On the Role of Biorthonormality in Representation of Random Processes

Victor A. Segalescu, Meir Zibulski, and Yehoshua Y. Zeevi

Abstract—The representation of a random process by a set of uncorrelated random variables is examined. The main result indicates that a basis decorrelates a random process if and only if it satisfies an integral equation similar to the type satisfied by the Karhunen–Loève expansion, but by relaxing the requirement of orthogonality of the representation functions.

Index Terms— Karhunen–Loève expansion, biorthonormal bases, decorrelation of random processes.

I. INTRODUCTION

The representation of a random process $\{x(t), t \in T\}$ by a set of random variables is discussed in [1]–[5]. Given a real zero-mean, mean-square continuous random process $x(t)$, with autocorrelation $R(s,t)$ such that

$$\int_T R(t,t)\, dt < \infty$$

and given a basis of $L_2(T)$, the process can be represented in general in the following way:

$$x(t) = \text{l.i.m.} \sum_i a_i \phi_i(t)$$
$$a_i = \int_T x(t)\psi_i(t)\, dt \qquad (1)$$

where $\{\psi_i(t)\}$ is the biorthonormal basis of $\{\phi_i(t)\}$.

The question which naturally arises is under what conditions, if any, the basis $\{\phi_i(t)\}$ decorrelates the process $x(t)$, i.e., the set of variables $\{a_i\}$ is uncorrelated. It is well known that if $\{\phi_i(t)\}$ are the eigenfunctions of $R(s,t)$ then the expansion (1) possesses a double orthogonality: orthogonality of the expansion functions $\{\phi_i(t)\}$ and orthogonality of the random variables $\{a_i\}$ [6]. The eigenfunctions are the solution of the following equation:

$$\int_T R(s,t)\varphi(t)\, dt = \lambda\varphi(s) \qquad (2)$$

and $\{\lambda_i\}$ are the corresponding eigenvalues, $\lambda_i = E[a_i^2]$. The set of orthonormalized eigenfunctions is unique, and sometimes the functions so obtained are quite complicated. In some applications (see, e.g., [2]–[4]) it is desirable to relax the requirement of orthogonality of the functions and retain only the orthogonality of the random variables. This more general case of decorrelation of a random process is the subject of this correspondence.

II. THE CONTINUOUS CASE

The following theorem presents the necessary and sufficient condition for a basis to decorrelate a random process.

Theorem 1: Given a random process $\{x(t), t \in T\}$ which satisfies the conditions listed above, a basis $\{\phi_i(t)\}$ of $L_2(T)$ decorrelates $x(t)$ if and only if it satisfies the following equation:

$$\int_T R(s,t)\psi_i(t)\, dt = \sigma_i \phi_i(s), \quad \forall i \qquad (3)$$

where $\{\psi_i(t)\}$ is the biorthonormal basis of $\{\phi_i(t)\}$, i.e.

$$\int_T \phi_i(t)\psi_j(t)\, dt = \delta_{ij}, \quad \forall i,j.$$

Proof: Since $\{\phi_i(t)\}$ is a basis of $L_2(T)$, the following equations hold:

$$x(t) = \text{l.i.m.} \sum_i a_i \phi_i(t)$$
$$a_i = \int_T x(t)\psi_i(t)\, dt. \qquad (4)$$

i) Assume (3) holds, then

$$E[a_i a_j] = E\left[\int_T x(s)\psi_i(s)\, ds \int_T x(t)\psi_j(t)\, dt\right]$$
$$= \int_T \int_T R(s,t)\psi_i(t)\psi_j(s)\, ds\, dt$$
$$= \int_T \sigma_i \phi_i(t)\psi_j(t)\, dt = \sigma_i \delta_{ij}. \qquad (5)$$

ii) Suppose $E[a_i a_j] = \sigma_i \delta_{ij}$ for all i,j. Then

$$E[a_i a_j] = E\left[\int_T x(s)\psi_i(s)\, ds \int_T x(t)\psi_j(t)\, dt\right]$$
$$= \int_T \left(\int_T R(s,t)\psi_i(t)\, dt\right)\psi_j(s)\, ds$$
$$= \sigma_i \delta_{ij}, \quad \forall i,j. \qquad (6)$$

It follows that

$$\int_T R(s,t)\psi_i(t)\, dt = \sigma_i \phi_i(s)$$

where $\{\phi_i\}$ is the biorthonormal basis of $\{\psi_i\}$ [7]. $\square$

Equation (3) degenerates to (2) in the case of the Karhunen–Loève (KL) expansion. In this case, the basis is orthonormal. In other words, the orthonormal basis utilized in the KL expansion is a special case of a general type of decorrelating bases, all of which satisfy (3). We therefore refer to the general case as the pseudo Karhunen–Loève (PKL) expansion.

