Sensitivity Analysis for Functional of Gaussian Processes: Karhunen-Loève Representations and Malliavin Calculus Techniques

Eleni E. Roumelioti* and Michael A. Zazanis†
Department of Statistics
Athens University of Economics and Business
Athens 10434, Greece

Abstract

We consider the problem of sensitivity analysis of functionals of Gaussian processes with respect to perturbations in the mean and the covariance function. The approach taken to sensitivity analysis is essentially via anticipating versions of the Girsanov theorem. A representation of the Gaussian process via a Gaussian measure on a Hilbert space is used together with the Feldman-Hajek theorem on the equivalence of Gaussian measures in order to obtain the desired sensitivity estimators. As expected, a necessary condition for such an approach is that the perturbations lie in the Cameron-Martin space of the covariance operator of the Gaussian process.

1 Introduction

In this paper we consider the problem of obtaining sensitivity estimators for functionals of Gaussian processes. Suppose that $\{X_t, t \in [0,T]\}$ is a real-valued gaussian process defined in $[0,T]$, and suppose in general that its mean, $m(t)$, and its covariance, $R(s,t) := \mathsf{Cov}(X_s, X_t)$, depend on a real parameter $\theta \in I$ (where I is an appropriate interval). We will assume that $\{X_t\}$ has continuous paths with probability 1 and consider a bounded, measurable functional $F : C[0,T] \to \mathbb{R}$. Suppose that the mean $m(t)$ of the process is a continuous function of time and that its dependence on the parameter θ is smooth in the sense that $\frac{\partial}{\partial\theta} m(t) = \alpha(t)$ for all $t \in [0,T]$. We will also assume that the covariance function $R(s,t)$ depends on the parameter θ smoothly so that $\frac{\partial}{\partial\theta} R(s,t) = V(s,t)$. Consider the performance criterion $J(\theta) := \mathbb{E}[F(X)]$. In order to obtain an efficient estimate for the sensitivity of $\frac{d}{d\theta} J(\theta)$ with respect to the parameter θ, one possible approach is to use a change of measure argument. By considering a Hilbert space representation of the Gaussian process and using known results on the equivalence of Gaussian measures in Hilbert space and in particular the Feldman-Hajek theorem [4] we will examine the conditions under which the measures induced in Hilbert space by the processes $X_t(\theta)$ and $X_t(\theta+\epsilon)$ are equivalent. Provided that they are the

*Supported in part by AUEB Grant # -2626-01
†Corresponding Author

corresponding Radon-Nikodým derivative $\frac{d\mu^{\theta+\epsilon}}{d\mu^\theta}$ will be used to obtain an efficient estimate of the sensitivity as follows. Setting

$$H := \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\frac{d\mu^{\theta+\epsilon}}{d\mu^\theta}-1\right) \tag{1}$$

we may obtain

$$\begin{aligned} J'(\theta) &= \lim_{\epsilon\to 0}\frac{\mathbb{E}_{\theta+\epsilon}[F(X)]-\mathbb{E}[F(X)]}{\epsilon} = \mathbb{E}_\theta\left[F(X)\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\frac{d\mu^{\theta+\epsilon}}{d\mu^\theta}-1\right)\right] \\ &= \mathbb{E}[F(X)H] \end{aligned} \tag{2}$$

provided that the interchange between the limit and the expectation above is justified.

2 Sensitivity via the Likelihood Ratio for Cylindrical Functionals of a Gaussian Process

We begin with the relatively trivial case where the functional F is *cylindrical* namely $F(X) = f(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ where $0 < t_1 < t_2 \ldots < t_n < T$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a bounded Borel function. Denote by R_n the $n \times n$ symmetric matrix with elements $[R_n]_{ij} = \mathsf{Cov}(X_{t_i}, X_{t_j})$ and by $\mu_n \in \mathbb{R}^n$ the vector $(\mathbb{E}X_{t_1}, \ldots, \mathbb{E}X_{t_n})^\top$. Both R_n and μ_n are differentiable functions of the underlying parameter θ. We then have the following

Theorem 1. *If $J(\theta) = \mathbb{E}[f(X_{t_1}, \ldots, X_{t_n})]$, under the above assumptions for the Gaussian process $\{X_t\}$, and assuming that R_n has full rank,*

$$J'(\theta) = \mathbb{E}[f(X_{t_1}, \ldots, X_{t_n})\, H]$$

where

$$\begin{aligned} H &= \partial_\theta \mu_n^\top R_n^{-1}(X_n-\mu_n) \\ &\quad + \frac{1}{2}\Big((X_n-\mu_n)^\top R_n^{-1}(\partial_\theta R_n)R_n^{-1}(X_n-\mu_n) \;-\; tr(R_n^{-1}\partial_\theta R_n)\Big) \end{aligned} \tag{3}$$

where, in the above expression, $X = (X_{t_1}, X_{t_2}, \ldots, X_{t_n})^\top$.

We point out for future reference that if only the mean depends on the parameter θ

$$H_{\text{mean}} = \partial_\theta \mu_n^\top R_n^{-1}(X_n-\mu_n) \tag{4}$$

and if only the covariance depends on θ

$$H_{\text{cov}} = \frac{1}{2}(X_n-\mu_n)^\top R_n^{-1}(\partial_\theta R_n)R_n^{-1}(X_n-\mu_n) \;-\; \frac{1}{2}\text{tr}(R_n^{-1}\partial_\theta R_n). \tag{5}$$

Extending the above technique to more general functionals F is not trivial. We will consider gaussian processes which have continuous sample paths with probability 1. It is rather easy to see that a necessary and sufficient condition for a centered gaussian process $\{X_t; t \in [0,T]\}$ to be continuous in the mean square sense is that its covariance function R be continuous at the diagonal (and therefore continuous on $[0,T] \times [0,T]$). The conditions under which the sample paths of a gaussian process are continuous with probability 1 are more complicated (see [1]). A sufficient condition that ensures the a.s. continuity for the sample paths is that ([1, p.14]) $\mathbb{E}(X_t - X_s)^2$ is sufficiently small for $|t-s|$ in the following sense:

$$\text{for some } C > 0 \text{ and } \alpha, \eta > 0 \quad \sup_{|s-t|<\eta} \mathbb{E}(X_t - X_s)^2 \leq \frac{C}{|\log|t-s||^{1+\alpha}}.$$

3 Symmetric Operators in Hilbert Spaces and Mercer's Theorem

Here we present some standard results in functional analysis. For further background and proofs we refer the reader to [24] and [19]. Let H be a real separable Hilbert space and denote by $L(H)$ the Banach algebra of all continuous linear operators $T : H \to H$. An operator $T \in L(H)$ is called compact if, for every bounded sequence $\{x_n\}$ in H, $\{Tx_n\}$ contains a converging subsequence. T is called *symmetric* or *self-adjoint* if, for any $x, y \in H$, $\langle Tx, y\rangle = \langle x, Ty\rangle$. $T \in L(H)$ is called *Hilbert-Schmidt* if there exists a complete orthonormal sequence $\{e_n\}$ in H such that $\sum_{n=1}^{\infty} \|Te_n\|^2 < \infty$. If T is Hilbert-Schmidt then it is compact.

Theorem 2 (Spectral Theorem)**.** *Let K be a compact symmetric operator in $L(H)$. Then there exists a finite or infinite sequence of orthonormal eigenvectors of K, $\{\phi_k\}$ with corresponding real eigenvalues μ_k such that*

$$Kx = \sum_k \mu_k \langle \phi_k, x\rangle\, \phi_k \qquad \textit{for all } x \in H.$$

A symmetric operator K is called *positive* iff $\langle Kx, x\rangle > 0$ for all $x \neq 0$. In this case the spectral theorem ensures that $\mu_k > 0$ i.e. all eigenvalues are positive and that the corresponding normalized eigenvector sequence is a complete orthonormal system in H.

