STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **3**, June 2015, pp 160–188.
Published online in International Academic Press (www.IAPress.org)

Minimax-robust prediction problem for stochastic sequences with stationary increments and cointegrated sequences

Maksym Luz [1], Mikhail Moklyachuk [1,*]

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

Received: 27 February 2015; Accepted: 21 April 2015
Editor: David G. Yu

Abstract The problem of optimal estimation of the linear functionals $A\xi=\sum_{k=0}^{\infty}a(k)\xi(k)$ and $A_N\xi=\sum_{k=0}^{N}a(k)\xi(k)$ which depend on the unknown values of a stochastic sequence $\xi(m)$ with stationary nth increments is considered. Estimates are obtained which are based on observations of the sequence $\xi(m)+\eta(m)$ at points of time $m=-1,-2,\ldots$, where the sequence $\eta(m)$ is stationary and uncorrelated with the sequence $\xi(m)$. Formulas for calculating the mean-square errors and the spectral characteristics of the optimal estimates of the functionals are derived in the case of spectral certainty, where spectral densities of the sequences $\xi(m)$ and $\eta(m)$ are exactly known. These results are applied for solving extrapolation problem for cointegrated sequences. In the case where spectral densities of the sequences are not known exactly, but sets of admissible spectral densities are given, the minimax-robust method of estimation is applied. Formulas that determine the least favorable spectral densities and minimax spectral characteristics are proposed for some special classes of admissible densities.

Keywords Stochastic sequence with stationary increments, cointegrated sequences, minimax-robust estimate, mean square error, least favorable spectral density, minimax-robust spectral characteristic

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

DOI: 10.19139/soic.v3i2.132

1. Introduction

The theory of estimation of the unknown values of stationary processes based on a set of observations plays an important role in many practical applications. The development of the theory started from the classical works of Kolmogorov [18] and Wiener [42], 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 a stochastic sequence. 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 random 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 the book by Rozanov [41]. Most of results which have appeared since that time were based on the assumption that the spectral structure of the stationary process is known. After the main points of the new theory were established, scientists tried to generalize the concept of stationarity. One of the natural generalization was proposed by Yaglom [45], Pinsker [38], Yaglom and Pinsker [37]. They proposed a class of processes with stationary increments of nth order.

*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.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

They described basic properties of these processes, found the spectral representation of the stationary increment and solved the extrapolation problem for processes with stationary nth increments. Other generalizations of the concept of stationarity can be found in the books by Yaglom [43, 44].

One of the fields of practical applications of the stationary and related stochastic sequences is economical modeling 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. Models with trends and seasonal components are represented by integrated ARMA (ARIMA) sequences and seasonal time series. The spectral structure of these sequences has unit roots in the autoregressive part. These sequences are most simple examples of sequences with stationary increments. Such models have been inducing the interest of scientists for the last 30 years. 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]. Statistical investigations of real data found some specific relations between integrated sequences. In some cases linear combinations of such sequences appears to be stationary. Such property Grander [5] called cointegration. Cointegrated models found their application in applied and theoretical econometrics and financial time series [11].

As we have already mentioned the problem of estimation of unobserved values of the investigated time series is important in mathematical studies. The classical methods of extrapolation, interpolation and filtering relay on the exact information about the spectral densities of the investigated processes. However, in practise none of the methods of estimation can provide the exact representation of spectral structure of the process. In the case where spectral densities are not known exactly, but a set of admissible spectral densities are given, we can apply the minimax (robust) method of estimation, which allows us to determine estimates that minimize the value of mean-square error for all densities from a given class. Grenander [12] was the first one who applied this approach to the extrapolation problem for stationary processes. In the papers by Franke [13], Franke and Poor [14], Kassam and Poor [17] the minimax extrapolation and interpolation problem for stationary sequences was solved by using convex optimization techniques. In the works by Moklyachuk [27] - [33] problems of extrapolation, interpolation and filtering for stationary processes and sequences were studied. The minimax extrapolation problem for functionals which depend on the unknown values of stationary sequences from observations with noise is solved in the paper by Moklyachuk [26]. The corresponding problems for vector-valued stationary sequences and processes were investigated by Moklyachuk and Masyutka [30] - [35]. In the articles by Dubovets'ka and Moklyachuk [6] - [10] and the book by Golichenko and Moklyachuk [3] the minimax estimation problems were investigated for another generalization of stationary processes – periodically correlated stochastic sequences and random processes.

Luz and Moklyachuk investigated the classical and minimax extrapolation, interpolation and filtering problems for sequences and processes with nth stationary increments. They presented solutions of the filtering problem for the linear functionals $A\xi=\sum_{k=0}^{\infty}a(k)\xi(-k)$ and $A_N\xi=\sum_{k=0}^{N}a(k)\xi(-k)$ in the papers [21, 23]. The minimax interpolation problem for the linear functional $A_N\xi=\sum_{k=0}^{N}a(k)\xi(k)$ which depends on the unknown values of the sequence $\xi(k)$ based on observations with and without noise was investigated in papers [19, 20], and for the linear functional $A\xi=\int_0^{\infty}a(t)\xi(t)dt$ which depends on the unknown values of a random process $\xi(t)$ in the paper [24].

In papers by Luz and Moklyachuk [22, 25] the problem of optimal linear extrapolation of linear functionals which depend on the unknown values of stochastic sequences and random processes with nth stationary increments from the observations without noise is investigate. The classical extrapolation problem for a non-stationary sequence which is observed with a non-stationary noise was studied by Bell [1]. However, he showed that the problem can be solved under additional assumptions, particularly if we have an additional finite set of values of the sequence $\xi(m)$.

In the proposed paper we consider the extrapolation problem for the functionals $A\xi=\sum_{k=0}^{\infty}a(k)\xi(k)$ and $A_N\xi=\sum_{k=0}^{N}a(k)\xi(k)$ which depend on the unknown values of a stochastic sequence $\xi(k)$ with stationary nth increments based on observations of the sequence $\xi(k)+\eta(k)$ at points $k=-1,-2,\ldots$, where $\eta(k)$ is a stationary stochastic sequence uncorrelated with the sequence $\xi(k)$. Under the condition of stationarity of the noise $\eta(k)$ we solve the problem without additional assumptions described by Bell [1]. The obtained estimates give us a method of solvution the extrapolation problem for cointegrated sequences $\xi(m)$ and $\zeta(m)$ assuming stationarity of a linear combination of these sequences. The estimation problem is also solved in the case of spectral uncertainty where spectral densities of sequences are not exactly known but a set of admissible spectral densities is given. Formulas that determine the least favorable 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.

2. Stationary increment stochastic sequences. Spectral representation

Definition 1
For a given stochastic sequence $\{\xi(m), m \in \mathbb{Z}\}$ the sequence

$$\xi^{(n)}(m,\mu) = (1-B_\mu)^n \xi(m) = \sum_{l=0}^{n} (-1)^l \binom{n}{l} \xi(m-l\mu), \tag{1}$$

where $\binom{n}{l} = \frac{n!}{l!(n-l)!}$, B_μ is a backward shift operator with step $\mu \in \mathbb{Z}$, such that $B_\mu \xi(m) = \xi(m-\mu)$, is called stochastic nth increment sequence with step $\mu \in \mathbb{Z}$.

The stochastic nth increment sequence $\xi^{(n)}(m,\mu)$ admits the following relations

$$\xi^{(n)}(m,-\mu) = (-1)^n \xi^{(n)}(m+n\mu,\mu), \tag{2}$$

$$\xi^{(n)}(m,k\mu) = \sum\nolimits_{l=0}^{(k-1)n} A_l \xi^{(n)}(m-l\mu,\mu), \quad k \in \mathbb{N}, \tag{3}$$

where coefficients $\{A_l, l = 0,1,2,\ldots,(k-1)n\}$ are determined by the representation

$$(1+x+\ldots+x^{k-1})^n = \sum_{l=0}^{(k-1)n} A_l x^l.$$

Definition 2
The stochastic nth increment sequence $\xi^{(n)}(m,\mu)$ generated by stochastic sequence $\{\xi(m), m \in \mathbb{Z}\}$ is wide sense stationary if the mathematical expectations

$$\mathsf{E}\xi^{(n)}(m_0,\mu) = c^{(n)}(\mu),$$

$$\mathsf{E}\xi^{(n)}(m_0+m,\mu_1)\xi^{(n)}(m_0,\mu_2) = D^{(n)}(m,\mu_1,\mu_2)$$

exist for all $m_0, \mu, m, \mu_1, \mu_2$ and do not depend on m_0. The function $c^{(n)}(\mu)$ is called mean value of the nth increment sequence and the function $D^{(n)}(m,\mu_1,\mu_2)$ is called structural function of the stationary nth increment sequence (or structural function of nth order of the stochastic sequence $\{\xi(m), m \in \mathbb{Z}\}$).

The stochastic sequence $\{\xi(m), m \in \mathbb{Z}\}$ which determines the stationary nth increment sequence $\xi^{(n)}(m,\mu)$ by formula (1) is called sequence with stationary nth increments (or integrated sequence of order n).

Theorem 1
The mean value $c^{(n)}(\mu)$ and the structural function $D^{(n)}(m,\mu_1,\mu_2)$ of the stochastic stationary nth increment sequence $\xi^{(n)}(m,\mu)$ can be represented in the forms

$$c^{(n)}(\mu) = c\mu^n, \tag{4}$$

$$D^{(n)}(m,\mu_1,\mu_2) = \int_{-\pi}^{\pi} e^{i\lambda m}(1-e^{-i\mu_1\lambda})^n(1-e^{i\mu_2\lambda})^n \frac{1}{\lambda^{2n}} dF(\lambda), \tag{5}$$

where c is a constant, $F(\lambda)$ is a left-continuous nondecreasing bounded function with $F(-\pi)=0$. The constant c and the function $F(\lambda)$ are determined uniquely by the increment sequence $\xi^{(n)}(m,\mu)$.

On other hand, a function $c^{(n)}(\mu)$ which has form (4) with a constant c and a function $D^{(n)}(m,\mu_1,\mu_2)$ which has form (5) with a function $F(\lambda)$ which satisfies the indicated conditions are the mean value and the structural function of a stationary nth increment sequence $\xi^{(n)}(m,\mu)$.

Representation (5) and the Karhunen theorem [4, 16] give us the spectral representation of the stationary nth increment sequence $\xi^{(n)}(m,\mu)$:

$$\xi^{(n)}(m,\mu)=\int_{-\pi}^{\pi}e^{im\lambda}(1-e^{-i\mu\lambda})^n\frac{1}{(i\lambda)^n}dZ_{\xi^{(n)}}(\lambda), \tag{6}$$

where $Z_{\xi^{(n)}}(\lambda)$ is a random process with independent increments on $[-\pi,\pi)$ connected with the spectral function $F(\lambda)$ by the relation

$$\mathsf{E}|Z_{\xi^{(n)}}(t_2)-Z_{\xi^{(n)}}(t_1)|^2=F(t_2)-F(t_1)<\infty \quad \text{for all } -\pi\leq t_1<t_2<\pi. \tag{7}$$

Denote by $H(\xi^{(n)})$ a subspace generated in the Hilbert space $H=L_2(\Omega,\mathcal{F},\mathcal{P})$ by elements $\{\xi^{(n)}(m,\mu):m,\mu\in\mathbb{Z}\}$ and by $H^p(\xi^{(n)})$, $p\in\mathbb{Z}$, a subspace of the space $H(\xi^{(n)})$ generated by elements $\{\xi^{(n)}(m,\mu):m\leq p,\mu>0\}$. Let

$$S(\xi^{(n)})=\bigcap_{p\in\mathbb{Z}}H^p(\xi^{(n)}).$$

Since the space $S(\xi^{(n)})$ is a subspace in the Hilbert space $H(\xi^{(n)})$, the space $H(\xi^{(n)})$ admits the decomposition

$$H(\xi^{(n)})=S(\xi^{(n)})\oplus R(\xi^{(n)}),$$

where $R(\xi^{(n)})$ is the orthogonal complement of the subspace $S(\xi^{(n)})$ in the space $H(\xi^{(n)})$.

Definition 3
A stationary nth increment sequence $\xi^{(n)}(m,\mu)$ is called regular if $H(\xi^{(n)})=R(\xi^{(n)})$ and it is called singular if $H(\xi^{(n)})=S(\xi^{(n)})$.

Theorem 2
A wide-sense stationary stochastic increment sequence $\xi^{(n)}(m,\mu)$ admits a unique representation in the form

$$\xi^{(n)}(m,\mu)=\xi^{(n)}_r(m,\mu)+\xi^{(n)}_s(m,\mu), \tag{8}$$

where $\{\xi^{(n)}_r(m,\mu):m\in\mathbb{Z}\}$ is a regular increment sequence and $\{\xi^{(n)}_s(m,\mu):m\in\mathbb{Z}\}$ is a singular increment sequence. Moreover, the increment sequences $\xi^{(n)}_r(m,\mu)$ and $\xi^{(n)}_s(k,\mu)$ are orthogonal for all $m,k\in\mathbb{Z}$.

Components of representation (8) are defined by the formulas

$$\xi^{(n)}_s(m,\mu)=\mathsf{E}[\xi^{(n)}(m,\mu)|S(\xi^{(n)})],\quad \xi^{(n)}_r(m,\mu)=\xi^{(n)}(m,\mu)-\xi^{(n)}_s(m,\mu).$$

Consider a stochastic sequence $\{\varepsilon(m):m\in\mathbb{Z}\}$ of uncorrelated random variables such that $\mathsf{E}\varepsilon(m)=0$ and $\mathsf{D}\varepsilon_m^2=1$. Define by $H^p(\varepsilon)$ the Hilbert subspace generated by elements $\{\varepsilon(m):m\leq p\}$. We will call the sequence $\{\varepsilon(m):m\in\mathbb{Z}\}$ an innovation sequence for a regular stationary nth increment sequence $\xi^{(n)}(m,\mu)$ if the condition $H^p(\xi^{(n)})=H^p(\varepsilon)$ holds true for all $p\in\mathbb{Z}$.

Theorem 3
A stochastic stationary increment sequence $\xi^{(n)}(m,\mu)$ is regular if and only if there exists an innovation sequence $\{\varepsilon(m):m\in\mathbb{Z}\}$ and a sequence of complex functions $\{\varphi^{(n)}(m,\mu):m\geq 0\}$, $\sum_{k=0}^{\infty}|\varphi^{(n)}(k,\mu)|^2<\infty$ such that

$$\xi^{(n)}(m,\mu)=\sum_{k=0}^{\infty}\varphi^{(n)}(k,\mu)\varepsilon(m-k). \tag{9}$$

Representation (9) is called canonical moving average representation of the stochastic stationary increment sequence $\xi^{(n)}(m,\mu)$.

Corollary 1
A wide-sense stationary stochastic increment sequence $\xi^{(n)}(m,\mu)$ admits a unique representation

$$\xi^{(n)}(m,\mu)=\xi_s^{(n)}(m,\mu)+\sum_{k=0}^{\infty}\varphi^{(n)}(k,\mu)\varepsilon(m-k), \tag{10}$$

where $\sum_{k=0}^{\infty}|\varphi^{(n)}(k,\mu)|^2<\infty$ and $\{\varepsilon_m : m\in\mathbb{Z}\}$ is an innovation sequence.

If the stationary nth increment sequence $\xi^{(n)}(m,\mu)$ admit the canonical representation (9), then its spectral function $F(\lambda)$ has the spectral density $f(\lambda)$ admitting the canonical factorization

$$f(\lambda)=|\Phi(e^{-i\lambda})|^2,\quad \Phi(z)=\sum_{k=0}^{\infty}\varphi(k)z^k, \tag{11}$$

where the function $\Phi(z)=\sum_{k=0}^{\infty}\varphi(k)z^k$ has the convergence radius $r>1$ and does not have zeros in the unit disk $\{z:|z|\leq 1\}$. Define

$$\Phi_\mu(z)=\sum_{k=0}^{\infty}\varphi^{(n)}(k,\mu)z^k=\sum_{k=0}^{\infty}\varphi_\mu(k)z^k,$$

where $\varphi_\mu(k)=\varphi^{(n)}(k,\mu)$ are coefficients from the canonical representation (9). Then the following relation holds true:

$$\left|\Phi_\mu(e^{-i\lambda})\right|^2=\frac{|1-e^{-i\lambda\mu}|^{2n}}{\lambda^{2n}}f(\lambda). \tag{12}$$

In the next section we will use spectral representation (6) and canonical factorization (12) for finding the optimal mean square estimate of the unknown values of the stochastic sequence $\{\xi(m),m\in\mathbb{Z}\}$ with nth stationary increments.

