KYBERNETIKA — VOLUME 25 (1989), NUMBER 6

ON THE EQUIVALENCE OF TWO METHODS FOR INTERPOLATION

PETR BUDINSKÝ

Two methods for interpolation in stationary discrete processes are presented in the paper. First, the method proposed by Brubacher and Wilson, second Jaglom's method. There is given a proof in the paper that both methods give the same results.

1. INTRODUCTION AND PRELIMINARIES

Let $\{Y_t\}$ be a white noise with $\mathsf{E}Y_t = 0$, $\mathsf{E}Y_t^2 = \sigma^2$, $t = \ldots, -1, 0, 1, \ldots$, $\mathsf{E}Y_sY_t = 0$, $s \neq t$. Let π_j be real numbers satisfying the condition $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and assume that there exists a linear stationary discrete process $\{X_t\}$ given by

$$Y_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}\,. \tag{1.1}$$

Assume throughout the paper that the variables X_{s+t_i}, $i = 0, 1, \ldots, n$ $(t_0 = 0)$ are missing. Very important problem is to find the best linear interpolation of X_{s+t_i}, $i = 0, 1, \ldots, n$. Brubacher and Wilson proposed in [2] to minimize the sum of squares of Y_t with respect to the unknown variables X_{s+t_i}. The method given by Jaglom (see [3] and [4]) is based on the projection in the Hilbert space. It will be proved in the paper that both methods give the same results.

It is well known that the spectral density $f(\lambda)$ of $\{X_t\}$ given by (1.1) has the form

$$f(\lambda) = (\sigma^2/2\pi)\,\Big|\sum_{j=0}^{\infty} \pi_j\, \mathrm{e}^{-\mathrm{i}j\lambda}\Big|^{-2}\,. \tag{1.2}$$

Introduce numbers p_k by

$$p_k = \begin{cases} \sum_{j=0}^{\infty} \pi_j \pi_{j+k}\,, & k = 0, 1, 2, \ldots, \\ p_{-k}\,, & k = -1, -2, \ldots. \end{cases} \tag{1.3}$$

Introduce an $(n + 1) \times (n + 1)$ symmetric matrix

$$\boldsymbol{P} = \|p_{t_i - t_j}\|_{i,j=0,1,\ldots,n} .$$

Let $\boldsymbol{P}$ be regular.

2. BRUBACHER-WILSON'S METHOD

Brubacher and Wilson propose in [2] to minimize

$$Z_m = \sum_{t=-m}^{m} Y_t^2 = \sum_{t=-m}^{m} \left(\sum_{j=0}^{\infty} \pi_j X_{t-j}\right)^2$$

with respect to the unknown variables X_{s+t_i}, $i = 0, 1, \ldots, n$. Using

$$\frac{\partial Z_m}{\partial X_{s+t_i}} = 2 \sum_{t=s+t_i}^{m} \pi_{t-s-t_i} \left(\sum_{j=0}^{\infty} \pi_j X_{t-j}\right) = 0$$

we have for $m \to \infty$

$$\sum_{k=-\infty}^{\infty} p_k X_{s+t_i+k} = 0 , \quad i = 0, 1, \ldots, n . \tag{2.1}$$

Let $\tilde{X}_{s+t_i}$ be a solution of the linear equations (2.1).

Denote $V_i = - \sum_{\substack{l=-\infty \\ l \neq t_0, t_1, \ldots, t_n}}^{\infty} p_{l-t_i} X_{s+l}$. Then (2.1) can be written in the form

$$\sum_{j=0}^{n} p_{t_j - t_i} \tilde{X}_{s+t_j} = V_i , \quad i = 0, 1, \ldots, n .$$

Remark 1. For $t_i = i$, $i = 0, 1, \ldots, n$ we obtain

$$\sum_{j=0}^{n} p_{j-i} X_{s+j} = - \sum_{j \notin [0,n]} p_{j-i} X_{s+j} \tag{2.3}$$

which for $n = 0$ can be simplified to

$$\tilde{X}_s = - (1/p_0) \sum_{j \neq 0} p_j X_{s+j} .$$

3. JAGLOM'S METHOD

Let H be the Hilbert space generated by the variables $\{X_t\}_{t=-\infty}^{\infty}$. Let $K = \{s + t_0, s + t_1, \ldots, s + t_n\}$, $0 = t_0 < t_1 < \ldots < t_n$ and let H_K be the subspace of H generated by the variables X_k, $k \notin K$. The best linear interpolation $\hat{X}_{s+t_j}$ of X_{s+t_j} based on X_k, $k \notin K$, is defined as the projection of X_{s+t_j} onto H_K. Let the projection be given by

$$\hat{X}_{s+t_j} = \sum_{\substack{k=-\infty \\ k \neq t_0 - t_j, \ldots, t_n - t_j}}^{\infty} a_k X_{s+t_j+k} .$$

