Kybernetika

František Štulajter

The linear filtration and prediction of indirectly observed random processes

Kybernetika, Vol. 12 (1976), No. 5, (316)--327

Persistent URL: http://dml.cz/dmlcz/124747

KYBERNETIKA — VOLUME 12 (1976), NUMBER 5

The Linear Filtration and Prediction of Indirectly Observed Random Processes

FRANTIŠEK ŠTULAJTER

The RKHS methods are used to develop the theory of estimation of indirectly observed random processes. On the basis of this theory the comparison of Kalman-Bucy's and Parzen's methods of filtration is performed.

1. INTRODUCTION

The theory of linear estimation of random variables, especially the theory of linear filtration and prediction of random processes in the past few years has been developed in two main directions. There exists Parzen's theory [8; 9] based on the RKHS methods and the theory of Kalman and Bucy [5] based on stochastic differential equations.

The aim of this paper is to develop on the basis of Parzen's theory the theory of linear filtration and prediction of indirectly observed random processes and then to compare Parzen's and Kalman-Bucy's methods of linear filtration. As we shall show in section 5, Kalman-Bucy's treatment of the problem may be regarded as a special case of Parzen's methods applied to indirectly observed random processes.

Let $Y = \{Y(t), t \in T\}$, where T is a closed interval on the real line be a random process defined on the probability space $(\Omega, \mathsf{S}, \mathsf{P})$ such that $\mathsf{E}_m Y(t) = m(t)$; $t \in T$ and $R_Y(s, t) = \operatorname{Cov}\{Y(s), Y(t)\}$; $s, t \in T$ exist. It is further assumed that $m(\cdot) \in M$, where M is a given set containing $m \equiv 0$. Let $L_2\{Y(t), t \in T\}$ be a subspace of the Hilbert space $L_2(\Omega, \mathsf{S}, \mathsf{P})$ spanned by the random variables $\{Y(t), t \in T\}$ with the inner product $(U, V) = \mathsf{E}_0(U \cdot V)$.

Let Z be a random variable with the finite dispersion $\mathsf{D}^2 Z$ and denote by $\mathsf{E}_m Z$; $m \in M$ the mean value of Z.

Definition 1.1. We say that the random variable Z *is linearly unbiased estimable* (from observations of the random process Y) if there exists a random variable $U \in$

$\in L_2\{Y(t), t \in T\}$ such that $\mathsf{E}_m U = \mathsf{E}_m Z$ for all $m \in M$. We call the random variable U *the linear unbiased estimate* (LUE) of the random variable Z. Our aim is to find (if it exists) *the best linear unbiased estimate* (BLUE) $\tilde{Z}$ of the random variable Z, minimizing *the mean square error of estimation* $\mathsf{E}_0[Z - \tilde{Z}]^2$.

2. PRELIMINARIES REGARDING HILBERT SPACES

As it is well known, *the Hilbert space methods* have a great importance in the theory of linear estimation. The aim of this section is to present a short recapitulation of some basic properties of Hilbert spaces, especially the RKHS.

Let $K(\cdot, \cdot)$ be a symmetric, nonnegative definite real function, defined on a set $E \times E \subset E^2$. Then $K(\cdot, \cdot)$ generates the unique Hilbert space denoted by $H(K)$, having the following properties:

i) the elements of $H(K)$ are real function defined on E,

ii) $K(\cdot, t) \in H(K)$; $t \in E$,

iii) $f(t) = \langle f, K(\cdot, t)\rangle$; $f \in H(K)$, $t \in E$,

where $\langle\cdot, \cdot\rangle$ is the inner product in $H(K)$. For a detailed study of RKHS we refer to Aronszajn [1].

An isomorphism ψ of the Hilbert space H_1 with the inner product $\langle\cdot, \cdot\rangle_1$ onto an Hilbert space H_2 with the inner product $\langle\cdot, \cdot\rangle_2$ is a unitary operator, i.e. a one-to-one inner product preserving linear mapping of H_1 onto H_2:

i) $\psi : H_1 \xrightarrow{\text{onto}} H_2$,

ii) $\psi(\alpha f_1 + \beta f_2) = \alpha\, \psi(f_1) + \beta\, \psi(f_2)$; $\alpha, \beta \in E^1$; $f_1, f_2 \in H_1$,

iii) $\langle\psi(f_1), \psi(f_2)\rangle_2 = \langle f_1, f_2\rangle_1$; $f_1, f_2 \in H_1$.

We then say that H_1 and H_2 are *isomorphic Hilbert spaces*.

Lemma 2.1. Let H_1, H_2 be two Hilbert spaces spanned by the sets $G_1 \subset H_1$ and $G_2 \subset H_2$. Let ψ be a one-to-one mapping of G_1 onto G_2 preserving the inner product, i.t.

$$\langle\psi_1(f_1), \psi(f_2)\rangle_2 = \langle f_1, f_2\rangle_1 \; ; \quad f_1, f_2 \in G_1 \,.$$

Then ψ is an isomorphism of H_1 onto H_2.

The proof is given in [8].

Lemma 2.2. If the normed space X is reflexive, it is also weakly complete.

The proof of this lemma may be found in [2].

