A brief note on the Karhunen-Loève expansion

Alen Alexanderian*

Abstract

We provide a detailed derivation of the Karhunen-Loève expansion of a stochastic process. We also discuss briefly Gaussian processes, and provide a simple numerical study for the purpose of illustration.

Contents

1 Introduction

The purpose of this brief note is to provide a self-contained coverage of the idea of the Karhunen-Loève (KL) expansion of a stochastic process. The writing of this note was motivated by being exposed to the many applications of the KL expansion in uncertainty propagation through dynamical systems with random parameter functions (see e.g. in [3, 1]). Since a clear and at the same time rigorous coverage the KL exapnsion is not so simple to find in the literature, here we provide a simple exposition of the theoretical basis for the KL expansion, including a detailed proof of convergence. We will see that the KL expansion is obtained through an interesting application of the Spectral Theorem for compact normal operators, in conjunction with Mercer's theorem which connects the spectral representation of a Hilbert-Schmidt integral operator to the corresponding Hilbert-Schmidt kernel.

We begin by recalling some functional analytic basics on compact operators in Section 2. The material in that section are classical and can be found in many standard textbooks on the subject; see e.g., [5] for an accessible presentation. Next, Mercer's Theorem is recalled in Section 3. Then, we recall some basics regarding stochastic processes in Section 4. In that section, a basic result stating the equivalence of mean-square continuity of a stochastic process

*North Carolina State University, Raleigh, NC, USA. E-mail: alexanderian@ncsu.edu
Last revised: September 24, 2015

and the continuity of the corresponding autocorrelation function is mentioned also. In Section 5, we discuss in detail KL expansions of centered mean-square continuous stochastic processes including a proof of convergence. Finally, in Section 6, we provide a numerical example where the KL expansion of a Gaussian random field is studied.

2 Preliminaries on compact operators

Let us begin by recalling the notion of precompact and relatively compact sets.

Definition 2.1. *(Relatively Compact)*
Let X be a metric space; $A \subseteq X$ is relatively compact in X, if $\bar{A}$ is compact in X.

Definition 2.2. *(Precompact)*
Let X be a metric space; $A \subseteq X$ is precompact (also called totally bounded) if for every $\epsilon > 0$, there exist finitely many points $x_1, \ldots, x_N$ in A such that $\cup_1^N B(x_i, \epsilon)$ covers A.

The following Theorem shows that when we are working in a complete metric space, precompactness and relative compactness are equivalent.

Theorem 2.3. *Let X be a metric space. If $A \subseteq X$ is relatively compact then it is precompact. Moreover, if X is complete then the converse holds also.*

Then, we define a compact operator as below.

Definition 2.4. *Let X and Y be two normed linear spaces and $T : X \to Y$ a linear map between X and Y. T is called a compact operator if for all bounded sets $E \subseteq X$, $T(E)$ is relatively compact in Y.*

By the above definition 2.4, if $E \subset X$ is a bounded set, then $\overline{T(E)}$ is compact in Y. The following basic result shows a couple of different ways of looking at compact operators.

Theorem 2.5. *Let X and Y be two normed linear spaces; suppose $T : X \to Y$, is a linear operator. Then the following are equivalent.*

1. *T is compact.*
2. *The image of the open unit ball under T is relatively compact in Y.*
3. *For any bounded sequence $\{x_n\}$ in X, there exist a subsequence $\{Tx_{n_k}\}$ of $\{Tx_n\}$ that converges in Y.*

Let us denote by $B[X]$ the set of all bounded linear operators on a normed linear space space X:

$$B[X] = \{T : X \to X |\ T \text{ is a bounded linear transformation.}\}.$$

Note that equipped by the operator norm $B[X]$ is a normed linear space. It is simple to show that compact operators form a subspace of $B[X]$. The following result (cf. [5] for a proof) shows that the set of compact normal operators is in fact a closed subspace of $B[X]$.

Theorem 2.6. *Let $\{T_n\}$ be a sequence of compact operators on a normed linear space X. Suppose $T_n \to T$ in $B[X]$. Then, T is also a compact operator.*

Another interesting fact regarding compact linear operators is that they form an ideal of the ring of bounded linear mappings $B[X]$. This follows from the following basic result whose simple proof is also included for reader's convenience.

Lemma 2.7. *Let X be a normed linear space, and let T and S be in $B[X]$. If T is compact, then so are ST and TS.*

Proof. Consider the mapping ST. Let $\{x_n\}$ be a bounded sequence in X. Then, by Theorem 2.5(3), there exists a subsequence $\{Tx_{n_k}\}$ of $\{Tx_n\}$ that converges in X: $Tx_{n_k} \to y^* \in X$. Now, since S is continuous, it follows that $STx_{n_k} \to S(y^*)$; that is, $\{STx_{n_k}\}$ converges in X also, and so ST is compact. To show TS is compact, take a bounded sequence $\{x_n\}$ in X and note that $\{Sx_n\}$ is bounded also (since S is continuous). Thus, again by Theorem 2.5(3), there exists a subsequence $\{TSx_{n_k}\}$ which converges in X, and thus, TS is also compact. □

Remark 2.8. *A compact linear operator of an infinite dimensional normed linear space is not invertible in $B[X]$. To see this, suppose that T has an inverse S in $B[X]$. Now, applying the previous Lemma, we get that $I = TS = ST$ is also compact. However, this implies that the closed unit ball in X is compact, which is not possible since we assumed X is infinite dimensional.*

(Recall that the closed unit ball in a normed linear space X is compact if and only if X is finite dimensional.)

