S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports

http://semr.math.nsc.ru

Том 4, стр. 482–503 (2007) УДК 517.925.54; 517.962.27/.8

MSC 65F25; 15A03,09,23; 93E12

ORTHOGONALIZATION, FACTORIZATION, AND IDENTIFICATION AS TO THE THEORY OF RECURSIVE EQUATIONS IN LINEAR ALGEBRA

A. O. YEGORSHIN

Abstract. We outline theoretical foundations for the recurrent algorithms of computational linear algebra based on counter orthogonalization processes over an ordered system of vectors; we also show the importance of these processes for analysis and applications. We present some important applications of counter orthogonalization processes related to some approximation problems and signal processing as well as recent applications related to the so called homogeneous structures and Toeplitz systems. In particular, these applications contain operators and inversion of matrices, $\mathbb{QDR}$– and $\mathbb{QDL}$–decompositions, $\mathbb{RDL}$– and $\mathbb{LDR}$–factorizations, solutions of integral equations and of systems of algebraic equations, signal estimation on based on approximation models in the form of differential and difference equations and variational identification (coefficients estimation) of the latter.

1. Introduction

1. Structure of the article. In section 2, we show that the recurrent equations of the Gram–Schmidt process of counter orthogonalization over a set of vectors in an inner product (Hilbert) space can be used to obtain the following well–known recurrent algorithms of linear algebra: the inversion, factorization, and decomposition algorithms for matrices, which not require the knowledge on eigenvectors.

In Section 2, we mainly discuss the methodological aspects. By this, we give the proofs for main facts and assertions only.

Yegorshin, A.O., Orthogonalization, factorization, and identification as to the theory of recursive equations in linear algebra .

Supported by Russian Foundation for Basic Research (Grant No. 06-01-00776).

Received September, 11, 2006, published December, 6, 2007.

In sections 3 and 4, we apply the Gram–Schmidt counter orthogonalization scheme to homogeneous sets of vectors in a Hilbert space and deduce the new recurrent algorithms for the counter orthogonalization (Theorem 3). This Theorem for the Hardy space $\mathbf{H}^2$ was earlier published by the author in [1,2] without a proof.

It was noted there that the equations of the Gram–Schmidt counter orthogonalization process give rise to a new class of recurrent algorithms without Riccati equation in various theoretical and practical fields: orthogonalization of polynomials on the unit circle, the (block) Toeplitz matrix inversion, solving integral equations of some types, and some others. These equations also arise in the theory of wave propagation in a homogeneous medium (astro– and geophysics), in stationary systems signal processing, etc.

In section 4, we describe the author's numerical solution to the variational problem of optimization of coefficients of the mathematical model approximating given data.

The consideration is carried out in the Euclidean space of rows of elements of a Hilbert space H. This construction is a direct generalization of matrices and operators acting on a finite–dimensional unitary space.

2. Premises. In the first part of the theory under consideration, the important sequential procedures of computational algebra (inversions and matrix triangular factorizations) are shown to have a general fundamental basis, namely the Gram–Schmidt orthogonalization procedures of some (block in general case) system of independent vectors in an abstract Hilbert space. Here the counter (forward and backward) orthogonalization procedures play an important role. In the paper we present a concise scheme of basic algorithms of computational algebra based on the above–mentioned orthogonalization procedure.

In particular, we consider the inference schemes for the well–known recurrent inversion algorithms for bordered matrices (Frobenius formulas) and summarized matrices (Riccati matrix equations), $\mathbb{QDR}$– and $\mathbb{QDL}$–decompositions, as well as for the sequential $\mathbb{RDL}$– and $\mathbb{LDR}$–factorization formulas. For certainty, they are called the Frobenius factorization (because they result from the Frobenius inversion of the bordered matrix) and the Cholessky factorization (because they originate from the Cholessky algorithms) correspondingly. Here $\mathbb{R}$, $\mathbb{L}$, $\mathbb{D}$ denote, in general, sets of block right–(upper–), left– (lower–) triangular and diagonal matrix respectively.

The most interesting applications (both in theory, e.g., orthogonal polynomials on the circumference calculation, and in practice, e.g., geophysics, signal processing, (block) Toeplitz matrix inversion) are associated with homogeneous systems of vectors. This name was given to systems of vectors which are obtained from a given set of vectors by means of a power of an isometric or of a partially isometric operator.

The specific property of the Gram–Schmidt orthogonalization procedures which explains their fundamental significance is that they are triangular, i.e. they are equivalent to transformations of some systems of vectors. These transformations preserve the (forward or backward) chains of embedded subspaces related to this system. Finite segments of the generating system of vectors form the bases in these subspaces. By this, they are preserved by equivalent right– and left– triangular transformations.

The fundamental importance of the triangular transformations is explained by the fact that they are abstract mathematical models of causal (i.e. physically realized) transformations of a given system under direct or reverse change of an independent variable. This variable can be considered to be time (in dynamic systems) or some other coordinate, e.g., a spacial one (in geo– or astrophysics).

