S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports

http://semr.math.nsc.ru

Том 17, стр. 1322–1351 (2020) УДК 517.925.54,517.962.27

DOI 10.33048/semi.2020.17.098 MSC 65F25, 15A03, 15A09, 15A23, 93E12

COUNTER EQUATIONS: SMOOTHING, FILTRATION, IDENTIFICATION

A.O. EGORSHIN

Abstract. Inverse problems of analysis, mathematical modeling, and identification of dynamical systems and processes are studied on the base of linear stationary models. The method of analysis is dynamical approximation of signals and functions determined on uniform grids on finite intervals. One class of approximating functions and their models is transition processes of linear difference or differential equations with constant, possibly unknown, coefficients. In the latter case, their estimation (identification) is performed on the basis of the least square method of approximation of observed processes and specified functions. We show that all the considered problems may be effectively solved using computer algorithms based on counter equations of bilateral orthogonalization of homogeneous vector systems, generated by approximating models.

Keywords: counter equations, bilateral orthogonalization, smoothing, filtration, identification, piecewise-linear approximation, estimation, nonlinear optimization, optimization of subspaces, difference equations, matrix inversion, Riccati matrix equation, renewal equation, renewal process, homogeneous vector system.

1. Introduction

1.1. **Preliminary data and renewal equations.** The tasks of studying (smoothing and identification) dynamical processes and functions on the base of grid data (realizations) are considered. The tasks are set as problems of approximation of finite sequences $\mathbf{y} = \{y_i\}_0^L$ with the help of dynamical models of special class.

Egorshin, A.O., Counter equations: smoothing, filtration, identification.

The work is supported by RFFI (grant 19-01-00754 a).
Received January, 19, 2018, published September, 11, 2020.

The main instrument used to solve the mentioned tasks of analysis and identification of dynamical processes by their realizations $\mathbf{y} = \mathbf{y}_{[L]}$ are the solutions $\widehat{\mathbf{y}}_{[k]}$ of the problems of smoothing of subrealizations $\mathbf{y}_{[k]}$ of those processes for $k = \overline{0, L}$. The realizations $\mathbf{y}$ are vectors on Euclidian space $E = E^{L+1}$. The vectors $\mathbf{y}$ can be, for example, sequences of samplings of a studied dynamical process on a given grid: $\mathbf{y} = \mathbf{y}_{[L]} = \{y_i\}_0^L \in E^{L+1} = E$. The process of sequential solution of these tasks of smoothing is implemented using recurrent equations, which in this paper are referred to as renewal equations.

The term and notion of renewal were introduced by T. Kailath [1] for characteristic of correction summands in recurrent equations. These summands are defined depending on prediction errors in sequential calculations.

The point of such calculations in the terminology used here is as follows. On a regular step $(k+1)$, corresponding to applying of the sampling y_{k+1} of the initial realization $\mathbf{y}$, the regular realization $\widehat{\mathbf{y}}_{[k+1]} = \{\widehat{y}_i\}_0^{k+1}$, which is referred to as smoothed, is calculated with the help of renewal equation. That is done by renewing the previous (the current, or the one corresponding to the current value k of the process parameter) smoothed realization $\widehat{\mathbf{y}}_{[k]}$. The calculations are made on the basis of a regular component π_{k+1} of the renewal process π. That component $\pi_{k+1} = y_{k+1} - \widehat{y}_{k+1/k}$ of the renewal process is defined depending on the deviation of the sampling y_{k+1} of the initial realization $\mathbf{y}$ from its prediction $\widehat{y}_{k+1/k}$. The prediction is calculated using the process model on the basis of the previous smoothed realization $\widehat{\mathbf{y}}_{[k]}$.

Initial sequences of arbitrary numbers $\mathbf{y}$ are considered as $(L+1)$-vectors on a bounded region of the Euclidian space $E^{L+1} = E$. The tasks of smoothing (approximation) are set and solved as problems of orthogonal projecting of the vector $\mathbf{y}$ on some subspaces in E. They are defined depending on which linear model of the analyzed process is used.

Sequences $\mathbf{y}$ can be the results of some experiments or the sequences of samplings of the analyzed functions and studied processes on a given grid. They may be considered as samplings of solutions of some differential or difference equations on this grid. Then the tasks of approximated identification of those equations on the basis of a chosen class of models can be set. In such cases, the possibility that the samplings contain accidental and undefined errors has to be taken into account.

In this paper, we also obtain filtration equations including identification equations. We will show that sequential identification equations within nonlinear filtration equations are also renewal equations.

1.2. Dynamical models in the method of least squares.

A class of approximating models is described by the following equations:

$$\mathbf{D}\widehat{y}(t) = \sum_{i=0}^{n} a_i^* \widehat{y}^{(i)}(t) = 0 \qquad t \in I_T = [0, T], \quad T = Lh. \tag{1}$$

We employ the samplings of these equations on a uniform grid I_h in an interval I_T. To describe the samplings of solutions of these equations on grids, the following difference equations are used:

(2)
$$\mathbb{D}\widehat{\mathbf{y}} = \mathbb{D}_N(\alpha)\widehat{\mathbf{y}}_{[L]} = \left\{ \sum_{i=0}^{n} \alpha_i^* \widehat{y}_{k+i} \right\}_{k=0}^{N}$$
$$= |\alpha^* \widehat{v}_{k+i}|_{i=0}^{n} = \alpha^* V_N^*(\widehat{\mathbf{y}}_{[L]}) = \alpha^* \widehat{V}^* = 0,$$

where $\widehat{v}_k = \left\{\widehat{y}_{k+i}\right\}_{i=0}^{n}$, $V_N^* = V^* = |v_i|_0^N = |v_0, \cdots, v_N|$, P° $N = L - n$.

The coefficients

(3)
$$\alpha^* = |\boldsymbol{\alpha}^*, 1| = \left|\alpha_0^*, \ldots, \alpha_{n-1}^*, 1\right|$$

of the difference operator in $\mathbb{D}$ in (2) are considered constant.

We now clarify the used notations. The superscript $()^*$ marks double involution: transpose and complex conjugation. Application of such index to a scalar denotes its complex conjugation. Elements in straight brackets with possible indices designate the row containing those elements, for example, $|x_i|_k^l$. A similar construction with braces marks the column of the said elements, for example, $\{x_i\}_k^l$. Straight brackets and braces are used to present and designate matrices, for example, $\begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}$ or $\{x_{ij}\}_{m,n}^{k,l}$.

By equations (2) we describe on given h-grids (with L cells of length h) sequences of $L+1$ samplings of studied continual dynamical processes. Equations (1) and (2) belong to the same class of equations which unites linear, ordinary difference, and differential equations with constant coefficients.

The class consists of groups of difference and differential equations, whose samplings of solutions coincide on given grids. Of possible differential equations, we choose an equation with a solution closest to a degree polynomial of order L, interpolating the samplings on a given grid .

1.3. **Formulation of the problem.** To approximate the initial realizations $\mathbf{y}$ with the smoothing sequences $\widehat{\mathbf{y}}$, the least squares method (LSM) is used. An accepted criterion of LSM and conditions for its minimization with respect to the smoothing realizations $\widehat{\mathbf{y}}$ are as follows:

(4)
$$J = \|\mathbf{y} - \widehat{\mathbf{y}}\|^2 = \sum_0^L (y_i - \widehat{y}_i)^*(y_i - \widehat{y}_i) \qquad \text{given} \qquad \mathbb{D}(\alpha)\widehat{\mathbf{y}} = 0 \quad \text{and} \quad \alpha_n \neq 0.$$

When solving problems of smoothing (coefficients $\alpha = \widehat{\alpha}$ are known), minimization of the functional J is performed by n components of the realization $\widehat{\mathbf{y}}$. Those might be its initial values as solutions of the difference equation (2), (4) of nth order. Also those may be any n of its components assigned as parameters of solution of the equation (2), (4). As a result of minimization of the functional J with respect to these parameters, we obtain the value $\widehat{J}$ of this functional on a projection $\widehat{\mathbf{y}} = \widehat{\mathbf{y}}(\widehat{\alpha}) = P(\mathcal{D})\mathbf{y}$ $\big(P(\mathcal{D})$ is an orthogonal projector on the subspace $\mathcal{D}\big)$ of the source realization $\mathbf{y}$ on the kernel $\mathcal{D}(\widehat{\alpha})$ of the operator $\mathbb{D}(\widehat{\alpha})$. The value $\widehat{J} = \widehat{J}(\widehat{\alpha})$ of the functional J on the projection $\widehat{\mathbf{y}} = \widehat{\mathbf{y}}(\widehat{\alpha})$ becomes equal $\|\mathbf{h}\|^2$, which is a square of a perpendicular distance $\mathbf{h} = \mathbf{h}(\widehat{\alpha}) = \widehat{\mathbf{y}} - \widehat{\mathbf{y}}(\widehat{\alpha})$ from the vector $\mathbf{y}$ to the kernel $\mathcal{D}(\widehat{\alpha})$ of the operator $\mathbb{D}(\widehat{\alpha})$.

Suppose that all or some of coefficients α in (3) of the operator $\mathbb{D}$ in (4) are unknown. Then, following the minimization of the functional J with respect to n components of the smoothing realization $\widehat{\mathbf{y}}$, a minimization with respect to α of its value $\widehat{J} = \widehat{J}(\widehat{\alpha})$ on the projection $\widehat{\mathbf{y}}$ is performed. That is a minimization of a square of a perpendicular distance $\mathbf{h} = \mathbf{h}(\widehat{\alpha}) = \widehat{\mathbf{y}} - \widehat{\mathbf{y}}(\widehat{\alpha})$ with respect to the unknown coefficients α of the operator $\mathbb{D}$ from equation (2). As a result, we obtain a solution $\widehat{\alpha}$ of so called identification problem (of the equation in (2) and the operator $\mathbb{D}$ in (4)).

Therefore, if not all coefficients of the operator $\mathbb{D}$ are known, then minimization of the functional J is performed with respect to $n + m$ variables: n components of the realization $\widehat{\mathbf{y}}$ (its initial or other values) and $m \leqslant n$ of unknown coefficients of the difference operator $\mathbb{D}$ (4) in the equation (2).

The problems of approximation and identification of the form (4) are called variational. For clarity of the aims of the paper, in the problems (4) we use homogeneous equations of models (1), (2).

Remark 1. It is possible to formulate and solve problems similar to (4), based on non-homogeneous equations and systems of equations with matrix coefficients. In paper [2], technical possibilities for formulating and solving problems of the form (4) for the mentioned generalizations are presented. These possibilities have to do with the use of special block constructions. The said possibilities are not considered in this paper. First, these generalizations lead to models of functions different from both physical and mathematical points of view. Second, analysis of the models of the problems studied here requires different and significantly more sophisticated approaches and methods. Third, the main goal of this paper is to represent the problems of approximation of finite sequences based on the models (1), (2), which we solve here, as direct generalizations of classical problems of polynomial approximation based on the most simple models of the form (1), (2): $y^{(n)}(t) = 0$ or $\Delta^n y_k = 0$. □

1.4. **Filtration problems.** Apart from solving problems of the form (4) of smoothing of the realizations $\mathbf{y}$ and identifying equations of a mentioned class which approximate those realizations, problems of filtration of states of dynamical systems of order n, represented by difference equations of the form (2) are also considered.

One of the well-known approaches to solving linear filtration problems in state-space of dynamical systems of general form, including non-stationary was proposed by R. Kalman and his followers [3]. T. Kailath has modified Kalman's equations into fast algorithms for stationary systems [1,3]. In this paper, problems of filtration are considered for stationary systems as well; however, as seen in (3), we deal with spaces E of source data $\mathbf{y} = \{y_i\}_0^L$ [2,4], not with state-space. Vectors in E — in particular, $\mathbf{y}$, $\widehat{\mathbf{y}}$ — are referred to as source and smoothing (or smoothed) realizations, respectively.

The author of paper [4], having formulated a variational problem in realization space, solves it in a state-space. In our work, studying and solving problems of smoothing, filtration, and identification is also performed in space of realizations of source data. The fact that we consider the questions of smoothing, filtration, and identification in space of realizations of source data (for problems of the form (4)) for stationary systems allows us to obtain qualitatively new results. With the help of such approach we do not only get to obtain more simple and fast algorithms compared to those by T. Kailath, but also to realize nonlinear filtration equations.

Those include evaluation equations for coefficients of approximating stationary models in real time.

The basis of the results represented in the paper are counter equations of bilateral orthogonalization of homogeneous vector systems [5].

2. Preliminary information

2.1. **Three directions.** There are three main research directions forming the basis of studies developed in our paper. The author has obtained some particular results in every direction. The first of them deals with applications of the method of least squares in identification problems of dynamical systems, the second one – with propagation equations in homogeneous (isotropic or stationary) media and systems, and the last direction covers minimization on a sphere of pseudo quadratic forms $\rho^2 = \rho^2(\alpha) = \alpha^* Q(\alpha)\alpha$ which do not depend on the length of a generating vector. Here $Q(\alpha)$ is a positive-definite matrix nonlinearly dependent on α.

For the first direction (LSM), we find it necessary to mention (apart from the fundamental works by C.F. Gauss [6]) the book by Y.V. Linnik [7]. In the area of identification, the LSM is often treated in a simplified way [8].

Our results in the area of identification by the LSM considered in this paper include formulating and solving variational problems of form (4) [2,9,10]. These are problems of signal smoothing and identification of difference and differential equations of form (1), (2) from the represented class.

In our paper, the focus of attention is given to recurrent methods of solution of problems (4). We solve them sequentially based on growth of the dimension $k = \overline{0, L}$ of subspace of the models $\mathcal{D}_k$ in the space $E = E^{L+1}$ of the source data $\mathbf{y}$. A vector of source data $\mathbf{y}$ is projected on these subspaces, and their parameters are optimized in identification problems which we consider. We want to show that solving processes of variational problems (4) are limited to two types of problems and algorithms.

The first type is the above mentioned problems of sequential projecting. The realizations $\mathbf{y} \in E = E^{L+1}$ of the source data $\mathbf{y} = \mathbf{y}_L$ are projected on the kernels $\mathcal{D}_k = \mathcal{D}_k(\alpha)$ of the difference operators $\mathbb{D}_k = \mathbb{D}_k(\alpha)$ of form (2) in Euclidean space E. The second one includes minimization problems of squares of corresponding perpendicular distances of the form $\rho^2(\alpha)$ shown above. They are denoted as identification functionals $\widehat{J}_3$. (The functionals J_1, J_2 belong to more simple identification problems and will be characterized below.)

The identification functionals $\widehat{J}_3 = \rho^2(\alpha)$ are the values of functionals J in (4) on the projections $\widehat{\mathbf{y}}(\alpha)$ of the source realizations $\mathbf{y}$ on the kernels $\mathcal{D} = \mathcal{D}(\alpha)$ of the operators $\mathbb{D} = \mathbb{D}(\alpha)$ in (2,4). To solve the problems mentioned above, the two other research directions are used.

First, these are propagation equations (for waves, radiation, information) in isotropic media and stationary systems. The corresponding propagation equations are obtained by several groups of researchers in different application areas: from astro- and geophysics [11,12,13] to computer science and computational mathematics [14]. In particular, they are obtained as a result of analysis of propagation processes in homogeneous media as in layered structures [15].

Following T. Kailath [1], the algorithms of parameter optimization in approximation problems which do not use Ricatty equations in stationary dynamical systems are referred to as fast. Those are also known as Krein–Levinson algorithms [16,17] and Redheffer equations [15]. Corresponding algorithms are used in computational mathematics as well, e.g., in calculations of polynomials orthogonal on the unit circle [18], inversions of Toeplitz matrices, solving integral equations with kernels dependent on difference of arguments [17,19].

In [5], we have shown that different propagation equations in stationary and isotropic media and systems have a common foundation. All of them are consequences of equations of bilateral (counter) orthogonalization of homogeneous vector systems X in Hilbert spaces [20]. Systems of the form $X = |x_i|_0^L = \left|U^i x_0\right|_0^L$, where U are isometric operators on generated system X, are refereed to in this way. We have obtained these equations in paper [5] and called those counter equations. From those equations we obtained, in particular, new fast computational algorithms of smoothing, filtration, and identification in stationary dynamical systems. Some of the results are presented in this study.