3. Extrapolation problem

Consider a stochastic sequence $\{\xi(m),m\in\mathbb{Z}\}$ which generates a stationary nth increment sequence $\xi^{(n)}(m,\mu)$ with absolutely continuous spectral function $F(\lambda)$ and spectral density $f(\lambda)$. Let $\{\eta(m),m\in\mathbb{Z}\}$ be an uncorrelated with the sequence $\xi(m)$ stationary stochastic sequence with absolutely continuous spectral function $G(\lambda)$ and spectral density $g(\lambda)$. From now we will assume that mean values of the increment sequence $\xi^{(n)}(m,\mu)$ and stationary sequence $\eta(m)$ equal to 0. We will also consider the increment step $\mu>0$.

In this section our purpose is to solve the problem of linear mean-square optimal estimation of the functionals

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

which depend on unknown values of the sequence $\xi(m)$ based on observations of the sequence $\zeta(m)=\xi(m)+\eta(m)$ at points $m=-1,-2,\ldots$.

First of all we indicate some conditions which are necessary for solving the considered problem. Assume that coefficients $a(k)$, $k\geq 0$, and the linear transformation D^μ which is defined in the following part of the section satisfy the conditions

$$\sum_{k=0}^{\infty}|a(k)|<\infty,\quad \sum_{k=0}^{\infty}(k+1)|a(k)|^2<\infty, \tag{13}$$

and

$$\sum_{k=0}^{\infty}|(D^\mu\mathbf{a})_k|<\infty,\quad \sum_{k=0}^{\infty}(k+1)|(D^\mu\mathbf{a})_k|^2<\infty. \tag{14}$$

Assume also that spectral densities $f(\lambda)$ and $g(\lambda)$ satisfy the minimality condition

$$\int_{-\pi}^{\pi} \frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))} d\lambda < \infty. \tag{15}$$

In order to find an estimate of the functional $A\xi$ we have to formulate the extrapolation problem in terms of linear functionals of some stationary sequences. The functional $A\xi$ can be presented as

$$A\xi = A\zeta - A\eta,$$

where $A\zeta = \sum_{k=0}^{\infty} a(k)\zeta(k)$, $A\eta = \sum_{k=0}^{\infty} a(k)\eta(k)$. Under conditions (13) the functional $A\eta$ has finite second moment.

We will exploit the representation of the functional $A\zeta$ which is proposed in [25] and is described in the following lemma.

Lemma 1
A linear functional $A\zeta = \sum_{k=0}^{\infty} a(k)\zeta(k)$ admits the representation $A\zeta = B\zeta - V\zeta$, where

$$B\zeta = \sum_{k=0}^{\infty} b_\mu(k)\zeta^{(n)}(k,\mu), \quad V\xi = \sum_{k=-\mu n}^{-1} v_\mu(k)\zeta(k),$$

$$v_\mu(k) = \sum_{l=[-\frac{k}{\mu}]'}^{n} (-1)^l \binom{n}{l} b_\mu(l\mu+k), \quad k=-1,-2,\dots,-\mu n, \tag{16}$$

$$b_\mu(k) = \sum_{m=k}^{\infty} a(m)d_\mu(m-k) = (D^\mu \mathbf{a})_k, \ k \geq 0. \tag{17}$$

where $[x]'$ denotes the least integer number among numbers which are greater or equal to x, $\{d(k) : k \geq 0\}$ are coefficients determined by the relation $\sum_{k=0}^{\infty} d_\mu(k)x^k = \left(\sum_{j=0}^{\infty} x^{\mu j}\right)^n$, D^μ is a linear operator determined by elements $D^\mu_{k,j} = d_\mu(j-k)$ if $0 \leq k \leq j$, and $D^\mu_{k,j} = 0$ if $j < k$, the vector $\mathbf{a} = (a(0), a(1), a(2), \dots)'$.

Corollary 2
The linear functional $A_N\zeta$ admits the representation $A_N\zeta = B_N\zeta - V_N\zeta$, where

$$B_N\zeta = \sum_{k=0}^{N} b_{\mu,N}(k)\zeta^{(n)}(k,\mu), \quad V_N\zeta = \sum_{k=-\mu n}^{-1} v_{\mu,N}(k)\zeta(k),$$

where coefficients $v_{\mu,N}(k)$, $k=-1,-2,\dots,-\mu n$, are calculated by the formulas

$$v_{\mu,N}(k) = \sum_{l=[-\frac{k}{\mu}]'}^{\min\{[\frac{N-k}{\mu}],n\}} (-1)^l \binom{n}{l} b_{\mu,N}(l\mu+k), \quad k=-1,-2,\dots,-\mu n,$$

$$b_{\mu,N}(k) = \sum_{m=k}^{N} a(m)d_\mu(m-k) = (D^\mu_N \mathbf{a}_N)_k, \quad k=0,1,\dots,N,$$

D^μ_N is a linear operator with elements $(D^\mu_N)_{k,j} = d_\mu(j-k)$ if $0 \leq k \leq j \leq N$, and $(D^\mu_N)_{k,j} = 0$ if $j < k$ or $j, k > N$, the vector $\mathbf{a}_N = (a(0), a(1), a(2), \dots, a(N), 0, \dots)'$.

From Lemma 1 we get the following representation of the functional $A\xi$:

$$A\xi = A\zeta - A\eta = B\zeta - A\eta - V\zeta = H\xi - V\zeta,$$

where $H\xi = B\zeta - A\eta$. Denote by $\Delta(f,g;\widehat{A}\xi) = \mathsf{E}|A\xi - \widehat{A}\xi|^2$ the mean-square error of the optimal estimate $\widehat{A}\xi$ of the functional $A\xi$ and by $\Delta(f,g;\widehat{H}\xi) = \mathsf{E}|H\xi - \widehat{H}\xi|^2$ the mean-square error of the optimal estimate $\widehat{H}\eta$ of the functional $H\eta$. Since the functional $V\zeta$ is determined by the observed values of $\zeta(k)$ at points $k = -\mu n, -\mu n+1, \dots, -1$, the following relations hold true

$$\widehat{A}\xi = \widehat{H}\xi - V\zeta, \tag{18}$$

$$\Delta(f,g;\widehat{A}\xi) = \mathsf{E}|A\xi - \widehat{A}\xi|^2 = \mathsf{E}|H\xi - V\zeta - \widehat{H}\xi + V\zeta|^2 = \mathsf{E}|H\xi - \widehat{H}\xi|^2 = \Delta(f,g;\widehat{H}\xi).$$

To find the mean-square optimal estimate of the functional $H\xi$ we apply the Hilbert space orthogonal projection method proposed by Kolmogorov [18]. The stationary stochastic sequence $\eta(m)$ admits the spectral representation

$$\eta(m) = \int_{-\pi}^{\pi} e^{i\lambda m} dZ_{\eta}(\lambda),$$

where $Z_{\eta}(\lambda)$ is a random process with independent increments on $[-\pi,\pi)$ corresponding to the spectral function $G(\lambda)$. The random processes $Z_{\eta}(\lambda)$ and $Z_{\eta^{(n)}}(\lambda)$ are connected by the relation $dZ_{\eta^{(n)}}(\lambda) = (i\lambda)^n dZ_{\eta}(\lambda)$, $\lambda \in [-\pi,\pi)$, obtained in [23]. The spectral density $p(\lambda)$ of the sequence $\zeta(m)$ is determined by spectral densities $f(\lambda)$ and $g(\lambda)$ by the relation

$$p(\lambda) = f(\lambda) + \lambda^{2n} g(\lambda).$$

The functional $H\xi$ admits the following spectral representation:

$$H\xi = \int_{-\pi}^{\pi} B_{\mu}(e^{i\lambda}) \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda) - \int_{-\pi}^{\pi} A(e^{i\lambda}) dZ_{\eta}(\lambda),$$

where

$$B_{\mu}(e^{i\lambda}) = \sum_{k=0}^{\infty} b_{\mu}(k) e^{i\lambda k} = \sum_{k=0}^{\infty} (D^{\mu}\mathbf{a})_k e^{i\lambda k}, \quad A(e^{i\lambda}) = \sum_{k=0}^{\infty} a(k) e^{i\lambda k}.$$

Denote by $H^{0-}(\xi_{\mu}^{(n)} + \eta_{\mu}^{(n)})$ a closed linear subspace of the Hilbert space $H = L_2(\Omega, \mathfrak{F}, \mathsf{P})$ of random variables having finite second moments which is generated by values $\{\xi^{(n)}(k,\mu) + \eta^{(n)}(k,\mu) : k \leq -1\}$, $\mu > 0$. Denote by $L_2^{0-}(f(\lambda) + \lambda^{2n} g(\lambda))$ a closed linear subspace of the Hilbert space $L_2(f(\lambda) + \lambda^{2n} g(\lambda))$ which is generated by the set of functions

$$\left\{ e^{i\lambda k}(1-e^{-i\lambda\mu})^n (i\lambda)^{-n} : k \leq -1 \right\}.$$

The representation

$$\xi^{(n)}(k,\mu) + \eta^{(n)}(k,\mu) = \int_{-\pi}^{\pi} e^{i\lambda k}(1-e^{-i\lambda\mu})^n \frac{1}{(i\lambda)^n} dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda)$$

yields a one to one correspondence between elements $e^{i\lambda k}(1-e^{-i\lambda\mu})^n (i\lambda)^{-n}$ of the space $L_2^{0-}(f(\lambda) + \lambda^{2n} g(\lambda))$ and elements $\xi^{(n)}(k,\mu) + \eta^{(n)}(k,\mu)$ of the space $H^{0-}(\xi_{\mu}^{(n)} + \eta_{\mu}^{(n)})$.

Every linear estimate $\widehat{A}\xi$ of the functional $A\xi$ admits the representation

$$\widehat{A}\xi = \int_{-\pi}^{\pi} h_{\mu}(\lambda) dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda) - \sum_{k=-\mu n}^{-1} v_{\mu}(k)(\xi(k) + \eta(k)), \tag{19}$$

where $h_{\mu}(\lambda)$ is the spectral characteristic of the estimate $\widehat{H}\xi$. The mean square optimal estimate $\widehat{H}\xi$ can be found as a projection of the element $H\xi$ on the subspace $H^{0-}(\xi_{\mu}^{(n)} + \eta_{\mu}^{(n)})$. This projection is determined by two conditions:

1) $\widehat{H}\xi \in H^{0-}(\xi_{\mu}^{(n)}+\eta_{\mu}^{(n)})$;
2) $(H\xi-\widehat{H}\xi) \perp H^{0-}(\xi_{\mu}^{(n)}+\eta_{\mu}^{(n)})$.
The second condition implies the following relations which hold true for all $k \le -1$:

$$\mathsf{E}(H\xi-\widehat{H}\xi)\overline{(\xi^{(n)}(k,\mu)+\eta^{(n)}(k,\mu))}$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(B_{\mu}(e^{i\lambda})(1-e^{-i\lambda\mu})^n-A(e^{i\lambda})-(i\lambda)^n h_{\mu}(\lambda)\right)e^{-i\lambda k}(1-e^{i\lambda\mu})^n g(\lambda)d\lambda$$
$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(B_{\mu}(e^{i\lambda})\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}-h_{\mu}(\lambda)\right)e^{-i\lambda k}(1-e^{i\lambda\mu})^n\frac{1}{(-i\lambda)^n}f(\lambda)d\lambda=0.$$

These relations can be represented in the form

$$\int_{-\pi}^{\pi}\left[\left(B_{\mu}(e^{i\lambda})\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}-h_{\mu}(\lambda)\right)p(\lambda)-A(e^{i\lambda})g(\lambda)(-i\lambda)^n\right]\frac{(1-e^{i\lambda\mu})^n}{(-i\lambda)^n}e^{-i\lambda k}d\lambda=0,\ k\le -1.$$

which allows us to derive the spectral characteristic $h_{\mu}(\lambda)$ of the estimate $\widehat{H}\xi$. It has the form

$$h_{\mu}(\lambda)=B_{\mu}(e^{i\lambda})\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}-A(e^{i\lambda})\frac{(-i\lambda)^n g(\lambda)}{f(\lambda)+\lambda^{2n}g(\lambda)}-\frac{(-i\lambda)^n C_{\mu}(e^{i\lambda})}{(1-e^{i\lambda\mu})^n(f(\lambda)+\lambda^{2n}g(\lambda))},$$

$$C_{\mu}(e^{i\lambda})=\sum_{k=0}^{\infty}c_{\mu}(k)e^{i\lambda k},$$

where $c_{\mu}(k)$, $k\ge 0$, are unknown coefficients which we need to determine. It follows from condition 1) that the spectral characteristic $h_{\mu}(\lambda)$ admits the representation

$$h_{\mu}(\lambda)=h(\lambda)(1-e^{-i\lambda\mu})^n\frac{1}{(i\lambda)^n},\quad h(\lambda)=\sum_{k=1}^{\infty}s(k)e^{-i\lambda k},$$

where

$$\int_{-\pi}^{\pi}|h(\lambda)|^2|1-e^{i\lambda\mu}|^{2n}\frac{f(\lambda)+\lambda^{2n}g(\lambda)}{\lambda^{2n}}d\lambda<\infty,\quad \frac{(i\lambda)^n h_{\mu}(\lambda)}{(1-e^{-i\lambda\mu})^n}\in L_2^{0-},$$

which leads to the conditions

$$\int_{-\pi}^{\pi}\left[B_{\mu}(e^{i\lambda})-\frac{A(e^{i\lambda})\lambda^{2n}g(\lambda)}{(1-e^{-i\lambda\mu})^n(f(\lambda)+\lambda^{2n}g(\lambda))}-\frac{\lambda^{2n}C_{\mu}(e^{i\lambda})}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}\right]e^{-i\lambda l}d\lambda=0,\ l\ge 0. \quad (20)$$

Determine for every $k,j\in\mathbb{Z}$ the Fourier coefficients of the corresponding functions

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

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

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

Using these Fourier coefficients we can represent equation (20) in terms of the system of linear equations

$$b_{\mu}(l)-\sum_{m=0}^{\infty}T_{l,m}^{\mu}a_{\mu}(m)=\sum_{k=0}^{\infty}P_{l,k}^{\mu}c_{\mu}(k),\quad l\ge 0,$$

where

$$a_\mu(m)=\sum_{l=0}^{\min\{n;[\frac{m}{\mu}]\}}(-1)^l\binom{n}{l}a(m-\mu l),\quad m\geq 0, \tag{21}$$

This system of equations can be written in the form

$$D^\mu\mathbf{a}-\mathbf{T}_\mu\mathbf{a}_\mu=\mathbf{P}_\mu\mathbf{c}_\mu,$$

where $\mathbf{c}_\mu=(c_\mu(0),c_\mu(1),c_\mu(2),\dots)'$, $\mathbf{a}_\mu=(a_\mu(0),a_\mu(1),a_\mu(2),\dots)'$; $\mathbf{P}_\mu$ and $\mathbf{T}_\mu$ are linear operators in the space ℓ_2 defined by the matrices with elements $(\mathbf{P}_\mu)_{l,k}=P^\mu_{l,k}$, $l,k\geq 0$ and $(\mathbf{T}_\mu)_{l,m}=T^\mu_{l,k}$, $l,k\geq 0$; the linear transformation D^μ is defined in Lemma 1. Consequently, the unknown coefficients $c_\mu(k)$, $k\geq 0$, which determine the spectral characteristic $h_\mu(\lambda)$ are calculated by the formula

$$c_\mu(k)=(\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k,\quad k\geq 0,$$

where $(\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k$, $k\geq 0$, is the kth element of the vector $\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu$. Thus, the spectral characteristic $h_\mu(\lambda)$ of the optimal estimate $\widehat{H}\xi$ of the functional $H\xi$ is calculated by the formula

$$h_\mu(\lambda)=B_\mu(e^{i\lambda})\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}-A(e^{i\lambda})\frac{(-i\lambda)^n g(\lambda)}{f(\lambda)+\lambda^{2n}g(\lambda)}-\frac{(-i\lambda)^n\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k e^{i\lambda k}}{(1-e^{i\lambda\mu})^n(f(\lambda)+\lambda^{2n}g(\lambda))}. \tag{22}$$

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

$$\Delta(f,g;\widehat{A}\xi)=\Delta(f,g;\widehat{H}\xi)=\mathsf{E}|H\xi-\widehat{H}\xi|^2$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n f(\lambda)-\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k e^{i\lambda k}\right|^2}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))^2}g(\lambda)d\lambda$$

$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n\lambda^{2n}g(\lambda)+\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k e^{i\lambda k}\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))^2}f(\lambda)d\lambda$$

$$=\langle D^\mu\mathbf{a}-\mathbf{T}_\mu\mathbf{a}_\mu,\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu\rangle+\langle\mathbf{Q}\mathbf{a},\mathbf{a}\rangle, \tag{23}$$

where $\mathbf{Q}$ is a linear operator in the space ℓ_2 defined by the matrix with elements $(\mathbf{Q})_{l,k}=Q_{l,k}$, $l,k\geq 0$; $\langle x,y\rangle=\sum_{k=0}^{\infty}x(k)\overline{y}(k)$ for vectors $x=(x(0),x(1),x(2),\dots)'$, $y=(y(0),y(1),y(2),\dots)'$.