3. GENERAL THEORY OF UNBIASED LINEAR ESTIMATION

In this section we shall give a Hilbert space description of linear estimation following Parzen [8; 9].

Let the observed random process $Y = \{Y(t), t \in T\}$ and the random variable Z have the properties described in section 1.

The following lemmas are very important for the estimation theory.

Lemma 3.1. Let the observed random process $Y = \{Y(t), t \in T\}$ have an unknown mean value function $m(t) = \mathsf{E}_m X(t)$; $t \in T$ belonging to a given subset M of $H(R_Y)$, where $R_Y(\cdot, \cdot)$ is the known covariance function of the process Y. Then there exists the mapping

$$\langle Y, \cdot \rangle : H(R_Y) \xrightarrow{\text{onto}} L_2\{Y(t), t \in T\}$$

with the following properties:

i) $\langle Y, R_Y(\cdot, t) \rangle = Y(t)$; $t \in T$,

ii) $E_m\langle Y, g \rangle = \langle m, g \rangle_{H(R_Y)}$; $m \in M$, $g \in H(R_Y)$,

iii) $\operatorname{Cov}\{\langle Y, g \rangle, \langle Y, h \rangle\} = \langle g, h \rangle_{H(R_Y)}$; $g, h \in H(R_Y)$.

Proof. Parzen [8; 9] or with greater detail in Pázman [10].

Lemma 3.2. Let Z be the random variable with a finite dispersion. Then for the function $\varrho_Z(t) = \operatorname{Cov}\{Z, Y(t)\}$; $t \in T$ there holds:

i) $\varrho_Z(\cdot) \in H(R_Y)$,

ii) $\operatorname{Cov}\{Z, \langle Y, g \rangle\} = \langle \varrho_Z, g \rangle_{H(R_Y)}$; $g \in H(R_Y)$.

Proof. Parzen [8; 9].

According to Lemma 3.1 we may reformulate definition 1.1 as follows.

Definition 3.1. The random variable Z with the finite dispersion $\mathsf{D}^2 Z$ *is estimable* iff there exists $g \in H(R_Y)$ such that

$$\mathsf{E}_m\langle Y, g \rangle = \langle m, g \rangle = \mathsf{E}_m Z \quad \text{for all} \quad m \in M\,.$$

Remark. If we denote by $f_Z(m) = \mathsf{E}_m Z$; $m \in M$, then Z is estimable iff $f_Z(\cdot)$ is a continuous linear functional on $\overline{M}$ – the subspace of $H(R_Y)$ spanned by the elements of the set $M \subset H(R_Y)$.

The well-known theorem on linear estimation, the proof of which may be found in [8; 9] is the following:

Theorem 3.1. Let $Y = \{Y(t); t \in T\}$ is an observed random process with $\mathsf{E}_m Y(t) = m(t)$; $t \in T$, where $m(\cdot) \in M \subset H(R_Y)$. Let Z be an estimable random variable with a known finite dispersion and with the known function $\varrho_Z(t) = \operatorname{Cov}\{Z, Y(t)\}$; $t \in T$. Then there exists the BLUE $\tilde{Z}$ of Z,

$$\tilde{Z} = \langle Y, \varrho_Z \rangle + \langle Y, E^*[z - \varrho_z \mid \overline{M}] \rangle\,,$$

where $z \in H(R_Y)$ is any element such that $f_Z(m) = \langle m, z \rangle_{H(R_Y)}$; $m \in \overline{M}$ and $E^*[\cdot \mid \overline{M}]$ is the projection in $H(R_Y)$ on the subspace $\overline{M}$ of $H(R_Y)$. For the mean square error of estimation we have:

$$\mathsf{E}[Z - \tilde{Z}]^2 = \mathsf{D}^2 Z - \|\varrho_Z\|^2_{H(R_Y)} + \|E^*[z - \varrho_Z \mid \overline{M}]\|^2_{H(R_Y)} .$$

We shall now apply this general theory to the problems of the linear filtration and prediction with filtration of indirectly observed random processes.

4. LINEAR ESTIMATION OF INDIRECTLY OBSERVED RANDOM PROCESSES

Now we shall study the problems of linear filtration and the problems of prediction with filtration for the model of the *indirectly observed random process* $X = \{X(t), t \in T\}$. Let

$$Y = \left\{Y(t) = \int_T a(t, s)\, X(s)\, \mathrm{d}s + Z(t);\ t \in T\right\} \tag{4.1}$$

be the observed random process. It is assumed that the covariance functions $R_X(\cdot, \cdot)$ and $R_Z(\cdot, \cdot)$ of the independent random processes X and Z exist and are known. Let $\mathsf{E}_m\, X(t) = m(t);\ t \in T,\ m(\cdot) \in M,\ M \subset H(R_X)$ and $\mathsf{E}\, Z(t) \equiv 0;\ t \in T$. From observations of the process Y we have to find (if it exists) the BLUE of $X(t_0)$ for $t_0 \in T$ (filtration) and for $t_0 \notin T$ (prediction with filtration).

