



Minimax-robust estimation problems for stationary stochastic sequences

Mikhail Moklyachuk

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

(Received: 4 October 2015; Accepted: 9 October 2015)

Abstract This survey provides an overview of optimal estimation of linear functionals which depend on the unknown values of a stationary stochastic sequence. Based on observations of the sequence without noise as well as observations of the sequence with a stationary noise, estimates could be obtained. Formulas for calculating the spectral characteristics and the mean-square errors of the optimal estimates of functionals are derived in the case of spectral certainty, where spectral densities of the sequences are exactly known. In the case of spectral uncertainty, where spectral densities of the sequences are not known exactly while sets of admissible spectral densities are given, the minimax-robust method of estimation is applied. Formulas that determine the least favourable spectral densities and the minimax spectral characteristics of estimates are presented for some special classes of admissible spectral densities.

Keywords Stochastic sequence, minimax-robust estimate, mean square error, least favourable spectral density, minimax-robust spectral characteristic

AMS 2010 subject classifications. Primary: 60G10, 60G25, 60G35, Secondary: 62M20, 62P20, 93E10, 93E11

DOI: 10.19139/soic.v3i4.173

Contents

1	Introduction	349
2	Stationary stochastic sequences. Spectral representation	351
2.1	Definitions. Examples	351
2.2	Spectral representation of stationary sequences	353
2.3	Linear filters of stationary sequences	354
2.4	Wold expansion of stationary sequences	359
2.5	Conclusions	363
3	Estimates for functionals of stationary sequences	363
3.1	The maximum value of the mean-square error of estimate of the functional $A_N \xi$	364
3.2	The maximum value of the mean square error of estimate of the functional $A \xi$	368
3.3	Conclusions	372
4	Extrapolation problem for functionals of stationary sequences	372
4.1	The classical Hilbert space projection method of linear extrapolation	373
4.2	Minimax-robust method of linear extrapolation of functionals	375
4.3	Least favourable spectral densities in the class D_0	377

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

4.4	Least favourable spectral densities in the class D_M	379
4.5	Least favourable spectral densities in the class D_v^u	380
4.6	Least favourable spectral densities in the class D_ε	382
4.7	Least favourable spectral densities in the class $D_{1\varepsilon}$	383
4.8	Least favourable spectral densities in the class $D_{2\varepsilon}$	384
4.9	Conclusions	385
5	Extrapolation problem for stationary sequences from observations with noise	386
5.1	The classical Hilbert space projection method of linear extrapolation	386
5.2	Minimax-robust method of extrapolation	390
5.3	Least favourable spectral densities in the class $D_f^0 \times D_g^0$	391
5.4	Least favourable spectral densities in the class $D_v^u \times D_\varepsilon$	393
5.5	Least favourable spectral densities in the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$	394
5.6	Conclusions	395
6	Interpolation problem for stationary sequences	395
6.1	The classical Hilbert space projection method of linear interpolation	396
6.2	Minimax-robust method of interpolation	400
6.3	Least favourable spectral densities in the class D_0^-	401
6.4	Least favourable spectral densities in the class D_M	404
6.5	Least favourable spectral densities in the class D_v^{u-}	405
6.6	Conclusions	406
7	Interpolation problem for stationary sequences from observations with noise	406
7.1	The classical Hilbert space projection method of linear interpolation	406
7.2	Minimax-robust method of interpolation	410
7.3	Least favourable spectral densities in the class $D_f^0 \times D_g^0$	412
7.4	Least favourable spectral densities in the class $D_v^u \times D_\varepsilon$	413
7.5	Least favourable spectral densities in the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$	414
7.6	Conclusions	416
8	Conclusion Remarks	416

1. Introduction

Theory of estimation of the unknown values of stationary stochastic processes based on a set of observations of the processes plays an important role in many practical applications. The development of the theory started from the classical works of Kolmogorov [24] and Wiener [77], in which they presented methods of solution of the extrapolation and interpolation problems for stationary processes. The interpolation problem considered by Kolmogorov means estimation of the missed values of stochastic sequences. The prediction problem consists in estimation the future values of the process based on observations of the process in the past. The third classical problem is filtering of stochastic processes which consists in estimation the original values of the signal process from observations of the process with noise. All these problems for stationary sequences and processes are clearly described in books by Gikhman and Skorokhod [12], Hannan [17], Rozanov [73], Yaglom [80, 81].

One of the fields of practical applications of the stationary and related stochastic sequences is economical modelling and financial time series. Most simple examples of stationary linear models are moving average (MA) sequences, autoregressive (AR) and autoregressive-moving average (ARMA) sequences, state space models, all of which refer to stationary sequences with rational spectral function without unit AR-roots. The main results concerning the model description, parameter estimation, forecasting and further investigations are described in the

classical book by Box, Jenkins and Reinsel [2]. Most of results which have appeared since that time were based on the assumption that the spectral structure of the stationary sequence is known.

Traditional methods of solution of linear extrapolation, interpolation and filtering problems for stationary stochastic sequences may be employed under the condition that spectral densities of the sequences are known exactly (see, for example, selected works of Kolmogorov [24], survey article by Kailath [21], books by Rozanov [73], Wiener [77], Yaglom [80, 81]). In practice, however, complete information on the spectral densities is impossible in most cases. To solve the problem the parametric or nonparametric estimates of the unknown spectral densities are found or these densities are selected by other reasoning. Then the classical estimation method is applied provided that the estimated or selected densities are the true one. This procedure can result in a significant increasing of the value of the error as Vastola and Poor [76] have demonstrated with the help of some examples. This is a reason to search estimates which are optimal for all densities from a certain class of the admissible spectral densities. These estimates are called minimax since they minimize the maximal value of the error. Many investigators have been interested in minimax extrapolation, interpolation and filtering problems for stationary stochastic sequences. See, for example, the papers by Hosoya [19] and Taniguchi [75]. A survey of results in minimax (robust) methods of data processing can be found in the paper by Kassam and Poor [23]. The paper by Grenander [15] should be marked as the first one where the minimax approach to extrapolation problem for stationary processes was proposed. Franke [9, 10], Franke and Poor [11] investigated the minimax extrapolation and filtering problems for stationary sequences with the help of convex optimization methods. This approach makes it possible to find equations that determine the least favourable spectral densities for various classes of admissible densities.

In the works by Moklyachuk [35] – [58] problems of extrapolation, interpolation and filtering for stationary processes and sequences were studied. The corresponding problems for vector-valued stationary sequences and processes were investigated by Moklyachuk and Masyutka [59] – [64]. In the articles by Dubovets'ka and Moklyachuk [3] - [7] and in the book by Golichenko and Moklyachuk [13] the minimax estimation problems were investigated for another generalization of stationary processes – periodically correlated stochastic sequences and processes. Luz and Moklyachuk [26] – [32], [34] investigated the classical and minimax extrapolation, interpolation and filtering problems for sequences and processes with n th stationary increments.

In the proposed paper we deal with the problem of estimation of the functional $A\xi = \sum_{k=0}^{\infty} a(k)\xi(k)$ which depends on the unknown values of a stationary stochastic sequence $\xi(k)$ based on observations of the sequence $\xi(k) + \eta(k)$, where $\eta(k)$ is an uncorrelated with the sequence $\xi(k)$ stationary stochastic sequence. Formulas for calculating the mean-square errors and the spectral characteristics of the optimal estimate of the functional are derived in the case of spectral certainty, where the spectral densities of the sequences $\xi(n)$ and $\eta(n)$ are exactly known. The estimation problem is solved in the case of spectral uncertainty, where the spectral densities of sequences are not exactly known, but, instead, a set of admissible spectral densities is given. The minimax-robust estimation method is applied in this case. Formulas that determine the least favourable spectral densities and the minimax-robust spectral characteristic of the optimal linear estimates of the functional $A\xi$ are derived in the case of spectral uncertainty for some concrete classes of admissible spectral densities.

The paper is organized as follows. The spectral representations of stationary stochastic sequences and their correlation functions are described in Section 2. Factorizations of spectral densities with relations to one-sided moving average representations and applications to the problem of prediction of stationary sequences are discussed in this section. In section 3 the problem of the mean square optimal linear estimation of the functionals which depend on the unknown ‘future’ values of a stationary stochastic sequence from a class of stationary stochastic sequences is investigated. Estimates are based on observations of the sequence in the ‘past’. Following the Grenander [15] approach to investigation the problem of optimal linear estimation of the linear functional we consider the problem as a two-person zero-sum game. It is show that this game has equilibrium point. In Section 4 the extrapolation problem for functionals of stationary sequences is investigated in the case of spectral certainty, where the spectral density of the stationary stochastic sequence is exactly known, as well as in the case of spectral uncertainty, where the spectral density of the sequence is not exactly known, but a class of admissible spectral densities is given. The classical Hilbert space projection method and the minimax-robust procedure to extrapolation of the functional are applied. The corresponding results of investigation of the extrapolation problem for functionals

of stationary sequences from observations with noise are presented in Section 5. In Section 6 and Section 7 results of investigation of the interpolation problem for functionals of stationary sequences from observations without noise and with noise are presented. The problem is investigated in the case of spectral certainty as well as in the case of spectral uncertainty. Formulas are proposed for calculation the value of the mean square errors and the spectral characteristics of the interpolation in the case of spectral certainty. Relations which determine the least favourable spectral densities and the minimax spectral characteristics are proposed in the case of spectral uncertainty for some special classes of spectral densities.

2. Stationary stochastic sequences. Spectral representation

In this section properties of stationary stochastic sequences illustrated with the help of some examples are presented. The spectral representation of stationary stochastic sequences and their correlation functions are described. Basic properties of linear transformations of stationary stochastic sequences are proposed. The Wold representation of stationary stochastic sequences is studied with application to the problem of prediction of stationary sequences.

2.1. Definitions. Examples

Let $H = L_2(\Omega, \mathcal{F}, P)$ be the space of all complex valued random variables $\xi = \xi(\omega), \omega \in \Omega$, such that $E|\xi|^2 < \infty$. Covariation of the random variables $\xi, \eta \in H$ is defined as the quantity

$$\text{cov}(\xi, \eta) = E(\xi - E\xi)\overline{(\eta - E\eta)}. \tag{1}$$

Definition 2.1. A sequence of complex valued random variables $\xi(n), n \in \mathbb{Z}$, with values in the space $H = L_2(\Omega, \mathcal{F}, P)$ is called *wide sense stationary* if for all $n \in \mathbb{Z}$ and $k \in \mathbb{Z}$ the following relations hold true

$$\begin{aligned} E\xi(n) &= E\xi(0) = C = \text{const}, \\ \text{cov}(\xi(k+n), \xi(k)) &= \text{cov}(\xi(n), \xi(0)). \end{aligned} \tag{2}$$

In the following we will assume that $E\xi(n) = 0, n \in \mathbb{Z}$. This assumption gives us a possibility to consider the covariation as the inner product in the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ and use the powerful methods and results of the Hilbert space theory.

Definition 2.2. The function

$$R(n) = \text{cov}(\xi(n), \xi(0)), \quad n \in \mathbb{Z}, \tag{3}$$

is called the *correlation function* of the wide sense stationary sequence $\xi(n)$.

Example 1

White noise sequence. The white noise sequence $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ is a sequence of orthogonal random variables such that $E\varepsilon(n) = 0, E\varepsilon(i)\overline{\varepsilon(j)} = \delta_{ij}$, where δ_{ij} is the Kroneker symbol. This sequence is stationary with the correlation function

$$R(n) = \int_{-\pi}^{\pi} e^{i\lambda n} dF(\lambda), \tag{4}$$

$$F(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{\lambda} f(v)dv, \quad f(v) = 1, \quad -\pi < v < \pi. \tag{5}$$

The white noise sequence has the absolutely continuous spectral function $F(\lambda)$ with the constant spectral density function $f(v) = 1$.

Example 2

Moving average sequence. Let $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ be a white noise sequence. Consider the sequence

$$\xi(n) = \sum_{k=-\infty}^{\infty} a(k)\varepsilon(n-k), \quad (6)$$

where $a(k)$ are complex valued numbers such that $\sum_{k=-\infty}^{\infty} |a(k)|^2 < \infty$. By the Parseval equality

$$\text{cov}(\xi(n+m), \xi(m)) = \text{cov}(\xi(n), \xi(0)) = \sum_{k=-\infty}^{\infty} a(n+k)\overline{a(k)}.$$

For this reason $\xi = \{\xi(n), n \in \mathbb{Z}\}$ is a stationary sequence, which is called (two-sided) moving average (MA) sequence generated by the white noise sequence $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$.

In the case where $a(k) = 0$ for all $k = -1, -2, \dots$, the sequence

$$\xi(n) = \sum_{k=0}^{\infty} a(k)\varepsilon(n-k), \quad (7)$$

is called one-sided moving average (MA) sequence generated by the white noise sequence $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$.

In the case where $a(k) = 0$ for all $k = -1, -2, \dots$ and all $k = p+1, p+2, \dots$, the sequence

$$\xi(n) = a(0)\varepsilon(n) + a(1)\varepsilon(n-1) + \dots + a(p)\varepsilon(n-p), \quad (8)$$

is called moving average $MA(p)$ sequence of order p generated by the white noise sequence $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$.

The correlation function $R(n)$ of the moving average sequence $MA(p)$ is of the form

$$R(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda n} f(\lambda) d\lambda,$$

with the spectral density function

$$f(\lambda) = |P(e^{-i\lambda})|^2, \quad (9)$$

where

$$P(z) = a(0) + a(1)z + \dots + a(p)z^p.$$

is a polynomial of order p .

Example 3

The autoregressive sequence $AR(q)$. Let $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ be a white noise sequence. A sequence $\xi = \{\xi(n), n \in \mathbb{Z}\}$ is called the autoregressive sequence of order q if it satisfies the equation

$$\xi(n) + b(1)\xi(n-1) + \dots + b(q)\xi(n-q) = \varepsilon(n). \quad (10)$$

In the case where all zeros of the polynomial

$$Q(z) = 1 + b_1z + \dots + b_qz^q \quad (11)$$

are outside the unit disk $D = \{z : |z| < 1\}$, equation (10) has a stationary solution which can be represented as one-sided moving average sequence. In this case the correlation function $R(n)$ is of the form

$$R(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda n} f(\lambda) d\lambda, \quad (12)$$

$$f(\lambda) = \frac{1}{|Q(e^{-i\lambda})|^2}.$$

Example 4

The autoregressive and moving average $ARMA(p, q)$ sequence. Suppose that in the right-hand side of the equation (10) instead of $\varepsilon(n)$ we put the sequence $a_0\varepsilon(n) + a(1)\varepsilon(n - 1) + \dots + a(p)\varepsilon(n - p)$. We get the so called autoregressive and moving average $ARMA(p, q)$ sequence of order (p, q)

$$\xi(n) + b(1)\xi(n - 1) + \dots + b(q)\xi(n - q) = a(0)\varepsilon(n) + a(1)\varepsilon(n - 1) + \dots + a(p)\varepsilon(n - p). \tag{13}$$

Under the same conditions on zeros of the polynomial $Q(z)$ as in the previous example, equation (13) has a stationary solution $\xi = \{\xi(n), n \in \mathbb{Z}\}$. The correlation function $R(n)$ of the sequence $\xi = \{\xi(n), n \in \mathbb{Z}\}$ is of the form

$$R(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda n} f(\lambda) d\lambda, \tag{14}$$

$$f(\lambda) = \left| \frac{P(e^{-i\lambda})}{Q(e^{-i\lambda})} \right|^2.$$

2.2. Spectral representation of stationary sequences

Theorem 1

Herglotz theorem. Let $R(n)$ be the covariance function of a (wide sense) stationary stochastic sequence $\xi = \{\xi(n), n \in \mathbb{Z}\}$ with zero mean value. Then on the space $([-\pi, \pi], \mathcal{B}([-\pi, \pi]))$ there is a finite measure $F = F(B)$, $B \in \mathcal{B}([-\pi, \pi])$, such that for every $n \in \mathbb{Z}$

$$R(n) = \int_{-\pi}^{\pi} e^{i\lambda n} F(d\lambda), \tag{15}$$

where $\int_{-\pi}^{\pi} e^{i\lambda n} F(d\lambda)$ is the Lebesgue-Stieltjes integral.

Remark 2.1. The measure $F = F(B)$ in the representation (15) is called the spectral measure, and the function $F(\lambda) = F([-\pi, \lambda])$ is called the spectral function of the stationary sequence with the covariance function $R(n)$.

Remark 2.2. In the case where $\sum_{n=-\infty}^{\infty} |R(n)| < \infty$ the spectral function $F(\lambda)$ has the spectral density $f(\lambda)$, which is determined by the formula

$$f(\lambda) = \sum_{n=-\infty}^{\infty} e^{-in\lambda} R(n).$$

The following theorem provides the corresponding spectral representation of the stationary stochastic sequence $\xi = \{\xi(n), n \in \mathbb{Z}\}$ itself.

Theorem 2

Let $\xi = \{\xi(n), n \in \mathbb{Z}\}$ be a stationary stochastic sequence with zero mean value $E\xi(n) = 0$, $n \in \mathbb{Z}$. There is an orthogonal stochastic measure $Z = Z(\Delta)$, $\Delta \in \mathcal{B}([-\pi, \pi])$, such that for every $n \in \mathbb{Z}$ the following spectral representation is valid

$$\xi(n) = \int_{-\pi}^{\pi} e^{i\lambda n} Z(d\lambda) \left(= \int_{[-\pi, \pi]} e^{i\lambda n} Z(d\lambda) \right). \tag{16}$$

Moreover, $E|Z(\Delta)|^2 = F(\Delta)$.

Let $\xi = \{\xi(n), n \in \mathbb{Z}\}$ be a stationary stochastic sequence with the spectral representation (16). Let $H(\xi)$ be the closed in the mean-square sense linear manifold spanned by the random variables $\xi(n)$, $n \in \mathbb{Z}$. Let $\eta \in H(\xi)$. The following theorem describes the structure of such random variables.

Theorem 3

For any $\eta \in H(\xi)$ there is a function $\varphi \in L^2(F)$ such that

$$\eta = \int_{-\pi}^{\pi} \varphi(\lambda)Z(d\lambda). \tag{17}$$

Remark 2.3. Let $H_0(\xi)$ and $L_0^2(F)$ be the corresponding closed linear manifolds spanned by the variables $\{\xi(n), n = 0, -1, -2, \dots\}$ and by the functions $\{e_n = e^{in\lambda}, n = 0, -1, -2, \dots\}$. Then for $\eta \in H_0(\xi)$ there is a function $\varphi \in L_0^2(F)$ such that

$$\eta = \int_{-\pi}^{\pi} \varphi(\lambda)Z(d\lambda).$$

2.3. Linear filters of stationary sequences

Consider a special but important class of linear transformations that are defined by means of the so called linear filters. Suppose that, at instant of time m , a system (filter) receives as input a signal $x(m)$, and that the output of the system is, at instant of time n , the signal $h(n - m)x(m)$, where $h = h(m)$, $m \in \mathbb{Z}$, is a complex valued function called the *impulse response function* of the filter. The total signal $y(n)$ obtained from the input $x(m)$, $m \in \mathbb{Z}$, can be represented in the form

$$y(n) = \sum_{m=-\infty}^{\infty} h(n - m)x(m). \tag{18}$$

For physically realizable systems values of the output at instant of time n are determined only by the "past" values of the signal, i.e. the values $x(m)$ at instants $m \leq n$. It is therefore natural to call a filter with the impulse response function $h = h(m)$ physically realizable (causal filter), if $h(m) = 0$ for all $m < 0$, in other words if

$$y(n) = \sum_{m=-\infty}^n h(n - m)x(m) = \sum_{m=0}^{\infty} h(m)x(n - m). \tag{19}$$

An important spectral characteristic of a filter with the impulse response function $h = h(m)$ is its Fourier transform

$$\varphi(\lambda) = \sum_{m=-\infty}^{\infty} e^{-im\lambda}h(m), \tag{20}$$

which is called the *spectral characteristic* of the filter.

Let us now describe conditions under which the series in (18) and (20) convergence. Let us suppose that the input is a stationary stochastic sequence $\xi = \{\xi(n), n \in \mathbb{Z}\}$ with the covariance function $R(n)$ and spectral decomposition (16). If the following condition holds true

$$\sum_{k,l=-\infty}^{\infty} h(k)R(k - l)\overline{h(l)} < \infty, \tag{21}$$

then the series $\sum_{m=-\infty}^{\infty} h(n - m)\xi(m)$ converges in the mean-square sense and therefore there is a stationary sequence $\eta = \{\eta(n), n \in \mathbb{Z}\}$ such that

$$\eta(n) = \sum_{m=-\infty}^{\infty} h(n - m)\xi(m) = \sum_{m=-\infty}^{\infty} h(m)\xi(n - m). \tag{22}$$

In terms of the spectral measure condition (21) is equivalent to the condition that the spectral characteristic of the filter $\varphi(\lambda) \in L^2(F)$, that is

$$\int_{-\pi}^{\pi} |\varphi(\lambda)|^2 F(d\lambda) < \infty. \tag{23}$$

Under the condition (21) or condition (23) from (22) and (16) we find the spectral representation of the sequence $\eta = \{\eta(n), n \in \mathbb{Z}\}$:

$$\eta(n) = \int_{-\pi}^{\pi} e^{i\lambda n} \varphi(\lambda) Z(d\lambda). \tag{24}$$

The correlation function $R_\eta(n)$ of the sequence $\eta = \{\eta(n), n \in \mathbb{Z}\}$ is determined by the formula

$$R_\eta(n) = \int_{-\pi}^{\pi} e^{i\lambda n} |\varphi(\lambda)|^2 F(d\lambda). \tag{25}$$

In particular, if the input to a filter with the spectral characteristic $\varphi = \varphi(\lambda)$ is taken to be white noise $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$, then the output will be a moving average stationary sequence

$$\eta(n) = \sum_{m=-\infty}^{\infty} h(m)\varepsilon(n - m) \tag{26}$$

with the spectral density

$$f_\eta(\lambda) = |\varphi(\lambda)|^2.$$

The following theorem shows that every stationary stochastic sequence with a spectral density can be represented as a moving average sequence.

Theorem 4

Let $\eta = \{\eta(n), n \in \mathbb{Z}\}$ be a stationary stochastic sequence with the spectral density $f_\eta(\lambda)$. Then (possibly on the extended probability space) we can find a white noise sequence $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ and a filter such that the representation (26) is valid.

Proof

For a given (nonnegative) spectral density function $f_\eta(\lambda)$ we can find a function $\varphi(\lambda)$ such that

$$f_\eta(\lambda) = |\varphi(\lambda)|^2.$$

Since

$$\int_{-\pi}^{\pi} f_\eta(\lambda) d\lambda < \infty,$$

we have $\varphi(\lambda) \in L^2(\lambda)$, where λ is the Lebesgue measure on $[-\pi, \pi]$. Hence the function $\varphi(\lambda)$ can be represented as a Fourier series (20) with

$$h(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{im\lambda} \varphi(\lambda) d\lambda,$$

where convergence is understood in the sense that

$$\int_{-\pi}^{\pi} \left| \varphi(\lambda) - \sum_{|m| \leq n} e^{-im\lambda} h(m) \right|^2 d\lambda \rightarrow 0, \quad n \rightarrow \infty.$$

Let

$$\eta(n) = \int_{-\pi}^{\pi} e^{in\lambda} Z(d\lambda), \quad n \in \mathbb{Z}.$$

In addition to the random measure $Z = Z(\Delta)$ we introduce another (independent) orthogonal stochastic measure $\tilde{Z} = \tilde{Z}(\Delta)$ with $M|\tilde{Z}(a, b)|^2 = \frac{b-a}{2\pi}$ and construct the measure

$$\bar{Z}(\Delta) = \int_{\Delta} \varphi^{\oplus}(\lambda) Z(d\lambda) + \int_{\Delta} [1 - \varphi^{\oplus}(\lambda)\varphi(\lambda)] \tilde{Z}(d\lambda),$$

where a^{\oplus} is the pseudoinverse operation

$$a^{\oplus} = \begin{cases} a^{-1}, & a \neq 0, \\ 0, & a = 0. \end{cases}$$

The stochastic measure $\bar{Z} = \bar{Z}(\Delta)$ is a measure with orthogonal values, and for every $\Delta = (a, b]$ we have

$$E|\bar{Z}(\Delta)|^2 = \frac{1}{2\pi} \int_{\Delta} |\varphi^{\oplus}(\lambda)|^2 |\varphi(\lambda)|^2 d\lambda + \frac{1}{2\pi} \int_{\Delta} |1 - \varphi^{\oplus}(\lambda)\varphi(\lambda)|^2 d\lambda = \frac{|\Delta|}{2\pi},$$

where $|\Delta| = b - a$. Therefore the stationary sequence

$$\varepsilon(n) = \int_{-\pi}^{\pi} e^{in\lambda} \bar{Z}(d\lambda), \quad n \in \mathbb{Z},$$

is a white noise sequence. Note that

$$\int_{-\pi}^{\pi} e^{in\lambda} \varphi(\lambda) \bar{Z}(d\lambda) = \int_{-\pi}^{\pi} e^{in\lambda} Z(d\lambda) = \eta(n), \quad (27)$$

and, from the other hand,

$$\begin{aligned} \int_{-\pi}^{\pi} e^{in\lambda} \varphi(\lambda) \bar{Z}(d\lambda) &= \int_{-\pi}^{\pi} e^{in\lambda} \left(\sum_{m=-\infty}^{\infty} e^{-im\lambda} h(m) \right) \bar{Z}(d\lambda) = \\ &= \sum_{m=-\infty}^{\infty} h(m) \int_{-\pi}^{\pi} e^{i(n-m)\lambda} \bar{Z}(d\lambda) = \sum_{m=-\infty}^{\infty} h(m) \varepsilon(n-m). \end{aligned}$$

This equality, together with (27), establishes the representation (26). This completes the proof of the theorem. \square

Corrolary 2.1

Let the spectral density function $f_{\eta}(\lambda) > 0$ (almost everywhere with respect to the Lebesgue measure) and let the following factorization of the function $f_{\eta}(\lambda) > 0$ holds true

$$f_{\eta}(\lambda) = |\varphi(\lambda)|^2, \quad (28)$$

$$\varphi(\lambda) = \sum_{k=0}^{\infty} e^{-i\lambda k} h(k), \quad \sum_{k=0}^{\infty} |h(k)|^2 < \infty. \quad (29)$$

Then the sequence $\eta = \{\eta(n), n \in \mathbb{Z}\}$ admits the representation as one-sided moving average sequence

$$\eta(n) = \sum_{m=0}^{\infty} h(m) \varepsilon(n-m).$$

In particular, if $P(z) = a(0) + a(1)z + \dots + a(p)z^p$ is a polynomial that has no zeros on $\{z : |z| = 1\}$, then the sequence $\eta = \{\eta(n), n \in \mathbb{Z}\}$ with the spectral density

$$f_{\eta}(\lambda) = |P(e^{-i\lambda})|^2$$

can be represented in the form

$$\eta(n) = a(0)\varepsilon(n) + a(1)\varepsilon(n - 1) + \dots + a(p)\varepsilon(n - p),$$

where $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ is a white noise sequence.

Corrolary 2.2

Let $\xi = \{\xi(n), n \in \mathbb{Z}\}$ be a sequence with rational spectral density

$$f(\lambda) = \left| \frac{P(e^{-i\lambda})}{Q(e^{-i\lambda})} \right|^2, \tag{30}$$

where polynomials $P(z) = a(0) + a(1)z + \dots + a(p)z^p$ and $Q(z) = 1 + b(1)z + \dots + b(q)z^q$ have no zeros on $\{z : |z| = 1\}$. Then there is a white noise sequence $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ such that the following equation holds true

$$\xi(n) + b(1)\xi(n - 1) + \dots + b(q)\xi(n - q) = a(0)\varepsilon(n) + a(1)\varepsilon(n - 1) + \dots + a(p)\varepsilon(n - p). \tag{31}$$

Conversely, every stationary stochastic sequence $\xi = \{\xi(n), n \in \mathbb{Z}\}$ that satisfies this equation with a white noise $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ and polynomial $Q(z) = 1 + b(1)z + \dots + b(q)z^q$ with no zeros on $\{z : |z| = 1\}$ has the spectral density (30).

