JOURNAL OF MULTIVARIATE ANALYSIS 51, 1–16 (1994)

Principal Component Analysis for a Stationary Random Function Defined on a Locally Compact Abelian Group

A. BOUDOU AND J. DAUXOIS

Université Paul Sabatier, 31062 Toulouse Cedex, France

When Z is a random L^2_H-valued measure, where H is a Hilbert space, we prove that there exists an $L^2_{\mathbb{C}^q}$-valued measure, which may depend on constraints and which best sums up the random measure Z according to a stationary criterion. Then a technique to reduce a random function is deduced from the above result. The random function is defined on a locally compact abelian group and is stationary and continuous. This work generalizes Brillinger's results on stationary time series. © 1994 Academic Press, Inc.

I. PRELIMINARIES

1. *Introduction*

Let H be a separable $\mathbb{C}$-Hilbert space and $\mathscr{X} = (X_t)_{t \in T}$ be a process, where each X_t is H-valued. To clarify the situation, it is very often of great interest to extract out from $\mathscr{X}$ a q-dimensional process that best summarizes $\mathscr{X}$ in a mathematical procedure to be specified later. For an hilbertian random function (r.f.), it is known that principal component analysis (PCA) techniques (see, for instance, [5a,b]) or those of harmonic analysis (cf. [6]) permit, under certain conditions, an answer to the above question (the conditions take into account a measure μ, defined on a measurable space $(T, \mathscr{E})$, that needs to use a σ-field of subsets of T). Nevertheless, these methods require assumptions concerning the second-order moments of X_t; such conditions are generally not satisfied for stationary processes (in this paper we only use weak stationarity). For instance, a basic hypothesis for PCA of an hilbertian r.f. $\mathscr{X}$ is that the mapping $t \mapsto \|X_t\|^2$ is μ-integrable; in the case of a stationary $\mathscr{X}$, this is possible only when $\mu(T)$ is finite.

In this paper we propose a method that allows a Hilbert space-valued continuous stationary r.f. $\mathscr{X}$ defined on a locally compact abelian group G

Received June 19, 1991; revised April 28, 1993.

AMS 1980 subject classifications: 62H25, 62M15.

Key words and phrases: principal component analysis, canonical analysis, random measure, stationary random function, frequency domain.

0047-259X/94 $6.00

to be summarized by means of a stationary q-dimensional r.f. $\mathscr{Y}$. For this purpose we will use a criterion closely related to stationarity. The intrinsic notion to be defined in such a method seems to be the PCA of a random measure associated with $\mathscr{X}$ (see [3]). Noting that, when $G = \mathbb{Z}$ or $G = \mathbb{R}^n$, the method includes Brillinger's work (see [4(a)–(b)]), we also use the term PCA, but it is not necessarily the best one, even if, in practice, the results derive from a family of classical PCA.

We also study PCA under constraints which enable us, for example, to eliminate certain frequencies; it is also possible to generalize what D. Brillinger calls canonical analysis for stationary time series.

Of course, one can apply these techniques when G is $\mathbb{R}^n$, $\mathbb{Z}^n$, or a finite group (thus it is possible to consider periodic stationary time series). When $G = \mathbb{R}^n$, discretization and convergency problems arise; we will not investigate them in this present work.

2. *Notation. The Space* $(H, H') - L^2(M)$

This section is particularly devoted to a version derived from [1] of notions developed in [10]. Given two separable $\mathbb{C}$-Hilbert spaces H and H', we denote by $\sigma_2(H, H')$ (resp. $\sigma_2(H)$) the $\mathbb{C}$-Hilbert space of Hilbert–Schmidt operators mapping H into H' (resp. H) with inner product defined by

$$\langle \cdot, \cdot \rangle_2 \colon (L, K) \in (\sigma_2(H, H'))^2 \mapsto \langle L, K \rangle_2 = \operatorname{tr} LK^*,$$

where K^* is the adjoint operator of K. We denote by $\hom(H, H')$ (resp. $\mathscr{L}(H, H')$) the vector space of the linear transformations (resp. continuous linear transformations) from H into H'.

Let us recall that, given $u \in H$ and $v \in H'$, $u \otimes v$ is the operator from H into H' defined by $u \otimes v(h) = \langle h, u \rangle v$.

"I" will denote the identity map of H (or of another Hilbert space according to the context).

Two functions $s \mapsto L_s$ and $s \mapsto K_s$ mapping an (additive) abelian group G into $\mathscr{L}(H, H')$ are said to be stationarily correlated if for every $(s, t) \in G^2$, one has $L_s K_t^* = L_{s-t} K_0^*$. A function that is stationarily correlated with itself is called a stationary function.

Given a probability space $(\Omega, \mathscr{A}, P)$, we denote by $L^2_H(P)$ (or, for the sake of brevity, L^2_H) the space of the measurable functions defined on $(\Omega, \mathscr{A}, P)$ and with values in H, such as $\|f\|^2$ being a P-integrable function. When $H = \mathbb{C}^p$, we will simply write L^2_p instead of $L^2_{\mathbb{C}^p}$. Throughout this paper, we assume that L^2 (and thus L^2_H) is a separable space. For any $X \in L^2_H$ and any $f \in L^2$, we put

$$\tilde{X}(f) = \mathbb{E}(fX);$$

then $\tilde{X}$ is in $\sigma_2(L^2, H)$.

Let M be a positive measure defined on the measurable space $(E, \mathscr{E})$ and taking values in $\sigma_2(H)$; M'_t denotes the (Radon–Nikodym) derivative of M with respect to $t = \operatorname{tr} M$. Let $\mathscr{G}$ denote the closed subspace

$$\mathscr{G} = \{\varphi \in L^2_{\sigma_2(H,H')}(E, \mathscr{E}, t);\ \forall e \in E,\ \exists \psi(e) \in \hom(H, H')\ \text{with}\ \varphi(e) = \psi(e)(M'_t(e))^{1/2}\}.$$

We denote by $\mathscr{M}$ the linear space of $\psi \in [\hom(H, H')]^E$ such that there exists $\varphi \in \mathscr{G}$ with $\varphi(e) = \psi(e)\ (M'_t(e))^{1/2} \in \sigma_2(H, H')$, for any $e \in E$, and we consider the quotient space of $\mathscr{M}$ by the kernel of the linear functional $\psi \in \mathscr{M} \mapsto \psi(M'_t)^{1/2} \in \mathscr{G}$. This quotient space will be denoted by $(H, H') - L^2(M)$.