These reasonings can be summarized in the form of the theorem.

Theorem 4

Let $\{\xi(m),m\in\mathbb{Z}\}$ be a stochastic sequence which defines stationary nth increment sequence $\xi^{(n)}(m,\mu)$ with absolutely continuous spectral function $F(\lambda)$ which has spectral density $f(\lambda)$. Let $\{\eta(m),m\in\mathbb{Z}\}$ be an uncorrelated with the sequence $\xi(m)$ stationary stochastic sequence with absolutely continuous spectral function $G(\lambda)$ which has spectral density $g(\lambda)$. Let the minimality condition (15) be satisfied. Let coefficients $\{a(k):k\geq 0\}$ satisfy conditions (13) – (14). The optimal linear estimate $\widehat{A}\xi$ of the functional $A\xi$ which depend on the unknown values of elements $\xi(m)$, $m\geq 0$, based on observations of the sequence $\xi(m)+\eta(m)$ at points $m=-1,-2,\dots$ is calculated by formula (19). The spectral characteristic $h_\mu(\lambda)$ of the optimal estimate $\widehat{A}\xi$ is calculated by formula (22). The value of the mean-square error $\Delta(f,g;\widehat{A}\xi)$ is calculated by formula (23).

Corollary 3

The spectral characteristic $h_\mu(\lambda)$ admits the representation $h_\mu(\lambda)=h^1_\mu(\lambda)-h^2_\mu(\lambda)$, where

$$h^1_\mu(\lambda)=B_\mu(e^{i\lambda})\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}-\frac{(-i\lambda)^n\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}D^\mu\mathbf{a})_k e^{i\lambda k}}{(1-e^{i\lambda\mu})^n(f(\lambda)+\lambda^{2n}g(\lambda))}, \tag{24}$$

$$h^2_\mu(\lambda)=A(e^{i\lambda})\frac{(-i\lambda)^n g(\lambda)}{f(\lambda)+\lambda^{2n}g(\lambda)}-\frac{(-i\lambda)^n\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k e^{i\lambda k}}{(1-e^{i\lambda\mu})^n(f(\lambda)+\lambda^{2n}g(\lambda))}. \tag{25}$$

Here $h^1_\mu(\lambda)$ and $h^2_\mu(\lambda)$ are the spectral characteristics of the optimal estimates $\widehat{B}\zeta$ and $\widehat{A}\eta$ of the functionals $B\zeta$ and $A\eta$ respectively based on observations $\xi(k)+\eta(k)$ at points $k=-1,-2,\dots$.

Theorem 4 allows us to obtain the optimal estimate $\widehat{A}_N\xi$ of the functional $A_N\xi$ which depend on the unknown values of elements $\xi(m)$, $m=0,1,2,\dots,N$, based on observations of the sequence $\xi(m)+\eta(m)$ at points $m=-1,-2,\dots$. Put $a(k)=0$, $k>N$. Then the spectral characteristic $h_{\mu,N}(\lambda)$ of the linear estimate

$$\widehat{A}_N\xi=\int_{-\pi}^{\pi}h_{\mu,N}(\lambda)dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda)-\sum_{k=-\mu n}^{-1}v_{\mu,N}(k)(\xi(k)+\eta(k)), \tag{26}$$

is calculated by the formula

$$\begin{aligned}h_{\mu,N}(\lambda)&=B^\mu_N(e^{i\lambda})\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}-A_N(e^{i\lambda})\frac{(-i\lambda)^n g(\lambda)}{f(\lambda)+\lambda^{2n}g(\lambda)}\\&-\frac{(-i\lambda)^n\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}D^\mu_N\mathbf{a}_N-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,N}\mathbf{a}_{\mu,N})_k e^{i\lambda k}}{(1-e^{i\lambda\mu})^n(f(\lambda)+\lambda^{2n}g(\lambda))},\end{aligned} \tag{27}$$

where

$$B^\mu_N(e^{i\lambda})=\sum_{k=0}^{N}(D^\mu_N\mathbf{a}_N)_k e^{i\lambda k},\quad A_N(e^{i\lambda})=\sum_{k=0}^{N}a(k)e^{i\lambda k},$$

$\mathbf{a}_N=(a(0),a(1),\dots,a(N),0,\dots)'$, $\mathbf{a}_{\mu,N}=(a_{\mu,N}(0),a_{\mu,N}(1),\dots,a_{\mu,N}(N+\mu n),0,\dots)'$,

$$a_{\mu,N}(m)=\sum_{l=\max\{[\frac{m-N}{\mu}]',0\}}^{\min\{[\frac{m}{\mu}],n\}}(-1)^l\binom{n}{l}a(m-\mu l),\quad 0\le m\le N+\mu n, \tag{28}$$

$\mathbf{T}_{\mu,N}$ is a linear operator in the space ℓ_2 defined by the matrix with elements $(\mathbf{T}_{\mu,N})_{l,m}=T^\mu_{l,m}$, $l\ge 0$, $0\le m\le N+\mu n$, and $(\mathbf{T}_{\mu,N})_{l,m}=0$, $l\ge0$, $m>N+\mu n$. Here $[x]'$ denotes the least integer number among numbers which are greater or equal to x. The mean-square error of the estimate $\widehat{A}_N\xi$ is calculated by the formula

$$\begin{aligned}\Delta(f,g;\widehat{A}_N\xi)&=\Delta(f,g;\widehat{H}_N\xi)=\mathsf{E}|H_N\xi-\widehat{H}_N\xi|^2\\&=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A_N(e^{i\lambda})(1-e^{i\lambda\mu})^n f(\lambda)-\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}D^\mu_N\mathbf{a}_N-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,N}\mathbf{a}_{\mu,N})_k e^{i\lambda k}\right|^2}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))^2}g(\lambda)d\lambda\\&+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A_N(e^{i\lambda})(1-e^{i\lambda\mu})^n\lambda^{2n}g(\lambda)+\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}D^\mu_N\mathbf{a}_N-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,N}\mathbf{a}_{\mu,N})_k e^{i\lambda k}\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))^2}f(\lambda)d\lambda\\&=\langle D^\mu_N\mathbf{a}_N-\mathbf{T}_{\mu,N}\mathbf{a}_{\mu,N},\mathbf{P}_\mu^{-1}D^\mu_N\mathbf{a}_N-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,N}\mathbf{a}_{\mu,N}\rangle+\langle\mathbf{Q}_N\mathbf{a}_N,\mathbf{a}_N\rangle,\end{aligned} \tag{29}$$

where $\mathbf{Q}_N$ is a linear operator in the space ℓ_2 defined by the matrix with elements $(\mathbf{Q}_N)_{l,k}=Q_{l,k}$, $0\le l,k\le N$, and $(\mathbf{Q}_N)_{l,k}=0$ otherwise.

The following theorem holds true.

Theorem 5

Let $\{\xi(m),m\in\mathbb{Z}\}$ be a stochastic sequence which defines stationary nth increment sequence $\xi^{(n)}(m,\mu)$ with an absolutely continuous spectral function $F(\lambda)$ which has spectral density $f(\lambda)$. Let $\{\eta(m),m\in\mathbb{Z}\}$ be an uncorrelated with the sequence $\xi(m)$ stationary stochastic sequence with an absolutely continuous spectral function $G(\lambda)$ which has spectral density $g(\lambda)$. Let the minimality condition (15) be satisfied. The optimal linear estimate

$\widehat{A}_N\xi$ of the functional $A_N\xi$ which depend on the unknown values of elements $\xi(k)$, $k=0,1,2,\dots,N$, from observations of the sequence $\xi(m)+\eta(m)$ at points $m=-1,-2,\dots$ is calculated by formula (26). The spectral characteristic $h_{\mu,N}(\lambda)$ of the optimal estimate $\widehat{A}_N\xi$ is calculated by formula (27). The value of the mean-square error $\Delta(f,g;\widehat{A}_N\xi)$ is calculated by formula (29).

A particular case of the considered problem is the problem of forecasting of an unobserved value of a stochastic sequence $\xi(p)$ at point p, $p\geq 0$, from observations of the sequence $\xi(k)+\eta(k)$ at points $k=-1,-2,\dots$. In this case the vector $\mathbf{a}_{\mu,N}$ has coefficients $a_{\mu,N}(m)=(-1)^l\binom{n}{l}$ if $m=p+\mu l$, $l=0,1,2,\dots,n$, $m\geq 0$, and $a_{\mu,N}(m)=0$ otherwise. Let us define a vector $a_n=(a_n(0),a_n(1),\dots,a_n(n),0,0,\dots)'$, where $a_n(k)=(-1)^k\binom{n}{k}$, $k=0,1,2,\dots,n$. If we choose $\mu>p\geq 0$, the spectral characteristic $h_{\mu,p}(\lambda)$ of the optimal estimate

$$\widehat{\xi}(p)=\int_{-\pi}^{\pi}h_{\mu,p}(\lambda)dZ_{\xi^{(n)}+\eta^{(n)}}(\lambda)-\sum_{l=1}^{n}(-1)^l\binom{n}{l}(\xi(p-\mu l)+\eta(p-\mu l)) \tag{30}$$

of the value $\xi(p)$, $p\geq 0$, can be calculated by the formula

$$h_{\mu,p}(\lambda)=\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}\sum_{k=0}^{p}d_\mu(p-k)e^{i\lambda k}-\frac{e^{i\lambda p}(-i\lambda)^n g(\lambda)}{f(\lambda)+\lambda^{2n}g(\lambda)}$$
$$-\frac{(-i\lambda)^n\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}\mathbf{d}_{\mu,p}-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,p}\mathbf{a}_n)_k e^{i\lambda k}}{(1-e^{i\lambda\mu})^n(f(\lambda)+\lambda^{2n}g(\lambda))}, \tag{31}$$

where $\mathbf{d}_{\mu,p}=(d_\mu(p),d_\mu(p-1),d_\mu(p-2),\dots,d_\mu(0),0,\dots)'$, $\mathbf{T}_{\mu,p}$ is a linear operator in the space ℓ_2 defined by the matrix with elements $(\mathbf{T}_{\mu,p})_{l,k}=T^{\mu}_{l,p+\mu k}$, $l\geq 0$, $0\leq k\leq n$, and $(\mathbf{T}_{\mu,p})_{l,k}=0$, $l\geq 0$, $k>n$. The mean-square error of the estimate is calculated by the formula

$$\Delta(f,g;\widehat{\xi}(p))=\Delta(f,g;\widehat{\eta}(p))=\mathsf{E}|\eta(p)-\widehat{\eta}(p)|^2$$
$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|e^{i\lambda p}(1-e^{i\lambda\mu})^n f(\lambda)-\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}\mathbf{d}_{\mu,p}-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,p}\mathbf{a}_n)_k e^{i\lambda k}\right|^2}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))^2}g(\lambda)d\lambda$$
$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|e^{i\lambda p}(1-e^{i\lambda\mu})^n\lambda^{2n}g(\lambda)+\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}\mathbf{d}_{\mu,p}-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,p}\mathbf{a}_n)_k e^{i\lambda k}\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))^2}f(\lambda)d\lambda$$
$$=\langle\mathbf{d}_{\mu,p}-\mathbf{T}_{\mu,p}\mathbf{a}_n,\mathbf{P}_\mu^{-1}\mathbf{d}_{\mu,p}-\mathbf{P}_\mu^{-1}\mathbf{T}_{\mu,p}\mathbf{a}_n\rangle+Q_{0,0}. \tag{32}$$

Thus, we have the following statement.

Corollary 4
The optimal linear estimate $\widehat{\xi}(p)$ of the unknown value $\xi(p)$, $p\geq 0$, of a stochastic sequence with nth stationary increments from observations of the sequence $\xi(k)+\eta(k)$ at points $k=-1,-2,\dots$ is calculated by formula (30). The spectral characteristic $h_{\mu,p}(\lambda)$ of the optimal estimate $\widehat{\xi}(p)$ is calculated by formula (31). The value of the mean-square error $\Delta(f,g;\widehat{\xi}(p))$ is calculated by formula (32).

Theorems 4, 5 and Corollary 4 determine solutions of the extrapolation problem for the linear functionals $A\xi$, $A_N\xi$ and the value $\xi(p)$, $p\geq 0$, using the Fourier coefficients of the functions

$$\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))},\quad\frac{\lambda^{2n}g(\lambda)}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}.$$

However, the problem of finding the inverse operator $(\mathbf{P}_\mu)^{-1}$ to the operator $\mathbf{P}_\mu$ defined by the Fourier coefficients of the function $\dfrac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}$ is complicated in most cases.

Fortunately, the proposed formulas can be simplified under the assumption that the functions

$$\frac{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}{\lambda^{2n}}, \quad \frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}, \quad g(\lambda) \tag{33}$$

admit the canonical factorizations

$$\frac{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}{\lambda^{2n}} = \left|\sum_{k=0}^{\infty}\theta_\mu(k)e^{-i\lambda k}\right|^2, \tag{34}$$

$$\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))} = \left|\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right|^2, \tag{35}$$

$$g(\lambda) = \sum_{k=-\infty}^{\infty} g(k)e^{i\lambda k} = \left|\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right|^2. \tag{36}$$

Let $\mathbf{G}$ be a linear operator in the space ℓ_2 defined by the matrix with elements $(\mathbf{G})_{l,k} = g(l-k)$, $l,k \geq 0$. The following lemmas give us representations of the functionals $\mathbf{P}_\mu$, $\mathbf{T}_\mu$ and $\mathbf{G}$.

Lemma 2

Suppose that the functions $\dfrac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}$ and $g(\lambda)$ admit factorizations (35) and (36) respectively. Let linear operators Ψ_μ and Φ in the space ℓ_2 be defined by matrices $(\Psi_\mu)_{k,j} = \psi_\mu(k-j)$ and $(\Phi)_{k,j} = \phi(k-j)$ for $0 \leq j \leq k$, $(\Psi_\mu)_{k,j} = 0$ and $(\Phi)_{k,j} = 0$ for $j > k$, $k,j \geq 0$. Then

a) The function $\dfrac{g(\lambda)\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}$ admits the factorization

$$\frac{g(\lambda)\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))} = \left|\sum_{k=0}^{\infty}\upsilon_\mu(k)e^{-i\lambda k}\right|^2, \tag{37}$$

where

$$\upsilon_\mu(k) = \sum_{j=0}^{k}\psi_\mu(j)\phi(k-j) = \sum_{j=0}^{k}\phi(j)\psi_\mu(k-j).$$

b) The linear operator Υ_μ in the space ℓ_2 defined by the matrix $(\Upsilon_\mu)_{k,j} = \upsilon_\mu(k-j)$ for $0 \leq j \leq k$, $(\Upsilon_\mu)_{k,j} = 0$ for $j > k$, $k,j \geq 0$, admits the representation

$$\Upsilon_\mu = \Psi_\mu\Phi = \Phi\Psi_\mu. \tag{38}$$

Proof. Statement a) follows from the equalities

$$\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\left(\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right) = \sum_{j=0}^{\infty}\sum_{k=j}^{\infty}\psi_\mu(j)\phi(k-j)e^{-i\lambda k} = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k}\psi_\mu(j)\phi(k-j)\right)e^{-i\lambda k}.$$

Statement b) follows from the equalities

$$\upsilon_\mu(i-j) = \sum_{k=0}^{i-j}\psi_\mu(k)\phi(i-j-k) = \sum_{p=j}^{i}\phi(i-p)\psi_\mu(p-j) = (\Phi\Psi_\mu)_{i,j} = (\Psi_\mu\Phi)_{i,j}, \quad i \geq j. \quad \square$$

Lemma 3
Suppose that functions (33) admit factorizations (34) – (36). Let linear operators Ψ_μ and Υ_μ in the space ℓ_2 be defined as in Lemma 2 and a linear operator Θ_μ in the space ℓ_2 be defined by the matrix $(\Theta_\mu)_{k,j}=\theta_\mu(k-j)$ for $0\leq j\leq k$, $(\Theta_\mu)_{k,j}=0$ for $j>k$, $k,j\geq 0$. Then

a) Linear operators $\mathbf{P}_\mu$, $\mathbf{T}_\mu$ and $\mathbf{G}$ in the space ℓ_2 admit the factorizations $\mathbf{P}_\mu=\Psi'_\mu\overline{\Psi}_\mu$, $\mathbf{T}_\mu=\Upsilon'_\mu\overline{\Upsilon}_\mu$ and $\mathbf{G}=\Phi'\overline{\Phi}$.

b) An inverse operator $\mathbf{V}_\mu=(\mathbf{P}_\mu)^{-1}$ admits the factorization $(\mathbf{P}_\mu)^{-1}=\overline{\Theta}_\mu\Theta'_\mu$ and elements of the matrix which determines the operator $\mathbf{V}_\mu$ are calculated by the formula

$$V^\mu_{k,j}=\sum_{p=0}^{\min(k,j)}\overline{\theta}_\mu(k-p)\theta_\mu(j-p),\quad k,j\geq 0.$$

Proof. We give a proof of statement a) for the linear operator $\mathbf{P}_\mu$ only. Factorization (35) implies

$$\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}=\sum_{m=-\infty}^{\infty}P^\mu(m)e^{i\lambda m}=\left|\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right|^2$$

$$=\sum_{m=-\infty}^{-1}\sum_{k=-m}^{\infty}\overline{\psi}_\mu(k)\psi_\mu(k+m)e^{i\lambda m}+\sum_{m=0}^{\infty}\sum_{k=0}^{\infty}\psi_\mu(k)\overline{\psi}_\mu(k+m)e^{i\lambda m}.$$

Thus, $P^\mu(m)=\sum_{k=0}^{\infty}\psi_\mu(k)\overline{\psi}_\mu(k+m)$, $m\geq 0$, and $P^\mu(-m)=\overline{P^\mu(m)}$, $m\geq 0$. For $i\geq j$ we have the equalities

$$P^\mu_{i,j}=P^\mu(i-j)=\sum_{l=i}^{\infty}\psi_\mu(l-i)\overline{\psi}_\mu(l-j)=(\Psi'_\mu\overline{\Psi}_\mu)_{i,j}$$

and for $i<j$ we have the equalities

$$P^\mu_{i,j}=P^\mu(i-j)=\overline{P^\mu(j-i)}=\sum_{l=j}^{\infty}\overline{\psi}_\mu(l-j)\psi_\mu(l-i)=(\overline{\Psi}'_\mu\Psi)_{i,j}.$$

that prove statement a).