Let for the function $a(\cdot, \cdot)$ defined on $T \times T$

$$\int_T \int_T a^2(s, t)\, \mathrm{d}t\, \mathrm{d}s < \infty \quad \text{hold}\,. \tag{4.2}$$

Let for $R_X(\cdot, \cdot)$ the relation (4.2) be also true. From the above assumptions we have: for all $m \in M$ the mean value function $n(\cdot)$ and the covariance function $R_Y(\cdot, \cdot)$ of the random process Y exist and

$$\mathsf{E}_m\, Y(t) = n(t) = \int_T a(t, s)\, m(s)\, \mathrm{d}s\,; \quad t \in T, \tag{4.3}$$

$$R_Y(s, t) = \int_T \int_T a(s, u)\, R_X(u, v)\, a(v, t)\, \mathrm{d}u\, \mathrm{d}v + R_Z(s, t)\,; \quad s, t \in T. \tag{4.4}$$

$R_X(\cdot, \cdot)$ is the kernel of the integral operator, say R_X, defined on the space $L_2(T) = \{f(\cdot) : f : T \to E^1, \int_T f^2(t)\, \mathrm{d}t < \infty\}$ by

$$R_X f(t) = \int_T R_X(t, s) f(s)\, \mathrm{d}s\,; \quad t \in T, \quad f \in L_2(T) \tag{4.5}$$

with a range in $L_2(T)$. Let A be the integral operator defined on $L_2(T)$ by (4.5) with the kernel $a(., .)$.

We will show now that for the integral operator $R_{Y'}$ defined by (4.5) with the covariance kernel $R_{Y'}(s, t) = \operatorname{Cov}\{Y'(s), Y'(t)\}$; $s, t \in T$ of the process

$$Y' = \left\{Y'(t) = \int_T a(t, s) X(s) \, ds; \ t \in T\right\} \tag{4.6}$$

$$R_{Y'} = A R_X A^* \quad \text{holds}, \tag{4.7}$$

A^* being the adjoint of A. Indeed: let $f \in L_2(T)$, then from (4.4) we have:

$$R_{Y'} f(t) = \int_T R_{Y'}(t, s) f(s) \, ds = \int_T \int_T \int_T a(s, u) \, R_X(u, v) \, a(t, v) f(s) \, du \, dv \, ds \, .$$

On the other hand from the definition of $A R_X A^*$ we have:

$$A R_X A^* f(t) = A R_X \int_T a^*(\cdot, s) f(s) \, ds(t) = A R_X \int_T a(s, \cdot) f(s) \, ds(t) =$$

$$= A \int_T R_X(\cdot, u) \int_T a(s, u) f(s) \, ds \, du(t) =$$

$$= \int_T a(t, v) \int_T R_X(u, v) \int_T a(s, u) f(s) \, ds \, dv \, du = R_{Y'} f(t) \, .$$

For the generating function $R_{Y'}(\cdot, t)$; $t \in T$ of the Hilbert space $H(R_{Y'})$ using (4.7) we may write:

$$R_{Y'}(\cdot, t) = \int_T a(\cdot, u) \int_T R_X(u, v) \, a(t, v) \, du \, dv = A \, g_t(\cdot) \, ; \quad t \in T, \tag{4.8}$$

where

$$g_t(u) = \int_T R_X(u, v) \, a(t, v) \, dv = R_X \, a(t, \cdot)(u) \, ; \quad u \in T; \quad t \in T. \tag{4.9}$$

Let $G = \{g_t(\cdot); \, t \in T\}$. We will show that $G \subset H(R_X)$ and for all $s, t \in T$

$$\begin{aligned} \langle g_s, g_t \rangle_{H(R_X)} &= R_{Y'}(s, t) = \langle R_{Y'}(\cdot, s), R_{Y'}(\cdot, t) \rangle_{H(R_{Y'})} = \\ &= \langle A g_s, A g_t \rangle_{H(R_{Y'})} \quad \text{holds}. \end{aligned} \tag{4.10}$$

Indeed: let $t \in T$; then $g_t(u) = \lim_{n \to \infty} \sum_{i=1}^{n} R_X(v_i, u) \, a(t, v_i)(v_{i+1} - v_i) = \lim_{n \to \infty} g_{n,t}(u)$ for all $u \in T$, where $g_{n,t}(\cdot) \in H(R_X)$; $n = 1, 2, \ldots$ $\{g_{n,t}\}_{n=1}^{\infty}$ is a Cauchy sequence in $H(R_X)$:

$$\|g_{nt} - g_{mt}\|_{H(R_X)} = \sum_{i,j=1}^{n} a(t, v_i) \, a(t, v_j) \, R_X(v_i, v_j) \, .$$

$$\cdot (v_{i+1} - v_i)(v_{j+1} - v_j) + \sum_{i,j=1}^{m} a(t, u_i)\, a(t, u_j)\, R_X(u_i, u_j)\,.$$

$$\cdot (u_{i+1} - u_i)(u_{j+1} - u_j) - 2\sum_{i=1}^{n}\sum_{j=1}^{m} a(t, v_i)\, a(t, u_j)\, R_X(v_i, u_j)\,.$$

