, N. and Masani, P. (1958). The prediction theory of multivariate stochastic processes, II. *Acta Math.* **99** 93–137.

Management Sciences Group D-55
The MITRE Corporation
P.O. Box 208
Bedford, Massachusetts 01730