Statement 2) comes from the relation $\Psi_\mu\Theta_\mu=\Theta_\mu\Psi_\mu=I$, which we need to prove. From factorizations (34) and (35) one can obtain

$$1=\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\left(\sum_{k=0}^{\infty}\theta_\mu(k)e^{-i\lambda k}\right)=\sum_{j=0}^{\infty}\left(\sum_{k=0}^{j}\psi_\mu(k)\theta_\mu(j-k)\right)e^{-i\lambda j}.$$

These equalities imply the following ones:

$$\delta_{i,j}=\sum_{k=0}^{i-j}\psi_\mu(k)\theta_\mu(i-j-k)=\sum_{p=j}^{i}\theta_\mu(i-p)\psi_\mu(p-j)=(\Theta_\mu\Psi_\mu)_{i,j}=(\Psi_\mu\Theta_\mu)_{i,j},\quad\square$$

Lemma 4
Suppose that the function $g(\lambda)$ admits factorization (36). Let a linear operator $\mathbf{S}$ in the space ℓ_2 be defined by a matrix with elements $(\mathbf{S})_{k,j}=g(k+j)$, $k,j\geq 0$, and a linear operator $\mathbf{K}$ in the space ℓ_2 be defined by a matrix with elements $(\mathbf{K})_{k,j}=\phi(k+j)$, $k,j\geq 0$. Then the operators $\mathbf{S}$ and $\mathbf{K}$ admit the relation

$$\mathbf{S}=\overline{\mathbf{K}}\Phi=\Phi'\overline{\mathbf{K}},$$

where the linear operator Φ is defined in Lemma 2.

Proof. In the same way as in the proof of Lemma 3 a) we obtaine the relation $g(m)=\sum_{k=0}^{\infty}\phi(k)\overline{\phi}(k+m)$, $m\geq 0$. Thus $g(i+j)=\sum_{l=j}^{\infty}=\overline{\phi}(i+l)\phi(l-j)=(\overline{\mathbf{K}}\Phi)_{i,j}$, $i,j\geq 0$. Since the matrices $\mathbf{S}$ and $\mathbf{K}$ are symmetric, we have $\mathbf{S}=\mathbf{S}'=\Phi'\overline{\mathbf{K}}$. □

Under the conditions of Lemma 2 and Lemma 3 on the spectral densities $f(\lambda)$ and $g(\lambda)$ formulas (22) and (23) can be simplified. These lemmas give us the factorizations of the linear functionals $\mathbf{T}_\mu$ and $\mathbf{P}_\mu^{-1}\mathbf{T}_\mu$:

$$\mathbf{T}_\mu=\Upsilon_\mu'\overline{\Upsilon}_\mu=\Psi_\mu'\Phi'\overline{\Phi\Psi}_\mu,$$

$$\mathbf{P}_\mu^{-1}\mathbf{T}_\mu=\overline{\Theta}_\mu\Theta_\mu'\Psi_\mu'\Phi'\overline{\Phi\Psi}_\mu=\overline{\Theta}_\mu\mathbf{G}\overline{\Psi}_\mu.$$

Denote $\mathbf{e}_\mu=\mathbf{G}\overline{\Psi}_\mu\mathbf{a}_\mu$. Factorization (35) allows us to make the following transformations:

$$\frac{\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k e^{i\lambda k}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}=\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\overline{\psi}_\mu(j)(\overline{\Theta}_\mu\mathbf{e}_\mu)_k e^{i\lambda(k+j)}$$

$$=\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\sum_{m=0}^{\infty}\sum_{p=0}^{m}\sum_{k=p}^{m}\overline{\psi}_\mu(m-k)\overline{\theta}_\mu(k-p)e_\mu(p)e^{i\lambda m}$$

$$=\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\sum_{m=0}^{\infty}\sum_{p=0}^{\infty}\delta_{m,p}e_\mu(p)e^{i\lambda m}=\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\sum_{m=0}^{\infty}e_\mu(m)e^{i\lambda m},$$

where $e_\mu(m)=(\mathbf{G}\overline{\Psi}_\mu\mathbf{a}_\mu)_m$, $m\geq 0$, is the mth element of the vector $\mathbf{e}_\mu=\mathbf{G}\overline{\Psi}_\mu\mathbf{a}_\mu$. Since

$$(\mathbf{G}\overline{\Psi}_\mu\mathbf{a}_\mu)_m=\sum_{j=0}^{\infty}\sum_{p=j}^{\infty}g(m-p)\overline{\psi}_\mu(p-j)a_\mu(j)=\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}g(m-j-l)\overline{\psi}_\mu(l)a_\mu(j),$$

the following equality holds true:

$$\frac{\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k e^{i\lambda k}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}=\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}g(m-j-l)\overline{\psi}_\mu(l)a_\mu(j)e^{i\lambda m}. \tag{39}$$

Using factorizations (35) and (36) we make the following transformations:

$$\frac{A_\mu(e^{i\lambda})\lambda^{2n}g(\lambda)}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}=\left|\sum_{k=0}^{\infty}\psi_\mu(k)e^{i\lambda k}\right|^2\sum_{j=0}^{\infty}\sum_{m=-\infty}^{\infty}g(m-j)a_\mu(j)e^{i\lambda m}$$

$$=\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\sum_{m=-\infty}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}g(m-j-l)\overline{\psi}_\mu(l)a_\mu(j)e^{i\lambda m}. \tag{40}$$

Equalities (39) and (40) let us rewrite expression (25) for the spectral characteristic $h_\mu^2(\lambda)$ of the optimal estimate $\widehat{A}\eta$ of the functional $A\eta$ as

$$h_\mu^2(\lambda)=\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}\left(\frac{A_\mu(e^{i\lambda})\lambda^{2n}g(\lambda)}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}-\frac{\lambda^{2n}\sum_{k=0}^{\infty}(\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu)_k e^{i\lambda k}}{|1-e^{i\lambda\mu}|^{2n}(f(\lambda)+\lambda^{2n}g(\lambda))}\right)$$

$$=\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}\left(\sum_{k=0}^{\infty}\psi_\mu(k)e^{-i\lambda k}\right)\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}\overline{g}(m+j+l)\overline{\psi}_\mu(l)a_\mu(j)e^{-i\lambda m}$$

$$= \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} \left(\sum_{k=0}^{\infty} \psi_\mu(k) e^{-i\lambda k} \right) \sum_{m=1}^{\infty} (\mathbf{C}_\mu \overline{\psi}_\mu)_m e^{-i\lambda m},$$

where $(\mathbf{C}_\mu \overline{\psi}_\mu)_m$, $m \geq 1$, is the mth element of the vector $\mathbf{C}_\mu \overline{\psi}_\mu = \overline{\Psi}'_\mu \overline{\mathbf{S}} \mathbf{a}_\mu$, $\psi_\mu = (\psi_\mu(0), \psi_\mu(1), \psi_\mu(2), \ldots)'$, $\mathbf{C}_\mu$ is a linear operator defined by a matrix with elements $(\mathbf{C}_\mu)_{k,j} = \mathbf{c}_\mu(k+j)$, $k, j \geq 0$. Here $\mathbf{c}_\mu = \overline{\mathbf{S}} \mathbf{a}_\mu$ is a vector, $\mathbf{S}$ is a linear operator defined by a matrix with elements $(\mathbf{S})_{k,j} = g(k+j)$, $k, j \geq 0$. From Lemma 4 the operator $\mathbf{S}$ admits representation $\mathbf{S} = \overline{\mathbf{K}} \Phi = \Phi' \overline{\mathbf{K}}$, where $\mathbf{K}$ is a linear operator defined by a matrix with elements $(\mathbf{K})_{k,j} = \phi(k+j)$, $k, j \geq 0$.

The spectral characteristic $h^1_\mu(\lambda)$ of the optimal estimate $\widehat{B}\xi$ of the functional $B\xi$ in the case where spectral densities admit canonical factorization (34) is of the form

$$h^1_\mu(\lambda) = B_\mu(e^{i\lambda}) \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} - \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} \left(\sum_{k=0}^{\infty} \psi_\mu(k) e^{-i\lambda k} \right) \sum_{m=0}^{\infty} (D^\mu \mathbf{A} \theta_\mu)_m e^{i\lambda m}$$

$$= \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} \left(\sum_{k=0}^{\infty} \psi_\mu(k) e^{-i\lambda k} \right) \sum_{m=1}^{\infty} ((\widetilde{\mathbf{B}}_\mu)' \theta_\mu)_m e^{-i\lambda m},$$

where $\theta_\mu = (\theta_\mu(0), \theta_\mu(1), \theta_\mu(2), \ldots)'$; $\mathbf{A}$ is a linear operator defined by a matrix with elements $(\mathbf{A})_{k,j} = a(k+j)$, $k, j \geq 0$; $\widetilde{\mathbf{B}}_\mu$ is a linear operator defined by the matrix with elements $(\widetilde{\mathbf{B}}_\mu)_{k,j} = \mathbf{b}_\mu(k-j)$ for $0 \leq j \leq k$, $(\widetilde{\mathbf{B}}_\mu)_{k,j} = 0$ for $j > k$, $k, j \geq 0$, $\mathbf{b}_\mu = D^\mu \mathbf{a}$. This representation of the spectral characteristic $h^1_\mu(\lambda)$ shows that the spectral characteristic $h_\mu(\lambda)$ of the estimate $\widehat{A}\xi$ can be calculated by the formula

$$h_\mu(\lambda) = \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} \left(\sum_{m=1}^{\infty} ((\widetilde{\mathbf{B}}_\mu)' \theta_\mu - \mathbf{C}_\mu \overline{\psi}_\mu)_m e^{-i\lambda m} \right) \sum_{k=0}^{\infty} \psi_\mu(k) e^{-i\lambda k}$$

$$= B_\mu(e^{i\lambda}) \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} - \widetilde{h}_\mu(\lambda), \tag{41}$$

$$\widetilde{h}_\mu(\lambda) = \frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n} \left(\sum_{m=1}^{\infty} (\mathbf{C}_\mu \overline{\psi}_\mu)_m e^{-i\lambda m} + \sum_{m=0}^{\infty} (D^\mu \mathbf{A} \theta_\mu)_m e^{i\lambda m} \right) \sum_{k=0}^{\infty} \psi_\mu(k) e^{-i\lambda k},$$

The mean square error of the estimate $\Delta(f, g; \widehat{A}\xi)$ is presented as follows:

$$\Delta(f, g; \widehat{A}\xi) = \Delta(f, g; \widehat{H}\xi) = \mathsf{E}|H\xi - \widehat{H}\xi|^2$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} |A(e^{i\lambda})|^2 g(\lambda) d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} |\widetilde{h}_\mu(e^{i\lambda})|^2 (f(\lambda) + \lambda^{2n} g(\lambda)) d\lambda$$

$$- \frac{1}{2\pi} \int_{-\pi}^{\pi} \widetilde{h}_\mu(e^{i\lambda}) \overline{A(e^{i\lambda})} (i\lambda)^n g(\lambda) d\lambda - \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{\widetilde{h}_\mu(e^{i\lambda})} A(e^{i\lambda}) (-i\lambda)^n g(\lambda) d\lambda$$

$$= \langle \overline{\Phi} \mathbf{a}, \overline{\Phi} \mathbf{a} \rangle + \langle \Theta'_\mu D^\mu \mathbf{a} - \mathbf{G} \overline{\Psi}_\mu \mathbf{a}_\mu, \Theta'_\mu D^\mu \mathbf{a} \rangle - \langle \overline{\Psi}'_\mu \overline{\Phi}' \mathbf{K} \mathbf{a}_\mu, \overline{\Psi}'_\mu \overline{\Phi}' \mathbf{K} \mathbf{a}_\mu \rangle_1 - \langle \overline{\Phi} \Theta'_\mu D^\mu \mathbf{a}, \overline{\Phi \Psi}_\mu \mathbf{a}_\mu \rangle$$

$$= \langle \mathbf{G} \mathbf{a}, \mathbf{a} \rangle + \langle D^\mu \mathbf{A} \theta_\mu - \mathbf{G} \widetilde{\mathbf{A}}_\mu \overline{\psi}_\mu, D^\mu \mathbf{A} \theta_\mu \rangle - \langle \mathbf{C}_\mu \overline{\psi}_\mu, \mathbf{C}_\mu \overline{\psi}_\mu \rangle_1 - \langle \mathbf{G} D^\mu \mathbf{A} \theta_\mu, \widetilde{\mathbf{A}}_\mu \overline{\psi}_\mu \rangle, \tag{42}$$

where $\widetilde{\mathbf{A}}_\mu$ is a linear operator in the space ℓ_2 defined as $(\widetilde{\mathbf{A}}_\mu)_{k,j} = a_\mu(k-j)$ for $0 \leq j \leq k$, $(\widetilde{\mathbf{A}}_\mu)_{k,j} = 0$ for $j > k$, $k, j \geq 0$; $\langle x, y \rangle_1 = \sum_{k=1}^{\infty} x(k) \overline{y}(k)$ for vectors $x = (x(0), x(1), x(2), \ldots)'$, $y = (y(0), y(1), y(2), \ldots)'$.

The obtained results are summarized in the following theorem.

Theorem 6
Let $\{\xi(m), m \in \mathbb{Z}\}$ be a stochastic sequence which defines a stationary nth increment sequence $\xi^{(n)}(m,\mu)$ with absolutely continuous spectral function $F(\lambda)$ which has spectral density $f(\lambda)$. Let $\{\eta(m), m \in \mathbb{Z}\}$ be an uncorrelated with the sequence $\xi(m)$ stationary stochastic sequence with absolutely continuous spectral function $G(\lambda)$ which has spectral density $g(\lambda)$. Suppose that spectral densities $f(\lambda)$ and $g(\lambda)$ admit canonical factorizations (35) – (36). Suppose also that coefficients $\{a(k) : k \geq 0\}$ satisfy conditions (13) – (14). Then the spectral characteristic $h_\mu(\lambda)$ of the optimal estimate $\widehat{A}\xi$ of the functional $A\xi$ which depend on the unknown values of elements $\xi(m)$, $m \geq 0$, based on observations of the sequence $\xi(m) + \eta(m)$ at points $m = -1, -2, \ldots$ can be calculated by formula (41). The value of the mean-square error $\Delta(f, g; \widehat{A}\xi)$ can be calculated by formula (42).