Second, to solve identification problems, we used special iterative procedures with inverse matrix of the form

$$\alpha_{k+1} = Q^{-1}(\alpha_k)\alpha_k / \|Q^{-1}(\alpha_k)\alpha_k\|, \qquad k = 0, 1, 2, \ldots \tag{5}$$

With the help of this kind of iterations, identificational functionals of the form $\alpha^* Q(\alpha)\alpha$ [21,22] and proportions of quadratic forms [23, p. 281] are minimized. Such functionals were mentioned above. These functionals arise in problems of form (4) given variational estimation of unknown components of vector of coefficients α of the difference operator $\mathbb{D}(\alpha)$. Iterative procedures of the mentioned type generalize the known procedures with inverse matrix. The latter are applied in methods of orthogonal regression [24].

2.2. **Identification of difference equations.** It is generally accepted that the history of estimating the coefficients of such equations starts with the paper by a French physicist Gaspard Riche de Prony [25] (published in 1795). He came to solving this applied problem wishing to estimate the exponential and sine frequency with the help of a set of samplings. To solve this problem, Prony suggested a minimal (not overdetermined) system of linear proportions for residuals of difference equation in order to calculate from it the coefficients of the latter. These results were significant for the time. Apparently, that is why all the evaluation methods for coefficients of difference equations became known as Prony's methods (generalized, modified). However, it is more correct to call them Prony's problem rather then Prony's methods.

The first stage in development of the methods for solving Prony's problem were overdetermined systems $\overline{V}\overline{\alpha} = -\overline{\mathbf{y}}$ of $N+1$ algebraic equations for residuals [26]. If we apply LSM to this system of linear equations, then the identification functional will have the form $J_1(\alpha) = \left\|\overline{V}\overline{\alpha} + \overline{\mathbf{y}}\right\|^2$. In other words, $J_1(\alpha) = \|V\alpha\|^2 = \alpha^* V^* V \alpha$ given $\alpha_n = 1$. The solution is as follows: $\overline{\alpha} = -\left(\overline{V}^*\overline{V}\right)^{-1}\overline{V}^*\overline{\mathbf{y}}.$

The matrix V of long $(n+1)$-samples $v_i^* = \left|y_i^* \cdot \cdots, y_{i+n}^*\right|$, $i = \overline{0, N}$ is a $\big((N+1) \times (n+1)\big)$-matrix $V = \left|\overline{V}, \overline{\mathbf{y}}\right| = \{v_i^*\}_0^N$. Then the system of equations (2) can be written in the form $v_i^*\alpha = 0$, $i = \overline{0, N}$, or $V\alpha = 0$.

The matrix of all the states of system (2) in L-realization $\mathbf{y}$ is a $\big((N+2) \times n\big)$-matrix $\widetilde{V} = \{\overline{v}_i^*\}_{-1}^N$ of short n-samples-states $\overline{v}_i^* = s_i^* = \left|y_{i+1}^* \cdot \cdots, y_{i+n}^*\right|$, $i = \overline{-1, N}$. It is convenient to start counting them from a number -1. The initial state of the system (2) is $s_{-1}^* = \overline{v}_{-1}^* = \left|y_0^*, \cdots, y_{n-1}^*\right|$. The matrix $\overline{V}$ in the functional J_1

is a $\big((N+1)\times n\big)$-matrix of states $s_i = \overline{v}_i$: $\overline{V} = \{s_i^*\}_{-1}^{N-1}$. The states determine a predicting (recursive) recording of ordinary linear difference equations of the form (2). For example, from (2) we obtain $y_{i+n} = -\overline{v}_{i-1}^*\boldsymbol{\alpha} = -s_{i-1}^*\boldsymbol{\alpha}$, $i=\overline{0,N}$.

The state $\widehat{s}_k = \left|\widehat{y}_{k+1}^*, \cdots, \widehat{y}_{k+n}^*\right|^*$ for the current step k and the current smoothing realization $\widehat{\mathbf{y}}_{k+n}$ is also called current. What we mean is that the smoothing realization $\widehat{\mathbf{y}}_{k+n}$ contains the last smoothed sampling $\widehat{y}_{k+n}$. The corresponding recurrent process of sequential smoothing of the initial realization $\mathbf{y} = \{y_i\}_0^L$ is described in our paper.

The identification methods based on systems of linear algebraic equations of the form $V\alpha = 0$ became known as «generalized» Prony's methods. We also call them algebraic or open-loop. In this methods, differential equations are considered as a set of algebraic proportions which are not interrelated. Inverse connections and closeness of dynamical models described by difference equations are not taken into account.

Algebraic methods were widely used in 1950–70s. Problems of identifications were considered in a simplified way, and the capabilities of computing were limited. Therefore, ordinary and linear open-loop methods of identification attracted practitioners at the time. These methods are still widespread nowadays.

The second stage in formulating and solving mathematical problems of coefficient estimation for difference equations were the methods of orthogonal regression type. In these methods, along with errors in the right-hand sides $\overline{\mathbf{y}}$ of the system of linear equations $\overline{V}\overline{\alpha} = -\overline{\mathbf{y}}$ for residuals, the ones in the matrix $\overline{V}$ of a system $V\alpha = \left|\overline{V},\overline{\mathbf{y}}\right|\alpha = 0$ are also taken into account [27]. The problems of orthogonal regression by LSM can be formulated in the following way: minimize $J_{or} = J_2 = \left\|V - \widehat{V}\right\|^2$ given $\widehat{V}\alpha = 0$.

It is interesting that the transition to this class of methods of solving Prony's problem only requires a substitution of «identity» normalization of vector α in the problem formulation: from $\alpha_n = 1$ to $\|\alpha\| = 1$. This was done by Pisarenko [28]. The simplest form of identification functional in methods of orthogonal regression type $\big($the value of functional $J_{or} = J_2$ on «projections» $\widehat{V}\big)$ has the form $J_2(\alpha) = \alpha V^* V\alpha \big/ \alpha^*\alpha$. In other words, $J_2(\alpha) = \alpha V^* V\alpha$ given $\|\alpha\| = 1$. The solution (if it exists and is unique) is obtained by iterations of the form (5) with inversible matrix $Q = V^*V$.

Many authors came to the methods of solving the systems of linear equations from different perspectives, taking into account unrelated errors in the matrices of those systems. Although there are different terms used to refer to those, all of them denote the same kind of mathematical problems. They belong to the mentioned problems of orthogonal regression type [24] and are characterized by the main property: taking into account unrelated errors in the matrix of the system $V\alpha = \left|\overline{V},\overline{\mathbf{y}}\right|\alpha = 0$. Note the well-known terms: total LSM [29], Pisarenko method [28].

2.3. **Variational identification.** The next significant step in development of methods for solving Prony's problem — following the methods of orthogonal regression type [24], taking into account unrelated errors in components of matrices$V$ in systems of the form $V\alpha = 0$ — was the emergence of variational [2,21,22], or, in other terms, structural [30] methods. Such methods use additional prior information, in particular, structural dependences of rows and columns of matrices V, consisting of components of the realization $\mathbf{y}$ are considered. That helps to increase the reliability of evaluations of vector α of coefficients of the operator $\mathbb{D}$ in (4). The term 'reliability of evaluations' refers to the stability of the process of evaluating with respect to the level of errors in initial data $\mathbf{y}$. In problems of identification of the realization $\mathbf{y}$, depending on order of count of components, the matrices V are either Toeplitz or Hankel (equal numbers on some diagonals).

The mentioned facts result in problems of identification with a special non-quadratic identification functional of the form J_3. Those have been mentioned above and are referred to as pseudo quadratic. Consider the corresponding identification problems. Let $A = A(\alpha)$ be a band Toeplitz $\big((L+1)\times(N+1)\big)$–matrix with a generating vector α such that $A^*\mathbf{y} = V\alpha$. With the help of this matrix, a system of $(N+1)$ difference equation (2) for smoothed realizations $\widehat{\mathbf{y}}$ will be written in the form $\widehat{V}\alpha = 0 = A^*\widehat{\mathbf{y}}$. The identification functionals J_3 get the form $J_3(\alpha) = \alpha V^*(A^*A)^{-1}V\alpha$. The functional $J_3(\alpha)$ was denoted above as $\widehat{J}(\alpha) = \rho^2(\alpha) = \alpha^*Q(\alpha)\alpha$.

Such identification functional for solving Prony's problem was approached to from different directions independently and almost at the same time (around the year 1970) by three groups of researchers: Osborne [31], Aoki and Yue [32], and the author in paper [22]. In literature, this approach to identifying systems of the form (2) are referred to (in our opinion, unreasonably) as «modified» Prony's method.

We obtained the preudo quadratic functional J_3 [2],[21],[33] as a result of solving variational problems of the type (4) and called these finite-dimensional variational problems piecewise linear dynamical approximation. It is linear because linear dynamical models (2) are used. The sequences of samplings $\mathbf{y}$ are approximated by solutions $\widehat{\mathbf{y}}(\alpha)$ of linear difference equations in (2), (4) with constant coefficients $\alpha^* = \big|\alpha_0^*, \dots, \alpha_{n-1}^*, 1\big| = |\boldsymbol{\alpha}^*, 1|$ along with optimization of these coefficients, if all or part of them are unknown.

Formulation of this kind of problems is the following: to minimize with respect to $\widehat{\mathbf{y}}$ and to α the functional

$$J = \|\mathbf{y} - \widehat{\mathbf{y}}\|^2 \quad \text{given} \quad \mathbb{D}\widehat{\mathbf{y}} = 0,$$

$$\text{where} \quad \mathbb{D}\widehat{\mathbf{y}} = \left\{\sum_{i=0}^{n} \alpha_i^* \widehat{y}_{k+i}\right\}_0^N \longrightarrow \widehat{V}\alpha = A^*\widehat{\mathbf{y}} = 0. \tag{6}$$

Here $\mathbb{D} = \sum_0^n \alpha_i^* z^i$ is a difference operator, where z is an operator of the shift $zy_k = y_{k+1},\ \ k = \overline{0, N}$.

The matrices A and V are determined above. The matrix V is a Henkel matrix of samples, and the matrix A is a band Toeplitz matrix with a generating vector α. Such matrices guarantee that a simple but important for this paper identity holds: $A^*\mathbf{y} = V\alpha$. It is used in the last equality in (6). The image of the matrix A is shown below in the representation (8).

Let $\mathcal{D} = \mathcal{D}(\alpha)$ – a n-dimensional subspace – be the kernel $\mathcal{D}$ of the operator $\mathbb{D} = \mathbb{D}(\alpha)$. The solution of problem (6) constitutes a subspace $\widehat{\mathcal{D}} = \mathcal{D}(\widehat{\alpha}) \subseteq \Omega$, nearest to the initial realization $\mathbf{y}$. It is determined by the vector $\widehat{\alpha}$. The second element of the solution is a projection $\widehat{\mathbf{y}} = \widehat{\mathbf{y}}(\widehat{\alpha})$ of the vector $\mathbf{y}$ on this subspace. Of course, the search for such solutions is performed in reverse order. First, all $(\forall\alpha \in \omega)$ of the projections $\widehat{\mathbf{y}}(\widehat{\alpha})$ on subspaces $\mathcal{D}(\alpha)$, $\forall\alpha \in \omega$, are determined.

Next, among perpendiculars $\mathbf{y} - \widehat{\mathbf{y}}(\widehat{\alpha})$, the most minimal one with respect to the length $\|\mathbf{y} - \widehat{\mathbf{y}}(\widehat{\alpha})\|$ in Euclidean space $E = E^{L+1}$ is found.

Example 1. In R^3, we have that $\mathbf{y} = |y_0, y_1, y_2|$. The equation of the model has the form $\widehat{y}_1 + \overline{\alpha}\widehat{y}_0 = 0$, $\widehat{y}_2 + \overline{\alpha}\widehat{y}_1 = 0$. In other words, $A^T\widehat{\mathbf{y}} = V\alpha = 0$. Here $A^T = \begin{vmatrix} \overline{\alpha} & 1 & 0 \\ 0 & \overline{\alpha} & 1 \end{vmatrix}$, $V = \begin{vmatrix} y_0 & y_1 \\ y_1 & y_2 \end{vmatrix}$, $\alpha^T = |\overline{\alpha}, 1|$. It is obvious that $A_\perp(\alpha) \sim \mathcal{D}(\alpha) \sim \left|1, -\overline{\alpha}, \overline{\alpha}^2\right| \sim \widehat{\mathbf{y}}(\alpha)$. The latter means that $\widehat{\mathbf{y}} = \mathcal{D}(\alpha)x = \left|1, -\overline{\alpha}, \overline{\alpha}^2\right|^T x$, where $x \neq 0$ is a real number. Hence, in this case the functional $J_3 = \|\mathbf{y} - \widehat{\mathbf{y}}\|^2 = \alpha^T V^T \left(A^T A\right)^{-1} V\alpha = \mathbf{y}^T A \left(A^T A\right)^{-1} A^T \mathbf{y} = \mathbf{y}^T P(\alpha)\mathbf{y}$ must be minimized with respect to two variables: $\overline{\alpha}$ and x.

The set Ω of possible solutions of problem (6) is generated by all admissible subspaces of the model. They are given by the directions $\mathcal{D}(\alpha) = \left|1, -\alpha, \alpha^2\right|^T$ for all values of the vector α from the set of its admissible values ω. Let $\omega = (-\infty, \infty)$. Then $\forall \alpha \in \omega$, and the set Ω of subspaces of the model is a surface of a right circular cone. Its central angle equals 90^0, and the central axis is set by the direction $|1, 0, 1|^T$.

For an arbitrary vector of initial data $\mathbf{y}$, the nearest point on the surface Ω must be found. This point determines both generatrix $\mathcal{D}(\alpha)$ (and therefore, the vector α, which is the solution of the identification problem), and the projection $\widehat{\mathbf{y}} = \widehat{\mathbf{y}}(\alpha)$ on this generatrix which is a solution of smoothing problem. The image of this example is provided in papers [20], [33].

3. Homogeneous systems

3.1. **Systems of vectors.** Let the rows $X = X_N = |x_i|_0^N = |x_0, \ldots, x_N|$ denote the systems of $N+1$ elements of some space E with a scalar product. This may be both functional space and the space E^{M+1} of sequences of length $M+1$. Special cases: the spaces L^2 or l^2 respectively.

The elements x_i, $i = \overline{0, N}$, can consist of single vectors (functions) or vector m-blocks $x_i = |x_{ij}|_{j=1}^m$. Then $M + 1 \geqslant (N+1)m$. If all $(N+1)m$ elements of the system X are linearly independent, it is called nondegenerate.

From the denotations suggested above, we have that $X_{\overline{k,l}} = |x_k, \ldots, x_l| = |x_i|_k^l$ and $X_k = |x_0, \ldots, x_k| = |x_i|_0^k$, $k = \overline{0, N}$. Linear spans of subsystems $X_{\overline{k,l}}$ are denoted as $S_{\overline{k,l}} = S\left(X_{\overline{k,l}}\right)$ and $S_k = S\left(X_k\right)$. We denote the projectors on these subspaces as $P_{\overline{k,l}}$, P_k respectively. The projections on their orthogonal complements are denoted as $P_{\overline{k,l}} = I - \Pi_{\overline{k,l}}$, $\Pi_k = I - P_k$.

We determine scalar products of the vectors x and y in E as $\langle x, y\rangle$. If the elements x, y are finite sets of numbers, then we can assume that $\langle x, y\rangle_E = y^* x$. The arguments here can be both systems X of functions from E, and matrices X with columns from E. In all the cases, Gramian matrices of the systems X and its subsystems can be represented in the form $\Gamma = \langle X, X\rangle = \{\gamma_{ij}\}_0^N = \{x_j, x_i\}_0^N$, $\Gamma_{\overline{k,l}} = \left\langle X_{\overline{k,l}}, X_{\overline{k,l}}\right\rangle$, $\Gamma_k = \langle X_k, X_k\rangle$.