$$\cdot (v_{i+1} - v_i)(u_{j+1} - u_j) \xrightarrow[n,m\to\infty]{} 0\,,$$

and so it converges to the function $h_t(\cdot) \in H(R_X)$. Further $|g_{nt}(u) - h_t(u)| = |\langle g_{nt} - h_t, R_X(\cdot, u)\rangle_{H(R_X)}| \leqq \|R_X(\cdot, u)\|_{H(R_X)} \cdot \|h_t - g_{nt}\|_{H(R_X)}$ for all $u \in T$ and $h_t(\cdot) = \lim_{n\to\infty} g_{n,t}(\cdot) = g_t(\cdot) \in H(R_X)$ for all $t \in T$. Next, for all $s, t \in T$

$$\langle g_s, g_t\rangle_{H(R_X)} = \lim_{n,m\to\infty} \langle g_{ns}, g_{mt}\rangle_{H(R_X)} =$$

$$= \int_T\int_T a(s, u)\, R_X(u, v)\, a(t, v)\, \mathrm{d}u\, \mathrm{d}v = R_{Y'}(s, t)$$

and we have proved (4.10). Let H_G *be the subspace of* $H(R_X)$ *spanned by the elements of the set* $G \subset H(R_X)$. According to Lemma 2.1 we have:

Lemma 4.1. $H(R_{Y'}) = AH_G$. The integral operator A defined by (4.5) is an isomorphism of H_G onto $H(R_{Y'})$.

Solving the problem of linear estimation of the random variable $X(t_0)$ for $t_0 \in T$ (or for $t_0 \notin T$), we have to distinguish the two important cases:

$$H_G = H(R_X) \quad \text{and} \tag{4.11}$$

$$H_G \neq H(R_X)\,. \tag{4.12}$$

(4.11) is true if the integral operator $A : H(R_X) \xrightarrow{\text{onto}} H(R_{Y'})$ (see Lemma 4.2 bellow) has its inverse: $Ag = 0$ iff $g = 0$; $g \in H(R_X)$.

Let us first examine the case (4.11). The random variable $X(t_0)$, $t_0 \in T$ *is estimable* from observations of the process Y given by (4.1). Indeed: let $N = AM$; then $N \subset H(R_{Y'})$ and (as shown in Aronszajn [1]) $N \subset H(R_Y)$. We have to show that $f_{t_0}(n) = \mathsf{E}_n X(t_0) = \mathsf{E}_{Am} X(t_0) = m(t_0)$ is a linear continuous functional on the subspace $\overline{N}$ of $H(R_Y) : m(t_0) = f_{t_0}(n) = \langle Am, h_{t_0}\rangle_{H(R_Y)}$ for all $m \in \overline{M}$ and some $h_{t_0} \in H(R_Y)$. Since $n \in \overline{N}$ implies $n \in H(R_Y)$ the following relations are true:

$$n(t) = \langle n, R_{Y'}(\cdot, t)\rangle_{H(R_{Y'})} = \langle n, R_Y(\cdot, t)\rangle_{H(R_Y)} \tag{4.13}$$

and from this we may conclude: to any $g \in L\{R_{Y'}(\cdot, t);\ t \in T\}$ there exists an $h \in L\{R_Y(\cdot, t);\ t \in T\}$ such that

$$\langle n, g\rangle_{H(R_{Y'})} = \langle n, h\rangle_{H(R_Y)}\,. \tag{4.14}$$

 The following equalities are obvious:

$$f_{t_0}(n) = m(t_0) = \langle m, R_X(\cdot, t_0)\rangle_{H(R_X)} = \langle Am, AR_X(\cdot, t_0)\rangle_{H(R_{Y'})} =$$
$$= \langle n, AR_X(\cdot, t_0)\rangle_{H(R_{Y'})} = \lim_{k\to\infty} \langle n, g_k\rangle_{H(R_{Y'})} = \lim_{k\to\infty} \langle n, h_k\rangle_{H(R_Y)},$$

where $g_k \in L\{R_{Y'}(\cdot, t); t \in T\}$; $k = 1, 2, \ldots$ are such that $\lim_{k\to\infty} \|g_k - AR_X(\cdot, t_0)\|_{H(R_{Y'})} = 0$ (we have used (4.14)). Therefore, applying Lemma 2.2, we obtain that the Cauchy sequence $\{h_k\}_{k=1}^{\infty}$ weakly converges in $H(R_Y)$ to some $h_{t_0} \in H(R_Y)$, and so we have

$$f_{t_0}(n) = m(t_0) = \lim_{k\to\infty} \langle n, h_k\rangle_{H(R_Y)} = \langle n, h_{t_0}\rangle_{H(R_Y)} \tag{4.15}$$

for all $n \in \bar{N}$ and the random variable $X(t_0)$, $t_0 \in T$ is estimable. The BLUE is of the form

$$\tilde{X}(t_0) = \langle Y, \varrho_{t_0}\rangle + \langle Y, E^*_{H(R_Y)}[h_{t_0} - \varrho_{t_0} \mid \bar{N}]\rangle, \quad t_0 \in T, \tag{4.16}$$

where $\varrho_{t_0}(t) = \operatorname{Cov}\{Y(t), X(t_0)\} = AR_X(\cdot, t_0)(t)$; $t \in T$.