2.1 Hilbert-Schmidt operators

Let $D \subset \mathbb{R}^n$ be a bounded domain. We call a function $k : D \times D \to \mathbb{R}$ a Hilbert-Schmidt kernel if

$$\int_D \int_D |k(x,y)|^2 \, dx \, dy < \infty,$$

that is, $k \in L^2(D \times D)$ (note that one special case is when k is a continuous function on $D \times D$). Define the integral operator K on $L^2(D)$, $K : u \to Ku$ for $u \in L^2(D)$, by

$$[Ku](x) = \int_D k(x,y)u(y) \, dy. \tag{2.1}$$

It is simple to show that K is a bounded operator on $L^2(D)$. Linearity is clear. As for boundedness, we note that for every $u \in L^2(D)$,

$$\begin{aligned} \|Ku\|^2_{L^2(D)} = \int_D \Big|(Ku)(x)\Big|^2 \, dx &= \int_D \Big| \int_D k(x,y)u(y) \, dy\Big|^2 \, dx \\ &\leq \int_D \Big(\int_D |k(x,y)|^2 \, dy\Big)\Big(\int_D |u(y)|^2 \, dy\Big) \, dx \qquad \text{(Cauchy-Schwarz)} \\ &= ||k||_{L^2(D\times D)}||u||_{L^2(D)} < \infty. \end{aligned}$$

An integral operator K as defined above is called a *Hilbert-Schmidt operator*. The following result which is usually proved using Theorem 2.6 is very useful.

Lemma 2.9. *Let D be a bounded domain in R^n and let $k \in L^2(D \times D)$ be a Hilbert-Schmidt kernel. Then, the integral operator $K : L^2(D) \to L^2(D)$ given by $[Ku](x) = \int_D k(x,y)u(y) \, dy$ is a compact operator.*

2.2 Spectral theorem for compact self-adjoint operators

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$. A linear operator $T : H \to H$ is called self adjoint if

$$\langle Tx, y \rangle = \langle x, Ty \rangle , \quad \forall x, y \in H.$$

Example 2.10. *Let us consider a Hilbert-Schmidt operator K on $L^2([a,b])$ as in* (2.1) *(where for simplicity we have taken $D = [a,b] \subset \mathbb{R}$). Then, it is simple to show that K is self-adjoint if and only if $k(x,y) = k(y,x)$ on $[a,b] \times [a,b]$.*

A linear operator $T : H \to H$, is called positive if $\langle Tx, x \rangle \geq 0$ for all x in H. Recall that a scalar $\lambda \in \mathbb{R}$ is called an eigenvalue of T if there exists a non-zero $x \in H$ such that $Tx = \lambda x$. Note that the eigenvalues of a positive operator are necessarily non-negative.

Compact self-adjoint operators on infinite dimensioal Hilbert spaces resemble many properties of the symmetric matrices. Of particular interest is the spectral decomposition of a compact self-adjoint operator as given by the following:

Theorem 2.11. *Let H be a (real or complex) Hilbert space and let $T : H \to H$ be a compact self-adjoint operator. Then, H has an orthonormal basis $\{e_i\}$ of eigenvectors of T corresponding to eigenvalues λ_i. In addition, the following holds:*

1. *The eigenvalues λ_i are real having zero as the only possible point of accumulation.*
2. *The eigenspaces corresponding to distinct eigenvalues are mutually orthogonal.*
3. *The eigenspaces corresponding to non-zero eigenvalues are finite-dimensional.*

In the case of a positive compact self-adjoint operator, we know that the eigenvalues are non-negative. Hence, we may order the eigenvalues as follows

$$\lambda_1 \geq \lambda_2 \geq ... \geq 0.$$

Remark 2.12. *Recall that for a linear operator A on a finite dimensional linear space, we define its spectrum $\sigma(A)$ as the set of its eigenvalues. On the other hand, for a linear operator T on an infinite dimensional (real) normed linear space the spectrum $\sigma(T)$ of T is defined by,*

$$\sigma(T) = \{\lambda \in \mathbb{R} : T - \lambda I \text{ is not invertible in } B[X]\},$$

and $\sigma(T)$ is the disjoint union of the point spectrum (set of eigenvalues), contiuous spectrum, and residual spectrum (see [5] for details). As we saw in Remark 2.8, a compact operator T on an infinite dimensional space X cannot be invertible in $B[X]$; therefore, we always have $0 \in \sigma(T)$. However, not much can be said on whether $\lambda = 0$ is in point spectrum (i.e. an eigenvalue) or the other parts of the spectrum.

