



# Extrapolation Problem for Stationary Sequences with Missing Observations

Mikhail Moklyachuk \*, Maria Sidei,

*Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, Ukraine*

(Received: 5 March 2017; Accepted: 20 June 2017)

**Abstract** In this paper, we consider the problem of the mean square optimal estimation of linear functionals which depend on unknown values of a stationary stochastic sequence based on observations of the sequence with a stationary noise. Formulas for calculating the mean-square error and the spectral characteristic of the optimal linear estimate of the functional are derived under the condition of spectral certainty, where spectral densities of the sequences are exactly known. The minimax (robust) method of estimation is applied in the case of spectral uncertainty, where spectral densities of the sequences are not known exactly while sets of admissible spectral densities are given. Formulas that determine the least favorable spectral densities and the minimax spectral characteristics are proposed for some special sets of admissible densities.

**Keywords** Stationary sequence, mean square error, minimax-robust estimate, least favorable spectral density, minimax spectral characteristic

**AMS 2010 subject classifications.** Primary: 60G10, 60G25, 60G35, Secondary: 62M20, 93E10, 93E11

**DOI:** 10.19139/soic.v5i3.284

## 1. Introduction

The problem of estimation of unknown values of stochastic processes is of great interest both in the theory of stochastic processes and applications of this theory to the data analysis in such fields of science as oceanography, meteorology, astronomy, radio physics, statistical hydromechanics etc. Efficient solution methods of estimation problems (interpolation, extrapolation and filtering) of stationary sequences were developed by Kolmogorov (see selected works by Kolmogorov [15]). Constructive methods of solution of the estimation problems for stationary stochastic processes were proposed by Wiener [43] and Yaglom [44, 45]. Detailed description and further development of the methods can be found in the books by Rozanov [41], Hannan [10], Brockwell and Davis [3], Pourahmadi [37].

The crucial assumption of most of the methods of estimation of the unobserved values of stochastic processes is that the spectral densities of the stochastic processes considered are exactly known. However, in practice, complete information on the spectral densities is impossible in most cases. To solve the problem in this case one usually finds parametric or nonparametric estimates of the unknown spectral densities and then, under assumption that the estimated densities are the true ones, one applies the traditional estimation methods. This procedure can result in significant increasing of the value of error as Vastola and Poor [42] have demonstrated with the help of some examples. To avoid this effect it is reasonable to search estimates which are optimal for all densities from a certain class of admissible spectral densities. These estimates are called minimax since they minimize the maximum value of the error. This method was first proposed in the paper by Grenander [9] where this approach was applied to extrapolation problem for stationary processes. In the survey paper by Kassam and Poor [14] the minimax-robust

---

\*Correspondence to: Mikhail Moklyachuk (Email: Moklyachuk@gmail.com). Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, Volodymyrska 64 Str., Kyiv 01601, Ukraine.

methods of data processing are described and several models of spectral uncertainty are analysed. A method of solution of problems of minimax extrapolation and interpolation of stationary sequences which is based on convex optimization methods is proposed in the works by Franke [5], Franke and Poor [6]. This approach makes it possible to find equations that determine the least favorable spectral densities for different classes of densities.

In the papers by Moklyachuk [23]-[27] the problem of optimal estimation of the functionals which depend on the unknown values of stationary processes were investigated. Solution methods of interpolation, extrapolation and filtering problems for periodically correlated stochastic processes are described in the book by Moklyachuk and Golichenko [28]. The corresponding estimation problems for vector-valued stationary processes are described in the book by Moklyachuk and Masyutka [29]. Estimation problems for functionals which depend on the unknown values of stochastic sequences with stationary increments were investigated by Luz and Moklyachuk [17]-[22]. The problems of estimation of stationary sequences and processes with missing values were investigated by Moklyachuk and Sidei [30]-[35]. Prediction problems for stationary processes with missing observations were investigated in works by Cheng et al. [4], Bondon [1, 2], Kasahara et al. [13], Pourahmadi et al. [38], Pelagatti [36].

In this paper we present results of investigation of the problem of the mean-square optimal estimation of the linear functional  $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$  which depends on the unknown values of a stationary sequence  $\{\xi(j), j \in \mathbb{Z}\}$  from observations of the sequence  $\xi(j) + \eta(j)$  at points  $j \in \mathbb{Z}_- \setminus S = \{\dots, -2, -1\} \setminus S$ ,  $S = \bigcup_{l=1}^s \{-M_l - N_l, -M_l - N_l + 1, \dots, -M_l\}$ ,  $M_l = \sum_{k=0}^l (N_k + K_k)$ ,  $N_0 = 0$ ,  $K_0 = 0$ . The problem is investigated in the case of spectral certainty where the spectral densities of the sequences  $\{\xi(j), j \in \mathbb{Z}\}$  and  $\{\eta(j), j \in \mathbb{Z}\}$  are exactly known and in the case of spectral uncertainty where the spectral densities are not exactly known while a set of admissible spectral densities is given.

## 2. Hilbert space projection method of extrapolation

Let  $\{\xi(j), j \in \mathbb{Z}\}$  and  $\{\eta(j), j \in \mathbb{Z}\}$  be stationary stochastic sequences with zero mean values:  $E\xi(j) = 0$ ,  $E\eta(j) = 0$  and correlation functions which admit the spectral decomposition (see Gikhman and Skorohod [8])

$$R_{\xi}(k) = E\xi(j+k)\overline{\xi(j)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda, \quad R_{\xi\eta}(k) = E\xi(j+k)\overline{\eta(j)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f_{\xi\eta}(\lambda) d\lambda,$$

$$R_{\eta\xi}(k) = E\eta(j+k)\overline{\xi(j)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f_{\eta\xi}(\lambda) d\lambda, \quad R_{\eta}(k) = E\eta(j+k)\overline{\eta(j)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} g(\lambda) d\lambda,$$

where  $f(\lambda)$ ,  $f_{\xi\eta}(\lambda)$ ,  $f_{\eta\xi}(\lambda)$ ,  $g(\lambda)$  are spectral densities of the stationary sequences such that the minimality condition holds true

$$\int_{-\pi}^{\pi} \frac{1}{f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)} d\lambda < \infty. \tag{1}$$

This condition guarantees that the mean-square error of the mean-square optimal estimate of the functional is nonzero (see Rozanov [41]).

Stationary sequences  $\xi(j)$  and  $\eta(j)$  admit spectral decomposition (see Gikhman and Skorohod [8]; Karhunen [12])

$$\xi(j) = \int_{-\pi}^{\pi} e^{ij\lambda} Z_{\xi}(d\lambda), \quad \eta(j) = \int_{-\pi}^{\pi} e^{ij\lambda} Z_{\eta}(d\lambda), \tag{2}$$

where  $Z_\xi(d\lambda)$  and  $Z_\eta(d\lambda)$  are orthogonal stochastic measures such that the following relations hold true

$$EZ_\xi(\Delta_1)\overline{Z_\xi(\Delta_2)} = \frac{1}{2\pi} \int_{\Delta_1 \cap \Delta_2} f(\lambda) d\lambda, \quad EZ_\xi(\Delta_1)\overline{Z_\eta(\Delta_2)} = \frac{1}{2\pi} \int_{\Delta_1 \cap \Delta_2} f_{\xi\eta}(\lambda) d\lambda,$$

$$EZ_\eta(\Delta_1)\overline{Z_\xi(\Delta_2)} = \frac{1}{2\pi} \int_{\Delta_1 \cap \Delta_2} f_{\eta\xi}(\lambda) d\lambda, \quad EZ_\eta(\Delta_1)\overline{Z_\eta(\Delta_2)} = \frac{1}{2\pi} \int_{\Delta_1 \cap \Delta_2} g(\lambda) d\lambda.$$

Consider the problem of mean-square optimal linear estimation of the functional

$$A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$$

which depends on the unknown values of the sequence  $\{\xi(j), j \in \mathbb{Z}\}$  from observations of the sequence  $\xi(j) + \eta(j)$  at points  $j \in \mathbb{Z}_- \setminus S$ , where  $S = \bigcup_{l=1}^s \{-M_l - N_l, \dots, -M_l\}$ .

Making use of the spectral decomposition (2) of the sequence  $\xi(j)$  we can represent the functional  $A\xi$  in the form

$$A\xi = \int_{-\pi}^{\pi} A(e^{i\lambda})Z_\xi(d\lambda), \quad A(e^{i\lambda}) = \sum_{j=0}^{\infty} a(j)e^{ij\lambda}.$$

We will suppose that the coefficients  $\{a(j), j = 0, 1, \dots\}$  which determine the functional  $A\xi$  are such that the following conditions

$$\sum_{k=0}^{\infty} |a(k)| < \infty, \quad \sum_{k=0}^{\infty} (k+1) |a(k)|^2 < \infty \quad (3)$$

are satisfied. The first condition ensures that the functional  $A\xi$  has a finite second moment since  $E|A\xi|^2 \leq R_\xi(0) \left( \sum_{k=0}^{\infty} |a(k)| \right)^2$ . The second condition ensures the compactness in  $\ell_2$  of the operators that will be defined below.

Denote by  $\hat{A}\xi$  the optimal linear estimate of the functional  $A\xi$  from the known observations of the sequence  $\xi(j) + \eta(j)$  at points  $j \in \mathbb{Z}_- \setminus S$ . Since the spectral densities of the stationary sequences  $\xi(j)$  and  $\eta(j)$  are known, we can use the Hilbert space method proposed by A. N. Kolmogorov (see selected works by Kolmogorov [15]) to find the estimate  $\hat{A}\xi$ .

Consider values  $\xi(j)$  and  $\eta(j)$  of the sequences as elements of the Hilbert space  $H = L_2(\Omega, \mathcal{F}, P)$  generated by random variables  $\xi$  with 0 mathematical expectations,  $E\xi = 0$ , finite variations,  $E|\xi|^2 < \infty$ , and inner product  $(\xi, \eta) = E\xi\bar{\eta}$ . Denote by  $H^s(\xi + \eta)$  the closed linear subspace generated by elements  $\{\xi(j) + \eta(j) : j \in \mathbb{Z}_- \setminus S\}$  in the Hilbert space  $H = L_2(\Omega, \mathcal{F}, P)$ . Let  $L_2(f + g)$  be the Hilbert space of complex-valued functions that are square-integrable with respect to the measure whose density is  $f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)$ . Denote by  $L_2^s(f + g)$  the subspace of  $L_2(f + g)$  generated by functions  $\{e^{ij\lambda}, j \in \mathbb{Z}_- \setminus S\}$ .

The mean-square optimal linear estimate  $\hat{A}\xi$  of the functional  $A\xi$  from observations of the sequence  $\xi(j) + \eta(j)$  is of the form

$$\hat{A}\xi = \int_{-\pi}^{\pi} h(e^{i\lambda})(Z_\xi(d\lambda) + Z_\eta(d\lambda)), \quad (4)$$

where  $h(e^{i\lambda}) \in L_2^s(f + g)$  is the spectral characteristic of the estimate.

The mean-square error of the estimate  $\hat{A}\xi$  is given by the formula

$$\begin{aligned} \Delta(h; F) &= E \left| A\xi - \hat{A}\xi \right|^2 = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |A(e^{i\lambda}) - h(e^{i\lambda})|^2 f(\lambda) d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} |h(e^{i\lambda})|^2 g(\lambda) d\lambda \\ &\quad - \frac{1}{2\pi} \int_{-\pi}^{\pi} (A(e^{i\lambda}) - h(e^{i\lambda})) \overline{h(e^{i\lambda})} f_{\xi\eta}(\lambda) d\lambda - \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{(A(e^{i\lambda}) - h(e^{i\lambda}))} h(e^{i\lambda}) f_{\eta\xi}(\lambda) d\lambda \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (A(e^{i\lambda}) - h(e^{i\lambda}), -h(e^{i\lambda})) F(\lambda) \left( \overline{A(e^{i\lambda}) - h(e^{i\lambda})}, \overline{-h(e^{i\lambda})} \right)^T d\lambda, \end{aligned} \tag{5}$$

where  $F(\lambda) = \begin{pmatrix} f(\lambda) & f_{\xi\eta}(\lambda) \\ f_{\eta\xi}(\lambda) & g(\lambda) \end{pmatrix}$  is the spectral density matrix.

