

MINIMAX RESULTS FOR ESTIMATING INTEGRALS OF ANALYTIC PROCESSES

KARIM BENHENNI AND JACQUES ISTAS

ABSTRACT. The problem of predicting integrals of stochastic processes is considered. Linear estimators have been constructed by means of samples at N discrete times for processes having a fixed Hölderian regularity $s > 0$ in quadratic mean. It is known that the rate of convergence of the mean squared error is of order $N^{-(2s+1)}$.

In the class of analytic processes \mathbf{H}^p , $p \geq 1$, we show that among all estimators, the linear ones are optimal. Moreover, using optimal coefficient estimators derived through the inversion of the covariance matrix, the corresponding maximal error has lower and upper bounds with exponential rates. Optimal simple nonparametric estimators with optimal sampling designs are constructed in \mathbf{H}^2 and \mathbf{H}^∞ and have also bounds with exponential rates.

1. INTRODUCTION

In addressing problems involving stochastic processes such as estimation of various quantities from the sample path, parameter inference and control theory, one frequently has access to observations at a finite number of points rather than over an entire observation interval. The following questions then arise: what are the optimal estimators of integrals of random processes and what are the rates of convergence of the mean squared error of these estimators as the number of observations gets large?

This problem was widely studied by Sacks and Ylvisaker (1966, 1968, 1970, 1971), Cambanis (1985) for processes having 0 or 1 quadratic mean derivatives; and by Eubank, Smith and Smith (1982) for some class of processes including the K th order iterated integrals of Brownian motion. Benhenni and Cambanis (1992a, 1992b) considered processes having K quadratic mean derivatives, K a nonnegative integer, and showed that the rate of convergence for the mean squared error in the approximation of random integrals by linear estimators based on the Euler MacLaurin formula for regular sampling design of size N , is of order $N^{-(2K+2)}$.

Let $\bar{L}_N(X)$ the Euler MacLaurin estimator (see Benhenni and Cambanis (1992a)) of the random integral of the process X having a covariance function $R(t, s) = \mathbb{E}X(t)X(s)$:

$$I(X) = \int_a^b \phi(t)X(t)dt,$$

URL address of the journal: <http://www.emath.fr/ps/>

Received by the journal April 30, 1996. Revised November 18, 1997. Accepted for publication June 11, 1998.

© Société de Mathématiques Appliquées et Industrielles. Typeset by L^AT_EX.

where ϕ is a known nonrandom function in $\mathbb{L}^2([a, b])$.

The estimator $\bar{L}_N(X)$ is said to be asymptotically optimal since it has the same asymptotic performance as the optimal linear estimator (see Benhenni and Cambanis (1992a)):

$$\begin{aligned} & \lim_{N \rightarrow \infty} N^{2K+2} \mathbb{E}(I(X) - \bar{L}_N(X))^2 \\ &= \lim_{N \rightarrow \infty} N^{2K+2} \inf_{L_N} \mathbb{E}(I(X) - L_N(X))^2 \\ &= \frac{B_{2K+2}}{(2K+2)!} \int_a^b \frac{\alpha_K(t) \phi^2(t)}{h^{2K+2}(t)} dt, \end{aligned} \tag{1.1}$$

where B_m is the m th Bernoulli number (e.g. Abramowitz and Stegun (1965)), h is the density that generates the sampling design $\{t_1 < t_2 < \dots < t_N\}$, $\alpha_K(t)$ defined by

$$\alpha_K(t) = \lim_{s \uparrow t} \frac{\partial^{2K+2}}{\partial t^K \partial s^{K+1}} R(t, s) - \lim_{s \downarrow t} \frac{\partial^{2K+2}}{\partial t^K \partial s^{K+1}} R(t, s)$$

is assumed finite and positive on $[a, b]$, and the infimum \inf_{L_N} is taken among all linear estimators.

Istas and Laredo (1997), Stein (1995) extended the results to processes satisfying a Hölder condition, that is, there exists a real number s such that for every $t \in [a, b]$, $\mathbb{E}(X(t+h) - X(t))^2 |h|^{-2s}$ tends to a finite nonvanishing constant as $h \rightarrow 0$ and showed that the rate for the mean squared error is of order $N^{-(2s+1)}$. In particular, the asymptotic result in (1.1) is obtained by taking $s = K + 1/2$, see Stein (1995) for details.

What happens when the process $X(t)$ is analytic? The following questions are still open. What is the exact rate of convergence for the mean squared error? In fact it is faster than any power of the sample size N since in this case, the asymptotic constant in (1.1) is equal to 0 where $\alpha_K(t) = 0$ for all K . Do nonlinear estimators provide a significant improvement on the performance of linear estimators?

