MATH. SCAND. 9 (1961), 83–89

COMPLEMENTS TO THE KARHUNEN REPRESENTATION

I. FLEISCHER and A. KOOHARIAN

1.

Classically, a stochastic process is a probability space Ω whose points are complex-valued functions $x(t)$ of a real parameter interpreted physically as time. It is assumed that the functions on Ω obtained by fixing t are measurable; for the present purposes it is further required that they be square integrable. This permits them to be normalized to have mean zero, and it makes it meaningful to speak of the covariance function $r(s,t) = E\,x(s)\overline{x(t)}$.

Now if r is a continuous function of the difference of its arguments: $r(s,t) = \varrho(s-t)$, then ϱ is positive definite and therefore, by Bochner's theorem, is the Fourier transform of a finite Borel measure on the line:

$$\varrho(\tau) = \int_{-\infty}^{\infty} e^{it\lambda}\,\mu(d\lambda)\,.$$

In these circumstances it has been known for a long time that the process may be represented as

$$x(t) = \int_{-\infty}^{\infty} e^{it\lambda}\,Z(d\lambda)\,,$$

where Z is an "orthogonal increments" process, i.e., essentially a family of signed measures on the line parametrized by Ω and satisfying $E\,|Z(d\lambda)|^2 = \mu(d\lambda)$.

On the other hand, let Z_i be an orthonormal basis in the subspace of $L_2(\Omega)$ spanned by the $x(t)$. The fact that $x(t) = \sum f_i(t)\,Z_i$ is valid for each t in the mean, permits us to deduce the relation

$$r(s,t) = \sum f_i(s)\overline{f_i(t)}\,.$$

Conversely, the latter equality may be shown to imply the former.

Karhunen [4] has shown how to treat both these phenomena in a unified manner. Let (Λ,μ) be a measure space, and for each t let $f(t,\lambda)$ be square integrable on this space. If

Received June 30, 1960.

$$r(s,t) = \int_\Lambda f(s,\lambda)\overline{f(t,\lambda)}\,\mu(d\lambda)\,,$$

then there exists a process Z having properties analogous to those above, in terms of which

$$x(t) = \int_\Lambda f(t,\lambda)\,Z(d\lambda)\,.$$

The equation holds for each t almost everywhere in Ω; or, since square integrable elements over Ω stand on both sides, taking an inner product on both sides yields an identity in t.

Our point of departure is the desire to treat processes which need not be directly defined on the time axis: whose sample points, for example include limits of functions of a real variable or may consist of operators on such functions. Thus we are led to the study of processes parameterized not by a real variable but by a function space over this variable, or, quite abstractly, by a topological vector space; the sample points then belong effectively to the dual of this space. Although Karhunen's theorem makes no use of the real line structure, and is therefore immediately adaptable to the situation at hand, we find it necessary, in view of the applications we have in mind, to strengthen his conclusion by exhibiting uniformities in the convergence with respect to the stochastic variable, and to show how to attain the corresponding uniformities in the indexing variable. The assumed structure on the indexing parameter may then be exploited to bring the result into a compact and useful form: in the presence of the uniformities previously established, the integral representation is shown to be valid in a suitably topologized tensor product space.

Two applications are presented. In the first, the process is defined on a Hilbert space $\mathfrak{H}$. Although a representation may be established in terms of any realization of $\mathfrak{H}$ as a space of square integrable functions, we are particularly interested in the series representation obtained when the covariance function is a completely continuous bilinear form on $\mathfrak{H}$. This series converges in the mean square sense on the product of the probability space and the measure space over which $\mathfrak{H}$ is realized. In this way we rederive the result which constituted the basis of our former investigation, but in a more consistent and lucid manner, and without relying on an extraneous continuity assumption.

In the second application, we index the process with the space of infinitely differentiable functions having compact support; consequently, the sample "functions" are distributions. Thus we are in a position to treat processes whose covariances are "delta functions", which have long been familiar to electrical engineers under the picturesque appellation "white noise." Translated to this setting, the representation yields a generalization of a theorem of Itô [3].

2.

We shall be dealing with a family X_φ, $\varphi \in \Phi$, of random variables of finite variance defined on a probability space Ω with measure p. For convenience we take all means equal to zero.