When H is a finite-dimensional space, each element θ (a coset) in $(H, H') - L^2(M)$ contains an $\mathscr{E}$-measurable function in $\mathscr{M}$ which is also written θ.

With the inner product,

$$\langle \cdot, \cdot \rangle_{(H,H') - L^2(M)}: (\varphi, \psi) \mapsto \int \operatorname{tr}(\varphi M'_t \psi^*)\, dt,$$

$(H, H') - L^2(M)$ is an Hilbert space.

An orthogonal random measure Z, defined on $(E, \mathscr{E})$ and taking values in L^2_H, is a mapping from $\mathscr{E}$ into $L^2_H(P)$ which is σ-additive and satisfies for any $(A, B) \in \mathscr{E}^2$ that $A \cap B = \varnothing$:

$$\widetilde{Z(A)(Z(B))^*} = 0.$$

We will simply say that Z is a random measure. Given Z, M_Z denotes the (induced) positive measure $A \in \mathscr{E} \mapsto \widetilde{Z(A)(Z(A))^*} \in \sigma_2(H)$. We will write t_Z instead of $\operatorname{tr} M_Z$ and we will denote by M'_Z the derivative of M_Z with respect to t_Z. The closure in the Hilbert space $L^2_{H'}$ of the subspace spanned $\{LZ(A);\ A \in \mathscr{E},\ L \in \sigma_2(H, H')\}$ is denoted by $H^{H'}_Z$; it can be seen that this last space is also the closure of $\operatorname{span}\{LZ(A);\ A \in \mathscr{E},\ L \in \mathscr{L}(H, H')\}$. Then, there exists a unique isometry $\varphi \mapsto \int \varphi\, dZ$ from $(H, H') - L^2(M_Z)$ onto $H^{H'}_Z$ such that one has

$$\forall A \in \mathscr{E}, \forall L \in \mathscr{L}(H, H'), \qquad \int \mathbb{1}_A L\, dZ = LZ(A).$$

The random measures Z and $\mathscr{Z}$ defined on $(E, \mathscr{E})$ and with values respectively in L^2_H and $L^2_{H'}$ are said to be stationarily correlated if the condition $\widetilde{Z(A)(\mathscr{Z}(B))^*} = 0$ is satisfied for each A and B that are disjoint elements of $\mathscr{E}$.

Given $\varphi \in (H, H') - L^2(M_Z)$, $Z_\varphi: A \in \mathscr{E} \mapsto \int \mathbb{1}_A \varphi \, dZ \in L^2_{H'}$ is a random measure which is stationarily correlated with Z.

When t_Z is dominated by a σ-finite measure ν (i.e., there exists for t_Z a derivative with respect to ν), then $t_{Z_\varphi} = \operatorname{tr} M_{Z_\varphi}$ is also dominated by ν; furthermore, one has $dM_{Z_\varphi}/d\nu = \varphi M'_\nu \varphi^*$.

If $\psi \in (H', H'') - L^2(M_Z)$ is a measurable function from E into $\sigma_2(H', H'')$, $\psi(\cdot)\,\varphi(\cdot)$ is an element of $(H, H'') - L^2(M_Z)$ that satisfies

$$\int \psi \, dZ_\varphi = \int \psi\varphi \, dZ.$$

Let $A \in \mathscr{E}$ and $\varphi = \mathbb{1}_A$; then this last equality is true for any $\psi \in (H', H'') - L^2(M_{Z_{\mathbb{1}_A I}})$.

II. Principal Component Analysis of a Random Measure

1. *Some Basic Results*

Let K and $\mathscr{K}$ denote the maps defined by

$$K\colon x \in H \mapsto (x, 0) \in H \times H', \qquad \mathscr{K}\colon y \in H' \mapsto (0, y) \in H \times H'.$$

Given random measures Z and $\mathscr{Z}$, defined on $(E, \mathscr{E})$ with respective values in L^2_H and $L^2_{H'}$, which are stationarily correlated, then the mapping ζ from $\mathscr{E}$ into $L^2_{H \times H'}$ defined by $\zeta(A) = KZ(A) + \mathscr{K}\mathscr{Z}(A)$ is a random measure with values in $L^2_{H \times H'}$ and one has

$$Z = \zeta_{\mathbb{1}_E K^*}, \qquad \mathscr{Z} = \zeta_{\mathbb{1}_E \mathscr{K}^*}.$$

Denoting by Q and R the respective orthogonal projections from $L^2_{H'}$ onto $H^{H'}_Z$ and from $L^2_{H''}$ onto $H^{H''}_Z$, there exists a unique $\varphi \in (H, H') - L^2(M_Z)$ such that

$$Q\left(\int I \, d\mathscr{Z}\right) = \int \varphi \, dZ.$$

We can now prove the following result:

Lemma 1. *Suppose Z and $\mathscr{Z}$ are stationarily correlated. Then, for any $A \in \mathscr{E}$ and any $L \in \sigma_2(H', H'')$, we have*

$$R\left(\int \mathbb{1}_A L \, d\mathscr{Z}\right) = \int \mathbb{1}_A L \, dZ_\varphi.$$