3 Mercer's Theorem

Let $D = [a, b] \subset \mathbb{R}$. We have seen that given a continuous kernel $k : D \times D \to \mathbb{R}$, we can define a Hilbert-Schmidt operator through (2.1) which is compact and has a complete set of eigenvectors in $L^2(D)$. The following result by Mercer provides a series representation for the kernel k based on spectral representation of the corresponding Hilbert-Schmidt operator K. A proof of this result can be found for example in [2].

Theorem 3.1 (Mercer)**.** *Let $k : D \times D \to \mathbb{R}$ be a continuous function, where $D = [a, b] \subset \mathbb{R}$. Suppose further that the corresponding Hilbert-Schmidt operator $K : L^2(D) \to L^2(D)$ given by* (2.1) *is postive. If $\{\lambda_i\}$ and $\{e_i\}$ are the eigenvalues and eigenvectors of K, then for all $s, t \in D$,*

$$k(s, t) = \sum_i \lambda_i e_i(s) e_i(t), \tag{3.1}$$

where convergence is absolute and uniform on $D \times D$.

4 Stochastic processes

In what follows we consider a probability space $(\Omega, \mathcal{F}, P)$, where Ω is a sample space, $\mathcal{F}$ is an appropriate σ-algebra on Ω and P is a probability measure. A real valued random variable X on $(\Omega, \mathcal{F}, P)$ is an $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable mapping $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The expectation and variance of a random variable X is denoted by,

$$\mathrm{E}[X] := \int_\Omega X(\omega)\, dP(\omega), \quad \mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right].$$

$L^2(\Omega, \mathcal{F}, P)$ denotes the Hilbert space of (equivalence classes) of real valued square integrable random variables on Ω:

$$L^2(\Omega, \mathcal{F}, P) = \{X : \Omega \to \mathbb{R} : \int_\Omega |X(\omega)|^2\, dP(\omega) < \infty\}.$$

with inner product, $\langle X, Y\rangle = \mathrm{E}[XY] = \int_\Omega XY\, dP$ and norm $||X|| = \langle X, X\rangle^{1/2}$.

Let $D \subseteq \mathbb{R}$, a stochastic prcess is a mapping $X : D \times \Omega \to \mathbb{R}$, such that $X(t, \cdot)$ is measurable for every $t \in D$; alternatively, we may define a stochastic process as a family of random variables, $X_t : \Omega \to \mathbb{R}$ with $t \in D$, and refer to X as $\{X_t\}_{x \in D}$. Both of these points of view of a stochastic process are useful and hence we will be switching between them as appropriate.

A stochastic process is called *centered* if $\mathrm{E}[X_t] = 0$ for all $t \in D$. Let $\{Y_t\}_{t \in D}$ be an arbitrary stochastic process. We note that

$$Y_t = \mathrm{E}[Y_t] + X_t,$$

where $X_t = Y_t - \mathrm{E}[Y_t]$ and $\{X_t\}_{t \in D}$ is a centered stochastic process. Therefore, without loss of generality, we will focus our attention to centered stochastic processes.

We say a stochastic process is *mean-square continuous* if

$$\lim_{\varepsilon \to 0} \mathrm{E}\left[(X_{t+\varepsilon} - X_t)^2\right] = 0.$$

The following definition is also useful.

Definition 4.1 (Realization of a stochastic process)**.** *Let $X : D \times \Omega \to \mathbb{R}$ be a stochastic process. For a fixed $\omega \in \Omega$, we define $\hat{X} : D \to \mathbb{R}$ by $\hat{X}(t) = X_t(\omega)$. We call $\hat{X}$ a realization of the stochastic process.*

For more details on theory of stochastic processes please consult [8, 6, 7].

4.1 Autocorrelation function of a stochastic process

The autocorrelation function of a stochastic process $\{X_t\}_{t\in D}$ is given by $R_X : D \times D \to \mathbb{R}$ defined through

$$R_X(s,t) = \mathrm{E}\left[X_s X_t\right], \quad s,t \in D.$$

The following well-known result states that for a stochastic process the continuity of its autocorrelation function is a necessary and sufficient condition for the mean-square continuity of the process.

Lemma 4.2. *A stochastic process $\{X_t\}_{t\in[a,b]}$ is mean-square continuous if and only if its autocorrelation function R_X is continuous on $[a,b]\times[a,b]$.*

Proof. Suppose R_X is continuous, and note that

$$\mathrm{E}\left[(X_{t+\varepsilon} - X_t)^2\right] = \mathrm{E}\left[X_{t+\varepsilon}^2\right] - 2\mathrm{E}\left[X_{t+\varepsilon}X_t\right] + \mathrm{E}\left[X_t^2\right] = R_X(t+\varepsilon,t+\varepsilon) - 2R_X(t+\varepsilon,t) + R_X(t,t).$$

Therefore, since R_X is continuous,

$$\lim_{\varepsilon\to 0} \mathrm{E}\left[(X_{t+\varepsilon} - X_t)^2\right] = \lim_{\varepsilon\to 0} R_X(t+\varepsilon,t+\varepsilon) - 2R_X(t+\varepsilon,t) + R_X(t,t) = 0.$$