Consider the following equations of bilateral orthogonalizations of the vector systems X.

Lemma 1. *Suppose that there is a system of vectors X in a space E. The chain of orthoprojectors $\Pi_{k+1} = I - P_{k+1}$ on orthogonal complements of the chain of* *embedded supspaces* S_{k+1}, $k=\overline{-1,N-1}$ *is determined by the following counter equations:*

$$\Pi_{k+1}=\Pi_k-f_{k+1}a_{k+1}\langle\cdot\,,f_{k+1}\rangle=\Pi_{\overline{1,k+1}}-\widetilde{f}_{k+1}\widetilde{a}_{k+1}\left\langle\cdot\,,\widetilde{f}_{k+1}\right\rangle,\quad \Pi_{-1}=\Pi_{\overline{1,0}}=I.$$

Here Π_k *and* $\Pi_{\overline{k,l}}$ *are projectors on the orthogonal complements* $S_k^{\perp}$ *and* $S^{\perp}_{\overline{1,k+1}}$ *of the linear spans* S_k *and* $S_{\overline{1,k+1}}$, $f_{k+1}=\Pi_k x_{k+1}$, $\widetilde{f}_{k+1}=\Pi_{\overline{1,k+1}}x_0$, $a_{k+1}=\|f_{k+1}\|^{-2}$, $\widetilde{a}_{k+1}=\|\widetilde{f}_{k+1}\|^{-2}$.

Proof. The equations for orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$, provided in the lemma, are the Gram–Sonin–Schmidt formulas of bilateral orthogonalization [34]. The vector f_{k+1} is the last orthogonalizing vector in right orthogonalization of the system X_{k+1}. The vector $\widetilde{f}_{k+1}$ is the last orthogonalizing vector of counter orthogonalization of the system X_{k+1}. The equations of the lemma are based on the fact that projectors on a sum of orthohonal subspaces are sums of orthoprojectors on these subspaces. □

3.2. **Orthogonalization of homogeneous systems.** A system of vectors $X=X(B)$ of the form $X=X_N=|x_i|_0^N=|x_0,\dots,x_N|=|B^l x_0|_0^N$ in some space E is called Krylov basis. Systems of vectors $X=X(U)$ generated by operators $B=U$, isometric on $S=S(X)$, but in general, partially isometric in E, are referred to as homogeneous. The equations of counter orthogonalization of homogeneous systems are of a particular interest. They are obtained in work [5].

Theorem 1. *Orthogonalizing vectors* f_{k+1}, $\widetilde{f}_{k+1}$ *of equations from Lemma* **1** *of bilateral orthogonalization of a homogeneous nondegenerate vector system* $X=X(U)$ *satisfy the system of counter nonlinear difference equations.*

For $k=\overline{0,N}$, *this equations have the form*

(7)

a. $f_{k+1}=Uf_k-\widetilde{f}_k\widetilde{\theta}_k$, **b.** $\widetilde{f}_{k+1}=\widetilde{f}_k-Uf_k\theta_k$, *where*

c. $\theta_k=a_k\mu_k$, **d.** $\widetilde{\theta}_k=\widetilde{a}_k\widetilde{\mu}_k$, *moreover*

e. $a_{k+1}=(I-\theta_k\widetilde{\theta}_k)^{-1}a_k$, **f.** $\widetilde{a}_{k+1}=(I-\widetilde{\theta}_k\theta_k)^{-1}\widetilde{a}_k$, *and*

g. $\mu_k=\langle x_0,Uf_k\rangle=\left\langle\widetilde{f}_k,Uf_k\right\rangle$, **h.** $\widetilde{\mu}_k=\left\langle x_{k+1},\widetilde{f}_k\right\rangle=\left\langle Uf_k,\widetilde{f}_k\right\rangle=\mu_k^*$.

Initial conditions: $f_0=\widetilde{f}_0=x_0$, $a_0=\widetilde{a}_0=1\big/\|x_0\|^2$.

Corollary 1. *If* $m=1$, *then* $\widetilde{\theta}=\theta^*$, $\widetilde{a}_{k+1}=a_{k+1}=\left(1-\|\theta\|^2\right)^{-1}a_k$.

For simplicity, we assume that, in general, systems X are not block ($m=1$). Gramian matrices $\langle X,X\rangle$ of systems which are not block are Toeplitz ones.

We refer to as isometric operators U in E to operators that keep distances ρ in E: $\rho(Ux_1,Ux_2)=\rho(x_1,x_2)$. Examples include operators of rotation and reflection [23, p. 243]. Partially isometric operators U are the ones keeping distances $\rho(x_1,x_2)$ only for pairs of points x_1, x_2 from some subspace S_U in E. For example, such are the operators U concentrated in $S_U\subset E$ by domains $Dom\{U\}\subseteq S_U\subset E$ and ranges $Im\{D\}\subseteq S_{IU}\subset E$. Other examples include operators of transfer of subspace $S_U\subset E$ within E such that $S_U\subset E\to US_U=S_{IU}\subset E$. For the points outside of the subspace S_U, partially isometric operator acts as a zero one, hence, the subspace $E \ominus S_U$ is the kernel of partially isometric operator U. The generating operators U, isometric only on $S\big(X(U)\big)$, are called the minimal U_0.

The simplest partially isometric operators of transfer are operators of unit shift $I^1_M = I^1$ and their k-powers I^k. An operator of k-shift «downward» in $E = E^{M+1}$ has the form $I^k = \{\delta_{i,j+k}\}_0^M$. Its inverse (adjoint) operator I^{-k} of k-shift «upward» has the form $I^{-k} = \{\delta_{i+k,j}\}_0^M$. Its matrix is $(I^k)^* = (I^*)^k = I^{-k}$.

3.3. **Orthonormal homogeneous systems.** Here we propose a more general case provided in the beginning of section **3.1**. Nondegenerate homogeneous system $Y = Y_L(U) = \big|y_{[i]}\big|_0^L$ can be given either in some functional Hilbert space H (then $y_{[i]} = y_{[i]}(t) \in H$), or in Euclidean finite-dimensional space E (then $y_{[i]} \in E$). A system Y is called orthonormal, if its Gramian matrix Γ (in H or E) is an identity one: $\Gamma = \langle Y, Y\rangle = \big\{\big\langle y_{[i]}, y_{[j]}\big\rangle\big\}_0^L = \{\delta_{ij}\}_0^L$. It can be shown that orthonormal system is homogeneous [5].

Lemma 2. *Nondegenerate vector systems $Z = Z_L$ are reduced to homogeneous orthonormal systems $Y = Y_L$ and $\widetilde{Y} = \widetilde{Y}_L$ by direct F* (forward) *upper– and inverse $\widetilde{F}$* (backward) *lower–triangular transforms. For subspaces $S = S(Y)$ and $\widetilde{S} = S(\widetilde{Y})$, minimal partially isometric generating operators U_0 and $\widetilde{U}_0$ can be constructed. In particular, $U_0 = Y_{\overline{1,L}}\,\langle\,\cdot\,, Y_{L-1}\rangle$.*

Proof. Triangular transforms indicated in Lemma **2** represent the results of direct F and counter $\widetilde{F}$ processes of orthogonalization of the system Z by Gram–Sonin–Schmidt formulas shown in Lemma **1**. For example, the last column F_k of the matrix $F_{(k)}$, orthogonalizing the system $Z_k = |z_0, \cdots, z_k|$ by the mentioned formulas

$$Y_k = |f_0, \cdots f_k| = Z_k F_{(k)}$$

has the form $F_k = \begin{vmatrix} -\langle Z_{k-1}, Z_{k-1}\rangle^{-1}\langle z_k, Z_{k-1}\rangle \\ 1 \end{vmatrix}$, $k = \overline{1,L}$.

It corresponds to the direct Gram–Sonin–Schmidt transforms for calculating the orthogonalizing vectors $f_k = \Pi_{k-1} z_k = z_k - Z_{k-1}\langle Z_{k-1}, Z_{k-1}\rangle^{-1}$ $\langle z_k, Z_{k-1}\rangle$ from Lemma **1** for the system Z_k. The orthogonalization processes of system Z related to those can be written in the following way: $Y = ZF\mathbb{A}^{-1/2}$, and $\widetilde{Y} = Z\widetilde{F}\widetilde{\mathbb{A}}^{-1/2}$. Here $\mathbb{A} = \operatorname{diag}\{a_i\}_0^L = \langle ZF, ZF\rangle$, and $\widetilde{\mathbb{A}} = \operatorname{diag}\{\widetilde{a}_i\}_0^L = \Big\langle Z\widetilde{F}, Z\widetilde{F}\Big\rangle$.

We continue with an example of a direct transform. In paper [5], it is shown that the minimal for a homogeneous system $Z = Z_L$ partially isometric operator $U = U_0$ generating this system is determined by the formula $U_0 = Z_{\overline{1,L}}\Gamma^{-1}_{L-1}\langle\,\cdot\,, Z_{L-1}\rangle$. If the system Z is homogeneous, then $\Gamma_{L-1} = \langle Z_{L-1}, Z_{L-1}\rangle = \Big\langle Z_{\overline{1,L}}, Z_{\overline{1,L}}\Big\rangle = \Gamma_{\overline{1,L}}$. The converse is also true [5]. Let Y be a homogeneous orthonormal system. Then $\Gamma = I$ and the «minimal» partially isometric operator on the subspace $S(Y)$ is $U_0 = Y_{\overline{1,L}}\,\langle\,\cdot\,, Y_{L-1}\rangle$. □

Recall the following denotations. Functionals, adjoint vectors, and matrix constructions are marked with a superscript $()^*$, denoting a convolution of two involutions: transposition and complex conjugation. Adjoint spaces are marked with a prime $()'$.

Let $Z^* = \big\langle\,\cdot\,, Z\big\rangle = \{z_i^*\}_0^M = \big\{\big\langle\,\cdot\,, z_i\big\rangle\big\}_0^M \in E'$ be a linearly independent (nondegenerate) system of functionals z_i^*, $i = \overline{0,M}$ in E. As mentioned above,

E can be Euclidean or finite-dimensional functional space. In the latter case it can have elements from an independent system of functions in some Hilbert space H. The sequence $\mathbf{y}_Z = \{\langle y(t), z_i \rangle\}_0^M = \langle y(t), Z \rangle$ of values of functionals from Z^* on the function $y(t) \in E$ are called generalized samplings [9] of the function $y(t) \in E$ on a system (grid) of functionals $Z^* \in E'$. This $(M+1)$–vector $\mathbf{y} = \mathbf{y}_Z$ is referred to as a realization of the function $y(t) \in E$ on a grid Z^*.

Lemma 3. *Suppose that $Y = \left|y_{[i]}(t)\right|_0^L = Y(U) \subset E$ is an orthonormal system in $E = E^{M+1}$, $M \geqslant L$, and functions $\breve{y}(t)$ form its span – a subspace $S = S(Y) \subseteq E$. There exists an isomorphism $S \to E^{L+1}$, given which the system Y can be assosiated in $S = S_Y(Y) \sim E^{L+1}$ to the unit vector basis: $Y \leftrightarrow I_{L+1}$.*

Proof. Suppose that $\breve{y}(t) \in S(Y)$. Then by definition $S(Y)$, there exists a $(L+1)$-vector $\mathbf{y} \in E^{L+1}$ such that $Y\mathbf{y} = \breve{y}(t)$. Now suppose that $\mathbf{y} \in E^{L+1}$. Then we have $Y\mathbf{y} = \breve{y}(t) \in S(Y)$. Therefore, the linear subspace $S(Y)$ is isomorphic to the Euclidean space E^{L+1}. Let $\mathbf{y} = \langle y(t), Y \rangle$ be a realization of an arbitrary function $y(t) \in E$ on a grid of functionals $\langle\,\cdot\,, Y\rangle \in E'$. Then the mapping $y(t) \to \mathbf{y}$ has the inverse $\mathbf{y} \to Y\mathbf{y} = \breve{y}(t) \in E^{L+1}$. Hence, the mapping $\langle\,\cdot\,, Y\rangle$: $y(t) \to \mathbf{y}$ is the required isomorphism, since $\langle Y, Y\rangle = I_{L+1}$. □

Corollary 2. *The operator U in E is an operator U_Y of a shift «downwards» on a subspace $S = S_Y(Y) \sim E^{L+1} = E$. The matrix M_U of the operator U_Y in the basis Y on S is $M_U = I_{L+1}^{-1} = \begin{vmatrix} 0 & 0 \\ I_L & 0 \end{vmatrix} = \{\delta_{i,j+1}\}_0^L$. That is the minimal partially isometric operator U_0 on M_U.*

Proof. Taking into account the definition $M_U = \langle U y_{[j]}, y_{[i]} \rangle$ of the matrix of the operator of the shift $U_Y = \{U_{Y/ij}\}_0^L$ in the basis $Y = \left|y_{[i]}\right|_0^L$ through its action on basis vectors, we obtain $M_U = \langle U y_{[j]}, y_{[i]} \rangle = \langle y_{[j+1]}, y_{[i]} \rangle = \delta_{i,j+1}$. In other words, in the basis Y of the subspace $S = S_Y(Y) \sim E = E^{L+1}$ (Lemma **3**), we have $Y_{L-1} \leftrightarrow |e_0, \ldots, e_{L-1}|$, and $Y_{\overline{1,L}} \leftrightarrow |e_1, \ldots, e_L|$. By Lemma **2**, $U_0 = Y_{\overline{1,L}} \langle\,\cdot\,, Y_{L-1}\rangle = I_{L+1}^{-1} = \{\delta_{i,j+1}\}_0^L$. □

The following lemma is obvious.

Lemma 4. *Let $y(t)$ be an arbitrary function from E, and $\mathbf{y}_k = \langle y(t), Y_k\rangle \in E^{k+1}$, $k = \overline{0, L}$, be its generalized samplings on a grid of functionals $Y_k^* = \langle\,\cdot\,, Y_k\rangle$. Then the functions $\breve{y}_k(t) = Y_k \mathbf{y}_k$, are the projections of the function $y(t) \in E$ on the subspaces $S(Y_k)$.*

3.4. **Polynomial homogeneous systems.** Let the system $Y = Y_L(U) = \left|y_{[i]}\right|_0^L = \left|y_{[i]}(t)\right|_0^L \in E$, $L \leqslant M$ be orthonormal. The generating operator U is, for example, the minimal isometric operator U_0 from Lemma **2**. The homogeneous system $X = X_N = \left|x_{[i]}\right|_0^N = \left|U^i x_{[0]}\right|_0^N$, $x_{[i]} = x_{[i]}(t)$, $N = L - n$, is called polynomial (n-polynomial), if its initial vector $x_{[0]}$ is a result of an action of n-polynomial $\sum_0^n U^j \alpha_j$ from U on the initial vector $y_{[0]}$ of the system $Y \in E$.

Lemma 5. *Suppose that $\alpha_{[0]} = |\alpha_0^*, \ldots, \alpha_n^*, 0_N^T|^*$. Polynomial vectors $x_{[i]}(t) = U x_{[i-1]} = U^i x_{[0]}$ from $X = S_Y(Y) \sim E$, where $i = \overline{1,k}$, $k \in \overline{1,N}$, form in X a basis of sliding vector (with a band–Toeplitz type basis matrix), if $x_{[0]}$ in coordinates Y is $x_{Y[0]} = \alpha_{[0]}$.*

Proof. We assume that $\alpha_i = 0$, given $i > n$. Then the coordinates of the vector $x_{[0]}$ in the basis Y $\big($on the grid of the functionals $Y^* = \langle\,\cdot\,, Y\rangle\big)$ is

$$x_{Y[0]} = \big\langle x_{[0]}, Y\big\rangle = \big\{\big\langle x_{[0]}, y_{[i]}\big\rangle\big\}_0^L =$$
$$\Big\{\Big\langle \textstyle\sum_{j=0}^n y_{[j]}\alpha_j, y_{[i]}\Big\rangle\Big\}_{i=0}^L = \Big\{\textstyle\sum_{j=0}^n \delta_{ij}\alpha_j\Big\}_{i=0}^L = \{\alpha_i\}_0^L = \alpha_{[0]}.$$

The matrix of the operator U in the orthonormal basis Y is a matrix of the shift $I^1 = \{\delta_{i,j+1}\}_0^L$, such that $I^1 e_j = e_{j+1}$ (Corollary **2**). Hence, in the basis Y for n-polynomial systems we have that $X_{k/Y} = \langle X_k, Y\rangle = A_k = A(\alpha) = \big|I^j\alpha_{[0]}\big|_0^k$ and $X_k = YA_k$ in the basis E.