We give in this paper some answers to these questions. First we define the class of analytic processes \mathbf{H}^p , $1 \leq p \leq \infty$, from the Hardy space H^p of analytic deterministic functions (e.g. Rudin (1966)). Denote by $L_N(X)$ a linear estimator of $I(X)$ constructed from the observations $X(t_i)$, $i = 1, \dots, N$, at N sampling points (t_1, \dots, t_N) taken from the interval $] -1, 1[$. The maximal mean squared error in the class \mathbf{H}^p , $1 \leq p \leq \infty$, is given by $E_{L_N}(\mathbf{H}^p) = \sup_{X \in \mathbf{H}^p} \mathbb{E}(I(X) - L_N(X))^2$. We study then the minimax risk error $d(\mathbf{H}^p) = \inf_{L_N} E_{L_N}(\mathbf{H}^p)$ (cf. Ibragimov and Has'minskii (1981), Ch.1). We give in Theorem 3.1 a lower bound for $d(\mathbf{H}^p)$. We show in Theorem 3.2 that the maximal error corresponding to the estimator derived from the similar deterministic recovery problem in H^p achieves this bound and that this bound has an exponential rate of convergence. Moreover we show in Theorem 4.1 that among all estimators, the linear estimators are optimal. For a given process in \mathbf{H}^p , with known covariance function, we consider the best parametric linear estimator \hat{L}_N obtained through the inversion of the covariance matrix. We show in Theorem 5.1 that the corresponding maximal error $E_{\hat{L}_N}(\mathbf{H}^p) = d(\mathbf{H}^p)$, and thus has an exponential rate of convergence.

Finally we construct in Section 6 optimal nonparametric estimators with optimal sampling designs in \mathbf{H}^∞ and in \mathbf{H}^2 .

2. NOTATIONS AND DEFINITIONS

First recall the definition of analytic function (e.g. Cartan (1978)). A function $f(z)$ is said to be analytic on an open domain $D \subset \mathbb{C}$ if, for every $z \in D$, $\frac{f(z+h)-f(z)}{h}$ tends to a finite limit as $h \rightarrow 0$, $h \in \mathbb{C}$. An analytic function is equal to its Taylor expansion. From now on, let

$$\Omega = \{z \in \mathbb{C}, |z| < 1\}.$$

We recall the Hardy space H^p , $1 \leq p \leq \infty$, of analytic functions defined on the unit disc of \mathbb{C} is such that (e.g. Rudin (1966), Ch. 17):

$$f \in H^p, 1 \leq p < \infty, \text{ if } M_p(f, r) = \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(re^{i\theta})|^p d\theta \right\}^{\frac{1}{p}} \text{ is bounded as}$$

$$r \rightarrow 1, \text{ and then } \|f\|_{H^p} = \lim_{r \rightarrow 1} M_p(f, r),$$

$$f \in H^\infty \text{ if } M_\infty(f, r) = \sup_{-\pi \leq \theta \leq \pi} |f(re^{i\theta})| \text{ is bounded as}$$

$$r \rightarrow 1, \text{ and then } \|f\|_{H^\infty} = \lim_{r \rightarrow 1} M_\infty(f, r).$$

For a given second order process $X(z), z \in \Omega$ with covariance function R , we define $H(R)$ the Reproducing Hilbert space generated by $R(z_1, z_2), (z_1, z_2) \in \Omega^2$ with inner product $\langle f(\cdot), R(\cdot, z) \rangle_{H(R)} = f(z), z \in \Omega$ for $f \in H(R)$, (see Parzen (1959)).

Let $L^2(X)$ the L^2 -closure of the linear span of the random variables $\{X(z), z \in \Omega\}$. Then every $f \in H(R)$ is of the form $f(z) = \mathbb{E}_X(X(z)\bar{\eta})$ for some $\eta \in L^2(X)$; and the correspondence $f \leftrightarrow \eta$ is an isomorphism between $H(R)$ and $L^2(X)$ so that if $f_i \leftrightarrow \eta_i$ then $\langle f_1, f_2 \rangle_{H(R)} = \mathbb{E}_X(\eta_1\bar{\eta}_2)$.

DEFINITION 2.1. $X(z)$ is analytic in Ω if it is differentiable in the squared mean for all $z \in \Omega$. There exists a process $X'(z)$ such that, as $h \rightarrow 0, h \in \mathbb{C}$

$$\mathbb{E} \left(\frac{X(z+h) - X(z)}{h} - X'(z) \right)^2 \rightarrow 0.$$

It follows that $X(z)$ is weakly differentiable

$$\mathbb{E} \left(\frac{X(z+h) - X(z)}{h} \bar{\eta} \right) \rightarrow \mathbb{E}(X'(z)\bar{\eta}),$$

for all $\eta \in L^2(X)$, as $h \rightarrow 0, h \in \mathbb{C}$. Then, if $f \in H(R), f(z) = \mathbb{E}(X(z)\bar{\eta})$ for some $\eta \in L^2(X)$, we have

$$\frac{f(z+h) - f(z)}{h} = \mathbb{E} \left(\frac{X(z+h) - X(z)}{h} \bar{\eta} \right)$$

$$\rightarrow \mathbb{E}(X'(z)\bar{\eta}),$$

as $h \rightarrow 0, h \in \mathbb{C}$. Thus f is analytic in Ω .