Define

$$\Phi_j(\lambda) = \sum_{\substack{k=-\infty \\ k \neq t_0 - t_j, \dots, t_n - t_j}}^{\infty} a_k \, \mathrm{e}^{\mathrm{i}k\lambda}, \quad -\pi \leqq \lambda \leqq \pi$$

and

$$\Psi_j(\lambda) = (1 - \Phi_j(\lambda)) f(\lambda), \quad -\pi \leqq \lambda \leqq \pi. \tag{3.1}$$

The function $\Phi_j(\lambda)$ is called the spectral characteristic for interpolation. Since $f(\lambda)$ has the form (1.2) we can introduce the functions

$$f^*(\mathrm{e}^{\mathrm{i}\lambda}) = f(\lambda), \quad \Phi_j^*(\mathrm{e}^{\mathrm{i}\lambda}) = \Phi_j(\lambda),$$
$$\Psi_j^*(\mathrm{e}^{\mathrm{i}\lambda}) = \Psi_j(\lambda).$$

Let $U = \{z; |z| < 1\}$, $\bar{U} = \{z; |z| \leqq 1\}$.

Theorem 1. Let $\Phi_j^*(z) = \Omega_{j,0}^*(z) + \Omega_{j,1}^*(z) + \dots + \Omega_{j,n+1}^*(z)$, $j = 0, 1, \dots, n$, be functions of the complex variable satisfying the following conditions:

a) $\Omega_{j,0}^*(z)$ and $\Omega_{j,n+1}^*(z)$ are analytic functions on U^C and $\bar{U}$, respectively, and

$$\Omega_{j,i}^*(z) = \sum_{k=t_{i-1}-t_j+1}^{t_i - t_j - 1} a_k z^k, \quad i = 1, 2, \dots, n;$$

b) $\lim\limits_{z \to \infty} z^{t_j} \Omega_{j,0}^*(z) = \lim\limits_{z \to 0} z^{t_j - t_n} \Omega_{j,n+1}^*(z) = 0$;

c) function $\Psi_j^*(z)$ can be expressed in the form

$$\Psi_j^*(z) = \sum_{k=0}^{n} c_k z^{t_k - t_j}, \quad c_k \in \mathbb{R}. \tag{3.2}$$

Then the function $\Phi_j(\lambda) = \Phi_j^*(\mathrm{e}^{\mathrm{i}\lambda})$ is the spectral characteristic for interpolation of X_{s+t_j} $(j = 0, 1, \dots n)$ based on $\{X_{s+t_j+k}\}$, $k \neq t_0 - t_j, \dots, t_n - t_j$.

Proof. We use a method similar to that given in [1] for the case of extrapolation. From a) and b) we have that $\Omega_{j,0}^*(z)$ and $\Omega_{j,n+1}^*(z)$ can be expressed in the form

$$\Omega_{j,0}^*(z) = \sum_{k=-\infty}^{t_0 - t_j - 1} a_k z^k \quad (z \in U^C), \quad \Omega_{j,n+1}^*(z) = \sum_{k=t_n - t_j + 1}^{\infty} a_k z^k \quad (z \in \bar{U}).$$

Then there exist $d \in (0, 1)$ and $d' \in (1, \infty)$ that both sums converge for $d < |z| < d'$. Thus $\sum_{k=-\infty}^{t_0 - t_j - 1} a_k \, \mathrm{e}^{\mathrm{i}k\lambda}$ and $\sum_{k=t_n - t_j + 1}^{\infty} a_k \, \mathrm{e}^{\mathrm{i}k\lambda}$ converge in the quadratic mean with respect to $f(\lambda)$.

Denote

$$\hat{X}_{s+t_j}^{(0)} = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(s+t_j)\lambda} \sum_{k=-\infty}^{t_0 - t_j - 1} a_k \, \mathrm{e}^{\mathrm{i}k\lambda} \, \mathrm{d}Z(\lambda),$$

$$\hat{X}_{s+t_j}^{(n+1)} = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(s+t_j)\lambda} \sum_{k=t_n - t_j + 1}^{\infty} a_k \, \mathrm{e}^{\mathrm{i}k\lambda} \, \mathrm{d}Z(\lambda),$$

where $Z(\lambda)$ is the random measure corresponding to $\{X_t\}$ (see [1]). Then we obtain

$$\hat{X}_{s+t_j}^{(0)} = \underset{N \to \infty}{\mathrm{l.i.m.}} \sum_{k=-N}^{t_0 - t_j - 1} a_k X_{s+t_j+k},$$

$$\hat{X}_{s+t_j}^{(n+1)} = \underset{N \to \infty}{\mathrm{l.i.m.}} \sum_{k=t_n - t_j + 1}^{N} a_k X_{s+t_j+k}.$$