According to the Hilbert space orthogonal projection method the optimal linear estimate of the functional  $A\xi$  is a projection of the element  $A\xi$  of the subspace  $H$  on the space  $H^s(\xi + \eta)$ . The projection is determined by the following conditions:

- 1)  $\hat{A}\xi \in H^s(\xi + \eta)$ ,
- 2)  $A\xi - \hat{A}\xi \perp H^s(\xi + \eta)$ .

It follows from the second condition that the spectral characteristic  $h(e^{i\lambda})$  of the optimal linear estimate  $\hat{A}\xi$  for any  $j \in \mathbb{Z}_- \setminus S$  satisfies equations

$$\begin{aligned} E \left[ (A\xi - \hat{A}\xi) \left( \overline{\xi(j)} + \overline{\eta(j)} \right) \right] &= \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (A(e^{i\lambda}) - h(e^{i\lambda})) e^{-ij\lambda} f(\lambda) d\lambda - \frac{1}{2\pi} \int_{-\pi}^{\pi} h(e^{i\lambda}) e^{-ij\lambda} f_{\eta\xi}(\lambda) d\lambda + \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} (A(e^{i\lambda}) - h(e^{i\lambda})) e^{-ij\lambda} f_{\xi\eta}(\lambda) d\lambda - \frac{1}{2\pi} \int_{-\pi}^{\pi} h(e^{i\lambda}) e^{-ij\lambda} g(\lambda) d\lambda = 0. \end{aligned}$$

The last relation can be written in the form

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [A(e^{i\lambda})(f(\lambda) + f_{\xi\eta}(\lambda)) - h(e^{i\lambda})(f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda))] e^{-ij\lambda} d\lambda = 0, \quad j \in \mathbb{Z}_- \setminus S.$$

Hence the function  $[A(e^{i\lambda})(f(\lambda) + f_{\xi\eta}(\lambda)) - h(e^{i\lambda})(f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda))]$  is of the form

$$\begin{aligned} A(e^{i\lambda})(f(\lambda) + f_{\xi\eta}(\lambda)) - h(e^{i\lambda})(f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)) &= C(e^{i\lambda}), \\ C(e^{i\lambda}) &= \sum_{j \in T} c(j) e^{ij\lambda}, \end{aligned}$$

where  $T = S \cup \{0, 1, \dots\}$ , and  $c(j), j \in T$  are unknown coefficients that should be determined.

From the last relation we deduce that the spectral characteristic of the optimal linear estimate  $\hat{A}\xi$  is of the form

$$h(e^{i\lambda}) = A(e^{i\lambda}) \frac{f(\lambda) + f_{\xi\eta}(\lambda)}{f_{\zeta}(\lambda)} - C(e^{i\lambda}) \frac{1}{f_{\zeta}(\lambda)}, \tag{6}$$

where  $f_{\zeta}(\lambda) = f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)$ .

From the first condition,  $\hat{A}\xi \in H^s(\xi + \eta)$ , which determines the optimal estimate of the functional  $A\xi$ , it follows that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} h(e^{i\lambda}) e^{-ij\lambda} d\lambda = 0, \quad j \in T,$$

namely

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left( A(e^{i\lambda}) \frac{f(\lambda) + f_{\xi\eta}(\lambda)}{f_{\zeta}(\lambda)} - C(e^{i\lambda}) \frac{1}{f_{\zeta}(\lambda)} \right) e^{-ij\lambda} d\lambda = 0, \quad j \in T.$$

Disclose brackets and write the last equation in the form

$$\sum_{k=0}^{\infty} a(k) \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i(k-j)\lambda} (f(\lambda) + f_{\xi\eta}(\lambda))}{f_{\zeta}(\lambda)} d\lambda - \sum_{l \in T} c(l) \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i(l-j)\lambda}}{f_{\zeta}(\lambda)} d\lambda = 0, \quad j \in T. \tag{7}$$

Let us introduce the Fourier coefficients of the functions

$$\begin{aligned} b_{k-j} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k-j)\lambda} \frac{1}{f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)} d\lambda; \\ r_{k-j} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k-j)\lambda} \frac{f(\lambda) + f_{\xi\eta}(\lambda)}{f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)} d\lambda; \\ q_{k-j} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k-j)\lambda} \frac{f(\lambda)g(\lambda) - f_{\xi\eta}(\lambda)f_{\eta\xi}(\lambda)}{f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)} d\lambda. \end{aligned} \tag{8}$$

Denote by  $\vec{a} = (0, 0, \dots, 0, \vec{a})$  vector with zero first  $|S| = \sum_{k=1}^s (N_k + 1)$  components, and the last component  $\vec{a} = (a(0), a(1) \dots)$  is constructed from coefficients which define the functional  $A\xi$ .

Now we can represent relation (7) in the form

$$\mathbf{R}\vec{a} = \mathbf{B}\vec{c}, \tag{9}$$

where  $\vec{c}$  is the vector constructed from the unknown coefficients  $c(k), k \in T$ , and the linear operator  $\mathbf{B}$  in the space  $\ell_2$  is defined by the matrix

$$B = \begin{pmatrix} B_{s,s} & B_{s,s-1} & \dots & B_{s,1} & B_{s,n} \\ B_{s-1,s} & B_{s-1,s-1} & \dots & B_{s-1,1} & B_{s-1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{1,s} & B_{1,s-1} & \dots & B_{1,1} & B_{1,n} \\ B_{n,s} & B_{n,s-1} & \dots & B_{n,1} & B_{n,n} \end{pmatrix},$$

where elements in the last column and the last row are the matrices with the elements

$$\begin{aligned} B_{l,n}(k, j) &= b_{k-j}, \quad l = 1, 2, \dots, s; k = -M_l - N_l, \dots, -M_l; j = 0, 1, 2, \dots, \\ B_{n,m}(k, j) &= b_{k-j}, \quad m = 1, 2, \dots, s; k = 0, 1, 2, \dots; j = -M_m - N_m, \dots, -M_m, \\ B_{n,n}(k, j) &= b_{k-j}, \quad k, j = 0, 1, 2, \dots, \end{aligned}$$

and other elements of matrix  $B$  are matrices with elements of the form

$$B_{l,m}(j, k) = b_{k-j}, \quad l, m = 1, 2, \dots, s; k = -M_l - N_l, \dots, -M_l; j = -M_m - N_m, \dots, -M_m,$$

The linear operator  $\mathbf{R}$  in the space  $\ell_2$  is defined by the corresponding matrix in the same manner.

The unknown coefficients  $c(k), k \in T$ , which are determined by equation (9) can be calculated by the formula

$$c(k) = (\mathbf{B}^{-1}\mathbf{R}\vec{\mathbf{a}})_k,$$

where  $(\mathbf{B}^{-1}\mathbf{R}\vec{\mathbf{a}})_k$  is the  $k$  component of the vector  $\mathbf{B}^{-1}\mathbf{R}\vec{\mathbf{a}}$ . Hence the spectral characteristic  $h(e^{i\lambda})$  of the estimate  $\hat{A}\xi$  can be calculated by the formula

$$h(e^{i\lambda}) = A(e^{i\lambda}) \frac{f(\lambda) + f_{\xi\eta}(\lambda)}{f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)} - \frac{\sum_{k \in T} (\mathbf{B}^{-1}\mathbf{R}\vec{\mathbf{a}})_k e^{ik\lambda}}{f(\lambda) + f_{\xi\eta}(\lambda) + f_{\eta\xi}(\lambda) + g(\lambda)}. \tag{10}$$

The mean-square error of the estimate  $\hat{A}\xi$  can be calculated by the formula (5) which can be represented in the form

$$\begin{aligned} \Delta(h; F) &= \Delta(h; f, g, f_{\xi\eta}, f_{\eta\xi}) = E \left| A\xi - \hat{A}\xi \right|^2 = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| A(e^{i\lambda}) \frac{f_{\eta\xi}(\lambda) + g(\lambda)}{f_{\zeta}(\lambda)} + C(e^{i\lambda}) \frac{1}{f_{\zeta}(\lambda)} \right|^2 f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| A(e^{i\lambda}) \frac{f(\lambda) + f_{\xi\eta}(\lambda)}{f_{\zeta}(\lambda)} - C(e^{i\lambda}) \frac{1}{f_{\zeta}(\lambda)} \right|^2 g(\lambda) d\lambda \\ &- \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( A(e^{i\lambda}) \frac{f_{\eta\xi}(\lambda) + g(\lambda)}{f_{\zeta}(\lambda)} + C(e^{i\lambda}) \frac{1}{f_{\zeta}(\lambda)} \right) \left( \overline{A(e^{i\lambda})} \frac{f(\lambda) + f_{\eta\xi}(\lambda)}{f_{\zeta}(\lambda)} - \overline{C(e^{i\lambda})} \frac{1}{f_{\zeta}(\lambda)} \right) f_{\xi\eta}(\lambda) d\lambda \\ &- \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \overline{A(e^{i\lambda})} \frac{f_{\xi\eta}(\lambda) + g(\lambda)}{f_{\zeta}(\lambda)} + \overline{C(e^{i\lambda})} \frac{1}{f_{\zeta}(\lambda)} \right) \left( A(e^{i\lambda}) \frac{f(\lambda) + f_{\xi\eta}(\lambda)}{f_{\zeta}(\lambda)} - C(e^{i\lambda}) \frac{1}{f_{\zeta}(\lambda)} \right) f_{\eta\xi}(\lambda) d\lambda \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |A(e^{i\lambda})|^2 \frac{f(\lambda)g(\lambda) - f_{\xi\eta}(\lambda)f_{\eta\xi}(\lambda)}{f_{\zeta}(\lambda)} d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} |C(e^{i\lambda})|^2 \frac{1}{f_{\zeta}(\lambda)} d\lambda \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{j=0}^{\infty} a(j)e^{ij\lambda} \right|^2 \frac{f(\lambda)g(\lambda) - f_{\xi\eta}(\lambda)f_{\eta\xi}(\lambda)}{f_{\zeta}(\lambda)} d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k \in T} (\mathbf{B}^{-1}\mathbf{R}\vec{\mathbf{a}})_k e^{ik\lambda} \right|^2 \frac{1}{f_{\zeta}(\lambda)} d\lambda \\ &= \langle \mathbf{R}\vec{\mathbf{a}}, \mathbf{B}^{-1}\mathbf{R}\vec{\mathbf{a}} \rangle + \langle \mathbf{Q}\vec{\mathbf{a}}, \vec{\mathbf{a}} \rangle, \end{aligned} \tag{11}$$

where  $\langle a, c \rangle = \sum_k a_k \bar{c}_k$  is the inner product in the space  $\ell_2$ .

The linear operator  $\mathbf{Q}$  in the space  $\ell_2$  is defined by the corresponding matrix in the same manner as operator  $\mathbf{B}$  is defined.

Thus we obtain the following theorem.