Let us describe conditions under which a function $f(\lambda)$ admits the representation (28), (29). First, denote by H_2 the set of all analytic in the unit disk $D = \{z : |z| < 1\}$ functions $f(z)$ such that

$$|f(z)|^2 = \lim_{r \uparrow 1} \int_{-\pi}^{\pi} |f(re^{i\theta})|^2 d\theta < \infty.$$

For the function

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

we have

$$f(re^{i\theta}) = \sum_{n=0}^{\infty} a_n r^n e^{in\theta}.$$

It means that $a_n r^n$ are the Fourier coefficients of the function $f(re^{i\theta})$. From the Parseval equality it follows that

$$\int_{-\pi}^{\pi} |f(re^{i\theta})|^2 d\theta = 2\pi \sum_{n=0}^{\infty} |a_n|^2 r^{2n}.$$

We can conclude that $f(z) \in H_2$ if and only if

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty.$$

For any function $f(z) \in H_2$ we can determine a series

$$f(e^{i\theta}) = \sum_{n=0}^{\infty} a_n e^{in\theta},$$

which converges in $L_2[-\pi, \pi]$. The function $f(z), |z| < 1$, restores from the function $f(e^{i\theta})$ with the help of the Poisson formula

$$f(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{iu}) P(r, \theta, u) du,$$

where

$$P(r, \theta, u) = \frac{1 - r^2}{1 - 2r \cos(\theta - u) + r^2} = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in(\theta - u)}.$$

Theorem 5

Let $f(\lambda)$ be a nonnegative and integrable with respect to the Lebesgue measure on $[-\pi, \pi)$ function. In order that a function $g(z) \in H_2$ exists such that

$$f(\lambda) = |g(e^{i\lambda})|^2,$$

it is necessary and sufficient that the following condition holds

$$\int_{-\pi}^{\pi} |\ln f(\lambda)| d\lambda < \infty. \quad (32)$$

Proof

Let condition (32) holds true. Then the function

$$v(r, \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f(\lambda) P(r, \theta, \lambda) d\lambda$$

is harmonic in the unit disk $D = \{z : |z| < 1\}$. From the Jensen inequality it follows that

$$v(r, \theta) \leq \ln \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) P(r, \theta, \lambda) d\lambda \right\}.$$

Denote by $\varphi(z)$ the analytic in the unit disk $D = \{z : |z| < 1\}$ function with the real part $v(r, \theta)$. Consider the function $g(z) = e^{\varphi(z)/2}$. For this function

$$|g(re^{i\theta})|^2 = e^{\operatorname{Re}\varphi(z)} = e^{v(r, \theta)} \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) P(r, \theta, \lambda) d\lambda,$$

$$\int_{-\pi}^{\pi} |g(re^{i\theta})|^2 d\theta \leq \int_{-\pi}^{\pi} f(\lambda) d\lambda.$$

So, the function $g(z) \in H_2$ and

$$\lim_{r \uparrow 1} |g(re^{i\theta})|^2 = e^{\lim_{r \uparrow 1} v(r, \theta)} = f(\theta)$$

almost everywhere. □

Corollary 2.3

In order that a stochastic sequence $\eta = \{\eta(n), n \in \mathbb{Z}\}$ admits the representation as one-sided moving average sequence

$$\eta(n) = \sum_{m=0}^{\infty} a(m) \varepsilon(n-m), \quad \sum_{m=0}^{\infty} |a(m)|^2 < \infty.$$

where $\varepsilon = \{\varepsilon(n), n \in \mathbb{Z}\}$ is a white noise sequence, it is necessary and sufficient that the sequence $\eta = \{\eta(n), n \in \mathbb{Z}\}$ has an absolutely continuous spectral measure and its spectral density $f(\lambda)$ satisfies the condition

$$\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda > -\infty. \quad (33)$$

Determined in the theorem analytical in the unit disk $D = \{z : |z| < 1\}$ function $\varphi(z)$ with the real part $v(r, \theta)$ has the boundary value $\ln f(\lambda)$. By the Schwartz formula we have

$$\varphi(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f(\lambda) \frac{e^{i\lambda} + z}{e^{i\lambda} - z} d\lambda. \quad (34)$$

It follows from the decomposition of the function $g(z) = \exp\{\frac{\varphi(z)}{2}\}$ into a power series $g(z) = \sum_{n=0}^{\infty} b_n z^n$ that the coefficients $a_n = \sqrt{2\pi}b_n$. From the other hand, the function $g(z)$ can be represented in a different form. Since

$$\frac{e^{i\lambda} + z}{e^{i\lambda} - z} = 1 + \frac{2ze^{-i\lambda}}{1 - ze^{-i\lambda}} = 1 + 2 \sum_{k=1}^{\infty} z^k e^{-ik\lambda},$$

we have

$$\overline{g(z)} = \exp \left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln f(\lambda) d\lambda + \frac{1}{2\pi} \sum_{k=1}^{\infty} d_k \bar{z}^k \right\},$$

$$d_k = \int_{-\pi}^{\pi} e^{ik\lambda} \ln f(\lambda) d\lambda.$$

If we introduce the notation

$$P = \exp \left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln f(\lambda) d\lambda \right\},$$

$$\exp \left\{ \frac{1}{2\pi} \sum_{k=1}^{\infty} d_k z^k \right\} = \sum_{n=0}^{\infty} c_n z^n, \quad c_0 = 1,$$

we will have

$$\overline{g(z)} = P \sum_{n=0}^{\infty} c_n \bar{z}^n.$$

So

$$a_n = \sqrt{2\pi} P c_n.$$

2.4. Wold expansion of stationary sequences

In contrast to the representation (16) which gives an expansion of a stationary sequence in the frequency domain, the Wold [78, 78] expansion operates in the time domain. The main point of this expansion is that a stationary sequence $\xi(n)$ can be represented as a sum of two stationary sequences, one of which is completely predictable (in the sense that its values are completely determined by its “past), whereas the second does not have this property.

First we give some definitions. Let $H(\xi)$ and $H_n(\xi)$ be closed linear manifolds, spanned by all values of the stationary sequence $\{\xi(k), k \in \mathbb{Z}\}$ and values $\{\xi(k), k = n, n - 1, n - 2, \dots\}$ respectively. Let

$$S(\xi) = \bigcap_n H_n(\xi).$$

For every element $\eta \in H(\xi)$ denote by

$$\hat{\eta}_n = \text{Proj}(\eta|H_n(\xi))$$

the projection of the element η on the subspace $H_n(\xi)$. Denote also

$$\hat{\eta}_S = \text{Proj}(\eta|S(\xi)),$$

Every element $\eta \in H(\xi)$ can be represented in the form

$$\eta = \hat{\eta}_S + (\eta - \hat{\eta}_S),$$

where $\eta - \hat{\eta}_S \perp \hat{\eta}_S$. Therefore the space $H(\xi)$ can be represented as the orthogonal sum

$$H(\xi) = S(\xi) \oplus R(\xi),$$

where $S(\xi)$ consists of the elements $\hat{\eta}_S$ with $\eta \in H(\xi)$, and $R(\xi)$ consists of the elements of the form $\eta - \hat{\eta}_S$.

Definition 2.3. A stationary sequence $\xi(n)$ is called *regular*, if

$$H(\xi) = R(\xi),$$

and *singular*, if

$$H(\xi) = S(\xi).$$

Remark 2.4. *Singular sequence is also called deterministic, and regular sequence is called purely or completely nondeterministic. If $S(\xi)$ is a proper subspace of the space $H(\xi)$, then the sequence $\xi(n)$ is called nondeterministic.*

Theorem 6

Every wide sense stationary random sequence $\xi(n)$ has a unique decomposition

$$\xi(n) = \xi^r(n) + \xi^s(n), \quad (35)$$

where $\xi^r(n)$ is regular sequence and $\xi^s(n)$ is singular sequence. The sequences $\xi^r(n)$ and $\xi^s(n)$ are orthogonal.

Proof

Define

$$\xi^s(n) = \text{Proj}(\xi(n)|S(\xi)), \quad \xi^r(n) = \xi(n) - \xi^s(n).$$

Since $\xi^r(n) \perp S(\xi)$ for every n , we have $S(\xi^r) \perp S(\xi)$. On the other hand, $S(\xi^r) \subseteq S(\xi)$ and therefore $S(\xi^r)$ is trivial (contains only random sequences that coincide almost surely with zero). Consequently $\xi^r(n)$ is regular.

Moreover, $H_n(\xi) \subseteq H_n(\xi^s) \oplus H_n(\xi^r)$ and $H_n(\xi^s) \subseteq H_n(\xi)$, $H_n(\xi^r) \subseteq H_n(\xi)$. Therefore $H_n(\xi) = H_n(\xi^s) \oplus H_n(\xi^r)$, and hence for every n

$$S(\xi) \subseteq H_n(\xi^s) \oplus H_n(\xi^r). \quad (36)$$

Since $\xi_n^r \perp S(\xi)$, it follows from (36) that

$$S(\xi) \subseteq H_n(\xi^s),$$

and therefore $S(\xi) \subseteq S(\xi^s) \subseteq H(\xi^s)$. But $\xi_n^s \in S(\xi)$, hence $H(\xi^s) \subseteq S(\xi)$, and consequently

$$S(\xi) = S(\xi^s) = H(\xi^s),$$

which means that $\xi^s(n)$ is singular.

The orthogonality of $\xi^s(n)$ and $\xi^r(n)$ follows in an obvious way from $\xi^s(n) \in S(\xi)$ and $\xi^r(n) \perp S(\xi)$. \square

Definition 2.4. Let $\xi(n)$ be a nondegenerate stationary sequence. A random sequence $\varepsilon(n)$ is called the *innovation sequence* for $\xi(n)$, if the following conditions holds true:

- 1) $\varepsilon(n)$ consists of pairwise orthogonal random variables with $E\varepsilon(n) = 0$, $E|\varepsilon(n)|^2 = 1$;
- 2) $H_n(\xi) = H_n(\varepsilon)$ for all $n \in \mathbb{Z}$.

Remark 2.5. *The reason for the term "innovation" is that $\varepsilon(n+1)$ provides new "information", not contained in $H_n(\xi)$, that is needed for forming $H_{n+1}(\xi)$.*

The following fundamental theorem establishes a connection between one-sided moving average sequences and regular sequences.

Theorem 7

The necessary and sufficient condition for a nondegenerate stationary sequence $\xi(n)$ to be regular is that there exist an innovation sequence $\varepsilon = \{\varepsilon(n)\}$ and a sequence of complex numbers $\{a(n), n \geq 0\}$, with $\sum_{n=0}^{\infty} |a(n)|^2 < \infty$, such that

$$\xi(n) = \sum_{k=0}^{\infty} a(k)\varepsilon(n-k). \quad (37)$$

Proof

Necessity. Represent $H_n(\xi)$ in the form

$$H_n(\xi) = H_{n-1}(\xi) \oplus B_n(\xi),$$

where $B_n(\xi)$ is the space of random variables of the form $\beta \cdot \xi(n)$, where β is a complex number. Since $H_n(\xi)$ is generated by elements from $H_{n-1}(\xi)$ and elements of the form $\beta \cdot \xi(n)$, the dimension (dim) of the space $B_n(\xi)$ is either zero or one. But the space $H_n(\xi)$ cannot coincide with $H_{n-1}(\xi)$ for any value of n . In fact, if for some n the space $B_n(\xi)$ is trivial, then by stationarity $B_n(\xi)$ is trivial for all n , and therefore $H(\xi) = S(\xi)$, contradicting the assumption that the sequence $\xi(n)$ is regular. Thus the space $B_n(\xi)$ has the dimension 1. Let $\eta(n)$ be a nonzero element of $B_n(\xi)$. Take

$$\varepsilon_n = \frac{\eta(n)}{\|\eta(n)\|},$$

where $\|\eta(n)\|^2 = E|\eta(n)|^2 > 0$. For fixed n and $k \geq 0$ consider the decomposition

$$H_n(\xi) = H_{n-k}(\xi) \oplus B_{n-k+1}(\xi) \oplus \dots \oplus B_n(\xi).$$

The elements $\varepsilon(n-k), \dots, \varepsilon(n)$ is an orthogonal basis in $B_{n-k+1}(\xi) \oplus \dots \oplus B_n(\xi)$ and

$$\xi(n) = \sum_{j=0}^{k-1} a(j)\varepsilon(n-j) + \text{Proj}(\xi(n)|H_{n-k}(\xi)), \quad a(j) = E\xi(n)\overline{\varepsilon(n-j)}. \tag{38}$$

By the Bessel inequality

$$\sum_{j=0}^{\infty} |a(j)|^2 \leq \|\xi(n)\|^2 < \infty.$$

The series

$$\sum_{j=0}^{\infty} a(j)\varepsilon(n-j)$$

converges in the mean square, and then, by (38), relation (37) will be established as soon as we show that

$$\text{Proj}(\xi(n)|H_{n-k}(\xi)) \xrightarrow{L^2} 0, \quad k \rightarrow \infty.$$

Consider the case $n = 0$. Denote $\hat{\xi}_i = \text{Proj}(\xi(0)|H_i(\xi))$. Since

$$\hat{\xi}_{-k} = \hat{\xi}_0 + \sum_{i=1}^k [\hat{\xi}_{-i} - \hat{\xi}_{-i+1}],$$

and the terms that appear in this sum are orthogonal, we have for every $k \geq 0$

$$\begin{aligned} \sum_{i=1}^k \|\hat{\xi}_{-i} - \hat{\xi}_{-i+1}\|^2 &= \left\| \sum_{i=1}^k (\hat{\xi}_{-i} - \hat{\xi}_{-i+1}) \right\|^2 = \\ &= \|\hat{\xi}_{-k} - \hat{\xi}_0\|^2 \leq 4\|\xi(0)\|^2 < \infty. \end{aligned}$$

Therefore the limit $\lim_{k \rightarrow \infty} \hat{\xi}_{-k}$ exists (in the mean square sense). For each k the value $\hat{\xi}_{-k} \in H_{-k}(\xi)$, and therefore the limit in question must belong to the space $\bigcap_{k \geq 0} H_{-k}(\xi) = S(\xi)$. But, by assumption, $S(\xi)$ is trivial, and therefore

$$\text{Proj}(\xi(n)|H_{n-k}(\xi)) \xrightarrow{L^2} 0, \quad k \rightarrow \infty.$$

Sufficiency. Let the nondegenerate stationary sequence $\xi(n)$ have a representation (37), where $\varepsilon = \{\varepsilon(n)\}$ is an orthonormal system (not necessarily satisfying the condition $H_n(\xi) = H_n(\varepsilon)$, $n \in \mathbb{Z}$). Then $H_n(\xi) \subseteq H_n(\varepsilon)$ and therefore $S(\xi) = \bigcap_k H_k(\xi) \subseteq H_n(\varepsilon)$ for every n . But $\varepsilon_{n+1} \perp H_n(\varepsilon)$, and therefore $\varepsilon_{n+1} \perp S(\xi)$ and at the same time $\varepsilon = \{\varepsilon(n)\}$ is a basis in $H(\xi)$. It follows that $S(\xi)$ is trivial, and consequently $\xi(n)$ is regular. This completes the proof of the theorem. \square

Remark 2.6. It follows from the proof of the theorem 7 that a nondegenerate sequence $\xi(n)$ is regular if and only if it admits a representation as one-sided moving average

$$\xi(n) = \sum_{k=0}^{\infty} \tilde{a}_k \tilde{\varepsilon}_{n-k} \tag{39}$$

where $\tilde{\varepsilon} = \{\tilde{\varepsilon}(n)\}$ is an orthonormal system which (it is important to emphasize this !) does not necessarily satisfy the condition $H_n(\xi) = H_n(\tilde{\varepsilon})$, $n \in \mathbb{Z}$. In this sense the conclusion of the theorem 7 says more, and specifically that for a regular stationary sequence $\xi(n)$ there exist a sequence of numbers $a = \{a(n)\}$ and an orthonormal system of random variables $\varepsilon = \{\varepsilon(n)\}$, such that not only (39), but also (37), is satisfied, with $H_n(\xi) = H_n(\varepsilon)$, $n \in \mathbb{Z}$.

Theorem 8

Wold expansion. A nondegenerate stationary sequence $\xi = \{\xi(n)\}$ can be represented in the form

$$\xi(n) = \xi^s(n) + \sum_{k=0}^{\infty} a(k)\varepsilon(n - k), \tag{40}$$

where

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

and $\varepsilon = \{\varepsilon(n)\}$ is the innovation sequence (for $\xi^r(n)$).

The significance of the concepts of regular and singular sequences becomes clear if we consider the following (linear) extrapolation problem, for whose solution the Wold expansion (40) is especially useful.

Let $H_0(\xi)$ be the closed linear manifold spanned by values of the stationary sequence $\{\xi(k), k = 0, -1, -2, \dots\}$. Consider the problem of constructing an optimal (least-squares) linear estimator $\hat{\xi}(n)$ of the value $\xi(n)$ of the sequence at point $n > 0$ based on observations of the “past” $\{\xi(k), k = 0, -1, -2, \dots\}$. The estimate $\hat{\xi}(n)$ is a projection of $\xi(n)$ on the manifold $H_0(\xi)$:

$$\hat{\xi}(n) = \text{Proj}(\xi(n)|H_0(\xi)). \tag{41}$$

Since the sequences $\xi^s(n)$ and $\xi^r(n)$ are orthogonal and $H_0(\xi) \subseteq H_0(\xi^s) \oplus H_0(\xi^r)$, we obtain, by using (40)

$$\begin{aligned} \hat{\xi}(n) &= \text{Proj}(\xi^s(n) + \xi^r(n)|H_0(\xi)) = \text{Proj}(\xi^s(n)|H_0(\xi)) + \text{Proj}(\xi^r(n)|H_0(\xi)) = \\ &= \text{Proj}(\xi^s(n)|H_0(\xi^s) \oplus H_0(\xi^r)) + \text{Proj}(\xi_n^r|H_0(\xi^s) \oplus H_0(\xi^r)) = \\ &= \text{Proj}(\xi^s(n)|H_0(\xi^s)) + \text{Proj}(\xi^r(n)|H_0(\xi^r)) = \\ &= \xi^s(n) + \text{Proj}\left(\sum_{k=0}^{\infty} a(k)\varepsilon(n - k)|H_0(\xi^r)\right). \end{aligned}$$

In (40) the sequence $\varepsilon = \{\varepsilon(n)\}$ is an innovation sequence for $\xi^r = \{\xi^r(n)\}$ and $H_0(\xi^r) = H_0(\varepsilon)$. Therefore

$$\hat{\xi}(n) = \xi^s(n) + \text{Proj}\left(\sum_{k=0}^{\infty} a(k)\varepsilon(n - k)|H_0(\varepsilon)\right) = \xi^s(n) + \sum_{k=n}^{\infty} a(k)\varepsilon(n - k) \tag{42}$$

and the mean-square error of prediction $\xi(n)$ based on observations $\{\xi(k), k = 0, -1, -2, \dots\}$ equals

$$\sigma_n^2 = E|\xi(n) - \hat{\xi}(n)|^2 = \sum_{k=0}^{n-1} |a(k)|^2. \tag{43}$$

Two important conclusions follows from the presented results.

1) If the sequence $\xi(n)$ is singular, then for every $n \geq 1$ the error of the extrapolation σ_n^2 is zero; in other words, we can predict $\xi(n)$ without error from its "past" $\{\xi(k), k = 0, -1, -2, \dots\}$.

2) If the sequence $\xi(n)$ is regular, then $\sigma_n^2 \leq \sigma_{n+1}^2$ and

$$\lim_{n \rightarrow \infty} \sigma_n^2 = \sum_{k=0}^{\infty} |a(k)|^2. \tag{44}$$

Since

$$\sum_{k=0}^{\infty} |a(k)|^2 = E|\xi(n)|^2,$$

it follows from (44) and (43) that if n increases, the prediction of $\xi(n)$ based on observations $\{\xi(k), k = 0, -1, -2, \dots\}$ becomes trivial (reducing simply to $E\xi(n) = 0$).

2.5. Conclusions

In this section we describe properties of stationary stochastic sequence illustrated with the help of some examples. The spectral representation of stationary stochastic sequences and their correlation functions are described. Properties of linear transformations of stationary stochastic sequences are described. The Wold representation of stationary stochastic sequences is analysed. Application to the problem of prediction of stationary sequences is described.

For the detailed exposition of results of the theory of stationary stochastic sequences see books by Gikhman and Skorokhod [12], Hannan[17], Rozanov[73], Yaglom [80, 81].

3. Estimates for functionals of stationary sequences

In this section we deal with the problem of the mean square optimal linear estimation of the functionals

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

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

which depend on the unknown values of a stationary stochastic sequence $\xi(j)$ from the class Ξ of stationary stochastic sequences satisfying the conditions

$$E\xi(j) = 0, \quad E|\xi(j)|^2 \leq P. \tag{45}$$

Estimates are based on results of observations of the sequence $\xi(j)$ at points of time $j = -1, -2, \dots$.

Following the Ulf Grenander [15] approach to investigation the problem of optimal linear estimation of the functional $A_N \xi$ (as well estimation of the functional $A \xi$) we consider the problem as a two-person zero-sum game in which the first player chooses a stationary stochastic sequence $\xi(j)$ from the class Ξ of stationary stochastic sequences such that the value of the mean square error of estimate of the functional attains its maximum. The second player is looking for an estimate of the linear functional which minimizes the value of the mean square error. It is show that this game has equilibrium point. The maximum error gives a moving average stationary sequence which is least favourable in the given class of stationary sequences. The greatest value of the error and the least favourable sequence are determined by the largest eigenvalue and the corresponding eigenvector of the operator determined by the coefficients $a(j)$ which determine the functionals.

3.1. The maximum value of the mean-square error of estimate of the functional $A_N \xi$

In this subsection the maximum value of the mean-square error of the optimal linear estimate of the functional $A_N \xi$ which depend on the unknown values of a stationary stochastic sequence $\xi(j)$ from the class Ξ is found.

Let $\Delta(\xi, \hat{A}_N) = E \left| A_N \xi - \hat{A}_N \xi \right|^2$ denotes the mean-square error of an estimate $\hat{A}_N \xi$ of the functional $A_N \xi$. Denote by Λ the class of all linear estimates of the functional $A_N \xi$.

Theorem 1

The function $\Delta(\xi, \hat{A}_N)$ has a saddle point on the set $\Xi \times \Lambda$. The following relation holds true

$$\min_{\hat{A}_N \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N) = \max_{\xi \in \Xi} \min_{\hat{A}_N \in \Lambda} \Delta(\xi, \hat{A}_N) = P \nu_N^2.$$

The least favourable in the class Ξ of stationary stochastic sequences satisfying conditions (45) for the optimal linear estimation of the functional $A_N \xi$ is one-sided moving average sequence of order N , which is determined by the formula

$$\xi(j) = \sum_{u=j-N}^j \varphi(j-u) \varepsilon(u).$$

Here ν_N^2 is the largest eigenvalue of the selfadjoint compact operator Q_N in the space $\mathbb{C}^{(N+1)}$, determined by the matrix $Q_N = \{Q_N(p, q)\}_{p, q=0}^N$ that is constructed with the elements

$$Q_N(p, q) = \sum_{u=0}^{\min(N-p, N-q)} a(p+u) \bar{a}(q+u), \quad p, q = 0, 1, \dots, N,$$

$\varepsilon(u)$ is a stationary stochastic sequence with orthogonal values:

$$E \varepsilon(i) \bar{\varepsilon}(j) = \delta_{ij},$$

where δ_{ij} is the Kronecker symbol; $\{\varphi(u), u = 0, \dots, N\}$ is determined by the eigenvector that corresponds to ν_N^2 , and the condition $\|\xi(j)\|^2 = P$.

Proof. Lower bound. Denote by Ξ_R the class of all regular stationary sequences which satisfy conditions (45). Since $\Xi_R \subset \Xi$, we have

$$\max_{\xi \in \Xi} \min_{\hat{A}_N \in \Lambda} \Delta(\xi, \hat{A}_N) \geq \max_{\xi \in \Xi_R} \min_{\hat{A}_N \in \Xi_R} \Delta(\xi, \hat{A}_N). \quad (46)$$

A regular stationary sequence $\xi(j)$ admits the canonical representation as an one-sided moving average sequence

$$\xi(j) = \sum_{u=-\infty}^j \varphi(j-u) \varepsilon(u), \quad (47)$$

where $\varepsilon(u)$ is a standard stationary stochastic sequence with orthogonal values, $\{\varphi(u) : u = 0, 1, \dots\}$ are coefficients of the canonical representation (47). The sequence $\xi(j) \in \Xi_R$ is determined if there are determined coefficients $\{\varphi(u) : u = 0, 1, \dots\}$ such that

$$\begin{aligned} E|\xi(j)|^2 &= \left| \sum_{u=-\infty}^j \varphi(j-u) \varepsilon(u) \right|^2 = \\ &= \sum_{u, v=-\infty}^j \varphi(j-u) \overline{\varphi(j-v)} E \varepsilon(u) \bar{\varepsilon}(v) = \end{aligned}$$

$$= \sum_{u=-\infty}^j |\varphi(j-u)|^2 = \sum_{u=0}^{\infty} |\varphi(u)|^2 \leq P. \tag{48}$$

The value of the mean-square error $E \left| A_N \xi - \hat{A}_N \xi \right|^2$ attains its minimum if we choose an estimate $\hat{A}_N \xi$ of the form

$$\hat{A}_N \xi = \sum_{j=0}^N a(j) \hat{\xi}(j),$$

where $\hat{\xi}(j)$ is the optimal estimate of the value of $\xi(j)$ based on observations of the sequence $\xi(p)$ at points $p = -1, -2, \dots$. Taking into consideration the canonical representation (47) of the regular sequence and the form of the optimal estimates of its values

$$\hat{\xi}(j) = \sum_{u=-\infty}^{-1} \varphi(j-u) \varepsilon(u), \tag{49}$$

we can write

$$\begin{aligned} & \min_{\hat{A}_N \in \Lambda} E \left| A_N \xi - \hat{A}_N \xi \right|^2 = \\ &= \sum_{i,j=0}^N a(i) \overline{a(j)} \sum_{u=0}^i \sum_{v=0}^j \varphi(i-u) \overline{\varphi(j-v)} E \varepsilon(u) \overline{\varepsilon(v)} \\ &= \sum_{i,j=0}^N a(i) \overline{a(j)} \sum_{u=0}^{\min(i,j)} \varphi(i-u) \overline{\varphi(j-u)} = \\ &= \sum_{i,j=0}^N a(i) \overline{a(j)} R(i, j), \end{aligned} \tag{50}$$

where

$$R(i, j) = \sum_{u=0}^{\min(i,j)} \varphi(i-u) \overline{\varphi(j-u)}.$$

The change of variables $p = i - u, q = j - u$, gives us a possibility to write (50) in a different form

$$\min_{\hat{A}_N \in \Lambda} E \left| A_N \xi - \hat{A}_N \xi \right|^2 = \sum_{p,q=0}^N \varphi(p) \overline{\varphi(q)} Q_N(p, q), \tag{51}$$

where

$$Q_N(p, q) = \sum_{u=0}^{\min(N-p, N-q)} a(p+u) \overline{a(q+u)}. \tag{52}$$

Denote by Q_N the operator in the space $\mathbb{C}^{(N+1)}$, determined by the matrix $\{Q_N(p, q)\}_{p,q=0}^N$. The operator Q_N is selfadjoint (its matrix is Hermitian) bounded operator. It can be represented in the form $Q_N = A_N \cdot A_N^*$, where the operator A_N is determined by the matrix $\{A_N(p, q)\}_{p,q=0}^N$ with

$$A_N(p, q) = \begin{cases} a(p+q), & p+q \leq N, \\ 0, & p+q > N. \end{cases}$$

For this reason the operator Q_N has nonnegative eigenvalues [1], [71]. It follows from (51), that $\varphi(p) = 0$ for $p > N$. Denote by

$$\vec{\varphi} = \{\tilde{\varphi}(0), \tilde{\varphi}(1), \dots, \tilde{\varphi}(N)\}, \quad \tilde{\varphi}(p) = P^{-1/2} \varphi(p).$$

Condition (48) with this notation is of the form

$$\|\vec{\varphi}\|^2 = \sum_{p=0}^N |\tilde{\varphi}(p)|^2 \leq 1, \tag{53}$$

where $\|\vec{\varphi}\|$ is the norm in the space $\mathbb{C}^{(N+1)}$. Taking into consideration (51) and (53), we can write

$$\min_{\hat{A}_N \in \Lambda} \Delta(\xi, \hat{A}_N) = P \langle Q_N \vec{\varphi}, \vec{\varphi} \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the inner product in the space $\mathbb{C}^{(N+1)}$.

Taking into account (46), we get the following bound from below for the maximum value of the error [8], [14], [71]

$$\max_{\xi \in \Xi} \min_{\hat{A}_N \in \Lambda} \Delta(\xi, \hat{A}_N) \geq P \max_{\|\vec{\varphi}\| \leq 1} \langle Q_N \vec{\varphi}, \vec{\varphi} \rangle = P \nu_N^2, \tag{54}$$

where ν_N^2 is the greatest eigenvalue of the operator Q_N .