We consider in this paper the class of processes $\mathbf{H}^p, 1 \leq p \leq \infty$, defined as follows

- $1 \leq p < \infty$, X is analytic in Ω and

$$\lim_{r \rightarrow 1} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} R^{p/2}(re^{i\theta}, re^{i\theta}) d\theta \right)^{\frac{1}{p}} \leq 1,$$

- $p = \infty$, X is analytic in Ω and

$$\lim_{r \rightarrow 1} \sup_{\theta \in [-\pi, \pi]} |R(re^{i\theta}, re^{i\theta})| \leq 1.$$

Clearly, $R(re^{i\theta}, re^{i\theta}) = \mathbb{E}|X(re^{i\theta})|^2 \geq 0$, then, for any $\alpha > 0$, $R^\alpha(re^{i\theta}, re^{i\theta})$ is defined as the positive real number η such that $\eta^{1/\alpha} = R(re^{i\theta}, re^{i\theta})$.

We have the following relation between the two spaces \mathbf{H}^p and $H(R)$, $p \geq 1$.

LEMMA 2.2. *Let $X \in \mathbf{H}^p$ be a stochastic process with covariance function R , and let $H(R)$ be the associated Reproducing Hilbert space. Then*

$$H(R) \subset H^p \text{ and } \|\cdot\|_{H^p} \leq \|\cdot\|_{H(R)}.$$

Proof. Let $f \in H(R)$; there exists $\eta \in L^2(X(z), z \in \Omega)$ such that $f(z) = \mathbb{E}(X(z)\bar{\eta})$ and $\|f\|_{H(R)}^2 = \mathbb{E}|\eta|^2$. From Cauchy-Schwarz inequality, we have

$$\begin{aligned} |f(z)|^2 &\leq \mathbb{E}|X(z)|^2 \mathbb{E}|\eta|^2 \\ &= R(z, z) \|f\|_{H(R)}^2, \end{aligned}$$

so that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(re^{i\theta})|^p d\theta \leq \|f\|_{H(R)}^p \frac{1}{2\pi} \int_{-\pi}^{\pi} R^{p/2}(re^{i\theta}, re^{i\theta}) d\theta.$$

It follows that $f \in \mathbf{H}^p$ and $\|f\|_{\mathbf{H}^p} \leq \|f\|_{H(R)}$.

Likewise we can show that $H(R) \subset H^\infty$ and $\|f\|_{\mathbf{H}^\infty} \leq \|f\|_{H(R)}$. □

We wish to estimate the random integral of a real valued process defined on $] - 1, 1[$

$$I(X) = \int_{-1}^1 \phi(t) X(t) dt, \tag{2.1}$$

where ϕ is a weight nonrandom function by linear estimators $L_N(X) \triangleq L(X(t_1), \dots, X(t_N))$ that are based on observations of X at the sampling points $(-1 < t_1 < \dots < t_N < 1)$.

In what follows we assume that the weight function ϕ is analytic with $\sup|\phi| \leq 1$, then the weighted process $\phi X \in \mathbf{H}^p$ if $X \in \mathbf{H}^p$. Indeed, we have for $1 \leq p < \infty$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [R_{\phi X}(re^{i\theta}, re^{i\theta})]^{\frac{p}{2}} d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\phi|^p(re^{i\theta}) [R_X(re^{i\theta}, re^{i\theta})]^{\frac{p}{2}} d\theta,$$

and for $p = \infty$

$$\sup_{\theta \in [-\pi, \pi]} |R_{\phi X}(re^{i\theta}, re^{i\theta})| \leq \sup_{\theta \in [-\pi, \pi]} |\phi^2(re^{i\theta})| \sup_{\theta \in [-\pi, \pi]} |R_X(re^{i\theta}, re^{i\theta})|.$$

By Cauchy-Schwarz inequality, we have

$$\mathbb{E} \left(\int_{-1}^1 \phi(t) X(t) dt \right)^2 \leq 2 \int_{-1}^1 \phi^2(t) R(t, t) dt.$$

Since $H^p \subset L^1$ (e.g. Duren (1970)), we deduce that $\int_{-1}^1 \phi(t)X(t)dt$ is well-defined.

The maximal mean squared error corresponding to such estimators is defined by

$$E_{L_N}(\mathbf{H}^p) = \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - L_N(X))^2,$$

and the minimax linear risk error is given by

$$d(\mathbf{H}^p) = \inf_{L_N} E_{L_N}(\mathbf{H}^p). \tag{2.2}$$

We denote by \mathbf{V}_0 the class of processes in \mathbf{H}^p that vanish at the sampling points (t_1, \dots, t_N) :

$$\mathbf{V}_0 = \{X \in \mathbf{H}^p, X(t_i) = 0 \quad i = 1, \dots, N\}. \tag{2.3}$$

3. RATES OF CONVERGENCE FOR LINEAR ESTIMATORS

Introduce the following rates

$$R_N^p = \sup_{\|f\|_{H^p} \leq 1} \left(\int_{-1}^1 B_N(u) f(u) du \right)^2 \quad \text{for } 1 \leq p \leq \infty, \tag{3.1}$$

where $B_N(u) = \prod_{i=1}^N \frac{u-t_i}{1-ut_i}$ is the finite Blaschke product (cf. Duren (1970) p.20).

