KYBERNETIKA — VOLUME 16 (1980), NUMBER 4

An Optimal Property of the Best Linear Unbiased Interpolation Filter

FRANTIŠEK ŠTULAJTER

The RKHS methods are used to prove an optimal property of the best linear unbiased interpolation filter in the case of a sum of two independent Gaussian processes.

1. INTRODUCTION

Let us consider the well known problem of interpolation with filtration. Let $X(t) = S(t) + N(t)$; $t \in T$ be a signal plus noise observed random process with $S = \{S(t);\ t \in T\}$ and $N = \{N(t);\ t \in T\}$ independent Gaussian random processes defined on a measurable space $(\Omega, \mathscr{A})$. It will be assumed that we know the covariance functions $R_S(s, t)$ and $R_N(s, t)$; $s, t \in T$ of these processes. These covariance functions are assumed to be continuous on $T \times T$. Let the random process N have zero mean value. The mean value of S is unknown, it is assumed merely that it belongs to some subspace M of $H(R_X)$, where $H(R_X)$ is a reproducing kernel Hilbert space (RKHS) with a kernel given by $R_X(s, t) = R_S(s, t) + R_N(s, t)$; $s, t \in T$. The problem of finding the best linear unbiased estimate (BLUE) $\tilde{S}_M(t)$ of $S(t)$ given $X = \{X(t);\ t \in T\}$ for a fixed $t \in T$ was solved by Parzen [5]. Our aim is to show an optimal property of the process $\tilde{S}_M = \{\tilde{S}_M(t);\ t \in T\}$ in the case when M is finite-dimensional. It is well known, see Kallianpur [4], Parzen [5] that for a Gaussian process X we have $P(X(\cdot) \in H(R_X)) = 0$. Nevertheless, as was shown by Pitcher [6], Driscoll [3] and Baker [2], in the case considered some additional conditions on S assure that

$$P_m(S(.) \in m \oplus H(R_N)) = 1 \, .$$

It will be shown that, for the finite-dimensional M,

$$P_m(\tilde{S}_M(.) \in m \oplus H(R_N)) = 1 \quad \text{for all} \quad m \in M \, .$$

Next it will be proved that $\tilde{S}_M(.)$ is the best unbiased estimate of $S(.)$ given X under the generalized square-error function $L(a, b) = \|a - b\|^2_{H(R_N)}$. This result was announced by Driscoll [3], for the case $M = \{0\}$, too.

2. THE MAIN RESULT

Let the covariance functions R_N and R_S be of the form:

$$R_N(s, t) = \sum_{k=1}^{\infty} \lambda_k \varphi_k(s) \varphi_k(t) ; \quad s, t \in T, \quad \lambda_k > 0 , \quad \sum_{k=1}^{\infty} \lambda_k < \infty \tag{1}$$

and

$$R_S(s, t) = \sum_{k=1}^{\infty} \mu_k \lambda_k \varphi_k(s) \varphi_k(t) \quad \text{with} \quad \mu_k \geqq 0 , \quad \sum_{k=1}^{\infty} \mu_k < \infty , \tag{2}$$

where $\{\varphi_k\}_{k=1}^{\infty}$ is a complete orthonormal system (CONS) in $L^2[T]$. The condition $\sum_{k=1}^{\infty} \mu_k < \infty$ is sufficient to ensure $P_0(S(.) \in H(R_N)) = 1$, see Pitcher [6].

From the RKHS theory (see Aronszajn [1]) we know that $\{\psi_k(t) = \sqrt{(\lambda_k)}\, \varphi_k(t)\}_{k=1}^{\infty}$ is a CONS in $H(R_N)$. Further, because $R_X(s, t) = \sum_{k=1}^{\infty} (1 + \mu_k) \psi_k(s) \psi_k(t)$, the space $H(R_X)$ can be characterized by:

$$H(R_X) = \left\{ f \in H(R_N) : \sum_{k=1}^{\infty} \frac{\langle f, \psi_k \rangle^2_{H(R_N)}}{1 + \mu_k} < \infty \right\} .$$

The system of vectors $\{\sqrt{(1 + \mu_k)}\, \psi_k\}_{k=1}^{\infty}$, is a CONS in $H(R_X)$.

It was shown by Parzen [5] that

$$\tilde{S}_M(t) = \langle X, R_S(., t) \rangle_{H(R_X)} + \langle X, \mathscr{P}^M[R_N(., t)] \rangle_{H(R_X)} \tag{3}$$

is the BLUE of $S(t)$ given X for every fixed $t \in T$. Here $\langle X, g \rangle_{H(R_X)}$; $g \in H(R_X)$, denotes an isomorphic image of an element $g \in H(R_X)$ in the space $L^2[X(t);\ t \in T]$ (see Parzen [5]) and $\mathscr{P}^M$ is a projection operator to the subspace M defined on $H(R_X)$.