Suppose that a symmetric operator K is also Hilbert-Schmidt with normalized eigenvectros $\{\phi_k\}$ and corresponding eigenvalues μ_k. Denoting by W the range of K it holds that the ϕ_k that correspond to non-zero eigenvalues span W. The orthogonal complement of W, $W^\top$ is the kernel of K (since K is symmetric) and we can choose an orthornormal set $\{\phi_l\}$, orthogonal to W, which spans $W^\top$. Thus we have an complete orthonormal sequence $\{\phi_n\}$ which spans H. Hence, if a symmetric operator K is also Hilbert-Schmidt,

$$\sum_{k=1}^{\infty} \mu_k^2 = \sum_{k=1}^{\infty} \langle K\phi_k, K\phi_k\rangle = \sum_{k=1}^{\infty} \|K\phi_k\|^2 < \infty.$$

Given a complete orthonormal sequence (e_k), $k = 1, 2, \ldots$ in H, the *trace* of a self-adjoint, non-negative operator $T \in L(H)$ is defined as

$$\mathrm{Tr}(T) \;=\; \sum_{k=1}^{\infty} \langle Te_k, e_k\rangle \;=\; \sum_k \lambda_k$$

where λ_k are the eigenvalues of T, provided that the series converges. (Note that the series either converges or diverges to $+\infty$ since the eigenvalues are non-negative.) If the trace of a self-adjoint, non-negative operator is finite then it is called *a trace class* or *nuclear* operator. Trace class operators are of course also Hilbert-Schmidt but the converse does not hold.

Let us now identify the real Hilbert space H with $L^2[0,T]$, the space of square integrable real functions defined on $[0,T]$. Thus the inner product in H is defined as $\langle x,y\rangle = \int_0^T x(t)y(t)dt$. Suppose that $R : [0,T]\times[0,T] \to \mathbb{R}$ is a continuous, positive definite, symmetric kernel, i.e. $R(s,t) = R(t,s)$ for all $s,t \in [0,T]$ and

$$\langle Rf, g\rangle = \int_0^T \int_0^T R(s,t)f(s)g(t) > 0 \qquad \text{for all } f,g \in L^2[0,T].$$

Then the following theorem due to Mercer holds [19, p.245]

Theorem 3 (Mercer)**.** *If the transformation R generated by the continuous symmetric kernel $R(x,y)$ is positive, that is, if $(Rf,f) \geq 0$ for all f, or equivalently, if all the characteristic values $\mu_i \neq 0$ are positive, the development*

$$R(x,y) = \sum_i \mu_i \phi_i(x)\phi_i(y)$$

is uniformly and absolutely convergent and the eigenfunctions ϕ_i are continuous.

The covariance $R : [0,T]\times[0,T] \to \mathbb{R}$ of the Gaussian process is a kernel satisfying the assumptions of Mercer's theorem. Let $\{e_k\}$, $k = 1,2,\dots$ denote the sequence of normalized eigenfunctions of the covariance kernel R and λ_k, $k = 1,2,\dots$ the corresponding eigenvalues which are of course positive since R is positive definite. A well known direct consequence of Mercer's theorem (since the eigenfunctions are normalized) is that $\sum_{k=1}^\infty \lambda_k = \int_0^T R(t,t)dt < \infty$ (the boundedness of the integral following from the continuity of R) and therefore the covariance operator is a *trace class* operator. We will also assume for ease of exposition that the kernel of the operator R is trivial, i.e. that $Rf = 0$ implies $f = 0$ (as an element of H). The above imply of course that $\lim_{k\to\infty} \lambda_k = 0$ and thus, without loss of generality we may assume that the eigenvalues have been ordered, namely $\lambda_1 \geq \lambda_2 \geq \ldots > 0$.

4 Gaussian Measures on Hilbert Spaces

Let $L^+(H)$ denote the subset of $L(H)$ consisting of *symmetric positive operators:* $T \in L^+(H)$ if $\langle Tx,y\rangle = \langle x,Ty\rangle$ for all $x,y \in H$ and $\langle Tx,x\rangle \geq 0$ for all $x \in H$. Also let $L_1^+(H)$ denote the subset of $L^+(H)$ such that if (e_k) is a complete orthonormal system in H then $\text{Tr}(Q) := \sum_{k=1}^\infty \langle Qe_k, e_k\rangle < \infty$. Thus $L_1^+(H)$ denotes the set of *trace class positive symmetric operators.*

Let μ be a probability measure on $(H,\mathcal{B}(H))$ where H is a real (separable) Hilbert space H and $\mathcal{B}(H)$ the Borel $\sigma-$field on H.

Definition 4. *A measure μ defined on $(H,\mathcal{B}(H))$ is Gaussian if there exists $m \in H$ and $Q \in L_1^+(H)$ such that*

$$\int_H e^{i\langle h,x\rangle}\mu(dx) \;=\; e^{i\langle m,h\rangle - \frac{1}{2}\langle Qh,h\rangle}, \qquad \forall\, h \in H. \tag{6}$$

Theorem 5. *An element $m \in H$ and an operator $Q \in L_1^+(H)$ defines uniquely a Gaussian measure μ on $(H,\mathcal{B}(H))$. The converse also holds. Every Gaussian measure μ on $(H,\mathcal{B}(H))$ determines an element $m \in H$, the mean, and a covariance operator $Q \in L_1^+(H)$ so that (6) holds.*

Proof. Denote by $\mathcal{N}_{a,\sigma^2}$ the normal law on $\mathbb{R}$ with mean a and variance σ^2 which has density with respect to the Lebesgue measure

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Assume that $\mathrm{Ker}(Q) = \{0\}$ and let $\{e_k\}$ be the complete orthonormal sequence of eigenvectors of Q, i.e. $Qe_k = \lambda_k e_k$, $k = 1, 2, \ldots$. Also set $m_k = \langle e_k, m\rangle$ for $k = 1, 2, \ldots$ and note that, since $\{e_k\}$ is complete $m = \sum_{k=1}^{\infty} m_k e_k$.

Consider now the measurable space $(\mathbb{R}^\infty, \mathcal{B}(R^\infty))$ of all real sequences $x = (x_1, x_2, \ldots)$ with the product measure $\mu := \bigotimes_{k=1}^{\infty} \mathcal{N}_{m_k,\lambda_k}$ on $(\mathbb{R}^\infty, \mathcal{B}(R^\infty))$. Using a monotone convergence argument (see [5, p.11]) we have

$$\int_{\mathbb{R}^\infty} \left(\sum_{k=1}^{\infty} x_k^2\right) \mu(dx) = \sum_{k=1}^{\infty} (m_k^2 + \lambda_k) < \infty,$$

which implies that $\mu(x \in \mathbb{R}^\infty : \sum_{k=1}^{\infty} x_k^2 < \infty) = 1$ i.e. that the measure μ is concentrated on ℓ^2, the subspace of all square-summable real sequences. Using the natural isomorphism between ℓ^2 and H, $\gamma : H \to \ell^2$, (where $\ell^2 = \left\{(x_k)_{k\in\mathbb{N}} : \sum_{k=1}^{\infty} |x_k|^2 < \infty\right\}$) by $\gamma(x) = (x_k)$ we have thus constructed a measure on H (which we will again denote by μ).