The lower bound for the minimax error is given by the following result.

THEOREM 3.1. *For a fixed sampling $T_N = (t_1, \dots, t_N)$ of size N , we have*

$$d(\mathbf{H}^p) \geq R_N^p.$$

Proof. Let $X \in \mathbf{V}_0$ and L_N a linear estimator. Then

$$\begin{aligned} \mathbb{E}_X (I(X))^2 &= \mathbb{E}_X (I(X) - L_N(0, \dots, 0))^2 \\ &\leq \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - L_N(X(t_1), \dots, X(t_N)))^2. \end{aligned}$$

It follows that

$$\sup_{X \in \mathbf{V}_0} \mathbb{E}_X (I(X))^2 \leq \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - L_N(X(t_1), \dots, X(t_N)))^2.$$

Moreover, by Rudin (1966), Th.17.9, we have

$$\sup_{X \in \mathbf{V}_0} \mathbb{E}_X (I(X))^2 = \sup_{\eta \in \mathbf{H}^p} \mathbb{E}_\eta \left(\int_{-1}^1 B_N(u) \eta(u) du \right)^2.$$

In particular, let the process $\eta(t) = f(t)v$ where $f \in H^p, \|f\|_{H^p} \leq 1$ and v is a random variable with mean 0 and variance 1. Then we obtain

$$\begin{aligned} \sup_{\eta \in \mathbf{H}^p} \mathbb{E}_\eta \left(\int_{-1}^1 B_N(u) \eta(u) du \right)^2 &\geq \sup_{\|f\|_{H^p} \leq 1} \mathbb{E} \left(\int_{-1}^1 f(u) v B_N(u) du \right)^2 \\ &= \sup_{\|f\|_{H^p} \leq 1} \left(\int_{-1}^1 B_N(u) f(u) du \right)^2. \end{aligned}$$

□

The upper bound for the minimax error is given by the following result.

THEOREM 3.2. *For any fixed sampling design $T_N = (t_1, \dots, t_N)$ of size N , we have*

$$d(\mathbf{H}^p) \leq R_N^p,$$

and the best linear estimator in the class of stochastic process \mathbf{H}^p is given by the best linear estimator in the class of deterministic functions H^p .

Proof. For any linear estimator L_N and $X \in \mathbf{H}^p$ we have, Parzen (1959), Wahba (1990)

$$\begin{aligned} & \mathbb{E}_X (I(X) - L_N(X))^2 \\ &= \sup \left\{ \left(\int h - L_N(h) \right)^2, h \in H(R), \|h\|_{H(R)} \leq 1 \right\} \\ &\leq \sup \left\{ \left(\int h - L_N(h) \right)^2, h \in H^p, \|h\|_{H^p} \leq 1 \right\}. \end{aligned}$$

The last inequality follows from Lemma 2.2. Then

$$\begin{aligned} & \inf_{L_N} \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - L_N(X))^2 \\ &\leq \inf_{L_N} \sup \left\{ \left(\int h - L_N(h) \right)^2, h \in H^p, \|h\|_{H^p} \leq 1 \right\}. \end{aligned} \quad (3.2)$$

The linear estimator that minimizes the right-hand side of (3.2) is therefore the best linear estimator for deterministic functions in H^p . From Miccheli and Rivlin (1984), we have

$$d(\mathbf{H}^p) \leq \sup \left\{ \left(\int h \right)^2, h(t_i) = 0, i = 1, \dots, N, \|h\|_{H^p} \leq 1 \right\},$$

and the result follows from Rudin (1966), Th.17.9. \square

REMARK 3.3. It follows from Theorems 3.1 and 3.2 that $d(\mathbf{H}^p) = R_N^p = \sup_{X \in \mathbf{V}_0} \mathbb{E}_X (I(X))^2$. The optimal sampling design is therefore given by $(t_1^*, \dots, t_N^*) = \text{Argmin}_{(t_1, \dots, t_N)} R_N^p$. Newman (1979), Loeb and Werner (1974) and Miccheli and Rivlin (1984) gave upper and lower bounds for the rate of convergence of the minimax error associated with the optimal sampling design for approximating integrals of deterministic functions in H^p , $1 \leq p \leq \infty$. Define the linear optimal minimax risk associated with the optimal sampling design

$$\begin{aligned} d^*(\mathbf{H}^p) &= \inf_{(t_1, \dots, t_N)} d(\mathbf{H}^p) \\ &= \inf_{B_N} \sup_{\|f\|_{H^p} \leq 1} \left(\int_{-1}^1 B_N(u) f(u) du \right)^2. \end{aligned} \quad (3.3)$$

Then we obtain for the random case

for $1 \leq p < \infty$,

$$\frac{4}{9} \exp \left\{ -12 \sqrt{\frac{N(p-1)}{p}} \right\} \leq d^*(\mathbf{H}^p) \leq 121 \exp \left\{ -\sqrt{\frac{N(p-1)}{p}} \right\}$$

$$\exp \left\{ -10\pi \sqrt{\frac{N}{2}} \right\} \leq d^*(\mathbf{H}^\infty) \leq \exp \left\{ -2\pi \sqrt{\frac{N}{2}} \right\}.$$

As pointed out by Newman (1979) for approximating integrals of deterministic H^1 -functions, no quadrature formula “works” for $p = 1$ in the random case.