Proof. Let $(B, T) \in \mathscr{E} \times \sigma_2(H, H'')$, we have

$$\left\langle \int \mathbb{1}_A L \, d\mathscr{Z}, TZ(B) \right\rangle = \left\langle \int \mathbb{1}_A L \mathscr{K}^* \, d\zeta, \int \mathbb{1}_B TK^* \, d\zeta \right\rangle$$

$$= \int \operatorname{tr}(\mathbb{1}_{A \cap B} L \mathscr{K}^* M'_\zeta K T^*) \, dt_\zeta$$

$$= \langle \mathscr{K}^*, \mathbb{1}_{A \cap B} L^* T K^* \rangle_{(H \times H', H') - L^2(M_\zeta)};$$

since we have

$$\langle \mathscr{K}^*, \mathbb{1}_{A \cap B} L^* T K^* \rangle_{(H \times H', H') - L^2(M_\zeta)} = \left\langle \int I \, d\mathscr{Z}, \int \mathbb{1}_{A \cap B} L^* T \, dZ \right\rangle_{L^2_{H'}}$$

$$= \left\langle \int \varphi \, dZ, \int \mathbb{1}_{A \cap B} L^* T \, dZ \right\rangle$$

$$= \left\langle \int \mathbb{1}_A L \, dZ_\varphi, TZ(B) \right\rangle,$$

we deduce that

$$\left\langle \int \mathbb{1}_A L \, d\mathscr{Z}, TZ(B) \right\rangle = \left\langle \int \mathbb{1}_A L \, dZ_\varphi, TZ(B) \right\rangle$$

and the result follows. ∎

Using the same notation, from this lemma we derive

COROLLARY 1. *Z and $\mathscr{Z}$ are stationarily correlated if and only if*

$$\forall A \in \mathscr{E}, \qquad Q(\mathscr{Z}(A)) = Z_\varphi(A).$$

Proof. The necessary condition is an obvious application of the lemma with $H'' = H'$ and $L = I$. Conversely, if A and B are disjoint subsets in $\mathscr{E}$ and if $L \in \sigma_2(H', H)$, we have

$$\langle \widetilde{Z(A)}(\widetilde{\mathscr{Z}(B)})^*, L \rangle = \langle L^* Z(A), \mathscr{Z}(B) \rangle = \langle L^* Z(A), Z_\varphi B \rangle$$

$$= \int \operatorname{tr}(\mathbb{1}_{A \cap B} L^* M'_Z \varphi^*) \, dt_Z = 0. \quad ∎$$

COROLLARY 2. *Suppose Z and $\mathscr{Z}$ are stationarily correlated. Then, for each $\gamma \in (H', H'') - L^2(M_{\mathscr{Z}})$, $R(\int \gamma \, d\mathscr{Z})$ belongs to $H^{H''}_{Z_\varphi}$.*

Proof. Clearly, the result is true for $\gamma = \mathbb{1}_A L$; so, by standard continuity arguments, it is also true for any $\gamma \in (H', H'') - L^2(M_{\mathscr{Z}})$. ∎

We can also note:

LEMMA 2. *Suppose that* Z *and* $\mathscr{X}$ *are stationarily correlated and that* $\chi \in (H. H'') - L^2(M_Z)$; *then* Z_χ *and* $\mathscr{X}$ *are stationarily correlated.*

Proof. Let $(A, B) \in \mathscr{E}^2$ with $A \cap B = \varnothing$ and $L \in \sigma_2(H', H'')$. We have

$$\begin{aligned}\langle \widetilde{Z_\chi(A)}(\widetilde{\mathscr{X}(B)})^*, L\rangle &= \operatorname{tr}(\widetilde{Z_\chi(A)}(\widetilde{\mathscr{X}(B)})^* L^*) = \langle L^*\widetilde{Z_\chi(A)}, \widetilde{\mathscr{X}(B)}\rangle\\ &= \langle L^* Z_\chi(A), Q(\mathscr{X}(B))\rangle = \langle L^* Z_\chi(A), Z_\varphi(B)\rangle\\ &= \int \mathbb{1}_{A\cap B} \operatorname{tr}\left(L^*\chi \frac{dM_Z}{dt_Z} \varphi^*\right) dt_Z = 0\end{aligned}$$

and, thus, the required result. ∎

This lemma is easily generalized by

LEMMA 2′. *Given* $\chi \in (H', H'') - L^2(M_Z)$ *and* $\psi \in (H', H'') - L^2(M_{\mathscr{X}})$. *If* Z *and* $\mathscr{X}$ *are stationarily correlated, so are* Z_χ *and* $\mathscr{X}_\psi$.

Proof. Since Z_χ and $\mathscr{X}$ are stationarily correlated, it follows, again using Lemma 2, that so are Z_χ and $\mathscr{X}_\psi$. ∎

2. *Definition and Existence of Principal Component Analysis*

Given a random measure Z with values in L^2_H and $q \in \mathbb{N}^*$, we term every pair $(\mathscr{X}, \beta)$, where $\mathscr{X}$ is a $L^2_{\mathbb{C}^q}$-valued random measure such that Z and $\mathscr{X}$ are stationarily correlated and $\beta \in (\mathbb{C}^q, H) - L(M_{\mathscr{X}})$ such that $\|\int I\,dZ - \int \beta\,d\mathscr{X}\|$ is minimal, *principal component analysis* (PCA) *for* Z *of order* q.

For any random measure $\mathscr{X}$ with values in $L^2_{\mathbb{C}^q}$ which is stationarily correlated with Z and for any $\beta \in (\mathbb{C}^q, H) - L^2(M_{\mathscr{X}})$, we have

$$\left\| \int I\,dZ - \int \beta\,d\mathscr{X} \right\| \geqslant \left\| \int I\,dZ - \Pi\left(\int \beta\,d\mathscr{X}\right)\right\|, \tag{1}$$

where Π denotes the projection operator from L^2_H onto H^H_Z.

Let φ be the unique element of $(H, \mathbb{C}^q) - L^2(M_Z)$ which satisfies $\int \varphi\,dZ = Q(\int I\,d\mathscr{X})$. By using Corollary 2, $\Pi(\int \beta\,d\mathscr{X})$ belongs to $H^H_{Z_\varphi}$. Thus, there exists $\gamma_{\mathscr{X},\beta} \in (\mathbb{C}^q, H) - L^2(M_{Z\varphi})$ such that

$$\Pi\left(\int \beta\,d\mathscr{X}\right) = \int \gamma_{\mathscr{X},\beta}\,dZ_\varphi.$$

Equation (1) can now be rewritten:

$$\left\| \int I\,dZ - \int \beta\,d\mathscr{X} \right\| \geqslant \left\| \int I\,dZ - \int \gamma_{\mathscr{X},\beta}\,dZ_\varphi \right\|.$$

Since Z_φ and Z are stationarily correlated, finding a PCA of order q for Z is reduced to the search for $\varphi \in (H, \mathbb{C}^q) - L^2(M_Z)$ and $\gamma \in (\mathbb{C}^q, H) - L^2(M_{Z_\varphi})$ which minimize $\|\int I\, dZ - \int \gamma\, dZ_\varphi\|$.