That is X_t is mean-square continuous. Conversely, if X_t is mean-square continous we proceed as follows:

$$\begin{aligned}
&|R_X(t+\varepsilon,s+\nu) - R_X(t,s)| = |\mathrm{E}\left[X_{t+\varepsilon}X_{s+\nu}\right] - \mathrm{E}\left[X_tX_s\right]| \\
&= \left|\mathrm{E}\left[(X_{t+\varepsilon}-X_t)(X_{s+\nu}-X_s)\right] + \mathrm{E}\left[(X_{t+\varepsilon}-X_t)X_s\right] + \mathrm{E}\left[(X_{s+\nu}-X_s)X_t\right]\right| \\
&\le \left|\mathrm{E}\left[(X_{t+\varepsilon}-X_t)(X_{s+\nu}-X_s)\right]\right| + \left|\mathrm{E}\left[(X_{t+\varepsilon}-X_t)X_s\right]\right| + \left|\mathrm{E}\left[(X_{s+\nu}-X_s)X_t\right]\right| \\
&\le \mathrm{E}\left[(X_{t+\varepsilon}-X_t)^2\right]^{1/2}\mathrm{E}\left[(X_{s+\nu}-X_s)^2\right]^{1/2} + \mathrm{E}\left[(X_{t+\varepsilon}-X_t)\right]^{1/2}\mathrm{E}\left[X_s^2\right]^{1/2} \\
&\qquad + \mathrm{E}\left[(X_{s+\nu}-X_s)^2\right]^{1/2}\mathrm{E}\left[X_t^2\right]^{1/2},
\end{aligned}$$

where the last inequality follows from Cauchy-Schwarz inequality. Thus, we have,

$$\begin{aligned}
|R_X(t+\varepsilon,s+\nu) - R_X(t,s)| &\le \mathrm{E}\left[(X_{t+\varepsilon}-X_t)^2\right]^{1/2}\mathrm{E}\left[(X_{s+\nu}-X_s)^2\right]^{1/2} \\
&\quad + \mathrm{E}\left[(X_{t+\varepsilon}-X_t)\right]^{1/2}\mathrm{E}\left[X_s^2\right]^{1/2} + \mathrm{E}\left[(X_{s+\nu}-X_s)^2\right]^{1/2}\mathrm{E}\left[X_t^2\right]^{1/2}, \quad (4.1)
\end{aligned}$$

and therefore, by mean-square continuity of X_t we have that

$$\lim_{(\varepsilon,\nu)\to(0,0)} |R_X(t+\varepsilon,s+\nu) - R_X(t,s)| = 0. \qquad \square$$

5 Karhunen-Loève expansion

Let $D \subseteq \mathbb{R}$. In this section, we assume that $X : D \times \Omega \to \mathbb{R}$ is a *centered mean-square continuous* stochastic process such that $X \in L^2(D\times\Omega)$. With the technical tools from the previous sections, we are now ready to derive the KL expansion of X.

Define the integral operator $K : L^2(D) \to L^2(D)$ by

$$[Ku](s) = \int_D k(s,t)u(t)\,dt, \quad k(s,t) = R_X(s,t), \qquad (5.1)$$

The following lemma summarizes the properties of the operator K.

Lemma 5.1. *Let $K : L^2(D) \to L^2(D)$ be as in* (5.1)*. Then the following hold:*

1. *K is compact.*
2. *K is positive*
3. *K is self-adjoint.*

Proof. (1) Since the process X is mean-square continuous, Lemma 4.2 implies that $k(s,t) = R_X(s,t)$ is continuous. Therefore, by Lemma 2.9, K is compact.

(2) We need to show $\langle Ku, u\rangle \geq 0$ for every $u \in L^2(D)$, where $\langle\cdot,\cdot\rangle$ denotes the $L^2(D)$ inner product.

$$\begin{aligned}
\langle Ku, u\rangle = \int_D Ku(s)u(s)\,ds &= \int_D \Big(\int_D k(s,t)u(t)\,dt\Big)u(s)\,ds \\
&= \int_D \Big(\int_D \mathrm{E}\left[X_sX_t\right]u(t)\,dt\Big)u(s)\,ds \\
&= \mathrm{E}\left[\int_D\int_D X_sX_tu(t)u(s)\,dt\,ds\right] \\
&= \mathrm{E}\left[\Big(\int_D X_su(s)\,ds\Big)\Big(\int_D X_tu(t)\,dt\Big)\right] \\
&= \mathrm{E}\left[\Big(\int_D X_tu(t)\,dt\Big)^2\right] \geq 0,
\end{aligned}$$

where we used Fubini's Theorem to interchange integrals.