Remark 1
Since

$$\int_{-\pi}^{\pi} \left| \ln \frac{|1 - e^{-i\lambda\mu}|^{2n}}{\lambda^{2n}} \right| d\lambda < \infty$$

for every $n \geq 1$ and $\mu \geq 1$, there exists a function $w_\mu(z) = \sum_{k=0}^{\infty} w_\mu(k) z^k$ such that $\sum_{k=0}^{\infty} |w_\mu(k)|^2 < \infty$, $\dfrac{|1 - e^{-i\lambda\mu}|^{2n}}{\lambda^{2n}} = |w_\mu(e^{-i\lambda})|^2$ (see, for example, [4] for details). The function $w_\mu(z)$ can be calculated with the help of the relation

$$w_\mu(z) = \exp\left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{e^{i\lambda} + z}{e^{i\lambda} - z} \ln \frac{|1 - e^{-i\lambda\mu}|^{2n}}{\lambda^{2n}} d\lambda \right\}. \tag{43}$$

Provided factorization (34) or factorization (35) take place, the function $f(\lambda) + \lambda^{2n} g(\lambda)$ admits the canonical factorization

$$f(\lambda) + \lambda^{2n} g(\lambda) = \left| \sum_{k=0}^{\infty} \theta(k) e^{-i\lambda k} \right|^2 = \left| \sum_{k=0}^{\infty} \psi(k) e^{-i\lambda k} \right|^{-2}. \tag{44}$$

Let linear operators Θ, Ψ and $\mathbf{W}_\mu$ in the space ℓ_2 be defined as $(\Theta)_{k,j} = \psi(k-j)$, $(\Psi)_{k,j} = \psi(k-j)$ and $(\mathbf{W}_\mu)_{k,j} = \mathbf{W}_\mu(k-j)$ for $0 \leq j \leq k$, $(\Theta)_{k,j} = 0$, $(\Psi)_{k,j} = 0$ and $(\mathbf{U}_\mu)_{k,j} = 0$ for $j > k$, $k, j \geq 0$. Let $\mathbf{U}_\mu = \mathbf{W}_\mu^{-1}$. Then the operators Θ_μ, Ψ_μ and Θ, Ψ are connected by the relations

$$\Theta_\mu = \Theta \mathbf{W}_\mu = \mathbf{W}_\mu \Theta, \quad \Psi_\mu = \Psi \mathbf{U}_\mu = \mathbf{U}_\mu \Psi,$$

which is obtained in the same way as relation (38) in Lemma 2. What is more,

$$\theta_\mu = \mathbf{W}_\mu \theta, \quad \psi_\mu = \mathbf{U}_\mu \psi, \tag{45}$$

where $\theta = (\theta(0), \theta(1), \theta(2), \ldots)'$, $\psi = (\psi(0), \psi(1), \psi(2), \ldots)'$.

Example 1
Consider an ARIMA(0,1,1) sequence $\{\xi(m), m \in \mathbb{Z}\}$. In this case the first order increments $\xi^{(1)}(m,\mu)$ are stationary and increments $\xi^{(1)}(m,1)$ with step $\mu = 1$ form an one-sided moving average stochastic sequence of order 1 with parameter ω, $-1 < \omega < 1$. The sequence $\xi(m)$ has the spectral density

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

Let $\{\eta(m), m \in \mathbb{Z}\}$ be an uncorrelated with $\xi(m)$ moving average stochastic sequence of order 1 with parameter ϕ, $-1 < \phi < 1$, and spectral density $g(\lambda) = |1 - \phi e^{-i\lambda}|^2$. Then the stochastic sequence $\{\xi(m) + \eta(m), m \in \mathbb{Z}\}$ is an ARIMA(0,1,2) sequence with the spectral density

$$f(\lambda) + \lambda^2 g(\lambda) = \frac{x\lambda^2 |1 - ye^{-i\lambda} - ze^{-2i\lambda}|^2}{|1 - e^{-i\lambda}|^2},$$

where

$$x=\frac{\phi}{t},\quad y=\frac{t}{t-1}\frac{(1+\phi)^2-\omega}{\phi},\quad z=t,$$

where t is a solution of the equation

$$ct^3-(2c+3)t^2+(c+2)t-1=0,$$

$$c=\frac{1}{\phi}(\omega^2+\phi^2-(1+\phi)^2).$$

Suppose that $t=\phi$. This assumption holds true if the parameters ϕ and ω satisfy the equation

$$(1-\phi)^2(\omega^2-2\phi-2)=2\phi^2.$$

In this case we have

$$x=1,\quad y=\frac{\omega-(1+\phi)^2}{1-\phi},\quad z=\phi.$$

Consider the problem of finding the mean square optimal linear estimate of the functional $A_1\xi=a\xi(0)+b\xi(1)$ which depends on unknown values $\xi(0)$, $\xi(1)$ of the sequence $\xi(m)$ from observations of the sequence $\xi(m)+\eta(m)$ at points $m=-1,-2,\dots$. To calculate the spectral characteristic of the optimal estimate $\widehat{A}_1\xi$ of the functional $A_1\xi$ we use formula (41). Coefficients $\phi(k)$, $k\geq 0$, from factorization (36) are the following: $\phi(0)=1$, $\phi(1)=-\phi$, $\phi(i)=0$ for $l\geq 2$. Thus the operator $\mathbf{S}$ is determined by the matrix with elements $(\mathbf{S})_{0,0}=1+\phi^2$, $(\mathbf{S})_{0,1}=(\mathbf{S})_{1,0}=-\phi$ and $(\mathbf{S})_{k,j}=0$ otherwise. Coefficients $a_\mu(k)=a_1(k)$, $k\geq 0$, are the following: $a_1(0)=a$, $a_1(1)=b-a$, $a_1(2)=-b$ and $a_1(l)=0$ for $l\geq 3$. Thus the operator $\mathbf{C}_\mu=\mathbf{C}_1$ is determined by the matrix with elements $(\mathbf{C}_\mu)_{0,0}=a(1+\phi+\phi^2)-b\phi$, $(\mathbf{C}_\mu)_{0,1}=(\mathbf{C}_\mu)_{1,0}=-a\phi$ and $(\mathbf{C}_\mu)_{k,j}=0$ otherwise. Coefficients $\psi_\mu(k)=\psi_1(k)$, $k\geq 0$, from factorization (35) are found using the equality $\Psi_1\Theta_1=I$ from the proof of Lemma 3 putting $\theta_1(0)=1$, $\theta_1(1)=-x$, $\theta_1(2)=-y$ and $\theta_1(l)=0$ for $l\geq 3$:

$$\psi_1(0)=1,\quad \psi_1(1)=y,\quad \psi_1(l)=y\psi_1(l-1)+z\psi_1(l-2)\text{ for }l\geq 2.$$

Coefficients $b_1(k)$, $k\geq 0$, are: $b_1(0)=a+b$, $b_1(1)=b$, $b_1(l)=0$, $l\geq 2$. Thus the vector $(\widetilde{\mathbf{B}}_\mu)'\theta_\mu=((a+b)-bx,-x(a+b)-by,-y(a+b),0,\dots)'$. Finally, the spectral characteristic of the estimate $\widehat{A}_1\xi$ is calculated by the formula

$$h_{1,1}(\lambda)=\frac{1-e^{-i\lambda\mu}}{i\lambda}\left((a\psi-x(a+b)-by)e^{-i\lambda}-y(a+b)e^{-2i\lambda}\right)\sum_{k=0}^{\infty}\psi_1(k)e^{-i\lambda k}$$

$$=\frac{1-e^{-i\lambda\mu}}{i\lambda}\sum_{k=1}^{\infty}s(k)e^{-i\lambda k},$$

where

$$s(1)=(a\psi-x(a+b)-by)\psi_1(0),$$

$$s(k)=(a\psi-x(a+b)-by)\psi_1(k-1)-y(a+b)\psi_1(k-2),\quad k\geq 2.$$

The optimal estimate $\widehat{A}_1\xi$ of the functional $A_1\xi$ is calculated by the formula

$$\widehat{A}_1\xi=(a+b)(\xi(-1)+\eta(-1))+\sum_{k=1}^{\infty}s(k)(\xi^{(1)}(-k,1)+\eta^{(1)}(-k,1))$$

$$=(a+b+s(1))(\xi(-1)+\eta(-1))+\sum_{k=2}^{\infty}(s(k)-s(k-1))(\xi(-k)+\eta(-k)).$$

4. Extrapolation of cointegrated sequences

Let $\{\xi(m), m \in \mathbb{Z}\}$ and $\{\zeta(m), m \in \mathbb{Z}\}$ be two integrated stochastic sequences which define stationary nth increment sequences $\xi^{(n)}(m,\mu)$ and $\zeta^{(n)}(m,\mu)$ with absolutely continuous spectral functions $F(\lambda)$ and $P(\lambda)$ and spectral densities $f(\lambda)$ and $p(\lambda)$ correspondingly.

Definition 4
Two integrated stochastic sequences $\{\xi(m), m \in \mathbb{Z}\}$ and $\{\zeta(m), m \in \mathbb{Z}\}$ are called cointegrated if there exists a constant $\beta \neq 0$ such that the stochastic sequence $\{\zeta(m) - \beta\xi(m), m \in \mathbb{Z}\}$ is stationary.

The extrapolation problem for cointegrated stochastic sequences means that we have to find the mean-square optimal linear estimates of the functionals

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

which depend on the unknown values of the sequence $\xi(m)$ based on observations of the sequence $\zeta(m)$ at points $m = -1, -2, \ldots$. To find a solution to this problem we can use results presented in the preceding section provided the sequences $\xi(m)$ and $\zeta(m) - \beta\xi(m)$ are uncorrelated.

Let the following condition holds true:

$$\int_{-\pi}^{\pi} \frac{\lambda^{2n}}{|1 - e^{i\lambda\mu}|^{2n} p(\lambda)} d\lambda < \infty. \tag{46}$$

Then we can determine operators $\mathbf{P}_\mu^\beta$, $\mathbf{T}_\mu^\beta$, $\mathbf{Q}^\beta$ with the help of the Fourier coefficients

$$P_{k,j}^{\mu,\beta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\lambda(j-k)} \frac{\lambda^{2n}}{|1 - e^{i\lambda\mu}|^{2n} p(\lambda)} d\lambda;$$

$$T_{k,j}^{\mu,\beta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\lambda(j-k)} \frac{\lambda^{2n}(p(\lambda) - \beta^2 f(\lambda))}{|1 - e^{i\lambda\mu}|^{2n} p(\lambda)} d\lambda;$$

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

of the functions

$$\frac{\lambda^{2n}}{|1 - e^{i\lambda\mu}|^{2n} p(\lambda)}, \quad \frac{p(\lambda) - \beta^2 f(\lambda)}{|1 - e^{i\lambda\mu}|^{2n} p(\lambda)}, \quad \frac{f(\lambda)p(\lambda) - \beta^2 f^2(\lambda)}{\lambda^{2n} p(\lambda)} \tag{47}$$

in the same way as we defined operators $\mathbf{P}_\mu$, $\mathbf{T}_\mu$, $\mathbf{Q}$ in Section 3. It follows from Theorem 4 that the spectral characteristic $h_\mu^\beta(\lambda)$ of the optimal estimate

$$\widehat{A}\xi = \int_{-\pi}^{\pi} h_\mu^\beta(\lambda) dZ_{\zeta^{(n)}}(\lambda) - V_\mu \zeta, \tag{48}$$

of the functional $A\xi$ is calculated by the formula

$$h_\mu^\beta(\lambda) = B^\mu(e^{i\lambda}) \frac{(1 - e^{-i\lambda\mu})^n}{(i\lambda)^n} - A(e^{i\lambda}) \frac{p(\lambda) - \beta^2 f(\lambda)}{(i\lambda)^n p(\lambda)} - \frac{(-i\lambda)^n C_\mu^\beta(e^{i\lambda})}{(1 - e^{i\lambda\mu})^n p(\lambda)}. \tag{49}$$

where

$$C_\mu^\beta(e^{i\lambda}) = \sum_{k=0}^{\infty} ((\mathbf{P}_\mu^\beta)^{-1} D^\mu \mathbf{a} - (\mathbf{P}_\mu^\beta)^{-1} \mathbf{T}_\mu^\beta \mathbf{a}_\mu)_k e^{i\lambda k}.$$

The mean-square error of the estimate is calculated by the formula

$$\Delta(f,g;\widehat{A}\xi)$$
$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n\beta^2 f(\lambda)-\lambda^{2n}C^{\beta}_{\mu}(e^{i\lambda})\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}p^2(\lambda)}p(\lambda)d\lambda$$
$$-\frac{\beta^2}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n\beta^2 f(\lambda)-\lambda^{2n}C^{\beta}_{\mu}(e^{i\lambda})\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}p^2(\lambda)}f(\lambda)d\lambda$$
$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n(p(\lambda)-\beta^2 f(\lambda))+\lambda^{2n}C^{\beta}_{\mu}(e^{i\lambda})\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}p^2(\lambda)}f(\lambda)d\lambda$$
$$=\langle D^{\mu}\mathbf{a}-\mathbf{T}^{\beta}_{\mu}\mathbf{a}_{\mu},(\mathbf{P}^{\beta}_{\mu})^{-1}D^{\mu}\mathbf{a}-(\mathbf{P}^{\beta}_{\mu})^{-1}\mathbf{T}^{\beta}_{\mu}\mathbf{a}_{\mu}\rangle+\langle\mathbf{Q}^{\beta}\mathbf{a},\mathbf{a}\rangle, \tag{50}$$

Now we can summarize the obtained results in the following statement.

Theorem 7
Let the cointegrated stochastic sequences $\{\xi(m),m\in\mathbb{Z}\}$ and $\{\zeta(m),m\in\mathbb{Z}\}$ have absolutely continuous spectral functions $F(\lambda)$ and $G(\lambda)$ with the spectral densities $f(\lambda)$ and $p(\lambda)$. Let the spectral density $p(\lambda)$ satisfy condition (46) and let coefficients $\{a(k):k\geq 0\}$ satisfy conditions (13) – (14). If the sequences $\xi(m)$ and $\zeta(m)-\beta\xi(m)$ are uncorrelated, then the optimal linear estimate $\widehat{A}\xi$ of the functional $A\xi$ which depend on the unknown values of elements $\xi(m)$, $m\geq 0$, based on observations of the sequence $\zeta(m)$ at points $m=-1,-2,\dots$ is calculated by formula (48). The spectral characteristic $h^{\beta}_{\mu}(\lambda)$ of the optimal estimate $\widehat{A}\xi$ is calculated by formula (49). The value of the mean-square error $\Delta(f,g;\widehat{A}\xi)$ is calculated by formula (50).

Define operators $\mathbf{P}^{\beta}_{\mu}$, $\mathbf{T}^{\beta}_{\mu,N}$, $\mathbf{Q}^{\beta}_{N}$ determined by the Fourier coefficients of functions (47) in the same way as we defined operators $\mathbf{P}_{\mu}$, $\mathbf{T}_{\mu,N}$, $\mathbf{Q}_{N}$ in Section 3. Theorem 5 implies that the spectral characteristic $h^{\beta}_{\mu,N}(\lambda)$ of the optimal estimate

$$\widehat{A}_N\xi=\int_{-\pi}^{\pi}h^{\beta}_{\mu,N}(\lambda)dZ_{\zeta^{(n)}}(\lambda)-V_{\mu,N}\zeta \tag{51}$$

of the functional $A_N\xi$ is calculated by the formula

$$h^{\beta}_{\mu,N}(\lambda)=B^{\mu}_{N}(e^{i\lambda})\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}-A_N(e^{i\lambda})\frac{p(\lambda)-\beta^2 f(\lambda)}{(i\lambda)^n p(\lambda)}-\frac{(-i\lambda)^n C^{\beta}_{\mu,N}(e^{i\lambda})}{(1-e^{i\lambda\mu})^n p(\lambda)}, \tag{52}$$

where $C^{\beta}_{\mu,N}(e^{i\lambda})=\sum_{k=0}^{\infty}((\mathbf{P}^{\beta}_{\mu})^{-1}D^{\mu}_{N}\mathbf{a}_N-(\mathbf{P}^{\beta}_{\mu})^{-1}\mathbf{T}^{\beta}_{\mu,N}\mathbf{a}_{\mu,N})_k e^{i\lambda k}$. The mean-square error of the estimate $\widehat{A}_N\xi$ is calculated by the formula

$$\Delta(f,g;\widehat{A}_N\xi)$$
$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A_N(e^{i\lambda})(1-e^{i\lambda\mu})^n\beta^2 f(\lambda)-\lambda^{2n}C^{\beta}_{\mu,N}(e^{i\lambda})\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}p^2(\lambda)}p(\lambda)d\lambda$$
$$-\frac{\beta^2}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A_N(e^{i\lambda})(1-e^{i\lambda\mu})^n\beta^2 f(\lambda)-\lambda^{2n}C^{\beta}_{\mu,N}(e^{i\lambda})\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}p^2(\lambda)}f(\lambda)d\lambda$$
$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|A_N(e^{i\lambda})(1-e^{i\lambda\mu})^n(p(\lambda)-\beta^2 f(\lambda))+\lambda^{2n}C^{\beta}_{\mu,N}(e^{i\lambda})\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}p^2(\lambda)}f(\lambda)d\lambda$$
$$=\langle D^{\mu}_{N}\mathbf{a}_N-\mathbf{T}^{\beta}_{\mu,N}\mathbf{a}_{\mu,N},(\mathbf{P}^{\beta}_{\mu})^{-1}D^{\mu}_{N}\mathbf{a}_N-(\mathbf{P}^{\beta}_{\mu})^{-1}\mathbf{T}^{\beta}_{\mu,N}\mathbf{a}_{\mu,N}\rangle+\langle\mathbf{Q}^{\beta}_{N}\mathbf{a}_N,\mathbf{a}_N\rangle. \tag{53}$$

The following theorem holds true.