To show the converse suppose that μ is a gaussian probability measure on H. Then $\int_H |x|\mu(dx) < \infty$. Consider the linear functional $F : H \to \mathbb{R}$ defined by $F(h) := \int_H \langle x, h\rangle \mu(dx)$ for any $h \in H$. The continuity of F can be seen from the fact that $\langle x, h\rangle \le |x|\,|h|$ and hence

$$|F(h)| \;\le\; \int_H |\langle x, h\rangle| \mu(dx) \;\le\; |h| \int_H |x|\mu(dx).$$

By the Riesz representation theorem there exists $m \in H$ such that $\langle h, m\rangle = \int_H \langle h, x\rangle \mu(dx)$ for all $h \in H$. Furthermore, $\int_H |x|^2 \mu(dx) < \infty$. Define the bilinear form $G : H \times H \to \mathbb{R}$

$$G(h, k) = \int_H \langle h, x - m\rangle \langle k, x - m\rangle \mu(dx).$$

By the Riesz representation theorem there exists a unique bounded operator Q such that

$$\langle Qh, k\rangle = \int_H \langle h, x - m\rangle \langle k, x - m\rangle \mu(dx) \quad \text{for all } h, k \in H. \tag{7}$$

Q is the covariance operator. $Q \in L_1^+(H)$. Then there exists a complete orthonormal sequence (e_k) in H such that $Qe_k = \lambda_k e_k$, $k \in \mathbb{N}$. For any $x \in H$ we set $x_k = \langle x, e_k\rangle$. □

4.1 The Cameron-Martin Space of an $L_1^+(H)$ Operator

If Q is a symmetric, positive definite, trace class operator there exists a unique operator $T \in L^+(H)$ such that $T^2 = Q$. We will denote T by $Q^{1/2}$ and, of course,

$$Q^{1/2}x \;=\; \sum_{k=1}^{\infty} \sqrt{\lambda_k} \langle e_k, x\rangle e_k, \quad x \in H.$$

It is clear that $Q^{1/2}$ is a Hilbert Schmidt (and therefore compact), positive definite linear operator. It is clearly injective since $Q^{1/2}x = 0$ implies $\langle Q^{1/2}x, Q^{1/2}x\rangle = 0$. This in turn implies that $\langle Qx, x\rangle = 0$ and hence $x = 0$ since $\ker Q = \{0\}$.

The image of H under $Q^{1/2}$ is a subspace which we will denote by $Q^{1/2}(H)$. It is called the *Cameron-Martin space* of Q. One can see that $Q^{1/2}(H)$ is dense in H. Indeed, suppose that for some $x_0 \in H$ and for all $y \in H$, $\langle Q^{1/2}y, x_0\rangle = 0$. Take $y = Q^{1/2}x_0$. Then $\langle Q^{1/2}Q^{1/2}x_0, x_0\rangle = 0$ or $\langle Qx_0, x_0\rangle = 0$ and hence $x_0 = 0$.

Proposition 6. *The Cameron-Martin space is a dense subspace of H. Furthermore it holds that $\mu(Q^{1/2}(H)) = 0$.*

For the proof the reader is referred to [3].

For each $z \in Q^{1/2}(H)$ we define a linear mapping $W_z : H \to L^2(H, \mu)$ via $x \mapsto \langle Q^{-1/2}z, x\rangle$. Note that W_z is a centered Gaussian random variable. If $z_1, \dots, z_n \in Q^{1/2}(H)$ consider the linear functionals $W_{z_i}(x) = \langle Q^{-1/2}z_i, x\rangle$, $i = 1, \dots, n$ then the law of the $\mathbb{R}^n$-valued random vector $(W_{z_1}, \dots, W_{z_n})$ is $N_{(\langle z_i, z_j\rangle)_{i,j=1,\dots,n}}$.

Consider $z_1, z_2 \in Q^{1/2}(H)$. Then

$$\begin{aligned}\int_H W_{z_1}(x)W_{z_2}(x)\mu(dx) &= \int_H \langle Q^{-1/2}z_1, x\rangle\langle Q^{-1/2}z_2, x\rangle\mu(dx)\\ &= \langle QQ^{-1/2}z_1, Q^{-1/2}z_2\rangle = \langle z_1, z_2\rangle.\end{aligned}$$

The above shows that the mapping $Q^{1/2}(H) \to L^2(H, \mu)$ defined by $z \mapsto W_z(\cdot)$ is an isometry which, in view of the fact that $Q^{1/2}(H)$ is dense in H, can be uniquely extended into a map from the whole of H into $L^2(H, \mu)$. We will use the same notation for the extension so that we have now defined $W_y(\cdot)$ for any $y \in H$. By an abuse of notation we will still write $W_y(x) = \langle Q^{-1/2}y, x\rangle$ for $x, y \in H$ even though, interpreted literally, the expression $Q^{-1/2}y$ does not make sense when $y \notin Q^{1/2}(H)$. The random variables $W_y(\cdot)$ are known as white noise functions and will play an important rôle in the sequel.

An explicit expression for $W_y(x)$ may be obtained as follows. Given the complete orthonormal basis $\{e_k\}$ of eigenvectors of Q, consider the projection operators $P_n : H \to H$ defined by $P_n x = \sum_{k=1}^n \langle e_k, x\rangle e_k$ for all $x \in H$ and all $n \in \mathbb{N}$. It is easy to see that, for any $y \in H$, $P_n y \in Q^{1/2}(H)$. Furthermore $Q^{1/2}P_n y = \sum_{k=1}^n \lambda_k^{1/2}\langle e_k, y\rangle e_k$. Then we may define

$$\begin{aligned}W_y(x) &= \lim_{n\to\infty}\langle Q^{1/2}P_n y, x\rangle = \lim_{n\to\infty}\sum_{k=1}^n \lambda_k^{1/2}\langle y, e_k\rangle\langle x, e_k\rangle\\ &= \sum_{k=1}^\infty \lambda_k^{1/2}\langle y, e_k\rangle\langle x, e_k\rangle. \end{aligned} \tag{8}$$

4.2 Equivalence and Singularity of Gaussian Measures on a Hilbert Space

Suppose that H is a finite dimensional linear space and μ, ν, are two Gaussian measures, $\mathcal{N}_{Q,a}$, $\mathcal{N}_{R,b}$ with mean vectors a and b respectively and Q, R, the corresponding covariance operators. If both Q and R are of full rank then it is well known and easy to see that the measures μ and ν are equivalent.

The situation is of course different if H is an infinite dimensional, separable Hilbert space. Let us consider the simplest case first where Q and R have the same eigenvectors. Thus we will make the following

Assumption 4.1. *Suppose that $\{e_k\}$, $k = 1, 2, \ldots$ is a complete orthonormal sequence of elements in H such that $Qe_k = \lambda_k e_k$ and $Re_k = \rho_k e_k$, $k = 1, 2, \ldots$, $\lambda_k > 0$, $\rho_k > 0$, and $\sum_{k=1}^{\infty} \lambda_k < \infty$, $\sum_{k=1}^{\infty} \rho_k < \infty$. (Both Q and R are trace class operators.)*

The following theorem is essentially due to Kakutani (see [5, p.32]).

Theorem 7. *Suppose that μ, ν are two centered Gaussian measures $\mathcal{N}_Q$, $\mathcal{N}_R$, on H with covariance operators satisfying Assumption 4.1. Then*

$$\text{If } \sum_{k=1}^{\infty} \left(\frac{\lambda_k - \rho_k}{\lambda_k + \rho_k} \right)^2 < \infty, \ \mu \text{ and } \nu \text{ are equivalent and}$$

$$\frac{d\nu}{d\mu}(x) = \prod_{k=1}^{\infty} \exp\left(-\frac{(\lambda_k - \rho_k)}{2\lambda_k \rho_k} x_k^2 \right) \tag{9}$$

with $x_k := \langle x, e_k \rangle$.

$$\text{If } \sum_{k=1}^{\infty} \left(\frac{\lambda_k - \rho_k}{\lambda_k + \rho_k} \right)^2 = \infty, \quad \mu \text{ and } \nu \text{ are singular.} \tag{10}$$

One particular direct implication is that if $R = \gamma Q$ for some $\gamma > 0$, then clearly R and Q have the same eigenvectors and $\rho_k = \gamma \lambda_k$, $k = 1, 2, \ldots$. It is then clear that the series in (10) diverges and the corresponding Gaussian measures are singular.