(3) This follows trivially from $R_X(s,t) = R_X(t,s)$ and Fubini's theorem:

$$\langle Ku, v\rangle = \int_D Ku(s)v(s)\,ds = \int_D \Big(\int_D k(t,s)v(s)\,ds\Big)u(t)\,dt = \langle u, Kv\rangle\,. \qquad \square$$

Now, let K be defined as in (5.1) the previous lemma allows us to invoke the spectral theorem for compact self-adjoint operators to conclude that K has a complete set of eigenvectors $\{e_i\}$ in $L^2(D)$ and real eigenvalues $\{\lambda_i\}$:

$$Ke_i = \lambda_i e_i. \tag{5.2}$$

Moreover, since K is positive, the eigenvalues λ_i are non-negative (and have zero as the only possible accumulation point). Now, the stochastic process X which we fixed in the beginning of this section is assumed to be square integrable on $D \times \Omega$ and thus, we may use the basis $\{e_i\}$ of $L^2(D)$ to expand X_t as follows,

$$X_t = \sum_i x_i e_i(t), \quad x_i = \int_D X_t e_i(t)\,dt \tag{5.3}$$

The above equality is to be understood in mean square sense. To be most specific, at this point we have that the realizations $\hat{X}$ of the stochastic process X admit the expansion

$$\hat{X} = \sum_i x_i e_i$$

where the convergence is in $L^2(D \times \Omega)$. We will see shortly that the result is in fact stronger, and we have

$$\lim_{N\to\infty} \mathrm{E}\left[\Big(X_t - \sum_{i=1}^N x_i e_i(t)\Big)^2\right] = 0,$$

uniformly in D, and thus, as a consequence, we have that (5.3) holds for all $t \in D$. Before proving this, we examine the coefficients x_i in (5.3). Note that x_i are random variables on Ω. The following lemma summarizes the properties of the coefficients x_i.

Lemma 5.2. *The coefficients x_i in* (5.3) *satisfy the following:*

1. $\mathrm{E}\left[x_i\right] = 0$
2. $\mathrm{E}\left[x_ix_j\right] = \delta_{ij}\lambda_j$.
3. $\mathrm{Var}\left[x_i\right] = \lambda_i$.

Proof. To see the first assertion note that

$$\begin{aligned}
\mathrm{E}\,[x_i] &= \mathrm{E}\left[\int_D X_t e_i(t)\,dt\right] \\
&= \int_\Omega \int_D X_t(\omega) e_i(t)\,dt\,dP(\omega) \\
&= \int_D \int_\Omega X_t(\omega) e_i(t)\,dP(\omega)\,dt \quad \text{(Fubini)} \\
&= \int_D \mathrm{E}\,[X_t]\, e_i(t)\,dt = 0,
\end{aligned}$$

where the last conclusion follows from $\mathrm{E}\,[X_t] = 0$ (X is a centered process). To see the second assertion, we proceed as follows

$$\begin{aligned}
\mathrm{E}\,[x_i x_j] &= \mathrm{E}\left[\Big(\int_D X_s e_i(s)\,ds\Big)\Big(\int_D X_t e_j(t)\,dt\Big)\right] \\
&= \mathrm{E}\left[\int_D \int_D X_s e_i(s) X_t e_j(t)\,ds\,dt\right] \\
&= \int_D \int_D \mathrm{E}\,[X_s X_t]\, e_i(s) e_j(t)\,ds\,dt \\
&= \int_D \Big(\int_D k(s,t) e_j(t)\,dt\Big) e_i(s)\,ds \\
&= \int_D [K e_j](s) e_i(s)\,ds \qquad \text{(from (5.1))} \\
&= \langle K e_j, e_i\rangle \\
&= \langle \lambda_j e_j, e_i\rangle \\
&= \lambda_j \delta_{ij},
\end{aligned}$$

where again we have used Fubini's Theorem to interchange integrals and the last conclusion follows from orthonormality of eigenvectors of K. The assertion (3) of the lemma follows easily from (1) and (2):

$$\mathrm{Var}\,[x_i] = \mathrm{E}\left[(x_i - \mathrm{E}\,[x_i])^2\right] = \mathrm{E}\left[x_i^2\right] = \lambda_i. \qquad \square$$

Now, we have the technical tools to prove the following:

Theorem 5.3 (Karhunen-Loeve)**.** *Let $X : D \times \Omega \to \mathbb{R}$ be a centered mean-square continuous stochastic process with $X \in L^2(\Omega \times D)$. There exist a basis $\{e_i\}$ of $L^2(D)$ such that for all $t \in D$,*

$$X_t = \sum_{i=1}^{\infty} x_i e_i(t), \quad \text{in } L^2(\Omega),$$

where coefficients x_i are given by $x_i(\omega) = \int_D X_t(\omega) e_i(t)\,dt$ and satisfy the following.

1. $\mathrm{E}\,[x_i] = 0$
2. $\mathrm{E}\,[x_i x_j] = \delta_{ij}\lambda_j$.
3. $\mathrm{Var}\,[x_i] = \lambda_i$.