Theorem 8
Let the cointegrated stochastic sequences $\{\xi(m), m\in\mathbb{Z}\}$ and $\{\zeta(m), m\in\mathbb{Z}\}$ have absolutely continuous spectral functions $F(\lambda)$ and $G(\lambda)$ with the spectral densities $f(\lambda)$ and $p(\lambda)$. Let the spectral density $p(\lambda)$ satisfy condition (46). If the sequences $\xi(m)$ and $\zeta(m)-\beta\xi(m)$ are uncorrelated, then the optimal linear estimate $\widehat{A}_N\xi$ of the functional $A_N\xi$ which depend on the unknown values of elements $\xi(m)$, $0\leq m\leq N$, based on observations of the sequence $\zeta(m)$ at points $m=-1,-2,\dots$ is calculated by formula (51). The spectral characteristic $h^{\beta}_{\mu,N}(\lambda)$ of the optimal estimate $\widehat{A}_N\xi$ is calculated by formula (52). The value of the mean-square error $\Delta(f,g;\widehat{A}_N\xi)$ is calculated by formula (53).

Suppose that spectral densities $f(\lambda)$ and $p(\lambda)$ admit the following canonical factorizations:

$$\frac{|1-e^{i\lambda\mu}|^{2n}p(\lambda)}{\lambda^{2n}}=\left|\sum_{k=0}^{\infty}\theta^{\beta}_{\mu}(k)e^{-i\lambda k}\right|^2=\left|\sum_{k=0}^{\infty}\psi^{\beta}_{\mu}(k)e^{-i\lambda k}\right|^{-2}, \tag{54}$$

$$p(\lambda)=\left|\sum_{k=0}^{\infty}\theta^{\beta}(k)e^{-i\lambda k}\right|^2=\left|\sum_{k=0}^{\infty}\psi^{\beta}(k)e^{-i\lambda k}\right|^{-2}, \tag{55}$$

$$p(\lambda)-\beta^2f(\lambda)=\left|\sum_{k=0}^{\infty}\phi^{\beta}(k)e^{-i\lambda k}\right|^2. \tag{56}$$

Define operators $\mathbf{K}^{\beta}$, Ψ^{β} and Φ^{β} by coefficients of the canonical factorizations (54)-(56) in the same way as we defined operators $\mathbf{K}$, Ψ and Φ in Section 3. It follows from Theorem 6 that the spectral characteristic $h^{\beta}_{\mu}(\lambda)$ of the optimal estimate $\widehat{A}\xi$ of the functional $A\xi$ is calculated by the formula

$$h^{\beta}_{\mu}(\lambda)=\frac{(1-e^{-i\lambda\mu})^n}{(i\lambda)^n}\left(\sum_{m=1}^{\infty}((\widetilde{\mathbf{B}}_{\mu})'\theta^{\beta}_{\mu}-\mathbf{C}^{\beta}_{\mu}\overline{\psi}^{\beta}_{\mu})_m e^{-i\lambda m}\right)\sum_{k=0}^{\infty}\psi^{\beta}_{\mu}(k)e^{-i\lambda k}, \tag{57}$$

where $\mathbf{C}^{\beta}_{\mu}\overline{\psi^{\beta}_{\mu}}=\overline{\mathbf{U}}'_{\mu}\overline{\Psi^{\beta}}'\overline{\Phi^{\beta}}'\mathbf{K}^{\beta}\mathbf{a}_{\mu}$. The mean-square error of the estimate is calculated by the formula

$$\Delta(f,g;\widehat{A}\xi)=\langle\mathbf{G}^{\beta}\mathbf{a},\mathbf{a}\rangle+\langle D^{\mu}\mathbf{A}\theta^{\beta}_{\mu}-\mathbf{G}^{\beta}\mathbf{A}_{\mu}\overline{\psi}^{\beta}_{\mu},D^{\mu}\mathbf{A}\theta^{\beta}_{\mu}\rangle-\langle\mathbf{C}^{\beta}_{\mu}\overline{\psi}^{\beta}_{\mu},\mathbf{C}^{\beta}_{\mu}\overline{\psi}^{\beta}_{\mu}\rangle_1-\langle\mathbf{G}^{\beta}D^{\mu}\mathbf{A}\theta^{\beta}_{\mu},\mathbf{A}_{\mu}\overline{\psi}^{\beta}_{\mu}\rangle. \tag{58}$$

Theorem 9
Let the cointegrated stochastic sequences $\{\xi(m), m\in\mathbb{Z}\}$ and $\{\zeta(m), m\in\mathbb{Z}\}$ satisfy conditions of Theorem 7. If spectral densities $f(\lambda)$ and $p(\lambda)$ admit factorizations (54) – (56), then the spectral characteristic $h^{\beta}_{\mu}(\lambda)$ of the optimal linear estimate $\widehat{A}\xi$ of the functional $A\xi$ which depend on the unknown values of elements $\xi(m)$, $m\geq 0$, based on observations of the sequence $\zeta(m)$ at points $m=-1,-2,\dots$ can be calculated by formula (57). The value of the mean-square error $\Delta(f,g;\widehat{A}\xi)$ is calculated by formula (58).

5. Minimax-robust method of extrapolation

The values of the mean-square errors $\Delta(h_{\mu}(f,g);f,g):=\Delta(f,g;\widehat{A}\xi)$ and $\Delta(h_{\mu,N}(f,g);f,g):=\Delta(f,g;\widehat{A}_N\xi)$ and the spectral characteristics $h_{\mu}(f,g)$ and $h_{\mu,N}(f,g)$ of the optimal linear estimates $\widehat{A}\xi$ and $\widehat{A}_N\xi$ of the functionals $A\xi$ and $A_N\xi$ which depend on the unknown values of the sequence $\xi(m)$ based on observations of the stochastic sequence $\xi(k)+\eta(k)$ can be calculated by formulas (23), (22) and (29), (27) correspondingly under the condition that spectral densities $f(\lambda)$ and $g(\lambda)$ of stochastic sequences $\xi(m)$ and $\eta(m)$ are exactly known. Having canonical factorizations (35) and (36) we can calculate the values of mean-square errors $\Delta(h_{\mu}(f,g);f,g)$ and spectral characteristics $h_{\mu}(f,g)$ by formulas (42), (41) respectively. However, such situation does not appear in practice since we do not know exactly spectral densities of the observed sequences. If in this case we can determine a set $\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$ of admissible spectral densities, the minimax (robust) approach to estimation of linear functionals which depend on the unknown values of stochastic sequence with stationary increments can be applied. It consists in finding an estimate that minimizes the maximum value of the mean-square errors for all spectral densities from a given class $\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$ of admissible spectral densities simultaneously.

To formalize this approach we present the following definitions.

Definition 5
For a given class of spectral densities $\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$ spectral densities $f_0(\lambda)\in\mathcal{D}_f$, $g_0(\lambda)\in\mathcal{D}_g$ are called least favorable in the class $\mathcal{D}$ for the optimal linear extrapolation of the functional $A\xi$ if the following relation holds true:

$$\Delta(f_0,g_0)=\Delta(h(f_0,g_0);f_0,g_0)=\max_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g}\Delta(h(f,g);f,g).$$

Definition 6
For a given class of spectral densities $\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$ the spectral characteristic $h^0(\lambda)$ of the optimal linear estimate of the functional $A\xi$ is called minimax-robust if there are satisfied the conditions

$$h^0(\lambda)\in H_{\mathcal{D}}=\bigcap_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g}L_2^{0-}(f(\lambda)+\lambda^{2n}g(\lambda)),$$

$$\min_{h\in H_{\mathcal{D}}}\max_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g}\Delta(h;f,g)=\max_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g}\Delta(h^0;f,g).$$

Using the derived in the previous sections formulas and the introduced definitions we can conclude that the following statements hold true.

Lemma 5
Spectral densities $f^0\in\mathcal{D}_f$, $g^0\in\mathcal{D}_g$ which satisfy condition (15) are least favorable in the class $\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$ for the optimal linear extrapolation of the functional $A\xi$ if operators $\mathbf{P}_\mu^0$, $\mathbf{T}_\mu^0$, $\mathbf{Q}^0$ determined by the Fourier coefficients of the functions

$$\frac{\lambda^{2n}}{|1-e^{i\lambda\mu}|^{2n}(f^0(\lambda)+\lambda^{2n}g^0(\lambda))},\quad \frac{\lambda^{2n}g^0(\lambda)}{|1-e^{i\lambda\mu}|^{2n}(f^0(\lambda)+\lambda^{2n}g^0(\lambda))},\quad \frac{f^0(\lambda)g^0(\lambda)}{f^0(\lambda)+\lambda^{2n}g^0(\lambda)}$$

determine a solution of the constraint optimisation problem

$$\max_{(f,g)\in\mathcal{D}_f\times\mathcal{D}_g}(\langle D^\mu\mathbf{a}-\mathbf{T}_\mu\mathbf{a}_\mu,\mathbf{P}_\mu^{-1}D^\mu\mathbf{a}-\mathbf{P}_\mu^{-1}\mathbf{T}_\mu\mathbf{a}_\mu\rangle)$$
$$=\langle D^\mu\mathbf{a}-\mathbf{T}^0_\mu\mathbf{a}_\mu,(\mathbf{P}^0_\mu)^{-1}D^\mu\mathbf{a}-(\mathbf{P}^0_\mu)^{-1}\mathbf{T}^0_\mu\mathbf{a}_\mu\rangle+\langle\mathbf{Q}^0\mathbf{a},\mathbf{a}\rangle. \tag{59}$$

The minimax spectral characteristic is determined as $h^0=h_\mu(f^0,g^0)$ if $h_\mu(f^0,g^0)\in H_{\mathcal{D}}$.

Lemma 6
Spectral densities $f^0\in\mathcal{D}_f$, $g^0\in\mathcal{D}_g$ which admit canonical factorizations (35) and (36) are least favorable in the class $\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$ for the optimal linear extrapolation of the functional $A\xi$ if coefficients $\{\theta^0(k),\psi^0(k),\phi^0(k):k\geq 0\}$ of factorizations

$$f^0(\lambda)+\lambda^{2n}g^0(\lambda)=\left|\sum_{k=0}^{\infty}\theta^0(k)e^{-i\lambda k}\right|^2=\left|\sum_{k=0}^{\infty}\psi^0(k)e^{-i\lambda k}\right|^{-2},\quad g^0(\lambda)=\left|\sum_{k=0}^{\infty}\phi^0(k)e^{-i\lambda k}\right|^2. \tag{60}$$

determine a solution to the constraint optimisation problem

$$\langle\mathbf{G}\mathbf{a},\mathbf{a}\rangle+\langle D^\mu\mathbf{A}\theta_\mu-\mathbf{G}\mathbf{A}_\mu\overline{\psi}_\mu,D^\mu\mathbf{A}\theta_\mu\rangle-\langle\mathbf{C}_\mu\overline{\psi}_\mu,\mathbf{C}_\mu\overline{\psi}_\mu\rangle_1-\langle\mathbf{G}D^\mu\mathbf{A}\theta_\mu,\mathbf{A}_\mu\overline{\psi}_\mu\rangle\to\sup, \tag{61}$$

$$f(\lambda)=\left|\sum_{k=0}^{\infty}\theta(k)e^{-i\lambda k}\right|^2-\lambda^{2n}\left|\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right|^2=\left|\sum_{k=0}^{\infty}\psi(k)e^{-i\lambda k}\right|^{-2}-\lambda^{2n}\left|\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right|^2\in\mathcal{D}_f,$$

$$g(\lambda)=\left|\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right|^2\in\mathcal{D}_g.$$

The minimax spectral characteristic is determined as $h^0=h_\mu(f^0,g^0)$ if $h_\mu(f^0,g^0)\in H_{\mathcal{D}}$.

Lemma 7
Spectral density $g^0\in\mathcal{D}_g$ which admit canonical factorization (36) with the known spectral density $f(\lambda)$ is least favorable in the class $\mathcal{D}_g$ for the optimal linear extrapolation of the functional $A\xi$ if coefficients $\{\theta^0(k),\psi^0(k),\phi^0(k):k\geq 0\}$ of factorizations

$$f(\lambda)+\lambda^{2n}g^0(\lambda)=\left|\sum_{k=0}^{\infty}\theta^0(k)e^{-i\lambda k}\right|^2=\left|\sum_{k=0}^{\infty}\psi^0(k)e^{-i\lambda k}\right|^{-2},\quad g^0(\lambda)=\left|\sum_{k=0}^{\infty}\phi^0(k)e^{-i\lambda k}\right|^2. \tag{62}$$

determine a solution to the constraint optimisation problem

$$\langle\mathbf{G}\mathbf{a},\mathbf{a}\rangle+\langle D^\mu\mathbf{A}\theta_\mu-\mathbf{G}\mathbf{A}_\mu\overline{\psi}_\mu,D^\mu\mathbf{A}\theta_\mu\rangle-\langle\mathbf{C}_\mu\overline{\psi}_\mu,\mathbf{C}_\mu\overline{\psi}_\mu\rangle_1-\langle\mathbf{G}D^\mu\mathbf{A}\theta_\mu,\mathbf{A}_\mu\overline{\psi}_\mu\rangle\to\sup, \tag{63}$$

$$g(\lambda)=\left|\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right|^2\in\mathcal{D}_g.$$

The minimax spectral characteristic is determined as $h^0=h_\mu(f,g^0)$ if $h_\mu(f,g^0)\in H_{\mathcal{D}}$.

If spectral density $g(\lambda)$ is known and admits canonical factorization (36), extremum problem (61) is an extremum problem with respect to variables $\{\theta(k),\psi(k):k\geq 0\}$.

Lemma 8
Spectral density $f^0\in\mathcal{D}_f$ which admit canonical factorization (35) with the known spectral density $g(\lambda)$ is least favorable in the class $\mathcal{D}_f$ for the optimal linear extrapolation of the functional $A\xi$ if coefficients $\{\theta^0(k),\psi^0(k):k\geq 0\}$ of the factorization

$$f^0(\lambda)+\lambda^{2n}g(\lambda)=\left|\sum_{k=0}^{\infty}\theta^0(k)e^{-i\lambda k}\right|^2=\left|\sum_{k=0}^{\infty}\psi^0(k)e^{-i\lambda k}\right|^{-2}, \tag{64}$$

determine a solution to the constraint optimisation problem

$$\langle\mathbf{C}_\mu\overline{\psi}_\mu,\mathbf{C}_\mu\overline{\psi}_\mu\rangle_1+\langle\mathbf{G}D^\mu\mathbf{A}\theta_\mu,\mathbf{A}_\mu\overline{\psi}_\mu\rangle-\langle D^\mu\mathbf{A}\theta_\mu-\mathbf{G}\mathbf{A}_\mu\overline{\psi}_\mu,D^\mu\mathbf{A}\theta_\mu\rangle\to\inf, \tag{65}$$

$$f(\lambda)=\left|\sum_{k=0}^{\infty}\theta(k)e^{-i\lambda k}\right|^2-\lambda^{2n}\left|\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right|^2=\left|\sum_{k=0}^{\infty}\psi(k)e^{-i\lambda k}\right|^{-2}-\lambda^{2n}\left|\sum_{k=0}^{\infty}\phi(k)e^{-i\lambda k}\right|^2\in\mathcal{D}_f.$$

The minimax spectral characteristic is determined as $h^0=h_\mu(f^0,g)$ if $h_\mu(f^0,g)\in H_{\mathcal{D}}$.