Since $\mathbb{C}^q$ is a finite dimensional space, we can choose for γ a measurable function from E into $\sigma_2(\mathbb{C}^q, H)$ and the term which is to be minimized is $\|\int I\, dZ - \int \gamma\varphi\, dZ\|$.

Given a σ-finite measure ν which dominates t_Z, the derivative $dM_Z/d\nu$ of M_Z with respect to ν admits a measurable Schmidt decomposition:

$$\frac{dM_Z}{d\nu}(\cdot) = \sum_{j \in J} \mu_j^2(\cdot)\, a_j(\cdot) \otimes a_j(\cdot).$$

Let us recall that "Schmidt decomposition" means that, for any $e \in E$, $a_j(e) \in H$ is an eigenvector associated with the eigenvalue $\mu_j^2(e)$ $(\mu_j(e) \geqslant 0)$ of $(dM_Z/d\nu)(e)$, $(\mu_j^2)_{j \in J}$ is a decreasing sequence, and $\{a_j(e);\, j \in J\}$ is an orthonormal system in H; the word "measurable" here refers to the fact that a_j and μ_j can be chosen as $\mathscr{E}$-measurable functions (see [7], for instance). Denoting by $\{f_1, ..., f_j, ..., f_q\}$ the standard basis of $\mathbb{C}^q$,

$$\alpha = \sum_{j=1}^{q} a_j \otimes f_j, \qquad \beta = \sum_{j=1}^{q} f_j \otimes a_j$$

respectively belong to $(H, \mathbb{C}^q) - L^2(M_Z)$ and $(\mathbb{C}^q, H) - L^2(M_{Z_\alpha})$.

From classical properties in operator approximation (see [8], for instance), it follows that for any $e \in E$, any $\varphi \in (H, \mathbb{C}^q) - L^2(M_Z)$, and any $\gamma \in (\mathbb{C}^q, H) - L^2(M_{Z_\varphi})$, we have

$$\begin{aligned}
&\left\| \left(\frac{dM_Z}{d\nu}(e)\right)^{1/2} - \beta(e)\,\alpha(e) \left(\frac{dM_Z}{d\nu}(e)\right)^{1/2} \right\|^2 \\
&= \left\| \left(\frac{dM_Z}{d\nu}(e)\right)^{1/2} - \sum_{j=1}^{q} \mu_j(e)\, a_j(e) \otimes a_j(e) \right\|^2 \\
&\leqslant \left\| \left(\frac{dM_Z}{d\nu}(e)\right)^{1/2} - \gamma(e)\,\varphi(e) \left(\frac{dM_Z}{d\nu}(e)\right)^{1/2} \right\|^2
\end{aligned}$$

and, thus, by integration with respect to ν, we obtain

$$\left\| \int I\, dZ - \int \beta\alpha\, dZ \right\| \leqslant \left\| \int I\, dZ - \int \gamma\varphi\, dZ \right\|$$

(let us recall that one has $\|\int \psi\, dZ\|^2 = \int \|\psi((M_Z)'_\nu)^{1/2}\|^2\, d\nu$). Thus, we have

PROPOSITION 1. *Let Z be a random measure with values in L^2_H and v a σ-finite measure dominating t_Z. Given a measurable Schmidt decomposition $\sum_{j\in J}\mu_j^2 a_j\otimes a_j$ of dM_Z/dv, we put*

$$\alpha=\sum_{j=1}^{q} a_j\otimes f_j, \qquad \beta=\sum_{j=1}^{q} f_j\otimes a_j.$$

Then, (Z_α, β) is a PCA of order q for Z.

III. PCA UNDER CONSTRAINTS

Let Z be a random measure with values in L^2_H. We consider here a given element $\delta\in H-L^2(M_Z)$ (we will write $H-L^2(M_Z)$ instead of $(H,H)-L^2(M_Z)$). Given $Y\in L^2_{H'}$, we are now looking for (φ,ψ), with $\varphi\in(H,\mathbb{C}^r)-L^2(M_{Z_\delta})$ and $\psi\in(\mathbb{C}^r,H')-L^2(M_{(Z_\delta)_\varphi})$ which minimize $\|Y-\int\psi d(Z_\delta)_\varphi\|$. Such a (φ,ψ) will be termed a *PCA of order r for Z with respect to Y under the constraint $H^{\mathbb{C}^r}_{Z_\delta}$*.

Let Π denote the projection from $L^2_{H'}$ onto $H^{H'}_{Z_\delta}$. For every $\varphi\in(H,\mathbb{C}^r)-L^2(M_{Z_\delta})$ and $\psi\in(\mathbb{C}^r,H')-L^2(M_{(Z_\delta)_\varphi})$, $\int\psi(Z_\delta)_\varphi$ belongs by hypothesis to $H^{H'}_{Z_\delta}$; thus we have

$$\left\|Y-\int\psi d(Z_\delta)_\varphi\right\|^2=\|Y-\Pi(Y)\|^2+\left\|\Pi(Y)-\int\psi d(Z_\delta)_\varphi\right\|^2.$$

Denoting by γ the unique element in $(H,H')-L^2(M_{Z_\delta})$ which satisfies

$$\Pi(Y)=\int\gamma\, dZ_\delta=\int Id(Z_\delta)_\gamma,$$

then the problem under study is reduced to the minimization of

$$\left\|\int Id(Z_\delta)_\gamma-\int\psi d(Z_\delta)_\varphi\right\|.$$

Let $(((Z_\delta)_\gamma)_\alpha,\beta)$ be a PCA of order r for the random measure $(Z_\delta)_\gamma$. We have

$$\left\|\int Id(Z_\delta)_\gamma-\int\beta d((Z_\delta)_\gamma)_\alpha\right\|\leqslant\left\|\int Id(Z_\delta)_\gamma-\int\psi d(Z_\delta)_\varphi\right\|$$

or

$$\left\|\int Id(Z_\delta)_\gamma-\int\beta d(Z_\delta)_{\alpha\gamma}\right\|\leqslant\left\|\int Id(Z_\delta)_\gamma-\int\psi d(Z_\delta)_\varphi\right\|.$$