Proof. Let K be the Hilbert-Schmidt operator defined as in (5.1). We know that K has a complete set of eigenvectors $\{e_i\}$ in $L^2(D)$ and non-negative eigenvalues $\{\lambda_i\}$. Note that $x_i(\omega) = \int_D X_t(\omega) e_i(t)\,dt$ satisfy the the properties (1)-(3) by Lemma 5.2. Next, consider

$$\varepsilon_n(t) := \mathrm{E}\left[\Big(X_t - \sum_{i=1}^{n} x_i e_i(t)\Big)^2\right].$$

The rest of the proof amounts to showing $\lim_{n\to\infty} \varepsilon_n(t) = 0$ uniformly (and hence pointwise) in D.

$$\begin{aligned}
\varepsilon_n(t) &= \mathrm{E}\left[\Big(X_t - \sum_{i=1}^{n} x_i e_i(t)\Big)^2\right] \\
&= \mathrm{E}\left[X_t^2\right] - 2\mathrm{E}\left[X_t \sum_{i=1}^{n} x_i e_i(t)\right] + \mathrm{E}\left[\sum_{i,j=1}^{n} x_i x_j e_i(t) e_j(t)\right]
\end{aligned} \tag{5.4}$$

Now, $\mathrm{E}\left[X_t^2\right] = k(t,t)$ with k as in (5.1),

$$\begin{aligned}\mathrm{E}\left[X_t \sum_{i=1}^{n} x_i e_i(t)\right] &= \mathrm{E}\left[X_t \sum_{i=1}^{n} \Big(\int_D X_s e_i(s)\, ds\Big) e_i(t)\right] \\ &= \sum_{i=1}^{n} \Big(\int_D \mathrm{E}\left[X_t X_s\right] e_i(s)\, ds\Big) e_i(t) \\ &= \sum_{i=1}^{n} \Big(\int_D k(t,s) e_i(s)\, ds\Big) e_i(t) = \sum_{i=1}^{n} [Ke_i](t) e_i(t) = \sum_{i=1}^{n} \lambda_i e_i(t)^2. \end{aligned} \tag{5.5}$$

Through a similar argument, we can show that

$$\mathrm{E}\left[\sum_{i,j=1}^{n} x_i x_j e_i(t) e_j(t)\right] = \sum_{i=1}^{n} \lambda_i e_i(t)^2 \tag{5.6}$$

Therefore, by (5.4), (5.5), and (5.6) we have

$$\varepsilon_n(t) = k(t,t) - \sum_{i=1}^{n} \lambda_i e_i(t) e_i(t),$$

invoking Theorem 3.1 (Mercer's Theorem) we have

$$\lim_{n\to\infty} \varepsilon_n(t) = 0,$$

uniformly; this completes the proof. □

Remark 5.4. *Suppose $\lambda_k = 0$ for some k, and consider the coefficient x_k in the expansion* (5.3). *Then, we have by the above Theorem $\mathrm{E}\left[x_k\right] = 0$ and $\mathrm{Var}\left[x_k\right] = \lambda_k = 0$, and therefore, $x_k = 0$. That is, the coefficient x_k corresponding to a zero eigenvalue is zero. Therefore, only x_i corresponding to postive eigenvalues λ_i appear in KL expansion of a square integrable, centered, and mean-square continous stochastic process.*

In the view of the above remark, we can normalize the coefficients x_i in a KL expansion and define $\xi_i = \frac{1}{\sqrt{\lambda_i}} x_i$. This leads to the following, more familiar, version of Theorem 5.3.

Corollary 5.5. *Let $X : D \times \Omega \to \mathbb{R}$ be a centered mean-square continuous stochastic process with $X \in L^2(\Omega \times D)$. There exist a basis $\{e_i\}$ of $L^2(D)$ such that for all $t \in D$,*

$$X(t,\omega) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i(\omega) e_i(t) \quad \text{in } L^2(\Omega). \tag{5.7}$$

where ξ_i are centered mutually uncorrelated random variables with unit variance and are given by,

$$\xi_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_D X_t(\omega) e_i(t)\, dt.$$

The KL expansion of a Gaussian process has the further property that ξ_i are independent standard normal random variables (see e.g. [3, 1]). The latter is a useful property in practical applications; for instance, this is used extensively in the method of stochastic finite element [1]. Moreover, in the case of a Gaussian process, the series representation in (5.7) converges almost surely [4].

6 A classical example

Here we consider the KL decomposition of a Gaussian random field X, which is characterized by its variance σ^2 and an autocorrelation function $R_X(s,t)$ given by,

$$R_X(s,t) = \sigma^2 \exp\Big(- \frac{|s-t|}{L_c}\Big). \tag{6.1}$$

We show in Figure 1 a plot of $R_X(s,t)$ over $[0,1] \times [0,1]$.

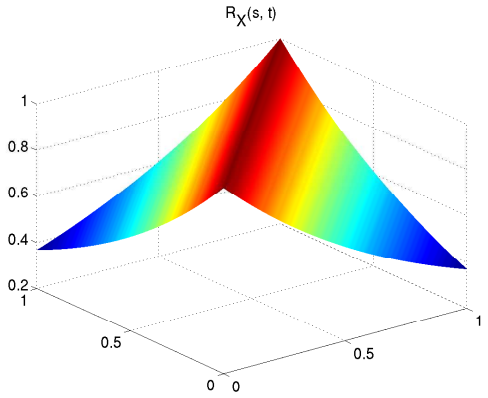


Figure 1: The autocorrelation function.