Now we shall look for the BLUE of the random variable $X(t_0)$, $t_0 \in T$ *for the regression model*. It will be assumed that $M = \bar{M}$ is the of all linear combinations of q known linearly independent functions $\varphi_1(\cdot), \ldots, \varphi_q(\cdot) \in H(R_X)$. Then for all t in T and for $t_0 \in T$, also $m(t) = \sum_{i=1}^{q} \beta_i \varphi_i(t)$ for some coefficients $\beta_1, \ldots, \beta_q$. In this case $f_{t_0}(n) = m(t_0) = \langle Am, h_{t_0}\rangle_{H(R_Y)}$, where $h_{t_0} \in N = AM \subset H(R_Y)$. Indeed: let $h_{t_0} = \sum_{j=1}^{q} \alpha_j \psi_j$, where $\psi_j = A\varphi_j$; $j = 1, 2, \ldots, q$. Then for the vector $\beta = (\beta_1, \ldots, \beta_q)'$ we have the condition:

$$m(t_0) = \sum_{i=1}^{q} \beta_i \varphi_i(t_0) = \langle \sum_{i=1}^{q} \beta_i \psi_i(\cdot), \sum_{j=1}^{q} \alpha_j \psi_j(\cdot)\rangle_{H(R_Y)}$$

for all $\beta = (\beta_1, \ldots, \beta_q)'$. This condition may be written in the form

$$\varphi(t_0) = \mathbf{F}\alpha,$$

where $\varphi(t_0) = (\varphi_1(t_0), \ldots, \varphi_q(t_0))'$ and $\mathbf{F}$ is the matrix $\mathbf{F} = \|\langle \psi_i, \psi_j\rangle_{H(R_Y)}\|_{i,j=1}$.

From this we have for $h_{t_0}(\cdot)$:

$$h_{t_0}(t) = \varphi'(t_0)\mathbf{F}^{-1}\psi(t), \quad \text{where} \quad \psi(t) = (\psi_1(t), \ldots, \psi_q(t))'.$$
$$\varrho_{t_0}(t) = AR_X(\cdot, t_0)(t) \quad \text{and} \quad \varrho^*_{t_0} = E^*_{H(R_Y)}[\varrho_{t_0} \mid N] =$$
$$= \langle AR_X(\cdot, t_0), \psi\rangle'_{H(R_Y)} \mathbf{F}^{-1}\psi(t).$$

Using Theorem 3.1 we have proved the following:

Theorem 4.1. Let us observe the process Y given by (4.1) and let (4.11) hold. Then for the random variable $X(t_0)$, $t_0 \in T$ the BLUE $\tilde{X}(t_0)$ *exists* and is given by (4.16). *For the regression model* of the mean value of X the BLUE $\tilde{X}(t_0)$ of $X(t_0)$ exists for $t_0 \in T$ and $t_0 \notin T$ and $\tilde{X}(t_0)$ is given by

$$(4.17) \quad \tilde{X}(t_0) = \langle Y, AR_X(\cdot, t_0)\rangle + (\varphi(t_0) - \langle AR_X(\cdot, t_0), \psi\rangle_{H(R_Y)})' \mathbf{F}^{-1}\langle Y, \psi\rangle .$$

For the mean square error of the estimation we have

$$(4.18) \qquad \mathsf{E}[X(t_0) - \tilde{X}(t_0)]^2 = \mathsf{D}^2 X(t_0) - \|AR_X(\cdot, t_0)\|^2_{H(R_Y)} + \\ + (\varphi(t_0) - \langle AR_X(\cdot, t_0), \psi\rangle_{H(R_Y)})' \mathbf{F}^{-1}(\varphi(t_0) - \langle AR_X(\cdot, t_0), \psi\rangle_{H(R_Y)}) .$$

From the point of view of the calculation of $\tilde{X}(t_0)$, the finite dimensional approximation $\tilde{X}_n(t_0)$ of $\tilde{X}(t_0)$ is very important. We may calculate $\tilde{X}_n(t_0)$ from the data given by the random vector $Y_n = (Y(t_1), \ldots, Y(t_n))'$; $t_i \in T$, $i = 1, 2, \ldots, n$ according to the formula (in the case of the regression model)

$$(4.19) \qquad \tilde{X}_n(t_0) = A\, R_X(t_0)' \,(\mathbf{AR}_X\mathbf{A}^* + \mathbf{R}_Z)^{\#}\, Y_n + \\ + [\varphi'(t_0) - A\, R_X(t_0)' \,(\mathbf{AR}_X\mathbf{A}^* + \mathbf{R}_Z)^{\#}\, \mathbf{T}^*] \,.\, [\mathbf{T}(\mathbf{AR}_X\mathbf{A}^* + \mathbf{R}_Z)^{\#}\, \mathbf{T}^*]^{-1} \,. \\ .\mathbf{T}(\mathbf{AR}_X\mathbf{A}^* + \mathbf{R}_Z)^{\#}\, Y_n \,,$$

where $\mathbf{AR}_X\mathbf{A}^*$ is the symbol for the matrix $\|AR_XA^*(t_i, t_j)\|_{i,j=1}^n$, $\mathbf{R}_Z = \|R_Z(t_i, t_j)\|_{i,j=1}^n$ and $\#$ is used for denotation of the generalized inverse of a matrix, $AR_X(t_0) = (AR_X(\cdot, t_0)(t_1), \ldots, AR_X(\cdot, t_0)(t_n))'$ and

$$\mathbf{T} = \begin{Vmatrix} \psi_1(t_1), & \ldots, & \psi_1(t_n) \\ \vdots & & \vdots \\ \psi_q(t_1), & \ldots, & \psi_q(t_n) \end{Vmatrix} .$$

It is proved in [11] that for $R_Y(\cdot, \cdot)$ continuous by some conditions on $t_i \in T$; $i = 1, 2, \ldots, n$ $\lim_{n\to\infty} \mathsf{E}[\tilde{X}_n(t_0) - \tilde{X}(t_0)]^2 = 0$.