Thus we see that the rate of convergence for the optimal minimax error for estimating integrals of analytic processes is of order exponential and can even be faster for $p > 1$. This generalizes the case when the processes to integrate have a s -Hölder smoothness in quadratic mean studied by Sacks and Ylvisaker (1966, 1968, 1970), Benhenni and Cambanis (1992a), Istas and Laredo (1997), Stein (1995) where the rate of convergence of the mean squared error is of order N^{2s+1} . It is a difficult task to find the optimal sampling design of size N that minimizes the minimax error for any $1 \leq p < \infty$. The case $p = \infty$ is covered by Bojanov (1974) and is treated in Section 6.2. It remains the problem of finding optimal linear estimators. They are not available when integrals of deterministic functions in H^p , $1 \leq p < \infty$, $p \neq 2$, are approximated.

4. RATES OF CONVERGENCES FOR NON-LINEAR ESTIMATORS

We wish now to estimate $I(X)$ by non-linear estimators, and we want to know if these estimators could improve actually the performance of linear estimators. Let $\alpha_N(X) \triangleq \alpha(X(t_1), \dots, X(t_N))$ be a (general) estimator. The “non-linear minimax” risk is defined by

$$e(\mathbf{H}^p) = \inf_{\alpha_N} \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \alpha_N(X))^2.$$

The linear estimators are optimal among all the estimators.

THEOREM 4.1. *For any fixed sampling design $T_N = (t_1, \dots, t_N)$ of size N , we have*

$$e(\mathbf{H}^p) = d(\mathbf{H}^p) = R_N^p.$$

Proof. For all $X \in \mathbf{V}_0$,

$$\begin{aligned} \mathbb{E}_X (I(X))^2 &= \mathbb{E}_X (I(X) - \alpha_N(0, \dots, 0) + \alpha_N(0, \dots, 0))^2 \\ &= \mathbb{E}_X (I(X) - \alpha_N(0, \dots, 0))^2 - \alpha_N^2(0, \dots, 0) \\ &\leq \mathbb{E}_X (I(X) - \alpha_N(0, \dots, 0))^2 \\ &\leq \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \alpha_N(X))^2, \end{aligned}$$

then

$$\sup_{X \in \mathbf{V}_0} \mathbb{E}_X (I(X))^2 \leq \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \alpha_N(X))^2,$$

$$\sup_{X \in \mathbf{V}_0} \mathbb{E}_X (I(X))^2 \leq \inf_{\alpha_N} \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \alpha_N(X))^2.$$

The proof then follows from the proofs of Theorems 3.1 and 3.2 and the straightforward bound $e(\mathbf{H}^p) \leq d(\mathbf{H}^p)$. □

5. PARAMETRIC OPTIMAL LINEAR ESTIMATOR IN $\mathbf{H}^p, 1 \leq p \leq \infty$

Let $f \in H(R)$ of the form $f(t) = \int_{-1}^1 \phi(s)R(s,t)ds$, and for a sampling design $T_N = \{t_1, \dots, t_N\}$ of size N we define the covariance matrix, assumed non-singular, $R_{T_N} = \{R(t_i, t_j)\}_{1 \leq i, j \leq N}$. Then it is known, Sacks and Ylvisaker (1966, 1968, 1970), Benhenni and Cambanis (1992a), that the estimator defined by $\widehat{L}_N(X) = f'_{T_N} R_{T_N}^{-1} X_{T_N}$ minimizes the mean squared error $\mathbb{E}_X (I(X) - L_N(X))^2$ with respect to the class of all linear estimators, where $f'_{T_N} = (f(t_1), \dots, f(t_N))$ and $X'_{T_N} = (X(t_1), \dots, X(t_N))$. The linear estimators $\widehat{L}_N(X)$, defined in the class $\mathbf{H}^p, p \geq 1$, are optimal in the sense of minimax error. This is given by the following result.

THEOREM 5.1. *For a fixed sampling design $T_N = (t_1, \dots, t_N)$ of size N , we have*

$$\sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2 = R_N^p.$$

Proof. For any linear estimator L_N and any process $X \in \mathbf{H}^p$, we have

$$\begin{aligned} \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2 &\leq \mathbb{E}_X (I(X) - L_N(X))^2, \\ \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2 &\leq \inf_{L_N} \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - L_N(X))^2 \\ &= R_N^p, \end{aligned}$$

by Theorems 3.1 and 3.2.