The basis matrices $A_k = A_k(\alpha)$ are band–Toeplitz type matrices. They are referred to as matrices of a sliding vector α (MSV α). We now show the matrices A_k^*:

$$A_k^*(\alpha) = \left\|\begin{array}{c|c} \begin{matrix} \alpha_0^* & \dots & \alpha_n^* & & & & \\ & \ddots & \vdots & \ddots & & \mathbf{0} & \\ & & \alpha_0^* & \ddots & \alpha_n^* & & \\ & \mathbf{0} & & \ddots & \vdots & \ddots & \\ & & & & \alpha_0^* & \dots & \alpha_n^* \end{matrix} & \mathbf{0} \end{array}\right\| \in \big((k+1)\times(L+1)\big). \tag{8}$$

The blocks $\|0\|$ in the matrix A_k (8) are zero. The matrix A_k represents the system of vectors in E and the operator $A_k : E^{k+1} \to E = E^{L+1}$, $k = \overline{0,N}$. We introduce similar systems and operators in E^{k+n+1}. In those, zero blocks of coordinates are absent.

We denote first $(k+n+1)$ rows of an identity matrix I_{L+1} by $\widetilde{I}_{k+n+1}$. The matrix $\widetilde{A}_k = \widetilde{I}_{k+n+1}A_k$ is a system of $(k+1)$ vectors in E^{k+n+1}. This is also a matrix of the sliding vector (MSV α). For basis matrices of n-polynomial systems, we have the following equalities:

$$\begin{gathered} X_k = YA_k = Y_{k+n}\widetilde{A}_k, \quad \text{where} \\ A_k = \left|\begin{matrix}\widetilde{A}_k \\ \{0\}^T\end{matrix}\right|, \quad \widetilde{A}_k = \widetilde{I}_{k+n+1}A_k\,, \quad \text{and} \quad \widetilde{I}_{k+n+1} = |I_{k+n+1}, \{0\}|\,. \quad \Box \end{gathered} \tag{9}$$

3.5. **Difference equations.** From (2), we have that

$$\mathbb{D}_k(\alpha)\widehat{\mathbf{y}}_{k+n} = \big\{D(\alpha)\widehat{y}_i\big\}_0^k = \left\{\sum_{j=0}^n \widehat{y}_{i+j}\alpha_j^*\right\}_{i=0}^k = \left\{\sum_{j=0}^n U^{i+j}\widehat{y}_0\alpha_j^*\right\}_{i=0}^k = 0. \tag{10}$$

Suppose that $Y_{k+n}(U) = |y_i|_0^{k+n} \in E = E^{M+1}$, $k = \overline{0,N}$, $N = L-n$, $L \leqslant M$, is an orthonormal system, and $S\,(Y_{k+n}) \subseteq E$ is its linear span. Let $X_k = X_k(\alpha) = \big|x_{[i]}\big|_{i=0}^k = \big|U^iD(\alpha)y_{[0]}\big|_{i=0}^k \in S\,(Y_{k+n})$ be a polynomial system, determined in the previous section, and $S_k = S_k(\alpha) = S(X_k) \subseteq S\,(Y_{k+n})$ be a subspace corresponding to it. The «geometrical» meaning of equations of the form (10) is given by the following propositions.

Theorem 2. *A function $\check{y}(t) \in S\,(Y_{k+n})$ is orthogonal to a polynomial subspace S_k, if and only if a subsequence $\mathbf{y}_{k+n}$ of its generalized samplings is such that $\mathbf{y}_{k+n} = \widehat{\mathbf{y}}_{k+n}$, and $\widehat{\mathbf{y}}_{k+n} = \langle\check{y}(t), Y_{k+n}\rangle \in E^{k+n+1}$ satisfies the equation* (10): $\mathbb{D}_k(\alpha)\widehat{\mathbf{y}}_{k+n} = 0$.

We determine a scalar product in E as $\langle\cdot,\cdot\rangle$.

Proof. Necessity. Suppose that $\langle\breve{y}(t), X_k\rangle = 0$. That is, the fact that the functions $\breve{y}(t)$ are orthogonal to the subspaces S_k for $k = \overline{0,N}$ is valid. Both functions $\breve{y}(t)$ and projections $\breve{y}(t) = P_{k+n}y(t)$ of the functions $y(t) \in E$ on the subspaces $S(Y_{k+n})$ can be represented in the form $\breve{y}(t) = Y_{k+n}\mathbf{y}_{k+n}$ (Lemma **4**), where (if $\breve{y}_{k+n}(t)$ is a projection) $\mathbf{y}_{k+n} = \langle y(t), Y_{k+n}\rangle \in E^{k+n+1}$. Moreover, as shown above (9), $X_k = YA_k = Y_{k+n}\widetilde{A}_k$.

Taking into account definitions (8) and (9), the condition $\breve{y}_{k+n}(t) \perp X_k$ means that

$$\left\langle Y_{k+n}\mathbf{y}_{k+n}, Y_{k+n}\widetilde{A}_k\right\rangle = \widetilde{A}_k^* \langle Y_{k+n}, Y_{k+n}\rangle \mathbf{y}_{k+n} = \left\langle \mathbf{y}_{k+n}, \widetilde{A}_k\right\rangle = 0. \tag{11}$$

This implies equation (10) for the realization $\mathbf{y}_{k+n}$.

Sufficiency. Suppose that equations (10) hold for the realization $\widehat{\mathbf{y}}_{k+n}$. This means that $\left\langle \widehat{\mathbf{y}}_{k+n}, \widetilde{A}_k\right\rangle = 0$, or

$$\widetilde{A}_k^*\widehat{\mathbf{y}}_{k+n} = \widetilde{A}_k^* \langle Y_{k+n}, Y_{k+n}\rangle \widehat{\mathbf{y}}_{k+n} = \left\langle Y_{k+n}\,\widehat{\mathbf{y}}_{k+n}\,, Y_{k+n}\widetilde{A}_k\right\rangle = \langle\breve{y}(t), X_k\rangle = 0.$$

Then it follows that $Y_{k+n}\,\widehat{\mathbf{y}}_{k+n} = \breve{y}(t) \perp X_k$, which was required. □

Corollary 3. *The matrices of the operators $\mathbb{D}_k : E^{k+n+1} \to E^{k+1}$, $k = \overline{0,N}$, are the constructed above in* (3) *matrices $\widetilde{A}_k^*$ of the sliding vector α^*.*

Proof. The result follows from equalities (9), equation (10) and Theorem **2**. □

It is obvious that the matrices $\widetilde{A}_k(\alpha)$ from (8) are matrices of the operators $\mathbb{D}'_k : E^{k+1} \to E^{k+n+1}$.

Corollary 4. *The subspace $S = S(Y_{k+n}) \subseteq E$ is a direct sum of two orthogonal subspaces (polynomial and operator): $S(Y_{k+n}) = S(X_k) \oplus Y_{k+n}\mathcal{D}_k$. Here $\mathcal{D}_k$ are the kernels of the operators $\mathbb{D}_k :\ E^{k+n+1} \to E^{k+1}$ of equations* (10).

Proof. The elements of the kernels $\mathcal{D}_k \subset E^{k+n+1}$ of the operators $\mathbb{D}_k$: $E^{k+n+1} \to E^{k+1}$ are the vectors $\widehat{\mathbf{y}}_{k+n} \in E^{k+n+1}$, satisfying equations (10). The kermels $\mathcal{D}_k$ have dimension n, which is the number of initial conditions in the solutions of equations (10), and $\dim S(\widetilde{A}_k) = k+1$ (Lemma **5**). From equations (10), it is clear that $S \sim E^{k+n+1}$ (Lemma **3**) consists of orthogonal complements: $E^{k+n+1} = S\big(\widetilde{A}_k\big) \oplus \mathcal{D}_k$. Due to the fact that the systems Y_{k+n} are orthonormal, we have that $S(Y_{k+n}) = Y_{k+n}E^{k+n+1}$. Then it follows that $S(Y_{k+n}) = S(X_k) \oplus Y_{k+n}\mathcal{D}_k$. □

Further, we assume that $Y = I$, $S(Y) = E$ and $X_k = A_k$.

3.6. **Variational problems on homogeneous systems.** We denote difference operators of the form (10) by $\mathbb{D}_k$, and their kernels (subspaces of dimension n in E) by $\mathcal{D}_k$. Suppose that $\mathbf{y} = \{y_i\}_0^L$. It is possible, in particular, that $\mathbf{y} = \langle y(t), Y\rangle$. Here $Y = Y_L = \left|y(t)_{[i]}\right|_0^L$ is an independent vector system in $E = E[T]$ on a finite interval $I_T = [0,T]$, and $y(t) \in E$. We introduce so called smoothing realizations $\widehat{\mathbf{y}}_{k+n} = \left\{\widehat{y}_{i/k}\right\}_{i=0}^{k+n}$. They minimize the following functionals given the mentioned conditions:

$$J_k = \|\mathbf{y} - \widehat{\mathbf{y}}_{k+n}\|^2, \quad \text{where} \quad \widehat{\mathbf{y}}_{k+n} = \left|\widehat{\mathbf{y}}^*_{[k+n]}, \mathbf{y}^*_{\overline{k+n+1,L}}\right|^*$$

$$\widehat{\mathbf{y}}_{[k+n]} \in \mathcal{D}_k \;\to\; \mathbb{D}_k(\alpha)\widehat{\mathbf{y}}_{[k+n]} = 0, \tag{12}$$

$$\alpha \in \omega \subseteq E^{n+1}, \quad k = \overline{-1, N}, \quad N = L - n.$$

The realizations $\widehat{\mathbf{y}}_{[k+n]}$ lying in the smoothing realizations $\widehat{\mathbf{y}}_{k+n}$ are called smoothed.

There are two variational problems in equalities (12): a linear and nonlinear ones.

Problems 1 (smoothing and filtration). *For $k = \overline{-1, N}$, find the minimum $\widehat{J}_k(\alpha)$ of functionals $J_k(\widehat{s}_0, \alpha)$ from* (12) *given the conditions mentioned in* (12). *The minima are found based on initial conditions (states) $\widehat{s}_{-1} = \widehat{s}_{-1/k}$ of the smoothed realizations $\widehat{\mathbf{y}}_{[k+n]}(\widehat{s}_{-1})$ which are the solutions of equation in* (12). *The coefficients α of equations* (12), (10) *in these cases are considered known.*

Problems **1** for the steps $k = \overline{-1, N-1}$ are the elements of a sequential solution of problem **1** for the step $k = N$. The functionals $\widehat{J}_k(\alpha)$ are the values of the functionals J_k on the solutions of problems **1** which are the smoothing realizations $\widehat{\mathbf{y}}_{k+n}(\widehat{s}_0)$.

Problems 2 (identification). *For $k = \overline{k_{\min}, N}$, find in a region ω of admissible coefficients α of the operators $\mathbb{D}_k$ the arguments $\widehat{\alpha}_{(k)}$ of global minima by α of the functionals $\widehat{J}_k(\alpha)$ in* (12)

Here k_{min} is the minimal value of k, with which **Problem 2** has a unique solution. From the system of equations (10),(12), we can see that the minimal value of k_{min} is $2n$. This is valid in the case when first $2n$ samplings of the initial realization $\mathbf{y}$ are such that the minimal (of n equations) system of (10), (12), formed from these samplings, has a unique solution.

Problems **1** and **2** are referred to as piecewise linear dynamical approximation problems. These are the problems of approximation of functions on finite intervals by solutions of linear dynamical models.

Problems **1** are called smoothing. These are problems of orthogonal projecting on subspaces, i.e. the kernels $\mathcal{D}_k(\alpha)$ of the operators $\mathbb{D}_k(\alpha)$ (10) and (12). In this case, both smoothing realizations $\widehat{\mathbf{y}}_{k+n}$ and n-vectors $\widehat{s}_i = \left\{\widehat{y}_{i+j/k}\right\}_{j=1}^{n}$, $i = \overline{-1, k}$, can be calculated. The latter ones are the states of the model (10), (12) at time i. The vectors of states of this model are described in Section **2.2**. The value $i = -1$ gives the initial state $\widehat{s}_{-1} = \left|\widehat{y}^*_0, \cdots, \widehat{y}^*_{n-1}\right|^*$ of realizations $\widehat{\mathbf{y}}_{k+n}(\widehat{s}_{-1})$. With $i = k$, we obtain filtration problems, i.e. calculations of the current states $\widehat{s}_{k/k} = \widehat{s}_k$ of model (10), (12) from realizations $\widehat{\mathbf{y}}_{k+n}(\widehat{s}_{-1})$. Note that the initial state on the step -1 of solving problems (3) consists of initial samplings: $\widehat{s}_{-1/-1} = s_{-1/-1} = \left|y^*_0, \cdots, y^*_{n-1}\right|^*$. On step k, we have that $\widehat{s}_{-1/k} = \left|\widehat{y}^*_{0/k} \cdots \widehat{y}^*_{(n-1)/k}\right|^*$, $\widehat{s}_{k/k} = \left|\widehat{y}^*_{(k+1)/k} \cdots \widehat{y}^*_{(k+n)/k}\right|^*$.

Problems **2**, referred to as identification, search for a subspace $\mathcal{D}_k(\alpha)$ with parameters $\alpha \in \omega$ from their admissible set ω, nearest to the realization $\mathbf{y}$. Problems **2** constitute minimization of the lengths ρ_k of corresponding perpendiculars. It is clear that $\rho_k^2 = \widehat{J}_k(\alpha)$.

Remark 2. Problems **2** can be treated as problems of generalized projecting. By generalized projecting in a normed space E we mean searching for an element in some set $\Omega \subset E$, nearest to the given vector $\mathbf{y}$; by projection of an element $\mathbf{y} \in E$ on a set $\Omega \subset E$ we mean a subspace $\widehat{\mathcal{D}}(\widehat{\alpha})$ and an element $\widehat{\alpha}$ from $\widehat{\mathcal{D}} \subset \Omega$, nearest to $\mathbf{y}$. In problems **2**, for every value of k, it is a set $\Omega = \Omega_k$ of subspaces $\mathcal{D}_k(\alpha)$ for $\forall \alpha \in \omega$.

We would like do make several remarks on identification problems (12). In Remarks **3** and **4**, we indicate equaivalent (12) formulations of variational problems of smoothing and identification of stationary systems. Remark **6** specifies the open-loop algebraic formulations of problems of estimation of coefficients α of difference equations (12), mentioned in Section **2.2**.

Remark 3. Suppose that $x_{k+1} = Bx_k$, $y_{k+1} = b^* x_{k+1}$. Then, similarly to (12), the problem of smoothing of the sequence $\mathbf{y} = \{y_i]\}_0^L$ and identification of the dynamical system $\{B, b\}$ can be formulated in the following way: minimize J from (12) given $\widehat{\mathbf{y}} = \left\{b^* B^i\right\}_0^L x_0 = W x_0$. The solution of the smoothing problems is as follows: $\widehat{x}_0 = (W^* W)^{-1} W^* \mathbf{y}$, $\widehat{\mathbf{y}} = W \widehat{x}_0$. This formulation of problems of smoothing and identification is equivalent to (12). Solving problems of identification (optimization of B, b to achieve the best approximation) is complicated.

Remark 4. Formulations of optimization problems for a stationary system mentioned in Paragraph 1, closest to (12), can be written the following way. Let $S(*)$ be a subspace spanned by arguments in $(*)$ and suppose that $k = N$. We denote by $W_{N,\perp} = W_\perp$ any system of $N = L - n$ vectors in the subspace $S_\perp(W)$, that is, such that $W_\perp W = 0$. For example, $W_\perp = \left|-\overline{W} W_0^{-1}, I_N\right|$. Here W_0 is an initial $(n \times n)$-block in W, and $\overline{W}$ is a remaining part of the matrix W, that is, $W^* = \left|W_0^*, \overline{W}^*\right|$. Now problem (12) can be equivalently formulated the following way: minimize J given the condition that $W_\perp \widehat{\mathbf{y}} = 0$. Solving identification problems here is also complicated.