6.1 Spectral decomposition of the autocorrelation function

For this particular example, the eigenfunctions $e_i(t)$ and eigenvalues λ_i can be computed analytically. The analytic expression for eigenvalues and eigenvectors can be found for example in [1, 3]. We consider the case of $\sigma^2 = 1$ and $L_c = 1$ in (6.1). In Figure 2, we show the first few eigenfunctions and eigenvalues of the autocorrelation function defined in (6.1). To get an idea

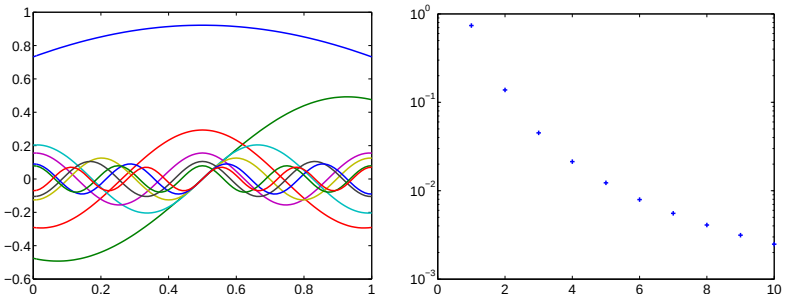


Figure 2: The first few eigenfunctions (a) and eigenvalues (b) of the autocorrelation function.

of how fast the approximation,

$$R_X^N(s,t) = \sum_{i=1}^{N} \lambda_i e_i(s) e_i(t)$$

converges to $R_X(s,t)$ we show in Figure 3 the plots of $R_X^N(s,t)$ for $N = 2, 4, 6, 8$. In Figure 4, we see that with $N = 6$, absolute error is bounded by 8×10^{-2}.

6.2 Simulating the random field

Having the eigenvalues and eigenfunctions of $R_X(t,\omega)$ at hand, we can simulate the random field $X(t,\omega)$ with a truncated KL expansion,

$$X_{\text{trunc}}^N(t,\omega) := \sum_{i=1}^{N} \sqrt{\lambda_i}\xi_i(\omega) e_i(t).$$

As discussed before, in this case, ξ_i are independent standard normal variables. In Figure 5(a), we plot a few realizations of the truncated KL expansion of $X(t,\cdot)$, $t \in [0,1]$ and in Figure 5(b), we show the distribution of $X(t,\cdot)$ at $t = 1/2$ versus standard normal distribution. For this experiment we used a low oreder KL expansion with $N = 6$ terms.

6.3 A final note regarding practical applications of KL expansions

In practice, when using KL expansions to model uncertainties in mathematical models, a premature *a priori* truncation of the KL expansion could potentially lead to misleading results,

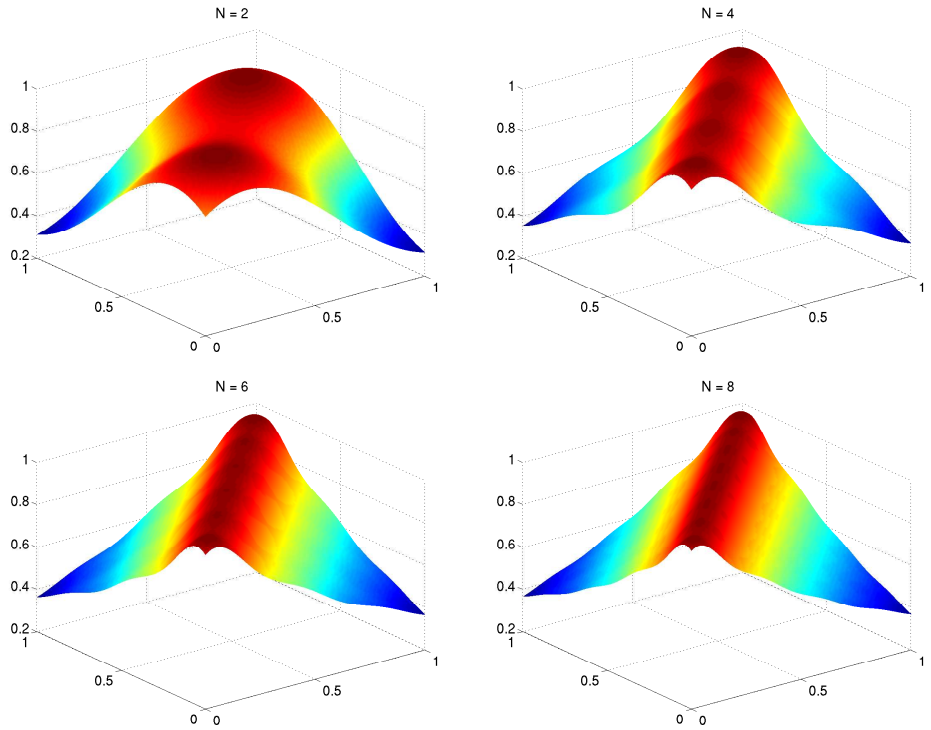
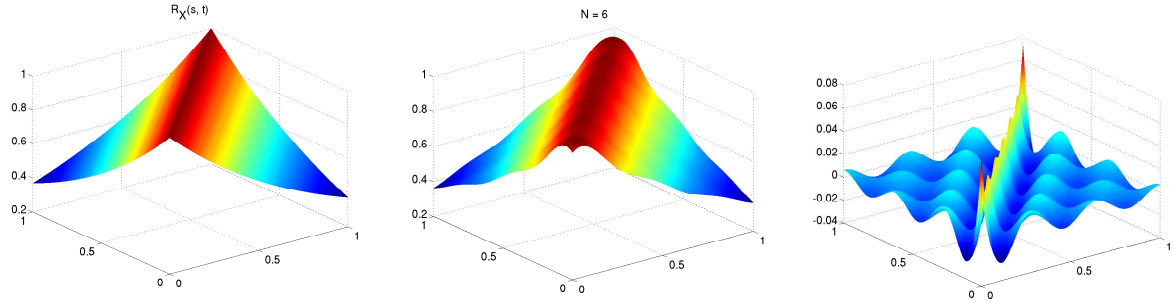