Upper bound. To find the upper bound of the minimax values of the error we use the inequality

$$\min_{\hat{A}_N \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N) \leq \min_{\hat{A}_N \in \Lambda_1} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N), \tag{55}$$

where Λ_1 is the class of all linear estimates of the functional $A_N \xi$, which have the form

$$\hat{A}_N \xi = \sum_{j=-\infty}^{-1} c(j) \xi(j). \tag{56}$$

Here $c(j)$ are complex-valued coefficients such that

$$\sum_{j=-\infty}^{-1} |c(j)|^2 < \infty.$$

Taking into consideration the spectral representations of the stationary stochastic sequence $\xi(j)$ and the correlation function of the sequence $\xi(j)$, we can write

$$\begin{aligned} \Delta(\xi, \hat{A}_N) &= E \left| A_N \xi - \hat{A}_N \xi \right|^2 = E \left| \sum_{j=0}^N a(j) \xi(j) - \sum_{j=-\infty}^{-1} c(j) \xi(j) \right|^2 = \\ &= \int_{-\pi}^{\pi} |A_N(e^{i\lambda}) - C(e^{i\lambda})|^2 F_{\xi}(d\lambda), \end{aligned}$$

where

$$A_N(e^{i\lambda}) = \sum_{j=0}^N a(j) e^{ij\lambda}, \quad C(e^{i\lambda}) = \sum_{j=-\infty}^{-1} c(j) e^{ij\lambda}.$$

Here $F_{\xi}(d\lambda)$ is the spectral measure of the stationary stochastic sequence. Restriction (45) is equivalent to the restriction

$$\max_{\xi \in \Xi} \int_{-\pi}^{\pi} F_{\xi}(d\lambda) \leq P. \tag{57}$$

So we can write

$$\max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N) = \max_{\xi \in \Xi} \int_{-\pi}^{\pi} |A_N(e^{i\lambda}) - C(e^{i\lambda})|^2 F_{\xi}(d\lambda) \leq$$

$$\begin{aligned} &\leq \left(\max_{\lambda \in [-\pi, \pi]} |A_N(e^{i\lambda}) - C(e^{i\lambda})|^2 \right) \max_{\xi \in \Xi} \int_{-\pi}^{\pi} F_{\xi}(d\lambda) \\ &\leq P \max_{\lambda \in [-\pi, \pi]} |A_N(e^{i\lambda}) - C(e^{i\lambda})|^2. \end{aligned}$$

To calculate

$$\max_{\lambda \in [-\pi, \pi]} |A_N(e^{i\lambda}) - C(e^{i\lambda})|^2$$

consider the class of all power series

$$f(z) = \sum_{n=0}^{\infty} \alpha(n)z^n,$$

which are regular in the region $|z| < 1$ and have fixed first $N + 1$ coefficients

$$\alpha(n) = d(n), \quad n = 0, 1, \dots, N.$$

Denote by ρ_N^2 the greatest eigenvalue of the matrix

$$H = \{H(p, q)\}_{p, q=0}^N,$$

$$H(p, q) = \sum_{j=0}^{\min(p, q)} d(p-j)\bar{d}(q-j), \quad p, q = 0, 1, \dots, N.$$

It follows from the properties of the power series, that [16]

$$\min_{\{\alpha(n): n \geq N+1\}} \max_{|z|=1} |f(z)|^2 = \rho_N^2.$$

Since in our case

$$d(p) = a(N - p), \quad p = 0, 1, \dots, N,$$

we have to determine the greatest eigenvalue of the matrix

$$G_N = \{G_N(p, q)\}_{p, q=0}^N,$$

$$G_N(p, q) = \sum_{u=0}^{\min(p, q)} a(N - p + u)\bar{a}(N - q + u).$$

Denote this greatest eigenvalue by ω_N^2 . With this notations we have

$$\min_{\hat{A}_N \in \Lambda_1} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N) \leq P\omega_N^2.$$

Taking into account (55), we get

$$\min_{\hat{A}_N \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N) \leq P\omega_N^2. \tag{58}$$

Note, that

$$G_N(N - p, N - q) = Q_N(p, q).$$

For this reason $\omega_N^2 = \nu_N^2$. Comparing (54) and (58), we get

$$\min_{\hat{A}_N \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N) \leq \max_{\xi \in \Xi} \min_{\hat{A}_N \in \Lambda} \Delta(\xi, \hat{A}_N). \tag{59}$$

Since the opposite inequality always holds true, only equality is possible in (59). Proof is complete.

From the proof of the theorem a construction of the optimal minimax estimate follows.

Corollary 1

The optimal minimax estimate $\hat{A}_N \xi$ of the functional $A_N \xi$ is of the form

$$\hat{A}_N \xi = \sum_{j=0}^N a(j) \left(\sum_{u=j-N}^{-1} \varphi(j-u) \varepsilon(u) \right),$$

where $\varepsilon(u)$ is a stationary sequence with orthogonal values, the sequence $\{\varphi(u), u = 0, 1, \dots, N\}$ is uniquely determined by coordinates of the eigenvector of the operator Q_N that corresponds to the greatest eigenvalue ν_N^2 and condition $E |\xi(j)|^2 = P$.

Example 1

Consider the problem of optimal linear stimulation of the functional

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

that depends on the unknown values of a stationary sequence $\xi(j)$, that satisfies the conditions

$$E \xi(j) = 0, \quad E |\xi(j)|^2 \leq 1,$$

based on observations of the sequence $\xi(j)$ at points $j = -1, -2, \dots$.

Eigenvalues of the operator Q_1 , determined by equation (52), are equal to $3 \pm \sqrt{5}$. So the greatest eigenvalue is $\nu_1^2 = 3 + \sqrt{5}$. The eigenvector corresponding to the eigenvalue $\nu_1^2 = 3 + \sqrt{5}$ is of the form $\vec{\varphi} = \{\varphi(0), \varphi(1)\}$, where

$$\varphi(0) = \sqrt{(5 + \sqrt{5})/10}, \quad \varphi(1) = \sqrt{(5 - \sqrt{5})/10}.$$

The least favourable stationary sequence $\xi(j)$ is a moving average sequence of the form

$$\begin{aligned} \xi(j) &= \varphi(0) \varepsilon(j) + \varphi(1) \varepsilon(j-1) = \\ &= \sqrt{(5 + \sqrt{5})/10} \varepsilon(j) + \sqrt{(5 - \sqrt{5})/10} \varepsilon(j+1). \end{aligned}$$

The optimal linear minimax estimate $\hat{A}_1 \xi$ of the functional $A_1 \xi$ is of the form

$$\hat{A}_1 \xi = \varphi(1) \varepsilon(-1) = \sqrt{(5 - \sqrt{5})/10} \varepsilon(-1).$$

The mean-square error of the optimal estimate of the functional $A_1 \xi$ does not exceed $3 + \sqrt{5}$.

3.2. The maximum value of the mean square error of estimate of the functional $A \xi$

In this subsection the maximum value of the mean square error of the optimal linear estimate of the functional

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

which depend on unknown values of a stationary stochastic sequence $\xi(j)$ from the class Ξ is found.

We will suppose that the sequence $\{a(j) : j = 0, 1, \dots\}$ which determines the functional $A \xi$ satisfies conditions

$$\sum_{j=0}^{\infty} |a(j)| < \infty, \quad \sum_{j=0}^{\infty} (j+1) |a(j)|^2 < \infty. \quad (60)$$

Theorem 9

The function $\Delta(\xi, \hat{A}) = E |A\xi - \hat{A}\xi|^2$ has a saddle point on the set $\Xi \times \Lambda$. The following relation holds true

$$\min_{\hat{A} \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}) = \max_{\xi \in \Xi} \min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}) = P\nu^2.$$

The least favourable in the class Ξ of stationary stochastic sequences satisfying conditions (45) for the optimal linear estimation of the functional $A\xi$ is a moving average sequence

$$\xi(j) = \sum_{u=-\infty}^j \varphi(j-u)\varepsilon(u).$$

Here ν^2 is the greatest eigenvalue and $\vec{\varphi} = \{\varphi(u) : u = 0, 1, \dots\}$ is the corresponding eigenvector of the selfadjoint compact operator in the space ℓ_2 determined by the matrix

$$Q = \{Q(p, q)\}_{p, q=0}^{\infty}, \quad Q(p, q) = \sum_{u=0}^{\infty} a(p+u)\overline{a(q+u)},$$

$\varepsilon(u)$ is a stationary sequence with orthogonal values.

Proof. Lower bound. Let $\xi \in \Xi_R$. In this case the following inequality holds true

$$\max_{\xi \in \Xi} \min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}) \geq \max_{\xi \in \Xi_R} \min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}). \tag{61}$$

Making use the canonical representation of the regular stationary sequence (47) as a moving average sequence and form (49) of the optimal estimate we have

$$\begin{aligned} \min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}) &= \min_{\hat{A} \in \Lambda} E |A\xi - \hat{A}\xi|^2 = \\ &= E \left| \sum_{j=0}^{\infty} a(j) \sum_{u=0}^j \varphi(j-u)\varepsilon(u) \right|^2 = \\ &= \sum_{p, q=0}^{\infty} \varphi(p)\overline{\varphi(q)}Q(p, q), \end{aligned} \tag{62}$$

where

$$Q(p, q) = \sum_{u=0}^{\infty} a(p+u)\overline{a(q+u)}. \tag{63}$$

Denote by Q the operator in the space ℓ_2 , determined by the matrix $Q = \{Q(p, q)\}_{p, q=0}^{\infty}$. Since the second condition from (60) is satisfied and

$$\begin{aligned} \sum_{p, q=0}^{\infty} |Q(p, q)|^2 &= \sum_{p, q=0}^{\infty} \left| \sum_{u=0}^{\infty} a(p+u)\overline{a(q+u)} \right|^2 \leq \\ &\leq \sum_{p, q=0}^{\infty} \left(\sum_{u=0}^{\infty} |a(p+u)|^2 \cdot \sum_{u=0}^{\infty} |a(q+u)|^2 \right) = \\ &= \left(\sum_{p=0}^{\infty} \sum_{u=0}^{\infty} |a(p+u)|^2 \right)^2 = \left(\sum_{p=0}^{\infty} (p+1) |a(p)|^2 \right)^2, \end{aligned}$$

we have

$$\|Q\| \leq N(Q) \leq \sum_{p=0}^{\infty} (p+1) |a(p)|^2 < \infty,$$

where $N(Q)$ is the Hilbert-Schmidt norm of the operator Q . The operator Q is selfadjoint (its matrix is Hermitian) Hilbert-Schmidt operator. For these reasons the operator Q is selfadjoint continuous operator [1],[71]. It can be represented in the form $Q = A \cdot A^*$, where the operator A is determined by the matrix

$$A = \{A(p, q)\}_{p, q=0}^{\infty}, \quad A(p, q) = a(p+q), \quad p, q = 0, 1, \dots, \infty.$$

The operator Q has real nonnegative eigenvalues. Note that the operator A is a Hilbert-Schmidt operator and its Hilbert-Schmidt norm equals to

$$N(A) = \left(\sum_{p=0}^{\infty} (p+1) |a(p)|^2 \right)^{1/2}.$$

Let us introduce the notation

$$\vec{\varphi} = \{\tilde{\varphi}(u) : u = 0, 1, \dots\}, \quad \tilde{\varphi}(u) = P^{-1/2} \varphi(u).$$

Making use the introduced notations (62) can be written in the form

$$\min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}) = P \langle Q \vec{\varphi}, \vec{\varphi} \rangle.$$

Taking into consideration (48), we get

$$\max_{\xi \in \Xi} \min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}) = P \max_{\|\vec{\varphi}\|=1} \langle Q \vec{\varphi}, \vec{\varphi} \rangle = P \nu^2,$$

where ν^2 is the greatest eigenvalue of the operator Q and $\langle \cdot, \cdot \rangle$ is the inner product in the space ℓ_2 . Making use (61), we can estimate the maximin value of the error

$$\max_{\xi \in \Xi} \min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}) \geq P \nu^2. \quad (64)$$

Upper bound. Consider the sequence of operators Q_N , determined by matrices (52), and the operator Q determined by the matrix with elements (63). Since the second condition from (60) is satisfied, then

$$N(Q - Q_N) = \sum_{p=N+1}^{\infty} (p+1) |a(p)|^2 \rightarrow 0,$$

as $N \rightarrow \infty$. Having in mind that

$$\|Q - Q_N\| \leq N(Q - Q_N),$$

we have

$$\lim_{N \rightarrow \infty} \|Q - Q_N\| = 0.$$

It means that the sequence of operators Q_N converge uniformly to the operator Q . That is why [8], [14]

$$\lim_{N \rightarrow \infty} \nu_N^2 = \nu^2,$$

where ν_N^2 is the greatest eigenvalue of the operator Q_N , and ν^2 is the greatest eigenvalue of the operator Q . Making use the statement of Theorem 1, we can write

$$\min_{\hat{A} \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}) = \quad (65)$$

$$= \lim_{N \rightarrow \infty} \min_{\hat{A}_N \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}_N) = \lim_{N \rightarrow \infty} P\nu_N^2 = P\nu^2.$$

Comparing (65) and (64), we get

$$\min_{\hat{A} \in \Lambda} \max_{\xi \in \Xi} \Delta(\xi, \hat{A}) = P\nu^2 \leq \max_{\xi \in \Xi} \min_{\hat{A} \in \Lambda} \Delta(\xi, \hat{A}),$$

where only equality is possible.

This completes the proof of the theorem.

Corollary 2

The optimal minimax linear estimate $\hat{A}\xi$ of the functional $A\xi$ is of the form

$$\hat{A}\xi = \sum_{j=0}^{\infty} a(j) \left[\sum_{u=-\infty}^{-1} \varphi(j-u)\varepsilon(u) \right],$$

where $\varepsilon(u)$ is a standard stationary sequence with orthogonal values, the sequence $\{\varphi(u) : u = 0, 1, \dots\}$ is uniquely determined by coordinates of the eigenvector of the operator Q that corresponds to the greatest eigenvalue ν^2 and condition $E \|\xi(j)\|^2 = P$.

Example 2

Consider the problem of optimal linear estimation of the functional

$$A\xi = \sum_{j=0}^{\infty} e^{-\lambda j} \xi(j),$$

where $\lambda > 0$, which depends on the unknown values of a stationary sequence $\xi(j)$, that satisfies the condition

$$E\xi(j) = 0, \quad E|\xi(j)|^2 \leq 1,$$

based on observations of the sequence at points $j = -1, -2, \dots$. Conditions (60) are satisfied. Elements of the matrix which determines the operator Q , determined by the equation (63), are of the form

$$Q(p, q) = \sum_{u=0}^{\infty} a(p+u)\overline{a(q+u)} = e^{-\lambda(p+q)}(1 - e^{-2\lambda})^{-1}.$$

Eigenvalues of the operator Q are determined by the system of equations

$$\mu\varphi(p) = \sum_{s=0}^{\infty} e^{-\lambda(p+s)}(1 - e^{-2\lambda})^{-1}\varphi(s), \quad p = 0, 1, \dots$$

It follows from this system of equations that $\varphi(p)$ are of the form

$$\varphi(p) = Ce^{-\lambda p}, \quad p = 0, 1, \dots$$

The constant C is determined by the condition

$$\sum_{p=0}^{\infty} |\varphi(p)|^2 = 1.$$

So we have

$$C = (1 - e^{-2\lambda})^{1/2}, \quad \varphi(p) = (1 - e^{-2\lambda})^{1/2}e^{-\lambda p}.$$

Substitution of these expressions to the system of equations gives us

$$\mu = (1 - e^{-2\lambda})^{-2}.$$

We can conclude that the least favourable in the class Ξ stationary sequence $\xi(j)$ is a moving average sequence of the form

$$\xi(j) = (1 - e^{-2\lambda})^{1/2} e^{-\lambda j} \sum_{u=-\infty}^j e^{\lambda u} \varepsilon(u),$$

where $\varepsilon(u)$ is a stationary sequence with orthogonal values.

The optimal linear minimax estimate $\hat{A}\xi$ of the functional $A\xi$ is as follows

$$\hat{A}\xi = (1 - e^{-2\lambda})^{1/2} \sum_{j=0}^{\infty} e^{-2\lambda j} \left[\sum_{u=-\infty}^{-1} e^{\lambda u} \varepsilon(u) \right].$$

The value of the mean-square error does not exceed

$$\Delta = (1 - e^{-2\lambda})^{-2}.$$

This value of the mean-square error gives the least favourable stationary sequence.

3.3. Conclusions

In this section we propose a method of solution of the mean square optimal linear estimation of the functionals $A_N\xi = \sum_{j=0}^N a(j)\xi(j)$ and $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ which depend on the unknown values of a stationary stochastic sequence $\xi(j)$ from the class Ξ of stationary stochastic sequences satisfying the conditions $E\xi(j) = 0$, $E|\xi(j)|^2 \leq P$. Estimates are based on results of observations of the sequence $\xi(j)$ at points of time $j = -1, -2, \dots$.

Inspired by the Ulf Grenander [15] approach to investigation the problem of optimal linear estimation of the functionals which depend on the unknown values of a stationary stochastic process we consider the problem as a two-person zero-sum game. It is show that this game has an equilibrium point. The maximum error gives a moving average stationary sequence which is least favourable in the given class of stationary sequences. The greatest value of the error and the least favourable sequence are determined by the greatest eigenvalue and the corresponding eigenvector of the operator determined by the coefficients $a(j)$ which determine the functional.

For the corresponding results for stationary stochastic processes with values in a Hilbert space see papers by Moklyachuk [35] – [39], [57].

4. Extrapolation problem for functionals of stationary sequences

In this section we deal with the problem of the mean-square optimal estimation of the linear functionals

$$A_N\xi = \sum_{j=0}^N a(j)\xi(j), \quad A\xi = \sum_{j=0}^{\infty} a(j)\xi(j),$$

which depend on the unknown values of a stationary stochastic sequence $\xi(j)$, $j \in \mathbb{Z}$, based on observations of the sequence $\xi(j)$ at points of time $j = -1, -2, \dots$.

The problem is investigated in the case of spectral certainty, where the spectral density of the stationary stochastic sequence $\xi(j)$ is exactly known. In this case the classical Hilbert space projection method of linear estimation of the functional is applied. Formulas are derived for calculation the value of the mean square error and the spectral characteristic of the mean-square optimal estimate of the linear functional. In the case of spectral uncertainty, where the spectral density of the stationary stochastic sequence is not exactly known, but a class of admissible spectral densities is given, the minimax-robust procedure to linear estimation of the functional is applied. Relations which determine the least favourable spectral densities and the minimax spectral characteristics are proposed for some special sets of admissible spectral densities.

4.1. The classical Hilbert space projection method of linear extrapolation

Let $\xi(j), j \in \mathbb{Z}, E\xi(j) = 0$, be a (wide sense) stationary stochastic sequence. We will consider values of $\xi(j), j \in \mathbb{Z}$, as elements of the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ of complex valued random variables with zero first moment, $E\xi = 0$, finite second moment, $E|\xi|^2 < \infty$, and the inner product $(\xi, \eta) = E\xi\bar{\eta}$. The correlation function $R(k) = (\xi(j+k), \xi(j)) = E\xi(j+k)\bar{\xi}(j)$ of the stationary stochastic sequence $\xi(j), j \in \mathbb{Z}$, admits the spectral representation

$$R(k) = \int_{-\pi}^{\pi} e^{ik\lambda} F(d\lambda),$$

where $F(d\lambda)$ is the spectral measure of the sequence. We will consider regular stationary stochastic sequences with absolutely continuous spectral measures and the correlation functions of the form

$$R(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda,$$

where $f(\lambda)$ is the spectral density function of the sequence $\xi(j)$ that satisfies the regularity condition

$$\int_{-\pi}^{\pi} \ln(f(\lambda)) d\lambda > -\infty. \tag{66}$$

This condition is necessary and sufficient in order that the error-free extrapolation of the unknown values of the sequence is impossible [12].

Suppose that coefficients $a(j)$, which determine the functional $A\xi$, satisfy conditions

$$\sum_{j=0}^{\infty} |a(j)| < \infty, \quad \sum_{j=0}^{\infty} (j+1) |a(j)|^2 < \infty. \tag{67}$$

In this case the functional $A\xi$ has the second moment and the Hilbert-Schmidt operator Q , determined in the previous section, has its Hilbert-Schmidt norm finite.

Let the sequence $\xi(j)$ admits the canonical representation as a moving average sequence

$$\xi(j) = \sum_{u=-\infty}^j d(j-u)\varepsilon(u), \tag{68}$$

where the sequence $d(u)$ satisfies condition

$$\sum_{u=0}^{\infty} |d(u)|^2 < \infty,$$

and where $\varepsilon(u)$ is a standard stationary white noise sequence

$$E\varepsilon(i)\bar{\varepsilon}(j) = \delta_{ij},$$

δ_{ij} is the Kronecker symbol.

In this case the spectral density $f(\lambda)$ of the stationary sequence $\xi(j)$ admits the canonical factorization

$$f(\lambda) = \varphi(\lambda)\overline{\varphi(\lambda)}, \quad \varphi(\lambda) = \sum_{k=0}^{\infty} d(k)e^{-ik\lambda}. \tag{69}$$

Denote by $L_2(f)$ the Hilbert space of functions $a(\lambda)$ such that

$$\int_{-\pi}^{\pi} a(\lambda) \overline{a(\lambda)} f(\lambda) d\lambda < \infty.$$

Denote by $L_2^-(f)$ the subspace of the space $L_2(f)$ generated by functions

$$\{e^{in\lambda}, n = -1, -2, \dots\}.$$

Every linear estimate $\hat{A}\xi$ of the functional $A\xi$ based on observations of the sequence $\xi(j)$ at points $j = -1, -2, \dots$ is of the form

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

where $Z_{\xi}(\Delta)$ is the orthogonal random measure of the sequence $\xi(j)$:

$$E \left(Z_{\xi}(\Delta_1) \overline{Z_{\xi}(\Delta_2)} \right) = \frac{1}{2\pi} \int_{\Delta_1 \cap \Delta_2} f(\lambda) d\lambda,$$

$h(e^{i\lambda})$ is the spectral characteristic of the estimate $\hat{A}\xi$ which belongs to the subspace $L_2^-(f)$. The mean-square error of the linear estimate $\hat{A}\xi$ of the functional $A\xi$ based on observations of the sequence $\xi(j)$ at points $j = -1, -2, \dots$, is calculated by the formula

$$\begin{aligned} \Delta(h, f) &= M \left| A\xi - \hat{A}\xi \right|^2 = \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |A(e^{i\lambda}) - h(e^{i\lambda})|^2 f(\lambda) d\lambda, \end{aligned}$$

where

$$A(e^{i\lambda}) = \sum_{j=0}^{\infty} a(j) e^{ij\lambda}.$$

If the stationary stochastic sequence $\xi(j)$ admits the canonical representation in the form of a moving average sequence (68), then the optimal estimate of the functional $A\xi$ is determined by the spectral characteristic $h(f) \in L_2^-(f)$ such that

$$\Delta(h(f), f) = \min_{h \in L_2^-(f)} \Delta(h, f) = \|Ad\|^2, \quad (70)$$

where

$$\|Ad\|^2 = \sum_{k=0}^{\infty} \|(Ad)_k\|^2, \quad (Ad)_k = \sum_{l=0}^{\infty} a(k+l)d(l).$$

Note, that $\|Ad\|^2 < \infty$ under the conditions (67). The spectral characteristic $h(f)$ of the optimal estimate is calculated by the formula

$$h(f) = A(e^{i\lambda}) - \varphi^{-1}(\lambda) r(e^{i\lambda}), \quad (71)$$

where

$$r(e^{i\lambda}) = \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda}. \quad (72)$$

For the functional $A_N\xi$ the mean-square error and the spectral characteristic of the optimal estimate of the functional is calculated by the formulas

$$\Delta_N(h(f), f) = \|A_N d\|^2, \quad (73)$$

$$h_N(f) = A_N(e^{i\lambda}) - \varphi^{-1}(\lambda) r_N(e^{i\lambda}), \quad (74)$$

where

$$A_N(e^{i\lambda}) = \sum_{j=0}^N a(j)e^{ij\lambda}, \quad r_N(e^{i\lambda}) = \sum_{k=0}^N (A_N d)_k e^{ik\lambda}, \tag{75}$$

$$\|A_N d\|^2 = \sum_{k=0}^N \|(A_N d)_k\|^2, \quad (A_N d)_k = \sum_{l=0}^{N-k} a(k+l)d(l), \quad k = 0, 1, \dots, N.$$

As a corollary, from the formula (73) one can find the following formula for calculation the mean-square error of optimal estimates $\hat{\xi}(j)$ of the unknown values $\xi(j)$

$$E \left| \xi(j) - \hat{\xi}(j) \right|^2 = \sum_{u=0}^j |d(u)|^2, \tag{76}$$

where $d(u)$ are determined from equations of factorization (69) of the density $f(\lambda)$.

Thus we came to conclusion that the following theorem holds true.

Theorem 10

If conditions (67) are satisfied and the density $f(\lambda)$ admits the canonical factorization (69), then the mean-square error of the optimal linear estimate of the functional $A\xi$ based on observations of the sequence $\xi(j)$ at points $j = -1, -2, \dots$, is calculated by formula (70) (by formula (73) for the functional $A_N\xi$). The spectral characteristic of the optimal linear estimate of the functional is calculated by formula (71) (by formula (74) for the functional $A_N\xi$).

4.2. Minimax-robust method of linear extrapolation of functionals

Formulas (70) – (76) can be applied to calculate the spectral characteristic and the mean-square error of the optimal linear estimate of the functional $A\xi$ only in the case where the spectral density $f(\lambda)$ of the stationary stochastic sequence $\xi(j)$ is exactly known. In the case where the spectral density $f(\lambda)$ is not exactly known, but, instead, a set D of admissible spectral densities is specified, the minimax approach to the problem of the optimal linear estimate of the functional which depends on the unknown values of the stationary sequence is reasonable. Under this approach one finds an estimate of the functional which is optimal and minimize the mean-square error for all spectral densities from a given class D simultaneously.

Definition 1

A spectral density $f^0(\lambda)$ is called the least favourable in the class D for the optimal linear extrapolation of the functional $A\xi$ if the following relation holds true

$$\Delta(h(f^0), f^0) = \max_{f \in D} \Delta(h(f), f) = \max_{f \in D} \min_{h \in L_2^-(f)} \Delta(h, f).$$

Taking into consideration relations (70) – (76), we can verify that the following statements hold true

Theorem 11

The spectral density $f^0(\lambda) \in D$ is the least favourable in the class D for the optimal linear extrapolation of the functional $A\xi$, if it admits the canonical factorization

$$f^0(\lambda) = \left(\sum_{k=0}^{\infty} d^0(k)e^{-ik\lambda} \right) \cdot \overline{\left(\sum_{k=0}^{\infty} d^0(k)e^{-ik\lambda} \right)}, \tag{77}$$

where $d^0 = \{d^0(k) : k = 0, 1, \dots\}$ is a solution to the constrained optimization problem

$$\|Ad\|^2 \rightarrow \max, \tag{78}$$

$$f(\lambda) = \left(\sum_{k=0}^{\infty} d(k)e^{-ik\lambda} \right) \cdot \overline{\left(\sum_{k=0}^{\infty} d(k)e^{-ik\lambda} \right)} \in D.$$

The sequence $\xi(j)$ in this case admits the canonical one-sided moving average representation

$$\xi(j) = \sum_{u=-\infty}^j d^0(j-u)\varepsilon(u). \tag{79}$$

Theorem 12

The spectral density $f^0(\lambda) \in D$ is the least favourable in the class D for the optimal linear extrapolation of the functional $A_N \xi$, if it admits the canonical factorization

$$f^0(\lambda) = \left(\sum_{k=0}^N d^0(k)e^{-ik\lambda} \right) \cdot \overline{\left(\sum_{k=0}^N d^0(k)e^{-ik\lambda} \right)}, \tag{80}$$

where $d^0 = \{d^0(k) : k = 0, 1, \dots, N\}$ is a solution to the constrained optimization problem

$$\|A_N d\|^2 \rightarrow \max, \tag{81}$$

$$f(\lambda) = \left(\sum_{k=0}^N d(k)e^{-ik\lambda} \right) \cdot \overline{\left(\sum_{k=0}^N d(k)e^{-ik\lambda} \right)} \in D.$$

The sequence $\xi(j)$ in this case admits the canonical one-sided moving average representation of order $N + 1$:

$$\xi(j) = \sum_{u=j-N}^j d^0(j-u)\varepsilon(u). \tag{82}$$

Definition 2

The spectral characteristic $h^0(e^{i\lambda})$ of the optimal linear extrapolation of the functional $A\xi$ is called minimax (robust) if the following conditions hold true

$$h^0(e^{i\lambda}) \in H_D = \bigcap_{f \in D} L_2^-(f), \quad \min_{h \in H_D} \max_{f \in D} \Delta(h, f) = \max_{f \in D} \Delta(h^0, f).$$

The least favourable spectral density $f^0(\lambda) \in D$ and the minimax (robust) spectral characteristic $h^0(e^{i\lambda}) \in H_D$ form a saddle point of the function $\Delta(h, f)$ on the set $H_D \times D$. The saddle point inequalities

$$\Delta(h, f^0) \geq \Delta(h^0, f^0) \geq \Delta(h^0, f) \quad \forall f \in D \quad \forall h \in H_D$$

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

$$\Delta(h(f^0), f^0) = \max_{f \in D} \Delta(h(f^0), f).$$

If we have found a solution $f^0(\lambda)$ to this problem, then the minimax(robust) spectral characteristics can be calculated by formulas (71), (74) if the condition $h(f^0) \in H_D$ holds true.