**Theorem 2.1**

Let  $\{\xi(j), j \in \mathbb{Z}\}$  and  $\{\eta(j), j \in \mathbb{Z}\}$  be stationary stochastic sequences with the spectral density matrix  $F(\lambda) = \begin{pmatrix} f(\lambda) & f_{\xi\eta}(\lambda) \\ f_{\eta\xi}(\lambda) & g(\lambda) \end{pmatrix}$  and let the minimality condition (1) be satisfied. The spectral characteristic  $h(e^{i\lambda})$  and the

mean-square error  $\Delta(h; F)$  of the optimal linear estimate of the functional  $A\xi$  which depends on the unknown values of the sequence  $\xi(j)$  based on observations of the sequence  $\xi(j) + \eta(j)$ ,  $j \in \mathbb{Z}_- \setminus S$  can be calculated by formulas (10), (11).

The corresponding results can be obtained for the uncorrelated sequences. In this case the spectral densities  $f_{\xi\eta}(\lambda) = 0$ ,  $f_{\eta\xi}(\lambda) = 0$  and we get the following corollary.

*Corollary 2.1*

Let  $\{\xi(j), j \in \mathbb{Z}\}$  and  $\{\eta(j), j \in \mathbb{Z}\}$  be uncorrelated stationary stochastic sequences with spectral densities  $f(\lambda)$  and  $g(\lambda)$  which satisfy the minimality condition

$$\int_{-\pi}^{\pi} (f(\lambda) + g(\lambda))^{-1} d\lambda < \infty. \quad (12)$$

The spectral characteristic  $h(e^{i\lambda})$  and the mean-square error  $\Delta(f, g)$  of the optimal linear estimate of the functional  $A\xi$  which depends on unknown values of the sequence  $\xi(j)$  based on observations of the sequence  $\xi(j) + \eta(j)$ ,  $j \in \mathbb{Z}_- \setminus S$  can be calculated by the formulas

$$h(e^{i\lambda}) = A(e^{i\lambda}) \frac{f(\lambda)}{f(\lambda) + g(\lambda)} - \frac{\sum_{k \in T} (\mathbf{B}^{-1} \mathbf{R} \mathbf{a})_k e^{ik\lambda}}{f(\lambda) + g(\lambda)}, \quad (13)$$

$$\begin{aligned} \Delta(h; f, g) &= E |A\xi - \hat{A}\xi|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A(e^{i\lambda})g(\lambda) + \sum_{k \in T} (\mathbf{B}^{-1} \mathbf{R} \mathbf{a})_k e^{ik\lambda} \right|^2}{(f(\lambda) + g(\lambda))^2} f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A(e^{i\lambda})f(\lambda) - \sum_{k \in T} (\mathbf{B}^{-1} \mathbf{R} \mathbf{a})_k e^{ik\lambda} \right|^2}{(f(\lambda) + g(\lambda))^2} g(\lambda) d\lambda \\ &= \langle \mathbf{R} \mathbf{a}, \mathbf{B}^{-1} \mathbf{R} \mathbf{a} \rangle + \langle \mathbf{Q} \mathbf{a}, \mathbf{a} \rangle, \end{aligned} \quad (14)$$

where  $\mathbf{B}$ ,  $\mathbf{R}$ ,  $\mathbf{Q}$  are linear operators in the space  $\ell_2$  with elements  $b_{k-j}$ ,  $r_{k-j}$ ,  $q_{k-j}$  respectively which are defined by the Fourier coefficients of the functions

$$\begin{aligned} b_{k-j} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k-j)\lambda} \frac{1}{f(\lambda) + g(\lambda)} d\lambda; \\ r_{k-j} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k-j)\lambda} \frac{f(\lambda)}{f(\lambda) + g(\lambda)} d\lambda; \\ q_{k-j} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k-j)\lambda} \frac{f(\lambda)g(\lambda)}{f(\lambda) + g(\lambda)} d\lambda. \end{aligned} \quad (15)$$

Consider the case where stationary sequence  $\{\xi(j), j \in \mathbb{Z}\}$  is observed without noise. Since in this case  $g(\lambda) = 0$ , the spectral characteristic of the estimate  $\hat{A}\xi$  is of the form

$$\begin{aligned} h(e^{i\lambda}) &= A(e^{i\lambda}) - C(e^{i\lambda})f^{-1}(\lambda), \\ C(e^{i\lambda}) &= \sum_{j \in T} c(j)e^{ij\lambda}, \end{aligned} \quad (16)$$

and the system of equations (9) can be represented in the form

$$\vec{a} = \mathbf{B}\vec{c}, \tag{17}$$

where  $\mathbf{B}$  is the linear operator in the space  $\ell_2$  which is constructed with the help of the Fourier coefficients of the function  $(f(\lambda))^{-1}$  and is of the similar form as operators defined before.

Hence, the unknown coefficients  $c(j), j \in T$ , can be calculated by the formula

$$c(j) = (\mathbf{B}^{-1}\vec{a})_j,$$

where  $(\mathbf{B}^{-1}\vec{a})_j$  is the  $j$  component of the vector  $\mathbf{B}^{-1}\vec{a}$ , and the spectral characteristic of the estimate  $\hat{A}\xi$  is defined by the formula

$$h(e^{i\lambda}) = \left( \sum_{j=0}^{\infty} a(j)e^{ij\lambda} \right) - \left( \sum_{j \in T} (\mathbf{B}^{-1}\vec{a})_j e^{ij\lambda} \right) f^{-1}(\lambda). \tag{18}$$

The mean-square error of the estimate  $\hat{A}\xi$  is determined by the formula

$$\Delta(h; f) = \langle \mathbf{B}^{-1}\vec{a}, \vec{a} \rangle. \tag{19}$$

Let us summarize the obtained result in the form of corollary.

*Corollary 2.2*

Let  $\{\xi(j), j \in \mathbb{Z}\}$  be a stationary stochastic sequence with the spectral density  $f(\lambda)$  which satisfy the minimality condition

$$\int_{-\pi}^{\pi} f^{-1}(\lambda)d\lambda < \infty. \tag{20}$$

The spectral characteristic  $h(e^{i\lambda})$  and the mean-square error  $\Delta(h, f)$  of the optimal linear estimate  $\hat{A}\xi$  of the functional  $A\xi$  from observations of the sequence  $\xi(j)$  at points  $j \in \mathbb{Z}_- \setminus S$ , where  $S = \bigcup_{l=1}^s \{-M_l - N_l, \dots, -M_l\}$ , can be calculated by formulas (18), (19).

Let  $\xi(j)$  and  $\eta(j)$  be uncorrelated stationary sequences. Consider the problem of the mean-square optimal linear extrapolation of the functional

$$A_N\xi = \sum_{j=0}^N a(j)\xi(j)$$

which depends on unknown values of the sequence  $\xi(j)$  from observations of the sequence  $\xi(j) + \eta(j)$  at points of time  $j \in \mathbb{Z}_- \setminus S$ , where  $S = \bigcup_{l=1}^s \{-M_l - N_l, -M_l - N_l + 1, \dots, -M_l\}$ . In order to find the spectral characteristic  $h_N(e^{i\lambda})$  of the estimate

$$\hat{A}_N\xi = \int_{-\pi}^{\pi} h_N(e^{i\lambda})(Z_{\xi}(d\lambda) + Z_{\eta}(d\lambda))$$

and the mean-square error  $\Delta(h_N; f, g)$  of the estimate of the functional  $A_N\xi$ , we define the vector  $\vec{a}_N = (0, 0, \dots, 0, \vec{a}_N)$  which has zero first  $|S| = \sum_{k=1}^s (N_k + 1)$  components and the last component is  $\vec{a}_N = (a(0), a(1), \dots, a(N), 0, 0, \dots)$ .

Consider the linear operator  $\mathbf{R}_N$  in the space  $\ell_2$  which is defined as follows  $\mathbf{R}_N(k, j) = \mathbf{R}(k, j), j \leq N, \mathbf{R}_N(k, j) = 0, j > N$ .



Thus the spectral characteristic of the optimal estimation  $\hat{A}_N \xi$  can be calculated by the formula

$$h_N(e^{i\lambda}) = A_N(e^{i\lambda}) \frac{f(\lambda)}{f(\lambda) + g(\lambda)} - \frac{\sum_{k \in T} (\mathbf{B}^{-1} \mathbf{R}_N \vec{\mathbf{a}}_N)_k e^{ik\lambda}}{f(\lambda) + g(\lambda)}. \quad (21)$$

The mean-square error of the estimate  $\hat{A}_N \xi$  is defined by the formula

$$\begin{aligned} \Delta(h_N; f, g) &= E \left| A_N \xi - \hat{A}_N \xi \right|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{A_N(e^{i\lambda})g(\lambda) + \sum_{k \in T} (\mathbf{B}^{-1} \mathbf{R}_N \vec{\mathbf{a}}_N)_k e^{ik\lambda}}{(f(\lambda) + g(\lambda))^2} \right|^2 f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{A_N(e^{i\lambda})f(\lambda) - \sum_{k \in T} (\mathbf{B}^{-1} \mathbf{R}_N \vec{\mathbf{a}}_N)_k e^{ik\lambda}}{(f(\lambda) + g(\lambda))^2} \right|^2 g(\lambda) d\lambda \\ &= \langle \mathbf{R}_N \vec{\mathbf{a}}_N, \mathbf{B}^{-1} \mathbf{R}_N \vec{\mathbf{a}}_N \rangle + \langle \mathbf{Q}_N \vec{\mathbf{a}}_N, \vec{\mathbf{a}}_N \rangle, \end{aligned} \quad (22)$$

where  $\mathbf{Q}_N$  is the linear operator in the space  $\ell_2$ ,  $\mathbf{Q}_N(k, j) = \mathbf{Q}(k, j)$ ,  $k, j \leq N$ ,  $\mathbf{Q}_N(k, j) = 0$ ,  $k, j > N$ . Notice that linear operators  $\mathbf{B}$ ,  $\mathbf{R}$ ,  $\mathbf{Q}$  are defined in Corollary 2.1.

#### Corollary 2.3

Let  $\xi(j)$ ,  $\eta(j)$  be uncorrelated stationary stochastic sequences with spectral densities  $f(\lambda)$  and  $g(\lambda)$  which satisfy the minimality condition (1). The spectral characteristic  $h_N(e^{i\lambda})$  and the mean-square error  $\Delta(h_N; f, g)$  of the optimal linear estimate of the functional  $A_N \xi$  which depends on unknown values of the sequence  $\xi(j)$  based on observations of the sequence  $\xi(j) + \eta(j)$ ,  $j \in \mathbb{Z}_- \setminus S$  can be calculated by formulas (21), (22).

In the case where the sequence is observed without noise we have the following corollary.

#### Corollary 2.4

Let  $\xi(j)$  be a stationary stochastic sequence with the spectral density  $f(\lambda)$  which satisfy the minimality condition (20). The spectral characteristic  $h_N(e^{i\lambda})$  and the mean-square error  $\Delta(h_N, f)$  of the optimal linear estimate  $\hat{A}_N \xi$  of the functional  $A_N \xi$  can be calculated by the formulas (23), (24)

$$h_N(e^{i\lambda}) = \left( \sum_{j=0}^N a(j) e^{ij\lambda} \right) - \left( \sum_{j \in T} (\mathbf{B}^{-1} \vec{\mathbf{a}}_N)_j e^{ij\lambda} \right) f^{-1}(\lambda), \quad (23)$$

$$\Delta(h_N; f) = \langle \mathbf{B}^{-1} \vec{\mathbf{a}}_N, \vec{\mathbf{a}}_N \rangle. \quad (24)$$

The linear operator  $\mathbf{B}$  is defined in Corollary 2.2.

In order to demonstrate the developed techniques we propose the following example.