Now we suppose there is given a family f_φ of square integrable functions on a measure space Λ with measure μ, which is in one-to-one correspondence with the family X_φ:

$$f_\varphi \leftrightarrow X_\varphi \qquad \text{for all} \quad \varphi \in \Phi , \tag{1}$$

in such a way that for $\varphi, \psi \in \Phi$,

$$R(\varphi,\psi) = EX_\varphi \overline{X}_\psi = \langle f_\varphi, f_\psi \rangle_\mu = \int_\Lambda f_\varphi \overline{f_\psi}\, d\mu . \tag{2}$$

In other words, the correspondence (1) is a Hilbert space isometry between the X_φ in $\mathfrak{H}_p$ and the f_φ in $\mathfrak{H}_\mu$. Such an isometry is known to be uniquely extendable to the closed linear subspaces $\mathfrak{M}_X$ and $\mathfrak{M}_f$ spanned by these families in their respective Hilbert spaces. Thus for every $Y \in \mathfrak{M}_X$ which corresponds to $g \in \mathfrak{M}_f$, we have

$$EX_\varphi \overline{Y} = \langle f_\varphi, g \rangle_\mu = \int_\Lambda f_\varphi \bar{g}\, d\mu . \tag{3}$$

In the special case $\mathfrak{M}_f = \mathfrak{H}_\mu$ we may in particular take for g the characteristic function χ_S of a measurable set S of finite measure in Λ. If we denote the corresponding Y by $Z(S)$, this family is readily seen to satisfy $EZ(S)\overline{Z(S')} = \mu(S \cap S')$ and to span $\mathfrak{M}_X$. Regarding $\bar{g}\mu(d\lambda)$ as a signed measure, we may write (3)

$$EX_\varphi \overline{Y} = \int_\Lambda f_\varphi \,[\bar{g}\,\mu(d\lambda)] = \int_\Lambda f_\varphi EZ(d\lambda)\overline{Y}$$

or, symbolically,

$$X_\varphi = \int_\Lambda f_\varphi Z(d\lambda) . \tag{4}$$

This is the content of Karhunen's theorem.

The right side of (4) is effectively an integral over Λ with values in $\mathfrak{M}_X$ which is asserted to exist in the weak topology. It may actually be shown to exist in the strong (mean square) topology; or, in the general formulation, the right side of (3), considered as the integral of f_φ with respect to the signed measure $\bar{g}d\mu$, may be approximated by a linear combination of characteristic functions uniformly for g bounded in $\mathfrak{M}_f$. Indeed,

$$\left|\int_A f_\varphi \bar{g}\, d\mu - \int_A \sum_i \langle f_\varphi, \chi_{S_i}\rangle \chi_{S_i} \bar{g}\, d\mu\right|^2 \leqq \int_A |g|^2\, d\mu \int_A \left| f_\varphi - \sum_i \langle f_\varphi, \chi_{S_i}\rangle \chi_{S_i}\right|^2 d\mu$$

whence the result, since for fixed φ the Hilbert space distance of f_φ to the linear space spanned by finitely many characteristic functions of μ-finite measurable sets converges to zero on the net of these spaces ordered by inclusion.

The same inequality shows further that if the mean square (μ) distance converges to zero on this net uniformly for $\varphi \in \Psi$, then there will be a simultaneous uniformity (for $\varphi \in \Psi$ and g bounded) in the convergence to zero of the mean $\bar{g}d\mu$ distance of f_φ to these subspaces.

3.

Henceforth, Φ will be a topological vector space and X_φ will be a linear, Ω-weakly continuous mapping of Φ into $\mathfrak{M}_X$. It follows that $R(\varphi, \psi)$ is a separately continuous bilinear form on Φ. The mapping adjoint to X then associates to every square integrable function $\mu \in \mathfrak{M}_X$ an element $\langle X_\varphi, u\rangle$ in the dual Φ' of Φ. In this generalized sense, one may say that the sample points of the process belong to Φ'.