Now we shall investigate the second important case of the problem of linear estimation of the indirectly observed random process X, namely the case (4.12) when $H_G \neq H(R_X)$. H_G is now a proper subspace of $H(R_X)$. From the projection theorem we have for all $g \in H(R_X) : g = g^* + (g - g^*)$, where $g^* = E^*_{H(R_X)}[g \mid H_G]$. In the following we need the following lemma:

Lemma 4.2. $Ag = 0$; $g \in H(R_X)$ iff $g \perp H_G$.

Proof: Is evident from the following equalities:

$$A\, g(t) = \int_T a(t, s)\, g(s)\, \mathrm{d}s = \int_T a(t, s) \langle g, R_X(\cdot, s)\rangle_{H(R_X)}\, \mathrm{d}s = \\ = \left\langle g, \int_T a(t, s)\, R_X(\cdot, s)\, \mathrm{d}s \right\rangle_{H(R_X)} = \langle g, g_t\rangle_{H(R_X)} \quad \text{for all} \quad t \in T.$$

From this Lemma we have: $Ag = Ag^*$ for all $g \in H(R_X)$ and

$$m(t_0) = \langle m, R_X(\cdot, t_0)\rangle_{H(R_X)} =$$
$$= \langle m^*, R_X^*(\cdot, t_0)\rangle_{H(R_X)} + \langle m - m^*, R_X(\cdot, t_0) - R_X^*(\cdot, t_0)\rangle_{H(R_X)} =$$
$$= \langle Am^*, AR_X^*(\cdot, t_0)\rangle_{H(R_Y)} + \langle m - m^*, R_X(\cdot, t_0) - R_X^*(\cdot, t_0)\rangle_{H(R_X)} =$$
$$= \langle Am, h_{t_0}\rangle_{H(R_Y)} + \langle m - m^*, R_X(\cdot, t_0) - R_X^*(\cdot, t_0)\rangle_{H(R_X)}\,; \quad t_0 \in T$$

and we see that the random variable $X(t_0)$, $t_0 \in T$ may not have a linear unbiased estimate, because $\langle m - m^*, R_X(\cdot, t_0) - R_X^*(\cdot, t_0)\rangle_{H(R_X)}$ is not a continuous linear functional on N if $M \not\subset H_G$. By Theorem 3.1 and from the decomposition $R_X(\cdot, t) = R_X^*(\cdot, t) + (R_X(\cdot, t) - R_X^*(\cdot, t))$ we have:

$$X(t) = \langle X, R_X^*(\cdot, t)\rangle + \langle X, R_X(\cdot, t) - R_X^*(\cdot, t)\rangle = X^*(t) + (X(t) - X^*(t))$$

for all $t \in T$. The random variable $X^*(t_0)$, $t_0 \in T$ has the BLUE, because $\mathsf{E}_m X^*(t) = m^*(t_0) = \langle m^*, R_X(\cdot, t_0)\rangle_{H(R_X)} = \langle Am^*, h_{t_0}\rangle_{H(R_Y)}$ is a continuous linear functional on $\bar{N} \subset H(R_Y)$ and $\operatorname{Cov}\{X^*(t_0), Y(t)\} = \langle AR_X^*(\cdot, t_0), R_Y(\cdot, t)\rangle_{H(R_Y)} = AR_X(\cdot, t_0)(t) = \operatorname{Cov}\{X(t_0), Y(t)\}$. The random variable $X(t_0) - X^*(t_0)$, $t_0 \in T$ has not a linear unbiased estimate, since $\operatorname{Cov}\{X(t_0) - X^*(t_0), Y(t)\} = 0$; $t \in T$ and its mean value function $m(t) - m^*(t)$; $t \in T$ is not a linear continuous functional on $\bar{N}$. Thus from Theorem 3.1 we have:

Theorem 4.2. If we observe the random process Y given by (4.1) and if (4.12) is true, then *the random variable* $X(t_0)$, $t_0 \in T$ *has not in general* LUE. The best linear estimate $\tilde{X}^*(t_0)$ of the random variable $X(t_0)$, $t \in T$ is given by

$$\tilde{X}^*(t_0) = \langle Y, AR_X(\cdot, t_0)\rangle_{H(R_Y)} + \langle Y, E^*_{H(R_Y)}[h_{t_0} - AR_X(\cdot, t_0) \mid \bar{N}]\rangle\,, \tag{4.20}$$

$$\mathsf{E}_m \tilde{X}^*(t_0) = m^*(t_0) \tag{4.21}$$

and

$$\mathsf{E}_m[X(t_0) - \tilde{X}^*(t_0)]^2 = \mathsf{D}^2 X(t_0) - \|AR_X(\cdot, t_0)\|^2_{H(R_Y)} + \tag{4.22}$$
$$+ \|E^*_{H(R_Y)}[h_{t_0} - AR_X(\cdot, t_0) \mid \bar{N}]\|^2_{H(R_Y)} + |m(t_0) - m^*(t_0)|^2 \quad \text{holds}\,.$$