As it was mentioned, homogeneous systems and counter processes in such systems are regarded to be the important applications of the theory under consideration. These systems are abstract mathematical models of homogeneous (stationary, isotropic) systems (media). In systems like these, propagations (transformations) conditions do not depend on the independent variable (coordinate) shift.

The second part of the paper is devoted to the applications of the “general” results of the first part to the homogeneous systems. Here both well–known and relatively new (variational identification) applications of the theory under discussion are briefly described.

3. Remarks about the bibliography. Some elements of recurrent algorithms in computational linear algebra (both for continuous and discrete variables) were studied in papers [3,4] by T. Kailath and his coauthors.

Whereas in our paper abstract mathematical objects and their formal transformations (namely, systems of vectors and their orthogonalization or, more general, biorthogonalization based on the Gram–Schmidt procedure) are considered to be the initial object of the theory, in [3,4] the authors study mathematical models of more “perceptible” objects: layer structures, media, and applied problems of propagating in them. The latter ones are solved by using Redheffer’s composition (see, for example, [5] or [6]). This model is also particulary interesting, especially for homogeneous structures (for instance, for isotropic media and stationary systems).

Equations for the problems of propagating in homogeneous structures have been known for a long time. They were obtained by different authors in various fields. We think the first such papers here were written by the astrophysicists V. Ambartsumian [7] and C. Chandrasekhar [8].

In geophysics, we can point out the E. Robinson’s papers [9,10].

In the field of signal processing, we could mention A. Lindquist’s papers on integral equations and filters [11,12], T. Kailath′s papers on “fast” algorithms of Kalman’s filtration [13,14,15], and the papers on variational approach to the filtration in identification problems [1,2,16,17] by the author of the present paper.

It is worth to mention some “mathematical” applications of the theory of homogeneous structures and counter processes in them. First we mention the polynomial orthogonalization on the unit circle [18,19], on homogeneous system of vectors in a Hilbert space (see the current paper) and, in particular, in a Hardy space [1,2]. Secondly we mention the method of solving the integral equations with a kernel depending on the difference of a variable [20,21].

The well–known N. Levinson’s article [21] can be considered to be historically important. Here the author gave the method of solving integral Wiener–Hopf equations with discrete variable, and actually described the inversion algorithm for Toeplitz matrices. For the case of continuous variable, one should list Krein’s fundamental papers [20,22,23] (see also [24,25]).

The papers by N. Levinson, S. Chandrasekhar, and M. Krein were the starting point for a series of papers on the so called “fast”, “lattice”, and “ladder” recurrent computational algorithms and schemes (see, for example, [11,26,27]). Only some of these papers were listed above, but they can serve as sources for references to other papers in this field.

2. Elements of Computational Algebra

1. Systems of vectors. Consider a system of vectors

$$X = X_N = |x_0, \cdots, x_N| = \{x_i\}_0^N \in H \otimes E,$$

$$E = E_N = \otimes_0^N E_{(i)}$$

with elements x_i $(i = \overline{0, N})$ in a Hilbert space H. An element is a sequence consisting of m_i vectors x_{ij}, $j = \overline{1, m_i}$

$$x_i = |x_{i1}, \cdots, x_{im_i}|;$$

the number m_i is called its *size*.

A system is called *one–sized* if all the numbers $m_i = m$ are the same; in case $m = 1$, a system is called *one–dimensional.*

A system X and an element x_i are viewed as rows. Then the Gram matrices of system X and its elements x_i are written in the form:

$$\Gamma = \Gamma(X) = (X, X), \quad \Gamma(x_i) = \gamma_{ii} = (x_i, x_i).$$

If $\det \gamma_{ii} = 0$ then the corresponding element is called *singular*. If $\det \Gamma = 0$ then the system of vectors X is called singular as well. The segment $|x_l, \cdots, x_k|$ of a sequence X is denoted by $X_{l,k}$, $(k = \overline{0, N}$, $l = \overline{0, k})$, its initial segment $|x_0, \cdots, x_k|$ is denoted by X_k.

Denote the linear span of vectors $x, \cdots, y$ by $S(x, \cdots, y)$; in particular,

$$S = S(X), \quad S_k = S(X_k), \quad S_{l,k} = S(X_{l,k}), \quad S_{(k)} = S(x_k).$$

Let $P_{l,k} = P(S_{i,k})$ be a projection operator on $S_{l,k}$, and $\Pi_{l,k} = I - P_{l,k}$ be a projection operator on $S_{l,k}^{\perp} = H \ominus S_{l,k}$, which is an orthogonal supplement of subspace $S_{l,k}$. In particular, $\Pi_{0,k} = \Pi_k$. We will assume for the convenience that

$$\Pi_{-1} = \Pi_{1,0} = I.$$

2. Counter orthogonalization processes.

Definition 1 (*of the elements* $\mathbf{f}$ *and* $\widetilde{\mathbf{f}}$ *of counter orthogonalization processes*). *Let*

a) $\quad f_k = \Pi_{k-1} x_k, \qquad h_k = (f_k, f_k)$ — *forward process*;

b) $\quad f_{i/k} = \Pi_{1+i,k} x_i, \qquad h_{i/k} = (f_{i/k}, f_{i/k}),$

$\quad \widetilde{f}_k = \Pi_{1,k} x_0 = f_{0/k}, \qquad \widetilde{h}_k = (\widetilde{f}_k, \widetilde{f}_k) = h_{0/k}$ — *backward process*;

$$k = \overline{0, N} \qquad i = \overline{k, 0}$$

$$f_{k/k} = x_k, \quad \Pi_{-1} = I, \qquad \Pi_{k+1,k} = I.$$

Lemma 1. *Let X be a nonsingular system of vectors. Then the following conditions are satisfied:*