The PKL expansion behaves a great deal like the KL expansion. For one thing, we can order the representation functions according to their contribution. Define the approximation error function $\epsilon_N(t)$

$$\epsilon_N(t) = E\left[\left(x(t) - \sum_{i=1}^N a_i \phi_i(t)\right)^2\right]. \qquad (7)$$

$\epsilon_N(t)$ is a monotonic decreasing function (of N) for all t, since by a straightforward calculation, one can easily show that

$$\epsilon_N(t) = R(t,t) - \sum_{i=1}^N \sigma_i \phi_i^2(t). \qquad (8)$$

Define the global approximation error function $\varepsilon_N(T)$

$$\varepsilon_N(T) = \int_T \epsilon_N(t)\, dt. \qquad (9)$$

Based on (8), it follows that we can order the basis functions, $\{\phi_i(t)\}$, according to the value of $\sigma_i\|\phi_i(t)\|^2$ (where $\|\cdot\|$ is the usual norm in $L_2(T)$), i.e., according to their contribution to the mapping of the process $x(t)$.

Naturally, if the process $x(t)$ is Gaussian, $\{a_i\}$ is a set of zero-mean independent random variables, where $E[a_i^2] = \sigma_i$.

This property of bases which decorrelate a random process, suggests the following iterative algorithm for generating a decorrelating basis.

Algorithm 1:

Step 1: Choose a function $\psi_1(t)$. Define $\phi_1(t)$ in the following manner:

$$\int_T R(s,t)\psi_1(t)\,dt = \sigma_1\phi_1(s)$$

$$\int_T \phi_1(t)\psi_1(t)\,dt = 1.$$

Step 2: Find $\psi_2(t)$ orthogonal to $\phi_1(t)$. Define $\phi_2(t)$ in the following manner:

$$\int_T R(s,t)\psi_2(t)\,dt = \sigma_2\phi_2(s)$$

$$\int_T \psi_2(t)\phi_2(t)\,dt = 1.$$

The function $\phi_2(t)$ defined in this manner is orthogonal to $\psi_1(t)$.

Step $k+1$: Find a function $\psi_{k+1}(t)$ orthogonal to $\{\phi_1(t), \phi_2(t), \cdots, \phi_k(t)\}$. Define $\phi_{k+1}(t)$ in the following manner:

$$\int_T R(s,t)\psi_{k+1}(t)\,dt = \sigma_{k+1}\phi_{k+1}(s)$$

$$\int_T \psi_{k+1}(t)\phi_{k+1}(t)\,dt = 1.$$

End: when a sufficient number of pseudo eigenfunctions are obtained.

(For example, sufficient under the criterion that the summation over the k pseudo eigenfunctions, determined during the k iterations of the algorithm, results in a representation with a mean-squared stochastic error smaller than ϵ).

Note that there exist several methods for finding a function $\psi_{k+1}(t)$ orthogonal to $\{\phi_1(t), \phi_2(t), \cdots, \phi_k(t)\}$. For example, if a function $f(t)$ is known such that $f \notin \mathrm{span}\,\{\phi_i\}_{i=1}^k$ then, we can choose

$$\psi_{k+1} = f - \sum_{i=1}^{k}\langle f, \phi_i\rangle\psi_i$$

where $\langle\cdot,\cdot\rangle$ is the inner product in $L_2(T)$. Also note that throughout the algorithm we assume that the integral operator $\boldsymbol{R}$, with the kernel $R(s,t)$, is positive-definite. If not, the functions ψ_i should not belong to the null-space of $\boldsymbol{R}$.

The following result presented in [2] is of relevance and importance to the main result presented in this note in that it considers a special application where the simultaneous decorrelation of two random processes is desirable (and therefore cannot employ the KL expansion).