Remark 5. The conditions for realization $\widehat{\mathbf{y}}$ in the process of minimization of functional J, mentioned in Paragraphs **1** and **2** of this remark, are equivalent to the conditions formulated in problems (12). However, they require more knowledge about the dynamical system used as a model. Thanks to Cayley–Hamilton theorem, in problems (12) it is only required to know the vector of coefficients of characteristic equation of the system. Moreover, relatively uncomplicated method of its optimization can be suggested (**problem 2**), which will be shown in this article.

Remark 6. Another alternative to problems (12) are algebraic methods using the following non-approximation criterion of optimality: $J = \|V\alpha\|^2$. Formulations of this kind arise when considering sequences of dependent difference relations in (12) as algebraic system of linear independent equations. This system is solved with respect to vector α by formulas of a linear LSM, mentioned in Section **2.2**. □

4. Orthogonal projecting

4.1. **Analytical solution to smoothing problems.**

Theorem 3. a. *The components* $\widehat{y}_{i/k+n}$ *for* $i = \overline{k+n+1, L}$ $\;0 \leqslant k \leqslant N-1$ *of smoothing realizations* $\widehat{\mathbf{y}}_{k+n}$ *which are the solutions of problems* (12) *equal the samplings* y_i *of initial realizations* $\mathbf{y} = \{y_j\}_0^L$*:* $\widehat{y}_{i/k+n} = y_j$, *if* $j = i = \overline{k+n+1, L}$.

b. *The solutions of problems of smoothing and filtration (problems* **1**) *are given by the formulas of orthogonal projecting of the vector* $\mathbf{y}$ *on the kernels* $\mathcal{D}_k$ *of the operators* $\mathbb{D}_k = \mathbb{D}_k(\alpha) : E^{k+n+1} \to E^{k+1}$.

Proof. **P°.** The components $\widehat{y}_{i/k+n}$ of the vector $\widehat{\mathbf{y}}_{k+n}$ for $i = \overline{k+n+1, L}$ given $0 \leqslant k \leqslant N-1$ do not fall under the restrictions of minimization of the functional J_k in (12). It follows from equation (10). Hence, the minimization of the functionals J_k in E^{L+1} with respect to conditions (12) provides equality of these components to the ones of the initial realization $\mathbf{y}$.

b. The solutions $\widehat{\mathbf{y}}_{k+n} = \widehat{\mathbf{y}}_{k+n}(\alpha)$ of problem **1** are orthogonal projections of the vector $\mathbf{y}$ on the kernels $\mathcal{D}_k = \mathcal{D}_k(\alpha)$ of the operators $\mathbb{D}_k = \mathbb{D}_k(\alpha)$ in E $(10, 12)$. This is true because that is the way optimization problem **1** is formulated. A projection on $\mathcal{D}_k = S(\Psi)$, where Ψ is an arbitrary basis in $\mathcal{D}_k$, can be found in the form $\widehat{\mathbf{y}} = \Psi x$ in the process of minimization of the functional $\|\mathbf{y} - \widehat{\mathbf{y}}\|^2$. Differentiating it by x, we obtain the known formula of the LSM: $x = \langle \Psi, \Psi \rangle^{-1} \langle \mathbf{y}, \Psi \rangle$. Then, it follows that $\widehat{\mathbf{y}} = P\mathbf{y}$, where $P(\Psi) = \Psi \langle \Psi, \Psi \rangle^{-1} \langle \cdot\,, \Psi \rangle$. □

The next result provides us with a possibility to use the counter equations from Lemma **1** and Theorem **1** when solving problems **1**.

Lemma 6. *Solutions of problem* **1** *are determined by the following projectors and projections on the orthogonal complements $\mathcal{D}_k^{\perp}$ of the kernels $\mathcal{D}_k$:*

$$\begin{gathered} P_k = P(\Psi_k) = I - P(A_k) = \Pi(A_k) = \Pi_k, \ \longrightarrow \\ \widehat{\mathbf{y}}_k(\alpha) = \Pi_k \mathbf{y} = \mathbf{y} - P(A_k)\mathbf{y} \ \Rightarrow \ \Delta\widehat{\mathbf{y}}_k(\alpha) = \mathbf{y} - \widehat{\mathbf{y}}_k(\alpha) = P(A_k)\mathbf{y}. \end{gathered} \tag{13}$$

Proof. It follows from Theorem **2** and Corollary **4** that the kernels $\mathcal{D}_k$ of difference operators $\mathbb{D}_k$ are the orthogonal complements $S^{\perp}\big(A_k(\alpha)\big) \subset E^{k+n+1} = E_k$ of linear spans $S\big(A_k(\alpha)\big) \subset E^{k+n+1} = E_k$ of homogeneous systems $A_k(\alpha)$ determined in Lemma **5**. □

Solving problems of smoothing based on counter equations. Formulas (13) imply the possibility of using the counter equation of Theorem **1** for projecting, because the bases A_k are generated by operators of shift.

Theorem 4. *Counter equations* (7) *of bilateral orthogonalization of homogeneous systems imply recurrent equations for sequential solutions of smoothing problems.*

Proof. We denote (4) $\alpha^* = |\boldsymbol{\alpha}^*, 1|$ and suppose that $k \in \overline{-1, N-1}$. From formulas (13) and equations from Lemma **1** for the smoothing realization $\widehat{\mathbf{y}}_{k+n+1}$, we obtain the following equations:

$$\begin{aligned} &\widehat{\mathbf{y}}_{k+n+1} = \Pi_{k+1}\mathbf{y} = \widehat{\mathbf{y}}_{k+n} - f_{k+1}a_{k+1}\pi_{k+1}, && \text{where} && \pi_{k+1} = \langle \mathbf{y}, f_{k+1} \rangle \implies \\ &\pi_{k+1} = \langle \widehat{\mathbf{y}}_k, x_{k+1} \rangle = y_{k+n+1} - \widehat{y}_{(k+n+1/k)}, && \text{and} && \widehat{y}_{(k+n+1/k)} = -\boldsymbol{\alpha}^*\widehat{s}_k. \end{aligned} \tag{14}$$

The latter formula is a consequence of Lemma **5**, and $\widehat{s}_k = \big|\widehat{y}^*_{k+1}, \cdots, \widehat{y}^*_{k+1}\big|^*$ is a current condition of system (2).

The structure of the variable π_{k+1}, $k = \overline{-1, N-1}$ and its place in equations of smoothing (14) allow to consider the sequence of variables π_{k+1} as a renewal process. The first equation in (14) is a renewal equation in a smoothing problem. The variable π_{k+1} contains new information, i.e. a difference between a new sampling y_{k+n+1} and its prediction. It determines the renewing summand $f_{k+1}a_{k+1}\pi_{k+1}$, which corrects the smoothing realization $\widehat{\mathbf{y}}_{k+n}$ under renewal to calculate the renewed realization $\widehat{\mathbf{y}}_{k+n+1}$.

Formulas (14) imply that both the process of renewal of π_{k+1} and the components $\widehat{y}_{i/k+n+1}$, $i \leqslant k+n+1$ of the realization $\widehat{\mathbf{y}}_{k+n+1}$ in the left-hand side of equation (14) are determined by the counter equations from Theorem **1** for the orthogonalizing vectors f_{k+1}. □

Partially smoothed realizations. It follows from Theorems **3** and **4** that the realizations $\widehat{\mathbf{y}}_{k+n+1}$ under renewal in equations (14) are partially smoothed realizations. They have the form, indicated in (12):

$$\widehat{\mathbf{y}}_{k+n+1} = \left| \left|\widehat{y}^*_{i/k+1}\right|_{i=0}^{k+n+1}, \left|y^*_i\right|_{k+n+2}^{L} \right|^* = \left|\widehat{\mathbf{y}}^*_{[k+n+1]}, \mathbf{y}^*_{\overline{k+n+2,L}}\right|^*, \tag{15}$$

where $-1 \leqslant k \leqslant N-1$. We will refer to the corrected part $\widehat{\mathbf{y}}_{[k+n+1]}$ of the smoothing realization $\widehat{\mathbf{y}}_{k+n+1}$ under renewal in equation (14) as a smoothed realization. The remaining part $\mathbf{y}_{\overline{k+n+2,L}}$ of the smoothing realization $\widehat{\mathbf{y}}_{k+n+1}$ consists of initial samplings (Theorem **3**). It constitutes the initial part of the smoothing realization.

Equations (14) and formula (15) imply that the components of the orthogonalizing vectors f_{k+1}, whose numbers $\overline{r+n+2,L}$ correspond to unchangeable components of the smoothing realizations $\widehat{\mathbf{y}}_{k+n+1}$ in equations (14),(15) should be zero. Then it follows that the orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$ have the structure similar to that of the smoothing realizations $\widehat{\mathbf{y}}_{k+n+1}$ shown in (15). In particular, the orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$ consist of two subvectors, similarly to the smoothing realizations $\widehat{\mathbf{y}}_{k+n+1}$. However, the remaining subvectors of the orthogonalizing vectors consist of components which equal zero.

$$f_{k+1} = \left|f^*_{[k+1]}, 0^T_{\overline{k+n+2,L}}\right|^*, \qquad \widetilde{f}_{k+1} = \left|\widetilde{f}^*_{[k+1]}, 0^T_{\overline{k+n+2,L}}\right|^*. \tag{16}$$

We call the first subvectors of the orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$ (of direct and counter orthogonalization) vectors of orthogonalizations and denote them, similarly to the smoothed realizations, by indices in straight brackets.

By the action of equation (14), the smoothed part $\widehat{\mathbf{y}}_{[k+n+1]} = \{\widehat{y}_{j/k+n+1}\}_{j=0}^{k+1+n}$ of the smoothing realization $\widehat{\mathbf{y}}_{k+n+1} = |\widehat{\mathbf{y}}^*_{[k+n+1]}, \mathbf{y}^*_{\overline{k+n+2,L}}|^*$ elongates by one position, and the part $\mathbf{y}_{\overline{k+n+2,L}} = \{y_j\}_{k+n+2}^{L}$ of initial samplings becomes one position shorter compared to the realization $\widehat{\mathbf{y}}_{k+n}$. The same can be said about the orthogonalizaing vectors f_{k+1} and $\widetilde{f}_{k+1}$ from equalities (16).

4.2. **Filtration mode of orthogonalization equations.** Equation (2) can be considered as an equation for a variable y of a dynamical system of order n with a canonical transformation matrix, which is a Frobenius matrix observed in discrete time [23,34]. Our further goal is to obtain filtration equations for the states $\widehat{s}_{k+1} = \left\{\left|\widehat{y}^*_{k+1+i}\right|_{i=1}^{n}\right\}^*$, $k = \overline{-1, N-1}$, of such system. The first component of the state $\widehat{s}_{k+1}$ equals the sampling $\widehat{y}_{k+2}$ of the smoothing realization $\widehat{\mathbf{y}}_{k+n+1}$. Therefore, the state $\widehat{s}_{k+1}$ constitutes the last n components of the smoothed realization $\widehat{\mathbf{y}}_{[k+n+1]}$ within this realization.

As it follows from (14), derivation of equations for the states s_{k+1} assumes figuring out the possibility and obtaining counter equations for the samples h_{k+1} under renewal from the orthogonalizing vectors f_{k+1} based on equations (7). Equations (7) imply that in order to do that, it is also necessary to introduce n-samples $\widetilde{h}_{k+1}$ from the vectors $\widetilde{f}_{k+1}$ of counter orthogonalization. The states h_{k+1} and $\widetilde{h}_{k+1}$ of the orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$ are the last n components in the vectors of orthogonalization $f_{[k+1]}$ and $\widetilde{f}_{[k+1]}$ of the orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$. These structures are shown in formulas (16).

Hence, for the step $k+1$ we introduce the n-vectors $h_{k+1} = \left\{f_{k+1+j/k+1}\right\}_{j=1}^{n}$ and $\widetilde{h}_{k+1} = \left\{\widetilde{f}_{k+1+j/k+1}\right\}_{j=1}^{n}$ of states of the vectors of orthogonalization $f_{[k+1]} = \{f_{j/k+1}\}_{j=0}^{k+n+1}$ and $\widetilde{f}_{[k+1]} = \{\widetilde{f}_{j/k+1}\}_{j=0}^{k+n+1}$.

Theorem 5. *Suppose that in $E = E^{L+1}$ a system X in coordinates Y is $X_Y = |U^i x_{Y[0]}|_0^N = |I^i \alpha_{[0]}|_0^N = A$. Here I^1 is an operator of a shift «downwards», and the vector has the form $x_{Y[0]} = \alpha_{[0]} = |\alpha^*, 0_N^T|^*$, and α is a $(n+1)$-vector. Then the following facts are valid:*

a. *Equations of counter orthogonalization of the form* (7) *hold for the states h_{k+1}, $\widetilde{h}_{k+1}$ of the vectors of orthogonalization $f_{[k+1]}$, $\widetilde{f}_{[k+1]}$, $k = \overline{0, N}$.*

b. *Equations for the states h_{k+1} and $\widetilde{h}_{k+1}$ of the vectors of orthogonalization have the form*

$$
(17)\qquad
\begin{array}{llll}
\mathbf{a}. & h_{k+1} = h_k - I_n^1 \widetilde{h}_k \widetilde{\theta}_k & \mathbf{b}. & \widetilde{h}_{k+1} = I_n^1 \widetilde{h}_k - h_k \theta_k, \qquad \textit{where} \\
\mathbf{c}. & \theta_k = a_k \mu_k, & \mathbf{d}. & \widetilde{\theta}_k = \widetilde{a}_k \widetilde{\mu}_k, \qquad \textit{moreover} \\
\mathbf{e}. & a_{k+1} = (I - \theta_k \widetilde{\theta}_k)^{-1} a_k, & \mathbf{f}. & \widetilde{a}_{k+1} = (I - \widetilde{\theta}_k \theta_k)^{-1} \widetilde{a}_k, \qquad \textit{and} \\
\mathbf{g}. & \mu_k = \boldsymbol{\alpha}^* \widetilde{h}_k, & \mathbf{h}. & \widetilde{\mu}_k = \widetilde{h}_k^* \boldsymbol{\alpha} = \mu_k^*,
\end{array}
$$

c. *The number n, which is the dimension of the vectors h_{k+1}, $\widetilde{h}_{k+1}$ of the states of the vectors of orthogonalization $f_{[k+1]}$, $\widetilde{f}_{[k+1]}$, is the minimal one for which a closed system of equations* (7) *of bilateral orthogonalization of homogeneous band systems of the vectors $X = A$ exists.*

Proof. From the construction of basis matrices A_k (8) of the homogeneous system $X_k = A_k(\alpha) \in E^{L+1}$ in the basis Y, described in the theorem, it follows that the orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$ of the system $X = A$ in equations from Lemma **1** and Theorem **1** consist of two subvectors, namely significant and zero ones: $f_{k+1} = \left|f_{[k+1]}^*, 0_{N-k-1}^T\right|^*$, $\widetilde{f}_{k+1} = \left|\widetilde{f}_{[k+1]}^*, 0_{N-k-1}^T\right|^*$. In the theorem, the significant subvectors $f_{[k+1]}$ and $\widetilde{f}_{[k+1]}$ of the orthogonalizing vectors f_{k+1} and $\widetilde{f}_{k+1}$ are referred to as vectors of orthogonalization. Two subvectors of the orthogonalizing vectors f_{k+1} correspond to the two subvectors (smoothed and initial ones) of the renewed smoothing realizations $\widehat{\mathbf{y}}_{k+n+1}$ in the left-hand side of equations (14).