The spectral density $f^0(\lambda)$ is a solution to the following constrained optimization problem

$$\Delta(f) = -\Delta(h(f^0), f) \rightarrow \inf, \quad f(\lambda) \in D, \tag{83}$$

$$\Delta(h(f^0), f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|r(e^{i\lambda})|^2}{|\varphi^0(\lambda)|^2} f(\lambda) d\lambda,$$

where $r(e^{i\lambda})$ is calculated by formulas (72), (75) with $f(\lambda) = f^0(\lambda)$.

The constrained optimization problem (83) is equivalent to the following unconstrained optimization problem [58],[70]

$$\Delta_D(f) = -\Delta(h(f^0), f) + \delta(f|D) \rightarrow \inf, \tag{84}$$

where $\delta(f|D)$ is the indicator function of the set D . Solution of the problem (84) is determined by the condition $0 \in \partial\Delta_D(f^0)$, where $\partial\Delta_D(f^0)$ is the subdifferential on the convex functional $\Delta_D(f)$ at the point f^0 . With the help of conditions (83), (84) we can find the least favourable spectral densities for concrete classes of spectral densities.

Note, that the form of the functional $\Delta(h(f_0), f)$ is convenient for application the Lagrange method of indefinite multipliers for finding solution to the problem (83). Making use the method of Lagrange multipliers and the form of subdifferentials of the indicator functions of certain classes of spectral densities we describe relations that determine least favourable spectral densities in some special classes of spectral densities [57], [64].

4.3. Least favourable spectral densities in the class D_0

Consider the problem of the optimal estimation of the functionals $A\xi = \sum_{j=1}^{\infty} a(j)\xi(j)$ and $A_N\xi = \sum_{j=1}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j)$ at points of time $j = -1, -2, \dots$ in the case where the spectral density $f(\lambda)$ is from the class D_0 of spectral densities which are characterized by restrictions on the first moment of the density

$$D_0 = \left\{ f(\lambda) : \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda)d\lambda \leq P_0 \right\},$$

where $P_0, P_0 > 0$, is a given number. This class of spectral densities describes stationary sequences with restriction on the dispersion $E|\xi(j)|^2 \leq P_0$.

We can apply the method of Lagrange multipliers to find solution to the optimization problem (83). We get the following relations that determine the least favourable spectral density $f^0 \in D_0$

$$|r(e^{i\lambda})|^2 = \alpha^2 |\varphi^0(\lambda)|^2, \tag{85}$$

where α^2 is the Lagrange multiplier.

This relation can be rewritten in the following way

$$\left(\sum_{k=0}^{\infty} (Ad^0)_k e^{ik\lambda} \right) \cdot \overline{\left(\sum_{k=0}^{\infty} (Ad^0)_k e^{ik\lambda} \right)} = \alpha^2 \left(\sum_{k=0}^{\infty} d^0(k) e^{-ik\lambda} \right) \cdot \overline{\left(\sum_{k=0}^{\infty} d^0(k) e^{-ik\lambda} \right)}. \tag{86}$$

The unknown α^2 and $d^0 = \{d^0(k) : k = 0, 1, \dots\}$ are determined with the help of equations of the canonical factorization (77) of the density $f^0(\lambda)$, solution of the constrained optimization problem (78) and restrictions imposed on densities from the class of admissible spectral densities D_0 .

For all solutions $d = \{d(k) : k = 0, 1, \dots\}$ of the equation

$$Ad = \alpha \bar{d}, \alpha \in \mathbb{C}, \tag{87}$$

which satisfy condition

$$\|d\|^2 = \sum_{k=0}^{\infty} |d(k)|^2 = P_0, \tag{88}$$

the following equality holds true

$$f^0(\lambda) = \left| \sum_{k=0}^{\infty} d(k) e^{-ik\lambda} \right|^2 = \left| c \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda} \right|^2. \tag{89}$$

Solution $d^0 = \{d^0(k) : k = 0, 1, \dots\}$ of the equation (87) which satisfies condition (88) and gives the maximum value $\|Ad^0\|^2 = \nu_0 P_0$ of the quantity $\|Ad\|^2$ determines the least favourable spectral density

$$f^0(\lambda) = \left| \sum_{k=0}^{\infty} d^0(k) e^{-ik\lambda} \right|^2, \quad (90)$$

which is the spectral density of the one-sided moving-average sequence

$$\xi(j) = \sum_{u=-\infty}^j d^0(j-u) \varepsilon(u). \quad (91)$$

Thus the following statement fulfilled.

Theorem 13

The least favourable in the class D_0 for the optimal extrapolation of the functional $A\xi$ is the spectral density (90) of the one-sided moving-average sequence (91) which is determined by solution $d^0 = \{d^0(k) : k = 0, 1, \dots\}$ of the equation (87) which satisfies condition (88) and gives the maximum value $\|Ad^0\|^2 = \nu^2 P_0$ of the quantity $\|Ad\|^2$. The minimax spectral characteristic of the optimal estimate of the functional $A\xi$ is calculated by the formula (71).

For the functional $A_N \xi = \sum_{j=1}^N a(j) \xi(j)$ the corresponding relation has the following form

$$|r_N(e^{i\lambda})|^2 = \alpha^2 |\varphi^0(\lambda)|^2, \quad (92)$$

and the equality

$$|r_N(e^{i\lambda})|^2 = \left| \sum_{k=0}^N (A_N d)_k e^{ik\lambda} \right|^2 = \left| \sum_{k=0}^N (\tilde{A}_N d)_k e^{-ik\lambda} \right|^2, \quad (93)$$

holds true, where

$$(\tilde{A}_N d)_k = \sum_{u=0}^k a(N-k+u) d(u), \quad k = 0, 1, \dots, N.$$

For all solutions $d = \{d(k) : k = 0, 1, \dots, N\}$ to equations

$$A_N d = \alpha \bar{d}, \quad \alpha \in \mathbb{C}, \quad (94)$$

$$\tilde{A}_N d = \beta d, \quad \beta \in \mathbb{C}, \quad (95)$$

which satisfy condition

$$\|d\|^2 = \sum_{k=0}^N |d(k)|^2 = P_0, \quad (96)$$

the following equality holds true

$$f^0(\lambda) = \left| \sum_{k=0}^N d(k) e^{-ik\lambda} \right|^2 = \left| c \sum_{k=0}^N (Ad)_k e^{ik\lambda} \right|^2. \quad (97)$$

Solutions $d^0 = \{d^0(k) : k = 0, 1, \dots, N\}$ of the equations (94), (95) which satisfy condition (96) and gives the maximum values $\|A_N d^0\|^2 = \|\tilde{A}_N d^0\|^2 = \nu_N^2 P_0$ of the quantity $\|A_N d\|^2$ determines the least favourable spectral density

$$f^0(\lambda) = \left| \sum_{k=0}^N d^0(k) e^{-ik\lambda} \right|^2, \quad (98)$$

which is the spectral density of the one-sided moving-average sequence of order N

$$\xi(j) = \sum_{u=j-N}^j d^0(j-u)\varepsilon(u). \tag{99}$$

Thus the following statement holds true.

Theorem 14

The least favourable in the class D_0 for the optimal extrapolation of the functional $A_N\xi$ is the spectral density (98) of the one-sided moving-average sequence (99) which is determined by solutions $d^0 = \{d^0(k) : k = 0, 1, \dots, N\}$ of the equations (94), (95) which satisfy condition (96) and gives the maximum value $\|A_N d^0\|^2 = \|\tilde{A}_N d^0\|^2 = \nu_N^2 P_0$ of the quantity $\|A_N d\|^2$. The minimax spectral characteristic of the optimal estimate of the functional $A_N\xi$ is calculated by the formula (74).

Corollary 3

The spectral density (98) of the one-sided moving-average sequence (99), where the sequence $d^0 = \{d^0(k) : k = 0, 1, \dots, N\}$ satisfy condition (96), is the least favourable in the class D_0 for the optimal extrapolation of the functional $a(N)\xi(N)$.

4.4. Least favourable spectral densities in the class D_M

Consider the problem of the optimal estimation of the functionals $A\xi = \sum_{j=1}^{\infty} a(j)\xi(j)$ and $A_N\xi = \sum_{j=1}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j)$ at points of time $j = -1, -2, \dots$ in the case where the spectral density $f(\lambda)$ is from the class D_M of spectral densities which are characterized by restrictions on the moments of the density

$$D_M = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) \cos(m\lambda) d\lambda = r(m), m = 0, 1, \dots, M \right. \right\},$$

where $r(0) = P_0$ and $r(0), r(1), \dots, r(M)$ is a strictly positive sequence. The moment problem in this case has not uniquely determined solutions and the class D_M contains an infinite number of densities [25].

We can apply the method of Lagrange multipliers to find solution to the optimization problem (83) in the case $D = D_M$. We get the following relations that determine the least favourable spectral density $f^0 \in D_M$

$$|r(e^{i\lambda})|^2 (f^0(\lambda))^{-1} = \sum_{m=0}^M \alpha_m \cos(m\lambda) = \left| \sum_{m=0}^M p(m) e^{im\lambda} \right|^2, \tag{100}$$

where $\alpha_m, m = 0, 1, \dots, M$ are the Lagrange multipliers.

It follows from this relation that the least favourable spectral density $f^0 \in D_M$ is of the form

$$f_0(\lambda) = \frac{\left| \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda} \right|^2}{\left| \sum_{m=0}^M p(m) e^{-im\lambda} \right|^2}. \tag{101}$$

The unknown $\{p(m) : m = 0, 1, \dots, M\}$ and $d = \{d(k) : k = 0, 1, \dots\}$ are determined by equations of the canonical factorization (77) of the density $f^0(\lambda)$, solution of the constrained optimization problem (78) and restrictions imposed on densities from the class of admissible spectral densities D_M .

Denote by $\nu_M P_0$ the maximum value $\|Ad^0\|^2$ of the quantity $\|Ad\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and are determined by equations of the canonical factorization (77) of the density $f^0(\lambda)$ and restrictions imposed on densities from the class of admissible spectral densities D_M .

Denote by $\nu_M^+ P_0$ the maximum value $\|Ad^0\|^2$ of the quantity $\|Ad\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and are determined by equations of canonical factorization (77) of the density (101) and restrictions imposed on densities from the class of admissible spectral densities D_M .

The following statement is true.

Theorem 15

If there exists a sequence $d^0 = \{d^0(k) : k = 0, 1, \dots\}$ which satisfies condition (88) and such that $\nu_M P_0 = \nu_M^+ P_0 = \|Ad^0\|^2$, then the least favourable in the class D_M for the optimal extrapolation of the functional $A\xi$ is the spectral density (90) of the one-sided moving-average sequence (91). If $\nu_M < \nu_M^+$, then the least favourable in the class D_M for the optimal extrapolation of the functional $A\xi$ is the spectral density (101) which admits the canonical factorization (77). The unknown $\{p(m) : m = 0, 1, \dots, M\}$ and $d = \{d(k) : k = 0, 1, \dots\}$ are determined by equations of canonical factorization (77) of the density $f^0(\lambda)$, solution of the constrained optimization problem (78) and restrictions imposed on densities from the class of admissible spectral densities D_M . The minimax spectral characteristic of the optimal estimate of the functional $A\xi$ is calculated by the formula (71).

Consider now the problem of the optimal estimation of the functional $A_N\xi = \sum_{j=1}^N a(j)\xi(j)$. It follows from the relation (100) that in this case the least favourable spectral density $f^0(\lambda) \in D_M$ is of the form

$$f_0(\lambda) = \frac{\left| \sum_{k=0}^N (A_N d)_k e^{ik\lambda} \right|^2}{\left| \sum_{m=0}^M p(m) e^{-im\lambda} \right|^2}. \tag{102}$$

These densities are spectral densities of the autoregressive-moving-average $ARMA(M, N)$ sequences

$$\sum_{m=0}^M p(m)\xi(n - m) = \sum_{k=0}^N (A_N d)_k \varepsilon(n - k). \tag{103}$$

The unknown $\{p(m) : m = 0, 1, \dots, M\}$ and $d = \{d(k) : k = 0, 1, \dots, N\}$ are determined by equations of canonical factorization (80) of the density $f^0(\lambda)$, solution of the constrained optimization problem (81) and restrictions imposed on densities from the class of admissible spectral densities D_M .

Denote by $\nu_{MN} P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots, N\}$ are determined by equations (94), (95), condition (96) and equations of canonical factorization (80) of the density $f^0(\lambda) \in D_M$.

Denote by $\nu_{MN}^+ P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots, N\}$ satisfies condition (96), equations of canonical factorization (77) of the density (102) and restrictions imposed on densities from the class of admissible spectral densities $f^0(\lambda) \in D_M$.

Thus the following statement holds true.

Theorem 16

If there exists a solution $d^0 = \{d^0(k) : k = 0, 1, \dots, N\}$ of the equation (94), or the equation (95), which satisfy condition (96) and such that $\nu_{MN} P_0 = \nu_{MN}^+ P_0 = \|A_N d^0\|^2$, then the least favourable in the class D_M for the optimal extrapolation of the functional $A_N\xi$ is the spectral density (98) of the one-sided moving-average sequence (99). If $\nu_{MN} < \nu_{MN}^+$, then the least favourable in the class D_M for the optimal extrapolation of the functional $A_N\xi$ is the spectral density (102) of the autoregressive-moving-average $ARMA(M, N)$ sequences (103). The unknown $\{p(m) : m = 0, 1, \dots, M\}$ and $d = \{d(k) : k = 0, 1, \dots\}$ are determined by equations of canonical factorization (77) of the density $f^0(\lambda)$, solution of the constrained optimization problem (78) and restrictions imposed on densities from the class of admissible spectral densities D_M . The minimax spectral characteristic of the optimal estimate of the functional $A_N\xi$ is calculated by the formula (74).

4.5. Least favourable spectral densities in the class D_v^u

Consider the problem of minimax estimation of the functionals $A\xi$ and $A_N\xi$ which depend on the unknown values of a stationary stochastic sequence $\xi(j)$ for the sets of spectral densities that describe the “strip” model of stationary

stochastic sequences

$$D_v^u = \left\{ f(\lambda) \left| v(\lambda) \leq f(\lambda) \leq u(\lambda), \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) d\lambda = P_0 \right. \right\},$$

where $v(\lambda), u(\lambda)$ are given bounded spectral densities.

From the condition $0 \in \partial \Delta_D(f_0)$ for $D = D_v^u$ we find the following equation which determines the least favourable spectral density for the optimal estimation of the functional $A\xi$

$$\left| \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda} \right|^2 = (\psi_1(\lambda) + \psi_2(\lambda) + \alpha_0) |\varphi^0(\lambda)|^2, \tag{104}$$

where $\psi_1(\lambda) \geq 0$ and $\psi_1(\lambda) = 0$ if $f_0(\lambda) \geq v(\lambda)$; $\psi_2(\lambda) \leq 0$ and $\psi_2(\lambda) = 0$ if $f_0(\lambda) \leq u(\lambda)$.

From this equation we find that the least favourable spectral density for the optimal estimation of the functional $A\xi$ is of the form

$$f_0(\lambda) = \max \left\{ v(\lambda), \min \left\{ u(\lambda), c \left| \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda} \right|^2 \right\} \right\}. \tag{105}$$

Denote by $\nu_u P_0$ the maximum value $\|Ad^0\|^2$ of the quantity $\|Ad\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ are solutions of the equation (87) which satisfy condition (88), the inequality

$$v(\lambda) \leq \left| \sum_{k=0}^{\infty} d(k) e^{-ik\lambda} \right|^2 \leq u(\lambda), \tag{106}$$

and determine the canonical factorization (77) of the density $f^0(\lambda) \in D_v^u$.

Denote by $\nu_u^+ P_0$ the maximum value $\|Ad^0\|^2$ of the quantity $\|Ad\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and determine the canonical factorization (77) of the density (105) from the class of admissible spectral densities D_v^u .

Thus the following theorem holds true.

Theorem 17

If there exists a solution of the equation (87) which satisfy condition (88) and such that $\nu_u P_0 = \nu_u^+ P_0 = \|Ad^0\|^2$, then the spectral density (90) of the one-sided moving-average sequence (91) is the least favourable in the set D_v^u for the optimal extrapolation of the functional $A\xi$. If $\nu_u < \nu_u^+$, then the least favourable in the class D_v^u for the optimal extrapolation of the functional $A\xi$ is the spectral density (105) which admits the canonical factorization (90). The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by the optimisation problem (78) and restrictions imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A\xi$ is calculated by the formula (71).

For the functional $A_N \xi$ the least favourable spectral density for the optimal estimation of the functional is of the form

$$f_0(\lambda) = \max \left\{ v(\lambda), \min \left\{ u(\lambda), c \left| \sum_{k=0}^N (A_N d)_k e^{ik\lambda} \right|^2 \right\} \right\}. \tag{107}$$

Denote by $\nu_{uN} P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots, N\}$ are solutions of the equations (94), (95), which satisfy condition (96), the inequality

$$v(\lambda) \leq \left| \sum_{k=0}^N d(k) e^{-ik\lambda} \right|^2 \leq u(\lambda), \tag{108}$$

and determine the canonical factorization (80) of the density $f^0(\lambda) \in D_v^u$.

Denote by $\nu_{uN}^+ P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and determine the canonical factorization (77) of the density (107) from the class of admissible spectral densities D_v^u .

The following theorem holds true.

Theorem 18

If there exists a solution of the equation (94), or equation (95), which satisfy condition (96) and such that $\nu_{uN} P_0 = \nu_{uN}^+ P_0 = \|A_N d^0\|^2$, then the spectral density (98) of the one-sided moving-average sequence (99) is the least favourable in the set D_v^u for the optimal extrapolation of the functional $A_N \xi$. If $\nu_{uN} < \nu_{uN}^+$, then the least favourable in the class D_v^u for the optimal extrapolation of the functional $A_N \xi$ is the spectral density (107) which admits the canonical factorization (90). The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by the optimisation problem (78) and restrictions imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A_N \xi$ is calculated by the formula (74).

4.6. Least favourable spectral densities in the class D_ε

Consider the problem of minimax estimation of the functionals $A\xi$ and $A_N \xi$ which depend on the unknown values of a stationary stochastic sequence $\xi(j)$ for the set of spectral densities that describes the “ ε -contamination” model of stationary stochastic sequences

$$D_\varepsilon = \left\{ f(\lambda) \mid f(\lambda) = (1 - \varepsilon)w(\lambda) + \varepsilon u(\lambda), \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) d\lambda = P_0 \right\},$$

where $w(\lambda)$ is a known spectral density, and $u(\lambda)$ is an unknown spectral density. From the condition $0 \in \partial \Delta_D(f^0)$ for the functional $A\xi$ we find the following equations which determine the least favourable spectral densities for the optimal estimation of the functional $A\xi$ for the given set of admissible spectral densities

$$\left| \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda} \right|^2 = (\psi_1(\lambda) + \alpha_0^{-1}) |\varphi^0(\lambda)|^2, \tag{109}$$

where $\psi_1(\lambda) \geq 0$ and $\psi_1(\lambda) = 0$ if $f_0(\lambda) \geq (1 - \varepsilon)w(\lambda)$.

From this equation we find that the least favourable spectral density for the optimal estimation of the functional $A\xi$ is of the form

$$f_0(\lambda) = \max \left\{ (1 - \varepsilon)w(\lambda), \alpha_0 \left| \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda} \right|^2 \right\}. \tag{110}$$

Denote by $\nu_\varepsilon P_0$ the maximum value $\|Ad^0\|^2$ of the quantity $\|Ad\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ are solutions of the equation (87) which satisfy condition (88), the inequality

$$\left| \sum_{k=0}^{\infty} d(k) e^{-ik\lambda} \right|^2 \geq (1 - \varepsilon)w(\lambda), \tag{111}$$

and determine the canonical factorization (77) of the density $f^0(\lambda) \in D_\varepsilon$.

Denote by $\nu_\varepsilon^+ P_0$ the maximum value $\|Ad^0\|^2$ of the quantity $\|Ad\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and determine the canonical factorization (77) of the density (110) from the class of admissible spectral densities D_ε .

The following theorem holds true.

Theorem 19

If there exists a solution of the equation (87) which satisfy condition (88) and such that $\nu_\varepsilon P_0 = \nu_\varepsilon^+ P_0 = \|Ad^0\|^2$, then the spectral density (90) of the one-sided moving-average sequence (91) is the least favourable in the set D_ε for

the optimal extrapolation of the functional $A\xi$. If $\nu_\varepsilon < \nu_\varepsilon^+$, then the least favourable in the class D_ε for the optimal extrapolation of the functional $A\xi$ is the spectral density (110) which admits the canonical factorization (90). The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by the optimisation problem (78) and restrictions imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A\xi$ is calculated by the formula (71).

For the functional $A_N\xi$ the least favourable spectral density for the optimal estimation of the functional is of the form

$$f_0(\lambda) = \max \left\{ (1 - \varepsilon)w(\lambda), \alpha_0 \left| \sum_{k=0}^N (A_N d)_k e^{ik\lambda} \right|^2 \right\}. \tag{112}$$

Denote by $\nu_\varepsilon^N P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots, N\}$ are solutions of the equations (94), (95), which satisfy condition (96), the inequality

$$\left| \sum_{k=0}^N d(k) e^{-ik\lambda} \right|^2 \geq (1 - \varepsilon)w(\lambda), \tag{113}$$

and determine the canonical factorization (80) of the density $f^0(\lambda) \in D_\varepsilon$.

Denote by $\nu_\varepsilon^{N+} P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and determine the canonical factorization (77) of the density (107) from the class of admissible spectral densities D_ε .

The following theorem holds true.

Theorem 20

If there exists a solution of the equation (94), or the equation, (95), which satisfy condition (96) and such that $\nu_\varepsilon^N P_0 = \nu_\varepsilon^{N+} P_0 = \|A_N d^0\|^2$, then the spectral density (98) of the one-sided moving-average sequence (99) is the least favourable in the set D_ε for the optimal extrapolation of the functional $A_N\xi$. If $\nu_\varepsilon^N < \nu_\varepsilon^{N+}$, then the least favourable in the class D_ε for the optimal extrapolation of the functional $A_N\xi$ is the spectral density (112) which admits the canonical factorization (90). The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by the optimisation problem (78) and restrictions imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A_N\xi$ is calculated by the formula (74).

4.7. Least favourable spectral densities in the class $D_{1\varepsilon}$

Consider the problem of minimax estimation of the functionals $A\xi$ and $A_N\xi$ which depend on the unknown values of a stationary stochastic sequence $\xi(j)$ for the set of spectral densities that describes the model of “ ε -neighborhood in the space L_1 of a stationary stochastic sequence

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

where ε is a given numbers, $v(\lambda)$ is a given spectral density.

From the condition $0 \in \partial\Delta_D(f^0)$ we find that the least favourable spectral density for the optimal estimation of the functional $A\xi$ is of the form

$$f_0(\lambda) = \max \left\{ v(\lambda), c \left| \sum_{k=0}^{\infty} (A d)_k e^{ik\lambda} \right|^2 \right\}. \tag{114}$$

Denote by $\nu_{1\varepsilon} P_0$ the maximum value $\|A d^0\|^2$ of the quantity $\|A d\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ are solutions of the equation (87) which satisfy condition (88), the inequality

$$\left| \sum_{k=0}^{\infty} d(k) e^{-ik\lambda} \right|^2 \geq v(\lambda), \tag{115}$$

and determine the canonical factorization (77) of the density $f^0(\lambda) \in D_{1\varepsilon}$.

Denote by $\nu_{1\varepsilon}^+ P_0$ the maximum value $\|Ad^0\|^2$ of the quantity $\|Ad\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and determine the canonical factorization (77) of the density (114) from the class of admissible spectral densities $D_{1\varepsilon}$.

The following theorem holds true.

Theorem 21

If there exists a solution of the equation (87) which satisfy condition (88) and such that $\nu_{1\varepsilon} P_0 = \nu_{1\varepsilon}^+ P_0 = \|Ad^0\|^2$, then the spectral density (90) of the one-sided moving-average sequence (91) is the least favourable in the set $D_{1\varepsilon}$ for the optimal extrapolation of the functional $A\xi$. If $\nu_{1\varepsilon} < \nu_{1\varepsilon}^+$, then the least favourable in the class $D_{1\varepsilon}$ for the optimal extrapolation of the functional $A\xi$ is the spectral density (114) which admits the canonical factorization (90). The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by the optimisation problem (78) and restrictions imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A\xi$ is calculated by the formula (71).

For the functional $A_N\xi$ the least favourable spectral density for the optimal estimation of the functional is of the form

$$f_0(\lambda) = \max \left\{ v(\lambda), c \left| \sum_{k=0}^N (A_N d)_k e^{ik\lambda} \right|^2 \right\}. \quad (116)$$

Denote by $\nu_{1\varepsilon}^N P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots, N\}$ are solutions of the equations (94), (95), which satisfy condition (96), the inequality

$$\left| \sum_{k=0}^N d(k) e^{-ik\lambda} \right|^2 \geq v(\lambda), \quad (117)$$

and determine the canonical factorization (80) of the density $f^0(\lambda) \in D_{1\varepsilon}$.

Denote by $\nu_{1\varepsilon}^{N+} P_0$ the maximum value $\|A_N d^0\|^2$ of the quantity $\|A_N d\|^2$, where $d = \{d(k) : k = 0, 1, \dots\}$ satisfies condition (88) and determine the canonical factorization (77) of the density (116) from the class of admissible spectral densities $D_{1\varepsilon}$.

The following theorem holds true.

Theorem 22

If there exists a solution of the equation (94), or equation (95), which satisfy condition (96) and such that $\nu_{1\varepsilon}^N P_0 = \nu_{1\varepsilon}^{N+} P_0 = \|A_N d^0\|^2$, then the spectral density (98) of the one-sided moving-average sequence (99) is the least favourable in the set $D_{1\varepsilon}$ for the optimal extrapolation of the functional $A_N\xi$. If $\nu_{1\varepsilon}^N < \nu_{1\varepsilon}^{N+}$, then the least favourable in the class $D_{1\varepsilon}$ for the optimal extrapolation of the functional $A_N\xi$ is the spectral density (116) which admits the canonical factorization (90). The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by the optimisation problem (78) and restrictions imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A_N\xi$ is calculated by the formula (74).

4.8. Least favourable spectral densities in the class $D_{2\varepsilon}$

Consider the problem of minimax estimation of the functionals $A\xi$ and $A_N\xi$ which depend on the unknown values of a stationary stochastic sequence $\xi(j)$ for the set of spectral densities that describes the model of “ ε -neighborhood in the space L_2 of a stationary stochastic sequence

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

where ε is a given numbers, $v(\lambda)$ is a given spectral density.

From the condition $0 \in \partial \Delta_D(f^0)$ we find that the least favourable spectral density for the optimal estimation of the functional $A\xi$ is of the form

$$f_0(\lambda) = \frac{v(\lambda)}{2} + \left(\frac{(v(\lambda))^2}{4} + \left| \sum_{k=0}^{\infty} (Ad)_k e^{ik\lambda} \right|^2 \right)^{1/2}. \tag{118}$$

The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by equations of the canonical factorization (90) of the density (118), optimisation problem (78) and restriction

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\lambda) - v(\lambda)|^2 d\lambda = \varepsilon. \tag{119}$$

The following theorem holds true.

Theorem 23

The least favourable in the set $D_{2\varepsilon}$ spectral density for the optimal extrapolation of the functional $A\xi$ is determined by equation (118). The sequence $d = \{d(k) : k = 0, 1, \dots\}$ is determined by equations of the canonical factorization (90) of the density (118), optimisation problem (78) and restriction (119) imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A\xi$ is calculated by the formula (71).