Lemma. Let M be a finite-dimensional subspace of $H(R_X)$ and let the conditions (1) and (2) are satisfied. Then

$$P_m(\tilde{S}_M(.) \in m \oplus H(R_N)) = 1 \quad \text{for every} \quad m \in M .$$

Proof. It is enough to prove that $P_0(\tilde{S}_M(.) \in H(R_N)) = 1$. Because $\tilde{S}_M(t) = \tilde{S}(t) + \tilde{N}_M(t)$; $t \in T$, where we have used the notations $\tilde{S}(t) = \langle X, R_S(., t) \rangle_{H(R_X)}$ and $\tilde{N}_M(t) = \langle X, \mathscr{P}^M[R_N(., t)] \rangle_{H(R_X)}$, the lemma will be proved by showing that $P_0(\tilde{S}(.) \in H(R_N)) = 1$ and $P_0(\tilde{N}_M(.) \in H(R_N)) = 1$. To do this we can write:

$$\tilde{S}(t) = \langle X, R_S(., t) \rangle_{H(R_X)} = \langle X, \sum_{k=1}^{\infty} \mu_k \psi_k(t) \psi_k(.) \rangle_{H(R_X)} =$$

$$= \sum_{k=1}^{\infty} \mu_k \langle X, \psi_k \rangle_{H(R_X)} \cdot \psi_k(t) \,; \quad t \in T\,.$$

Moreover, we have $P_0\left(\sum_{k=1}^{\infty} \mu_k^2 \langle X, \psi_k \rangle_{H(R_X)}^2 < \infty\right) = 1$, because

$$\sum_{k=1}^{\infty} \mu_k^2 E_0[\langle X, \psi_k \rangle_{H(R_X)}^2] = \sum_{k=1}^{\infty} \frac{\mu_k^2}{1 + \mu_k} < \infty$$

and thus $P_0(\tilde{S}(.) \in H(R_N)) = 1$. Further

$$\tilde{N}_M(t) = \langle X, \mathscr{P}^M[R_N(\cdot, t)] \rangle_{H(R_X)} = \sum_{k=1}^{\infty} \langle X, \mathscr{P}^M[\psi_k] \rangle_{H(R_X)} \cdot \psi_k(t); \; t \in T\,.$$

The series $\sum_{k=1}^{\infty} \langle X, \mathscr{P}^M[\psi_k] \rangle_{H(R_X)}^2$ converges P_0-almost surely, because

$$\sum_{k=1}^{\infty} E[\langle X, \mathscr{P}^M[\psi_k] \rangle_{H(R_X)}^2] =$$

$$= \sum_{k=1}^{\infty} \langle \mathscr{P}^M[\psi_k], \psi_k \rangle_{H(R_X)} = \sum_{k=1}^{\infty} \frac{1}{1 + \mu_k} \langle \mathscr{P}^M[\sqrt{(1 + \mu_k)}\, \psi_k], \sqrt{(1 + \mu_k)}\, \psi_k \rangle_{H(R_X)} \leqq$$

$$\leqq \operatorname{tr} \mathscr{P}^M < \infty$$

if M is finite-dimensional and the lemma is proved.

Remarks. (1) If $M = \{0\}$, then $\tilde{S}(t)$ is the BLUE of $S(t)$ for every fixed $t \in T$.

(2) Because $X = S + N$, where S and N are independent Gaussian processes, $\tilde{S}(t) = E[S(t) \mid \mathscr{B}_X]$; $t \in T$, where $\mathscr{B}_X$ denotes a completion of a sub σ-algebra of $\mathscr{A}$ generated by the random process $X = \{X(t); t \in T\}$.

From this lemma we clearly have $P_m(S(.) - \tilde{S}_M(.)) \in H(R_N)) = 1$ for every $m \in M$. Thus almost all sample paths of the Gaussian process $\{S(t) - \tilde{S}_M(t); t \in T\}$ belong to $H(R_N)$. This process generates a Gaussian measure $\tilde{\mu}_M$ in $H(R_N)$ uniquely determined by its covariance operator $\tilde{\boldsymbol{R}}_M$, for which we have:

$$E_0[\|S(.) - \tilde{S}_M(.)\|_{H(R_N)}^2] = \operatorname{tr} \tilde{\boldsymbol{R}}_M = \sum_{k=1}^{\infty} \langle \tilde{\boldsymbol{R}}_M \psi_k, \psi_k \rangle_{H(R_N)} < \infty\,.$$

For these results, see Driscoll [3].