The function h^0 and the pair (f^0,g^0) form a saddle point of the function $\Delta(h;f,g)$ on the set $H_{\mathcal{D}}\times\mathcal{D}$. The saddle point inequalities

$$\Delta(h;f^0,g^0)\geq\Delta(h^0;f^0,g^0)\geq\Delta(h^0;f,g)\quad\forall f\in\mathcal{D}_f,\forall g\in\mathcal{D}_g,\forall h\in H_{\mathcal{D}}$$

hold true if $h^0=h_\mu(f^0,g^0)$ and $h_\mu(f^0,g^0)\in H_{\mathcal{D}}$, where (f^0,g^0) is a solution of the constraint optimisation problem

$$\widetilde{\Delta}(f,g)=-\Delta(h_\mu(f^0,g^0);f,g)\to\inf,\quad(f,g)\in\mathcal{D}, \tag{66}$$

where

$$\Delta(h_\mu(f^0,g^0);f,g)$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n f^0(\lambda) - \lambda^{2n}\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a} - (\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|^2}{|1-e^{i\lambda\mu}|^{2n}(f^0(\lambda)+\lambda^{2n}g^0(\lambda))^2} g(\lambda)d\lambda$$

$$+\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n \lambda^{2n}g^0(\lambda) + \lambda^{2n}\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a} - (\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|^2}{\lambda^{2n}|1-e^{i\lambda\mu}|^{2n}(f^0(\lambda)+\lambda^{2n}g^0(\lambda))^2} f(\lambda)d\lambda$$

or the constraint optimisation problem

$$\widetilde{\Delta}(f,g) = -\Delta(h_\mu(f^0,g^0);f,g) \to \inf, \quad (f,g)\in\mathcal{D}, \tag{67}$$

where

$$\Delta(h_\mu(f^0,g^0);f,g) =$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\left|r_{\mu,g}^0(e^{-i\lambda})\right|^2}{f^0(\lambda)+\lambda^{2n}g^0(\lambda)} f(\lambda)d\lambda + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\lambda^{2n}\left|r_{\mu,f}^0(e^{-i\lambda})\right|^2}{f^0(\lambda)+\lambda^{2n}g^0(\lambda)} g(\lambda)d\lambda.$$

Here

$$r_{\mu,g}^0(e^{-i\lambda}) = B_\mu(e^{i\lambda})\sum_{k=0}^{\infty}\theta_\mu^0(k)e^{-i\lambda k} - \sum_{k=1}^{\infty}(\mathbf{C}_\mu^0\overline{\psi}_\mu^0 + (\widetilde{\mathbf{B}}_\mu)'\theta_\mu^0)_k e^{-i\lambda k},$$

$$r_{\mu,f}^0(e^{-i\lambda}) = \left(B_\mu(e^{i\lambda}) - \frac{A(e^{i\lambda})}{(1-e^{-i\lambda\mu})^n}\right)\sum_{k=0}^{\infty}\theta_\mu^0(k)e^{-i\lambda k} - \sum_{k=1}^{\infty}(\mathbf{C}_\mu^0\overline{\psi}_\mu^0 + (\widetilde{\mathbf{B}}_\mu)'\theta_\mu^0)_k e^{-i\lambda k}.$$

These constraint optimisation problems (66),(67) are equivalent to the unconstraint optimisation problem [39]

$$\Delta_{\mathcal{D}}(f,g) = \widetilde{\Delta}(f,g) + \delta(f,g|\mathcal{D}_f\times\mathcal{D}_g) \to \inf, \tag{68}$$

where $\delta(f,g|\mathcal{D}_f\times\mathcal{D}_g)$ is the indicator function of the set $\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$. Solution (f^0,g^0) to this unconstraint optimisation problem is characterized by the condition $0\in\partial\Delta_{\mathcal{D}}(f^0,g^0)$, where $\partial\Delta_{\mathcal{D}}(f^0,g^0)$ is the subdifferential of the functional $\Delta_{\mathcal{D}}(f,g)$ at point $(f^0,g^0)\in\mathcal{D}=\mathcal{D}_f\times\mathcal{D}_g$, that is the set of all continuous linear functionals Λ on $L_1\times L_1$ which satisfy the inequality $\Delta_{\mathcal{D}}(f,g)-\Delta_{\mathcal{D}}(f^0,g^0)\geq\Lambda((f,g)-(f^0,g^0)), (f,g)\in\mathcal{D}$ (see books [15, 39, 40] for more details).

The form of the functional $\Delta(h_\mu(f^0,g^0);f,g)$ is convenient for application the Lagrange method of indefinite multipliers for finding solution to the problem (68). Making use the method of Lagrange multipliers and the form of subdifferentials of the indicator functions $\delta(f,g|\mathcal{D}_f\times\mathcal{D}_g)$ of the set $\mathcal{D}_f\times\mathcal{D}_g$ of spectral densities we describe relations that determine least favourable spectral densities in some special classes of spectral densities (see books [15, 33, 35] for additional details).

6. Least favorable spectral densities in the class $\mathcal{D}_f^0\times\mathcal{D}_g^0$

Consider the problem of minimax extrapolation of the functional $A\xi$ which depend on unobserved values of a stochastic sequence $\xi(m)$ with stationary nth increments based on observations of the sequence $\xi(m)+\eta(m)$ at points of time $m=-1,-2,\ldots$ under the condition that spectral densities of the sequences are not known exactly, but the set of admissible spectral densities $\mathcal{D}=\mathcal{D}_f^0\times\mathcal{D}_g^0$ is given, where

$$\mathcal{D}_f^0 = \left\{f(\lambda)|\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\lambda)d\lambda\leq P_1\right\}, \quad \mathcal{D}_g^0 = \left\{g(\lambda)|\frac{1}{2\pi}\int_{-\pi}^{\pi}g(\lambda)d\lambda\leq P_2\right\}.$$

Let us assume that densities $f^0\in\mathcal{D}_f^0$, $g^0\in\mathcal{D}_g^0$ and the functions

$$h_{\mu,f}(f^0,g^0) = \frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n\lambda^{2n}g^0(\lambda) + \lambda^{2n}\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a} - (\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|}{|\lambda|^n|1-e^{i\lambda\mu}|^n(f^0(\lambda)+\lambda^{2n}g^0(\lambda))}, \tag{69}$$

$$h_{\mu,g}(f^0,g^0)=\frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n f^0(\lambda)-\lambda^{2n}\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a}-(\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|}{|1-e^{i\lambda\mu}|^n(f^0(\lambda)+\lambda^{2n}g^0(\lambda))} \tag{70}$$

are bounded. In this case the functional $\Delta(h_\mu(f^0,g^0);f,g)$ is continuous and bounded in the $\mathcal{L}_1\times\mathcal{L}_1$ space. The condition $0\in\partial\Delta_{\mathcal{D}}(f^0,g^0)$ leads to the equation for the least favorable densities $f^0\in\mathcal{D}_f^0$, $g^0\in\mathcal{D}_g^0$:

$$\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n\lambda^{2n}g^0(\lambda)+\lambda^{2n}\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a}-(\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|$$

$$=\alpha_1|\lambda|^n|1-e^{i\lambda\mu}|^n(f^0(\lambda)+\lambda^{2n}g^0(\lambda)), \tag{71}$$

$$\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n f^0(\lambda)-\lambda^{2n}\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a}-(\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|$$

$$=\alpha_2|1-e^{i\lambda\mu}|^n(f^0(\lambda)+\lambda^{2n}g^0(\lambda)), \tag{72}$$

where $\alpha_1\geq 0$ and $\alpha_2\geq 0$ are such constants that $\alpha_1\neq 0$ if $\frac{1}{2\pi}\int_{-\pi}^{\pi}f^0(\lambda)d\lambda=P_1$ and $\alpha_2\neq 0$ if $\frac{1}{2\pi}\int_{-\pi}^{\pi}g^0(\lambda)d\lambda=P_2$.

Thus, we have the following statements.

Theorem 10
Let spectral densities $f^0(\lambda)\in\mathcal{D}_f^0$ and $g^0(\lambda)\in\mathcal{D}_g^0$ satisfy condition (15), let functions $h_{\mu,f}(f^0,g^0)$ and $h_{\mu,g}(f^0,g^0)$ be bounded. The spectral densities $f^0(\lambda)$ and $g^0(\lambda)$ determined by equations (71), (72) are least favorable in the class $\mathcal{D}=\mathcal{D}_f^0\times\mathcal{D}_g^0$ for the optimal linear estimation of the functional $A\xi$ if they determine a solution of extremum problem (59). The function $h_\mu(f^0,g^0)$ determined by formula (22) is minimax spectral characteristic of the optimal estimate of the functional $A\xi$.

Theorem 11
Suppose that spectral density $f(\lambda)$ is known, spectral density $g^0(\lambda)\in\mathcal{D}_g^0$ and conditions (15) is satisfied. Let the function $h_{\mu,g}(f,g^0)$ be bounded. Spectral density $g^0(\lambda)$ is least favorable in the class $\mathcal{D}_g^0$ for the optimal linear extrapolation of the functional $A\xi$ if it is of the form

$$g^0(\lambda)=\frac{1}{\lambda^{2n}}\max\left\{0,\frac{\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n f(\lambda)-\lambda^{2n}\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a}-(\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|}{\alpha_2|1-e^{i\lambda\mu}|^n}-f(\lambda)\right\}$$

and the pair (f,g^0) determines a solution to extremum problem (59). The function $h_\mu(f,g^0)$ determined by formula (22) is minimax spectral characteristic of the optimal estimation of the functional $A\xi$.

Theorem 12
Let spectral density $g(\lambda)$ be known, spectral density $f^0(\lambda)\in\mathcal{D}_f^0$ and condition (15) be satisfied. Let the function $h_{\mu,f}(f^0,g)$ be bounded. Spectral density $f^0(\lambda)$ is least favorable in the class $\mathcal{D}_f^0$ for the optimal linear extrapolation of the functional $A\xi$ if it is of the form

$$f^0(\lambda)=\max\left\{0,\frac{|\lambda|^n\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n g(\lambda)+\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a}-(\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|}{\alpha_1|1-e^{i\lambda\mu}|^n}-\lambda^{2n}g(\lambda)\right\}$$

and the pair (f^0,g) determines a solution to extremum problem (59). The function $h_\mu(f^0,g)$ determined by formula (22) is minimax spectral characteristic of the optimal estimation of the functional $A\xi$.

7. Least favorable spectral densities which admit factorization in the class $\mathcal{D}_f^0 \times \mathcal{D}_g^0$

Consider the problem of minimax extrapolation of the functional $A\xi$ from observations $\xi(k)+\eta(k)$, $k \leq -1$, provided that spectral densities $f(\lambda)$ and $g(\lambda)$ admit canonical factorizations (35) – (36) and belong to the set $\mathcal{D} = \mathcal{D}_f^0 \times \mathcal{D}_g^0$, where

$$\mathcal{D}_f^0 = \left\{ f(\lambda) | \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) d\lambda \leq P_1 \right\}, \quad \mathcal{D}_g^0 = \left\{ g(\lambda) | \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) d\lambda \leq P_2 \right\}.$$

The condition $0 \in \partial\Delta_{\mathcal{D}}(f^0, g^0)$ implies that least favorable densities $f^0 \in \mathcal{D}_f^0$, $g^0 \in \mathcal{D}_g^0$ satisfy equations

$$f^0(\lambda) + \lambda^{2n} g^0(\lambda) = \alpha_1 \left| B_\mu(e^{i\lambda}) \sum_{k=0}^{\infty} \theta_\mu(k) e^{-i\lambda k} - \sum_{k=1}^{\infty} (\mathbf{C}_\mu \overline{\psi}_\mu + (\widetilde{\mathbf{B}}_\mu)' \theta_\mu)_k e^{-i\lambda k} \right|^2, \tag{73}$$

$$f^0(\lambda) + \lambda^{2n} g^0(\lambda) = \alpha_2 \lambda^{2n} \left| r_{\mu,f}(e^{-i\lambda}) \right|^2, \tag{74}$$

$$r_{\mu,f}(e^{-i\lambda}) = \left(B_\mu(e^{i\lambda}) - \frac{A(e^{i\lambda})}{(1-e^{-i\lambda\mu})^n} \right) \sum_{k=0}^{\infty} \theta_\mu(k) e^{-i\lambda k} - \sum_{k=1}^{\infty} (\mathbf{C}_\mu \overline{\psi}_\mu + (\widetilde{\mathbf{B}}_\mu)' \theta_\mu)_k e^{-i\lambda k},$$

where coefficients $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, the matrix $\mathbf{C}_\mu$, vectors $\theta_\mu = (\theta_\mu(0), \theta_\mu(1), \theta_\mu(2), \ldots)'$ and $\psi_\mu = (\psi_\mu(0), \psi_\mu(1), \psi_\mu(2), \ldots)'$ are determined by factorizations (36) and (44) of the functions $g^0(\lambda)$ and $f^0(\lambda) + \lambda^{2n} g^0(\lambda)$, relation (45) and the conditions

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) d\lambda = P_1, \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) d\lambda = P_2. \tag{75}$$

Thus, the following statements come true.

Theorem 13
Least favorable spectral densities $f^0(\lambda) \in \mathcal{D}_f^0$ and $g^0(\lambda) \in \mathcal{D}_g^0$ for the optimal linear estimation of the functional $A\xi$ which admit canonical factorizations (36), (44) are determined by equations (73) – (74), relation (45), conditions (61) and (75).

Theorem 14
Assume that spectral density $f(\lambda)$ is known and admits the canonical factorization. Then spectral density

$$g^0(\lambda) = \frac{1}{\lambda^{2n}} \left[\alpha_2 \lambda^{2n} \left| r_{\mu,f}(e^{-i\lambda}) \right|^2 - f(\lambda) \right]_+$$

from the class $\mathcal{D}_g^0$ is the least favorable spectral density for the optimal linear estimation of the functional $A\xi$. The coefficient $\alpha_2 \geq 0$, matrix $\mathbf{C}_\mu$, vectors $\theta_\mu = (\theta_\mu(0), \theta_\mu(1), \theta_\mu(2), \ldots)'$ and $\psi_\mu = (\psi_\mu(0), \psi_\mu(1), \psi_\mu(2), \ldots)'$ are determined by canonical factorizations (36), (44) of the functions $g^0(\lambda)$ and $f(\lambda) + \lambda^{2n} g^0(\lambda)$, relation (45), conditions (63) and $\int_{-\pi}^{\pi} g(\lambda) d\lambda = 2\pi P_2$.

Theorem 15
Assume that spectral density $g(\lambda)$ is known and admits the canonical factorization. Then the spectral density

$$f^0(\lambda) = \left[\alpha_1 \left| B_\mu(e^{i\lambda}) \sum_{k=0}^{\infty} \theta_\mu(k) e^{-i\lambda k} - \sum_{k=1}^{\infty} (\mathbf{C}_\mu \overline{\psi}_\mu + (\widetilde{\mathbf{B}}_\mu)' \theta_\mu)_k e^{-i\lambda k} \right|^2 - \lambda^{2n} g(\lambda) \right]_+$$

from the class $\mathcal{D}_f^0$ is the least favorable spectral density for the optimal linear estimation of the functional $A\xi$. The matrix $\mathbf{C}_\mu$ is known and defined by coefficients of factorization (36) of the spectral density $g(\lambda)$. The coefficient $\alpha_1 \geq 0$ and vectors $\theta_\mu = (\theta_\mu(0), \theta_\mu(1), \theta_\mu(2), \ldots)'$, $\psi_\mu = (\psi_\mu(0), \psi_\mu(1), \psi_\mu(2), \ldots)'$ are determined by canonical factorization (44) of the function $f^0(\lambda) + \lambda^{2n} g(\lambda)$, relation (45), conditions (65) and $\int_{-\pi}^{\pi} f(\lambda) d\lambda = 2\pi P_1$.

8. Least favorable densities in the class $\mathcal{D} = \mathcal{D}_v^u \times \mathcal{D}_\varepsilon$

Consider the problem of optimal linear extrapolation of the functional $A\xi$ for the set of spectral densities $\mathcal{D} = \mathcal{D}_v^u \times \mathcal{D}_\varepsilon$, where

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

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

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

Let spectral densities $f^0 \in \mathcal{D}_u^v$, $g^0 \in \mathcal{D}_\varepsilon$ be such that functions $h_{\mu,f}(f^0, g^0)$ and $h_{\mu,g}(f^0, g^0)$ determined by formulas (69), (70) are bounded. From the condition $0 \in \partial\Delta_{\mathcal{D}}(f^0, g^0)$ we get the following equations defining least favorable densities

$$\left| A(e^{i\lambda})(1-e^{i\lambda\mu})^n \lambda^{2n} g^0(\lambda) + \lambda^{2n} \sum_{k=0}^{\infty} ((\mathbf{P}_\mu^0)^{-1} D^\mu \mathbf{a} - (\mathbf{P}_\mu^0)^{-1} \mathbf{T}_\mu^0 \mathbf{a}_\mu)_k e^{i\lambda k} \right|$$
$$= |\lambda|^n |1 - e^{i\lambda\mu}|^n (f^0(\lambda) + \lambda^{2n} g^0(\lambda))(\gamma_1(\lambda) + \gamma_2(\lambda) + \alpha_1^{-1}), \tag{76}$$

$$\left| A(e^{i\lambda})(1-e^{i\lambda\mu})^n f^0(\lambda) - \lambda^{2n} \sum_{k=0}^{\infty} ((\mathbf{P}_\mu^0)^{-1} D^\mu \mathbf{a} - (\mathbf{P}_\mu^0)^{-1} \mathbf{T}_\mu^0 \mathbf{a}_\mu)_k e^{i\lambda k} \right|$$
$$= |1 - e^{i\lambda\mu}|^n (f^0(\lambda) + \lambda^{2n} g^0(\lambda))(\beta(\lambda) + \alpha_2^{-1}), \tag{77}$$

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

The following statements hold true.