In order to obtain equations (17) for the states h_{k+1}, $\widetilde{h}_{k+1}$ of the vectors of orthogonalization $f_{[k+1]}$ and $\widetilde{f}_{[k+1]}$ from equations (7), we extract the corresponding $2n$ equalities from equations (7a,b)

$$
(18)\qquad f_{k+1} = U f_k - \widetilde{f}_k \widetilde{\theta}_k, \qquad \widetilde{f}_{k+1} = \widetilde{f}_k - U f_k \theta_k
$$

for the orthogonalizing vectors $f_{k+1} = \Pi_k x_{k+1}$ and $\widetilde{f}_{k+1} = \Pi_{\overline{1,k+1}} x_0$. Due to substitution of the vectors f, $\widetilde{f}$ by h, $\widetilde{h}$ and the fact that $\mu_k = \mu_k^* = \langle x_0, U f_k\rangle = \left\langle \Pi_{\overline{1,k}} x_0, \Pi_{\overline{1,k}} x_{k+1}\right\rangle$, equations (7g,h) also change. Recall that $a_k = \|f_k\|^{-2}$, $\widetilde{a}_k = \left\|\widetilde{f}_k\right\|^{-2}$.

After performing the necessary substitutions, we obtain the equations mentioned in the theorem from system (7) for $k = \overline{0, N}$. Here the definitions of operators of shift from the last paragraph of Section **3.2** are taken into account. The initial conditions are as follows: $a_0 = \widetilde{a}_0 = 1\big/\|x_0\|^2$, $h_0 = \widetilde{h}_0 = \widetilde{\boldsymbol{\alpha}}$. The vector $\widetilde{\boldsymbol{\alpha}}$ is such that $\alpha^* = \left|\alpha_0^*, \widetilde{\boldsymbol{\alpha}}^*\right|$.

We now verify that system of equations (17) is closed. First, we use equations (7h) and (17h):

(19)

$$\widetilde{\mu}_k = \mu_k^* = \left\langle x_{k+1}, \widetilde{f}_k \right\rangle = \sum_{i=1}^{n} \widetilde{f}_{k+i/k}^{*} \alpha_{i-1} = \widetilde{h}_k^* \boldsymbol{\alpha}, \qquad \text{where} \qquad \widetilde{h}_k = \left\{ \widetilde{f}_{k+j/k} \right\}_{j=1}^{n}.$$

From equation (19), it follows that for a number which is less than number n of current components of vectors of orthogonalization a closed system of counter equations of orthogonalization does not exist. □

We call the equations from Theorem **5** a filtration mode of orthogonalization equations from Theorem **1**. The equations from Theorem **5** for the current states h_k and $\widetilde{h}_k$ of the vectors of orthogonalization f_k and $\widetilde{f}_k$ are the principal ones not only for equations of filtration in the context of solving problems of smoothing (14). In Section **5** we will show that filtration mode (17) of equation (7) is necessary and sufficient for solving identification problems **2**, which are the problems of evaluation of coefficients of equation (10).

4.3. **Filtration equations.**

Theorem 6. *The system of filtration equations for evaluation of states of model* (10) *consists of n extracted equations*

(20)

$$\widehat{s}_{k+1} = \widehat{s}_{[k]} - h_{k+1} a_{k+1} \pi_{k+1}, \qquad \textit{where} \qquad \pi_{k+1} = \langle \widehat{\mathbf{y}}_k, x_{k+1} \rangle \qquad \Longrightarrow$$

$$\pi_{k+1} = \alpha^* \begin{vmatrix} \widehat{y}_{k+1/k} \\ \widehat{s}_{[k]} \end{vmatrix} = \alpha^* \begin{vmatrix} \widehat{s}_k \\ y_{k+1+n} \end{vmatrix} = y_{k+1+n} - \widehat{y}_{k+1+n/k}, \quad \widehat{y}_{k+1+n/k} = -\boldsymbol{\alpha}^* \widehat{s}_k$$

in the system of smoothing equations (14) *and also counter equations* (17) *obtained in Theorem* **5** *for the states h_{k+1}, $\widetilde{h}_{k+1}$ of the vectors of orthogonalization $f_{[k+1]}$ and $\widetilde{f}_{[k+1]}$.*

Proof. For the last formula in equations (20), recall the designation (4).

We consider the current states $\widehat{s}_k = \left\{\widehat{y}_{k+j/k}\right\}_{j=1}^{n}$ of model (10) for $k = \overline{-1, N}$. Using them, by the recurrent formula which equation (10) implies, we calculate the next samplings $\widehat{y}_{(k+n+1)/k}$ of the smoothing realization. Further, we must apply renewal equations (20),(21) of this theorem and its Corollary **5**.

We introduce partially smoothed n-vectors $\widehat{s}_{[k]}$. These are n-samples, shifted by one position «downwards» with respect to the states $\widehat{s}_k$ in the vector $\widehat{\mathbf{y}}_{k+n}$ from (14). Therefore, as it follows from the representation of structure of the realization $\widehat{\mathbf{y}}_{k+n}$ in (12), the last one, that is the n-th sampling of vector $\widehat{s}_{[k]}$, equals the initial sampling y_{k+n+1}. With $k = -1$, the states $\widehat{s}_{-1}$ and $\widehat{s}_{[-1]}$ are also the initial samplings: $\widehat{s}_{-1} = s_{-1} = \{y_j\}_0^{n-1}$, $\widehat{s}_{[-1]} = s_{[-1]} = \{y_j\}_1^{n}$. The vector s_{-1} is the initial condition of the process of sequential (recurrent) smoothing of the realization $\mathbf{y}$.

In order to obtain the equations for the next smoothed states

$$\widehat{s}_{k+1} = \left\{\widehat{y}_{(k+j+1)/(k+1)}\right\}_{j=1}^{n}$$

of model (10) from equations (14), we perform «cutting» of the last n from $k+n+1$ of the smoothing equations from system (14). Here $k = \overline{-1, N-1}$. We mean the last n components in the renewed smoothed realizations $\widehat{\mathbf{y}}_{[k+n+1]}$, introduced in Section **2.3.** The equations being «cut» describe the current states $\widehat{s}_{k+1}$ in the filtration problem. We obtain the system of equations shown in the theorem.

Renewal equations (20) along with the equations mentioned in Theorem **5** are the equations of variational fast filter. □

It is well-known that the matrix $\Phi = I_n^1 - e_n\boldsymbol{\alpha}^*$ (4) is called a canonical Frobenius matrix [6]. Here $I_n^1 = \{\delta_{i+1,j}\}_1^n$, and $e_n = \{\delta_{i,n}\}_1^n = |0_{n-1}^T, 1|^T$.

Corollary 5. *Variational equations of filtration* (20) *in problems of smoothing* **1** *are a dynamical system with a canonical transitional Frobenius matrix:*

$$\begin{aligned} \widehat{s}_{k+1} = \Phi\widehat{s}_k + K_{k+1}\pi_{k+1}, \quad \textit{where} \quad \pi_{k+1} = \langle\widehat{\mathbf{y}}_k, x_{k+1}\rangle \quad \longrightarrow \\ \pi_{k+1} = y_{k+n+1} - e_n^T\Phi\widehat{s}_k, \quad \textit{and} \quad K_{k+1} = e_n - h_{k+1}a_{k+1}. \end{aligned} \tag{21}$$

Proof. Here $\Phi\widehat{s}_k = \widehat{s}_{k+1/k}$ are predictions of the states $\widehat{s}_{k+1}$, and Φ is a Frobenius matrix.

A system, equivalent to the system of equations (20), can be written in the form $\widehat{s}_{k+1} = \Phi\widehat{s}_k$, $\widehat{y}_{k+1+n} = e_n^T\widehat{s}_{k+1}$. Equations (21) follow from (20) and the following equalities: $\widehat{s}_{[k]} = I^1\widehat{s}_k + e_n y_{k+n+1} \pm e_n\boldsymbol{\alpha}^*\widehat{s}_k = \Phi\widehat{s}_k + e_n\pi_{k+1}$. □

5. Optimization of subspaces

5.1. Identification functional.

Theorem 7. *The solutions of identification problems (problems* **2**, *mentioned in Introduction) are the vectors $\widehat{\alpha}_{(k)}$, which are the arguments of the minima $\widehat{J}_k(\widehat{\alpha}_{(k)})$ of the identification functionals $\widehat{J}_k(\alpha)$. The latter ones are the values of the identification functionals J_k* (12) *on the projections $\widehat{\mathbf{y}}_{k+n}(\alpha)$* (12).

Proof. A generalized projection of some point $\mathbf{y} \in E$ on an arbitrary closed set $\Omega \subseteq E$ is a point $\widehat{\mathbf{y}} \in \Omega$, nearest to the initial $\mathbf{y}$. In this case, Ω is a set of subspaces $\mathcal{D}_k(\alpha) = S\big(\Psi_k(\alpha)\big)$, determined by n independent parameters of vectors $\alpha \in \omega$, where ω is a set of admissible coefficients of equations (10).

Independent parameters of equation (10) can be choosen, for example, as vectors $\boldsymbol{\alpha}$ from the representation of the vector of the coefficients of equation (10) in the form: $\alpha = |\boldsymbol{\alpha}^*, 1|^*$ (4). The identification functionals determined in the theorem are the squared distances from the realization $\mathbf{y}$ to the subspaces $\mathcal{D}_k(\alpha)$.

A minimization of the identification functionals with respect to the vector $\boldsymbol{\alpha}$ is a search for this minimim on a hyperplane $\alpha_n = 1$. A minimization with respect to the vector α given $\|\alpha\| = const$ is a minimization with respect to a direction, for example, a search for a point on a sphere. The latter case is more preferred, as the search is performed in a limited region.

From Theorem **3**, we obtain

$$\begin{aligned} \widehat{\alpha}_{(k)} = \arg\min_{\alpha\in\omega} \widehat{J}_k(\alpha) = \arg\min_{\alpha\in\omega} J\Big(P\big(\Psi_k(\alpha)\big)\mathbf{y}\Big), \qquad \text{where} \\ \widehat{J}_k(\alpha) = \|\Delta\widehat{\mathbf{y}}_k(\alpha)\|^2 = \|(I-P_k)\mathbf{y}\|^2 = \langle\mathbf{y}, (I-P_k)\mathbf{y}\rangle, \quad P_k = P(\Psi_k). \end{aligned} \tag{22}$$

□

Corollary 6. *The renewal equation for identification functionals is the following quadratic equation with respect to errors in prediction:* $\widehat{J}_{k+1} = \widehat{J}_k + \pi^*_{k+1} a_{k+1} \pi_{k+1}$, *where* π_{k+1} *are errors in prediction from the equations of smoothing and filtration* (14), (20), (21).

Proof. The equation for identification functional $\widehat{J}_{k+1}$ follows from equations of smoothing (14) from Theorem **4**. From formulas (22), we have that

$$\widehat{J}_{k+1} = \langle \mathbf{y}, P(A_{k+1})\mathbf{y} \rangle$$

. We use equations of smoothing (14) and get that

$$\widehat{J}_{k+1} = \|\mathbf{y} - \widehat{\mathbf{y}}_{k+1}\|^2 = \left\| \left(P(A_k)\mathbf{y} + f_{k+1} a_{k+1} \langle \mathbf{y}, f_{k+1} \rangle\right) \right\|^2.$$

The summands of the perpendicular $\mathbf{y} - \widehat{\mathbf{y}}_{k+1}(\alpha)$ are pairwise orthogonal: $P(A_k)\mathbf{y} \perp f_{k+1} = \Pi(A_k)x_{k+1}$. We take into account the expression for π in (14),(20),(21) and obtain the required equation. □

Theorem 8. *Let* $\overline{\alpha}$ *be a vector of unknown* (*and being subject to approximational optimization* (12)) *complex conjugate coefficients of system of equations* (10). *For every* $k = \overline{0, N}$ *there exists a matrix* V_k *of samples of samplings, such that the identification functional* $\widehat{J}_k$ *of equation* (10) *has the form*

$$\widehat{J}_k = \widehat{J}_k(\alpha) = \overline{\alpha}^* V_k^* C_k^{-1}(\alpha) V_k \overline{\alpha}$$

, *moreover, if* $c \neq 0$, *then* $\widehat{J}(c\alpha) = \widehat{J}(\alpha)$. *For scalar equations with unknown coefficients the matrices* V_k *are Henkel ones, and* $\overline{\alpha} = \alpha^{*T} \in E^{n+1}$.

Proof. Only scalar equations (10) will be considered. We determine the functions in the following way: $m_k(\mathbf{y}, \alpha) = \mathbb{D}_k(\alpha)\mathbf{y}$. They are called a residual of equation (10). If errors in $\mathbf{y}$ are small, then sometimes $m \approx 0$. This fact is used for the mentioned algebraic identification of equation (10). From Theorem **2** and Corollary **3**, it follows that $m_k = A^*_k(\alpha)\mathbf{y}$. We will show that in the case mentioned in the theorem there exist matrices V_k such that $m_k = V_k(\mathbf{y})\alpha^{*T}$.

From (10), we can see that those are Henkel $\big((k+1) \times (n+1)\big)$-matrices $V_k = \{v^*_l\}_0^k$, generated by $(n+1)$-samples $v^*_l = |y_{l+j}|_{j=0}^n$ from the realization $\mathbf{y}$. Note also $\big((k+1) \times n\big)$-matrices $\overline{V} = \overline{V}_k$ of n-samples. Their rows are the states of model (10). The matrices $\overline{V}$ are the matrices V without the last column $\overline{\mathbf{y}} = \mathbf{y}_{\overline{n,L}} = \{y_j\}_n^L$. It corresponds to the identity component of the vector of coefficients (4) in the formula for the residual $m = V\alpha^{*T}$. The matrices $\overline{V}$ from the mentioned approximated equality $m \approx 0$ form a system of equations for algebraic identification of the vector of unknown coefficients of equation (10) $\boldsymbol{\alpha}$: $\overline{V}\boldsymbol{\alpha}^{*T} \approx -\overline{\mathbf{y}}$. Hence, for $k = N$ we have

$$V = \begin{vmatrix} y_0, & y_1, & \dots, & y_{n-1}, & y_n \\ y_1, & y_2, & \dots, & y_n, & y_{n+1} \\ y_2, & y_3, & \dots, & y_{n+1}, & y_{n+2} \\ y_3, & y_4, & \dots, & y_{n+2}, & y_{n+3} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ y_N, & y_{N+1}, & \dots, & y_{L-1}, & y_L \end{vmatrix}, \tag{23}$$

$$\overline{V} = \begin{vmatrix} y_0, & y_1, & \dots, & y_{n-1} \\ y_1, & y_2, & \dots, & y_n \\ y_2, & y_3, & \dots, & y_{n+1} \\ y_3, & y_4, & \dots, & y_{n+2} \\ \vdots & \vdots & \dots & \vdots \\ y_N, & y_{N+1}, & \dots, & y_{L-1} \end{vmatrix}.$$

For $\overline{\alpha} = \alpha^{*T} \in E^{n+1}$, we come to the required identity for residuals m_k.

$$m_k = A_k^* \mathbf{y} = V_k \overline{\alpha} = \overline{\mathbf{y}} + \overline{V}_k \boldsymbol{\alpha}^{*T} \qquad \longrightarrow \qquad \widehat{J}_k = m_k^* \langle A_k, A_k \rangle^{-1} m_k. \tag{24}$$

Here all the designations used in formulas (4),(9),(15),(23) are taken into account. □

In this simple form (24), the identity $A^*\mathbf{y} = V\alpha^{*T}$ for the residuals m of equation (10) with the matrix V (23) and the vector $\alpha^{*T} \in E^{n+1}$ of coefficients of this equation only holds for scalar equations (6),(12). For systems of equations and non-commuting matrix coefficients α_j, $j = \overline{0,n}$, the matrices V in (23) and the vectors $\overline{\alpha}$, $\boldsymbol{\alpha}^{*T}$ in (24) have more complex constructions.