Thus, we obtain

PROPOSITION 2. *Let Z be a random measure with values in L^2_H, $\delta \in H - L^2(M_Z)$ and $Y \in L^2_{H'}$. Denoting by γ the unique element in $(H, H') - L^2(M_{Z_\delta})$ such that $\int \gamma\, dZ_\delta$ is the orthogonal projection of Y onto $H^{H'}_{Z_\delta}$ and let $(((Z_\delta)_\gamma)_\alpha, \beta)$ be a PCA of order r for $(Z_\delta)_\gamma$, then $(\alpha\gamma, \beta)$ is a PCA of order r for Z with respect to Y under the constraint $H^{\mathbb{C}^r}_{Z_\delta}$.*

IV. PCA OF A STATIONARY RANDOM FUNCTION

1. *Second-Order Stationary Random Function*

Let G denote a locally compact abelian group and $\hat{G}$ its dual group; these groups are provided with their respective Borel fields. For any character $\gamma \in \hat{G}$ and any $g \in G$, as usual we write (γ, g) instead of $\gamma(g)$.

We are here concerned with an L^2_H-valued random measure Z defined on $(\hat{G}, \mathscr{B}_{\hat{G}})$ such that t_Z is a regular measure; let us recall that t_Z is regular (see, for example, [11]) if, denoting by $|t_Z|$ the total variation of t_Z, we have

$$|t_Z|(E) = \sup |t_Z|(K) = \inf |t_Z|(V),$$

for every Borel set E, where K ranges over all compact subsets of E and V ranges over all open supersets of E. The random function $\mathscr{X} = (X_g)_{g \in G}$ defined by

$$X_g = \int (\cdot, g) I\, dZ, \tag{2}$$

is stationary (that is to say, so is the mapping $g \in G \mapsto \tilde{X}_g \in \sigma_2(L^2, H)$). Moreover, $g \mapsto X_g$ is a continuous mapping from G into L^2_H; we will simply say that $\mathscr{X}$ is continuous. Since $\operatorname{span}\{(\cdot, g);\ g \in G\}$ is dense in $L^2(\hat{G}, \mathscr{B}_{\hat{G}}, t_Z)$, we may note that $H^{H'}_Z$ is the closure of the spanned subspaces: $\operatorname{span}\{KX_g;\ g \in G,\ K \in \sigma_2(H, H')\}$ and $\operatorname{span}\{KX_g;\ g \in G,\ K \in \mathscr{L}(H, H')\}$.

Conversely, given a continuous stationary L^2_H-valued random function $\mathscr{X} = (X_g)_{g \in G}$, there exists a unique random measure $Z(\mathscr{X})$ (we will say *the random measure associated with* $\mathscr{X}$), defined on $(\hat{G}, \mathscr{B}_{\hat{G}})$, and taking values into L^2_H, such that $t_{Z(\mathscr{X})}$ is regular and (2) is true for every $g \in G$.

Let $\mathscr{Y} = (Y_g)_{g \in G}$, the random function associated with a random measure $\mathscr{Z}$ defined on $(\hat{G}, \mathscr{B}_{\hat{G}})$ and with values in $L^2_{H'}$, such that $t_{\mathscr{Z}}$ is regular. In order that $\mathscr{X}$ and $\mathscr{Y}$ are stationarily correlated (i.e., the mappings $g \mapsto \tilde{X}_g$ and $g \mapsto \tilde{Y}_g$ are stationarily correlated), it is necessary and sufficient that we have the same property for Z and $\mathscr{Z}$.

For any $\varphi \in (H, H') - L^2(M_Z)$, the family

$$\mathscr{Y} = \left(\int (\cdot, g) \varphi \, dZ \right)_{g \in G}$$

is the image of the stationary random function $\mathscr{X}$ by the filter (whose transfer function is) φ. The $L^2_{H'}$-valued random function $\mathscr{Y}$ is again stationary and continuous; its associated random measure $Z(\mathscr{Y})$ is Z_φ.

2. *PCA*

Given a random function $\mathscr{X} = (X_g)_{g \in G}$ which is a family of elements belonging to L^2_H defined on a locally compact abelian group G, it can be of interest to find how to best summarize $\mathscr{X}$ by an approximating random function $\mathscr{Y} = (Y_g)_{g \in G}$: $g \in G \mapsto Y_g \in L^2_{\mathbb{C}^q}$. The choice for $\mathscr{Y}$ and its determination first need the introduction of a criterion that allows the quality of the summarizing function to be estimated.

When $\mathscr{X}$ is a continuous stationary random function, it is of interest to seek a random function $\mathscr{Y}$ in order that the random function $\mathscr{W} = (W_g)_{g \in G}$: $g \in G \mapsto W_g \in L^2_H$ obtained by filtering $\mathscr{Y}$ be "as close as possible to $\mathscr{X}$." Then it seems "natural" to require that $\mathscr{Y}$ have the same properties of stationarity and continuity as $\mathscr{X}$ has. If $\mathscr{Y}$ is filtered by $\psi \in (\mathbb{C}^q, H) - L^2(M_{Z(\mathscr{Y})})$, $\mathscr{W}$ is a stationary random function. If we choose $\mathscr{X}$ and $\mathscr{Y}$ to be stationarily correlated, it follows that $\mathscr{W}$ and $\mathscr{X}$ are also stationarily correlated. Thus, we have, for each $g \in G$,

$$\langle X_g, W_g \rangle = \operatorname{tr} \tilde{X}_g (\tilde{W}_g)^* = \operatorname{tr} \tilde{X}_0 (\tilde{W}_0)^* = \langle X_0, W_0 \rangle$$

and, hence, we have

$$\|X_g - W_g\| = \|X_0 - W_0\|.$$

We can now evaluate the quality of the summarizing random function $\mathscr{Y}$ by

$$\inf \left\{ \left\| X_0 - \int \psi \, dZ(\mathscr{Y}) \right\| ; \psi \in (\mathbb{C}^q, H) - L^2(M_{Z(\mathscr{Y})}) \right\}$$

and we are led to the definition

> A PCA of order q for the continuous stationary L^2_H-valued random function $\mathscr{X}$ is a pair $(\mathscr{Y}, \psi)$, where $\mathscr{Y}$ is a continuous stationary $L^2_{\mathbb{C}^q}$-valued random function such that $\mathscr{X}$ and $\mathscr{Y}$ are stationarily correlated and $\psi \in (\mathbb{C}^q, H) - L^2(M_{Z(\mathscr{Y})})$, which minimize $\|X_0 - \int \psi \, dZ(\mathscr{Y})\|$.