Theorem 8. *Assume that μ, ν, are two Gaussian measures on H that are not singular. Then there exists a symmetric Hilbert-Schmidt operator S on H such that*

$$R = Q^{1/2}(I - S)Q^{1/2}.$$

The converse is also true:

Theorem 9. *Assume that there exists a symmetric Hilbert-Schmidt operator such that $R = Q^{1/2}(I - S)Q^{1/2}$. Then μ and ν are equivalent.*

The theorems (8) and (9) together constitute the Hajek-Feldman theorem (see [3, p.55]).

Finally we state the following fundamental theorem which extends the classical Cameron-Martin theorem to general gaussian processes.

Theorem 10 (Cameron-Martin)**.** *Suppose that $\mu = N_{Q,a}$ and $\nu = N_{Q,b}$. Then μ and ν are equivalent if $a - b \in Q^{1/2}(H)$ and singular if not. In the first case*

$$\frac{d\mu}{d\nu}(x) \ = \ \exp\left(-\frac{1}{2} \|Q^{-1/2}(a-b)\|^2 + \langle Q^{-1/2}(a-b), Q^{-1/2}x \rangle \right), \qquad x \in H. \tag{11}$$

4.3 Karhunen-Loève Representation of Gaussian Process

Suppose that $\{X_t; t \in [0,T]\}$ is a centered gaussian process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with covariance function $R(s,t) := \mathbb{E}[X_s X_t] = R(s,t)$. Suppose that H is the Hilbert space $L^2[0,T]$ with inner product defined for square integrable functions, $\phi, \psi \in L^2[0,T]$ via $\langle \phi, \psi \rangle = \int_0^T \phi(s)\psi(s)ds$. Denoting by R both the kernel of the integral operator and the operator itself, since no confusion can arise,

$$Rf(t) := \int_0^T R(s,t)f(s)ds$$

for any $f \in L^2[0,T]$. R is a symmetric bounded, positive definite operator. Let $\{e_k\}$ be a sequence of normalized eigenfunctions associated with R. We then have $Re_k = \lambda_k e_k$, $k = 1, 2, \dots$ with $\lambda_k > 0$ and we shall assume that the kernel of R is trivial. Mercer's theorem implies that

$$R(s,t) = \sum_{k=1}^{\infty} \lambda_k e_k(s) e_k(t).$$

The convergence is uniform and it holds that the trace of the operator R is

$$\operatorname{tr} R := \sum_{k=1}^{\infty} \lambda_k = \int_0^T R(t,t)dt.$$

If $\{Z_k\}$ is an i.i.d. sequence of standard normal random variables then the following representation holds

$$X_t = \sum_{k=1}^{\infty} \lambda_k^{1/2} Z_k e_k(t). \tag{12}$$

This is known as the Karhunen-Loève representation of a Gaussian process.

If the process $\{X_t\}$ is not centered, let $m(t) = \mathbb{E}X_t$, $t \in [0,T]$, denote the mean function of the process which we assume to be continuous on $[0,T]$. Hence $m(t)$ belongs to $L^2[0,T]$ and therefore, since the eigenfunctions of the kernel R form a complete base of $L^2[0,T]$, we have the expansion

$$m(t) = \sum_{k=1}^{\infty} m_k e_k(t)$$

with

$$m_k = \langle m, e_k \rangle = \int_0^T m(t) e_k(t) dt, \qquad k = 1, 2, \dots.$$

Thus, for a non-central gaussian process with mean function $m(t)$,

$$X_t = \sum_{k=1}^{\infty} \left(\lambda_k^{1/2} Z_k + m_k \right) e_k(t). \tag{13}$$

5 Sensitivity Analysis for Gaussian Processes

In section 5.2 we sketched briefly the problem of obtaining sensitivity estimators for cylindrical functionals of Gaussian processes. Since the problem is essentially finite dimensional, the approach is well known and can be carried out entirely using elementary arguments. The situation is different when we examine more general functionals of a Gaussian process.

5.1 Perturbation of the Mean of a Gaussian Process

In the above framework we consider a Gaussian measure on the Hilbert space H with covariance operator R. If (e_k) is a complete orthonormal sequence of eigenvectors or R and $x \in H$, then, if $x_k := \langle x, e_k \rangle$, $\{x_k\}$ are independent normal random variables $\mathcal{N}(0, \lambda_k)$. We will assume that Ker $R = \{0\}$.

Suppose that X_t is a Gaussian process with covariance kernel R. We will assume without loss of generality that its mean is zero. (With trivial modifications, the analysis below will apply to a process with given mean function $m(t)$.) Let $\alpha(t)$ be a continuous function on $[0, T]$ and define the family of processes $X_t^\epsilon := X_t + \epsilon\alpha(t)$. Finally let F be a bounded Borel functional $F : C[0, T] \to \mathbb{R}$.

Theorem 11. *Suppose that $X_t^\epsilon = X_t + \epsilon\alpha(t)$. If $\alpha \in R^{1/2}(H)$, the reproducing kernel Hilbert space associated with R, then*

$$\lim_{\epsilon \to 0} \epsilon^{-1} \left(\mathbb{E}(F(X^\epsilon)) - \mathbb{E}[F(X)] \right) = \mathbb{E}[F(X) H_{mean}]$$

with

$$H_{mean} = \langle R^{-1/2}\alpha, R^{-1/2}x \rangle. \tag{14}$$

Proof. Let μ denote the measure on H corresponding to X_t and μ_ϵ the measure corresponding to X_t^ϵ. Then from Theorem 10 μ and μ_ϵ are absolutely continuous if and only if $\alpha \in R^{1/2}(H)$. In this case, the likelihood ratio is

$$\rho_\epsilon(x) = \exp\left(\epsilon \langle R^{-1/2}\alpha, R^{-1/2}x \rangle - \frac{\epsilon^2}{2} \|R^{-1/2}\alpha\|^2 \right). \tag{15}$$

If $J(\epsilon) := \mathbb{E}[F(X^\epsilon)]$ then

$$J'(0) = \lim_{\epsilon \to 0} \epsilon^{-1} \mathbb{E}[F(X)(\rho_\epsilon - 1)]. \tag{16}$$

In order to prove the result it suffices to interchange the expectation and the limiting operation, a procedure that here requires care because of the fact that, as we have seen $\langle R^{-1/2}\alpha, R^{-1/2}x \rangle$ has been defined via an isometry. For this reason we will establish the result in two steps.