Proof: It is necessary to prove (4.22) only. But by Lemmas 3.1 and 3.2

$$\mathsf{E}_m[X(t_0) - \tilde{X}^*(t_0)]^2 = \mathsf{D}^2 X(t_0) + \mathsf{E}_m[\tilde{X}^*(t_0) - m^*(t_0)]^2 -$$
$$- 2\,\mathsf{E}_m[X(t_0) - m(t_0)]\,[\tilde{X}^*(t_0) - m^*(t_0)] + |m(t_0) - m^*(t_0)|^2 =$$
$$= \mathsf{D}^2 X(t_0) + \|AR_X(\cdot, t_0) + E^*_{H(R_Y)}[h_{t_0} - AR_X(\cdot, t_0) \mid \bar{N}]\|^2_{H(R_Y)} -$$
$$- 2\langle AR_X(\cdot, t_0), AR_X(\cdot, t_0) + E^*_{H(R_Y)}[h_{t_0} - AR_X(\cdot, t_0) \mid \bar{N}]\rangle_{H(R_Y)} +$$
$$+ |m(t_0) - m^*(t_0)|^2\,.$$

Remark. If (4.12) holds, but if $M \subset H_G$, then the estimate given by (4.20) is the BLUE. If $M \not\subset H_G$, we see from (4.21) that the BLE of $X(t_0)$, $t_0 \in T$ is not unbiased; its bias is equal to $|m(t_0) - m^*(t_0)|$ and the mean square error of estimation is given by (4.22). The value $|m(t_0) - m^*(t_0)|$ of the bias is not linearly estimable from observations of the random process Y.

5. COMPARISON OF KALMAN-BUCY'S AND PARZEN'S METHODS OF LINEAR FILTRATION

First of all we shall now briefly explain the Kalman-Bucy method of filtration following the book of Lipcer and Shirjajev [7]. Let (X, Y) be a two-dimensional Gaussian random process defined for $t \in [0, T]$; (X, Y) fulfils the stochastic differential equations

$$\mathrm{d}X(t) = a(t)\, X(t)\, \mathrm{d}t + b(t)\, \mathrm{d}W_1(t) \quad \text{and} \tag{5.1}$$

$$\mathrm{d}Y(t) = A(t)\, X(t)\, \mathrm{d}t + B(t)\, \mathrm{d}W_2(t)\,; \quad t \in [0, T]\,. \tag{5.2}$$

W_1, W_2 are independent Wiener processes. For the deterministic functions $a(\cdot)$, $A(\cdot)$, $b(\cdot)$ and $B(\cdot)$ it is assumed that

$$\int_0^T |a(t)|\, \mathrm{d}t < \infty\,, \quad \int_0^T b^2(t)\, \mathrm{d}t < \infty\,, \quad \int_0^T |A(t)|\, \mathrm{d}t < \infty\,, \quad \int_0^T B^2(t)\, \mathrm{d}t < \infty\,.$$

Equation (5.1) has one continuous solution given by

$$X(t) = \exp\left\{\int_0^t a(u)\, \mathrm{d}u\right\}\left[X_0 + \int_0^t \exp\left\{-\int_0^s a(u)\, \mathrm{d}u\right\} b(s)\, \mathrm{d}W_1(s)\right]; \quad 0 \leqq t \leqq T \tag{5.3}$$

and the equation (5.2) may be written in the form

$$Y(t) = Y_0 + \int_0^t A(s)\, X(s)\, \mathrm{d}s + \int_0^t B(s)\, \mathrm{d}W_2(s)\,; \quad 0 \leqq t \leqq T. \tag{5.4}$$

The aim is, on the basis of observations of the process $Y_0^t = \{Y(s), 0 \leqq s \leqq t\}$ to find the BLUE of the random variable $X(t)$; $t \in [0, T]$. Kalman and Bucy [5] proved that if $\int_0^T A^2(t)\, \mathrm{d}t < \infty$; $B^2(t) \geqq c > 0$; $0 \leqq t \leqq T$, then the BLUE $\tilde{X}(t)$ exists and they derived a stochastic differential equation for $\tilde{X}(t)$. We now show that *the model given by Kalman and Bucy is a special case of ours* modified a little the theory of linear estimation of the indirectly observed random process X.

Let $t \in [0, T]$ be fixed. Denote by $R_X^t(\cdot, \cdot)$ the restriction of the function $R_X(\cdot, \cdot)$ defined on $[0, T] \times [0, T]$ on $[0, t] \times [0, t]$. Then $R_X^t(\cdot, \cdot)$ is a symmetric, non-negative definite function generating $H(R_X^t)$.