#### Example 2.1

Consider the problem of the optimal linear estimation of the functional

$$A_1 \xi = a(0) \xi(0) + a(1) \xi(1)$$

which depends on the unknown values of a stationary sequence  $\xi(j)$  from observations of the sequence  $\xi(j)$  at points  $j \in \mathbb{Z}_- \setminus S$ , where  $S = \{-3, -2\}$ . Let the spectral density of the sequence  $\xi(j)$  be of the form  $f(\lambda) = |1 - \alpha e^{-i\lambda}|^{-2}$ . In this case the function

$$f^{-1}(\lambda) = |1 - \alpha e^{-i\lambda}|^2 = b_{-1} e^{-i\lambda} + b_0 + b_1 e^{i\lambda},$$

where  $b_0 = 1 + |\alpha|^2$ ,  $b_{-1} = -\alpha$ ,  $b_1 = -\bar{\alpha}$ ,  $b_p = 0$ ,  $|p| > 1$  are the Fourier coefficients of the function  $f^{-1}(\lambda)$ . According to the Corollary 2.4 the spectral characteristic of the optimal estimate  $\hat{A}_1 \xi$  of the functional  $A_1 \xi$  is

calculated by the formula

$$h_1(e^{i\lambda}) = (a(0) + a(1)e^{i\lambda}) - \left( \sum_{j \in T} (\mathbf{B}^{-1} \vec{\mathbf{a}}_1)_j e^{ij\lambda} \right) (b_{-1}e^{-i\lambda} + b_0 + b_1e^{i\lambda}),$$

where vector  $\vec{\mathbf{a}}_1 = (0, 0, a(0), a(1), 0, 0, \dots)$ .

To find the unknown coefficients  $c(j) = (\mathbf{B}^{-1} \vec{\mathbf{a}}_1)_j, j \in T = S \cup \{0, 1, 2, \dots\}$ , we use equation (17), where  $\vec{\mathbf{c}} = (c(-3), c(-2), c(0), c(1), c(2), c(3), \dots)$ . The operator  $\mathbf{B}$  is defined by the matrix

$$B = \begin{pmatrix} b_0 & b_{-1} & 0 & 0 & 0 & 0 & 0 & \dots \\ b_1 & b_0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & b_0 & b_{-1} & 0 & 0 & 0 & \dots \\ 0 & 0 & b_1 & b_0 & b_{-1} & 0 & 0 & \dots \\ 0 & 0 & 0 & b_1 & b_0 & b_{-1} & 0 & \dots \\ 0 & 0 & 0 & 0 & b_1 & b_0 & b_{-1} & \dots \\ 0 & 0 & 0 & 0 & 0 & b_1 & b_0 & \dots \\ \dots & & & & & & & \dots \end{pmatrix}.$$

We have to find the inverse matrix  $B^{-1}$  which defines the inverse operator  $\mathbf{B}^{-1}$ . We first represent the matrix  $B$  in the form

$$B = \begin{pmatrix} B_{00} & 0 \\ 0 & B_{11} \end{pmatrix},$$

where

$$B_{00} = \begin{pmatrix} b_0 & b_{-1} \\ b_1 & b_0 \end{pmatrix},$$

$$B_{11} = \begin{pmatrix} b_0 & b_{-1} & 0 & 0 & 0 & \dots \\ b_1 & b_0 & b_{-1} & 0 & 0 & \dots \\ 0 & b_1 & b_0 & b_{-1} & 0 & \dots \\ 0 & 0 & b_1 & b_0 & b_{-1} & \dots \\ 0 & 0 & 0 & b_1 & b_0 & \dots \\ \dots & & & & & \dots \end{pmatrix}.$$

Making use of the indicated representation we may conclude that the matrix  $B^{-1}$  can be represented in the form

$$B^{-1} = \begin{pmatrix} B_{00}^{-1} & 0 \\ 0 & B_{11}^{-1} \end{pmatrix},$$

where  $B_{00}^{-1}, B_{11}^{-1}$  are inverse matrices to matrices  $B_{00}, B_{11}$  respectively. The matrix  $B_{00}^{-1}$  can be found in the form

$$B_{00}^{-1} = \begin{pmatrix} \frac{b_0}{b_0^2 - b_1 b_{-1}} & \frac{-b_{-1}}{b_0^2 - b_1 b_{-1}} \\ \frac{-b_1}{b_0^2 - b_1 b_{-1}} & \frac{b_0}{b_0^2 - b_1 b_{-1}} \end{pmatrix}.$$

In order to find the matrix  $(B_{11})^{-1}$  we use the following method. The matrix  $B_{11}$  is constructed with the help of the Fourier coefficients of the function  $f^{-1}(\lambda)$

$$B_{11}(k, j) = b_{k-j}, \quad k, j = 0, 1, 2, \dots$$

The density  $f^{-1}(\lambda) = |1 - \alpha e^{-i\lambda}|^2$  admits the factorization

$$\frac{1}{f(\lambda)} = \sum_{p=-\infty}^{\infty} b_p e^{ip\lambda} = \left| \sum_{j=0}^{\infty} \psi_j e^{-ij\lambda} \right|^2 = \left| \sum_{j=0}^{\infty} \theta_j e^{-ij\lambda} \right|^{-2},$$

$$b_0 = 1 + |\alpha|^2, b_{-1} = -\alpha, b_1 = -\bar{\alpha}, b_p = 0, |p| > 1,$$

$$\psi_0 = 1, \psi_1 = -\alpha, \psi_j = 0, j > 1,$$

$$\theta_j = \alpha^j, j \geq 0.$$

Hence  $b_p = \sum_{k=0}^{\infty} \psi_k \bar{\psi}_{k+p}$ ,  $p \geq 0$ , and  $b_{-p} = \bar{b}_p$ ,  $p \geq 0$ . Then  $b_{i-j} = \sum_{l=\max(i,j)}^{\infty} \psi_{l-i} \bar{\psi}_{l-j}$ .

Denote by  $\Psi$  and  $\Theta$  linear operators in the space  $\ell_2$  determined by matrices with elements  $\Psi_{i,j} = \psi_{i-j}$ ,  $\Theta_{i,j} = \theta_{i-j}$ , for  $0 \leq j \leq i$ ,  $\Psi_{i,j} = 0$ ,  $\Theta_{i,j} = 0$ , for  $0 \leq i < j$ . Then elements of the matrix  $B_{11}$  can be represented in the form  $B_{11}(i, j) = (\Psi' \bar{\Psi})_{i,j}$ . Since  $\Psi \Theta = \Theta \Psi = I$ , elements of the matrix  $B_{11}^{-1}$  can be calculated by the formula  $B_{11}^{-1}(i, j) = (\bar{\Theta} \Theta')_{i,j} = \sum_{l=0}^{\min(i,j)} \bar{\theta}_{i-l} \theta_{j-l}$ , and the matrix  $(B_{11})^{-1}$  is of the form

$$B_{11}^{-1} = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \dots \\ \bar{\alpha} & \bar{\alpha}\alpha + 1 & \bar{\alpha}\alpha^2 + \alpha & \bar{\alpha}\alpha^3 + \alpha^2 & \dots \\ \alpha^2 & \alpha^2\alpha + \bar{\alpha} & \alpha^2\alpha^2 + \bar{\alpha}\alpha + 1 & \alpha^2\alpha^3 + \bar{\alpha}\alpha^2 + \alpha & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

From equation (17) we can find the unknown coefficients  $c(j)$ ,  $j \in T$ ,

$$\begin{aligned} c(-3) &= 0, \\ c(-2) &= 0, \\ c(0) &= a(0) + a(1)\alpha, \\ c(1) &= a(0)\bar{\alpha} + a(1)(\bar{\alpha}\alpha + 1), \\ c(2) &= a(0)\bar{\alpha}^2 + a(1)(\bar{\alpha}^2\alpha + \bar{\alpha}), \\ &\dots \\ c(i) &= a(0)\bar{\alpha}^i + a(1)(\bar{\alpha}^i\alpha + \bar{\alpha}^{i-1}), \quad i > 2. \end{aligned}$$

Hence the spectral characteristic of the optimal estimate is calculated by the formula

$$\begin{aligned} h_1(e^{i\lambda}) &= (a(0) + a(1)e^{i\lambda}) - (c(-3)e^{-i3\lambda} + c(-2)e^{-i2\lambda} + c(0) + c(1)e^{i\lambda} + c(2)e^{i2\lambda} \\ &\quad + \sum_{j>2} c(j)e^{ij\lambda}) (b_{-1}e^{-i\lambda} + b_0 + b_1e^{i\lambda}) = -c(0)b_{-1}e^{-i\lambda} - c(1)b_1e^{i2\lambda} \\ &\quad - c(2)b_0e^{i2\lambda} - c(2)b_1e^{i3\lambda} - \sum_{j>2} c(j)e^{ij\lambda} (b_{-1}e^{-i\lambda} + b_0 + b_1e^{i\lambda}). \end{aligned}$$

Since coefficients  $b_1c(j-1) + b_0c(j) + b_{-1}c(j+1)$  for  $j \geq 2$  are zero, the spectral characteristic of the estimate  $\hat{A}_1\xi$  is of the form

$$h_1(e^{i\lambda}) = -c(0)b_{-1}e^{-i\lambda} = (a(0) + a(1)\alpha)\alpha e^{-i\lambda}.$$

The mean-square error of the estimate of the functional  $A_1\xi$  is calculated by the formula

$$\Delta(h_1; f) = \langle B^{-1}\bar{\mathbf{a}}_1, \bar{\mathbf{a}}_1 \rangle = (a(0))^2 + a(0)a(1)\alpha + a(0)a(1)\bar{\alpha} + (a(1))^2|\alpha|^2 + (a(1))^2.$$

With the help of the next example we demonstrate the impact of missing values on the mean-square error of the optimal estimate of the functional.

### Example 2.2

Let us investigate the mean-square error of the optimal estimation of the functional  $A\xi = \sum_{k=0}^{\infty} a(k)\xi(k)$  from observations of the sequence  $\xi(j)$ :

1. at points  $j \in Z_-$  with no missing observations and
2. at points  $j \in Z_- \setminus S$ , where there are missing observations.

Thus, we can see the impact of missing values on the mean-square error of the optimal estimate of the functional. For easy computation, consider the case where only one value,  $S = \{-n\}$ , is missed. Suppose that the function  $f^{-1}(\lambda)$  admits the factorization

$$\frac{1}{f(\lambda)} = \sum_{p=-\infty}^{\infty} b_p e^{ip\lambda} = \left| \sum_{j=0}^{\infty} \psi_j e^{-ij\lambda} \right|^2 = \left| \sum_{j=0}^{\infty} \theta_j e^{-ij\lambda} \right|^{-2}.$$

We first find the mean-square error of the extrapolation of the functional  $A\xi$  in the case where all observations of the sequence at points  $j \in Z_-$  are given. In this case by the formula proposed by Moklyachuk [23], [27]

$$\Delta_1(h; f) = \langle \mathbf{B}_1^{-1} \vec{a}, \vec{a} \rangle = \langle \mathbf{A} \vec{\theta}, \mathbf{A} \vec{\theta} \rangle,$$

where  $\mathbf{B}_1$  is the linear operator in the space  $\ell_2$  defined by the matrix  $B_1$  with elements that are the Fourier coefficients of the function  $f^{-1}(\lambda)$ :

$$B_1 = \begin{pmatrix} b_0 & b_{-1} & b_{-2} & \dots \\ b_1 & b_0 & b_{-1} & \dots \\ b_2 & b_1 & b_0 & \dots \\ \dots & & & \end{pmatrix},$$

$\mathbf{A}$  is the linear operator in the space  $\ell_2$  defined by the relation

$$(\mathbf{A} \vec{\theta})_k = \sum_{l=0}^{\infty} a(k+l) \theta_l,$$

the vector  $\vec{a} = (a(0), a(1), a(2), \dots)$ , and the vector  $\vec{\theta} = (\theta_0, \theta_1, \dots)$ . Note, that under condition (3) the operator  $\mathbf{A}$  is compact.