For any process $X \in \mathbf{V}_0$, we have

$$\begin{aligned} \mathbb{E}_X (I(X))^2 &= \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2, \\ \sup_{X \in \mathbf{V}_0} \mathbb{E}_X (I(X))^2 &\leq \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2, \\ R_N^p &\leq \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2, \end{aligned}$$

by Theorem 3.1. □

Then we deduce from Remark 3.3 the rate of convergence for the maximal error when optimal estimators $\widehat{L}_N(X)$ associated with optimal sampling design are used in the class \mathbf{H}^p ,

$$\begin{aligned} \frac{4}{9} \exp \left\{ -12 \sqrt{\frac{N(p-1)}{p}} \right\} &\leq \inf_{T_N} \sup_{X \in \mathbf{H}^p} \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2 \\ &\leq 121 \exp \left\{ -\sqrt{\frac{N(p-1)}{p}} \right\}, 1 \leq p < \infty, \\ \exp \left\{ -10\pi \sqrt{\frac{N}{2}} \right\} &\leq \inf_{T_N} \sup_{X \in \mathbf{H}^\infty} \mathbb{E}_X (I(X) - \widehat{L}_N(X))^2 \leq \exp \left\{ -2\pi \sqrt{\frac{N}{2}} \right\}. \end{aligned}$$

These estimators also do not “work” for $p = 1$. Therefore the rate of convergence is of order exponential or faster in the class of analytic process in \mathbf{H}^p , $p > 1$. These optimal estimators are highly parametric because they depend on the covariance function assumed known. Besides they are unstable since they require the inversion of the covariance matrix.

6. CONSTRUCTION OF OPTIMAL ESTIMATORS

There are no simple estimators available for any $1 \leq p < \infty, p \neq 2$. However, there are estimators for the cases $p = 2$ and $p = \infty$.

6.1. NONPARAMETRIC OPTIMAL ESTIMATORS IN \mathbf{H}^2

Let $f(z) = \sum_{p \geq 0} f_p z^p$ be an analytic function. The space H^2 is a Hilbert space with norm $\|f\|_{H^2}^2 = \sum_{p \geq 0} f_p^2$ (Rudin (1966), ch.17). For any sampling design $T_N = \{-1 < t_1 < \dots < t_N < 1\}$, define the subspace

$$V_0 = \{f \in H^2, f(t_i) = 0, i = 1, \dots, N\},$$

and let W be the orthogonal supplement of V_0 in H^2 : $V_0 \oplus W = H^2, V_0 \perp W$. An inspection of the kernel of the linear application from H^2 onto $\mathbb{R}^N, f \rightarrow (f(t_1), \dots, f(t_N))$ makes it clear that $\dim(W) = N$. Consider the following N functions of H^2

$$g_i(z) = \sum_{p \geq 0} t_i^p z^p \quad z \in [-1, 1] \quad i = 1, \dots, N.$$

For any function $f \in V_0, \langle f, g_i \rangle_{H^2} = \sum_{p \geq 0} t_i^p f_p = f(t_i) = 0$, and therefore $V_0 \perp \text{span}\{g_i, i = 1, \dots, N\}$. Moreover, $\{g_i, i = 1, \dots, N\}$ is an independent family so that $\text{span}\{g_i, i = 1, \dots, N\} = W$. The orthogonal projection P of H^2 onto W can be written as

$$P(f)(z) = \sum_{i,j=1}^N \lambda_{i,j} f(t_j) g_i(z),$$

where the coefficients $\lambda_{i,j}$ are determined using $P(g_k) = g_k$ and $g_k(t_j) = \frac{1}{1-t_k t_j}$. They are given by the following system

$$\sum_{j=1}^N \lambda_{i,j} \frac{1}{1-t_k t_j} = \delta_{i,k} \text{ for } i, k = 1, \dots, N.$$

We then construct the estimator through $\bar{L}_N(f) = \int_{-1}^1 P(f)$ and we obtain

$$\bar{L}_N(f) = \sum_{i,j=1}^N \lambda_{i,j} f(t_j) \frac{\text{Log}\left(\frac{1+t_i}{1-t_i}\right)}{t_i}. \tag{6.1}$$

Following Theorems 3.1 and 3.2, Miccheli and Rivlin (1977), Th.3, the optimality of (6.1) is obtained, and is stated by the following Theorem.

THEOREM 6.1. For any sampling design $T_N = (t_1, \dots, t_N)$ of size N , we have

$$\sup_{X \in \mathbf{H}^2} \mathbb{E}_X (I(X) - \bar{L}_N(X))^2 = d(\mathbf{H}^2) = R_N^2.$$

The estimators $\bar{L}_N(X)$ are nonparametric as they do not require knowledge of the covariance function.

It follows from Remark 3.3 with $p = 2$ that these estimators have also an exponential rate of convergence.

EXAMPLE 6.2. Let X be a centered process with covariance function $R(z_1, z_2) = \exp\{-(z_1 - \bar{z}_2)^2\}$, $|z_1|, |z_2| < 1$. Then the covariance function R can be expanded as

$$R(z_1, z_2) = \sum_{n,m \geq 0} r_{n,m} z_1^n \bar{z}_2^m, \text{ with } \sum_{n,m \geq 0} |r_{n,m}| < \infty.$$

It follows that the process can be expanded as $X(z) = \sum_{n \geq 0} x_n z^n$, where $\{x_n, n \geq 0\}$ are centered random variables satisfying $\mathbb{E} x_n x_m = r_{n,m}$, and thus belongs to \mathbf{H}^2 . Let $T_N = (\frac{i}{N}, i = -N+1, \dots, 0, \dots, N-1)$ be a regular sampling. The normalized mean squared error $\mathbb{E}_X (I(X) - \bar{L}_N(X))^2 / \mathbb{E} I^2(X)$ was computed for T_N with several values of N (cf. Figure 1 in the Appendix). We see that the mean squared error decreases quickly even with a very small sample size N .