Denote

$$\hat{X}^{(i)}_{s+t_j} = \int_{-\pi}^{\pi} e^{i(s+t_j)\lambda}\, \Omega_{j,i}(\lambda)\, dZ(\lambda) = \sum_{k=t_{i-1}-t_j+1}^{t_i-t_j-1} a_k X_{s+t_j+k}\,, \quad i = 1, 2, \ldots, n\,.$$

Then

$$\hat{X}^{(0)}_{s+t_j} \in H_K\,, \quad \hat{X}^{(n+1)}_{s+t_j} \in H_K \quad \text{and} \quad \hat{X}^{(i)}_{s+t_j} \in H_K\,, \quad i = 1, 2, \ldots, n\,.$$

Thus

$$\hat{X}_{s+t_j} = \sum_{i=0}^{n+1} \hat{X}^{(i)}_{s+t_j} \in H_K\,,$$

$$(X_{s+t_j} - \hat{X}_{s+t_j}, X_{s+k}) = \mathsf{E} \int_{-\pi}^{\pi} e^{i(s+t_j)\lambda} (1 - \Phi_j(\lambda))\, dZ(\lambda)\, \overline{\int_{-\pi}^{\pi} e^{i(s+k)\lambda}\, dZ(\lambda)} =$$

$$= \int_{-\pi}^{\pi} e^{i(t_j-k)\lambda}\, \Psi_j(\lambda)\, d\lambda = \int_{-\pi}^{\pi} \sum_{l=0}^{n} c_l\, e^{i(t_l-k)\lambda}\, d\lambda\,.$$

Hence

$$(X_{s+t_j} - \hat{X}_{s+t_j}, X_{s+k}) = 0 \quad \text{for} \quad k \neq t_0, t_1, \ldots, t_n\,.$$

This implies $(X_{s+t_j} - \hat{X}_{s+t_j}) \perp H_K$. Thus we have proved that $\hat{X}_{s+t_j}$ is the projection of X_{s+t_j} onto H_K. □

Remark 2. In the special case when $K = \{s\}$ we have

$$\Psi_0^*(z) = (1 - \Phi^*(z)) f^*(z) = c_0 \quad \text{where } c_0 \text{ is a real constant}\,.$$

Remark 3. For finding the best linear interpolation $\hat{X}_{s+t_j}$ it is sufficient to determine numbers $c_0, c_1, \ldots, c_n$ (depending on j) in (3.2) and then to express $\Phi_j^*(z)$ from (3.1) in the form

$$\Phi_j^*(z) = 1 - (1/f^*(z)) \sum_{k=0}^{n} c_k z^{t_k - t_j}\,. \tag{3.3}$$

Especially if $K = \{s\}$ then

$$\Phi_0^*(z) = 1 - (c_0/f^*(z))\,.$$

Theorem 2. Define a vector $\mathbf{e}_j = (\delta_{0,j}, \ldots, \delta_{n,j})'$, where $\delta_{k,j}$ is Kronecker's δ. Then

$$\hat{X}_{s+t_j} = - \sum_{\substack{l=-\infty \\ l \neq 0, t_1, \ldots, t_n}}^{\infty} X_{s+l}(c_0^* p_l + c_1^* p_{l-t_1} + \ldots + c_n^* p_{l-t_n})\,, \tag{3.4}$$

where $\mathbf{c}^* = (c_0^*, c_1^*, \ldots, c_n^*)'$ is a solution of the equations

$$\mathbf{Pc}^* = \mathbf{e}_j\,, \quad j = 0, 1, \ldots, n\,. \tag{3.5}$$

Proof. We can write $f(z) = (\sigma^2/2\pi)/\sum_{l=-\infty}^{\infty} p_l z^l$.

Using (3.3) we get further

$$\Phi_j^*(z) = 1 - \sum_{l=-\infty}^{\infty} p_l z^l \sum_{k=0}^{n} c_k^* z^{t_k - t_j} = \tag{3.6}$$

$$= 1 - \sum_{l=-\infty}^{\infty} \sum_{k=0}^{n} p_l c_k^* z^{l + t_k - t_j}\,,$$

where $c_k^* = (2\pi/\sigma^2)\, c_k$.