When $((Z(\mathscr{X}))_\alpha, \beta)$ is a PCA of order q for $Z(\mathscr{X})$ (of course, $\alpha \in (H, \mathbb{C}^q) - L^2(M_{Z(\mathscr{X})})$ and $\beta \in (\mathbb{C}^q, H) - L^2(M_{Z(\mathscr{X})_\alpha})$), for any continuous stationary $L^2_{\mathbb{C}^q}$-valued random function $\mathscr{Y}$ which is stationarily correlated with $\mathscr{X}$ and for any $\psi \in (\mathbb{C}^q, H) - L^2(M_{Z(\mathscr{Y})})$, we have that

$$\left\| X_0 - \int \beta d(Z(\mathscr{X}))_\alpha \right\| \leqslant \left\| X_0 - \int \psi \, dZ(\mathscr{Y}) \right\|$$

and so we have

PROPOSITION 3. *Let $((Z(\mathscr{X}))_\alpha, \beta)$ be a PCA of order q for the random measure $Z(\mathscr{X})$ associated with a continuous stationary random function $\mathscr{X}$. Then*

$$\left(\left(\int (\cdot, g)\alpha \, dZ(\mathscr{X})\right)_{g \in G}, \beta\right)$$

is a PCA of order q for $\mathscr{X}$.

We may note with the previous notation that we have for each $g \in G$ that

$$\begin{aligned} Y_g &= \int (\cdot, g)\alpha \, dZ(\mathscr{X}) = \sum_{j=1}^{q} \int (\cdot, g) \, a_j \otimes f_j \, dZ(\mathscr{X}) \\ &= \sum_{j=1}^{q} \int (\cdot, g)(1 \otimes f_j)(a_j \otimes 1) \, dZ(\mathscr{X}) \\ &= \sum_{j=1}^{q} (1 \otimes f_j) \left(\int (\cdot, g) \, a_j \otimes 1 \, dZ(\mathscr{X}) \right) \\ &= \sum_{j=1}^{q} \left(\int (\cdot, g) \, a_j \otimes 1 \, dZ(\mathscr{X}) \right) f_j, \end{aligned}$$

which shows that the first q components at step $q+1$ are those of step q.

3. *The Case of Stationary Time Series*

Let us now proceed to the investigation of a particular case. If we choose $G = \mathbb{Z}$ ($\hat{G}$ is the one-dimensional torus which may be identified to $[-\pi, \pi[$) and $H = \mathbb{C}^p$, we can of course use the previous definition and thus a notion of PCA for time series arises. An analogous notion is introduced in [4(b)] for a time series $\mathscr{X} = (X_n)_{n \in \mathbb{Z}}$, requiring that the summarizing function $\mathscr{Y}$ is, *a priori*, obtained by filtering and with the additional restrictive hypothesis of summability for the family $(\|\tilde{X}_n(\tilde{X}_0)^*\|)_{n \in \mathbb{Z}}$. The PCA we expose here does not need the first part of this last assumption, but in fact leads, *a posteriori*, to a filtered summary;

under Brillinger assumptions it can be seen that both forms of PCA coincide and are obtained by means of spectral analysis of the spectral density of $\mathscr{X}$,

$$\lambda \in [-\pi, \pi[\mapsto \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} e^{-i\lambda n} \tilde{X}_n (\tilde{X}_0)^* \in \sigma_2(\mathbb{C}^p),$$

and that $\mathscr{Y}$ is an infinite moving average (cf. [2]).

V. PCA under Constraints for a Stationary Random Function

The two following examples are of some interest in this framework.

1. *Canonical Analysis of Two Stationary Random Functions*

Let $\mathscr{X} = (X_g)_{g \in G}$ and $\mathscr{Y} = (Y_g)_{g \in G}$ be two continuous stationary random functions. We assume that the L^2_H-valued random function $\mathscr{X}$ and the $L^2_{H'}$-valued random function $\mathscr{Y}$ are stationarily correlated. Let us consider a PCA (α, β) of order r for $Z(\mathscr{X})$ with respect to Y_0 under the constraint $H^{\mathbb{C}^r}_{Z(\mathscr{X})}$. Such a PCA allows the minimization $\| Y_0 - \int \psi d(Z(\mathscr{X}))_\varphi \|$ (or $\| Y_g - \int (\cdot, g) \psi d(Z(\mathscr{X}))_\varphi \|$), the infimum being taken over all $\varphi \in (H, \mathbb{C}^r) - L^2(M_{Z(\mathscr{X})})$ and $\psi \in (\mathbb{C}^r, H') - L^2(M_{Z(\mathscr{X})_\varphi})$.

Such a PCA may be called a *canonical analysis* (CA) *of order r for the random function $\mathscr{X}$ with respect to $\mathscr{Y}$*. We must point out that $\mathscr{X}$ and $\mathscr{Y}$ play asymmetrical roles.

In this CA, the image of $\mathscr{X}$ by the filter α is stationary, continuous, and stationarily correlated with $\mathscr{X}$ and with $\mathscr{Y}$. It is a best summarizing function for $\mathscr{Y}$ by means of an $L^2_{\mathbb{C}^r}$-valued random function which takes into account the random function $\mathscr{X}$, since, if $\mathscr{W} = (W_g)_{g \in G}$ denotes the image of $(\int (\cdot, g) I \, dZ(\mathscr{X})_\alpha)_g$ by the filter β, then $\| Y_0 - W_0 \|$ $(= \| Y_g - W_g \|)$ is minimal.