For the functional $A_N\xi$ the least favourable spectral density for the optimal estimation of the functional is of the form

$$f_0(\lambda) = \frac{v(\lambda)}{2} + \left(\frac{(v(\lambda))^2}{4} + \left| \sum_{k=0}^N (A_N d)_k e^{ik\lambda} \right|^2 \right)^{1/2}. \tag{120}$$

The sequence $d = \{d(k) : k = 0, 1, \dots, N\}$ is determined by equations of the canonical factorization (90) of the density (120), optimisation problem (78) and restriction (119).

Theorem 24

The least favourable in the set $D_{2\varepsilon}$ spectral density for the optimal extrapolation of the functional $A_N\xi$ is determined by equation (120). The sequence $d = \{d(k) : k = 0, 1, \dots, N\}$ is determined by equations of the canonical factorization (90) of the density (120), optimisation problem (78) and restriction (119) imposed on densities by the given set of admissible spectral densities. The minimax spectral characteristic of the optimal estimate of the functional $A\xi$ is calculated by the formula (74).

4.9. Conclusions

In this section we propose methods of solution of the problem of the mean-square optimal linear estimation of the functionals $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ and $A_N\xi = \sum_{j=0}^N a(j)\xi(j)$ which depend on the unknown values of the stationary stochastic sequence $\xi(j)$. Estimates are based on observations of the sequence $\xi(j)$ at points of time $j = -1, -2, \dots$. The problem is investigated in the case of spectral certainty, where the spectral densities of the stationary stochastic sequence $\xi(j)$ is exactly known. In this case the classical Hilbert space projection method of linear estimation is applied. Formulas are derived for calculation the value of the mean square errors and the spectral characteristics of the mean-square optimal estimates of the linear functionals. In the case of spectral uncertainty, where the spectral density of the stationary stochastic sequence is not exactly known, but a class of admissible spectral densities is given the minimax-robust procedure to linear estimation of the functionals is applied. Relations which determine the least favourable spectral densities and the minimax spectral characteristics are proposed for some given sets of admissible spectral densities.

The minimax-robust approach to the problem of one step ahead optimal prediction of the stationary stochastic sequences as well as estimation of one missed value of the sequences based on convex optimization methods was

initiated in papers by Franke [9, 10], Franke and Poor [11]. See also papers by Hosoya [19], Taniguchi [75], and survey paper by Kassam and Poor [23].

For the relative results on the mean-square optimal linear extrapolation of linear functionals for stationary stochastic sequences and processes see papers by Moklyachuk [39] – [45], [57], book by Moklyachuk and Masyutka [64].

5. Extrapolation problem for stationary sequences from observations with noise

In this section we deal with the problem of the mean-square optimal estimation of the linear functionals $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ and $A_N\xi = \sum_{j=0}^N a(j)\xi(j)$ which depend on the unknown values of a stationary stochastic sequence $\xi(j), j \in \mathbb{Z}$, from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$, where $\eta(j)$ is an uncorrelated with $\xi(j)$ stationary stochastic sequence. The problem is investigated in the case of spectral certainty, where the spectral densities of the stationary stochastic sequences $\xi(j)$ and $\eta(j)$ are exactly known. In this case the classical Hilbert space projection method of linear estimation is applied. Formulas are derived for calculation the value of the mean square errors and the spectral characteristics of the mean-square optimal estimates of the linear functionals. In the case of spectral uncertainty, where the spectral densities of the stationary stochastic sequences are not exactly known, but a class of admissible spectral densities is given the minimax-robust procedure to linear estimation of the functional is applied. Relations which determine the least favourable spectral densities and the minimax spectral characteristics are proposed for some special sets of admissible spectral densities.

5.1. The classical Hilbert space projection method of linear extrapolation

Let $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, be (wide sense) stationary stochastic sequences with zero mathematical expectations $E\xi(j) = 0, E\eta(j) = 0$. The correlation functions $R_\xi(k) = E\xi(j+k)\overline{\xi(j)}$ and $R_\eta(k) = E\eta(j+k)\overline{\eta(j)}$ of stationary stochastic sequences $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, admit the spectral representations

$$R_\xi(k) = \int_{-\pi}^{\pi} e^{ik\lambda} F(d\lambda), \quad R_\eta(k) = \int_{-\pi}^{\pi} e^{ik\lambda} G(d\lambda),$$

where $F(d\lambda)$ and $G(d\lambda)$ are spectral measures of the sequences. We will consider stationary stochastic sequences with absolutely continuous spectral measures $F(d\lambda)$ and $G(d\lambda)$ and the correlation functions of the form

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

where $f(\lambda)$ and $g(\lambda)$ are the spectral density functions of the sequences $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, correspondingly.

We will suppose that the spectral density functions $f(\lambda)$ and $g(\lambda)$ satisfy the minimality condition

$$\int_{-\pi}^{\pi} \frac{1}{f(\lambda) + g(\lambda)} d\lambda < \infty. \quad (121)$$

Under this condition the error-free extrapolation of the unknown values of the sequence $\xi(j) + \eta(j)$ is impossible.

The stationary stochastic sequences $\xi(j)$ and $\eta(j)$ admit the spectral representations

$$\xi(j) = \int_{-\pi}^{\pi} e^{ij\lambda} dZ_\xi(\lambda), \quad \eta(j) = \int_{-\pi}^{\pi} e^{ij\lambda} dZ_\eta(\lambda),$$

where $Z_\xi(d\lambda)$ and $Z_\eta(d\lambda)$ are orthogonal stochastic measures of the sequences $\xi(j)$ and $\eta(j)$ such that

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

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

Consider 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 stochastic sequence $\xi(j), j = 0, 1, \dots$, based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$.

We will suppose that the sequence $\{a(j) : j = 0, 1, \dots\}$ which determines the functional $A\xi$ satisfies the following conditions

$$\sum_{j=0}^{\infty} |a(j)| < \infty, \quad \sum_{j=0}^{\infty} (j+1) |a(j)|^2 < \infty. \tag{122}$$

It follows from the spectral representation of the sequence $\xi(j)$ that we can represent the functional $A\xi$ in the form

$$A\xi = \int_{-\pi}^{\pi} A(e^{i\lambda})Z_\xi(d\lambda), \tag{123}$$

where

$$A(e^{i\lambda}) = \sum_{j=0}^{\infty} a(j)e^{ij\lambda}.$$

Denote by $\hat{A}\xi$ the mean square optimal linear estimate of the functional $A\xi$ from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$. Denote by $\Delta(f, g) = E |A\xi - \hat{A}\xi|^2$ the mean square error of the estimate $\hat{A}\xi$. To find the estimate $\hat{A}\xi$ we will use the Hilbert space projection method proposed by Kolmogorov [24]. We will consider $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, as elements of the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ of complex valued random variables with zero first moment, $E\xi = 0$, finite second moment, $E|\xi|^2 < \infty$, and the inner product $(\xi, \eta) = E\xi\bar{\eta}$.

Denote by $H^0(\xi + \eta)$ the subspace of the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ generated by elements $\{\xi(j) + \eta(j) : j = -1, -2, \dots\}$. Denote by $L_2(f + g)$ the Hilbert space of complex-valued functions that are square-integrable with respect to the measure whose density is $f(\lambda) + g(\lambda)$. Denote by $L_2^0(f + g)$ the subspace of $L_2(f + g)$ generated by functions $\{e^{ij\lambda} : j = -1, -2, \dots\}$.

The mean square optimal linear estimate $\hat{A}\xi$ of the functional $A\xi$ based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$ is an element of the $H^0(\xi + \eta)$. It can be represented in the form

$$\hat{A}\xi = \int_{-\pi}^{\pi} h(e^{i\lambda})(Z_\xi(d\lambda) + Z_\eta(d\lambda)), \tag{124}$$

where $h(e^{i\lambda}) \in L_2^0(f + g)$ is the spectral characteristic of the estimate $\hat{A}\xi$.

The mean square error $\Delta(h; f, g)$ of the estimate $\hat{A}\xi$ is given by the formula

$$\Delta(h; f, g) = E |A\xi - \hat{A}\xi|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |A(e^{i\lambda}) - h(e^{i\lambda})|^2 f(\lambda)d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} |h(e^{i\lambda})|^2 g(\lambda)d\lambda.$$

The Hilbert space projection method proposed by A. N. Kolmogorov [24] makes it possible to find the spectral characteristic $h(e^{i\lambda})$ and the mean square error $\Delta(h; f, g)$ of the optimal linear estimate of the functional $A\xi$ in the case where the spectral densities $f(\lambda)$ and $g(\lambda)$ of the sequences $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, are exactly known and the minimality condition (121) is satisfied. The spectral characteristic is determined by the following conditions:

- 1) $\hat{A}\xi \in H^0(\xi + \eta)$,
- 2) $A\xi - \hat{A}\xi \perp H^0(\xi + \eta)$.

It follows from the second condition that the following equations should be satisfied

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

The last equations are equivalent to equations

$$\int_{-\pi}^{\pi} [A(e^{i\lambda})f(\lambda) - h(e^{i\lambda})(f(\lambda) + g(\lambda))] e^{-ij\lambda} d\lambda = 0, \quad j = -1, -2, \dots$$

It follows from these equations that the function $[A(e^{i\lambda})f(\lambda) - h(e^{i\lambda})(f(\lambda) + g(\lambda))]$ is of the form

$$A(e^{i\lambda})f(\lambda) - h(e^{i\lambda})(f(\lambda) + g(\lambda)) = C(e^{i\lambda}), \quad (125)$$

$$C(e^{i\lambda}) = \sum_{j=0}^{\infty} c(j)e^{ij\lambda},$$

where $c(j), j = 0, 1, \dots$ are unknown coefficients that we have to find.

From the relation (125) we deduce that the spectral characteristic $h(e^{i\lambda})$ of the optimal linear estimate of the functional $A\xi$ is of the form

$$\begin{aligned} h(e^{i\lambda}) &= \frac{A(e^{i\lambda})f(\lambda) - C(e^{i\lambda})}{f(\lambda) + g(\lambda)} = \\ &= A(e^{i\lambda}) - \frac{A(e^{i\lambda})g(\lambda) + C(e^{i\lambda})}{f(\lambda) + g(\lambda)}. \end{aligned} \quad (126)$$

It follows from the first condition, which determines the spectral characteristic $h(e^{i\lambda}) \in L_2^0(f + g)$ of the optimal linear estimate of the functional $A\xi$ that the Fourier coefficients of the function $h(e^{i\lambda})$ are equal to zero for $j = 0, 1, \dots$, namely

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

Using the last relations and (126) we get the following system of equations

$$\int_{-\pi}^{\pi} \left(A(e^{i\lambda}) \frac{f(\lambda)}{f(\lambda) + g(\lambda)} - \frac{C(e^{i\lambda})}{f(\lambda) + g(\lambda)} \right) e^{-ij\lambda} d\lambda = 0, \quad j = 0, 1, \dots$$

These equations can be represented in the form

$$\sum_{k=0}^{\infty} a(k) \int_{-\pi}^{\pi} \frac{e^{i(k-j)\lambda} f(\lambda)}{f(\lambda) + g(\lambda)} d\lambda - \sum_{k=0}^{\infty} c(k) \int_{-\pi}^{\pi} \frac{e^{i(k-j)\lambda}}{f(\lambda) + g(\lambda)} d\lambda = 0, \quad j = 0, 1, \dots \quad (127)$$

Let us introduce the following notations

$$R_{j,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)\lambda} \frac{f(\lambda)}{f(\lambda) + g(\lambda)} d\lambda;$$

$$B_{j,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)\lambda} \frac{1}{f(\lambda) + g(\lambda)} d\lambda;$$

$$Q_{j,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)\lambda} \frac{f(\lambda)g(\lambda)}{f(\lambda) + g(\lambda)} d\lambda.$$

Making use the introduced notations we can write equations (127) in the form

$$\sum_{k=0}^{\infty} R_{j,k} a(k) = \sum_{k=0}^{\infty} B_{j,k} c(k), \quad j = 0, 1, \dots$$

The derived equations can be written in the form

$$\mathbf{R}\mathbf{a} = \mathbf{B}\mathbf{c},$$

where $\mathbf{a} = (a(0), a(1), \dots)$ is a vector constructed from the coefficients that determine the functional $A\xi$, $\mathbf{c} = (c(0), c(1), \dots)$ is a vector constructed from the unknown coefficients $c(k), k = 0, 1, \dots$, \mathbf{B} and \mathbf{R} are linear operators in ℓ_2 , which are determined by matrices with elements $(\mathbf{B})_{j,k} = B_{j,k}, (\mathbf{R})_{j,k} = R_{j,k}, j, k = 0, 1, \dots$

We get the formula

$$\mathbf{c} = \mathbf{B}^{-1}\mathbf{R}\mathbf{a}, \quad (128)$$

Hence, the unknown coefficients $c(j), j = 0, 1, \dots$, are calculated by the formula

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

where $(\mathbf{B}^{-1}\mathbf{R}\mathbf{a})_j$ is the j -th component of the vector $\mathbf{B}^{-1}\mathbf{R}\mathbf{a}$, and the formula for calculating the spectral characteristic of the estimate $\hat{A}\xi$ is of the form

$$h(e^{i\lambda}) = A(e^{i\lambda}) \frac{f(\lambda)}{f(\lambda) + g(\lambda)} - \frac{\sum_{k=0}^{\infty} (\mathbf{B}^{-1}\mathbf{R}\mathbf{a})_k e^{ik\lambda}}{f(\lambda) + g(\lambda)}. \quad (129)$$

The mean square error of the estimate of the function can be calculated by the formula

$$\begin{aligned} \Delta(h; f, g) &= E |A\xi - \hat{A}\xi|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A(e^{i\lambda})g(\lambda) + \sum_{k=0}^{\infty} (\mathbf{B}^{-1}\mathbf{R}\mathbf{a})_k e^{ik\lambda} \right|^2}{(f(\lambda) + g(\lambda))^2} f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A(e^{i\lambda})f(\lambda) - \sum_{k=0}^{\infty} (\mathbf{B}^{-1}\mathbf{R}\mathbf{a})_k e^{ik\lambda} \right|^2}{(f(\lambda) + g(\lambda))^2} g(\lambda) d\lambda \\ &= \langle \mathbf{R}\mathbf{a}, \mathbf{B}^{-1}\mathbf{R}\mathbf{a} \rangle + \langle \mathbf{Q}\mathbf{a}, \mathbf{a} \rangle, \end{aligned} \quad (130)$$

where $\langle \mathbf{a}, \mathbf{c} \rangle = \sum_{k=0}^{\infty} a(k)\overline{c(k)}$ is the inner product in the space ℓ_2 and \mathbf{Q} is a linear operator in ℓ_2 , which is determined by the matrix with elements $(\mathbf{Q})_{j,k} = Q_{j,k}, j, k = 0, 1, \dots$

Let us summarize our results and present them in the form of a theorem.

Theorem 5.1

Let $\xi(j)$ and $\eta(j)$ be stationary stochastic sequences with the spectral densities $f(\lambda)$ and $g(\lambda)$ that satisfy the minimality condition (121). Let conditions (122) be satisfied. The spectral characteristic $h(e^{i\lambda})$ and the mean square error $\Delta(h; f, g)$ of the optimal linear estimate $\hat{A}\xi$ of the functional $A\xi$ from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$ can be calculated by formulas (129), (130).

5.2. Minimax-robust method of extrapolation

The traditional methods of estimation of the functional $A\xi$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ can be applied in the case where the spectral densities $f(\lambda)$ and $g(\lambda)$ of the considered stochastic sequences $\xi(j)$ and $\eta(j)$ are exactly known. In practise, however, we do not have complete information on spectral densities of the sequences. For this reason we apply the minimax(robust) method of estimation of the functional $A\xi$, that is we find an estimate that minimizes the maximum of the mean square errors for all spectral densities from the given class of admissible spectral densities D .

Definition 5.1. For a given class of spectral densities $D = D_f \times D_g$ the spectral densities $f_0(\lambda) \in D_f, g_0(\lambda) \in D_g$ are called the least favourable in D for the optimal linear estimation of the functional $A\xi$ 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 5.2. 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 $A\xi$ is called minimax-robust if

$$h^0(e^{i\lambda}) \in H_D = \bigcap_{(f,g) \in D_f \times D_g} L_2^0(f + g),$$

$$\min_{h \in H_D} \max_{(f,g) \in D} \Delta(h; f, g) = \sup_{(f,g) \in D} \Delta(h^0; f, g).$$

It follows from the introduced definitions and the obtained formulas that the following statement holds true.

Lemma 5.1

The spectral densities $f_0(\lambda) \in D_f, g_0(\lambda) \in D_g$ are the least favourable in the class of admissible spectral densities $D = D_f \times D_g$ for the optimal linear estimate of the functional $A\xi$ if the Fourier coefficients of the 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 $\mathbf{B}^0, \mathbf{R}^0, \mathbf{Q}^0$ that determine a solution to the optimization problem

$$\begin{aligned} & \max_{(f,g) \in D_f \times D_g} \langle \mathbf{R}\mathbf{a}, \mathbf{B}^{-1}\mathbf{R}\mathbf{a} \rangle + \langle \mathbf{Q}\mathbf{a}, \mathbf{a} \rangle \\ & = \langle \mathbf{R}^0\mathbf{a}, (\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a} \rangle + \langle \mathbf{Q}^0\mathbf{a}, \mathbf{a} \rangle. \end{aligned} \tag{131}$$

The minimax spectral characteristic $h^0 = h(f_0, g_0)$ can be calculated by the formula (129) if $h(f_0, g_0) \in H_D$.

The least favourable 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

$$\Delta(h; f_0, g_0) \geq \Delta(h^0; f_0, g_0) \geq \Delta(h^0; f, g)$$

$$\forall h \in H_D, \forall f \in D_f, \forall g \in D_g$$

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

$$\sup_{(f,g) \in D_f \times D_g} \Delta(h(f_0, g_0); f, g) = \Delta(h(f_0, g_0); f_0, g_0), \tag{132}$$

where

$$\begin{aligned} \Delta(h(f_0, g_0); f, g) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A(e^{i\lambda})g_0(\lambda) + \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right|^2}{(f_0(\lambda) + g_0(\lambda))^2} f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A(e^{i\lambda})f_0(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right|^2}{(f_0(\lambda) + g_0(\lambda))^2} g(\lambda) d\lambda, \end{aligned} \tag{133}$$

The constrained optimization problem (132) is equivalent to the unconstrained optimization problem

$$\Delta_D(f, g) = -\Delta(h(f_0, g_0); f, g) + \delta(f, g | D_f \times D_g) \rightarrow \inf, \tag{134}$$

where $\delta(f, g | D_f \times D_g)$ is the indicator function of the set $D = D_f \times D_g$. Solution (f_0, g_0) to the problem (134) is characterized by the condition $0 \in \partial\Delta_D(f_0, g_0)$, where $\partial\Delta_D(f_0, g_0)$ is the subdifferential of the convex functional $\Delta_D(f, g)$ at point (f_0, g_0) . This condition makes it possible to find the least favourable spectral densities in some special classes of spectral densities D (see books [20], [70], [72] for additional details).

Note, that the form of the functional $\Delta(h(f_0, g_0); f, g)$ (133) is convenient for application the Lagrange method of indefinite multipliers for finding solution to the problem (132). Making use the method of Lagrange multipliers and the form of subdifferentials of the indicator functions we describe relations that determine least favourable spectral densities in some special classes of spectral densities (see books [13, 58, 64] for additional details).

Lemma 5.2

Let (f_0, g_0) be a solution to the optimization problem (134). The spectral densities $f_0(\lambda), g_0(\lambda)$ are the least favourable in the class $D = D_f \times D_g$, and the spectral characteristic $h^0 = h(f_0, g_0)$ is minimax for the optimal estimate of the functional $A\xi$, in the case where $h(f_0, g_0) \in H_D$.

5.3. Least favourable spectral densities in the class $D_f^0 \times D_g^0$

Consider the problem of the optimal estimation of the functional $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$ in the case where the spectral densities $f(\lambda), g(\lambda)$ are from the class $D = D_f^0 \times D_g^0$, where

$$\begin{aligned} D_f^0 &= \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) d\lambda \leq P_1 \right. \right\}, \\ D_g^0 &= \left\{ g(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) d\lambda \leq P_2 \right. \right\}. \end{aligned}$$

Let the densities $f_0(\lambda) \in D_f^0, g_0(\lambda) \in D_g^0$ and the functions $h_f(f_0, g_0), h_g(f_0, g_0)$, determined by the relations

$$h_f(f_0, g_0) = \frac{\left| A(e^{i\lambda})g_0(\lambda) + \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right|^2}{(f_0(\lambda) + g_0(\lambda))^2}, \tag{135}$$

$$h_g(g_0, g_0) = \frac{\left| A(e^{i\lambda})f_0(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right|^2}{(f_0(\lambda) + g_0(\lambda))^2}, \tag{136}$$

be bounded. In this case the 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 linear and continuous on the space $L_1 \times L_1$ and we can apply the method of Lagrange multipliers to find solution to the optimization problem (134). We get the following relations that determine least favourable spectral densities $f^0 \in D_f^0, g^0 \in D_g^0$

$$\begin{aligned} & -\frac{1}{2\pi} \int_{-\pi}^{\pi} h_f(f_0, g_0) \rho(f(\lambda)) d\lambda - \frac{1}{2\pi} \int_{-\pi}^{\pi} h_g(f_0, g_0) \rho(g(\lambda)) d\lambda \\ & + \alpha_1 \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho(f(\lambda)) d\lambda + \alpha_2 \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho(g(\lambda)) d\lambda = 0, \end{aligned}$$

where $\rho(f(\lambda))$ and $\rho(g(\lambda))$ are variations of the functions $f(\lambda)$ and $g(\lambda)$, the constants $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$. From this relation we get that the least favourable spectral densities $f_0(\lambda) \in D_f^0, g_0(\lambda) \in D_g^0$ satisfy equations

$$\left| A(e^{i\lambda})g_0(\lambda) + \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right| = \alpha_1 (f_0(\lambda) + g_0(\lambda)), \tag{137}$$

$$\left| A(e^{i\lambda})f_0(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right| = \alpha_2 (f_0(\lambda) + g_0(\lambda)). \tag{138}$$

Note, that $\alpha_1 \neq 0$ in the case, where

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_0(\lambda) d\lambda = P_1,$$

and $\alpha_2 \neq 0$ in the case, where

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} g_0(\lambda) d\lambda = P_2.$$

Summing up our reasoning we come to conclusion that the following theorem holds true.

Theorem 5.2

Let the spectral densities $f_0(\lambda) \in D_f^0$ and $g_0(\lambda) \in D_g^0$ satisfy the minimality condition (121) and let the functions $h_f(f_0, g_0)$ and $h_g(f_0, g_0)$, determined by the formulas (135), (136), be bounded. The functions $f_0(\lambda)$ and $g_0(\lambda)$, which give solution to the system of equations (137), (138) are the least favourable spectral densities in the class $D = D_f^0 \times D_g^0$ for the optimal estimation of the functional $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$, if they determine a solution to the optimization problem (131). The function $h^0(e^{i\lambda})$, determined by the formula (129), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A\xi$.

Theorem 5.3

Let the spectral density $f(\lambda)$ be known and fixed and the spectral density $g_0(\lambda) \in D_g^0$. Let the functions $f(\lambda)$ and $g_0(\lambda)$ be such that the function $(f(\lambda) + g_0(\lambda))^{-1}$ is integrable and let the function $h_g(f, g_0)$, determined by the formula (136), be bounded. The spectral density $g_0(\lambda)$ is the least favourable spectral densities in the class D_g^0 for the optimal estimation of the functional $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$, if it is of the form

$$g_0(\lambda) = \max \left\{ 0, \alpha_2^{-1} \left| A(e^{i\lambda})f(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1} \mathbf{R}^0 \mathbf{a})_j e^{ij\lambda} \right| - f(\lambda) \right\}$$

and the functions $f(\lambda)$ and $g_0(\lambda)$ determine a solution to the optimization problem (131). The function $h^0(e^{i\lambda})$, determined by the formula (129), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A\xi$.

5.4. Least favourable spectral densities in the class $D_v^u \times D_\varepsilon$

Consider the problem for the optimal estimation of the functional $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$ in the case where the spectral densities $f(\lambda), g(\lambda)$ are from the class $D = D_v^u \times D_\varepsilon$, where

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

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

Here the spectral densities $v(\lambda), u(\lambda), g_1(\lambda)$ are known and fixed and the densities $v(\lambda)$ and $u(\lambda)$ are bounded.

Let the densities $f^0(\lambda) \in D_v^u, g^0(\lambda) \in D_\varepsilon$ determine bounded functions $h_f(f_0, g_0), h_g(f_0, g_0)$ with the help of formulas (135), (136). Then from the condition $0 \in \partial \Delta_{D_{f,g}}(f_0, g_0)$ we derive the following equations that determine the least favourable spectral densities

$$\left| A(e^{i\lambda})g_0(\lambda) + \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1} \mathbf{R}^0 \mathbf{a})_j e^{ij\lambda} \right| = (f_0(\lambda) + g_0(\lambda))(\gamma_1(\lambda) + \gamma_2(\lambda) + \alpha_1), \tag{139}$$

$$\left| A(e^{i\lambda})f_0(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1} \mathbf{R}^0 \mathbf{a})_j e^{ij\lambda} \right| = (f_0(\lambda) + g_0(\lambda))(\varphi(\lambda) + \alpha_2), \tag{140}$$

where $\gamma_1 \leq 0$ and $\gamma_1 = 0$ in the case $f_0(\lambda) \geq v(\lambda)$; $\gamma_2 \geq 0$ and $\gamma_2 = 0$ in the case $f_0(\lambda) \leq u(\lambda)$; $\varphi(\lambda) \leq 0$ and $\varphi(\lambda) = 0$ in the case $g_0(\lambda) \geq (1 - \varepsilon)g_1(\lambda)$.

The following theorems hold true.

Theorem 5.4

Let the spectral densities $f_0(\lambda) \in D_v^u$ and $g_0(\lambda) \in D_\varepsilon$ satisfy the minimality condition (121) and let the functions $h_f(f_0, g_0)$ and $h_g(f_0, g_0)$, determined by the formulas (135), (136), be bounded. The functions $f_0(\lambda), g_0(\lambda)$, which give solution to the system of equations (139), (140) are the least favourable spectral densities in the class $D_v^u \times D_\varepsilon$, if they determine a solution to the optimization problem (131). The function $h^0(e^{i\lambda})$, determined by the formula (129), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A\xi$.

Theorem 5.5

Let the spectral density $f(\lambda)$ be known and fixed and let the spectral density $g_0(\lambda) \in D_\varepsilon$. Let the function $f(\lambda) + g_0(\lambda)$ satisfy the minimality condition (121), and let the function $h_g(f, g_0)$, determined by the formula

(136), be bounded. The spectral density $g_0(\lambda)$ is the least favourable spectral densities in the class D_ε for the optimal linear estimate of the functional $A\xi$, based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$, if it is of the form

$$g_0(\lambda) = \max \left\{ (1 - \varepsilon)g_1(\lambda), \alpha_2 \left| A(e^{i\lambda})f(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right| - f(\lambda) \right\}$$

and the functions $f(\lambda), g_0(\lambda)$ determine a solution to the optimization problem (131). The function $h^0(e^{i\lambda})$, determined by the formula (129), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A\xi$.

5.5. Least favourable spectral densities in the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$

Consider the problem for the optimal estimation of the functional $A\xi = \sum_{j=0}^{\infty} a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$ in the case where the spectral densities $f(\lambda), g(\lambda)$ are from the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$, which describe the model of “ ε -neighbourhood of spectral densities in the space $L_2 \times L_1$ ”. Let

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

be the “ ε -neighbourhood” in the space L_2 of a given bounded spectral density $f_1(\lambda)$ and let

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

be the “ ε -neighbourhood” in the space L_1 of a given bounded spectral density $g_1(\lambda)$.

Let the spectral densities $f_0(\lambda) \in D_{2\varepsilon_1}, g_0(\lambda) \in D_{1\varepsilon_2}$ determine the bounded functions $h_f(f_0, g_0), h_g(f_0, g_0)$ with the help of formulas (135), (136). Then from the condition $0 \in \partial\Delta_{D_{f,g}}(f_0, g_0)$ for $D = D_{2\varepsilon_1} \times D_{1\varepsilon_2}$ we derive the following equations that determine the least favourable spectral densities

$$\left| A(e^{i\lambda})g_0(\lambda) + \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right|^2 = (f_0(\lambda) + g_0(\lambda))^2 (f_0(\lambda) - f_1(\lambda))\alpha_1, \tag{141}$$

$$\left| A(e^{i\lambda})f_0(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right|^2 = (f_0(\lambda) + g_0(\lambda))^2 \Psi(\lambda)\alpha_2, \tag{142}$$

where $|\Psi(\lambda)| \leq 1$ and $\Psi(\lambda) = \text{sign}(g_0(\lambda) - g_1(\lambda))$, in the case $g_0(\lambda) \neq g_1(\lambda)$, α_1, α_2 are constants.