Figure 3: Improvements of the approximations to $R_X(s,t)$ as the expansion order is increased.



(a) (b) (c)

Figure 4: (a) The autocorrelation function $R_X(s,t)$, (b) the approximation $R_X^N(s,t)$ with $N = 6$, and (c) pointwise difference between $R_X(s,t)$ and $R_X^N(s,t)$ with $N = 6$.

because the effect of the higher order oscillatory modes on the output of a physical system could be significant. Also, sampling such a low-order KL expansion results in realizations of the random field that might look artificially smooth; see for example the realizations of a low-order KL expansion reported in Figure 5. In Figure 6 we illustrate the influence of the higher order modes on the realizations of the truncated KL expansion, in the context of the same example; in the figure, we consider two fixed realizations of the process, and for each realization we plot $X_{\text{trunc}}^N(t,\omega)$ with successively larger values of N.

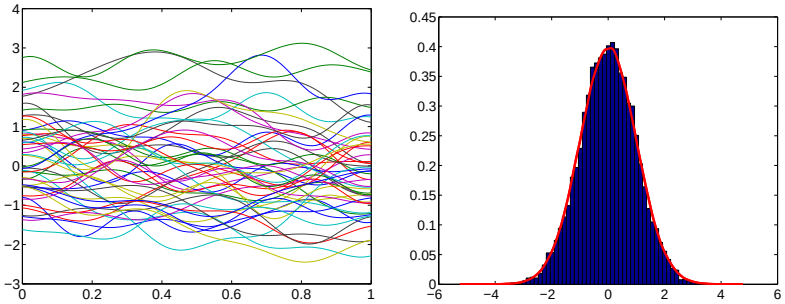


Figure 5: (a) A few realizations of the random field $X(t,\cdot)$ approximated by a truncated KL expansion with $N = 6$ terms. (b) distribution of $X(t,\omega)$ at $t = 1/2$ (blue) versus a standard normal distribution (red).

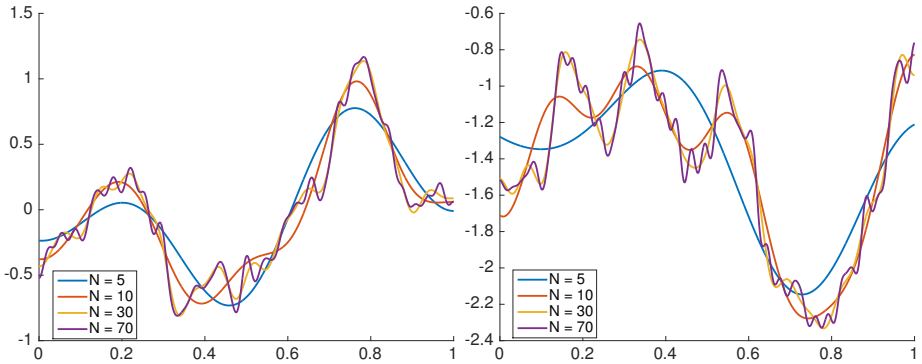


Figure 6: Two realizations of the random field $X(t,\cdot)$ simulated via a truncated KL expansion. To see the influence of the higher order oscillations captured by higher order KL modes, we successively increase the truncation order N.

References

[1] Roger G. Ghanem and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag New York, Inc., New York, NY, USA, 1991.

[2] Israel Gohberg, Seymour Goldberg, and M. A. Kaashoek. *Basic classes of linear operators*. 2004.

[3] Olivier P. Le Maitre and Omar M. Knio. *Spectral Methods for Uncertainty Quantification With Applications to Computational Fluid Dynamics*. Scientific Computation. Springer, 2010.

[4] Michel Loève. *Probability theory. I*. Springer-Verlag, New York-Heidelberg, fourth edition, 1977. Graduate Texts in Mathematics, Vol. 45.

[5] Arch W. Naylor and George Sell. *Linear operator theory in engineering and science*. Springer-Verlag, New York, 1982.

[6] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Foundations, Reprint of the second (1994) edition.

[7] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[8] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.