A Particular Case. Let $\mathscr{X} = (X_n)_{n \in \mathbb{Z}}$ and $\mathscr{Y} = (Y_n)_{n \in \mathbb{Z}}$ be time series respectively with values in $L^2_{\mathbb{C}^p}$ and $L^2_{\mathbb{C}^q}$. Using the same notations as in Section II (but we have here $H = \mathbb{C}^p$ and $H' = \mathbb{C}^q$), we assume that $(T_n = KX_n + \mathscr{K} Y_n)_n$ is a stationary time series and that the family $(\tilde{T}_n (\tilde{T}_0)^*)_n$ is absolutely summable in $\sigma_2(\mathbb{C}^{p+q})$. The first condition can be seen as being equivalent to stationarity and cross-stationarity for $\mathscr{X}$ and $\mathscr{Y}$. Thus, it is able to handle the CA of order r (with $r \leqslant \min(p, q)$) for $\mathscr{X}$ with respect to $\mathscr{Y}$. Under the additional (but theorically very restrictive) Brillinger hypothesis that the spectral density of $\mathscr{X}$ has an inverse for each $\lambda \in [-\pi, \pi[$, we find again the results for the CA proposed in [4(b)].

2. *Pass-Band PCA and Periodical PCA*

Let us illustrate the PCA under constraints by another example. Given a continuous stationary L^2_H-valued random measure $\mathscr{X} = (X_g)_{g \in G}$, for each f which is a bounded measurable function from $\hat{G}$ into $\mathbb{C}$, we set

$$A = f^{-1}(\{0\}), \qquad \mathscr{E}_{H'}(f) = \left\{ \varphi \in (H, H') - L^2(M_{Z(\mathscr{X})}); \int f\varphi \, dZ(\mathscr{X}) = 0 \right\}.$$

It is easy to see that $\varphi \in (H, H') - L^2(M_{Z(\mathscr{X})})$ belongs to $\mathscr{E}_{H'}(f)$ if and only if $\int \varphi \, dZ(\mathscr{X})$ belongs to $H^{H'}_{(Z(\mathscr{X}))_{\mathbb{1}_A I}}$. So, seeking $\varphi \in \mathscr{E}_{\mathbb{C}^q}$ and $\psi \in (\mathbb{C}^q, H) - L^2(M_{(Z(\mathscr{X}))_{\mathbb{1}_A I}})$ such that

$$\left\| \int I \, dZ(\mathscr{X}) - \int \psi d(Z(\mathscr{X})_\varphi) \right\|$$

is as small as possible is the same as looking for the PCA of order q of $Z(\mathscr{X})$ with respect to X_0 under the constraint $H^{\mathbb{C}^q}_{(Z(\mathscr{X}))_{\mathbb{1}_A I}}$. So, we are led to the PCA of the random measure $(Z(\mathscr{X}))_{\mathbb{1}_A I}$.

Two choices for f offer, *a priori*, an interesting development:

a. $f = \mathbb{1}_B$, which furnishes a PCA with the constraint on the filter to be a pass-band one, and thus eliminating "frequencies" corresponding to B. In particular, for $G = \mathbb{Z}^2$, a q-dimensional summary in the form $(Y^1_n + Y^2_m)_{n,m}$ can be obtained from the two-parameter stationary time series $(X_{n,m})_{(n,m) \in \mathbb{Z}^2}$. It suffices to choose for B the complement of $(\{0\} \times [-\pi, \pi[) \cup ([-\pi, \pi[\times \{0\})$.

b. $f = 1 - (\cdot, g_0)$, which leads to a periodical summarizing function with period g_0.

VI. Stationary Random Function with Real-Valued Components

1. *Image of a Random Measure and Conjugate Random Measure*

Let $\mathscr{L}$ be a measurable transformation from $(E, \mathscr{E})$ onto itself and let Z be an L^2_H-valued random function; the mapping $\mathscr{L}(Z): A \in \mathscr{E} \mapsto Z(\mathscr{L}^{-1}(A)) \in L^2_H$ is again a random measure which is said to be the image of the random measure Z by $\mathscr{L}$. It can be shown easily that the image given by $\mathscr{L}$ of the real measure t_Z is the real measure $t_{\mathscr{L}(Z)}$: $A \in \mathscr{E} \mapsto \operatorname{tr} M_{\mathscr{L}(Z)}(A) \in \mathbb{R}_+$. We make here the additional assumption that $\mathscr{L}$ is an involution (that is to say, $\mathscr{L}^2 = I$). Then, if $\varphi \in (H, H') - L^2((M_{\mathscr{L}(Z)})$, it can be seen that $\varphi(\mathscr{L}(\cdot)) \in (H, H') - L^2(M_Z)$ and that one has

$$\int \varphi \, d\mathscr{L}(Z) = \int \varphi(\mathscr{L}(\cdot)) \, dZ.$$

Denoting an orthonormal basis of H by $\{h_j;\, j\in J\}$, clearly

$$\Gamma_H\colon x\in H\mapsto\left(\sum_{j\in J}\langle h_j, x\rangle\, h_j\right)\in H,$$

is an antilinear involution which keeps each basis vector invariant and preserves the norm. We can also observe that for every $(x, y)\in H^2$, we have

$$\langle\Gamma_H(x), y\rangle=\langle\Gamma_H(y), x\rangle$$

and, of course,

$$\langle\Gamma_H(x), \Gamma_H(y)\rangle=\langle y, x\rangle.$$

The mapping $\bar{Z}\colon A\in\mathscr{E}\mapsto\Gamma_H Z(A)\in L^2_H$ is also a random measure called the *conjugate random measure associated with* Z. It satisfies for each $A\in\mathscr{E}$:

$$M_{\bar{Z}}(A)=\Gamma_H M_Z(A)\,\Gamma_H,\qquad t_{\bar{Z}}(A)=t_Z(A).$$

Furthermore, given $\varphi\in(H, H')-L^2(M_{\bar{Z}})$, $\varphi_c\colon e\in E\mapsto\Gamma_{H'}\varphi(e)\,\Gamma_H\in$ hom(H, H') is in $(H, H')-L^2(M_{\bar{Z}})$ and one has

$$\int\varphi\, d\bar{Z}=\Gamma_{H'}\left(\int\varphi_c\, dZ\right).$$

2. *PCA for a "Real" Random Function*

Let us now consider a continuous stationary random function $\mathscr{X}=(X_g)_{g\in G}$, defined on a locally compact abelian group and with values in L^2_H. With an obvious misuse, we say that $\mathscr{X}$ is a real random function if for each $g\in G$, one has $\Gamma_H X_g=X_g$. When $H=\mathbb{C}^p$ with its standard basis, it can be seen that $\mathscr{X}$ is real if and only if the components of X_g are real random variables. Since $t_{Z(\mathscr{X})}$ is a regular measure, it is so for the measure image $t_{\mathscr{L}(Z(\mathscr{X}))}$ of $t_{Z(\mathscr{X})}$ by the continuous map $\mathscr{L}$ from $\hat{G}$ into itself defined by $\mathscr{L}(\gamma)=\gamma^{-1}$.