Equations (141), (142) with the optimization problem (131) and the normalizing conditions

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\lambda) - f_1(\lambda)|^2 d\lambda = \varepsilon_1 \tag{143}$$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |g(\lambda) - g_1(\lambda)| d\lambda = \varepsilon_2 \tag{144}$$

determine the least favourable spectral densities in the class $D = D_{2\varepsilon_1} \times D_{1\varepsilon_2}$.

The following theorems hold true.

Theorem 5.6

Let the spectral densities $f_0(\lambda) \in D_{2\varepsilon_1}$, $g_0(\lambda) \in D_{1\varepsilon_2}$ satisfy the minimality condition (121) and let the functions $h_f(f_0, g_0)$ and $h_g(f_0, g_0)$, determined by the formulas (135), (136), be bounded. The spectral densities $f_0(\lambda), g_0(\lambda)$, which give solution to the system of equations (141)–(144) are the least favourable spectral densities in the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$, if they determine a solution to the optimization problem (131). The function $h^0(e^{i\lambda})$, determined by the formula (129), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A\xi$.

Theorem 5.7

Let the spectral density $f(\lambda)$ be known, and let the spectral density $g_0(\lambda) \in D_{1\varepsilon_2}$. Let the function $f(\lambda) + g_0(\lambda)$ satisfy the minimality condition (121), and let the function $h_g(f, g_0)$, determined by the formula (136), be bounded. The spectral density $g_0(\lambda)$ is the least favourable spectral densities in the class $D_{1\varepsilon_2}$ for the optimal linear estimate of the functional $A\xi$, based on observations of the sequence $\xi(j)$ at points of time $j = -1, -2, \dots$, if it is of the form

$$g_0(\lambda) = \max \left\{ g_1(\lambda), \alpha_2 \left| A(e^{i\lambda})f(\lambda) - \sum_{j=0}^{\infty} ((\mathbf{B}^0)^{-1}\mathbf{R}^0\mathbf{a})_j e^{ij\lambda} \right| - f(\lambda) \right\}$$

and the functions $f(\lambda), g_0(\lambda)$ determine a solution to the optimization problem (131). The function $h^0(e^{i\lambda})$, determined by the formula (129), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A\xi$.

5.6. Conclusions

In this section we propose methods of solution of the problem of the 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 stationary stochastic sequence $\xi(j)$.

Estimates are based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j = -1, -2, \dots$. Here $\eta(j)$ is an uncorrelated with $\xi(j)$ stationary sequence. We provide formulas for calculating the values of the mean square error and the spectral characteristic of the optimal linear estimate of the functional in the case of spectral certainty, where the spectral densities $f(\lambda)$ and $g(\lambda)$ of the sequences $\xi(j)$ and $\eta(j)$ are exactly known. In the case of spectral uncertainty, where the spectral densities $f(\lambda)$ and $g(\lambda)$ are not known, but a set of admissible spectral densities is given, the minimax approach is applied to estimation of the functionals. We obtain formulas that determine the least favourable spectral densities and the minimax spectral characteristics of the optimal linear estimates of the functional $A\xi$ for concrete classes of admissible spectral densities.

For the relative results on the mean-square optimal linear extrapolation of linear functionals for stationary stochastic sequences and processes based on observations with noise see papers by Moklyachuk [46] – [48], [52] – [57], book by Moklyachuk and Masyutka [64].

6. Interpolation problem for stationary sequences

In this section we deal with the problem of the mean-square optimal estimation of the linear functional $A_N\xi = \sum_{j=0}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j), j \in \mathbb{Z}$, based on observations of the sequence at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$. The problem is investigated in the case of spectral certainty, where the spectral density of the stationary stochastic sequence $\xi(j)$ is exactly known. In this case the classical Hilbert space projection method of linear estimation of the functional is applied. Formulas are derived for calculation the value of the mean square error and the spectral characteristic of the mean-square optimal estimate of the linear functional. In the case of spectral uncertainty, where the spectral density of the stationary stochastic sequence is not exactly known, but a class of admissible spectral densities is given the minimax-robust procedure to linear estimation of the functional is applied. Relations which determine the least favourable spectral densities and the minimax spectral characteristics are proposed for some special sets of admissible spectral densities.

6.1. The classical Hilbert space projection method of linear interpolation

Let $\xi(j)$, $j \in \mathbb{Z}$, be a (wide sense) stationary stochastic sequence. We will consider values of $\xi(j)$, $j \in \mathbb{Z}$, as elements of the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ of complex valued random variables with zero first moment, $E\xi = 0$, finite second moment, $E|\xi|^2 < \infty$, and the inner product $(\xi, \eta) = E\xi\bar{\eta}$. The correlation function $R(k) = (\xi(j+k), \xi(j)) = E\xi(j+k)\bar{\xi(j)}$ of the stationary stochastic sequence $\xi(j)$, $j \in \mathbb{Z}$, admits the spectral representation [12]

$$R(k) = \int_{-\pi}^{\pi} e^{ik\lambda} F(d\lambda),$$

where $F(d\lambda)$ is the spectral measure of the sequence. We will consider stationary stochastic sequences with absolutely continuous spectral measures and the correlation functions of the form

$$R(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda,$$

where $f(\lambda)$ is the spectral density function of the sequence $\xi(j)$ that satisfies the minimality condition

$$\int_{-\pi}^{\pi} f^{-1}(\lambda) d\lambda < \infty. \quad (145)$$

Under this condition the error-free interpolation of the unknown values of the sequence is impossible [73]. The stationary stochastic sequence $\xi(j)$, $j \in \mathbb{Z}$, admits the spectral representation [12, 22]

$$\xi(j) = \int_{-\pi}^{\pi} e^{ij\lambda} Z(d\lambda), \quad (146)$$

where $Z(\Delta)$ is the orthogonal stochastic measure of the sequence such that

$$EZ(\Delta_1)\overline{Z(\Delta_2)} = F(\Delta_1 \cap \Delta_2) = \frac{1}{2\pi} \int_{\Delta_1 \cap \Delta_2} f(\lambda) d\lambda.$$

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

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

which depends on the unknown values of a stationary stochastic sequence $\xi(j)$, $j \in \mathbb{Z}$, from observations of the sequence at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$.

It follows from the spectral representation (146) of the sequence $\xi(j)$ that we can represent the functional $A_N \xi$ in the form

$$A_N \xi = \int_{-\pi}^{\pi} A_N(e^{i\lambda}) Z(d\lambda), \quad (147)$$

where

$$A_N(e^{i\lambda}) = \sum_{j=0}^N a(j) e^{ij\lambda}.$$

Denote by $H^N(\xi)$ the subspace of the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ generated by elements $\{\xi(j) : j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}\}$. Let $L_2(f)$ be the Hilbert space of complex-valued functions that are square-integrable with

respect to the measure whose density is $f(\lambda)$. Denote by $L_2^N(f)$ the subspace of $L_2(f)$ generated by functions $\{e^{ij\lambda}, j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}\}$. The mean square optimal linear estimate $\hat{A}_N \xi$ of the functional $A_N \xi$ from observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ is an element of the $H^N(\xi)$. It can be represented in the form

$$\hat{A}_N \xi = \int_{-\pi}^{\pi} h(e^{i\lambda}) Z(d\lambda), \tag{148}$$

where $h(e^{i\lambda}) \in L_2^N(f)$ is the spectral characteristic of the estimate $\hat{A}_N \xi$.

The mean square error $\Delta(h; f)$ of the estimate $\hat{A}_N \xi$ is given by the formula

$$\Delta(h; f) = E \left| A_N \xi - \hat{A}_N \xi \right|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |A_N(e^{i\lambda}) - h(e^{i\lambda})|^2 f(\lambda) d\lambda.$$

The Hilbert space projection method proposed by A. N. Kolmogorov [24] makes it possible to find the spectral characteristic $h(e^{i\lambda})$ and the mean square error $\Delta(h; f)$ of the optimal linear estimate of the functional $A_N \xi$ in the case where the spectral density $f(\lambda)$ of the sequence $\xi(j), j \in \mathbb{Z}$, is exactly known and the minimality condition (145) is satisfied. The spectral characteristic can be found from the following conditions:

- 1) $h(e^{i\lambda}) \in L_2^N(f)$,
- 2) $A_N(e^{i\lambda}) - h(e^{i\lambda}) \perp L_2^N(f)$.

It follows from the second condition that for any $\eta \in H^N(\xi)$ the following equation should be satisfied

$$(A_N \xi - \hat{A}_N \xi, \eta) = E(A_N \xi - \hat{A}_N \xi) \bar{\eta} = 0.$$

The last relation is equivalent to equations

$$E(A_N \xi - \hat{A}_N \xi) \bar{\xi}_k = 0, \quad k \in \mathbb{Z} \setminus \{0, 1, \dots, N\}.$$

By using representations (147), (148) and definition of the inner product in the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ we get

$$\begin{aligned} & E \left(\int_{-\pi}^{\pi} (A_N(e^{i\lambda}) - h(e^{i\lambda})) Z(d\lambda) \cdot \int_{-\pi}^{\pi} e^{-ik\lambda} \overline{Z(d\lambda)} \right) = \\ & = \frac{1}{2\pi} \int_{-\pi}^{\pi} (A_N(e^{i\lambda}) - h(e^{i\lambda})) f(\lambda) e^{-ik\lambda} d\lambda = 0, \quad k \in \mathbb{Z} \setminus \{0, 1, \dots, N\}. \end{aligned}$$

It follows from these equations that the function $(A_N(e^{i\lambda}) - h(e^{i\lambda}))f(\lambda)$ is of the form

$$(A_N(e^{i\lambda}) - h(e^{i\lambda}))f(\lambda) = C_N(e^{i\lambda}), \tag{149}$$

$$C_N(e^{i\lambda}) = \sum_{j=0}^N c(j) e^{ij\lambda},$$

where $c(j), j = 0, 1, \dots, N$, are unknown coefficients that we have to find.

From the relation (149) we deduce that the spectral characteristic $h(e^{i\lambda})$ of the optimal linear estimate of the functional $A_N \xi$ is of the form

$$h(e^{i\lambda}) = A_N(e^{i\lambda}) - C_N(e^{i\lambda})f^{-1}(\lambda). \tag{150}$$

To find equations for determination the unknown coefficients $c(j), j = 0, 1, \dots, N$, we use the decomposition of the function $f^{-1}(\lambda)$ into the Fourier series

$$f^{-1}(\lambda) = \sum_{m=-\infty}^{\infty} r_m e^{im\lambda}, \quad (151)$$

where r_m are the Fourier coefficients of the function $f^{-1}(\lambda)$

$$r_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{-1}(\lambda) e^{-im\lambda} d\lambda.$$

Inserting (151) into (150) we obtain the following representation of the spectral characteristic

$$h(e^{i\lambda}) = \left(\sum_{j=0}^N a(j) e^{ij\lambda} \right) - \left(\sum_{j=0}^N c(j) e^{ij\lambda} \right) \left(\sum_{m=-\infty}^{\infty} r_m e^{im\lambda} \right). \quad (152)$$

It follows from the first condition, $h(e^{i\lambda}) \in L_2^N(f)$, which determines the spectral characteristic, that the Fourier coefficients of the function $h(e^{i\lambda})$ are equal to zero for $j = 0, 1, \dots, N$, namely

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

Using the last relations and (152) we get the following system of equations that determine the unknown coefficients $c(j), j = 0, 1, \dots, N$,

$$\begin{aligned} a(0) - \sum_{j=0}^N c(j) r_{-j} &= 0; \\ a(1) - \sum_{j=0}^N c(j) r_{1-j} &= 0; \\ &\dots \\ a(N) - \sum_{j=0}^N c(j) r_{N-j} &= 0. \end{aligned} \quad (153)$$

Denote by $\vec{a}_N = (a(0), a(1), \dots, a(N))$ and let $\vec{c}_N = (c(0), c(1), \dots, c(N))$ be a vector constructed from the unknown coefficients $c(j), j = 0, 1, \dots, N$. Let B_N be an $(N+1) \times (N+1)$ matrix

$$B_N = \begin{pmatrix} B_{00} & B_{01} & \dots & B_{0s} \\ B_{10} & B_{11} & \dots & B_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ B_{N0} & B_{N1} & \dots & B_{NN} \end{pmatrix},$$

with elements that are Fourier coefficients of the function $f^{-1}(\lambda)$:

$$B_{kj} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{-1}(\lambda) e^{-i(k-j)\lambda} d\lambda = r_{k-j}, \quad k = 0, 1, \dots, N, \quad j = 0, 1, \dots, N.$$

Making use the introduced notations we can write equations (153) in the form of equation

$$\vec{\mathbf{a}}_N = B_N \vec{\mathbf{c}}_N, \tag{154}$$

Since the matrix B_N is reversible [74], we get the formula

$$\vec{\mathbf{c}}_N = B_N^{-1} \vec{\mathbf{a}}_N. \tag{155}$$

Hence, the unknown coefficients $c(j), j = 0, 1, \dots, N$, are calculated by the formula

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

where $(B_N^{-1} \vec{\mathbf{a}}_N)_j$ is the j component of the vector $B_N^{-1} \vec{\mathbf{a}}_N$, and the formula for calculating the spectral characteristic of the estimate $\hat{A}_N \xi$ is of the form

$$h(e^{i\lambda}) = \left(\sum_{j=0}^N a(j) e^{ij\lambda} \right) - \left(\sum_{j=0}^N (B_N^{-1} \vec{\mathbf{a}}_N)_j e^{ij\lambda} \right) \left(\sum_{m=-\infty}^{\infty} r_m e^{im\lambda} \right). \tag{156}$$

The mean square error of the estimate of the function can be calculated by the formula

$$\begin{aligned} \Delta(h; f) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |C_N(e^{i\lambda})|^2 f^{-1}(\lambda) d\lambda = \\ &= \int_{-\pi}^{\pi} \left(\sum_{k=0}^N c(k) e^{ik\lambda} \right) \left(\sum_{j=0}^N \overline{c(j)} e^{-ij\lambda} \right) \left(\sum_{m=-\infty}^{\infty} r_m e^{im\lambda} \right) d\lambda = \\ &= \langle \vec{\mathbf{c}}_N, \vec{\mathbf{c}}_N B_N \rangle = \langle B_N^{-1} \vec{\mathbf{a}}_N, \vec{\mathbf{a}}_N \rangle, \end{aligned} \tag{157}$$

where $\langle \cdot, \cdot \rangle$ is the inner product in $C^{(N+1)}$.

Let us summarize our results and present them in the form of a theorem.

Theorem 6.1

Let $\xi(j)$ be a stationary stochastic sequence with the spectral density $f(\lambda)$ that satisfies the minimality condition (145). The spectral characteristic $h(e^{i\lambda})$ and the mean square error $\Delta(h, f)$ of the optimal linear estimate $\hat{A}_N \xi$ of the functional $A_N \xi$ based on observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ can be calculated by formulas (156), (157).

Example 1. Consider the problem of linear estimation of the functional $A_1 \xi = a\xi(0) + b\xi(1)$ which depends on the unknown values $\xi(0), \xi(1)$ of the stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j)$ at points $j \in \mathbb{Z} \setminus \{0, 1\}$. In this case the spectral characteristic (152) of the estimate $\hat{A}_1 \xi$ can be calculated by the formula

$$h(f) = a + be^{i\lambda} - (c(0) + c(1)e^{i\lambda}) f^{-1}(\lambda), \tag{158}$$

where $f(\lambda)$ is a known spectral density, coefficients $c(0), c(1)$ are determined by the system of equations

$$\begin{aligned} a &= c(0)\alpha + c(1)\beta, \\ b &= c(0)\bar{\beta} + c(1)\alpha, \end{aligned}$$

where

$$\alpha = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{-1}(\lambda) d\lambda,$$

$$\beta = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda} f^{-1}(\lambda) d\lambda,$$

The matrix B_1 is of the form

$$B_1 = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \alpha \end{pmatrix},$$

The determinant $D = \det(B_1)$ of the matrix B_1 is as follows $D = \det(B_1) = \alpha^2 - |\beta|^2$
 We get the following formulas for calculating the coefficients $c(0), c(1)$

$$c(0) = (a\alpha - b\beta)/(\alpha^2 - |\beta|^2),$$

$$c(1) = (b\alpha - a\bar{\beta})/(\alpha^2 - |\beta|^2).$$

Thus, the unknown coefficients $c(0), c(1)$ in (158) are determined.
 The mean square error of the estimate is calculated by the formula

$$\Delta(f) = [(|a|^2 + |b|^2) \alpha - (\bar{a}b + a\bar{b})\beta] / (\alpha^2 - |\beta|^2). \tag{159}$$

6.2. Minimax-robust method of interpolation

The traditional methods of estimation of the functional $A_N \xi$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ can be applied in the case where the spectral density $f(\lambda)$ of the considered stochastic sequence $\xi(j)$ is exactly known. In practice, however, we do not have complete information on spectral density of the sequence. For this reason we apply the minimax(robust) method of estimation of the functional $A_N \xi$, that is we find an estimate that minimizes the maximum of the mean square errors for all spectral densities from a given class of admissible spectral densities D .

Definition 6.1. For a given class of spectral densities D a spectral density $f_0(\lambda) \in D$ is called the least favourable in D for the optimal linear estimation of the functional $A_N \xi$ if the following relation holds true

$$\Delta(f_0) = \Delta(h(f_0); f_0) = \max_{f \in D} \Delta(h(f); f).$$

Definition 6.2. For a given class of spectral densities D the spectral characteristic $h^0(e^{i\lambda})$ of the optimal linear estimate of the functional $A_N \xi$ is called minimax-robust if

$$h^0(e^{i\lambda}) \in H_D = \bigcap_{f \in D} L_2^N(f),$$

$$\min_{h \in H_D} \max_{f \in D} \Delta(h; f) = \sup_{f \in D} \Delta(h^0; f).$$

It follows from the introduced definitions and the obtained formulas that the following statement holds true.

Lemma 6.1

The spectral density $f_0(\lambda) \in D$ is the least favourable in a class of admissible spectral densities D for the optimal linear estimation of the functional $A_N \xi$ if the Fourier coefficients of the function $f_0^{-1}(\lambda)$ define a matrix B_N^0 that determines a solution to the optimization problem

$$\max_{f \in D} \langle B_N^{-1} \bar{\mathbf{a}}_N, \bar{\mathbf{a}}_N \rangle = \langle (B_N^0)^{-1} \bar{\mathbf{a}}_N, \bar{\mathbf{a}}_N \rangle. \tag{160}$$

The minimax spectral characteristic $h^0 = h(f_0)$ can be calculated by the formula (156) if $h(f_0) \in H_D$.

The least favourable spectral density f_0 and the minimax spectral characteristic h^0 form a saddle point of the function $\Delta(h; f)$ on the set $H_D \times D$. The saddle point inequalities

$$\Delta(h; f_0) \geq \Delta(h^0; f_0) \geq \Delta(h^0; f) \quad \forall f \in D, \forall h \in H_D$$

hold true if $h^0 = h(f_0)$ and $h(f_0) \in H_D$, where f_0 is a solution to the constrained optimization problem

$$\tilde{\Delta}(f) = -\Delta(h^0; f) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|C_N^0(e^{i\lambda})|^2}{f_0^2(\lambda)} f(\lambda) d\lambda \rightarrow \inf, \quad f(\lambda) \in D, \quad (161)$$

where

$$C_N^0(e^{i\lambda}) = \sum_{j=0}^N ((B_N^0)^{-1} \vec{a}_N)_j e^{ij\lambda}.$$

The constrained optimization problem (161) is equivalent to the unconstrained optimization problem

$$\Delta_D(f) = \tilde{\Delta}(f) + \delta(f|D) \rightarrow \inf,$$

where $\delta(f|D)$ is the indicator function of the set D . Solution f_0 to this problem is characterized by the condition $0 \in \partial\Delta_D(f_0)$, where $\partial\Delta_D(f_0)$ is the subdifferential of the convex functional $\Delta_D(f)$ at point f_0 . This condition makes it possible to find the least favourable spectral densities in some special classes of spectral densities D [20], [70], [72].

Note, that the form of the functional $\Delta(h^0; f)$ is convenient for application the Lagrange method of indefinite multipliers for finding solution to the problem (161). Making use the method of Lagrange multipliers and the form of subdifferentials of the indicator functions we describe relations that determine least favourable spectral densities in some special classes of spectral densities (see books [13, 58, 64] for additional details).

6.3. Least favourable spectral densities in the class D_0^-

Consider the problem of the optimal estimation of the functional $A_N \xi = \sum_{j=1}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ in the case where the spectral density is from the class

$$D_0^- = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{-1}(\lambda) d\lambda \geq P \right. \right\}.$$

Let the sequence $a(k), k = 0, 1, \dots, N$, that determines the functional $A_N \xi$, be strictly positive. To find solutions to the constrained optimization problem (161) we use the Lagrange multipliers method. With the help of this method we get the equation

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{|C_N^0(e^{i\lambda})|^2}{f_0^2(\lambda)} - p_0^2 \frac{1}{f_0^2(\lambda)} \right] \rho(f(\lambda)) d\lambda = 0,$$

where p_0^2 is a constant (the Lagrange multiplier), $\rho(f(\lambda))$ is a variation of the function $f(\lambda)$. From a generalization of the Lagrange lemma we get that the Fourier coefficients of the function f_0^{-1} satisfy the equation

$$\left| \sum_{k=0}^N c(k) e^{ik\lambda} \right|^2 = p_0^2, \quad (162)$$

where $c(k), k = 0, 1, \dots, N$, are components of the vector \vec{c}_N that satisfies the equation

$$B_N^0 \vec{c}_N = \vec{a}_N, \quad (163)$$

the matrix B_N^0 is determined by the Fourier coefficients of the function $f_0^{-1}(\lambda)$

$$B_N^0(k, j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_0^{-1}(\lambda) e^{-i(k-j)\lambda} d\lambda = r_{k-j}^0,$$

$$\begin{aligned} k &= 0, 1, \dots, N, \\ j &= 0, 1, \dots, N. \end{aligned}$$

The Fourier coefficients $r_k = r_{-k}, k = 0, 1, \dots, N$, satisfy both equation (162) and equation (163). These coefficients can be found from the equation

$$B_N^0 \vec{p}_N^0 = \vec{a}_N, \quad \vec{p}_N^0 = (p_0, 0, \dots, 0).$$

The last relation can be presented in the form of the system of equations

$$r_k p_0 = a(k), \quad k = 0, 1, \dots, N.$$

From the first equation of the system (for $k = 0$) we find the unknown value

$$p_0 = a(0)(r_0)^{-1}.$$

It follows from the extremum condition (160) and the restriction on the spectral densities from the class D_0^- that the Fourier coefficient

$$r_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_0^{-1}(\lambda) d\lambda = P.$$

Thus,

$$r_k = Pa(k)a^{-1}(0), \quad k = 0, 1, \dots, N.$$

We can represent the function $f_0^{-1}(\lambda)$ in the form

$$f_0^{-1}(\lambda) = \sum_{k=-N}^N r_{|k|} e^{ik\lambda}.$$

Since the sequence $a(k), k = 0, 1, \dots, N$, is strictly positive, the sequence $r_k, k = 0, 1, \dots, N$, is also strictly positive and the function $f_0^{-1}(\lambda)$ is positive, so it can be represented in the form [25]

$$f_0^{-1}(\lambda) = \left| \sum_{k=0}^N \gamma_k e^{-ik\lambda} \right|^2, \quad \lambda \in [-\pi, \pi],$$

Hence, $f_0(\lambda)$ is the spectral density of the autoregressive stochastic sequence of order N generated by the equation

$$\sum_{k=0}^N \gamma_k \xi(n-k) = \varepsilon_n, \tag{164}$$

where ε_n is a “white noise” sequence.

The minimax spectral characteristic $h(f_0)$ of the optimal linear estimate of the functional $A_N \xi$ can be calculated by the formula (150), where

$$C_N(e^{i\lambda}) = \sum_{k=0}^N c(k) e^{ik\lambda} = p_0 = P^{-1}a(0),$$

namely

$$h(f_0) = \sum_{k=0}^N a(k) e^{ik\lambda} - P^{-1}a(0) \sum_{k=-N}^N r_k e^{ik\lambda} = \sum_{k=1}^N a(k) e^{-ik\lambda}. \tag{165}$$

Let the sequence $a(N), a(N-1), \dots, a(0)$ that determines the functional $A_N \xi$ be strictly positive. In this case the Fourier coefficients $r_k = r_{-k}, k = 0, 1, \dots, N$, can be found from the equation

$$B_N^0 \vec{p}_N^0 = \vec{a}_N, \quad \vec{p}_N^0 = (0, \dots, 0, p_0).$$

From this equation we find the unknown values of

$$r_k = Pa(N - k)a^{-1}(N), \quad k = 0, 1, \dots, N.$$

The minimax spectral characteristic $h(f_0)$ of the optimal linear estimate of the functional $A_N\xi$ can be calculated by the formula

$$h(f_0) = \sum_{k=0}^N a(k)e^{ik\lambda} - P^{-1}a(N)e^{iN\lambda} \sum_{k=-N}^N r_{|k|}e^{ik\lambda} = \sum_{k=1}^N a(N - k)e^{-i(N+k)\lambda}. \quad (166)$$

Summing up our reasoning we come to conclusion that the following theorem holds true.

Theorem 6.2

The least favourable in the class D_0^- spectral density for the optimal linear estimation of the functional $A_N\xi$ determined by strictly positive sequence $a(0), a(1), \dots, a(N)$ is the spectral density of the autoregressive sequence (164) whose Fourier coefficients are $r_k = r_{-k} = Pa(k)a^{-1}(0), k = 0, 1, \dots, N$. The minimax spectral characteristics $h(f_0)$ is given by formula (165). The least favourable in the class D_0^- spectral density for the optimal linear estimation of the functional $A_N\xi$ determined by strictly positive sequence $a(N), a(N - 1), \dots, a(0)$ is the spectral density of the autoregressive sequence whose Fourier coefficients are $r_k = r_{-k} = Pa(N - k)a^{-1}(N), k = 0, 1, \dots, N$. The minimax spectral characteristics $h(f_0)$ is given by formula (166).

Example 2. Consider the problem of the optimal linear estimation of the functional $A_1\xi = a\xi(0) + b\xi(1)$ which depends on the unknown values $\xi(0), \xi(1)$ of the stationary stochastic sequence $\{\xi(j) : j \in \mathbb{Z}\}$ from observations of the sequence at points of time $\mathbb{Z} \setminus \{0, 1\}$.

The least favourable spectral density in the class D_0^- is of the form

$$f_0(\lambda) = 1/|x + ye^{i\lambda}|^2,$$

where

$$x = \pm \left(P \left(1 \pm (1 - 4(b/a)^2)^{1/2} \right) / 2 \right)^{1/2},$$

$$y = \pm \left(P \left(1 \mp (1 - 4(b/a)^2)^{1/2} \right) / 2 \right)^{1/2}$$

under the condition $|b/a| < 1/2$. For example, in the case of $a = 4, b = \sqrt{3}$ the least favourable spectral density $f_0(\lambda)$ and the minimax spectral characteristic are calculated by the formulas

$$f_0(\lambda) = 4/P|\sqrt{3} + e^{i\lambda}|^2,$$

$$h(f_0) = -\sqrt{3}e^{-i\lambda}.$$

Under the condition $|b/a| > 2$ the least favourable spectral density $f_0(\lambda)$ is as follows

$$f_0(\lambda) = 1/|x + ye^{i\lambda}|^2,$$

where

$$x = \pm \left(P \left(1 \pm (1 - 4(a/b)^2)^{1/2} \right) / 2 \right)^{1/2},$$

$$y = \pm \left(P \left(1 \mp (1 - 4(a/b)^2)^{1/2} \right) / 2 \right)^{1/2}.$$

For example, in the case of $a = \sqrt{3}, b = 4$ the least favourable spectral density $f_0(\lambda)$ and the minimax spectral characteristic are calculated by the formulas

$$f_0(\lambda) = 4/P|1 + \sqrt{3}e^{i\lambda}|^2,$$

$$h(f_0) = -\sqrt{3}e^{2i\lambda}.$$

6.4. Least favourable spectral densities in the class D_M

Consider the problem of the optimal estimation of the functional $A_N \xi = \sum_{j=1}^N a(j) \xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ in the case where the spectral density is from the set of spectral densities with restrictions on the moments of the function $f^{-1}(\lambda)$. Let

$$D_M = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{-1}(\lambda) \cos(m\lambda) d\lambda = r_m, m = 0, 1, \dots, M \right. \right\},$$

where $r_m, m = 0, 1, \dots, M$ is a strictly positive sequence. There is an infinite number of functions in the class D_M [25] and the function

$$f^{-1}(\lambda) = \sum_{m=-M}^M r_{|m|} e^{im\lambda} > 0, \lambda \in [-\pi, \pi].$$

To find solutions to the constrained optimization problem (161) for the set D_M of admissible spectral densities we use the Lagrange multipliers method and get the equation

$$\left| \sum_{k=0}^N c(k) e^{ik\lambda} \right|^2 = \sum_{m=0}^M \alpha_m \cos(m\lambda) = \left| \sum_{m=0}^M p(m) e^{im\lambda} \right|^2, \tag{167}$$

where $\alpha_m, m = 0, 1, \dots, M$ are the Lagrange multipliers and $c(k), k = 0, \dots, N$ are solutions to the equation $B_N^0 \vec{c}_N = \vec{a}_N$.