When $\mathfrak{M}_f = \mathfrak{H}_\mu$ and the convergence described at the end of the last section is valid uniformly for bounded subsets of Φ, (4) may be interpreted as an integral in the dual of $\Phi \otimes \mathfrak{M}_X$, topologized by uniform convergence on tensor products of bounded sets. In other words

$$X = \int f \otimes Z(d\lambda) \tag{5}$$

is the limit in the indicated sense of expressions

$$\sum \chi_{S_i} \int_{S_i} f_\varphi\, d\lambda \otimes Z(S_i)\ .$$

We shall refer to (5) as the strong representation for X in the tensor product topology.

An immediate consequence of (5) is that X is bounded on tensor products of bounded sets. More generally if the convergence is uniform on any set S on which the approximating sums are bounded, then X is again bounded on S. In certain contexts the existence of a topology on $\Phi \otimes \mathfrak{M}_X$ in which S is bounded leads to the continuity of X. Conversely if X together with its approximating sums are continuous in the topologized $\Phi \otimes \mathfrak{M}_X$ then the representation converges uniformly on precompact sets.

4.

In this section Φ is Hilbert space and, to begin with, R is thus only continuous on $\Phi \times \Phi$ in the product topology. Since $R(\varphi,\varphi) \geqq 0$, the operator A associated to R via $R(\varphi,\psi) = \langle A(\varphi),\psi\rangle$ is positive symmetric; in terms of its square root we obtain

$$R(\varphi,\psi) = \langle A^{\frac{1}{2}}(\varphi), A^{\frac{1}{2}}(\psi)\rangle . \tag{6}$$

Now if Φ is realized as some space of square integrable functions over a measure space Λ, we deduce

$$X_\varphi = \int_\Lambda A^{\frac{1}{2}}(\varphi)\, dZ . \tag{7}$$

In particular, if the range of $A^{\frac{1}{2}}$ has the basis $\{\chi_i \mid i \in \Lambda\}$, (7) can be given as the discrete form

$$X_\varphi = \sum_{i\varepsilon\Lambda} \langle A^{\frac{1}{2}}\varphi, \chi_i\rangle\, Z(i) , \tag{8}$$

where the series converges in the mean on Ω for each $\varphi \in \Phi$. In neither (7) nor (8), however, can we expect to conclude, in general, any uniformity of the representation with respect to its φ-dependence.

To obtain such uniformity we are led to assume that R derives from a completely continuous operator on Φ (equivalently, R may be a continuous linear functional on the tensor product $\Phi \otimes \Phi$ in the topology induced by the natural inner product in this space.) Since its definition implies the positive definiteness of R, the theory of such operators yields

$$R(\varphi,\psi) = \sum_{i=1}^{\infty} \langle\varphi,\chi_i\rangle\langle\psi,\chi_i\rangle/\lambda_i , \tag{9}$$

where χ_i is an orthonormal set in Φ and λ_i are positive numbers for which $\sum_i 1/\lambda_i^2$ is finite. Thus we are led to take for Λ the positive integers, and to set $\mu(i) = 1/\lambda_i$, $f_\varphi(i) = \langle\varphi,\chi_i\rangle$. In $f_{\chi_j}(i) = \delta_{ji}$ we have a set of generators for $\mathfrak{H}_\mu$; therefore, there exists an orthogonal family $Z(i)$, of norm $1/\lambda_i^{\frac{1}{2}}$, which spans $\mathfrak{H}_\mu$ and in terms of which

$$X_\varphi = \sum_{i=1}^{\infty} \langle\varphi,\chi_i\rangle\, Z(i) . \tag{10}$$

Furthermore, since

$$\sum_{i=N}^{\infty} \langle\varphi,\chi_i\rangle^2/\lambda_i \leqq \max_{i\geqq N} \|\varphi\|^2/\lambda_i ,$$

(10) converges uniformly for bounded sets in $\mathfrak{H}_\mu$ as well as for bounded sets in $\mathfrak{M}_X$; in other words, introducing the normalized vectors $Z_i = \lambda_i^{\frac{1}{2}} Z(i)$, we obtain

$$X = \sum_{i=1}^{\infty} \chi_i \otimes Z_i / \lambda_i^{\frac{1}{2}} \tag{11}$$

in the strong sense.

It follows from the remarks at the end of Sec. 3 that X is continuous in the topology induced by the natural inner product on $\Phi \otimes \mathfrak{M}_X$ if and only if $\sum_{i=1}^{\infty} 1/\lambda_i < \infty$.