To fulfil the conditions from Theorem 1 the coefficient by $z^{t_i - t_j}$ must be equal to $\delta_{i,j}$. For $k = m$ $(m = 0, 1, \ldots, n)$ and $l = t_i - t_m$ we have $z^{l + t_k - t_j} = z^{t_i - t_j}$ and the coefficient standing by $z^{t_i - t_j}$ is equal to $\sum_{k=0}^{n} c_k^* p_{t_i - t_k}$. Hence we have the linear equations $\sum_{k=0}^{n} c_k^* p_{t_i - t_k} = \delta_{i,j}$ which are equivalent to (3.5). But the formula (3.6) can be written in the form

$$\Phi_j^*(z) = - \sum_{\substack{l=-\infty \\ l \neq t_0 - t_j, \ldots, t_n - t_j}}^{\infty} \sum_{k=0}^{n} p_{l - t_k + t_j} c_k z^l$$

and thus

$$\hat{X}_{s+t_j} = \sum_{\substack{l=-\infty \\ l \neq t_0 - t_j, \ldots, t_n - t_j}}^{\infty} \sum_{k=0}^{n} p_{l - t_k + t_j} c_k^* X_{s+t_j+l} .$$

Substituting $l' = t_j + l$ we obtain (3.4). □

Remark 4. If $K = \{s\}$ we have

$$\hat{X}_s = -(1/p_0) \sum_{k=1}^{\infty} p_k (X_{s+k} + X_{s-k}) .$$

Denote P_{ij}^* the algebraic complement of p_{ij} in the matrix $\mathbf{P}$. Then

$$\hat{X}_{s+t_j} = -(1/\det \mathbf{P}) \sum_{\substack{l=-\infty \\ l \neq 0, t_1, \ldots, t_n}}^{\infty} X_{s+l} \sum_{k=0}^{n} P_{jk}^* p_{l-t_k} . \tag{3.7}$$

Remark 5. Especially for $t_i = i$, $i = 0, 1, \ldots. n$ we get

$$\hat{X}_{s+t_j} = -(1/\det \mathbf{P}) \left[\sum_{l=1}^{\infty} X_{s+n+l} \sum_{k=0}^{n} P_{jk}^* p_{l+n-k} + \sum_{l=1}^{\infty} X_{s-l} \sum_{k=0}^{n} P_{jk}^* p_{l+k} \right] .$$

4. COMPARISON OF BOTH METHODS

In this section we use the notation from the previous sections. The following theorem is the main result of our paper.

Theorem 4. Let $K = \{s + t_0, \ldots, s + t_n\}$. Then $\tilde{X}_{s+t_j} = \hat{X}_{s+t_j}$, $j = 0, 1, \ldots, n$.

Proof. Denote $\tilde{\mathbf{X}} = (\tilde{X}_{s+t_0}, \ldots, \tilde{X}_{s+t_n})'$ and $\mathbf{V} = (V_0, \ldots, V_n)'$. Using the notation from Section 2 we can write (2.2) in the form $\mathbf{P}\tilde{\mathbf{X}} = \mathbf{V}$. Thus $\tilde{\mathbf{X}} = \mathbf{P}^{-1}\mathbf{V}$. From here we obtain

$$\tilde{X}_{s+t_j} = (1/\det \mathbf{P}) \sum_{i=0}^{n} P_{ji}^* V_i .$$

Hence

$$\tilde{X}_{s+t_j} = -(1/\det \mathbf{P}) \sum_{\substack{l=-\infty \\ l \neq t_0, t_1, \ldots, t_n}}^{\infty} X_{s+l} \sum_{i=0}^{n} P_{ji}^* p_{l-t_i} ,$$

which corresponds to (3.7). Thus $\tilde{X}_{s+t_j} = \hat{X}_{s+t_j}$. □

(Received February 13, 1989.)

REFERENCES

[1] J. Anděl: Statistical Analysis of Time Series (in Czech). SNTL, Prague 1976.

[2] S. R. Brubacher and T. G. Wilson: Interpolating time series with application to the estimation of holiday effects on electricity demand. Appl. Statist. *25* (1976), 107—116.

[3] A. M. Jaglom: Vvedĕnije v tĕoriju stacionarnych slučajnych funkcij. Usp. mat. nauk *7* (1952), vyp. 5 (51), 3—168.

[4] Ju. A. Rozanov: Stacionarnyje slučajnyje processy. Gos. izd., Moskva 1963.

RNDr. Petr Budinský, katedra matematiky a konstruktivní geometrie strojní fakulty ČVUT (Department of Mathematics and Constructive Geometry, Faculty of Mechanical Engineering — Czech Technical University), Karlovo nám. 13, 121 35 Praha 2. Czechoslovakia.