 From (5.3) we have:

$$m^t(s) = \mathsf{E}X(s) = \mathsf{E}X_0 \,.\, \exp\left[\int_0^s a(u)\,\mathrm{d}u\right]; \quad 0 \leqq s \leqq t \quad \text{and} \tag{5.5}$$

$$R_X^t(s_1, s_2) = \exp\left[\int_0^{s_1} a(u)\,\mathrm{d}u + \int_0^{s_2} a(u)\,\mathrm{d}u\right] . \tag{5.6}$$

$$. \left[\mathsf{D}^2 X_0 + \int_0^{s_1 \wedge s_2} \exp\left\{-2\int_0^u (av)\,\mathrm{d}v\right\} b^2(u)\,\mathrm{d}u\right];$$

If $\mathsf{D}^2 X_0 > 0$, then $m^t(\cdot) = (\mathsf{E}X_0/\mathsf{D}^2 X_0)\, R_X^t(\cdot, 0) \in H(R_X^t)$. If $\mathsf{D}^2 X_0 = 0$ and if $\mathsf{E}X_0 \neq 0$, then $m^t(\cdot) \notin H(R_X^t)$, because $m(0) \neq 0$ and $h(\cdot) \in H(R_X^t)$ implies $h(0) = 0$ for R_X^t given by (5.6). But if $\mathsf{D}^2 X_0 = 0$, $X(t) = m^t(t) + V(t)$ according to (5.3), where $m^t(\cdot)$ is the known deterministic function and only

$$V(t) = \exp\left\{\int_0^t a(u)\,\mathrm{d}u\right\} \int_0^t \exp\left\{-\int_0^v a(u)\,\mathrm{d}u\right\} b(v)\,\mathrm{d}W_1(v)$$

have to be estimated from Y_0^t.

Now let us correct the theory of the section 4 to fit it with the Kalman-Bucy model. From (5.4) we get the form of the covariance function of the random process

$$Y_0'^t = \left\{Y'(s) = \int_0^s A(u)\, X(u)\,\mathrm{d}u; \quad 0 \leqq s \leqq t\right\};$$

$$R_{Y'}^t(s_1, s_2) = \int_0^{s_1}\int_0^{s_2} A(u)\, R_X(u, v)\, A(v)\,\mathrm{d}u\,\mathrm{d}v = A\, g_{s_1}(s_2)\,,$$

where $g_s(u) = \int_0^s R_X(u, v)\, A(v)\,\mathrm{d}v$; $0 \leqq u \leqq t$ and the operator A is now defined on $L_2(0, t)$ by $A\,f(s) = \int_0^s A(u) f(u)\,\mathrm{d}u$; $0 \leqq s \leqq t$. Let H_G^t be the subspace of $H(R_X^t)$ spanned by the elements of the set $G^t = \{g_s(\cdot);\ 0 \leqq s \leqq t\}$. Now for the BLUE $\tilde{X}(t)$ of $X(t)$; $t \in [0, T]$ we may use Theorem 4.1 or 4.2, because the results of section 4 are valid (with obvious modifications caused by the other definition of the operator A).

Remark 1. The Parzen method of calculation of BLUE has not a "recursive" form by altering $t \in [0, T]$.

Remark 2. If $\mathsf{D}^2 X_0 > 0$, it is possible to apply Theorem 4.1 to the case when $\mathsf{E}X(0)$ is unknown, too: indeed, from (5.5) we see that M^t – the space of the mean value functions of $X_0^t = \{X(s);\ 0 \leqq s \leqq t\}$ is a one-dimensional subspace of $H(R_X^t)$ spanned by the function $m^t(s) = \exp\left(\int_0^s a(u)\,\mathrm{d}u\right)$; $0 \leqq s \leqq t$. The operator A is one-to-one on M^t if $A \not\equiv 0$. The last fact follows from the definition of the operator given in this section and from the fact that $m^t(s) > 0$; $0 \leqq s \leqq t$.

(Received December 29, 1975.)

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

[2] G. Bachman, N. Narici: Functional Analysis. Academic Press, New York 1966.

[3] J. Hájek: On linear Statistical Problems in Stochastic Processes. Czech. Math. J. *12* (1962), 404—444.

[4] P. Halmos: Introduction to Hilbert Spaces. Chelsea, New York 1951.

[5] R. E. Kalman, R. S. Bucy: New Results in Linear Filtering and Prediction Theory. J. Basic Engrg. *83* (1961), 95—108.

[6] A. Kolmogorov: Interpolation and Extrapolation. Bull. Acad. Sc. URSS Ser. Math. (1941), 3—14.

[7] Р. Ш. Липцер, А. Н. Ширяев: Статистика случайных процессов. Наука, Москва 1974.

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

[9] E. Parzen: A New Approach to the Synthesis of Optimal Smoothing and Prediction Systems. Math. Optim. Techniques (1963), 75—108.

[10] A. Pázman: The Ordering of Experimental Design. Kybernetika *10* (1974), 373—388.

[11] F. Štulajter: Lineárna filtrácia a predikcia náhodných procesov. Report, Bratislava 1975.

[12] N. Wiener: The Extrapolation, Interpolation and Smoothing of Stationary Time Series. Wiley, New York 1949.

RNDr. František Štulajter, CSc, Ústav teórie merania (Institute for Measurement Theory — Slovak Academy of Sciences), Dúbravská cesta, 885 27 Bratislava. Czechoslovakia.