Consider two cases: $M \geq N$ and $M < N$. Let $M \geq N$. In this case the given Fourier coefficients r_m define the matrix B_N^0 and the optimization problem (160) is degenerate. If we take $p(N+1) = \dots = p(M) = 0$ and components $(p(0), \dots, p(N))$ of the vector \vec{p}_N find from the equation $B_N^0 \vec{p}_N = \vec{a}_N$ then the relation (167) holds true. Thus the least favorable is every density $f(\lambda) \in D_M$ and the density of the autoregression stochastic sequence

$$f_0(\lambda) = 1 / \sum_{m=-M}^M r_{|m|} e^{im\lambda} = 1 / \left| \sum_{k=0}^M \gamma_k e^{ik\lambda} \right| \tag{168}$$

is least favourable, too.

Let $M < N$. Then the matrix B_N is determined by the known $r_m, m = 0, 1, \dots, M$ and the unknown $r_m, m = M + 1, \dots, N$, Fourier coefficients of the function $f^{-1}(\lambda)$.

The unknown coefficients $p(k), k = 0, 1, \dots, M$, and $r_m, m = M + 1, \dots, N$, can be found from the equation $B_N \vec{p}_N = \vec{a}_N$ with $p_N^0 = (p(0), \dots, p(M), 0, \dots, 0)$, or from the equation $B_N \vec{p}_0^N = \vec{a}_N$ with $p_0^N = (0, \dots, 0, p(M), p(M-1), \dots, p(0))$.

If the sequence $r_m, m = 0, 1, \dots, N$, that is constructed from the strictly positive sequence $r_m, m = 0, 1, \dots, M$, and the calculated coefficients $r_m, m = M + 1, \dots, N$, is also strictly positive, then the least favourable spectral density $f_0(\lambda)$ is determined by the Fourier coefficients $r_m, m = 0, 1, \dots, N$, of the function $f_0^{-1}(\lambda)$

$$f_0(\lambda) = 1 / \sum_{m=-N}^N r_{|m|} e^{im\lambda} = 1 / \left| \sum_{k=0}^N \gamma_k e^{ik\lambda} \right| \tag{169}$$

Let us summarize our results and present them in the form of a theorem.

Theorem 6.3

The least favourable spectral density in the class D_M for the optimal linear estimation of the functional $A_N \xi$ in the case where $M \geq N$ is the spectral density (168) of the autoregression stochastic sequence of order M determined by coefficients $r_m, m = 0, 1, \dots, M$. If $M < N$ and solutions $r_m, m = M + 1, \dots, N$, to the equation $B_N \vec{p}_N^0 =$

\vec{a}_N with $p_N^0 = (p(0), \dots, p(M), 0, \dots, 0)$, or to the equation $B_N \vec{p}_0^N = \vec{a}_N$ with $p_0^N = (0, \dots, 0, p(M), p(M-1), \dots, p(0))$ together with coefficients $r_m, m = 0, 1, \dots, M$, form a strictly positive sequence, the least favourable spectral density in D_M is the density (169) of the autoregression stochastic sequence of the order N . The minimax characteristic of the estimate is calculated by formula (156).

Similar statement holds true for the set of spectral densities

$$D_{M, \mathcal{R}} = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{-1}(\lambda) \cos(w\lambda) d\lambda = r_m, m = 0, 1, \dots, M; \vec{r}_M \in \mathcal{R} \right. \right\},$$

where \mathcal{R} is a convex compact which have a strictly positive sequence as an interior point.

6.5. Least favourable spectral densities in the class D_v^{u-}

Consider the problem of the optimal estimation of the functional $A_N \xi = \sum_{j=1}^N a(j) \xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ in the case where the spectral density is from the set of spectral densities

$$D_v^{u-} = \left\{ f(\lambda) \left| 0 \leq v(\lambda) \leq f(\lambda) \leq u(\lambda), \frac{1}{2\pi} \int_{-\pi}^{\pi} f^{-1}(\lambda) d\lambda = P \right. \right\},$$

where $v(\lambda), u(\lambda)$ are given bounded spectral densities. Let the sequence $\{a(0), a(1), \dots, a(N)\}$ (or the sequence $\{a(N), a(N-1), \dots, a(0)\}$) be strictly positive. To find solutions to the constrained optimization problem (161) for the set D_v^{u-} of admissible spectral densities we use the condition $0 \in \partial \Delta_D(f_0)$. It follows from the condition $0 \in \partial \Delta_D(f_0)$ for $D = D_v^{u-}$ that the Fourier coefficients of the function f_0^{-1} satisfy both equation

$$B_N^0 \vec{c}_N = \vec{a}_N$$

and the equation

$$\left| \sum_{k=0}^N \left((B_N^0)^{-1} \vec{a}_N \right)_k e^{ik\lambda} \right|^2 = \psi_1(\lambda) + \psi_2(\lambda) + p_0^{-2},$$

where $\psi_1(\lambda) \geq 0$ and $\psi_1(\lambda) = 0$ if $f_0(\lambda) \geq v(\lambda)$; $\psi_2(\lambda) \leq 0$ and $\psi_2(\lambda) = 0$ if $f_0(\lambda) \leq u(\lambda)$.

Therefore, in the case where $v(\lambda) \leq f_0(\lambda) \leq u(\lambda)$, the function $f_0^{-1}(\lambda)$ is of the form

$$f_0^{-1}(\lambda) = \sum_{k=-N}^N r_{|k|} e^{ik\lambda} = \left| \sum_{k=0}^N \gamma_k e^{ik\lambda} \right|^2,$$

with $r_k = Pa(k)a^{-1}(0)$, in the case where the sequence $\{a(0), a(1), \dots, a(N)\}$ is strictly positive, and with $r_k = r_{-k} = Pa(N-k)a^{-1}(N)$, in the case where the sequence $\{a(N), a(N-1), \dots, a(0)\}$ is strictly positive.

The least favourable in the class $D = D_v^{u-}$ is the density of the autoregression stochastic sequence of the order N if the following inequality holds true

$$v^{-1}(\lambda) \geq \sum_{k=0}^N (r_k e^{ik\lambda} + r_{-k} e^{-ik\lambda}) = \left| \sum_{k=0}^N \gamma_k e^{ik\lambda} \right|^2 \geq u^{-1}(\lambda), \lambda \in [-\pi, \pi]. \tag{170}$$

In general case the least favourable density is of the form

$$f_0(\lambda) = \max \left\{ v(\lambda), \min \left\{ u(\lambda), \left| p_0 \sum_{k=0}^N \left((B_N^0)^{-1} \vec{a}_N \right)_k e^{ik\lambda} \right|^{-2} \right\} \right\}. \tag{171}$$

The following theorem holds true.

Theorem 6.4

If the sequence $\{a(0), a(1), \dots, a(N)\}$ is strictly positive and coefficients $r_k = r_{-k} = Pa(k)a^{-1}(0)$, $k = 0, \dots, N$, satisfy the inequality (170) (or the sequence $\{a(N), a(N-1), \dots, a(0)\}$ is strictly positive and coefficients $r_k = r_{-k} = Pa(N-k)a^{-1}(N)$ satisfy the inequality (170)), then the least favourable in the class D_v^{u-} spectral density for the optimal linear estimate of the functional $A_N \xi$ is density of the autoregression stochastic sequence (164) of order N . The minimax characteristic $h(f_0)$ of the estimate can be calculated by the formula (165). If the inequality (170) is not satisfied, then the least favourable spectral density in D_v^{u-} is determined by relation (171) and the extremum condition (160). The minimax characteristic of the estimate is calculated by formula (156).

6.6. Conclusions

In this section we propose methods of solution of the problem of the mean-square optimal linear estimation of the functional $A_N \xi = \sum_{j=0}^N a(j)\xi(j)$ which depends on the unknown values of the stationary stochastic sequence $\xi(j)$.

Estimates are based on observations of the sequence $\xi(j)$ at points $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$. We provide formulas for calculating the values of the mean square error and the spectral characteristic of the optimal linear estimate of the functional in the case of spectral certainty, where the spectral density of the sequence $\xi(j)$ is exactly known. In the case of spectral uncertainty, where the spectral density is unknown, but a set of admissible spectral densities is given, the minimax approach is applied to estimation of the functional. We obtain formulas that determine the least favourable spectral densities and the minimax spectral characteristics of the optimal linear estimates of the functional $A_N \xi$ for concrete classes of admissible spectral densities. It is shown that spectral densities the autoregressive stochastic sequences are the least favourable in some classes of spectral densities.

The minimax-robust approach to the problem of estimation of one missed value of the stationary stochastic sequences based on convex optimization methods was initiated in papers by Franke [9, 10]. See also papers by Hosoya [19], Taniguchi [75], and survey by Kassam and Poor [23].

For the relative results on the mean-square optimal linear interpolation of linear functionals for stationary stochastic sequences and processes see papers by Moklyachuk [50] – [57], book by Moklyachuk and Masyutka [64], papers by Moklyachuk and Sidei [67], Moklyachuk and Ostapenko [65].

7. Interpolation problem for stationary sequences from observations with noise

In this section we consider the problem of the mean-square optimal estimation of the linear functional $A_N \xi = \sum_{j=0}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$, $j \in \mathbb{Z}$, from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$. The problem is investigated in the case of spectral certainty, where the spectral densities of the stationary stochastic sequences $\xi(j)$ and $\eta(j)$ are exactly known. In this case the classical Hilbert space projection method of linear estimation of the functional is applied. Formulas are derived for calculation the value of the mean square error and the spectral characteristic of the mean-square optimal estimate of the linear functional. In in the case of spectral uncertainty, where the spectral densities of the stationary stochastic sequences $\xi(j)$ and $\eta(j)$ are not exactly known, but classes of admissible spectral densities are given, the minimax-robust procedure to linear estimation of the functional is applied. Relations which determine the least favourable spectral densities and the minimax spectral characteristics are proposed for some special sets of admissible spectral densities.

7.1. The classical Hilbert space projection method of linear interpolation

Let $\xi(j)$, $j \in \mathbb{Z}$, and $\eta(j)$, $j \in \mathbb{Z}$, be (wide sense) stationary stochastic sequences with zero mathematical expectations $E\xi(j) = 0$, $E\eta(j) = 0$. The correlation functions $R_\xi(k) = E\xi(j+k)\xi(j)$ and $R_\eta(k) = E\eta(j+k)\eta(j)$

$k)\overline{\eta(j)}$ of the stationary stochastic sequences $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, admit the spectral representations [12]

$$R_\xi(k) = \int_{-\pi}^{\pi} e^{ik\lambda} F(d\lambda), \quad R_\eta(k) = \int_{-\pi}^{\pi} e^{ik\lambda} G(d\lambda),$$

where $F(d\lambda)$ and $G(d\lambda)$ are the spectral measures of the sequences. We will consider stationary stochastic sequences with absolutely continuous spectral measures $F(d\lambda)$ and $G(d\lambda)$ and the correlation functions of the form

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

where $f(\lambda)$ and $g(\lambda)$ are the spectral density functions of the sequences $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, correspondingly.

We will suppose that the spectral density functions $f(\lambda)$ and $g(\lambda)$ satisfy the minimality condition

$$\int_{-\pi}^{\pi} \frac{1}{f(\lambda) + g(\lambda)} d\lambda < \infty. \tag{172}$$

Under this condition the error-free interpolation of the unknown values of the sequence $\xi(j) + \eta(j)$ is impossible [73].

The stationary stochastic sequences $\xi(j)$ and $\eta(j)$ admit the spectral representations [12, 22]

$$\xi(j) = \int_{-\pi}^{\pi} e^{ij\lambda} dZ_\xi(\lambda), \quad \eta(j) = \int_{-\pi}^{\pi} e^{ij\lambda} dZ_\eta(\lambda),$$

where $Z_\xi(d\lambda)$ and $Z_\eta(d\lambda)$ are orthogonal stochastic measure of the sequences $\xi(j)$ and $\eta(j)$ such that

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

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

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

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

which depends on the unknown values of a stationary stochastic sequence $\xi(j), j \in \mathbb{Z}$, from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$, where $\eta(j), j \in \mathbb{Z}$, is uncorrelated with $\xi(j), j \in \mathbb{Z}$, stationary stochastic sequence.

It follows from the spectral decomposition of the sequence $\xi(j)$ that we can represent the functional $A_N \xi$ in the form

$$A_N \xi = \int_{-\pi}^{\pi} A_N(e^{i\lambda}) Z_\xi(d\lambda), \tag{173}$$

where

$$A_N(e^{i\lambda}) = \sum_{j=0}^N a(j) e^{ij\lambda}.$$

Denote by $\hat{A}_N\xi$ the mean square optimal linear estimate of the functional $A_N\xi$ from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$. Denote by $\Delta(f, g) = E \left| A_N\xi - \hat{A}_N\xi \right|^2$ the mean square error of the estimate $\hat{A}_N\xi$. To find the estimate $\hat{A}_N\xi$ we will use the Hilbert space projection method proposed by A. N. Kolmogorov [24]. We will consider random values $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, as elements of the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ of complex valued random variables with zero first moment, $E\xi = 0$, finite second moment, $E|\xi|^2 < \infty$, and the inner product $(\xi, \eta) = E\xi\bar{\eta}$.

Denote by $H^N(\xi + \eta)$ the subspace of the Hilbert space $H = L_2(\Omega, \mathcal{F}, P)$ generated by elements $\{\xi(j) + \eta(j) : j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}\}$. Denote by $L_2(f + g)$ be the Hilbert space of complex-valued functions that are square-integrable with respect to the measure whose density is $f(\lambda) + g(\lambda)$. Denote by $L_2^N(f + g)$ the subspace of $L_2(f + g)$ generated by functions $\{e^{ij\lambda}, j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}\}$.

The mean square optimal linear estimate $\hat{A}_N\xi$ of the functional $A_N\xi$ from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ is an element of the $H^N(\xi + \eta)$. It can be represented in the form

$$\hat{A}_N\xi = \int_{-\pi}^{\pi} h(e^{i\lambda})(Z_{\xi}(d\lambda) + Z_{\eta}(d\lambda)), \tag{174}$$

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

The mean square error $\Delta(h; f, g)$ of the estimate $\hat{A}_N\xi$ is given by the formula

$$\Delta(h; f, g) = E \left| A_N\xi - \hat{A}_N\xi \right|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |A_N(e^{i\lambda}) - h(e^{i\lambda})|^2 f(\lambda)d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} |h(e^{i\lambda})|^2 g(\lambda)d\lambda.$$

The Hilbert space projection method proposed by A. N. Kolmogorov [24] makes it possible to find the spectral characteristic $h(e^{i\lambda})$ and the mean square error $\Delta(h; f, g)$ of the optimal linear estimate of the functional $A_N\xi$ in the case where the spectral densities $f(\lambda)$ and $g(\lambda)$ of the sequences $\xi(j), j \in \mathbb{Z}$, and $\eta(j), j \in \mathbb{Z}$, are exactly known and the minimality condition (172) is satisfied. The spectral characteristic can be found from the following conditions:

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

It follows from the second condition that for any $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ the following equations should be satisfied

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

The last equations are equivalent to equations

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [A_N(e^{i\lambda})f(\lambda) - h(e^{i\lambda})(f(\lambda) + g(\lambda))] e^{-ij\lambda} d\lambda = 0, \quad j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}.$$

It follows from these equations that the function $[A_N(e^{i\lambda})f(\lambda) - h(e^{i\lambda})(f(\lambda) + g(\lambda))]$ is of the form

$$A_N(e^{i\lambda})f(\lambda) - h(e^{i\lambda})(f(\lambda) + g(\lambda)) = C_N(e^{i\lambda}), \tag{175}$$

$$C_N(e^{i\lambda}) = \sum_{j=0}^N c(j)e^{ij\lambda},$$

where $c(j), j = 0, 1, \dots, N$, are unknown coefficients that we have to find.

From the relation (175) we deduce that the spectral characteristic $h(e^{i\lambda})$ of the optimal linear estimate of the functional $A_N \xi$ is of the form

$$\begin{aligned} h(e^{i\lambda}) &= \frac{A_N(e^{i\lambda})f(\lambda) - C_N(e^{i\lambda})}{f(\lambda) + g(\lambda)} = \\ &= A_N(e^{i\lambda}) - \frac{A_N(e^{i\lambda})g(\lambda) + C_N(e^{i\lambda})}{f(\lambda) + g(\lambda)}. \end{aligned} \tag{176}$$

It follows from the first condition, which determines the spectral characteristic $h(e^{i\lambda}) \in L_2^N(f + g)$ of the optimal linear estimate of the functional $A_N \xi$, that the Fourier coefficients of the function $h(e^{i\lambda})$ are equal to zero for $j = 0, 1, \dots, N$, namely

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

Using the last relations and (176) we get the following system of equations

$$\int_{-\pi}^{\pi} \left(A_N(e^{i\lambda}) \frac{f(\lambda)}{f(\lambda) + g(\lambda)} - \frac{C_N(e^{i\lambda})}{f(\lambda) + g(\lambda)} \right) e^{-ij\lambda} d\lambda = 0, \quad j = 0, 1, \dots, N.$$

The last equations can be written in the form

$$\sum_{k=0}^N a(k) \int_{-\pi}^{\pi} \frac{e^{i(k-j)\lambda} f(\lambda)}{f(\lambda) + g(\lambda)} d\lambda - \sum_{k=0}^N c(k) \int_{-\pi}^{\pi} \frac{e^{i(k-j)\lambda}}{f(\lambda) + g(\lambda)} d\lambda = 0, \quad j = 0, 1, \dots, N. \tag{177}$$

Let us introduce the following notations

$$\begin{aligned} R_{j,k}^N &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)\lambda} \frac{f(\lambda)}{f(\lambda) + g(\lambda)} d\lambda, \\ B_{j,k}^N &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)\lambda} \frac{1}{f(\lambda) + g(\lambda)} d\lambda, \\ Q_{j,k}^N &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)\lambda} \frac{f(\lambda)g(\lambda)}{f(\lambda) + g(\lambda)} d\lambda. \end{aligned}$$

Making use the introduced notations we can write equations (177) in the form

$$\sum_{k=0}^N R_{j,k}^N a(k) = \sum_{k=0}^N B_{j,k}^N c(k), \quad j = 0, 1, \dots, N.$$

The derived equations can be written in the matrix form

$$\mathbf{R}_N \mathbf{a}_N = \mathbf{B}_N \mathbf{c}_N,$$

where $\mathbf{a}_N = (a(0), a(1), \dots, a(N))$ is a vector constructed from the coefficients that determine the functional $A_N \xi$, $\mathbf{c}_N = (c(0), c(1), \dots, c(N))$ is a vector constructed from the unknown coefficients $c(k), k = 0, 1, \dots, N$,

\mathbf{B}_N and \mathbf{R}_N are linear operators in \mathbb{C}^{N+1} , which are determined by matrices with elements $(\mathbf{B}_N)_{j,k} = B_{j,k}^N$, $(\mathbf{R}_N)_{j,k} = R_{j,k}^N$, $j, k = 0, 1, \dots, N$.

Since the matrix \mathbf{B}_N is reversible [74], we get the formula

$$\mathbf{c}_N = \mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N, \quad (178)$$

Hence, the unknown coefficients $c(j)$, $j = 0, 1, \dots, N$, are calculated by the formula

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

where $(\mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N)_j$ is the j -th component of the vector $\mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N$, and the formula for calculating the spectral characteristic of the estimate $\hat{A}_N \xi$ is of the form

$$h(e^{i\lambda}) = A_N(e^{i\lambda}) \frac{f(\lambda)}{f(\lambda) + g(\lambda)} - \frac{\sum_{k=0}^N (\mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N)_k e^{ik\lambda}}{f(\lambda) + g(\lambda)}. \quad (179)$$

The mean square error of the estimate of the function can be calculated by the formula

$$\begin{aligned} \Delta(h; f, g) &= E \left| A_N \xi - \hat{A}_N \xi \right|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A_N(e^{i\lambda})g(\lambda) + \sum_{k=0}^N (\mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N)_k e^{ik\lambda} \right|^2}{(f(\lambda) + g(\lambda))^2} f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left| A_N(e^{i\lambda})f(\lambda) - \sum_{k=0}^N (\mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N)_k e^{ik\lambda} \right|^2}{(f(\lambda) + g(\lambda))^2} g(\lambda) d\lambda \\ &= \langle \mathbf{R}_N \mathbf{a}_N, \mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N \rangle + \langle \mathbf{Q}_N \mathbf{a}_N, \mathbf{a}_N \rangle, \end{aligned} \quad (180)$$

where \mathbf{Q}_N is a linear operator in \mathbb{C}^{N+1} , which is determined by matrix with elements $(\mathbf{Q}_N)_{j,k} = Q_{j,k}^N$, $j, k = 0, 1, \dots, N$.

Let us summarize our results and present them in the form of a theorem.

Theorem 7.1

Let $\xi(j)$ and $\eta(j)$ be uncorrelated stationary stochastic sequences with the spectral densities $f(\lambda)$ and $g(\lambda)$ that satisfy the minimality condition (172). The spectral characteristic $h(e^{i\lambda})$ and the mean square error $\Delta(h; f, g)$ of the optimal linear estimate $\hat{A}_N \xi$ of the functional $A_N \xi$ based on observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ can be calculated by formulas (179), (180).

7.2. Minimax-robust method of interpolation

The traditional methods of estimation of the functional $A_N \xi$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ can be applied in the case where the spectral densities $f(\lambda)$ and $g(\lambda)$ of the considered stochastic sequences $\xi(j)$ and $\eta(j)$ are exactly known. In practise, however, we do not have complete information on spectral densities of the sequences. For this reason we apply the minimax(robust) method of estimation of the functional $A_N \xi$, that is we find an estimate that minimizes the maximum of the mean square errors for all spectral densities from the given class of admissible spectral densities D .

Definition 7.1. For a given class of spectral densities $D = D_f \times D_g$ the spectral densities $f_0(\lambda) \in D_f$, $g_0(\lambda) \in D_g$ are called the least favourable in D for the optimal linear estimation of the functional $A_N \xi$ 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 7.2. 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 $A_N \xi$ is called minimax-robust if

$$h^0(e^{i\lambda}) \in H_D = \bigcap_{(f,g) \in D_f \times D_g} L_2^N(f+g),$$

$$\min_{h \in H_D} \max_{(f,g) \in D} \Delta(h; f, g) = \sup_{(f,g) \in D} \Delta(h^0; f, g).$$

It follows from the introduced definitions and the obtained formulas that the following statement holds true.

Lemma 7.1

The spectral densities $f_0(\lambda) \in D_f, g_0(\lambda) \in D_g$ are the least favourable in the class of admissible spectral densities $D = D_f \times D_g$ for the optimal linear estimate of the functional $A_N \xi$ if the Fourier coefficients of the 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 $\mathbf{B}_N^0, \mathbf{R}_N^0, \mathbf{Q}_N^0$ that determine a solution to the optimization problem

$$\begin{aligned} \max_{(f,g) \in D_f \times D_g} & \langle \mathbf{R}_N \mathbf{a}_N, \mathbf{B}_N^{-1} \mathbf{R}_N \mathbf{a}_N \rangle + \langle \mathbf{Q}_N \mathbf{a}_N, \mathbf{a}_N \rangle \\ & = \langle \mathbf{R}_N^0 \mathbf{a}_N, (\mathbf{B}_N^0)^{-1} \mathbf{R}_N^0 \mathbf{a}_N \rangle + \langle \mathbf{Q}_N^0 \mathbf{a}_N, \mathbf{a}_N \rangle. \end{aligned} \tag{181}$$

The minimax spectral characteristic $h^0 = h(f_0, g_0)$ can be calculated by the formula (179) if $h(f_0, g_0) \in H_D$.

The least favourable 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

$$\Delta(h; f_0, g_0) \geq \Delta(h^0; f_0, g_0) \geq \Delta(h^0; f, g)$$

$$\forall h \in H_D, \forall f \in D_f, \forall g \in D_g$$

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

$$\sup_{(f,g) \in D_f \times D_g} \Delta(h(f_0, g_0); f, g) = \Delta(h(f_0, g_0); f_0, g_0), \tag{182}$$

where

$$\begin{aligned} \Delta(h(f_0, g_0); f, g) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|A_N(e^{i\lambda})g_0(\lambda) + C_N^0(e^{i\lambda})|^2}{(f_0(\lambda) + g_0(\lambda))^2} f(\lambda) d\lambda \\ &+ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|A_N(e^{i\lambda})f_0(\lambda) - C_N^0(e^{i\lambda})|^2}{(f_0(\lambda) + g_0(\lambda))^2} g(\lambda) d\lambda, \\ C_N^0(e^{i\lambda}) &= \sum_{j=0}^N ((\mathbf{B}_N^0)^{-1} \mathbf{R}_N^0 \mathbf{a}_N)_j e^{ij\lambda}, \end{aligned}$$

The constrained optimization problem (182) is equivalent to the unconstrained optimization problem

$$\Delta_D(f, g) = -\Delta(h(f_0, g_0); f, g) + \delta(f, g | D_f \times D_g) \rightarrow \inf, \tag{183}$$

where $\delta(f, g | D_f \times D_g)$ is the indicator function of the set $D = D_f \times D_g$. Solution (f_0, g_0) to the problem (183) is characterized by the condition $0 \in \partial \Delta_D(f_0, g_0)$, where $\partial \Delta_D(f_0, g_0)$ is the subdifferential of the convex functional

$\Delta_D(f, g)$ at point (f_0, g_0) . This condition makes it possible to find the least favourable spectral densities in some special classes of spectral densities D [20], [70], [72].

Note, that the form of the functional $\Delta(h(f_0, g_0); f, g)$ is convenient for application the Lagrange method of indefinite multipliers for finding solution to the problem (182). Making use the method of Lagrange multipliers and the form of subdifferentials of the indicator functions we describe relations that determine least favourable spectral densities in some special classes of spectral densities (see books [13, 58, 64] for additional details).

Lemma 7.2

Let (f_0, g_0) be a solution to the optimization problem (183). The spectral densities $f_0(\lambda), g_0(\lambda)$ are the least favourable in the class $D = D_f \times D_g$, and the spectral characteristic $h^0 = h(f_0, g_0)$ is minimax for the optimal estimate of the functional $A_N \xi$ if $h(f_0, g_0) \in H_D$.

7.3. Least favourable spectral densities in the class $D_f^0 \times D_g^0$

Consider the problem of the optimal estimation of the functional $A_N \xi = \sum_{j=1}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ in the case where the spectral densities $f(\lambda), g(\lambda)$ are from the class $D = D_f^0 \times D_g^0$, where

$$D_f^0 = \left\{ f(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) d\lambda \leq P_1 \right. \right\},$$

$$D_g^0 = \left\{ g(\lambda) \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) d\lambda \leq P_2 \right. \right\}.$$

Let the densities $f_0(\lambda) \in D_f^0, g_0(\lambda) \in D_g^0$ and the functions $h_f(f_0, g_0), h_g(f_0, g_0)$, which are determined by the relations

$$h_f(f_0, g_0) = \frac{|A_N(e^{i\lambda})g_0(\lambda) + C_N^0(e^{i\lambda})|^2}{(f_0(\lambda) + g_0(\lambda))^2}, \tag{184}$$

$$h_g(g_0, g_0) = \frac{|A_N(e^{i\lambda})f_0(\lambda) - C_N^0(e^{i\lambda})|^2}{(f_0(\lambda) + g_0(\lambda))^2}, \tag{185}$$

be bounded. In this case the 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 linear and continuous on the space $L_1 \times L_1$ and we can apply the method of Lagrange multipliers to find solution to the optimization problem (183). We get the following relations that determine the least favourable spectral densities $f^0 \in D_f^0, g^0 \in D_g^0$

$$-\frac{1}{2\pi} \int_{-\pi}^{\pi} h_f(f_0, g_0) \rho(f(\lambda)) d\lambda - \frac{1}{2\pi} \int_{-\pi}^{\pi} h_g(f_0, g_0) \rho(g(\lambda)) d\lambda$$

$$+ \alpha_1 \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho(f(\lambda)) d\lambda + \alpha_2 \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho(g(\lambda)) d\lambda = 0,$$

where $\rho(f(\lambda))$ and $\rho(g(\lambda))$ are variations of the functions $f(\lambda)$ and $g(\lambda)$, the constants $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$. From this relation we get that the least favourable spectral densities $f_0(\lambda) \in D_f^0, g_0(\lambda) \in D_g^0$ satisfy equations

$$|A_N(e^{i\lambda})g_0(\lambda) + C_N^0(e^{i\lambda})| = \alpha_1 (f_0(\lambda) + g_0(\lambda)), \tag{186}$$

$$|A_N(e^{i\lambda})f_0(\lambda) - C_N^0(e^{i\lambda})| = \alpha_2(f_0(\lambda) + g_0(\lambda)). \tag{187}$$

Note, that $\alpha_1 \neq 0$ in the case where

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_0(\lambda) d\lambda = P_1,$$

and $\alpha_2 \neq 0$ in the case where

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} g_0(\lambda) d\lambda = P_2.$$

Summing up our reasoning we come to conclusion that the following theorem holds true.