5.

Now Φ is taken as the space $\mathscr{D}$ of infinitely differentiable functions vanishing outside compact sets (in a finite dimensional Euclidean space) topologized so as to render the adjoint of X an Ω-weakly continuous distribution-valued mapping. It follows that R is a distribution in two variables. Stationarity may be formulated as the requirement that R be expressible in terms of a distribution r in one variable and the operator τ of argument translation as

$$R(\varphi, \psi) = r(\langle \tau\varphi, \psi \rangle) . \tag{12}$$

From its definition $R(\varphi, \varphi) \geqq 0$ whence r is of positive type. By the distribution-theoretic version of Bochner's theorem, there exists a Borel measure $\mu \geqq 0$ such that

$$r(\varphi) = \int \mathfrak{F}(\varphi)\, d\mu . \tag{13}$$

Here $\mathfrak{F}$ is the Fourier transform operator which takes values in the space $\mathscr{D}^*$ of infinitely differentiable functions vanishing at infinity more rapidly than any polynomial.

It is immaterial for our purposes whether $\mathscr{D}$ or $\mathscr{D}^*$ is used as the indexing parameter. Indeed, since both $\mathfrak{F}$ and integration with respect to μ are continuous on $\mathscr{D}^*$, r is continuous in the topology induced on $\mathscr{D}$ by $\mathscr{D}^*$, and therefore has a unique extension to $\mathscr{D}^*$ satisfying (13). Similarly, in virtue of (2), X_φ is Ω-weakly continuous in the same topology and may also be extended to $\mathscr{D}^*$. Conversely, the possibility of restricting r and X from $\mathscr{D}^*$ to $\mathscr{D}$ is immediate.

It remains to verify that $\mathfrak{F}(\mathscr{D}^*)$ spans $\mathfrak{H}_\mu$ and that the representation for R arising from (13),

$$R(\varphi, \psi) = \int \mathfrak{F}(\varphi)\, \mathfrak{F}(\psi)\, d\mu ,$$

has the required uniformity properties. The former may be deduced from the facts that $\mathfrak{F}$ is onto $\mathscr{D}^*$, that elements of $\mathscr{D}^*$ approximate characteristic functions of compact sets uniformly, and that μ is a Borel measure. The latter follows because the $\mathfrak{H}_\mu$ distance of $\mathfrak{F}(\varphi)$ to the subspace spanned by finitely many characteristic functions of compact sets is continuous on $\mathscr{D}^*$ and, therefore, converges to zero uniformly for compact (=bounded) subsets of $\mathscr{D}^*$. Thus, the representation has the form

$$X_\varphi = \int_\Lambda \mathfrak{F}(\varphi) \otimes Z(d\lambda) \tag{14}$$

in the strong sense. Inasmuch as the injection of $\mathscr{D}$ in $\mathscr{D}^*$ is continuous and onto a dense subset, the same reasoning may be applied to obtain the strong representaion for $\mathscr{D}$.

This result was in large part originally obtained by Itô [3], without, however, remarking the relevance of the Karhunen theorem. His development requires Ω strong continuity, is valid only in $\mathscr{D}$, and does not establish the uniformities with respect to the φ variation.

REFERENCES

1. N. Bourbaki, *Éléments de mathématique*, Livre II, Algèbre, Chap. III, Algèbre multilinéaire, Paris, 1948.
2. I. Fleischer and A. Kooharian, *On the statistical treatment of stochastic processes*, Ann. Math. Stat. 29 (1958), 544–549.
3. K. Itô, *Stationary random distributions*, Mem. Coll. Sci. Univ. Kyoto Ser. A 28 (1954), 209–223.
4. K. Karhunen, *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Ann. Acad. Sci. Fennicae Ser. A I. Math.-Phys. no. 37(1947), 1–79.
5. J. L. Kelly, *General topology*, New York, 1955.
6. F. Riesz et B. Sz.-Nagy, *Leçons d'analyse fonctionnelle*, Budapest, 1952.
7. L. Schwartz, *Théorie des distributions* 1–2, Paris, 1950.

APPLIED RESEARCH LABORATORY, SYLVANIA ELECTRONIC SYSTEMS,
WALTHAM 54, MASS., U. S. A.