Step 1. Suppose first that α belongs in $R(H)$. Then

$$\langle R^{-1/2}\alpha, R^{-1/2}x \rangle = \langle R^{-1}\alpha, x \rangle$$

where the right hand side of the above equation has "literal meaning" for all $x \in H$. Thus from (15), using the inequality $|e^\theta - 1| \leq |\theta| e^{|\theta|}$ which holds for all $\theta \in \mathbb{R}$ and other similar elementary arguments we have that for some $C > 0$ and all $\epsilon \in (0, \epsilon_0)$

$$\epsilon^{-1} |\rho_\epsilon(x) - 1| \leq C \left| \langle R^{-1}\alpha, x \rangle \right| e^{\epsilon |\langle R^{-1}\alpha, x \rangle|} + \frac{\epsilon}{2} \|R^{-1/2}\alpha\|^2 \quad \text{for all } x \in H. \tag{17}$$

(The whole point of performing step 1 is to be able to claim in fact that the above inequality holds for all $x \in H$ and not just for $x \in R^{1/2}(H)$.) Since

$$\begin{aligned} \int_H e^{iu\langle R^{-1}\alpha, x \rangle} \mu(dx) &= e^{-\frac{u^2}{2} \langle R(R^{-1})\alpha, R^{-1}\alpha \rangle} \\ &= e^{-\frac{u^2}{2} \|R^{-1/2}\alpha\|^2} \quad \text{for all } u \in \mathbb{R} \end{aligned}$$

we conclude that $\langle R^{-1}\alpha, x\rangle$ is a centered normal random variable, say ξ, with variance $\|R^{-1/2}\alpha\|^2$ and hence $\int_H \left|\langle R^{-1}\alpha, x\rangle\right| e^{\epsilon|\langle R^{-1}\alpha,x\rangle|}\mu(dx) = \mathbb{E}[|\xi|e^{\epsilon\xi}] < \infty$ by an elementary argument. This implies that the random variable on the right hand side of (17) has finite expectation. Thus we can use the fact that

$$\lim_{\epsilon\to 0}\epsilon^{-1}[\rho_\epsilon(x)-1] \;=\; \langle R^{-1}\alpha, x\rangle \;=\; \langle R^{-1/2}\alpha, R^{-1/2}x\rangle \tag{18}$$

together with the dominated convergence theorem to establish (14) when $a \in R(H)$.

<u>*Step 2.*</u> To establish the result for $\alpha \in R^{1/2}(H)$, the Cameron-Martin space of H, we use the fact that $R(H)$, is dense in H and therefore in $R^{1/2}(H)$ as well. Hence we can pick a sequence of elements of $R(H)$, say $\{\alpha_n\}$ such that $\|\alpha_n - \alpha\| \to 0$ as $n \to \infty$. In fact we will use the family of projection operators $\{P_n\}$ with respect to the complete orthonormal basis $\{e_n\}$ defined in §5.4.1 to obtain the sequence $\alpha_n := P_n a$, $n = 1, 2, \ldots$ which converges in H to α. Obviously $\alpha_n \in P_n(H) \subset R(H)$ for all $n \in \mathbb{N}$. Also note that

$$\|R^{-1/2}\alpha_n\|^2 = \sum_{k=1}^{n}\frac{\langle \alpha, e_k\rangle^2}{\lambda_k}$$

and thus the sequence $\{\|R^{-1/2}\alpha_n\|\}$ is a non-negative *increasing* sequence which converges to $\|R^{-1/2}\alpha\|$.

Set

$$\begin{aligned}\rho_\epsilon^n(x) &= \exp\left(\epsilon\langle R^{-1/2}\alpha_n, R^{-1/2}x\rangle - \frac{\epsilon^2}{2}\|R^{-1/2}\alpha_n\|^2\right)\\ &= \exp\left(\epsilon\langle R^{-1}\alpha_n, x\rangle - \frac{\epsilon^2}{2}\|R^{-1/2}\alpha_n\|^2\right)\end{aligned}$$

and

$$\phi_n(\epsilon) := \epsilon^{-1}\mathbb{E}[F(X)(\rho_\epsilon^n - 1)]. \tag{19}$$

Then, since $|F| \le K$ for some $K > 0$

$$|\phi_n(\epsilon) - \phi_m(\epsilon)| = \epsilon^{-1}|\mathbb{E}[F(X)(\rho_\epsilon^n - \rho_\epsilon^m)]| \le K\epsilon^{-1}\mathbb{E}|\rho_\epsilon^n - \rho_\epsilon^m|. \tag{20}$$

Adding and subtracting a term we can write

$$\begin{aligned}\rho_\epsilon^n(x) - \rho_\epsilon^m(x) &= e^{\epsilon\langle R^{-1}\alpha_n,x\rangle - \frac{\epsilon^2}{2}\|R^{-1/2}\alpha_n\|^2} - e^{\epsilon\langle R^{-1}\alpha_m,x\rangle - \frac{\epsilon^2}{2}\|R^{-1/2}\alpha_n\|^2}\\ &\quad + e^{\epsilon\langle R^{-1}\alpha_m,x\rangle - \frac{\epsilon^2}{2}\|R^{-1/2}\alpha_n\|^2} - e^{\epsilon\langle R^{-1}\alpha_m,x\rangle - \frac{\epsilon^2}{2}\|R^{-1/2}\alpha_m\|^2}\end{aligned}$$

whence we obtain

$$\begin{aligned}\rho_\epsilon^n(x) - \rho_\epsilon^m(x) &= \left(e^{\epsilon\langle R^{-1}\alpha_n,x\rangle} - e^{\epsilon\langle R^{-1}\alpha_m,x\rangle}\right) e^{-\frac{\epsilon^2}{2}\|R^{-1/2}\alpha_n\|^2}\\ &\quad + e^{\epsilon\langle R^{-1}\alpha_m,x\rangle}\left(e^{-\frac{\epsilon^2}{2}\|R^{-1/2}\alpha_n\|^2} - e^{-\frac{\epsilon^2}{2}\|R^{-1/2}\alpha_m\|^2}\right)\end{aligned}$$

Supposing that $m < n$ and taking into account the fact that $\|R^{-1/2}\alpha_n\|$ is an increasing sequence we obtain

$$\begin{aligned}\epsilon^{-1}|\rho_\epsilon^n - \rho_\epsilon^m| &\le \frac{\epsilon}{2}\left(e^{\epsilon\langle R^{-1}\alpha_n,x\rangle} + e^{\epsilon\langle R^{-1}\alpha_m,x\rangle}\right)\left(\|R^{-1/2}\alpha_n\|^2 - \|R^{-1/2}\alpha_m\|^2\right)\\ &\quad + |\langle R^{-1}(\alpha_n - \alpha_m), x\rangle| e^{\epsilon\langle R^{-1}\alpha_m,x\rangle}\end{aligned}$$

Thus, using the Cauchy-Schwarz inequality,

$$\begin{aligned}\epsilon^{-1}\mathbb{E}|\rho_\epsilon^n - \rho_\epsilon^m| &\leq \frac{\epsilon}{2}\int_H \left(e^{\epsilon\langle R^{-1}\alpha_n, x\rangle} + e^{\epsilon\langle R^{-1}\alpha_m, x\rangle}\right)\mu(dx)\left(\|R^{-1/2}\alpha_n\|^2 - \|R^{-1/2}\alpha_m\|^2\right) \\ &\quad + \left(\int_H \langle R^{-1}(\alpha_n - \alpha_m), x\rangle^2 \mu(dx)\right)^{1/2}\left(\int_H e^{2\epsilon\langle R^{-1}\alpha_m, x\rangle}\mu(dx)\right)^{1/2}\end{aligned}$$

Thus, from (6) and (7)

$$\begin{aligned}\int_H \langle R^{-1}(\alpha_n - \alpha_m), x\rangle^2 \mu(dx) &= \langle RR^{-1}(\alpha_n - \alpha_m), R^{-1}(\alpha_n - \alpha_m)\rangle \\ &= \langle R^{-1/2}(\alpha_n - \alpha_m), R^{-1/2}(\alpha_n - \alpha_m)\rangle = \|R^{-1/2}(\alpha_n - \alpha_m)\|^2,\end{aligned}$$

$$\int_H e^{\epsilon\langle R^{-1}\alpha_n, x\rangle}\mu(dx) = e^{\frac{1}{2}\epsilon^2\|R^{-1/2}\alpha_n\|^2} \leq e^{\frac{1}{2}\epsilon^2\|R^{-1/2}\alpha\|^2}.$$

Therefore we have the bound

$$\begin{aligned}\epsilon^{-1}\mathbb{E}|\rho_\epsilon^n - \rho_\epsilon^m| &\leq \epsilon\, e^{\frac{1}{2}\epsilon^2\|R^{-1/2}\alpha\|^2}\left(\|R^{-1/2}\alpha_n\|^2 - \|R^{-1/2}\alpha_m\|^2\right) \\ &\quad + \|R^{-1/2}(\alpha_n - \alpha_m)\|^2 e^{2\epsilon^2\|R^{-1/2}\alpha\|^2}. \qquad (21)\end{aligned}$$