Corollary 7. *The functionals of LSM J_k from* (12) *on the projections $\widehat{\mathbf{y}}_k$ can be represented in the form of positively determined pseudo-quadratic forms $\widehat{J}_k$ of the vector $\overline{\alpha}$ of unknown coefficients of the difference equation in* (12). *The forms $\widehat{J}_k$ do not depend on the length of the vector α of coefficients of the equation, and their matrices have non-linear inverse quadratic dependencies on the vector α. The corresponding formulas have the form*

$$\begin{gathered} \widehat{J}_k(\alpha) = \|\Delta \widehat{\mathbf{y}}_k(\alpha)\|^2 = \overline{\alpha}^* Q_k(\alpha) \overline{\alpha}, \qquad \textit{where} \\ Q_k(\alpha) = Q_k = V_k^* C_k^{-1} V_k, \quad \textit{and} \quad C_k = C_k(\alpha) = \langle A_k, A_k \rangle. \end{gathered} \tag{25}$$

Proof. Equations (25) for perpendicular distances follow from formulas (22), (24). The mentioned formulas also imply the other statements of the corollary. □

5.2. **Matrix of realizations.**

Lemma 7. *Suppose that $(L+1)\times(n+1)$-matrix $|\mathbf{w}_j|_0^n = W = W_L$ is a set of realizations $\mathbf{w}_j$, such that $W \in S(A)$, and their residuals are $A^*W = V$. Then $W = AC^{-1}V$.*

Proof. The statement follows from that fact that the solution of the system $A^*W = \langle W, A\rangle = V$, determined as $W \in S(A) \to W = A\mathbf{x}$, is $W = AC^{-1}V = A^{-*}V$. □

We refer to the matrices $Q_k = \langle W_k, W_k \rangle$ as identifying.

Corollary 8. *The identifying matrices have a minimal trace among the Gramian matrices of the realizations $\widetilde{W}_k$, such that $\langle \widetilde{W}_k, A_k \rangle = V_k$.*

Proof. Let Λ be a $(N+1)\times(n+1)$-matrix of Lagrange multipliers. We search for the minimum of the value $\mathrm{Sp}\left\langle \widetilde{W}, \widetilde{W} \right\rangle + \mathrm{Sp}\left\langle \left(\left\langle \widetilde{W}, A \right\rangle - V \right), \Lambda \right\rangle$ by $\widetilde{W}$.

Differentiating it by $\widetilde{W}$, we obtain $\widetilde{W} + A\Lambda = 0 \to V + \langle A, A\rangle \Lambda = 0 \to \Lambda = \langle A, A\rangle^{-1} V \to \widehat{W} = A\langle A, A\rangle^{-1} V = AC^{-1}V = A^{-*}V = W.$ □

Corollary 9. *The identification functionals are the functions of* α*, such that* $\widehat{J}_k = \langle \mathbf{y}, P(A_k)\mathbf{y}\rangle = \alpha^T \langle W_k, W_k\rangle \alpha^{*T}$.

Theorem 9. *The minimal solutions* W_{k+1} *of algebraic equations* $\langle W_{k+1}, A_{k+1}\rangle = V_{k+1}$ *as functions of* $k = \overline{-1, N-1}$ *are described by renewal equations with orthogonalizing vectors from Lemma* **1** *as coefficients of strengthening of renewals:*
(26)
$$W_{k+1} = W_k + f_{k+1}a_{k+1}q^*_{k+1}, \quad \textit{where} \quad f_{k+1} = \Pi_k x_{k+1}, \quad a_{k+1} = \|f_{k+1}\|^{-2},$$
$$q^*_{k+1} = \left(v^*_{k+1} - \widehat{v}^*_{k+1/k} \right), \qquad \textit{and} \qquad \widehat{v}^*_{k+1/k} = \langle W_k, x_{k+1}\rangle; \qquad W_{-1} = 0.$$

Proof. Let $C^{-1}_{k/0}$ be an inverse matrix C^{-1}_k, bordered by zero column and row on the right and from below. We now perform simple transforms.

$$W_{k+1} = A_{k+1}C^{-1}_{k+1}V_{k+1} = A_{k+1}\left(C^{-1}_{k+1} - C^{-1}_{k/0} + C^{-1}_{k/0} \right) V_{k+1} = \tag{27}$$
$$W_k + A_{k+1}\left(C^{-1}_{k+1} - C^{-1}_{k/0} \right) \langle W_{k+1}, A_{k+1}\rangle.$$

We take into account that $A_{k+1}\left(C^{-1}_{k+1} - C^{-1}_{k/0} \right) \langle \cdot, A_{k+1}\rangle = P(A_{k+1}) - P(A_k) = f_{k+1}a_{k+1}\langle \cdot, f_{k+1}\rangle$ (Lemma **1**) and $\langle f_{k+1}, f_{k+1}\rangle = \langle f_{k+1}, x_{k+1}\rangle = x^*_{k+1}f_{k+1}$. Then,

$$f_{k+1}a_{k+1}\langle \cdot, f_{k+1}\rangle = f_{k+1}a_{k+1}\langle f_{k+1}, f_{k+1}\rangle a_{k+1}\langle \cdot, f_{k+1}\rangle =$$
$$f_{k+1}a_{k+1}x^*_{k+1}f_{k+1}a_{k+1}\langle \cdot, f_{k+1}\rangle =$$
$$f_{k+1}a_{k+1}x^*_{k+1}A_{k+1}\left(C^{-1}_{k+1} - C^{-1}_{k/0} \right)\langle \cdot, A_{k+1}\rangle =$$
$$f_{k+1}a_{k+1}\langle A_{k+1}, x_{k+1}\rangle \left(C^{-1}_{k+1} - C^{-1}_{k/0} \right)\langle \cdot, A_{k+1}\rangle.$$

We return to equations (23):

$$\langle A_{k+1}, x_{k+1}\rangle C^{-1}_{k+1}V_{k+1} = \left| 0^T_{k+1}, 1 \right| V_{k+1} = v^*_{k+1}, \tag{28}$$
$$\langle A_{k+1}, x_{k+1}\rangle C^{-1}_{k/0}V_{k+1} = \langle W_k, x_{k+1}\rangle = \widehat{v}^*_{k+1/k}.$$

Hence,

$$\langle A_{k+1}, x_{k+1}\rangle \left(C^{-1}_{k+1} - C^{-1}_{k/0} \right) V_{k+1} = \left(v^*_{k+1} - \widehat{v}^*_{k+1/k} \right) = q^*_{k+1}.$$

Therefore, the row q^*_{k+1} is a renewal. It is an error of the prediction $\widehat{v}^*_{k+1/k}$ of the system $\langle W_k, A_k\rangle = V_k$ on the $(k+1)$-th row v^*_{k+1} of the matrix of samples $V_{k+1} = \langle W_{k+1}, A_{k+1}\rangle$. □

5.3. **Equation for identifying matrix.** We denote the current $(n \times (n+1))$-blocks of the matrices of realizations W_k as w_k. These blocks are formed by rows with numbers $l = \overline{k+1, k+n}$ of the matrices W_k. They are the analogues of the current states h_k of the orthogonalizing vectors f_k, preceding the zero part of the latter ones, and the current states $\widehat{s}_k$ in the partially smoothed realizations $\widehat{\mathbf{y}}_{k+n}$, preceding the non-smoothed part of this realizations with initial samplings. Similar to the states h_k of the orthogonalizing vectors f_k, the blocks w_k precede the zero «lower» part of matrices W_k. The latter one correspond to the zero parts of the vectors f_k, described above (Theorems **5** and **9**).

Corollary 10. *The errors q^*_{k+1} of predictions $\widehat{v}^*_{k+1/k}$ in renewal equations* (26) *for the matrices of realizations W_{k+1} are determined by the current $(n \times (n+1))$-blocks w_k of the matrices W_k, if $X = A$ is a band matrix* (8) *with band width of $n+1$.*

Proof. From (26) and (28), we obtain the analogues of predictions in (14) and (20):

$$\widehat{v}^*_{k+1/k} = \langle W_k, x_{k+1} \rangle \;\longrightarrow\; \widehat{v}^*_{k+1/k} = \boldsymbol{\alpha}^* w_k. \tag{29}$$

Therefore, with the conditions mentioned in the statement of Corollary **10**, the prediction $\widehat{v}^*_{k+1/k}$ of the rows $\widehat{v}^*_{k+1/k+1} = \widehat{v}^*_{k+1}$ is determined by the current $(n \times (n+1))$-blocks w_k of the matrices W_k, multiplied by the row vector $\boldsymbol{\alpha}^*$. □

Lemma 8. *Let $X = A$ be a band Toeplitz matrix with a generating vector $\alpha \in E^{n+1}$* (7). *Then the system of renewal equations for the current blocks w_{k+1} of the matrices of realizations W_{k+1} for $k = \overline{-1, N-1}$ is determined by equations* (26) *from Theorem* **9** and (29) from Corollary **10**:

$$w_{k+1} = w_{[k]} + h_{k+1} a_{k+1} q^*_{k+1}, \qquad \text{where} \qquad q^*_{k+1} = v^*_{k+1} - \widehat{v}^*_{k+1/k},$$
$$\text{and} \qquad \widehat{v}^*_{k+1/k} = \langle W_k, x_{k+1} \rangle \qquad \longrightarrow \qquad \widehat{v}^*_{k+1/k} = \boldsymbol{\alpha}^* w_k,$$

and also by the system of counter equations (17) *from Theorem* **5**:

a. $h_{k+1} = h_k - I^1_n \widetilde{h}_k \widetilde{\theta}_k$	**b.** $\widetilde{h}_{k+1} = I^1_n \widetilde{h}_k - h_k \theta_k,$
c. $\theta_k = a_k \mu_k,$	**d.** $\widetilde{\theta}_k = \widetilde{a}_k \widetilde{\mu}_k,$
e. $a_{k+1} = (I - \theta_k \widetilde{\theta}_k)^{-1} a_k,$	**f.** $\widetilde{a}_{k+1} = (I - \widetilde{\theta}_k \theta_k)^{-1} \widetilde{a}_k,$
g. $\mu_k = \boldsymbol{\alpha}^* \widetilde{h}_k,$	**h.** $\widetilde{\mu}_k = \widetilde{h}^*_k \boldsymbol{\alpha} = \mu^*_k,$

Proof. We add the n extracted equalities from equations (26) of Theorem **9** to equations (29). This method of «cutting» was used when obtaining the filtration equations (20) from smoothing equations (14) from Theorem **4**.

For $k = \overline{-1, N-1}$ and given $w_{[-1]} = 0$, from (26), we get that

$$w_{k+1} = w_{[k]} + h_{k+1} a_{k+1} q^*_{k+1}, \quad \text{where} \quad q^*_{k+1} = v^*_{k+1} - \boldsymbol{\alpha}^* w_k. \tag{30}$$

The matrix $w_{[k]}$ is similar to the vector $\widehat{s}_{[k]}$ in filtration equation (20). It designates the same n rows in the W_k under renewal as the matrix w_{k+1} in the renewed matrix W_{k+1} in the first equation in (26). The last component of the vector $\widehat{s}_{[k]}$ in equations (20) equals the initial sampling. In $w_{[k]}$, it is the zero row, since $W_{-1} = 0$ in (22). We must add the equations of Theorem **5** to equations (26). □

Theorem 10. *The renewal equation of the identifying matrix $Q(\alpha)$ is quadratic, and with respect to errors in prediction, the equation*

$$Q_{k+1} = Q_k + q_{k+1} a_{k+1} q^*_{k+1}, \qquad \textit{where} \qquad q_{k+1} = v_{k+1} - w^*_k \boldsymbol{\alpha} \tag{31}$$

is the errors in prediction from Lemma **8** (30)*,* $k = \overline{k_{\min}, N-1}$*,* $Q_{-1} = 0$.

Proof. From Theorem **9**, we have (26):

$$Q_{k+1} = \langle W_{k+1}, W_{k+1} \rangle = \left\langle W_k + f_{k+1} a_{k+1} q^*_{k+1}, W_k + f_{k+1} a_{k+1} q^*_{k+1} \right\rangle .$$

Two summands in the renewal formula for W_{k+1} are orthogonal to each other. The required result is obtained, since $a_{k+1} = \|f_{k+1}\|^{-2}$. □

Corollary 11. *The Gramian matrix* $Q(\alpha) = \langle W, W \rangle$ *can be calculated using counter equations of orthogonalization in filtration mode. We add to equations* (30) *and* (31) *the equations and conditions of Theorem* **5** *or Lemma* **8***.*

6. Identification equations

6.1. **Prototype: a problem of orthogonal regression.** The process of characterizing the functional of variational identification (VI) $\widehat{J} = \widehat{J}_{vi}$ (25) as a pseudo-quadratic form, independent of the length of the vector determining it, results in obtaining the analogy between the problem of its minimization and the problem of search of the minimum of the positively determined quadratic form on a sphere [21,22]. The problems of «total» LSM [24] or orthogonal regression (OR) [23] lead us to the mentioned problems. We will use the term orthogonal regression to designate the class of problems related to the LSM. We now provide the justification for such preference.

Both linear LSM and OR suppose in this class of problems constructing a set of a given type (often smooth), approximating a set of points in some space. The set is constructed in a way as to minimize a particular distance from the points to the set. In linear LSM, the distance is calculated by sections from the points to the set along one of the coordinates. In OR problems, we calculate the distance by the shortest (orthogonal for smooth sets) sections connecting the points to the set.

A special case of OR problems, closest to the VI problem under consideration, can be formulated with the help of matrix constructions (23) and equalities (24).

Problem 3. *For every* k*, minimize with respect to* $\widehat{V}_k$ *and* α *the functional*

$$J_{or,k} = \|V - \widehat{V}_k\|^2 = \mathrm{Sp}\big((V - \widehat{V}_k)^*(V - \widehat{V}_k)\big), \quad \textit{if} \quad m_k(\widehat{V}_k, \alpha) = \widehat{V}_k \overline{\alpha} = 0. \tag{32a}$$

or with respect to $\overline{\alpha}$ *the functional*

$$\widehat{J}_{or,k} = \overline{\alpha}^* V_k^* V_k \overline{\alpha} / \overline{\alpha}^* \overline{\alpha}, \qquad \textit{where} \qquad \overline{\alpha} = \alpha^{*T}. \tag{32b}$$

We will show the identification functionals $\widehat{J}_{or}$ and $\widehat{J}_{vi}$ in a comparable form. Their common property is that their values do not depend on the length of the generating vector, that is, they are constant on the rays α:

(33)

$$\text{a)} \quad \widehat{J}_{or} = \overline{\alpha}^* Q_{or} \overline{\alpha}, \qquad \text{where} \qquad Q_{or}(\overline{\alpha}) = V^* V \big/ (\overline{\alpha}^* \overline{\alpha})^{-1} = \langle V, V \rangle_{E^{N+1}},$$

$$\text{b)} \quad \widehat{J}_{vi} = \overline{\alpha}^* Q_{vi} \overline{\alpha}, \qquad \text{where} \qquad Q_{vi}(\overline{\alpha}) = V^* \langle A, A \rangle^{-1} V = \langle W, W \rangle_{E^{L+1}} .$$

Then it follows that to study the properties of these functionals it is enough to limit their values on any sphere. For example, on a unit sphere.

It is well-known [23] that the minimum of a quadratic form on a sphere, for example, on the unit sphere S_{sph}, is reached at the eigenvector (we denote it by β_0) of the matrix of this form, corresponding to its minimal characteristic value (λ_0). That is why the matrices Q_{vi} and Q_{or} are referred to as identification in equalities

(33). The minimum is global if the characteristic value λ_0 is isolated. That is why in the OR problems of iteration

$$(34)\qquad \widetilde{\alpha}_{[j+1]} = Q^{-1}\left(\widehat{\alpha}_{[j]}\right)\widehat{\alpha}_{[j]}, \qquad \widehat{\alpha}_{[j]} = \widetilde{\alpha}_{[j]}/\|\widetilde{\alpha}_{[j]}\| \in S_{sph} \subset E^{n+1}, \\ \text{if} \qquad \widehat{\alpha}_{[0]} \not\perp \beta_0, \qquad \text{and} \qquad \lambda_1/\lambda_0 = \xi > 1$$

converge to β_0. The rate of convergence increases with larger values of ξ, that is, given larger isolation of the characteristic value λ_0 [10,21].