Theorem 7.2

Let the spectral densities $f_0(\lambda) \in D_f^0$ and $g_0(\lambda) \in D_g^0$ satisfy the minimality condition (172) and let the functions $h_f(f_0, g_0)$ and $h_g(f_0, g_0)$, determined by the formulas (184), (185), be bounded. The functions $f_0(\lambda), g_0(\lambda)$, which give solution to the system of equations (186), (187) are the least favourable spectral densities in the class $D = D_f^0 \times D_g^0$, if they determine a solution to the optimization problem (181). The function $h^0(e^{i\lambda})$, determined by the formula (176), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N\xi$.

Theorem 7.3

Let the spectral density $f(\lambda)$ be known and fixed, and let the spectral density $g_0(\lambda) \in D_g^0$. Let the functions $f(\lambda)$ and $g_0(\lambda)$ be such that the function $(f(\lambda) + g_0(\lambda))^{-1}$ is integrable and let the function $h_g(f, g_0)$, determined by the formula (185), be bounded. The spectral density $g_0(\lambda)$ is the least favourable spectral densities in the class D_g^0 for the optimal linear estimate of the functional $A_N\xi$, if it is of the form

$$g_0(\lambda) = \max \{0, \alpha_2^{-1} |A_N(e^{i\lambda})f(\lambda) - C_N^0(e^{i\lambda})| - f(\lambda)\}$$

and the functions $f(\lambda), g_0(\lambda)$ determine a solution to the optimization problem (181). The function $h^0(e^{i\lambda})$, determined by the formula (176), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N\xi$.

Theorem 7.4

Let the spectral density $f_0(\lambda) \in D_f^0$, let the function $f_0^{-1}(\lambda)$ be integrable, and let the function $h(f_0)$, determined by the formula (150), be bounded. The spectral density $f_0(\lambda)$ is the least favourable spectral densities in the class D_f^0 for the optimal linear estimate of the functional $A_N\xi$, based on observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$, if it satisfies the relation

$$f_0(\lambda) = \alpha_1 |C_N^0(e^{i\lambda})|$$

and $f_0(\lambda)$ determine a solution to the optimization problem (160). The function $h^0(e^{i\lambda})$, determined by the formula (150), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N\xi$.

7.4. Least favourable spectral densities in the class $D_v^u \times D_\varepsilon$

Consider the problem of the optimal estimation of the functional $A_N\xi = \sum_{j=1}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ in the case where the spectral densities $f(\lambda), g(\lambda)$ are from the class $D = D_v^u \times D_\varepsilon$, where

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

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

Here the spectral densities $v(\lambda), u(\lambda), g_1(\lambda)$ are known and fixed and the densities $v(\lambda)$ and $u(\lambda)$ are bounded.

Let the densities $f^0(\lambda) \in D_v^u, g^0(\lambda) \in D_\varepsilon$ determine the bounded functions $h_f(f_0, g_0), h_g(f_0, g_0)$ with the help of formulas (184), (185). Then from the condition $0 \in \partial \Delta_{D_{f,g}}(f_0, g_0)$ we derive the following equations that determine the least favourable spectral densities

$$|A_N(e^{i\lambda})g_0(\lambda) + C_N^0(e^{i\lambda})| = (f_0(\lambda) + g_0(\lambda))(\gamma_1(\lambda) + \gamma_2(\lambda) + \alpha_1^{-1}), \quad (188)$$

$$|A_N(e^{i\lambda})f_0(\lambda) - C_N^0(e^{i\lambda})| = (f_0(\lambda) + g_0(\lambda))(\varphi(\lambda) + \alpha_2^{-1}). \quad (189)$$

Here $\gamma_1 \leq 0$ and $\gamma_1 = 0$ in the case $f_0(\lambda) \geq v(\lambda)$; $\gamma_2 \geq 0$ and $\gamma_2 = 0$ in the case $f_0(\lambda) \leq u(\lambda)$; $\varphi(\lambda) \leq 0$ and $\varphi(\lambda) = 0$ in the case $g_0(\lambda) \geq (1 - \varepsilon)g_1(\lambda)$.

The following theorems hold true.

Theorem 7.5

Let the spectral densities $f_0(\lambda) \in D_v^u, g_0(\lambda) \in D_\varepsilon$ satisfy the minimality condition (172) and let the functions $h_f(f_0, g_0)$ and $h_g(f_0, g_0)$, determined by the formulas (184), (185), be bounded. The functions $f_0(\lambda), g_0(\lambda)$, which give solution to the system of equations (188), (189) are the least favourable spectral densities in the class $D_v^u \times D_\varepsilon$, if they determine a solution to the optimization problem (181). The function $h^0(e^{i\lambda})$, determined by the formula (176), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N \xi$.

Theorem 7.6

Let the spectral density $f(\lambda)$ be known, and let the spectral density $g_0(\lambda) \in D_\varepsilon$. Let the function $f(\lambda) + g_0(\lambda)$ satisfy the minimality condition (172), and let the function $h_g(f, g_0)$, determined by the formula (185), be bounded. The spectral density $g_0(\lambda)$ is the least favourable spectral densities in the class D_ε for the optimal linear estimate of the functional $A_N \xi$, if it is of the form

$$g_0(\lambda) = \max \{ (1 - \varepsilon)g_1(\lambda), \alpha_2 |A_N(e^{i\lambda})f(\lambda) - C_N^0(e^{i\lambda})| - f(\lambda) \}$$

and the functions $f(\lambda), g_0(\lambda)$ determine a solution to the optimization problem (181). The function $h^0(e^{i\lambda})$, determined by the formula (176), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N \xi$.

Theorem 7.7

Let the spectral density $f_0(\lambda) \in D_v^u$, let the function $f_0^{-1}(\lambda)$ be integrable, and let the function $h(f_0)$, determined by the formula (150), be bounded. The spectral density $f_0(\lambda)$ is the least favourable spectral densities in the class D_v^u for the optimal linear estimate of the functional $A_N \xi$, based on observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$, if it satisfies the relation

$$f_0(\lambda) = \max \{ v(\lambda), \min \{ u(\lambda), \alpha_1 |C_N^0(e^{i\lambda})| \} \}$$

and $f_0(\lambda)$ determine a solution to the optimization problem (160). The function $h^0(e^{i\lambda})$, determined by the formula (150), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N \xi$.

7.5. Least favourable spectral densities in the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$

Consider the problem of the optimal estimation of the functional $A_N \xi = \sum_{j=1}^N a(j)\xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$ from observations of the sequence $\xi(j) + \eta(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$ in the case where the spectral densities $f(\lambda), g(\lambda)$ are from the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$, which describe the models of “ ε -neighbourhood” of spectral densities in the space $L_2 \times L_1$. Let

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

be “ ε -neighbourhood” in the space L_2 of a given bounded spectral density $f_1(\lambda)$, and let

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

be “ ε -neighbourhood in the space L_1 of a given bounded spectral density $g_1(\lambda)$.”

Let the spectral densities $f_0(\lambda) \in D_{2\varepsilon_1}$, $g_0(\lambda) \in D_{1\varepsilon_2}$ determine the bounded functions $h_f(f_0, g_0)$, $h_g(f_0, g_0)$ with the help of formulas (184), (185). Then from the condition $0 \in \partial\Delta_{D_{f,g}}(f_0, g_0)$ for $D = D_{2\varepsilon_1} \times D_{1\varepsilon_2}$ we derive the following equations that determine the least favourable spectral densities

$$|A_N(e^{i\lambda})g_0(\lambda) + C_N^0(e^{i\lambda})|^2 = (f_0(\lambda) + g_0(\lambda))^2(f_0(\lambda) - f_1(\lambda))\alpha_1, \tag{190}$$

$$|A_N(e^{i\lambda})f_0(\lambda) - C_N^0(e^{i\lambda})|^2 = (f_0(\lambda) + g_0(\lambda))^2\Psi(\lambda)\alpha_2, \tag{191}$$

where $|\Psi(\lambda)| \leq 1$ and $\Psi(\lambda) = \text{sign}(g_0(\lambda) - g_1(\lambda))$, in the case $g_0(\lambda) \neq g_1(\lambda)$, α_1, α_2 are constants.

Equations (190), (191) with the optimization problem (160) and the normalising conditions

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\lambda) - f_1(\lambda)|^2 d\lambda = \varepsilon_1 \tag{192}$$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |g(\lambda) - g_1(\lambda)| d\lambda = \varepsilon_2 \tag{193}$$

determine the least favourable spectral densities in the class $D = D_{2\varepsilon_1} \times D_{1\varepsilon_2}$.

The following theorems hold true.

Theorem 7.8

Let the spectral densities $f_0(\lambda) \in D_{2\varepsilon_1}$, $g_0(\lambda) \in D_{1\varepsilon_2}$ satisfy the minimality condition (172) and let the functions $h_f(f_0, g_0)$ and $h_g(f_0, g_0)$, determined by the formulas (184), (185), be bounded. The spectral densities $f_0(\lambda), g_0(\lambda)$, which give solution to the system of equations (190)–(193) are the least favourable spectral densities in the class $D_{2\varepsilon_1} \times D_{1\varepsilon_2}$, if they determine a solution to the optimization problem (181). The function $h^0(e^{i\lambda})$, determined by the formula (176), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N\xi$.

Theorem 7.9

Let the spectral density $f(\lambda)$ be known, and let the spectral density $g_0(\lambda) \in D_{1\varepsilon_2}$. Let the function $f(\lambda) + g_0(\lambda)$ satisfy the minimality condition (172), and let the function $h_g(f, g_0)$, determined by the formula (185), be bounded. The spectral density $g_0(\lambda)$ is the least favourable spectral densities in the class $D_{1\varepsilon_2}$ for the optimal linear estimate of the functional $A_N\xi$, if it is of the form

$$g_0(\lambda) = \max \{g_1(\lambda), \alpha_2^{-1} |A_N(e^{i\lambda})f(\lambda) - C_N^0(e^{i\lambda})| - f(\lambda)\}$$

and the functions $f(\lambda), g_0(\lambda)$ determine a solution to the optimization problem (181). The function $h^0(e^{i\lambda})$, determined by the formula (176), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N\xi$.

Theorem 7.10

Let the spectral density $f_0(\lambda) \in D_{2\varepsilon_1}$, let the function $f_0^{-1}(\lambda)$ be integrable, and let the function $h(f_0)$, determined by the formula (150), be bounded. The spectral density $f_0(\lambda)$ is the least favourable spectral densities in the class $D_{2\varepsilon_1}$ for the optimal linear estimate of the functional $A_N\xi$, based on observations of the sequence $\xi(j)$ at points of time $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$, if it satisfies the relation

$$|C_N^0(e^{i\lambda})|^2 = (f_0(\lambda))^2(f_0(\lambda) - f_1(\lambda))\alpha_1$$

and $f_0(\lambda)$ determine a solution to the optimization problem (160). The function $h^0(e^{i\lambda})$, determined by the formula (150), is minimax-robust spectral characteristic of the optimal linear estimate of the functional $A_N\xi$.

7.6. Conclusions

In this section we propose methods of solution of the problem of the mean-square optimal linear estimation of the functional $A_N \xi = \sum_{j=0}^N a(j) \xi(j)$ which depends on the unknown values of a stationary stochastic sequence $\xi(j)$.

Estimates are based on observations of the sequence $\xi(j) + \eta(j)$ at points $j \in \mathbb{Z} \setminus \{0, 1, \dots, N\}$, where $\eta(j)$ is an uncorrelated with $\xi(j)$ stationary sequence. We provide formulas for calculating values of the mean square error and the spectral characteristic of the optimal linear estimate of the functional in the case of spectral certainty where the spectral densities $f(\lambda)$ and $g(\lambda)$ of the sequences $\xi(j)$ and $\eta(j)$ are exactly known. In the case of spectral uncertainty where the spectral densities $f(\lambda)$ and $g(\lambda)$ are not known, but a set of admissible spectral densities is given, the minimax approach is applied. We obtain formulas that determine the least favourable spectral densities and the minimax spectral characteristics of the optimal linear estimates of the functional $A_N \xi$ for concrete classes of admissible spectral densities.

For the relative results on the mean-square optimal linear interpolation of linear functionals for stationary stochastic sequences and processes based on observations with noise see papers by Moklyachuk [55] – [57], book by Moklyachuk and Masyutka [64].

8. Conclusion Remarks

In the proposed paper we describe methods of solution of the problems of the mean-square optimal linear extrapolation and interpolation of linear functionals which depend on the unknown values of a stationary stochastic sequence $\xi(k)$ based on observations of the sequence $\xi(k)$ as well as observations of the sequence $\xi(k) + \eta(k)$, where $\eta(k)$ is an uncorrelated with the sequence $\xi(k)$ stationary stochastic sequence. The corresponding methods of solution of the problem of the mean-square optimal linear filtering of stationary stochastic sequences are described in the paper by Luz and Moklyachuk [33].

Following the Ulf Grenander [15] approach to investigation the problem of optimal linear estimation of the functional which depends on the unknown values of the stationary stochastic continuous parameter process we consider the problem as a two-person zero-sum game in which the first player chooses a stationary stochastic sequence $\xi(j)$ from the class Ξ of stationary stochastic sequences with $E\xi(j) = 0$ and $E|\xi(j)|^2 = 1$ which maximizes the value of the mean square error of estimate. The second player is looking for an estimate of the linear functional which minimizes the value of the mean square error. It is show that this game has equilibrium point. The maximum error gives a one-sided moving average stationary sequence which is least favourable in the given class of stationary sequences. The greatest value of the error and the least favourable sequence are determined by the largest eigenvalue and the corresponding eigenvector of the operator determined by coefficients which determine the functionals. Note, that this approach can be applied to a specific class of estimation problems.

The second approach to the estimation problems we applied is based on the Kolmogorov [24] Hilbert space projection method which we apply in the case of spectral certainty, where the spectral densities of the sequences $\xi(n)$ and $\eta(n)$ are exactly known, and the convex optimization method proposed by Franke [9, 10] which is applied in the case of spectral uncertainty, where the spectral densities of the sequences are not exactly known, but, instead, a set of admissible spectral densities is given. Formulas for calculation the mean-square errors and the spectral characteristics of the optimal estimates of functionals are derived in the case of spectral certainty. In the case of spectral uncertainty the minimax-robust estimation method is applied. Formulas that determine the least favourable spectral densities and the minimax-robust spectral characteristics of the optimal linear estimates of the functionals are derived.

In the papers by Moklyachuk [35] – [58] problems of extrapolation, interpolation and filtering for stationary processes and sequences were studied. The corresponding problems for vector-valued stationary sequences and processes were investigated by Moklyachuk and Masyutka [59] – [64]. In the articles by Dubovets'ka and Moklyachuk [3] - [7] and in the book by Golichenko and Moklyachuk [13] the minimax estimation problems were investigated for another generalization of stationary processes – periodically correlated stochastic sequences and stochastic processes. Luz and Moklyachuk [26] – [32], [34] investigated the classical and minimax extrapolation,

interpolation and filtering problems for sequences and processes with n th stationary increments. Investigation of the mean-square optimal linear estimation problems for functionals of stationary stochastic sequences and processes with missing observations is started in the papers by Moklyachuk and Sidei [67], [68]. The minimax estimation problems for functionals of a generalization of stationary processes – harmonizable stable stochastic sequences and processes is started in the papers by Moklyachuk and Ostapenko [65], [66]. For the results for functionals of random fields see the book by Moklyachuk and Shchestyuk [69].

REFERENCES

1. N. I. Akhieser and I. M. Glasmann, *Theorie der linearen Operatoren im Hilbert-Raum*, Berlin: Akademie-Verlag, XVI, 1968.
2. G. E. P. Box, G. M. Jenkins and G. C. Reinsel, *Time series analysis. Forecasting and control. 3rd ed.*, Englewood Cliffs, NJ: Prentice Hall, 1994.
3. I. I. Dubovets'ka, O.Yu. Masyutka and M.P. Moklyachuk, *Interpolation of periodically correlated stochastic sequences*, Theory of Probability and Mathematical Statistics, vol. 84, pp. 43-56, 2012.
4. I. I. Dubovets'ka and M. P. Moklyachuk, *Filtration of linear functionals of periodically correlated sequences*, Theory of Probability and Mathematical Statistics, vol. 86, pp. 51-64, 2013.
5. I. I. Dubovets'ka and M. P. Moklyachuk, *Extrapolation of periodically correlated processes from observations with noise*, Theory of Probability and Mathematical Statistics, vol. 88, pp. 43-55, 2013.
6. I. I. Dubovets'ka and M. P. Moklyachuk, *Minimax estimation problem for periodically correlated stochastic processes*, Journal of Mathematics and System Science, vol. 3, no. 1, pp. 26-30, 2013.
7. I. I. Dubovets'ka and M. P. Moklyachuk, *On minimax estimation problems for periodically correlated stochastic processes*, Contemporary Mathematics and Statistics, vol.2, no. 1, pp. 123-150, 2014.
8. N. Dunford and J. T. Schwartz, *Linear Operators. I. General theory*, New York and London: Interscience Publishers, XIV, 1958.
9. J. Franke, *On the robust prediction and interpolation of time series in the presence of correlated noise*, J. Time Series Analysis, vol. 5, No. 4, pp. 227–244, 1984.
10. J. Franke, *Minimax robust prediction of discrete time series*, Z. Wahrscheinlichkeitstheor. Verw. Gebiete, vol. 68, pp. 337–364, 1985.
11. 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.
12. I. I. Gikhman and A. V. Skorokhod, *The theory of stochastic processes. I.*, Berlin: Springer, 2004.
13. I. I. Golichenko and M. P. Moklyachuk, *Estimates of functionals of periodically correlated processes*, Kyiv: NVP "Interservis", 2014.
14. S. H. Gould, *Variational methods for eigenvalue problems. An introduction to the Weinstein method of intermediate problems*, 2nd ed. Mathematical Expositions. 10. Toronto: University of Toronto Press; London: Oxford University Press XVI, 1966.
15. U. Grenander, *A prediction problem in game theory*, Arkiv för Matematik, vol. 3, pp. 371–379, 1957.
16. U. Grenander and G. Szegő, *Toeplitz forms and their applications*, Berkeley and Los Angeles: University of California Press VIII, 1958.
17. E. J. Hannan, *Multiple time series*, Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley & Sons, Inc. XI, 1970.
18. K. Hoffman, *Banach spaces of analytic functions*, Prentice–Hall Series in Modern Analysis. Englewood Cliffs, N.J.: Prentice–Hall, Inc. XIII, 1962.
19. Y. Hosoya, *Robust linear extrapolations of secondorder stationary processes*, The Annals of Probability, vol. 6, pp. 574-784, 1978.
20. 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.
21. T. Kailath, *A view of three decades of linear filtering theory*, IEEE Transactions on Information Theory, vol. 20, No. 2, pp. 146–181, 1974.
22. K. Karhunen, *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Annales Academiae Scientiarum Fennicae. Ser. A I, vol. 37, 1947.
23. 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.
24. 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.
25. M. G. Krein and A. A. Nudelman, *The Markov moment problem and extremal problems*, Translations of Mathematical Monographs. Vol. 50. Providence, R.I.: American Mathematical Society, 1977.
26. 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.
27. M. M. Luz and M. P. Moklyachuk, *Interpolation of functionals of stochastic sequences with stationary increments for observations with noise*, Prykl. Stat., Aktuarna Finans. Mat., no. 2, pp. 131-148, 2012.
28. 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.
29. M. Luz and M. Moklyachuk, *Robust extrapolation problem for stochastic processes with stationary increments*, Mathematics and Statistics, vol. 2, no. 2, pp. 78 - 88, 2014.

30. M. Luz and M. Moklyachuk, *Minimax-robust filtering problem for stochastic sequences with stationary increments and cointegrated sequences*, cStatistics, Optimization & Information Computing, vol. 2, no. 3, pp. 176 - 199, 2014.
31. M. Luz and M. Moklyachuk, *Minimax interpolation problem for random processes with stationary increments*, Statistics, Optimization & Information Computing, vol. 3, no. 1, pp. 30-41, 2015.
32. 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.
33. M. Luz and M. Moklyachuk, *Filtering problem for functionals of stationary sequences*, Statistics, Optimization & Information Computing (submitted), 2015.
34. M. Moklyachuk and M. Luz, *Robust extrapolation problem for stochastic sequences with stationary increments*, Contemporary Mathematics and Statistics, vol. 1, no. 3, pp. 123 - 150, 2013.
35. M. P. Moklyachuk, *Estimation of linear functionals of stationary stochastic processes and a two-person zero-sum game*, Stanford University Technical Report, No. 169, 1981.
36. M. P. Moklyachuk, *On a problem of game theory and the extrapolation of stochastic processes with values in a Hilbert space*, Theory of Probability and Mathematical Statistics, vol. 24, pp. 107-114, 1981.
37. M. P. Moklyachuk, *On an antagonistic game and prediction of stationary random sequences in a Hilbert space*, Theory of Probability and Mathematical Statistics, vol. 25, pp. 107-113, 1982.
38. M. P. Moklyachuk, *A filtration of transformations of random sequences*, Ukrainian Mathematical Journal, vol.37, no.6, pp. 597-601, 1985.
39. M. P. Moklyachuk, *On a property of one-sided moving average random sequences*. Theory of Probability and Mathematical Statistics, vol. 32, pp. 95-102, 1986.
40. M. P. Moklyachuk, *Minimax extrapolation and autoregressive-moving average processes*, Theory of Probability and Mathematical Statistics, vol. 41, pp. 77-84, 1990.
41. M. P. Moklyachuk, *Minimax extrapolation of random processes for models of ϵ -pollution*, Theory of Probability and Mathematical Statistics, vol. 42, pp. 113-121, 1991.
42. M. P. Moklyachuk, *Minimax filtering of stationary sequences with white noise*, Theory of Probability and Mathematical Statistics, vol. 43, pp. 109-122, 1991.
43. M. P. Moklyachuk, *Minimax filtering of linear transformations of stationary sequences*, Ukrainian Mathematical Journal, vol. 43, no. 1 pp. 92-99, 1991.
44. M. P. Moklyachuk, *Minimax filtering of linear transforms of stationary processes*, Theory of Probability and Mathematical Statistics, vol. 44, pp. 95-102, 1992.
45. M. P. Moklyachuk, *On linear prediction of random processes under conditions of uncertainty*, Theory of Probability and Mathematical Statistics, vol. 45, pp. 87-93, 1992.
46. M. P. Moklyachuk, *On the problem of minimax extrapolation of vector sequences perturbed by white noise*, Theory of Probability and Mathematical Statistics, vol. 46, pp. 89-102, 1993.
47. M. P. Moklyachuk, *On a filtering problem for vector-valued sequences*, Theory of Probability and Mathematical Statistics, vol. 47, pp. 107-118, 1993.
48. M. P. Moklyachuk, *On stochastic equation describing the one-sided moving average process and minimax estimation problem*, Random Operators and Stochastic Equations, vol. 1, no.4, pp. 329-343, 1993.
49. M. P. Moklyachuk, *On minimax filtration of vector processes*, Ukrainian Mathematical Journal, vol.45, no.3, pp. 414-423, 1993.
50. M. P. Moklyachuk, *Stochastic autoregressive sequences and minimax interpolation*, Theory of Probability and Mathematical Statistics, vol. 48, pp. 95-103, 1994.
51. M. P. Moklyachuk, *On interpolation problem for vector-valued stochastic sequences*, Random Operators and Stochastic Equations, vol. 3, no.1, pp. 63-74, 1995.
52. M. P. Moklyachuk, *Estimates of stochastic processes from observations with noise*, Theory of Stochastic Processes, vol.3(19), no.3-4, pp. 330-338, 1997.
53. M. P. Moklyachuk, *Extrapolation of stationary sequences from observations with noise*, Theory of Probability and Mathematical Statistics, vol. 57, pp. 133-141, 1998.
54. M. P. Moklyachuk, *Some problems of estimation from noisy data*, ZAMM, Z. Angew. Math. Mech., vol.78, Suppl.3, pp. 1021-1022, 1998.
55. M. P. Moklyachuk, *Robust procedures in time series analysis*, Theory of Stochastic Processes, vol. 6, no. 3-4, pp. 127-147, 2000.
56. 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.
57. M. P. Moklyachuk, *Robust estimations of functionals of stochastic processes*, Kyiv University, Kyiv, 2008.
58. M. P. Moklyachuk, *Nonsmooth analysis and optimization*, Kyiv University, Kyiv, 2008.
59. M. Moklyachuk and A. Masyutka, *Extrapolation of multidimensional stationary processes*, Random Operators and Stochastic Equations, vol. 14, no. 3, pp. 233-244, 2006.
60. M. Moklyachuk and A. Masyutka, *Robust estimation problems for stochastic processes*, Theory of Stochastic Processes, vol. 12, no. 3-4, pp. 88-113, 2006.
61. M. Moklyachuk and A. Masyutka, *Robust filtering of stochastic processes*, Theory of Stochastic Processes, vol. 13, no. 1-2, pp. 166-181, 2007.
62. M. Moklyachuk, and A. Masyutka, *Minimax prediction problem for multidimensional stationary stochastic sequences*, Theory of Stochastic Processes, vol. 14, no. 3-4, pp. 89-103, 2008.
63. M. Moklyachuk and A. Masyutka, *Minimax prediction problem for multidimensional stationary stochastic processes*, Communications in Statistics – Theory and Methods., vol. 40, no. 19-20, pp. 3700-3710, 2001.
64. M. Moklyachuk and O. Masyutka, *Minimax-robust estimation technique for stationary stochastic processes*, LAP LAMBERT Academic Publishing, 2012.

65. M. Moklyachuk and V. Ostapenko, *Minimax interpolation of harmonizable sequences*, Theory of Probability and Mathematical Statistics, vol. 92, pp. 125–136, 2015.
66. M. Moklyachuk and V. Ostapenko, *Minimax interpolation problem for harmonizable stable sequences with noise observations*, Journal of Applied Mathematics and Statistics (submitted), 2015.
67. M. Moklyachuk and M. Sidei, *Interpolation problem for stationary sequences with missing observations*, Statistics, Optimization & Information Computing, vol. 3, no. 3, pp. 259–275, 2015.
68. M. Moklyachuk and M. Sidei, *Interpolation of stationary sequences with missing observations observed with noise*, Theory of Probability and Mathematical Statistics (submitted), 2015.
69. M. Moklyachuk and N. Shchestyuk, *Estimates of functionals from random fields*, Uzhgorod: Autdor-Shark, 2013.
70. B. N. Pshenichnyj, *Necessary conditions of an extremum*, Pure and Applied mathematics. 4. New York: Marcel Dekker, 1971.
71. Frigyes Riesz and Béla Sz.–Nagy, *Functional analysis*, Dover Books on Advanced Mathematics. New York: Dover Publications, Inc. XII, 1990.
72. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1997.
73. Yu. A. Rozanov, *Stationary stochastic processes*, San Francisco-Cambridge-London-Amsterdam: Holden-Day, 1967.
74. H. Salehi, *Algorithms for linear interpolator and interpolation error for minimal stationary stochastic processes*, The Annals of Probability, Vol. 7, No. 5, pp. 840–846, 1979.
75. M. Taniguchi, *Robust regression and interpolation for time series*, J. Time Series Analysis, vol.2, pp. 53–62, 1981.
76. 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.
77. 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.
78. H. Wold, *A study in the analysis of stationary time series*, Thesis University of Stockholm, 1938.
79. H. Wold, *On prediction in stationary time series*, Annals of Mathematical Statistics, vol. 19, No. 4, pp. 558–567, 1948.
80. 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.
81. 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.