As $m, n \to \infty$ both $\|R^{-1/2}\alpha_n\|^2 - \|R^{-1/2}\alpha_m\|^2$ and $\|R^{-1/2}(\alpha_n - \alpha_m)\|^2$ converge to 0 and therefore we can see that for each $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that for all $\epsilon \in [-\epsilon_0, \epsilon_0]$ and all $m, n \geq n_0$, $\epsilon^{-1}\mathbb{E}|\rho_\epsilon^n - \rho_\epsilon^m| < \delta$. From (20) we conclude that $\phi_n(\epsilon)$ converges uniformly in $[-\epsilon_0, \epsilon_0]$ and hence

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\phi_n(\epsilon) = \lim_{n\to\infty}\lim_{\epsilon\to 0}\phi_n(\epsilon)\,. \qquad (22)$$

Since $\langle R^{-1/2}\alpha_n, R^{-1/2}x\rangle \to \langle R^{-1/2}\alpha, R^{-1/2}x\rangle$ and $\|R^{-1/2}\alpha_n\| \to \|R^{-1/2}\alpha\| \to \|R^{-1/2}\alpha\|$ as $n \to \infty$ the left hand side of (22) is equal to $J'(0)$ whereas the right hand side is equal to $\mathbb{E}[F\langle R^{-1/2}\alpha, R^{-1/2}x]$ thus completing the proof. □

It is interesting to compare (14) with the corresponding finite-dimensional result in (4) provided the latter is written in the form $\partial_\theta\mu^\top R_n^{-1/2}R_n^{-1/2}x$. Of course the principal difference is that, while for a full rank matrix R_n in the finite dimensional case the range of $R^{1/2}$ is the whole n-dimensional Euclidean space, in H it is the much smaller Cameron-Martin space.

To cast (14) in a form that is useful in simulation we begin with the representation of the perturbing function $\alpha(t)$ in terms of the eigenfunctions of R to obtain

$$\alpha(t) = \sum_{k=1}^{\infty}\alpha_k e_k(t)$$

where $\alpha_k := \int_0^T a(t)e_k(t)dt$, $k = 1, 2, \ldots$. Hence

$$R^{-1/2}\alpha = \sum_{k=1}^{\infty}\alpha_k\lambda_k^{-1/2}e_k(t).$$

The above expression is well defined because $\alpha \in R^{1/2}(H)$. Also,

$$X_t \;:=\; \sum_{k=1}^{\infty} x_k e_k(t) \quad \text{with} \quad x_k \;=\; \int_0^T X_t e_k(t)dt.$$

$$R^{-1/2}x = \sum_{k=1}^{\infty} \lambda_k^{-1/2} x_k e_k(t) = \sum_{k=1}^{\infty} Z_k e_k(t).$$

This does not converge in any sense unless x belongs to $R^{1/2}(H)$. Hence

$$\langle R^{-1/2}\alpha, R^{-1/2}x\rangle \;=\; \sum_{k=1}^{\infty} \alpha_k \lambda_k^{-1/2} Z_k.$$

This random variable has finite variance, provided that $\alpha \in R^{1/2}(H)$ since

$$\mathsf{Var}(\langle R^{-1/2}\alpha, R^{-1/2}x\rangle) \;=\; \sum_{k=1}^{\infty} \frac{\alpha_k^2}{\lambda_k} \;<\; \infty.$$

Implementation Issues. While the expression (14) provides in theory the appropriate weight H the implementation in practice provides numerical challenges since evaluating the expression

$$\langle R^{-1/2}\alpha, R^{-1/2}x\rangle = \sum_{k=1}^{\infty} \frac{\langle a, e_k\rangle\langle x, e_k\rangle}{\lambda_k}$$

in terms of the sample paths of the process would require the numerical evaluation of the expression

$$\sum_{k=1}^{\infty} \frac{1}{\lambda_k} \int_0^T \alpha(t)e_k(t)dt \int_0^T X_t e_k(t)dt.$$

In most situations this may not be practicable and the errors introduced in the truncation of the infinite series may be significant. The following alternative approach provides a solution which has better numerical properties.

Proposition 12. *Given $a \in R(H)$ let $\beta \in H$ be the unique solution of the Fredholm equation of the first type*

$$\alpha(t) \;=\; \int_0^T R(s,t)\beta(s)ds. \tag{23}$$

Then

$$H_{mean} \;=\; \int_0^T \beta(t)X_t dt. \tag{24}$$

Proof. Note that (23) implies that

$$\alpha_k = \langle \alpha, e_k\rangle = \langle R\beta, e_k\rangle = \langle \beta, Re_k\rangle = \lambda_k \beta_k$$

and hence

$$\begin{aligned} H_{\text{mean}} \;&=\; \langle R^{-1/2}\alpha, R^{-1/2}x\rangle \;=\; \sum_{k=1}^{\infty} \frac{\alpha_k x_k}{\lambda_k} \;=\; \sum_{k=1}^{\infty} \beta_k x_k \;=\; \sum_{k=1}^{\infty} \langle \beta, e_k\rangle\langle x, e_k\rangle \\ &=\; \langle \beta, x\rangle \;=\; \int_0^T X_t \beta(t)dt. \end{aligned}$$

$\square$

5.2 Perturbation of the Covariance of a Gaussian Process

In this section we will consider perturbations of the covariance operator R. Let $R_\epsilon := R+\epsilon V$ where V is also an $L_1^+(H)$ operator. Denote by μ and μ_ϵ respectively the Gaussian measures on H induced by R and R_ϵ. Then it follows from Theorems 8 and 9 that μ and μ_ϵ are equivalent if and only if $V = R^{1/2}SR^{1/2}$ where S is a Hilbert-Schmidt operator. We will make the stronger assumption that S is in *trace class*. The following theorem is a direct consequence of the Feldman-Hajek theorem. (see also [4, p.26]).

Theorem 13. *If $R, Q \in L_1^+(H)$ and $Q = R^{1/2}(I+S)R^{1/2}$, $S \in L_1(H)$ (i.e. if S is trace class) and $\|S\| < 1$ then, denoting by μ the measure $\mathcal{N}_R$ and by ν the measure $\mathcal{N}_Q$ on H,*

$$\frac{d\nu}{d\mu}(x) = \frac{1}{(\det(I+S))^{1/2}} \exp\left(\frac{1}{2}\langle S(I+S)^{-1}R^{-1/2}x, R^{-1/2}x\rangle\right), \quad x \in H. \tag{25}$$

In the above theorem, if $S \in L_1^+(H)$, denoting by γ_k, $k = 1, 2, \dots$ the eigenvalues of S then the determinant of the operator $I + S$ is *defined* as

$$\det(I+S) = \prod_{k\in K}(1+\gamma_k). \tag{26}$$

If the set of eigenvalues is finite then this is a finite product. Otherwise, if K is a countable set of indices, the infinite product $\det(I+S) = \prod_{k=1}^{\infty}(1+\gamma_k)$ converges because the series $\sum_{k=1}^{\infty}\gamma_k$ converges due to the assumption that S is trace class.

The following theorem of Varberg [23] will be useful in the sequel.