Using the introduced in the previous example linear operators  $\Psi$  and  $\Theta$ , we can represent elements of the matrix  $B_1$  in the form

$$B_1(i, j) = (\Psi' \bar{\Psi})_{i,j} = \sum_{l=\max(i,j)}^{\infty} \psi_{l-i} \bar{\psi}_{l-j},$$

and elements of the matrix  $B_1^{-1}$  which determines the operator  $\mathbf{B}_1^{-1}$  can be calculated by the formula

$$B_1^{-1}(i, j) = (\bar{\Theta} \Theta')_{i,j} = \sum_{l=0}^{\min(i,j)} \bar{\theta}_{i-l} \theta_{j-l} = \omega_{i,j}.$$

Hence, the mean-square error of the optimal estimate of the functional  $A\xi$  from observations without missing values is calculated by the formula

$$\Delta_1(h; f) = \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \omega_{k,j} a(j) \right) a(k) = \langle \mathbf{B}_1^{-1} \vec{a}, \vec{a} \rangle = \langle \Theta' \vec{a}, \Theta' \vec{a} \rangle.$$

Next we find the mean-square error of the optimal estimate of the functional based on observations of the sequence with one missing observation by the formula  $\Delta_2(\tilde{h}; f) = \langle \mathbf{B}_2^{-1} \vec{a}, \vec{a} \rangle$ , where  $\mathbf{B}_2$  is linear operator in the space  $\ell_2$  defined by the matrix  $B_2$  with elements  $B_2(i, j) = b_{i,j} = b_{i-j}$  that are the Fourier coefficients of the function  $f^{-1}(\lambda)$ :

$$B_2 = \begin{pmatrix} b_0 & b_{-n} & b_{-n-1} & b_{-n-2} & \dots \\ b_n & b_0 & b_{-1} & b_{-2} & \dots \\ b_{n+1} & b_1 & b_0 & b_{-1} & \dots \\ \dots & & & & \end{pmatrix},$$

and the vector  $\vec{a} = (0, a(0), a(1), a(2), \dots)$ .

In order to find the inverse matrix  $B_2^{-1}$  we use the Frobenius formula (see Gantmacher [7]). The matrix  $B_2$  can be represented in the form

$$B_2 = \begin{pmatrix} B_{00} & B_{01} \\ B_{10} & B_{11} \end{pmatrix},$$

$$B_{00} = (b_0), \quad B_{01} = (b_{-n} \quad b_{-n-1} \quad b_{-n-2} \quad \dots) = \vec{b}_n',$$

$$B_{10} = \begin{pmatrix} b_n \\ b_{n+1} \\ b_{n+2} \\ \dots \end{pmatrix} = \vec{b}_n, \quad B_{11} = \begin{pmatrix} b_0 & b_{-1} & b_{-2} & \dots \\ b_1 & b_0 & b_{-1} & \dots \\ b_2 & b_1 & b_0 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

So the inverse matrix  $B_2^{-1}$  is of the form

$$B_2^{-1} = \begin{pmatrix} V^{-1} & -V^{-1}B_{01}B_{11}^{-1} \\ -B_{11}^{-1}B_{10}V^{-1} & B_{11}^{-1} + B_{11}^{-1}B_{10}V^{-1}B_{01}B_{11}^{-1} \end{pmatrix},$$

where  $V = B_{00} - B_{01}B_{11}^{-1}B_{10}$ . The matrix  $B_{11}$  coincides with the matrix  $B_1$  from this example.

With the help of the relation

$$\Psi' \vec{\psi}_n = \begin{pmatrix} b_n \\ b_{n+1} \\ b_{n+2} \\ \dots \end{pmatrix} = \vec{b}_n, \quad \vec{\psi}_n = \begin{pmatrix} \psi_n \\ \psi_{n+1} \\ \dots \end{pmatrix},$$

we can represent the value  $V$  in terms of operators  $\Psi$  and  $\Theta$

$$V = b_0 - (\vec{b}_n)' \Theta \Theta' \vec{b}_n = b_0 - \langle \Theta' \Psi' \vec{\psi}_n, \Theta' \Psi' \vec{\psi}_n \rangle = b_0 - \langle \vec{\psi}_n, \vec{\psi}_n \rangle = \sum_{k=0}^{\infty} |\psi_k|^2 - \sum_{k=n}^{\infty} |\psi_k|^2 = \sum_{k=0}^{n-1} |\psi_k|^2.$$

Denote by  $\tilde{\omega}_{k,j} = (B_{11}^{-1} + B_{11}^{-1}B_{10}V^{-1}B_{01}B_{11}^{-1})_{k,j}$ . Then for  $k, j > 0$ ,

$$\tilde{\omega}_{k,j} = \omega_{k,j} + V^{-1} \sum_{i=0}^{\infty} \omega_{k,i} b_{i,-n} \sum_{l=0}^{\infty} b_{-n,l} \omega_{l,j} = \omega_{k,j} + \gamma_{k,j}. \quad (25)$$

Making use of the introduced notations we can represent  $B_{11}^{-1}B_{10}V^{-1}B_{01}B_{11}^{-1}$  and consequently  $\gamma_{k,j}$  in the following form

$$\begin{aligned} B_{11}^{-1}B_{10}V^{-1}B_{01}B_{11}^{-1} &= \\ &= V^{-1} \left( \Theta \Theta' \vec{b}_n \right) \left( \Theta \Theta' \vec{b}_n \right)^\top = V^{-1} \left( \Theta \Theta' \Psi' \vec{\psi}_n \right) \left( \Theta \Theta' \Psi' \vec{\psi}_n \right)^\top = V^{-1} \left( \Theta \vec{\psi}_n \right) \left( \Theta \vec{\psi}_n \right)^\top, \\ \gamma_{k,j} &= V^{-1} \left( \left( \Theta \vec{\psi}_n \right) \left( \Theta \vec{\psi}_n \right)^\top \right)_{k,j}. \end{aligned}$$

Now we calculate the mean-square error of extrapolation of the functional  $A\xi$  from observations with one missing value. In this case  $T = \{-n\} \cup \{0, 1, \dots\}$  and we have that

$$\begin{aligned} \Delta_2(\tilde{h}; f) &= \sum_{k \in T} \left( \sum_{j \in T} \tilde{\omega}_{k,j} a(j) \right) a(k) = \sum_{k \in S} \left( \sum_{j \in T} \tilde{\omega}_{k,j} a(j) \right) a(k) + \sum_{k=0}^{\infty} \left( \sum_{j \in T} \tilde{\omega}_{k,j} a(j) \right) a(k) \\ &= \sum_{k=0}^{\infty} \left( \sum_{j \in S} \tilde{\omega}_{k,j} a(j) + \sum_{j=0}^{\infty} \tilde{\omega}_{k,j} a(j) \right) a(k) = \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \tilde{\omega}_{k,j} a(j) \right) a(k). \end{aligned}$$

It follows from (25) that

$$\begin{aligned} \Delta_2(\tilde{h}; f) &= \Delta_1(h; f) + \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \gamma_{k,j} a(j) \right) a(k) = \\ &= \langle \Theta' \vec{a}, \Theta' \vec{a} \rangle + V^{-1} \left| \langle \vec{\psi}_n, \Theta' \vec{a} \rangle \right|^2 = \|\Theta' \vec{a}\|^2 + V^{-1} \|\vec{\psi}_n\|^2. \end{aligned}$$

So we can estimate the impact of the missing values on the mean-square error of the optimal estimate of the functional  $A\xi$ .

In particular, in the case of estimating of the value  $\xi(0)$  based on observations  $\xi(k), k = -1, -2, \dots$  with one missing value  $\xi(-n)$  the vector  $\vec{a} = (1, 0, 0, \dots)$  and the formula for the mean-square error is of the form

$$\Delta_2(\tilde{h}; f) = |\theta_0|^2 + V^{-1} |\theta_0|^2 |\psi_n|^2 = |\theta_0|^2 \left( 1 + \frac{|\psi_n|^2}{\sum_{k=0}^{n-1} |\psi_k|^2} \right).$$

Note, that the resulting formula for the same problem based on the AR representation of the stationary sequence was obtained by Bondon (see Bondon [1]). He found that

$$E|\xi(0) - \hat{\xi}(0)|^2 = \sigma^2 \left( 1 + \frac{\psi_n^2}{\sum_{i=0}^{n-1} \psi_i^2} \right), \quad \sigma^2 = \exp \left[ \int_{-\pi}^{\pi} \ln f(\lambda) d\lambda \right],$$

where  $\psi_k$  are the AR parameters of the sequence.

### 3. Minimax approach to extrapolation problem for stationary sequences with missing observations

Theorem 2.1 and its corollaries can be applied for finding solutions to extrapolation problem for stationary sequences with missing observations only in the case of spectral certainty, where spectral densities  $f(\lambda), f_{\xi\eta}(\lambda), f_{\eta\xi}(\lambda), g(\lambda)$ , which form the spectral density matrix  $F(\lambda)$ , are exactly known. If the complete information about spectral densities is impossible while a class of admissible spectral density matrices  $D = \{F(\lambda)\}$  is given, the minimax(robust) method of extrapolation is reasonable. It consists in finding an estimate which minimizes the value of the mean-square error for all spectral density matrices from the given class. For description of the minimax method we introduce the following definitions (see Moklyachuk [23]- [27]).

*Definition 3.1*

For a given class of spectral density matrices  $D = \{F(\lambda)\}$  the spectral density matrix  $F^0(\lambda) \in D$  is called the least favorable in the class  $D$  for the optimal linear extrapolation of the functional  $A\xi$  if the following relation holds true

$$\Delta(F^0) = \Delta(h(F^0); F^0) = \max_{F \in D} \Delta(h(F); F).$$

*Definition 3.2*

For a given class of spectral density matrices  $D = \{F(\lambda)\}$  the spectral characteristic  $h^0(e^{i\lambda})$  of the optimal linear estimate of the functional  $A\xi$  is called minimax-robust if there are satisfied conditions

$$\begin{aligned} h^0(e^{i\lambda}) &\in H_D = \bigcap_{F \in D} L_2^s(f + g), \\ \min_{h \in H_D} \max_{F \in D} \Delta(h; F) &= \max_{F \in D} \Delta(h^0; F). \end{aligned}$$

From the introduced definitions and formulas derived above we can obtain the following statement.