Theorem 16
Let spectral densities $f^0(\lambda) \in \mathcal{D}_u^v$, $g^0(\lambda) \in \mathcal{D}_\varepsilon$ satisfy condition (15). Let functions $h_{\mu,f}(f^0, g^0)$ and $h_{\mu,g}(f^0, g^0)$ determined by formulas (69), (70) be bounded. Spectral densities $f^0(\lambda)$ and $g^0(\lambda)$ determined by equations (76), (77) are least favorable in the class $\mathcal{D} = \mathcal{D}_v^u \times \mathcal{D}_\varepsilon$ for the optimal linear extrapolation of the functional $A\xi$ if they determine a solution of extremum problem (59). The function $h_\mu(f^0, g^0)$ determined by (22) is minimax spectral characteristic of the optimal estimate of the functional $A\xi$.

Theorem 17
Let spectral density $f(\lambda)$ be known, spectral density $g^0(\lambda) \in \mathcal{D}_\varepsilon$ and condition (15) be satisfied. Assume that the function $h_{\mu,g}(f, g^0)$ determined by formula (70) is bounded. Spectral density $g^0(\lambda)$ is least favorable in the class $\mathcal{D}_\varepsilon$ for the optimal linear extrapolation of the functional $A\xi$ if it is of the form

$$g^0(\lambda) = \frac{1}{\lambda^{2n}} \max\left\{(1-\varepsilon) g_1(\lambda), f_1(\lambda)\right\},$$

$$f_1(\lambda) = \frac{\alpha_2 \left| A(e^{i\lambda})(1-e^{i\lambda\mu})^n f(\lambda) - \lambda^{2n} \sum_{k=0}^{\infty} ((\mathbf{P}_\mu^0)^{-1} D^\mu \mathbf{a} - (\mathbf{P}_\mu^0)^{-1} \mathbf{T}_\mu^0 \mathbf{a}_\mu)_k e^{i\lambda k} \right|}{|1 - e^{i\lambda\mu}|^n} - f(\lambda),$$

and the pair (f, g^0) determines a solution to extremum problem (59). The function $h_\mu(f, g^0)$ determined by formula (22) is minimax spectral characteristic of the optimal estimate of the functional $A\xi$.

Theorem 18
Suppose spectral density $g(\lambda)$ is known, spectral density $f^0(\lambda) \in \mathcal{D}_v^u$ and condition (15) is satisfied. Let the function $h_{\mu,f}(f^0, g)$ be bounded. Spectral density $f^0(\lambda)$ is least favorable in the class $\mathcal{D}_v^u$ for the optimal linear extrapolation of the functional $A\xi$ if it is of the form

$$f^0(\lambda) = \min\left\{v(\lambda), \max\left\{u(\lambda), g_2(\lambda)\right\}\right\},$$

$$g_2(\lambda)=\frac{\alpha_1|\lambda|^n\left|A(e^{i\lambda})(1-e^{i\lambda\mu})^n g(\lambda)+\sum_{k=0}^{\infty}((\mathbf{P}_\mu^0)^{-1}D^\mu\mathbf{a}-(\mathbf{P}_\mu^0)^{-1}\mathbf{T}_\mu^0\mathbf{a}_\mu)_k e^{i\lambda k}\right|}{|1-e^{i\lambda\mu}|^n}-\lambda^{2n}g(\lambda)$$

and the pair (f^0,g) determines a solution to extremum problem (59). The function $h_\mu(f^0,g)$ determined by formula (22) is minimax spectral characteristic of the optimal estimation of the functional $A\xi$.

9. Least favorable spectral densities which admit factorization in the class $\mathcal{D}=\mathcal{D}_v^u\times\mathcal{D}_\varepsilon$

Consider the problem of minimax extrapolation of the functional $A\xi$ from observations $\xi(k)+\eta(k)$, $k\leq-1$, provided spectral densities $f(\lambda)$ and $d(\lambda)$ admit canonical factorizations (35) – (36) and belong to the set $\mathcal{D}=\mathcal{D}_v^u\times\mathcal{D}_\varepsilon$, where

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

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

Here spectral densities $u(\lambda)$, $v(\lambda)$, $g_1(\lambda)$ are known and fixed, spectral density $w(\lambda)$ is unknown. The condition $0\in\partial\Delta_{\mathcal{D}}(f^0,g^0)$ implies that least favorable densities $f^0\in\mathcal{D}_u^u$, $g^0\in\mathcal{D}_\varepsilon$ satisfy the equations

$$f^0(\lambda)+\lambda^{2n}g^0(\lambda)=\alpha_1\left|r_{\mu,g}(e^{-i\lambda})\right|^2(\gamma_1(\lambda)+\gamma_2(\lambda)+1)^{-1},\tag{78}$$

$$r_{\mu,g}(e^{-i\lambda})=B_\mu(e^{i\lambda})\sum_{k=0}^{\infty}\theta_\mu(k)e^{-i\lambda k}-\sum_{k=1}^{\infty}(\mathbf{C}_\mu\overline{\psi}_\mu+(\widetilde{\mathbf{B}}_\mu)'\theta_\mu)_k e^{-i\lambda k}$$

$$f^0(\lambda)+\lambda^{2n}g^0(\lambda)=\alpha_2\lambda^{2n}\left|r_{\mu,f}(e^{-i\lambda})\right|^2(\beta(\lambda)+1)^{-1},\tag{79}$$

$$r_{\mu,f}(e^{-i\lambda})=\left(B_\mu(e^{i\lambda})-\frac{A(e^{i\lambda})}{(1-e^{-i\lambda\mu})^n}\right)\sum_{k=0}^{\infty}\theta_\mu(k)e^{-i\lambda k}-\sum_{k=1}^{\infty}(\mathbf{C}_\mu\overline{\psi}_\mu+(\widetilde{\mathbf{B}}_\mu)'\theta_\mu)_k e^{-i\lambda k},$$

where $\gamma_1(\lambda)\leq0$ and $\gamma_1(\lambda)=0$ if $f^0(\lambda)\geq v(\lambda)$; $\gamma_2(\lambda)\geq0$ and $\gamma_2=0$ if $f^0(\lambda)\leq u(\lambda)$; $\beta(\lambda)\leq0$ and $\beta(\lambda)=0$ when $g^0(\lambda)\geq(1-\varepsilon)g_1(\lambda)$. Coefficients $\alpha_1\geq0$, $\alpha_2\geq0$, matrix $\mathbf{C}_\mu$, vectors $\theta_\mu=(\theta_\mu(0),\theta_\mu(1),\theta_\mu(2),\dots)'$ and $\psi_\mu=(\psi_\mu(0),\psi_\mu(1),\psi_\mu(2),\dots)'$ are determined by factorizations (36) and (44) of the functions $g^0(\lambda)$ and $f^0(\lambda)+\lambda^{2n}g^0(\lambda)$, relation (45) and conditions (75).

Thus, the following statements come true.

Theorem 19
Least favorable spectral densities $f^0(\lambda)\in\mathcal{D}_v^u$ and $g^0(\lambda)\in\mathcal{D}_\varepsilon$ for the optimal linear estimation of the functional $A\xi$ which allow canonical factorizations (36), (44) are determined by equations (78) – (79), relation (45), conditions (61) and (75).

Theorem 20
Assume that spectral density $f(\lambda)$ is known and admits canonical factorization. Then the spectral density

$$g^0(\lambda)=\frac{1}{\lambda^{2n}}\max\left\{\alpha_2\lambda^{2n}\left|r_{\mu,f}(e^{-i\lambda})\right|^2-f(\lambda),(1-\varepsilon)g_1(\lambda)\right\}$$

from the class $\mathcal{D}_\varepsilon$ is the least favorable spectral density for the optimal linear estimation of the functional $A\xi$. A coefficient $\alpha_2\geq0$, matrix $\mathbf{C}_\mu$, vectors $\theta_\mu=(\theta_\mu(0),\theta_\mu(1),\theta_\mu(2),\dots)'$ and $\psi_\mu=(\psi_\mu(0),\psi_\mu(1),\psi_\mu(2),\dots)'$ are determined by canonical factorizations (36), (44) of the functions $g^0(\lambda)$ and $f(\lambda)+\lambda^{2n}g^0(\lambda)$, relation (45), conditions (63) and $\int_{-\pi}^{\pi}g(\lambda)d\lambda=2\pi P_2$.

Theorem 21
Assume that spectral density $g(\lambda)$ is known and admits canonical factorization. Then spectral density

$$f^0(\lambda) = \min\left\{\max\left\{\alpha_1 \left|r_{\mu,g}(e^{-i\lambda})\right|^2 - \lambda^{2n} g(\lambda), v(\lambda)\right\} u(\lambda)\right\}$$

from the class $\mathcal{D}_v^u$ is the least favorable spectral density for the optimal linear estimation of the functional $A\xi$. Matrix $\mathbf{C}_\mu$ is known and defined by the coefficients of factorization (36) of the spectral density $g(\lambda)$. A coefficient $\alpha_1 \geq 0$ and vectors $\theta_\mu = (\theta_\mu(0), \theta_\mu(1), \theta_\mu(2), \ldots)'$ and $\psi_\mu = (\psi_\mu(0), \psi_\mu(1), \psi_\mu(2), \ldots)'$ are determined by canonical factorization (44) of the function $f^0(\lambda) + \lambda^{2n} g(\lambda)$, relation (45), conditions (65) and $\int_{-\pi}^{\pi} f(\lambda) d\lambda = 2\pi P_1$.

10. Conclusions

In this article we propose solutions to the extrapolation problem for the functionals $A\xi = \sum_{k=0}^{\infty} a(k)\xi(k)$ and $A_N\xi = \sum_{k=0}^{N} a(k)\xi(k)$, which depend on unobserved values of a stochastic sequence $\xi(m)$ with stationary nth increments. Estimates are based on observations of the sequence $\xi(m) + \eta(m)$ at points of time $m = -1, -2, \ldots$, where $\eta(m)$ is an uncorrelated with $\xi(m)$ stationary sequence. We derive formulas for calculation values of the mean-square errors and spectral characteristics of the optimal linear estimates of the functionals in the case of spectral certainty where spectral densities of the sequences are known. The obtained results are applied to finding solution to the extrapolation problem for cointegrated sequences. In the case of spectral uncertainty where spectral densities are not known exactly, but a set of admissible spectral densities is specified, the minimax-robust method is applied. Formulas that determine least favorable spectral densities and minimax (robust) spectral characteristics are derived for some special sets of admissible spectral densities. The extrapolation problem for ARIMA(0,1,1) sequence is analyzed as an example of application of the developed method.

Acknowledgement

The authors would like to thank the referees for careful reading of the article and giving constructive suggestions.

REFERENCES

1. W. Bell, *Signal extraction for nonstationary time series*, The Annals of Statistics, vol. 12, no. 2, pp. 646-664, 1984.
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. Golichenko, and M. P. Moklyachuk, *Estimates of functionals of periodically correlated processes*, Kyiv: NVP "Interservis", 2014.
4. I. I. Gikhman, and A. V. Skorokhod, *The theory of stochastic processes. I.*, Berlin: Springer, 2004.
5. C. W. J. Granger, *Cointegrated variables and error correction models*, UCSD Discussion paper, vol.83-13a, 1983.
6. 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.
7. 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.
8. 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.
9. I. I. Dubovets'ka, and M. P. Moklyachuk, *Minimaxestimation problem for periodically correlated stochastic processes*, Journal of Mathematics and System Science, vol. 3, no. 1, pp. 26-30, 2013.
10. 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.
11. R. F. Engle, and C. W. J. Granger, *Co-integration and error correction: Representation, estimation and testing*, Econometrica, vol. 55, pp. 251-276, 1987.
12. U. Grenander, *A prediction problem in game theory*, Arkiv för Matematik, vol. 3, pp. 371-379, 1957.
13. J. Franke, *Minimax robust prediction of discretetime series*, Z. Wahrsch. Verw. Gebiete, vol. 68, pp. 337-364, 1985.
14. J. Franke, and H. V. Poor, *Minimax-robust filteringand finite-length robust predictors*, Robust and Nonlinear Time Series Analysis. Lecture Notes in Statistics, Springer-Verlag, vol. 26, pp. 87-126, 1984.

15. 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.
16. K. Karhunen, *Uber lineare Methoden in der Wahrscheinlichkeitsrechnung*, Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys., no. 37, 1947.
17. 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.
18. A. N. Kolmogorov, *Selected works by A. N. Kolmogorov. Vol. II: Probability theory and mathematical statistics. Ed. by A. N. Shiryayev. Mathematics and Its Applications. Soviet Series. 26. Dordrecht etc.* Kluwer Academic Publishers, 1992.
19. M. M. Luz, and M. P. Moklyachuk, *Interpolation of functionals of stochactic sequanses with stationary increments*, Theory of Probability and Mathematical Statistics, vol. 87, pp. 117-133, 2013.
20. 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.
21. 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.
22. 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.
23. M. Luz, and M. Moklyachuk, *Minimax-robust filtering problem for stochastic sequences with stationary increments and cointegrated sequences*, Statistics, Optimization & Information Computing, vol. 2, no. 3, pp. 176 - 199, 2014.
24. 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.
25. 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.
26. M. P. Moklyachuk, *Minimax extrapolation and autoregressive-moving average processes*, Theory of Probability and Mathematical Statistics, vol. 41, pp. 77-84, 1990.
27. M. P. Moklyachuk, *Robust procedures in time series analysis*, Theory of Stochastic Processes, vol. 6, no. 3-4, pp. 127-147, 2000.
28. 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.
29. M. P. Moklyachuk, *Robust estimations of functionals of stochastic processes.*, Kyiv University, Kyiv, 2008.
30. M. Moklyachuk, and A. Masyutka, *Extrapolation of multidimensional stationary processes*, Random Operators and Stochastic Equations, vol. 14, no. 3, pp. 233-244, 2006.
31. M. Moklyachuk, and A. Masyutka, *Robust estimation problems for stochastic processes*, Theory of Stochastic Processes, vol. 12, no. 3-4, pp. 88-113, 2006.
32. M. Moklyachuk, and A. Masyutka, *Robust filtering of stochastic processes*, Theory of Stochastic Processes, vol. 13, no. 1-2, pp. 166-181, 2007.
33. 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.
34. 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.
35. M. Moklyachuk, and O. Masyutka, *Minimax-robust estimation technique for stationary stochastic processes*, LAP LAMBERT Academic Publishing, 2012.
36. M. P. Moklyachuk, *Nonsmooth analysis and optimization*, Kyiv University, Kyiv, 2008.
37. M. S. Pinsker, and A. M. Yaglom, *On linear extrapolaion of random processes with nth stationary incremens*, Doklady Akademii Nauk SSSR, vol. 94, pp. 385-388, 1954.
38. M. S. Pinsker, *The theory of curves with nth stationary incremens in Hilber spaces*, Izvestiya Akademii Nauk SSSR. Ser. Mat., vol. 19, no. 5, pp. 319-344, 1955.
39. B. N. Pshenichnyi, *Necessary conditions of an extremum*, Marcel Dekker, New York, 1971.
40. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1997.
41. Yu. A. Rozanov, *Stationary stochastic processes*, Holden-Day, San Francisco, 1967
42. N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series. Whis Engineering Applications*, The M. I. T. Press, Massachusetts Institute of Technology, Cambridge, Mass., 1966.
43. 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.
44. A. M. Yaglom, *Correlation theory of stationary and related random functions. Vol. 2: Suplementary notes and references*, Springer Series in Statistics, Springer-Verlag, New York etc., 1987.
45. A. M. Yaglom, *Correlation theory of stationary and related random processes with stationary nth increments*, Mat. Sbornik, vol. 37, no. 1, pp. 141-196, 1955.
46. A. M. Yaglom, *Some clases of random fields in n-dimentional space related with random stationary processes*, Teor. Veroyatn. Primen., vol. 11, no. 3, pp. 292-337, 1957.