Theorem 14 (Varberg). *Let $\{Z_i\}$, $i = 1, 2, \dots$ be a sequence of independent, identically distributed random variables with $\mathbb{E}Z_i = 0$ and $\mathsf{Var}(Z_i) = 1$. Suppose that $\{a_{ij}\}$, $i, j = 1, \dots,$ a double array of real numbers and set $S_N := \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}Z_iZ_j$. Then, if $\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} a_{jk}^2 < \infty$ and $\sum_{k=1}^{\infty}|a_{kk}| < \infty$ the sequence S_N converges almost surely to a finite random variable.*

Suppose again that $\{X_t, t \in [0,T]\}$ is a gaussian process which, without loss of generality, we will assume to be centered. Suppose that $R(s,t) = \mathbb{E}[X_sX_t]$ is the covariance function of the process and $F : C[0,T] \to \mathbb{R}$ a bounded Borel functional. We would like to estimate the sensitivity of the performance criterion $\mathbb{E}[F(X)]$ when the covariance function is perturbed by a non-negative, symmetric kernel V. Thus we will consider the family of processes X_t^ϵ, $t \in [0,T]$, with covariance $R^\epsilon = R+\epsilon V$. We will suppose that the perturbing kernel V is such that the measure μ of the original process and μ_ϵ of the perturbed process are equivalent. In view of Theorem 13 we will assume in fact that $V = R^{1/2}SR^{1/2}$ where S is in fact a symmetric, non-negative *trace class* operator.

Denote the eigenvectors of S by χ_k, with $S\chi_k = \gamma_k\chi_k$ $k = 1, 2, \dots$ and assume without loss of generality that $\gamma_1 \geq \gamma_2 \geq \dots \geq 0$. Starting with the elementary inequality $\log(1+x) \geq \frac{x}{1+x}$ which holds for all $x > -1$ we note that

$$\log\prod_{k=1}^{n}(1+x\gamma_k)^{-1/2} \leq -\frac{1}{2}\sum_{k=1}^{n}\frac{x\gamma_k}{1+x\gamma_k} \leq -Cx\sum_{k=1}^{n}\gamma_k$$

for some $C > 0$ and all $x \geq B > -\frac{1}{\gamma_1}$. Hence $\left(\prod_{k=1}^{n}(1+x\gamma_k)\right)^{-1/2} \leq e^{-Cx\sum_{k=1}^{n}\gamma_k} \leq e^{-Cx\gamma_1}$ and hence

$$(\det(I-\epsilon S)) \leq e^{-C\gamma_1\epsilon} \quad \text{for } \epsilon \text{ sufficiently small.}$$

Theorem 15. *Under the above assumptions*

$$J' \;:=\; \frac{d}{d\epsilon}\mathbb{E}_{\mu_\epsilon}[F(X_\epsilon)]\bigg|_{\epsilon=0} \;=\; \mathbb{E}_\mu[F(X)H_{cov}] \tag{27}$$

with

$$H_{cov} \;=\; \frac{1}{2}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\langle Se_k, e_l\rangle\langle e_k, R^{-1/2}x\rangle\langle e_l, R^{-1/2}x\rangle \;-\; \frac{1}{2}Tr(S)\,. \tag{28}$$

Proof. Set

$$\rho_\epsilon(x) := \frac{d\mu_\epsilon}{d\mu}(x) = \frac{1}{(\det(I+\epsilon S))^{1/2}}\,\exp\left(\frac{1}{2}\langle \epsilon S(I+\epsilon S)^{-1}R^{-1/2}x, R^{-1/2}x\rangle\right)$$

Set for convenience $T_\epsilon := S(I+\epsilon S)^{-1}$ and note that T_ϵ is trace class with eigenvectors χ_k and eigenvalues $\frac{\gamma_k}{1+\epsilon\gamma_k}$. Set $\langle e_k, x\rangle = \lambda_k Z_k$. $\{Z_k\}$ is a sequence of independent standard normal random variables. Noting that $R^{-1/2}x = \sum_{k=1}^{\infty}\frac{\langle x, e_k\rangle}{\lambda_k^{1/2}}e_k$

$$\begin{aligned}\langle T_\epsilon R^{-1/2}x, R^{-1/2}x\rangle \;&=\; \left\langle T_\epsilon\sum_{k=1}^{\infty}\frac{\langle x, e_k\rangle}{\lambda_k^{1/2}}e_k, \sum_{l=1}^{\infty}\frac{\langle x, e_l\rangle}{\lambda_l^{1/2}}e_l\right\rangle\\ &=\; \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\langle T_\epsilon e_k, e_l\rangle\frac{\langle x, e_k\rangle\langle x, e_l\rangle}{\lambda_k^{1/2}\lambda_l^{1/2}}.\end{aligned}$$

This is equal to

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\langle T_\epsilon e_k, e_l\rangle Z_k Z_l.$$

The convergence of this series is guaranteed from Varberg's theorem because T_ϵ is a trace class operator. Further, note that $\langle T_\epsilon e_k, e_l\rangle = \langle T_\epsilon\sum_{i=1}^{\infty}\langle e_k, \chi_i\rangle\chi_i, \sum_{j=1}^{\infty}\langle e_l, \chi_j\rangle e_j$ whence we obtain

$$\langle T_\epsilon e_k, e_l\rangle = \sum_{i=1}^{\infty}\frac{\gamma_i}{1+\epsilon\gamma_i}\langle e_k, \chi_i\rangle\langle e_l, \chi_i\rangle$$

Hence

$$\begin{aligned}\Theta_\epsilon \;:=\; \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\langle T_\epsilon e_k, e_l\rangle Z_k Z_l \;&=\; \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}Z_k Z_l\sum_{i=1}^{\infty}\frac{\gamma_i}{1+\epsilon\gamma_i}\langle e_k, \chi_i\rangle\langle e_l, \chi_i\rangle\\ &=\; \sum_{i=1}^{\infty}\frac{\gamma_i}{1+\epsilon\gamma_i}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}Z_k Z_l\langle e_k, \chi_i\rangle\langle e_l, \chi_i\rangle\\ &=\; \sum_{i=1}^{\infty}\frac{\gamma_i}{1+\epsilon\gamma_i}\left(\sum_{k=1}^{\infty}Z_k\langle e_k, \chi_i\rangle\right)^2.\end{aligned} \tag{29}$$

Set $Y_i := \sum_{k=1}^{\infty}Z_k\langle e_k, \chi_i\rangle$, $i = 1, 2, \ldots$ and note that the Y_i's are standard normal random variable since by Parseval's identity $\sum_{k=1}^{\infty}\langle e_k, \chi_i\rangle^2 = \|\chi_k\|^2 = 1$. Furthermore they are independent since

$$\begin{aligned}\mathbb{E}[Y_iY_j] \;&=\; \mathbb{E}\left[\sum_{k=1}^{\infty}Z_k\langle e_k, \chi_i\rangle\sum_{l=1}^{\infty}Z_l\langle e_l, \chi_j\rangle\right] \;=\; \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\mathbb{E}[Z_kZ_l]\langle e_k, \chi_i\rangle\langle e_l, \chi_j\rangle\\ &=\; \sum_{k=1}^{\infty}\langle e_k, \chi_i\rangle\langle e_k, \chi_j\rangle \;=\; \langle \chi_i, \chi_j\rangle \;=\; \delta_{ij}.\end{aligned}$$

Hence, since $\gamma_i \geq 0$ for all i we have with probability 1 that

$$0 \;<\; \Theta_\epsilon \;\leq\; \Theta_0 \;:=\; \sum_{i=1}^{\infty} \gamma_i \left(\sum_{k=1}^{\infty} Z_k \langle e_k, \chi_i \rangle \right)^2 \;<\infty \qquad \text{w.p. 1 for } 0 \leq \epsilon \leq \epsilon_0.$$

Also, obviously the determinant $\prod_{i=1}^{\infty}(1+\epsilon\gamma_i)$ is an increasing function of ϵ (since all the eigenvalues γ_i are non-negative).