Let $K_1(s,t)$ and $K_2(s,t)$, $-\tau \leq s, t \leq \tau$, be real continuous symmetric and square-integrable functions, and let $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ be the integral operators with the kernels $K_1(s,t)$ and $K_2(s,t)$. Assume that $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are positive-definite and nonnegative-definite, respectively. Assume also, that the operator $\boldsymbol{K}_1^{-1/2}\boldsymbol{K}_2\boldsymbol{K}_1^{-1/2}$ is densely defined, bounded and its extension to the entire $L_2(T)$, denoted by $\boldsymbol{S}$, has eigenfunctions which span $L_2([-\tau,\tau])$. Let $\{\varphi_i(t)\}$ and $\{\lambda_i\}$ be the set of eigenfunctions and the corresponding eigenvalues of $\boldsymbol{S}$. Then, for $x(t)$ a sample function of either one of two (separable and

measurable) zero-mean processes with autocorrelations $K_1(s,t)$ and $K_2(s,t)$, and associated probability measures P_1, P_2 we have

$$x(t) = \sum_i a_i\big(\boldsymbol{K}_1^{1/2}\varphi_i\big)(t)$$

in the stochastic mean, uniformly in t, where

$$a_i = \int_{-\tau}^{\tau} x(t)\big(\boldsymbol{K}_1^{-1/2}\varphi_i\big)(t)\,dt.$$

Moreover, the random variables $\{a_i\}$ are uncorrelated under both measures.

Let $\psi_i = \boldsymbol{K}_1^{-1/2}\varphi_i$, $\phi_i = \boldsymbol{K}^{1/2}\varphi_i$. One can easily check that, indeed, $\{\psi_i\}$, $\{\phi_i\}$ are biorthonormal (since $\boldsymbol{K}^{1/2}$ is self-adjoint and $\{\varphi_i\}$ is orthonormal) and (3) holds for either $R(s,t) = K_1(s,t)$ or $R(s,t) = K_2(s,t)$. Clearly, the above result guarantees the existence of an infinite number of bases which decorrelate the random process $x(t)$, since we can find an infinite number of operators satisfying the condition imposed on $K_1(s,t)$. Moreover, it provides an alternative way of generating the decorrelating bases.

A. An Example

The following example shows that the PKL expansion can prove itself as an efficient mathematical tool. We refer to the classical problem of detection of a known deterministic signal in Gaussian colored noise, which has been solved in various manners [8]. The two hypotheses are

$$r(t) = \begin{cases} f(t) + n(t), & 0 \leq t \leq T : H_1 \\ n(t), & 0 \leq t \leq T : H_0 \end{cases}$$

where $f(t)$ is the signal, and $n(t)$ is a real zero-mean, mean-square continuous Gaussian random process with autocorrelation function $R(s,t)$.

The optimal receiver can be constructed in this case, under the assumption that $f(t)$ belongs to the space spanned by the KL basis of $n(t)$, as follows. Construct a basis $\{\phi_i(t)\}$ which decorrelates $n(t)$ according to Theorem 1, s.t. $\phi_1(t) = f(t)$. This can be done by first solving

$$\int_0^T R(s,t)\psi_1(t)\,dt = \sigma_1 f(s) \tag{10}$$

for $\psi_1(t)$, and then proceeding according to Algorithm 1 from Step 2. We can express $r(t)$ in the following manner:

$$r(t) = \text{l.i.m.} \sum_i r_i\phi_i(t)$$

where

$$r_i = \begin{cases} \int_T^0 (f(t) + n(t))\psi_i(t)\,dt, & : H_1 \\ \int_T^0 n(t)\psi_i(t)\,dt, & : H_0. \end{cases}$$

Since $n(t)$ is a Gaussian random process and $\{n_i\}$

$$n_i = \int_0^T n(t)\psi_i(t)\,dt$$

is an uncorrelated set of variables, it follows that $\{n_i\}$, and therefore $\{r_i\}$ as well, are independent random variables. We have therefore to consider only r_1 and use only $\psi_1(t)$ which is obtained by solving (10). The optimal detector which emerges out of this analysis (Fig. 1) is in accord with standard decision-making schemes known in communication [8]. We stress again that this example is given only in order to illustrate how the concept of PKL expansion can be used in practical applications.
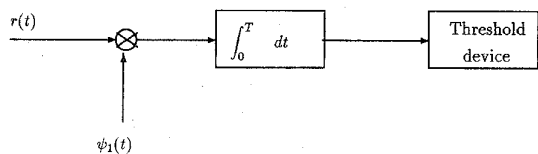
Fig. 1. An optimal detector of a deterministic signal with additive colored Gaussian noise.

Fig. 2. A schematic diagram illustrating the relationship of the various spaces of the representation bases.

III. THE DISCRETE CASE

Let $\{x_n, n \in M\}$, where M is a countable set, be a real discrete zero-mean random process, s.t.

$$E \sum x_n^2 < \infty.$$

Let R_x be its (possibly infinite) autocorrelation matrix. Given $\{\phi_k\}$ a basis of $\ell_2(M)$ which satisfies the following equation:

$$R_x \psi_k = \sigma_k \phi_k, \quad \forall k \tag{11}$$

where $\{\psi_k\}$ is the biorthonormal basis of $\{\phi_k\}$, i.e., $\phi_k \psi_l^T = \delta_{kl}, \forall k, l$, the expansion coefficients of x_n according to $\{\phi_k\}$ are uncorrelated.