Normalization of the result of each step of iterations (34) is preferred in order to avoid uncontrolled growth of the lengths of the vectors $\widetilde{\alpha}$ given a small value of λ_0. Normalization to a sphere is preferred over reduction of $\overline{\alpha}$, for example, to a hyperplane $\widehat{\alpha}_n = const$, as the choice of the former helps to avoid such complications as when during the process of iterations (34) it turns out that $\widehat{\alpha}_n \approx 0$.

The analogies between OR and VI problems (the basis of these analogies is identity (24) $A^*\mathbf{y} = V\overline{\alpha}$) and common properties (33) of the identification functionals in these problems suggest us that there should be analogous ways of search for the extremum of these functionals.

Iterations (34) of search for the extremum in VI problems should be more effective (in rate and region of convergence) compared to OR problems. Experiments prove this assumption. It is based, in particular, on the fact that problems of the form (32) have a significantly larger amount of optimizable variables compared to VI problems (12) [21].

In problems (32), apart from the vector α, $(k+1)n$ of the initial conditions $\widehat{s}_j$, $j = \overline{0,k}$, are evaluated for $k+1$ of rows of the matrices $\widehat{V}_k$. These evaluations are required for predicting the $(n+1)$-th element of every row with respect to equation (10). In problems (12), with all k, one initial condition (except for α) is optimized. That is the n-vector $\widehat{s}_{0/k}$ of the realization $\widehat{\mathbf{y}}_{[k]}$. Hence, there is a total of $2n$ parameters.

6.2. **Equation of inversion of sums of matrices.** The renewal equation for the identifying matrix $Q_{or,j}(\alpha)$ follows from the formula (33a). It has the form

$$(35)\qquad Q_{or,k+1}(\alpha) = Q_{or,k}(\alpha) + \Delta Q_{or,k+1} = Q_{or,k}(\alpha) + v_{k+1}a_0v^*_{k+1}/\|\alpha\|^2.$$

Here $k = \overline{-1, N-1}$, $a_0 = 1/\|\alpha\|^2$, v^*_{k+1} is the $(k+1)$th row of the matrix V from (14), and $Q_{or,-1} = 0$. Theorem **10** in equality (22) determines the similar form of the renewal $\Delta Q_{vi,k+1} = q_{k+1}a_{k+1}q^*_{k+1}$ of the identifying matrix Q_{vi} in VI problems.

The process of sequential inversion of sums of matrices of the forms (31), (35) is described by a well-known recurrent matrix equation [3]:

$$(36)\qquad (Q + uav^*)^{-1} = Q^{-1} - Q^{-1}u\left(a^{-1} + v^*Q^{-1}u\right)^{-1}v^*Q^{-1}.$$

A question arises regarding feasibility of using equation (36) for sequential inversion of the matrices Q taking into account their additive components of the form (35). The analysis of equation (36) results in the following answer. If the search with required accuracy of the eigenvector β_0 of the matrix Q needs several iterations (34), then the use of equation (36) for sequential inversion of the matrices Q taking into account their additive components is not feasible in terms of the computing expenses.

The sequential inversion of identifying matrix with the help of equations (36) is desirable when one iteration is enough to obtain the solution β_0 with required accuracy (34). As experiments show, that is not uncommon in many practically important cases. In particular, that is valid in problems of tracking of changes in current evaluations when the length of intervals of observation is changed.

Theoretical research on functionals of variational identification and iterations (34) in VI problems is problematic. The reasons can be seen in formulas (33b). In paper [35], formulas for the vector of the first and the matrix of the second derivatives of the functional $\widehat{J}_{vi}(\alpha)$ (33b) with respect to the vector α are obtained. The obtained formulas can be seen also in work [33]. These articles show the extent to which they are difficult to analyze.

6.3. **Equations of variational identification.**

Lemma 9. *Suppose that* $\widehat{\alpha}_{[0]} \in S_{sph} \subset E^{n+1}$. *The equations for the current* $\widehat{\alpha}_{(k+1)[1]} \in S_{sph}$, $k = \overline{0, N-1}$, *evaluations of coefficients of equation* (10) *on the basis of one iteration* (34) *are as follows:*

$$\widetilde{\alpha}_{(k+1)[1]} = Q_{k+1}^{-1}\left(\widehat{\alpha}_{[0]}\right)\widehat{\alpha}_{[0]} = Q_k^{-1}\left(\widehat{\alpha}_{[0]}\right)\widehat{\alpha}_{[0]}, \qquad \widehat{\alpha}_{(k+1)[1]} = \widetilde{\alpha}_{(k+1)[1]} / \|\widetilde{\alpha}_{k+1}\|$$

are the renewal equations. Here the matrices Q_k *are* $Q_{vi,k}$ *from* (31) *and* $Q_{or,k}$ *from* (35), $k = \overline{k_{min}, N-1}$. *The matrices* $Q_{k_{\min}}$ *are invertible.*

Proof. The system of equations of one step of iteration (34) of VI consists of Riccati equations of the form (36) and renewal equations for the matrices Q. These are equations (29–31) for the matrices $Q_{vi,k+1}\left(\widehat{\alpha}_{[0]}\right)$ and equation (35) for the matrices $Q_{or,k+1}$.

Suppose that $k = \overline{0, N-1}$. We use the normalization to identity length denoted as $\widehat{\alpha} \in S_{sph}$, for the current $\widehat{\alpha}_{(k+1)[1]} = \widetilde{\alpha}_{(k+1)[1]} / \|\widetilde{\alpha}_{(k+1)[1]}\|$ and preceding $\widehat{\alpha}_{(k)[1]} = \widetilde{\alpha}_{(k)[1]} / \|\widetilde{\alpha}_{(k)[1]}\|$ vectors of equation (30). We obtain the equation for the normalized vectors $\widehat{\alpha}_{(k+1)[1]} \in S_{sph}$:

$$\widehat{\alpha}_{(k+1)[1]} = \left(\widehat{\alpha}_{(k)[1]} - G_{k+1}\left(a_{k+1}^{-1} + q_{k+1}^{*} G_{k+1}\right)^{-1} \Delta_{k+1}\right) \chi_{k+1}. \tag{37}$$

Here

$$\Delta_{k+1} = q_{k+1}^{*}\widehat{\alpha}_{(k)[1]}, \quad \chi_{k+1} = \|\widetilde{\alpha}_{(k)[1]}\| / \|\widetilde{\alpha}_{(k+1)[1]}\|, \quad G_{k+1} = Q_k^{-1} q_{k+1},$$

$$\text{where} \qquad \widetilde{\alpha}_{(k)[1]} = Q_k^{-1}\alpha_{[0]}, \qquad \widetilde{\alpha}_{(k+1)[1]} = Q_{k+1}^{-1}\alpha_{[0]}.$$

From equation (37), it follows that $\Delta_{k+1} = \widehat{\alpha}_{n(k)[1]} q_{k+1}^{*} \overline{\alpha}_{(k)[1]}$. Recall that $\overline{\alpha} = \alpha^{*T} = \left|\boldsymbol{\alpha}^{*}, 1\right|^{T}$. Here $\widehat{\alpha}_{n(k)[1]}$ denotes the n-th component of the vector $\widehat{\alpha}_{(k)[1]} \in S_{sph}$. The values Δ_{k+1} are the renewals determined by errors in prediction of the «form π» (14),(20). Indeed, from (26) we know that $q_{k+1}^{*} = \left(v_{k+1} - \widehat{v}_{k+1/k}\right)^{*}$. Hence,

$$\Delta_{k+1} = \widehat{\alpha}_{n(k)[1]}^{*}\left(v_{k+1}^{*}\overline{\alpha}_{(k)[1]} - \widehat{v}_{k+1/k}^{*}\overline{\alpha}_{(k)[1]}\right) = \widehat{\alpha}_{n(k)[1]}^{*}\left(\pi_{or,k+1} - \widehat{\pi}_{k+1}\right), \quad \text{where}$$

$$\pi_{or,k+1} = y_{k+n+1} - \widehat{y}_{k+n+1/v}, \qquad \widehat{\pi}_{k+1} = \boldsymbol{\alpha}^{*} w_k \overline{\alpha}_{(k)[1]} = \widehat{y}_{k+n+1/k} - \widehat{y}_{k+n+1/k,w}.$$

Unlike (14) and (20), the prediction $\widehat{y}_{k+n+1/v}$ on the sampling y_{k+n+1} is performed not with respect to the current state $\widehat{s}_k$ from the smoothed realization $\widehat{\mathbf{y}}_{k+n}$, but to the row v_k^{*} of the matrix $\overline{V}$ (23) of samplings of the initial realization $\mathbf{y}$. This is a prediction from equations of smoothing of the form (14), but performed by the methods of OR (32). Using the methods of OR (35), unlike the methods of VI (31), we get that $a_{k+1} = a_k = 1$, $q_{k+1} = v_{k+1}$.

For the prediction $\widehat{y}_{k+n+1/w} = \widehat{v}^T_{k+1/k}\boldsymbol{\alpha}^{*T} = \boldsymbol{\alpha}^*\overline{w}_k\boldsymbol{\alpha}^{*T}$, equation (29) is used. The prediction is done with respect to the current block w_k of the matrix of realizations W_k. Without the last column, it is a $(n \times n)$-block. Here it is denoted by $\overline{w}_k$. This prediction $\widehat{y}_{k+n+1/w}$ on the sampling $\widehat{y}_{k+n+1/k}$ is done with respect to the prediction $\widehat{v}^T_{k+1/k}$ of this row in equations (29), (30). $\square$

7. Conclusion

We have obtained the systems of difference equations of solution of variation smoothing problems, the systems of equations of nonlinear filtration of signals, as well as the equations of identification of dynamical processes based on linear stationary models.

It has been shown that the basis of the obtained system of equations are the counter equations of bilateral orthogonalization of homogeneous vector systems.

We have demonstrated that in order to solve problems of evaluation of states of systems and to identify them it is sufficient to use counter equations of orthogonalization in filtration mode along with Riccati equations of inversion of sums of matrices.

References

[1] T. Kailath, *Some new algorithms for recursive estimation in constant linear systems*, IEEE Trans. Inf. Theory, **19**:6 (1973), 750–760. Zbl 0342.93053

[2] A.O. Egorshin, *Least square method and fast algorithms in variational problems of identification and filtration (VI method)*, Avtometriya, **1** (1988), 30–42.

[3] P.Ye. Elyasberg, *Definition of the Motion via Measurement Result*, URSS, Moscow, 2011.

[4] R.C.K. Lee, *Optimal Estimation, Identification and Control*, MIT Press, Cambridge, 1964. Zbl 0137.35703

[5] A.O. Egorshin, *On counter orthogonalization processes*, Numer. Analysis Appl., **5**:4 (2012), 307–319. Zbl 1299.65071

[6] C.F. Gauss, *Theoria Motus Corporum Coelestium*, Hamburg, 1809.

[7] Yu.V. Linnik, *Least Square Method and the Foundations of the Theory of Observation Processing*, Fizmatgiz, Moscow, 1958. 0161.15804

[8] A.O. Egorshin, *Equations of identification methods for linear differential equations*, Mat. Zamet. SVFU, **21**:3 (2014), 28–45. Zbl 1374.41025

[9] A.O. Egorshin, *On one variational smoothing problem*, Vestn. Udmurt. Univ. Mat. Mekh. Komp'yut. Nauki, **2011**:4 (2011), 9–22. Zbl 1299.65072

[10] A.O. Egorshin, *On one variational problem of the piecewise-linear dynamical approximation*, Vestn. Udmurt. Univ. Mat. Mekh. Komp'yut. Nauki, **2012**:4 (2012), 30–45. Zbl 1299.65073

[11] V.A. Ambartsumyan, *On the question of diffuse reflection of light by turbid medium*, Dokl. Akad. Nauk SSSR, **38**:8 (1943), 257–261.

[12] S. Chandrasekhar, *On radiate equilibrium of stellar atmosphere. XXI–XXII*, Astrophys. J., **106** and **107** (1947 and 1948), 152–216 and 48–72.

[13] E.A. Robinson, *Spectral approach to geophysical inversion by Lorentz, Fourier, and Radon transforms*, Proc. IEEE, **70**:9 (1982), 1039–1054.

[14] T. Kailath, *Some Chandrasekhar-type algorithms for quadratic regulators*, Proc. IEEE Conference on Detection and Control, New Orleans, LA, Dec. 1972, 219–223.

[15] R. Redheffer, *On the relation of transmission-line theory to scattering and transfer*, J. Math. Phys., **41** (1962), 1–41. Zbl 0104.43605

[16] M.G. Krein, *On a new method for solving linear integral equations of the first and second kinds*, Dokl. Akad. Nauk SSSR, **100**:3 (1955), 413–416. Zbl 0066.08801

[17] N. Levinson, *The Wiener RMS (root-mean-square) error criterion in filter design and prediction*, J. Math. Phys., **25** (1947), 261–278.

[18] N.I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Gos. Izd. Fiz.-Mat. Lit., Moscow, 1961. Zbl 0124.06202

[19] I.C. Gohberg, M.G. Krein, *Theory and Applications of Volterra Operators in Hilbert Space*, Amer. Math. Soc., Providence, 1970. Zbl 0194.43804
[20] A.O. Egorshin, *On a method for estimating coefficients of simulating equations for sequences*, Sib. Zh. Ind. Mat., **3**:2 (2000), 78–96. Zbl 1009.30019
[21] V.P. Budyanov, A.O. Egorshin, *Signal smoothing and estimation of dynamical parameters in automatic systems using a digital computer*, Avtometriya, **1973**:1 (1973), 78–82.
[22] A.O. Egorshin, *Closed computational methods of identification of linear objects*, Optimal and Self-Adjusting Systems, Institute of Automation and Electrometry SB RAS, Novosibirsk, 1971, 40–53.
[23] F.R. Gantmacher, *The Theory of Matrices*, Hirsch Chelsea Publishing Co., New York, 1959. MR0107649
[24] M. Kendall, A. Stuart, *The Advanced Theory of Statistics, Vol. 3: Design and Analysis, and Time-Series*, 3rd ed., Charles Griffin and Company Ltd, London - High Wycombe, 1976. Zbl 0361.62001
[25] Prony Estimation // StatSci.org (http://www.statsci.org/other/prony.html), 2000.
[26] A.S. Householder, *On Prony's method of fitting exponential decay curves and multiple-hit survival curves*, Oak Ridge National Lab. Report ORNL–455, Tennessee, 1950.
[27] M.J. Levin, *Estimation of a system pulse transfer function in the presence of noise*, IEEE Trans. Autom. Control, **9**:3 (1964), 229–235.
[28] V.F. Pisarenko, *On the estimation of spectra by means of non-linear functions of the covariance matrix*, Geophys. J.R. Astron. Soc., **28** (1972), 511–531. Zbl 0249.62090
[29] G.H. Golub, C.F. Van Loan, *An analysis of the total least squares problem*, SIAM J. Numer. Anal., **17** (1980), 883–893. Zbl 0468.65011
[30] B. De Moor, *Structured total least squares and L_2 approximation problems*, Linear Algebra Appl., **188-189** (1993), 163–205. Zbl 0781.65028
[31] M.R. Osborne, *A class of non-linear regression problems*, Data Represent., Proc. Semin. Aust. nat. Univ. 1969 (1970), 94–101. Zbl 0341.62058
[32] M. Aoki, P. Yue, *On a priori error estimates of some identification methods*, IEEE Trans. Autom. Control, **15**:5 (1970), 541–548.
[33] A.O. Egorshin, *Optimization of parameters of stationary models in a unitary space*, Autom. Remote Control, **65**:12 (2004), 1885–1903. Zbl 1095.93006
[34] I.M. Glazman, Yu.I. Lyubich, *Finite-Dimensional Linear Analysis in Tasks*, Nauka, Moscow, 1969. Zbl 0188.07401
[35] A.O. Egorshin, *On tracking extremum parameters in the identification variational problem*, J. Math. Sci., **195**:6 (2013), 791–804.

Alexey Olegovich Egorshin
Sobolev Institute of Mathematics,
4, Koptyuga ave.,
Novosibirsk, 630090, Russia
Email address: egorshin@math.nsc.ru