$$\begin{aligned} 0 < \frac{1}{\epsilon}(\rho_\epsilon - 1) \;\; &\leq \;\; \frac{1}{\epsilon}\left(\frac{e^{\frac{\epsilon}{2}\Theta_0}}{(\det(I+S_\epsilon))^{1/2}} - 1 \right) \leq \frac{1}{\epsilon}\left(e^{\frac{\epsilon}{2}\Theta_0} - 1 \right) \\ &\leq \;\; e^{\frac{\epsilon}{2}\Theta_0}\frac{1}{2}\Theta_0. \end{aligned} \tag{30}$$

From a standard computation,

$$\mathbb{E}[e^{t\Theta_0}] \;=\; \mathbb{E}\prod_{i=1}^{\infty} e^{t\gamma_i Y_i} \;=\; \prod_{i=1}^{\infty} \frac{1}{(1-2t\gamma_i)^{1/2}} \quad \text{for } t < 1/(2\gamma_1).$$

Thus, for $\epsilon < 1/\gamma_1$, $\mathbb{E}[e^{\frac{\epsilon}{2}\Theta_0}\frac{1}{2}\Theta_0]\infty$ and we can use the Dominated Convergence Theorem to obtain

$$J'(0) \;=\; \lim_{\epsilon\to 0} \mathbb{E}[F(X)\epsilon^{-1}(\rho_\epsilon - 1)] \;=\; \mathbb{E}[F(X) \lim \epsilon \to 0 \epsilon^{-1}(\rho_\epsilon - 1)].$$

Since

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(\det(I+\epsilon S)^{-1/2} - 1 \right) \;=\; -\frac{1}{2}\sum_{k=1}^{\infty}\gamma_k \;=\; -\frac{1}{2}\mathrm{Tr}(S) \tag{31}$$

and

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(\exp\left(\frac{\epsilon}{2} \langle S(I+\epsilon S)^{-1} R^{-1/2}x, R^{-1/2}x \rangle \right) - 1 \right) \;=\; \frac{1}{2}\langle SR^{-1/2}x, R^{-1/2}x \rangle \tag{32}$$

it follows that

$$H_{\text{cov}} \;:=\; \lim_{\epsilon\to 0} \epsilon^{-1}(\rho_\epsilon - 1) \;=\; \frac{1}{2}\langle SR^{-1/2}x, R^{-1/2}x \rangle - \frac{1}{2}\mathrm{Tr}(S). \tag{33}$$

The inner product on the right hand side of (32) is

$$\begin{aligned} \langle SR^{-1/2}x, R^{-1/2}x \rangle \;\; &= \;\; \left\langle S \sum_{k=1}^{\infty} \lambda_k^{-1/2} \langle e_k, x \rangle \, e_k \;,\; \sum_{l=1}^{\infty} \lambda_k^{-1/2} \langle e_l, x \rangle \, e_l \right\rangle \\ &= \;\; \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} (\lambda_k\lambda_l)^{-1/2} \langle Se_k, e_l \rangle \, \langle e_k, x \rangle \, \langle e_l, x \rangle \\ &= \;\; \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \langle Se_k, e_l \rangle \, \langle R^{-1/2}e_k, x \rangle \, \langle R^{-1/2}e_l, x \rangle. \end{aligned} \tag{34}$$

Thus (33) and (34) establish the proof of the theorem. □

Again it is interesting to compare the expression for H_{cov} in (28) and, better yet, in the alternative form obtained in (33), with the finite-dimensional version of (5). One would have to take of course the mean $\mu_n = 0$ in (5) since we have assumed here the process to be centered. Also the counterpart of $\partial_\theta R_n$ in (5) is $V := R^{1/2}SR^{1/2}$ here. Finally note that the trace term in (5) can be written as $\mathrm{tr}(R_n^{-1/2}R_n^{-1/2}\partial_\theta R_n) = \mathrm{tr}(R_n^{-1/2}(\partial_\theta R_n)R_n^{-1/2})$ in view of the well known identity $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ (where A, B are any two matrices for which the products AB and BA are defined).

6 Conclusion and Further Work

Hilbert space techniques have been extremely fruitful in the analysis of Gaussian processes for over five decades. In recent years there has been renewed interest in reproducing kernel Hilbert space techniques in relation to functional data analysis. Such techniques are also very useful in sensitivity analysis as the preliminary results obtained in this section indicate. We intend to further clarify the theoretical questions arising in this respect and in applying these techniques to the simulation of Gaussian processes in connection with applications.

References

[1] Adler, R.J. (1990). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes.* Institute of Mathematical Statistics, Lecture Notes Vol. 12.

[2] Asmussen, S. and M. Taksar (1997). Controlled diffusion models for optimal divident pay–out. *Insurance: Mathematics and Economics,* 20, 1–15.

[3] Da Prato, G. and J. Zabczyk (2014). *Stochastic Equations in Infinite Dimensions*. 2nd ed. Cambridge University Press.

[4] Da Prato, G. and J. Zabczyk (2004). *Second Order Partial Differential Equations in Hilbert Spaces.* Cambridge University Press.

[5] Da Prato, G. (2006). *An Introduction to Infinite-Dimensional Analysis.* Springer Verlag.

[6] Fournié, E., J.M. Lasry, J. Lebuchoux, P.L. Lions, and N. Touzi (1999). Applications of Malliavin calculus to Monte Carlo methods in finance, *Finance and Stochastics*, 3, 391–412.

[7] Fournié, E., J.M. Lasry, J. Lebuchoux, and P.L. Lions (2001). Applications of Malliavin Calculus to Monte-Carlo Methods in Finance. II, *Finance and Stochastics*, 5, 201–236.

[8] Friedman, Avner (1975). *Stochastic Differential Equations and Applications, Vol. 1* Academic Press, 1975.

[9] Glasserman P. (2004). *Monte Carlo Methods in Financial Engineering.* Springer–Verlag, New York.

[10] Gobet, E. and R. Munos (2005). Sensitivity analysis using Itô-Malliavin calculus and martingales, and application to stochastic optimal control. *SIAM J. of Control and Optimization*, 43, 1676–1713.

[11] Harrison, M. (1977). Ruin problems with compounding assets. *Stochastic Process. Appl,* 5, 67–79.

[12] Hsing, T. and R. Eubank (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*, John Wiley.

[13] Iglehart, D.L. (1969). Diffusion approximations in collective risk theory. *J. Appl. Probab.* 6, 285–292.

[14] Nualart, D. (2006). *Malliavin Calculus and Related Topics*, 2nd ed.; Springer Verlag, New York.

[15] Norberg, R. (1999). Ruin problems with assets and liabilities of diffusion type, *Stoch. Processes Appl.* 81, 255–269.

[16] Privault, N. and X. Wei (2004). A Malliavin calculus approach to sensitivity analysis in insurance, *Insurance: Mathematics and Economics,* **35**, 679–690.

[17] Protter P.E. (2004). *Stochastic Integration and Differential Equations*, 2nd ed.; Springer Verlag, New York.

[18] Reiman, M. and A. Weiss (1989). Sensitivity analysis for simulations via likelihood ratios. *Operations Research*, 37, 830–844.

[19] Riesz, F. and B. Sz.-Nagy (1990). *Functional Analysis*. Dover.

[20] Schott, J.R. (2005). *Matrix Analysis for Statistics*, Wiley, New York.

[21] Shepp, L.A. (1966). Radon-Nikodym Derivatives of Gaussian Measures, *Ann. Math. Stat.* 37, 2, 321-354.

[22] Varberg, D.E. (1966). Convergence of quadratic forms in independent random variables. *The Annals of Mathematical Statistics*, 37, 567–576.

[23] Varberg, D.E. (1968). Almost sure convergence of quadratic forms in independent random variables. *The Annals of Mathematical Statistics*, 39, 5, 1502–1506.

[24] Young N. (1988). *An Introduction to Hilbert Space.* Cambridge University Press.