*Lemma 3.1*

Spectral densities  $f^0(\lambda), f_{\xi\eta}^0(\lambda), f_{\eta\xi}^0(\lambda), g^0(\lambda)$ , satisfying the minimality condition (1), form the spectral density matrix  $F^0(\lambda) \in D$  which is least favorable in the class  $D$  for the optimal linear extrapolation of the functional  $A\xi$  if the Fourier coefficients (8) of functions

$$\frac{1}{f^0(\lambda) + f_{\xi\eta}^0(\lambda) + f_{\eta\xi}^0(\lambda) + g^0(\lambda)}, \quad \frac{f^0(\lambda) + f_{\xi\eta}^0(\lambda)}{f^0(\lambda) + f_{\xi\eta}^0(\lambda) + f_{\eta\xi}^0(\lambda) + g^0(\lambda)}, \quad \frac{f^0(\lambda)g^0(\lambda) - f_{\xi\eta}^0(\lambda)f_{\eta\xi}^0(\lambda)}{f^0(\lambda) + f_{\xi\eta}^0(\lambda) + f_{\eta\xi}^0(\lambda) + g^0(\lambda)}$$

define operators  $\mathbf{B}^0, \mathbf{R}^0, \mathbf{Q}^0$  which determine a solution to the constrain optimization problem

$$\begin{aligned} \max_{F \in D} \langle \mathbf{R}\vec{a}, \mathbf{B}^{-1}\mathbf{R}\vec{a} \rangle + \langle \mathbf{Q}\vec{a}, \vec{a} \rangle &= \\ &= \langle \mathbf{R}^0\vec{a}, (\mathbf{B}^0)^{-1}\mathbf{R}^0\vec{a} \rangle + \langle \mathbf{Q}^0\vec{a}, \vec{a} \rangle. \end{aligned} \tag{26}$$

The minimax spectral characteristic  $h^0 = h(F^0)$  is calculated by the formula (10) if  $h(F^0) \in H_D$ .

In the case of uncorrelated stationary sequences the corresponding definitions and lemmas are as follows.

*Definition 3.3*

For a given class of spectral densities  $D = D_f \times D_g$  spectral densities  $f^0(\lambda) \in D_f, g^0(\lambda) \in D_g$  are called the least favorable in the class  $D$  for the optimal linear extrapolation of the functional  $A\xi$  based on observations of the uncorrelated sequences if the following relation holds true

$$\Delta(f^0, g^0) = \Delta(h(f^0, g^0); f^0, g^0) = \max_{(f,g) \in D_f \times D_g} \Delta(h(f, g); f, g).$$

*Definition 3.4*

For a given class of spectral densities  $D = D_f \times D_g$  the spectral characteristic  $h^0(e^{i\lambda})$  of the optimal linear estimate of the functional  $A\xi$  based on observations of the uncorrelated sequences is called minimax-robust if there are satisfied conditions

$$\begin{aligned} h^0(e^{i\lambda}) \in H_D &= \bigcap_{(f,g) \in D_f \times D_g} L_2^s(f + g), \\ \min_{h \in H_D} \max_{(f,g) \in D} \Delta(h; f, g) &= \max_{(f,g) \in D} \Delta(h^0; f, g). \end{aligned}$$

*Lemma 3.2*

Spectral densities  $f^0(\lambda) \in D_f, g^0(\lambda) \in D_g$  satisfying the minimality condition (12) are the least favorable in the class  $D = D_f \times D_g$  for the optimal linear extrapolation of the functional  $A\xi$  based on observations of the uncorrelated sequences if the Fourier coefficients (15) of functions

$$(f^0(\lambda) + g^0(\lambda))^{-1}, \quad f^0(\lambda)(f^0(\lambda) + g^0(\lambda))^{-1}, \quad f^0(\lambda)g^0(\lambda)(f^0(\lambda) + g^0(\lambda))^{-1}$$

define operators  $\mathbf{B}^0, \mathbf{R}^0, \mathbf{Q}^0$  which determine a solution to the constrain optimization problem

$$\begin{aligned} \max_{(f,g) \in D_f \times D_g} \langle \mathbf{R}\vec{a}, \mathbf{B}^{-1}\mathbf{R}\vec{a} \rangle + \langle \mathbf{Q}\vec{a}, \vec{a} \rangle &= \\ &= \langle \mathbf{R}^0\vec{a}, (\mathbf{B}^0)^{-1}\mathbf{R}^0\vec{a} \rangle + \langle \mathbf{Q}^0\vec{a}, \vec{a} \rangle. \end{aligned} \tag{27}$$

The minimax spectral characteristic  $h^0 = h(f^0, g^0)$  is calculated by the formula (13) if  $h(f^0, g^0) \in H_D$ .

In the case of observations of the sequence without noise we obtain the following corollary.

*Corollary 3.1*

Let the spectral density  $f^0(\lambda) \in D_f$  be such that the function  $(f^0(\lambda))^{-1}$  is integrable. The spectral density  $f^0(\lambda) \in D_f$  is the least favorable in the class  $D_f$  for the optimal linear extrapolation of the functional  $A\xi$  if the Fourier coefficients of the function  $(f^0(\lambda))^{-1}$  define the operator  $\mathbf{B}^0$  which determines a solution to the optimization problem

$$\max_{f \in D_f} \langle \mathbf{B}^{-1} \vec{a}, \vec{a} \rangle = \langle (\mathbf{B}^0)^{-1} \vec{a}, \vec{a} \rangle. \tag{28}$$

The minimax spectral characteristic  $h^0 = h(f^0)$  is calculated by the formula (18) if  $h(f^0) \in H_{D_f}$ .

The least favorable spectral densities  $f^0(\lambda), g^0(\lambda)$  and the minimax spectral characteristic  $h^0 = h(f^0, g^0)$  form a saddle point of the function  $\Delta(h; f, g)$  on the set  $H_D \times D$ . The saddle point inequalities

$$\begin{aligned} \Delta(h; f^0, g^0) &\geq \Delta(h^0; f^0, g^0) \geq \Delta(h^0; f, g) \\ \forall h \in H_D, \forall f \in D_f, \forall g \in D_g \end{aligned}$$

hold true if  $h^0 = h(f^0, g^0)$  and  $h(f^0, g^0) \in H_D$ , where  $(f^0, g^0)$  is a solution to the constrained optimization problem

$$\sup_{(f, g) \in D_f \times D_g} \Delta(h(f^0, g^0); f, g) = \Delta(h(f^0, g^0); f^0, g^0), \tag{29}$$

$$\begin{aligned} \Delta(h(f^0, g^0); f, g) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|A(e^{i\lambda})g^0(\lambda) + C^0(e^{i\lambda})|^2}{(f^0(\lambda) + g^0(\lambda))^2} f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|A(e^{i\lambda})f^0(\lambda) - C^0(e^{i\lambda})|^2}{(f^0(\lambda) + g^0(\lambda))^2} g(\lambda) d\lambda, \end{aligned}$$

$$C^0(e^{i\lambda}) = \sum_{j \in T} ((\mathbf{B}^0)^{-1} \mathbf{R}^0 \vec{a})_j e^{ij\lambda}.$$

The constrained optimization problem (29) is equivalent to the unconstrained optimization problem (see Pshenichnyj [39]):

$$\Delta_D(f, g) = -\Delta(h(f^0, g^0); f, g) + \delta((f, g) | D_f \times D_g) \rightarrow \inf, \tag{30}$$

where  $\delta((f, g) | D_f \times D_g)$  is the indicator function of the set  $D = D_f \times D_g$ .

A solution of the problem (30) is determined by the condition  $0 \in \partial \Delta_D(f^0, g^0)$ , which is the necessary and sufficient condition under which the pair  $(f^0, g^0)$  belongs to the set of minimums of the convex functional  $\Delta(h(f^0, g^0); f, g)$ . Here the notion  $\partial \Delta_D(f^0, g^0)$  indicates the subdifferential of the convex functional  $\Delta_D(f, g)$  at point  $(f^0, g^0)$  which is a set of all linear bounded functionals  $\Lambda$  on  $L_1 \times L_1$  satisfying the inequality (see Rockafellar [40]).

$$\Delta_D(f, g) - \Delta_D(f^0, g^0) \geq \Lambda((f, g) - (f_0, g_0)), \forall (f, g) \in D_f \times D_g.$$

The form of the functional  $\Delta(h(f^0, g^0); f, g)$  allows us to find derivatives and differentials in the space  $L_1 \times L_1$ . Therefore the complexity of the optimization problem (30) is determined by the complexity of calculating of the subdifferential of the indicator function  $\delta((f, g) | D_f \times D_g)$  of the set  $D_f \times D_g$  (see, for example, Ioffe and Tihomirov [11], Moklyachuk [26]).

The following statement holds true.

*Lemma 3.3*

Let  $(f^0, g^0)$  be a solution to the optimization problem (30). The spectral densities  $f^0(\lambda), g^0(\lambda)$  are the least favorable in the class  $D = D_f \times D_g$  and the spectral characteristic  $h^0 = h(f^0, g^0)$  is the minimax of the optimal linear estimate of the functional  $A\xi$  if  $h(f^0, g^0) \in H_D$ .



#### 4. Least favorable spectral densities in the class $D = D_W \times D_0$

Consider the problem of extrapolation of the functional  $A\xi$  based on observations of the uncorrelated sequences in the case where spectral densities of the observed sequences belong to the class  $D = D_W \times D_0$ , where

$$D_W = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) \cos(w\lambda) d\lambda = r_w, w = 0, 1, \dots, W \right. \right\},$$

$$D_0 = \left\{ g(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) d\lambda \leq P_1 \right. \right\}.$$

Suppose that the sequence  $\{r_w, w = 0, 1, \dots, W\}$  is strictly positive. In this case the problem of moments has infinite many solutions and the set  $D_W$  contains infinite number of densities (see Krein and Nudelman [16]).

Let the densities  $f^0(\lambda) \in D_W, g^0(\lambda) \in D_0$  and let the functions defined by formulas

$$h_f(f^0, g^0) = \frac{|A(e^{i\lambda})g^0(\lambda) + C^0(e^{i\lambda})|^2}{(f^0(\lambda) + g^0(\lambda))^2}, \quad (31)$$

$$h_g(f^0, g^0) = \frac{|A(e^{i\lambda})f^0(\lambda) - C^0(e^{i\lambda})|^2}{(f^0(\lambda) + g^0(\lambda))^2}, \quad (32)$$

be bounded. In this case the linear functional

$$\Delta(h(f^0, g^0); f, g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h_f(f^0, g^0) f(\lambda) d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} h_g(f^0, g^0) g(\lambda) d\lambda$$

is continuous and bounded in the space  $L_1 \times L_1$ . We can use the method of Lagrange multipliers to solve the optimization problem (29) and find equations which the least favorable densities satisfy (see Pshenichnyj [39], Moklyachuk [26])

$$|A(e^{i\lambda})g_0(\lambda) + C^0(e^{i\lambda})| = \left( \sum_{w=0}^W \psi_w \cos(w\lambda) \right) (f^0(\lambda) + g^0(\lambda)), \quad (33)$$

$$|A(e^{i\lambda})f_0(\lambda) - C^0(e^{i\lambda})| = \alpha (f^0(\lambda) + g^0(\lambda)), \quad (34)$$

where constants  $\psi_w \geq 0, \alpha \geq 0$ .

##### Theorem 4.1

Let the spectral densities  $f^0(\lambda) \in D_W, g^0(\lambda) \in D_0$  be such that the minimality condition (12) holds true. Let the functions defined by (31), (32) be bounded. Then the functions  $f^0(\lambda), g^0(\lambda)$  determined by equations (33), (34) are the least favorable spectral densities in the class  $D_W \times D_0$  if they determine a solution to optimization problem (27). The function  $h(f^0, g^0)$  determined by formula (13) is the minimax spectral characteristic of the optimal estimate of the functional  $A\xi$ .

##### Theorem 4.2

Let the spectral densities  $f^0(\lambda) \in D_0, g^0(\lambda) \in D_0$  be such that the minimality condition (12) holds true and functions defined by (31), (32) be bounded. Then functions  $f^0(\lambda), g^0(\lambda)$  determined from the following equations

$$|A(e^{i\lambda})g^0(\lambda) + C^0(e^{i\lambda})| = \alpha_1 (f^0(\lambda) + g^0(\lambda)), \quad (35)$$

$$|A(e^{i\lambda})f^0(\lambda) - C^0(e^{i\lambda})| = \alpha_2 (f^0(\lambda) + g^0(\lambda)), \quad (36)$$

$\alpha_1 \geq 0, \alpha_2 \geq 0$ , are the least favorable densities in the class  $D_0 \times D_0$ , if they determine a solution to the optimization problem (27). Note that  $\alpha_1 \neq 0$ , if  $\frac{1}{2\pi} \int_{-\pi}^{\pi} f^0(\lambda)d\lambda = P_1$ , and  $\alpha_2 \neq 0$ , if  $\frac{1}{2\pi} \int_{-\pi}^{\pi} g^0(\lambda)d\lambda = P_1$ . The function  $h(f^0, g^0)$  calculated by formula (13) is the minimax spectral characteristic of the optimal estimate of the functional  $A\xi$ .