6.2. NONPARAMETRIC OPTIMAL ESTIMATORS IN \mathbf{H}^∞

For $p = \infty$, using Bojanov's formula (Bojanov (1974)), we can define simple linear optimal estimators that use only observations at the optimal sampling points generated by

$T_N^* = (t_1^*, \dots, t_N^*) = \text{Argmin}_{(t_1, \dots, t_N)} \int_{-1}^1 B_N^2(u) du$, and thus are nonparametric. They are of the form

$$\bar{L}_N(X) = \sum_{i=1}^N c_{i,N} (\phi X)(t_i^*),$$

where the coefficients $c_{i,N}$ are defined by

$$c_{i,N} = \int_{-1}^1 Q_{i,N}(x) dx,$$

where

$$Q_{i,N}(x) = \frac{\omega_i^2(x)}{\omega_i^2(t_i^*)} \left\{ 1 - W_i^4(x) - 2 \frac{\omega_i'(t_i^*)}{\omega_i(t_i^*) W_i'(t_i^*)} W_i(x) (1 - W_i^2(x)) \right\},$$

and

$$W_i(x) = \frac{x - t_i^*}{1 - x t_i^*},$$

$$\omega_i(x) = \prod_{k=1, k \neq i}^N W_k(x)$$

The optimality of these estimators is given by the following result.

THEOREM 6.3. For the optimal sampling T_N^* of size N , we have

$$\sup_{X \in \mathbf{H}^\infty} \mathbb{E}_X (I(X) - \bar{L}_N(X))^2 = \left(\int_{-1}^1 B_N^2(u) du \right)^2.$$

The optimal sampling design T_N^* is obtained numerically by minimizing the constant $\int_{-1}^1 B_N^2(u) du$. For instance when $N = 4$, $T_4^* = \{-t_1^*, -t_2^*, t_1^*, t_2^*\}$ where $t_1^* \sim 0.892$, $t_2^* \sim 0.401$ in which case $\int_{-1}^1 B_4^2(u) du \sim 0.0518$. When $N = 12$, $T_{12}^* = \{-t_1^*, -t_2^*, \dots, t_1^*, t_2^*\}$ where $t_1^* \sim 0.986$, $t_2^* \sim 0.963$, $t_3^* \sim 0.904$, $t_4^* \sim 0.777$, $t_5^* \sim 0.544$, $t_6^* \sim 0.198$ in which case $\int_{-1}^1 B_{12}^2(u) du \sim 8.02 \cdot 10^{-5}$. Round-off problems occur in computing the coefficients of Bojanov's estimator for large sample sizes and hence it is unstable.

Proof. We have from Lemma 2.2 and Bojanov (1974), for $X \in \mathbf{H}^\infty$,

$$\begin{aligned} & \mathbb{E}_X (I(X) - \bar{L}_N(X))^2 \\ &= \sup \left\{ \left(\int_{-1}^1 h - \bar{L}_N(h) \right)^2, h \in H(R), \|h\|_{H(R)} \leq 1 \right\} \\ &\leq \sup \left\{ \left(\int_{-1}^1 h - \bar{L}_N(h) \right)^2, h \in H^\infty, \|h\|_{H^\infty} \leq 1 \right\} \\ &= \left(\int_{-1}^1 B_N^2(u) du \right)^2. \end{aligned}$$

The equality follows from Theorem 3.1. □

Likewise, from Remark 3.3, we deduce upper and lower bounds for the rate of convergence of the maximal error when these estimators are used

$$\exp \left\{ -10\pi \sqrt{\frac{N}{2}} \right\} \leq \sup_{X \in \mathbf{H}^\infty} \mathbb{E}_X (I(X) - \bar{L}_N(X))^2 \leq \exp \left\{ -2\pi \sqrt{\frac{N}{2}} \right\}.$$

Therefore for $p = \infty$, using simple nonparametric optimal estimators, the rate of convergence for the maximal error is of the same exponential order as the optimal coefficient estimators $\hat{L}_N(X)$.

ACKNOWLEDGEMENTS

We thank GABRIEL LANG for his help in Section 6.1.