Let, for every $t \in T$, $\hat{S}_M(t)$ be any linear estimate of $S(t)$ given X such that $P_0(\hat{S}_M(.) \in H(R_N)) = 1$. Then we have:

$$\hat{S}_M(t) = \langle X, \mathscr{P}^M[R_X(., t)] \rangle_{H(R_X)} + \langle X, \hat{h}_t \rangle_{H(R_X)} =$$
$$= \tilde{S}_M(t) - \langle X, \mathscr{P}^{M\perp}[R_S(., t)] \rangle_{H(R_X)} + \langle X, \hat{h}_t \rangle_{H(R_X)} \,; \quad t \in T\,,$$

where $\hat{h}_t$ is any element of $M^{\perp}$. From this we get:

$$\langle \hat{\boldsymbol{R}}_M R_N(., s), R_N(., t) \rangle_{H(R_N)} = E_0[\hat{S}_M(s) - S(s)]\,[\hat{S}_M(t) - S(t)] =$$

$$= E_0[\tilde{S}_M(s) - S(s)]\,[\tilde{S}_M(t) - S(t)] +$$
$$+ E_0[\langle X, \mathscr{P}^{M\perp}[R_S(.,s)] - \hat{h}_s\rangle \,.\, \langle X, \mathscr{P}^{M\perp}[R_S(.,t)] - \hat{h}_t\rangle] =$$
$$= \langle \hat{\boldsymbol{R}}_M R_N(.,s), R_N(.,t)\rangle_{H(R_N)} + E_0[\langle X, \mathscr{P}^{M\perp}[R_S(.,t)] - \hat{h}_t\rangle_{H(R_X)} \,.$$
$$.\, \langle X, \mathscr{P}^{M\perp}[R_S(.,t)] - \hat{h}_s\rangle_{H(R_X)}]\,.$$

Now we can deduce that

$$\operatorname{tr} \hat{\boldsymbol{R}}_M = E_m[\|S(.) - \hat{S}_M(.)\|^2_{H(R_N)}] \geqq E_m[\|S(.) - \tilde{S}_M(.)\|^2_{H(R_N)}] = \operatorname{tr} \tilde{\boldsymbol{R}}_M\,.$$

We set

$$E_m[\|S(.) - \hat{S}_M(.)\|^2_{H(R_N)}] = +\infty \quad \text{if} \quad P_0[\hat{S}_M(.) \in H(R_N)] = 0\,.$$

The results obtained are formulated in the following theorem.

Theorem. Let $X(t) = S(t) + N(t)$; $t \in T$, where N and S are independent Gaussian processes with continuous covariance functions given by (1) and (2). Let $E[N(t)] \equiv 0$ and $E_m[S(t)] = m(t)$; $t \in T$, where $m(.) \in M$, M-finite-dimensional subspace of $H(R_X)$. Let $\tilde{S}_M(t)$ be the BLUE of $S(t)$ given X, given by (3) for every $t \in T$. Then

$$E_m[\|S(.) - \tilde{S}_M(.)\|^2_{H(R_N)}] \leqq E_m[\|S - \hat{S}_M\|^2_{H(R_N)}]$$

for any unbiased linear estimate $\hat{S}_M(t)$ of $S(t)$; $t \in T$ given X.

(Received March 31, 1979.)

REFERENCES

[1] N. Aronszajn: Theory of Reproducing Kernels. Trans. Amer. Math. Soc. *68* (1950), 337—404.

[2] Ch. R. Baker: Relations Between Mutual Information, Strong Equivalence, Signal-to-Noise Ratio and Signal Sample Path Properties. Trans. of 8-th Prague Conference ..., (M. Driml, ed.), Vol. A. Academia, Prague 1978, 85—92.

[3] M. F. Driscoll: The Reproducing Kernel Hilbert Space Structure of the Sample Paths of a Gaussian Process. Z. Wahrscheinlichkeitstheorie verw. Geb. *26* (1973), 309—316.

[4] G. Kallianpur: Zero-one Laws for Gaussian Processes. Trans. Amer. Math. Soc. *149* (1970), 199—211.

[5] E. Parzen: Time Series Analysis Papers. Holden-Day, San Francisco 1967.

[6] T. S. Pitcher: On the Sample Function of Processes, Which Can Be Added to a Gaussian Process. Annals of Math. Stat. *34* (1963), 329—333.

RNDr. František Štulajter, CSc., Ústav merania a meracej techniky SAV (Institute of Measurement and Measuring Technique — Slovak Academy of Sciences), Dúbravská cesta, 885 27 Bratislava, Czechoslovakia.