*Corollary 4.1*

Suppose that the density  $f(\lambda)$  is known, and the density  $g^0(\lambda) \in D_0$ . Let the function  $f(\lambda) + g^0(\lambda)$  satisfy the minimality condition (12) and let the function  $h_g(f, g^0)$  determined by (32) be bounded. The spectral density  $g^0(\lambda)$  is the least favorable in the class  $D_0$  for the optimal linear extrapolation of the functional  $A\xi$  if

$$g^0(\lambda) = \max \{0, \alpha^{-1} |A(e^{i\lambda})f(\lambda) - C^0(e^{i\lambda})| - f(\lambda)\},$$

and the pair  $(f(\lambda), g^0(\lambda))$  determines a solution to the optimization problem (27). The minimax spectral characteristic  $h(f, g^0)$  of the optimal linear estimate of the functional  $A\xi$  is calculated by the formula (13).

*Corollary 4.2*

Let the density  $g(\lambda)$  be known and let  $f^0(\lambda) \in D_W$ . Suppose the function  $f^0(\lambda) + g(\lambda)$  satisfies the minimality condition (12) and the function  $h_f(f^0, g)$  defined by (31) is bounded. The spectral density  $f^0(\lambda)$  is the least favorable in the class  $D_W$  for the optimal linear extrapolation of the functional  $A\xi$  if

$$f^0(\lambda) = \frac{|A(e^{i\lambda})g(\lambda) + C^0(e^{i\lambda})|}{\sum_{w=0}^W \psi_w \cos(w\lambda)} - g(\lambda),$$

and the pair  $(f^0(\lambda), g(\lambda))$  determines a solution to the optimization problem (27). The function  $h(f^0, g)$  calculated by (13) is the minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$ .

*Corollary 4.3*

Let the density  $g(\lambda)$  be known and  $f^0(\lambda) \in D_0$ . Let the function  $f^0(\lambda) + g(\lambda)$  satisfy the minimality condition (12) and let the function  $h_f(f^0, g)$  defined by (31) be bounded. The spectral density  $f^0(\lambda)$  is the least favorable in the class  $D_0$  if it is of the form

$$f^0(\lambda) = \max \{0, \alpha^{-1} |A(e^{i\lambda})g(\lambda) + C^0(e^{i\lambda})| - g(\lambda)\},$$

and the pair  $(f^0(\lambda), g(\lambda))$  determines a solution to the optimization problem (27). The function  $h(f^0, g)$  calculated by (13) is the minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$ .

*Corollary 4.4*

Let the sequence  $f(\lambda)$  be known and the spectral density  $g^0(\lambda) \in D_W$ . Suppose that the function  $f(\lambda) + g^0(\lambda)$  satisfies the minimality condition (12) and the function  $h_g(f, g_0)$  defined by (32) is bounded. The spectral density  $g^0(\lambda)$  is the least favorable in the class  $D_W$  if

$$g^0(\lambda) = \frac{|A(e^{i\lambda})f(\lambda) - C^0(e^{i\lambda})|}{\sum_{w=0}^W \psi_w \cos(w\lambda)} - f(\lambda),$$

and the pair  $(f(\lambda), g^0(\lambda))$  determines a solution to the optimization problem (27). The function  $h(f, g^0)$  calculated by (13) is the minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$ .

Consider the case where stationary sequence is observed without noise,  $g(\lambda) = 0$ . We obtain the following corollaries.

*Corollary 4.5*

Let the spectral density  $f^0(\lambda) \in D_0$  and satisfy the minimality condition (20). The spectral density  $f^0(\lambda)$  is the

least favorable in the class  $D_0$  for the optimal extrapolation of the functional  $A\xi$  if

$$f^0(\lambda) = \alpha^{-1} |C^0(e^{i\lambda})|,$$

and  $f^0(\lambda)$  determines a solution to the optimization problem (28). The function  $h(f^0)$  calculated by (18) is the minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$ .

*Corollary 4.6*

Let the density  $f^0(\lambda) \in D_W$  and satisfy the minimality condition (20). The spectral density  $f^0(\lambda)$  is the least favorable in the class  $D_W$  for the optimal extrapolation of the functional  $A\xi$  if

$$f^0(\lambda) = \frac{|C^0(e^{i\lambda})|}{\sum_{w=0}^W \psi_w \cos(w\lambda)},$$

and  $f^0(\lambda)$  determines a solution to the optimization problem (28). The function  $h(f^0)$  calculated by (18) is the minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$ .

### 5. Least favorable spectral densities in the class $D = D_v^u \times D_\varepsilon$

Consider the problem of extrapolation of the functional  $A\xi$  based on observations of the uncorrelated sequences in the case where spectral densities of the observed sequences belong to the class  $D = D_v^u \times D_\varepsilon$ ,

$$D_v^u = \left\{ f(\lambda) \left| v(\lambda) \leq f(\lambda) \leq u(\lambda), \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) d\lambda \leq P_1 \right. \right\},$$

$$D_\varepsilon = \left\{ g(\lambda) \left| g(\lambda) = (1 - \varepsilon)g_1(\lambda) + \varepsilon w(\lambda), \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) d\lambda \leq P_2 \right. \right\},$$

where spectral densities  $u(\lambda), v(\lambda), g_1(\lambda)$  are known and fixed, densities  $u(\lambda)$  and  $v(\lambda)$  are bounded. The class  $D_v^u$  describes the “strip” model of stochastic sequences, while the class  $D_\varepsilon$  describes the “ $\varepsilon$  - contamination” model of stochastic sequences.

Suppose that the functions  $f^0(\lambda) \in D_v^u, g^0(\lambda) \in D_\varepsilon$  determine bounded functions  $h_f(f^0, g^0), h_g(f^0, g^0)$  by formulas (31), (32). It follows from the condition  $0 \in \partial\Delta_D(f_0, g_0)$  that the least favorable spectral densities satisfy equations

$$|A(e^{i\lambda})g^0(\lambda) + C^0(e^{i\lambda})| = (f^0(\lambda) + g^0(\lambda))(\gamma_1(\lambda) + \gamma_2(\lambda) + \alpha_1^{-1}), \quad (37)$$

$$|A(e^{i\lambda})f^0(\lambda) - C^0(e^{i\lambda})| = (f^0(\lambda) + g^0(\lambda))(\varphi(\lambda) + \alpha_2^{-1}), \quad (38)$$

where  $\gamma_1 \leq 0$  and  $\gamma_1 = 0$  if  $f^0(\lambda) \geq v(\lambda)$ ;  $\gamma_2 \geq 0$  and  $\gamma_2 = 0$  if  $f^0(\lambda) \leq u(\lambda)$ ;  $\varphi(\lambda) \leq 0$  and  $\varphi(\lambda) = 0$  if  $g^0(\lambda) \geq (1 - \varepsilon)g_1(\lambda)$ .

Hence the following theorem holds true.

*Theorem 5.1*

Suppose that the spectral densities  $f^0(\lambda) \in D_v^u, g^0(\lambda) \in D_\varepsilon$  are such that the minimality condition (12) holds true. Let the functions determined by formulas (31), (32) be bounded. Then the functions  $f^0(\lambda), g^0(\lambda)$  determined by (37), (38) are the least favorable densities in the class  $D_v^u \times D_\varepsilon$  if they determine a solution to the optimization problem (27). The function  $h(f^0, g^0)$  calculated by (13) is the minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$ .

*Corollary 5.1*

Suppose that we know the form of spectral density  $f(\lambda)$  and spectral density  $g^0(\lambda) \in D_\varepsilon$ . Let the function  $f(\lambda) + g^0(\lambda)$  satisfy the minimality condition (12) and function  $h_g(f, g^0)$  determined by (32) be bounded. The spectral density  $g^0(\lambda)$  is the least favorable in the class  $D_\varepsilon$  for the optimal linear extrapolation of the functional  $A\xi$  if

$$g^0(\lambda) = \max \{ (1 - \varepsilon)g_1(\lambda), f_1(\lambda) \},$$

$$f_1(\lambda) = \alpha_2 |A(e^{i\lambda})f(\lambda) - C^0(e^{i\lambda})| - f(\lambda),$$

and the pair  $(f(\lambda), g^0(\lambda))$  determines a solution to the optimization problem (27). The minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$  is determined by (13).

*Corollary 5.2*

Let the spectral density  $g(\lambda)$  be known and the spectral density  $f^0(\lambda) \in D_v^u$ . Suppose that the function  $f^0(\lambda) + g(\lambda)$  satisfies the minimality condition (12) and the function  $h_f(f^0, g)$  determined by (31) is bounded. The spectral density  $f^0(\lambda)$  is the least favorable in the class  $D_v^u$  if

$$f^0(\lambda) = \min \{ \max \{ \alpha_1 |A(e^{i\lambda})g(\lambda) + C^0(e^{i\lambda})| - g(\lambda), v(\lambda) \}, u(\lambda) \},$$

and the pair  $(f^0(\lambda), g(\lambda))$  determines a solution to the optimization problem (27). The function  $h(f^0, g)$  calculated by (10) is the minimax spectral characteristic of the optimal linear estimate of the functional  $A\xi$ .

**6. Least favorable spectral densities in the class  $D = D_{\varepsilon_1}^1 \times D_{\varepsilon_2}$**

Consider the problem of minimax extrapolation of the functional  $A\xi$  based on observations of the uncorrelated sequences in the case where spectral densities of the observed sequences belong to the class  $D_{\varepsilon_1}^1 \times D_{\varepsilon_2}$ ,

$$D_{\varepsilon_1}^1 = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\lambda) - f_1(\lambda)| d\lambda \leq \varepsilon_1 \right. \right\},$$

$$D_{\varepsilon_2} = \left\{ g(\lambda) \left| g(\lambda) = (1 - \varepsilon_2)g_1(\lambda) + \varepsilon_2 w(\lambda), \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) d\lambda \leq P_1 \right. \right\},$$

where spectral densities  $f_1(\lambda), g_1(\lambda)$  are known and fixed. The class  $D_{\varepsilon_1}^1$  describes the model of “ $\varepsilon$ -neighbourhood” in the space  $L_1$  of a given bounded spectral density  $f_1(\lambda)$ .

Suppose that the densities  $f^0(\lambda) \in D_{\varepsilon_1}^1, g^0(\lambda) \in D_{\varepsilon_2}$  determine the bounded functions  $h_f(f^0, g^0), h_g(f^0, g^0)$  by formulas (31), (32). The condition  $0 \in \partial\Delta_D(f_0, g_0)$  determines equations which the least favorable spectral densities satisfy

$$|A(e^{i\lambda})g^0(\lambda) + C^0(e^{i\lambda})| = (f^0(\lambda) + g^0(\lambda))\Psi(\lambda)\alpha_1, \tag{39}$$

$$|A(e^{i\lambda})f^0(\lambda) - C^0(e^{i\lambda})| = (f^0(\lambda) + g^0(\lambda))(\varphi(\lambda) + \alpha_2^{-1}), \tag{40}$$

where  $|\Psi(\lambda)| \leq 1$  and  $\Psi(\lambda) = \text{sign}(f_0(\lambda) - f_1(\lambda))$  if  $f_0(\lambda) \neq f_1(\lambda)$ ;  $\alpha_1, \alpha_2$  are constants,  $\varphi(\lambda) \leq 0$ , and  $\varphi(\lambda) = 0$  if  $g_0(\lambda) \geq (1 - \varepsilon_2)g_1(\lambda)$ . Equations (39), (40) together with the extremum condition (27) and normality condition

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f^0(\lambda) - f_1(\lambda)| d\lambda = \varepsilon_1, \tag{41}$$

determine the least favorable spectral densities in the class  $D$ .