APPENDIX A. FIGURE

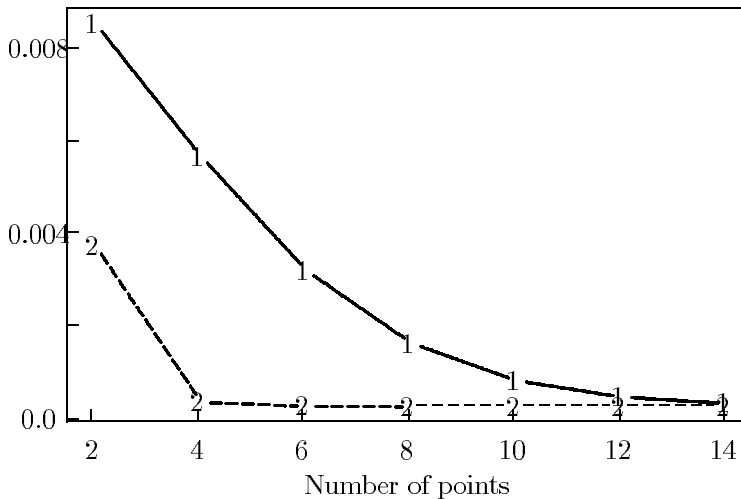


FIGURE 1. 1: Normalized error for simple estimator $\bar{L}_N(X)$ in \mathbf{H}^2 . 2: Normalized error for optimal parametric estimator $\hat{L}_N(X)$ in \mathbf{H}^2

REFERENCES

- ABRAMOWITZ, M., STEGUN, I. (1965). *Handbook of Mathematical Functions*. Dover, New York.
- BENHENNI, K., CAMBANIS, S. (1992a). Sampling designs for estimating integrals of stochastic processes. *Ann. Statist.* **20** 161–194.
- BENHENNI, K., CAMBANIS, S. (1992b). Sampling designs for estimating integrals of stochastic processes using quadratic mean derivatives. In *Approximation theory: proceedings of the sixth southeastern approximation theorists annual conference*, G.A. Anastassiou ed., 93–123.
- BOJANOV, B. (1974). On an optimal quadrature formula. *Comptes rendus de l'Académie bulgare des Sciences* **27** (5) 619–621.
- CAMBANIS, S. (1985). Sampling designs for time series. In *Time series in the time domain. Handbook of statistics*, E.J. Hannan, P.R. Krishnaiah, M.M. Rao, eds. **5** 337–362.
- CARTAN, H. (1978). *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Hermann, Paris.
- CRAMER, H., LEADBETTER, M.R. (1967). *Stationary and related stochastic processes*. Wiley, New-York.
- DUREN, P. (1970). *Theory of H^p spaces*. New-York and London.
- EUBANK, R. L., SMITH, P.L., SMITH, P.W. (1982). A note on optimal and asymptotically optimal designs for certain time series models. *Ann. Statist.* **10** 1295–1301.
- IBRAGIMOV, I., HAS'MINSKII, R. (1981). *Statistical estimation, Asymptotic theory*. Springer-Verlag, New-York.
- ISTAS, J., LAREDO, C. (1997). Estimating functionals of a stochastic process. *Adv. Appl. Prob.* **10** (1) 249–270.
- LOEB, H., WERNER, H. (1974). Optimal numerical quadrature in H^p spaces. *Math. Z.* **138** 111–117.
- MICCHELI, C., RIVLIN, T. (1977). Optimal estimation in approximation theory. In *A survey of optimal recovery* C.A. Miccheli, T.J. Rivlin eds, Plenum, N.Y., 1–54.
- MICCHELI, C., RIVLIN, T. (1984). Lectures on optimal recovery. In *Numerical Analysis, Lecture Notes in Math.* **1129** 21–93, Springer.
- NEWMAN, D. (1979). Quadrature formulae for H_p functions. *Math. Z.* **166** 111–115.

- PARZEN, E. (1959). Statistical inference on time series by Hilbert space methods. In *Time series analysis papers*, E. Parzen ed. Holden-Day, San Fransisco, 251–382.
- RUDIN, W. (1966). *Real and complex analysis*. McGraw-Hill, New-York.
- SACKS, J., YLVISAKER, D. (1966). Designs for regression problem problems with correlated errors. *Ann. Math. Stat.* **37** 66–89.
- SACKS, J., YLVISAKER, D. (1968). Designs for regression problem problems with correlated errors II. *Ann. Math. Stat.* **39** 49–69.
- SACKS, J., YLVISAKER, D. (1970). Designs for regression problem problems with correlated errors III. *Ann. Math. Stat.* **41** 2057–2074.
- SACKS, J., YLVISAKER, D. (1971). Statistical designs and integral approximation. In *Proceedings of the twelfth biennial canadian mathematical society seminar* R. Pyke eds. Canadian Mathematical Congress, 115–136.
- STEIN, M. (1995). Predicting integrals of stochastic processes. *Ann. Appl. Prob.* **5** (1) 158–170.
- WAHBA, G. (1990). *Spline models for observational data*. S.I.A.M., Philadelphia, Pennsylvania.

KARIM BENHENNI: LABORATOIRE DE STATISTIQUE ET ANALYSE DES DONNÉES, BSHM, UNIVERSITÉ PIERRE MENDÈS FRANCE, 1251 AV. CENTRALE, BP 47 38040 GRENOBLE CEDEX 09, FRANCE. E-MAIL: benhenni@labsad.upmf-grenoble.fr.

JACQUES ISTAS: LABORATOIRE DE BIOMÉTRIE, DOMAINE DE VILVERT, I.N.R.A., 78350 JOUY-EN-JOSAS, FRANCE. E-MAIL: Jacques.Istas@jouy.inra.fr.