From the results of VI.1, we obtain for each g

$$\int(\cdot, g)\, Id\overline{Z(\mathscr{X})}=\Gamma_H X_{-g},\qquad\int(\cdot,\cdot)I\, d\mathscr{L}(Z(\mathscr{X}))=X_{-g}.$$

It follows that $\mathscr{X}$ is real if and only if $\overline{Z(\mathscr{X})}$ and $\mathscr{L}(Z(\mathscr{X}))$ coincide on $\mathscr{B}_{\hat{G}}$. Then, if $\varphi\in(H, H')-L^2(M_{Z(\mathscr{X})})$ is such that, for any $\gamma\in\hat{G}$, one has $\Gamma_{H'}\varphi(\mathscr{L}(\gamma))\,\Gamma_H=\varphi(\gamma)$, the filtered $\mathscr{Y}$ of $\mathscr{X}$ given by φ is a real random function. Setting $t=t_Z$ $(=t_{\bar{Z}}=t_{\mathscr{L}(Z)})$, we can deduce that

$$\frac{dM_Z}{dt}(\gamma)=\Gamma_{H'}\frac{dM_Z}{dt}(\gamma^{-1})\,\Gamma_H.$$

Finally, we obtain the existence of a measurable Schmidt decomposition which, under the notations of II.2, satisfies $\Gamma_H a_j \mathscr{L} = a_j$. Therefore, the filter α obtained by the PCA of order q of $\mathscr{X}$ is such that for any $\gamma \in \hat{G}$: $\Gamma_{\mathbb{C}^q} \alpha(\mathscr{L}(\gamma))\, \Gamma_H = \alpha(\gamma)$ and we may claim that a filtered random function by α of $\mathscr{X}$ is again a real random function. Thus we have

PROPOSITION 4. *Let $\mathscr{X} = (X_g)_{g \in G}$ be a random function with associated random measure Z. We put $\mathscr{L}(\gamma) = \gamma^{-1}$, for any $\gamma \in \hat{G}$. The random function $\mathscr{X}$ is "real" if and only if $\overline{Z(\mathscr{X})} = \mathscr{L}(Z(\mathscr{X}))$. Under this condition, it is possible to find a "real" filtered random function $\mathscr{Y}$ in the PCA of order q of $\mathscr{X}$.*

3. *CA of Two Real Random Functions*

We have seen that the CA of order r for two continuous stationary random functions $\mathscr{X}$ and $\mathscr{Y}$ which are stationarily correlated is none other than a particular PCA. Thus, from VI.1 it follows that there exists a "real" CA of order r for the random function $\mathscr{X}$ with respect to $\mathscr{Y}$ when both the random functions are real.

ACKNOWLEDGMENTS

Special thanks are due to Professor J. Gower. The authors thank the referee for his valuable comments and suggestions.

REFERENCES

[1] AZENCOTT, R., AND DACUNHA-CASTELLE, D. (1984). *Séries d'Observations Irrégulières.* Masson, Paris.

[2] BOUDOU, A., AND DAUXOIS, J. (1988). Analyses de Séries Stationnaires Multidimensionnelles. In *Publications du Laboratoire de Statistique et Probabilités de l'Université Paul Sabatier*, 02-88, Toulouse.

[3] BOUDOU, A., AND DAUXOIS, J. (1989). Analyses de mesures aléatoires; Applications aux séries stationnaires. *C. R. Acad. Sci. Paris, Série I* **309** 391–394.

[4] BRILLINGER, D. R. (a) (1970). The frequency analysis of relations between stationary spatial series. In *12th Biennial Seminar of the Canadian Mathematical Congress*, Vol. X, pp. 39–81; (b) (1975). *Time Series Data Analysis and Theory*. Holt, Rinehart, & Winston, New York.

[5] DAUXOIS, J., AND POUSSE, A. (a) (1976). *Les Analyses Factorielles en Calcul des Probabilités et en Statistique: Essai d'étude synthétique*. Thèse de doctorat d'Etat. Université Paul Sabatier, Toulouse; (b) (1977). Some convergence problems in factor analyses. In *Recent Developments in Statistics* (Barra *et al.*, Eds.), North-Holland, Amsterdam.

[6] DEVILLE, J. C. (1981). Décomposition des Processus Aléatoires. Une comparaison entre les méthodes factorielles et l'analyse spectrale. *Statist. Anal. Données* **3** 5–33.

[7] FIEGER, W. (1963). Die Anwendung einiger Mass-Integration Theorischer Sätze auf Matrizielle Riemann-Stieljes-Integrale. *Math. Ann.* **150** 387–410.

[8] GOHBERG, I. C., AND KREIN, M. G. (1971). *Opérateurs Linéaires Non Autoadjoints dans un Espace Hilbertien.* Dunod, Paris.

[9] PAYEN, R. (1967). Fonctions aléatoires du second ordre à valeurs dans un espace de Hilbert. *Ann. Inst. H. Poincaré, Sect. B* **11**, No. 4 323–396.

[10] ROSENBERG, M. (1974). Operators as spectral integrals of operators-valued functions from the study of multivariate stationary stochastic processes. *J. Multivariate Anal.* **4** 166–209.

[11] RUDIN, W. (1967). *Fourier Analysis on Groups.* Wiley-Interscience, New York.