Note that, in accordance with the result presented in [2], for any symmetric, positive-definite matrix A, s.t. $A^{-1/2} R_x A^{-1/2}$ is bounded and its eigenvectors span $\ell_2(M)$, the set $\{A^{1/2}\phi_k\}$ decorrelates the random process $\{x_n, n \in M\}$.

Equation (11) can be written as

$$R_x \Psi = \Phi \Sigma \tag{12}$$

where the columns of Ψ, Φ are the sets $\{\psi_k\}, \{\phi_k\}$ respectively, and Σ is a diagonal matrix with the set $\{\sigma_k\}$ along its diagonal. Moreover, $\Psi^T \Phi = I$, where I is the identity matrix. In the finite case, i.e., M is a finite set, (12) can be expressed as the following congruence relation:

$$R_x = \Phi \Sigma \Phi^T. \tag{13}$$

A particular example of (13) relates to the so-called LU decomposition. Then, it is known that a symmetric positive-definite matrix has a unique decomposition of the type $R_x = L\Sigma L^T$, where L is a lower triangular matrix with 1's on its diagonal.

Note that in the finite case, the iterative algorithm presented above converges in exactly n steps, where n is the number of elements in the set M.

IV. CONCLUSION

The generalization of the basic KL expansion presented in this correspondence can be summarized as follows. Given $\{x(t), t \in T\}$, a real zero-mean, mean-square continuous random process, we can always express $x(t)$ as

$$x(t) = \text{l.i.m.} \sum_i a_i \phi_i(t)$$

provided $\{\phi_i(t)\}$ forms a basis of the space spanned by the KL basis of $x(t)$. Define $\mathcal{B}$ as the set of all possible bases of the space spanned by the above KL basis $\mathcal{O}$ as the set of all orthonormal bases in $\mathcal{B}$, $\mathcal{B}_{PKL}$ as the set of all PKL bases in $\mathcal{B}$, and $\mathcal{O}_{KL}$ as the KL basis. Fig. 2 depicts the relationships between all of these sets.

The KL expansion is a special case of the PKL expansion. The KL expansion is optimal under the least mean-square criterion [8]. Thus the PKL expansion proposed in this correspondence cannot be better in this respect. Any PKL basis does retain, however, most of the important properties of the KL basis: in particular a PKL basis represents a random process by a set of uncorrelated
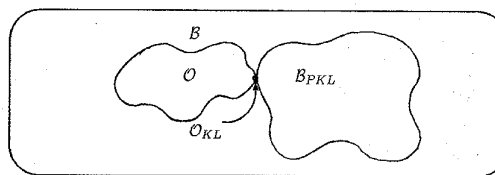
random variables, and we can order the representation functions according to their contribution to the mapping process. Furthermore, the PKL expansion offers a degree of freedom in the selection of the representation functions which can be advantageous in certain applications such as representation of certain classes of images or signals which can be considered as realization of stochastic processes, and for which certain constraints prohibit the utilization of the KL representation. Thus the PKL expansion can be a useful mathematical tool in a variety of applications where the specific properties of the KL expansion, but other functions, are desirable. Specific applications of the PKL expansion are currently under investigation.

REFERENCES

[1] B. H. Bharucha and T. T. Kadota, "On the representation of continuous parameter processes by a sequence of random variables," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 139–141, 1970.

[2] T. T. Kadota, "Simultaneous diagonalization of two covariance kernels and application to second order stochastic processes," *SIAM J. Appl. Math.*, vol. 15, pp. 1470–1480, 1967.

[3] ——, "Simultaneously orthogonal expansions of two stationary Gaussian processes—Examples," *Bell Syst. Tech. J.*, vol. 45, pp. 1071–1096, 1966.

[4] B. H. Bharucha, "A posteriori probability and detection theory," *Inform. Contr.*, vol. 14, pp. 98–132, 1969.

[5] K. Fukunaga and W. L. G. Koontz, "Representation of random processes using the finite Karhunen–Loève expansion," *Inform. Contr.*, vol. 16, pp. 85–101, 1970.

[6] M. Loève, *Probability Theory*, 2nd ed. Princeton, N.J.: Van Nostrand, 1960.

[7] J. R. Higgins, *Completeness and Basis Properties of Sets of Special Functions* Cambridge, UK: Cambridge Univ. Press, 1977.

[8] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.