The following theorem and corollaries hold true.

*Theorem 6.1*

Let the densities  $f^0(\lambda) \in D_{\varepsilon_1}^1$ ,  $g^0(\lambda) \in D_{\varepsilon_2}$  be such that the minimality condition (12) holds true and functions determined by (31), (32) be bounded. The spectral densities  $f^0(\lambda)$ ,  $g^0(\lambda)$  are the least favorable in the class  $D_{\varepsilon_1}^1 \times D_{\varepsilon_2}$  for the optimal linear extrapolation of the functional  $A\xi$  if they satisfy equations (39)–(41) and determine a solution to the optimization problem (27). The function  $h(f^0, g^0)$  calculated by the formula (13) is the minimax spectral characteristic of the estimate of the functional  $A\xi$ .

*Corollary 6.1*

Let the density  $g(\lambda)$  be known and the density  $f^0(\lambda) \in D_{\varepsilon_1}^1$ . Suppose the function  $f^0(\lambda) + g(\lambda)$  satisfies the minimality condition (12) and function  $h_f(f^0, g)$  determined by (31) is bounded. The spectral density  $f^0(\lambda)$  is the least favorable in the class  $D_{\varepsilon_1}^1$  if it is of the form

$$f^0(\lambda) = \max \{f_1(\lambda), \alpha_1 |A(e^{i\lambda})g(\lambda) + C^0(e^{i\lambda})| - g(\lambda)\},$$

and the pair  $(f^0(\lambda), g(\lambda))$  determines a solution to the optimization problem (27). The function  $h(f^0, g)$  calculated by the formula (13) is the minimax spectral characteristic of the estimate of the functional  $A\xi$ .

*Corollary 6.2*

Suppose  $f(\lambda)$  is known and the density  $g^0(\lambda) \in D_{\varepsilon_2}$ . Let the function  $f(\lambda) + g^0(\lambda)$  satisfy the minimality condition (12) and the function  $h_g(f, g^0)$  determined by (32) be bounded. The spectral density  $g^0(\lambda)$  is the least favorable in the class  $D_{\varepsilon_2}$  if

$$g^0(\lambda) = \max \{(1 - \varepsilon_2)g_1(\lambda), \alpha_2 |A(e^{i\lambda})f(\lambda) - C^0(e^{i\lambda})| - f(\lambda)\},$$

and the pair  $(f(\lambda), g^0(\lambda))$  determines a solution to the optimization problem (27). The function  $h(f, g^0)$  calculated by the formula (13) is the minimax spectral characteristic of the estimate of the functional  $A\xi$ .

**7. Conclusions**

In the article we propose methods of the mean-square optimal linear extrapolation of functionals which depend on the unknown values of a stationary sequence based on observations of the sequence with noise and missing values. In the case of spectral certainty where the spectral densities of stationary sequences are known we derive formulas for calculating the spectral characteristic and the mean-square error of the estimate of the functional. In the case of spectral uncertainty where certain sets of admissible densities are given we derive relations which determine the least favorable spectral densities and minimax spectral characteristics of the extrapolation.

## REFERENCES

1. P. Bondon, *Prediction with incomplete past of a stationary process*, Stochastic Process and their Applications. vol.98, pp. 67-76, 2002.
2. P. Bondon, *Influence of missing values on the prediction of a stationary time series*, Journal of Time Series Analysis, vol. 26, no. 4, pp. 519-525, 2005.
3. P. J. Brockwell and R. A. Davis, *Time series: Theory and methods (2nd ed.)*, New York: Springer, 1991.
4. R. Cheng, A.G. Miamee, M. Pourahmadi, *Some extremal problems in  $L^p(w)$* , Proc. Am. Math. Soc. vol.126, pp. 2333-2340, 1998.
5. J. Franke, *Minimax robust prediction of discrete time series*, Z. Wahrscheinlichkeitstheor. Verw. Gebiete, vol. 68, pp. 337-364, 1985.
6. J. Franke and H. V. Poor, *Minimax-robust filtering and finite-length robust predictors*, Robust and Nonlinear Time Series Analysis. Lecture Notes in Statistics, Springer-Verlag, vol. 26, pp. 87-126, 1984.
7. F. R. Gantmaher, *Applications of the theory of matrices*, Interscience publishers, inc., New York, 1959.
8. I. I. Gikhman and A. V. Skorokhod, *The theory of stochastic processes. I.*, Berlin: Springer, 2004.
9. U. Grenander, *A prediction problem in game theory*, Arkiv för Matematik, vol. 3, pp. 371-379, 1957.
10. E. J. Hannan, *Multiple time series*, Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley & Sons, Inc. XI, 1970.
11. A. D. Ioffe, and V. M. Tihomirov, *Theory of extremal problems*, Studies in Mathematics and its Applications, Vol. 6. Amsterdam, New York, Oxford: North-Holland Publishing Company. XII, 1979.
12. K. Karhunen, *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Annales Academiae Scientiarum Fennicae. Ser. A I, vol. 37, 1947.

13. Y. Kasahara, and M. Pourahmadi, and A. Inoue, *Duals of random vectors and processes with applications to prediction problems with missing values*, Statistics & Probability Letters, vol. 79, no. 14, pp. 1637-1646, 2009.
14. S. A. Kassam and H. V. Poor, *Robust techniques for signal processing: A survey*, Proceedings of the IEEE, vol. 73, no. 3, pp. 433-481, 1985.
15. A. N. Kolmogorov, *Selected works by A. N. Kolmogorov. Vol. II: Probability theory and mathematical statistics. Ed. by A. N. Shiriyayev*, Mathematics and Its Applications. Soviet Series. 26. Dordrecht etc. Kluwer Academic Publishers, 1992.
16. M. G. Krein and A. A. Nudelman, *The Markov moment problem and extremal problems. Ideas and problems of P.L. Cebysev and A.A. Markov and their further development. Translations of Mathematical Monographs. Vol. 50*. Providence, R.I.: American Mathematical Society (AMS), 1977.
17. M. M. Luz, and M. P. Moklyachuk, *Interpolation of functionals of stochastic sequences with stationary increments*, Theory of Probability and Mathematical Statistics, vol. 87, pp. 117-133, 2013.
18. M. Luz, and M. Moklyachuk, *Minimax-robust filtering problem for stochastic sequences with stationary increments and cointegrated sequences*, Statistics, Optimization & Information Computing, vol. 2, no. 3, pp. 176-199, 2014.
19. M. M. Luz, and M. P. Moklyachuk, *Minimax-robust filtering problem for stochastic sequence with stationary increments*, Theory of Probability and Mathematical Statistics, vol. 89, pp. 127-142, 2014.
20. M. Luz, and M. Moklyachuk, *Minimax-robust prediction problem for stochastic sequences with stationary increments and cointegrated sequences*, Statistics, Optimization & Information Computing, vol. 3, no. 2, pp. 160-188, 2015.
21. M. M. Luz, and M. P. Moklyachuk, *Filtering problem for functionals of stationary sequences*, Statistics, Optimization & Information Computing, vol. 4, no. 1 pp. 68-83, 2016.
22. M. M. Luz and M. P. Moklyachuk, *Estimates of functionals from processes with stationary increments and cointegrated sequences*, NVP "Interservis", 2016.
23. M. P. Moklyachuk, *Robust procedures in time series analysis*, Theory of Stochastic Processes, vol. 6, no. 3-4, pp. 127-147, 2000.
24. M. P. Moklyachuk, *Game theory and convex optimization methods in robust estimation problems*, Theory of Stochastic Processes, vol. 7, no. 1-2, pp. 253-264, 2001.
25. M. P. Moklyachuk, *Robust estimations of functionals of stochastic processes*, Kyiv University, Kyiv, 2008.
26. M. P. Moklyachuk, *Nonsmooth analysis and optimization*, Kyiv University, Kyiv, 2008.
27. M. P. Moklyachuk, *Minimax-robust estimation problems for stationary stochastic sequences*, Statistics, Optimization & Information Computing, vol. 3, no. 4, pp. 348-419, 2015.
28. M. Moklyachuk and I. Golichenko, *Periodically correlated processes estimates*. LAP Lambert Academic Publishing, 2016.
29. M. Moklyachuk and O. Masyutka, *Minimax-robust estimation technique for stationary stochastic processes*, LAP LAMBERT Academic Publishing, 2012.
30. M. P. Moklyachuk and M. I. Sidei, *Interpolation of stationary sequences observed with the noise*, Theory of Probability and Mathematical Statistics, vol. 93, pp. 143-156, 2015.
31. M. P. Moklyachuk and M. I. Sidei, *Interpolation problem for stationary sequences with missing observations*, Statistics, Optimization & Information Computing, vol. 3, no. 3, pp. 259-275, 2015.
32. M. Moklyachuk and M. Sidei, *Extrapolation problem for functionals of stationary processes with missing observations*, Bukovinian Mathematical Journal, vol. 4, no. 1-2, pp. 122-129, 2016.
33. M. P. Moklyachuk and M. I. Sidei, *Interpolation of functionals of stationary processes with missing observations*, Bulletin of Taras Shevchenko National University of Kyiv, no. 1, pp. 24-30, 2016.
34. M. P. Moklyachuk and M. I. Sidei, *Filtering problem for stationary sequences with missing observations*, Statistics, Optimization & Information Computing, vol. 4, no.4, pp. 308-325, 2016.
35. M. Moklyachuk and M. Sidei, *Filtering Problem for functionals of stationary processes with missing observations*, Communications in Optimization Theory, 2016, pp. 1-18, Article ID 21, 2016.
36. M. M. Pelagatti, *Time series modelling with unobserved components*, New York: CRC Press, 2015.
37. M. Pourahmadi, *Foundations of time series analysis and prediction theory*, New York: Wiley, 2001.
38. M. Pourahmadi, A. Inoue and Y. Kasahara, *A prediction problem in  $L^2(w)$* , Proceedings of the American Mathematical Society, vol. 135, no.4, pp 1233-1239, 2007.
39. B. N. Pshenichnyj, *Necessary conditions of an extremum*, Pure and Applied mathematics. 4. New York: Marcel Dekker, 1971.
40. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1997.
41. Yu. A. Rozanov, *Stationary stochastic processes*, San Francisco-Cambridge-London-Amsterdam: Holden-Day, 1967.
42. K. S. Vastola and H. V. Poor, *An analysis of the effects of spectral uncertainty on Wiener filtering*, Automatica, vol. 28, pp. 289-293, 1983.
43. N. Wiener, *Extrapolation, interpolation and smoothing of stationary time series. With engineering applications*, The M. I. T. Press, Massachusetts Institute of Technology, Cambridge, Mass., 1966.
44. A. M. Yaglom, *Correlation theory of stationary and related random functions. Vol. 1: Basic results*, Springer Series in Statistics, Springer-Verlag, New York etc., 1987.
45. A. M. Yaglom, *Correlation theory of stationary and related random functions. Vol. 2: Supplementary notes and references*, Springer Series in Statistics, Springer-Verlag, New York etc., 1987.