1. the elements f_l and $\widetilde{f}_l$ of orthogonalization processes from Definition 1 are nonsingular: $\det h_l \neq 0$, $\widetilde{h}_l \neq 0$, $l = \overline{0, N}$;.

2. the operators Π_l can be calculated with the use of elements f_l and $\widetilde{f}_l$ by means of recurrent formulas from the initial conditions

$$\Pi_{-1} = \Pi_{1,0} = I;$$

3. the following equalities hold for $k = \overline{-1, N-1}$:

$$\Pi_{k+1}(\cdot) = \Pi_k(\cdot) - f_{k+1}a_{k+1}(\,\cdot\,, f_{k+1}), \qquad a_{k+1} = h_{k+1}^{-1};$$

$$\Pi_{k+1}(\cdot) = \Pi_{1,k+1}(\cdot) - \widetilde{f}_{k+1}\widetilde{a}_{k+1}(\,\cdot\,, \widetilde{f}_{k+1}), \qquad \widetilde{a}_{k+1} = \widetilde{h}_{k+1}^{-1}.$$

3. Auxiliary results. Hereinafter we will discuss the block constructions only.

Definition 2. *Two systems of vectors are called equivalent (backward–equivalent) if they define the same forward (or backward) chain of subspaces.*

Definition 3. *We let $\mathbb{R}$ to be the set of all upper– (right–) triangular matrices (operators in E) with unit (or mixed–zero) diagonal blocks, $\mathbb{L}$ to be the set of all lower– (left–) triangular matrices (operators in E) with unit (or mixed–zero) diagonal blocks, and $\mathbb{D}$ to be the set of all block diagonal matrices (operators in E), i.e., matrices of kind $A = \operatorname{diag}\{a_k\}_0^N$.*

We write $C = FAG \in \mathbb{RDL}$, if $F \in \mathbb{R}$, $A \in \mathbb{D}$, and $G \in \mathbb{L}$; similar notations will be used for other combinations as well.

We will denote an arbitrary unitary or isometric operator by $\mathbb{Q}$.

Definition 4. *We consider the rows of X_k to be elements of E_k. We can also consider the natural row operators $X_k : E_k \to H$. Their zero extension in E are defined as*

$$\mathbf{X}_k = X_{k/N} = |X_k, 0_{N-k}| : E \to H; \qquad \mathbf{X} = X_{N/N} : E \to H.$$

Define some further notations, which will be referred to as notation (a), (b), etc.

a) We denote by $K_F = \operatorname{Ker} F$ the kernel of an operator F and we denote by $T_F = K_F^{\perp}$ its effective domain, which is the orthogonal complement of it; $S_F = \operatorname{Im} F$ stands for its image domain.

b) In particular,

$$T_{X_k} = T_k = (\operatorname{Ker} \boldsymbol{X}_k)^{\perp} = E_k \subset E, \qquad T = T_N = E,$$

$$S_k = \operatorname{Im} \boldsymbol{X}_k \subset H, \qquad S = S_N = H.$$

c) By $\boldsymbol{I}_k = I_{k/N} : E_k \to E$, $I_k = |e_0, \cdots, e_k|$ we denote the row operator of orthos in E.

d) The composite indices of the form "k/l" defined earlier, denote inessential (zero) operator extension from E_k into E_l (for $k<l$, see, e.g., Definition 4), or its narrowing (for $k>l$). The former means rows and matrices complement via zero blocks, the latter means blocks removal.

e) Changing an operator symbol to semibold face denotes its extension to E_N (see Definition 4).

f) An index k of the $(k+1)\times(k+1)$–matrix (of the operator) $\{F_{ij}\}_0^k$ in parenthesis denotes its last column (corresponding to ortho $e_{k/k}$): $F_{\cdot k} = F_{(k)}$. An index in square brackets is the first column (corresponding to $e_{0/k+1}$): $F_{\cdot 0} = F_{[k]}$. A composite index $k+1/k$ means that there is no diagonal block in these block–vectors: F_{kk} and F_{00} respectively (see (d)).

g) $A_k = \operatorname{diag}\{a_i\}_0^k \in \mathbb{D}$.

h) $\widetilde{A}_k = \operatorname{diag}\{a_{i/k}\}_0^k = \operatorname{diag}\{\widetilde{a}_k, \{a_{i/k}\}_1^k\} \in \mathbb{D}$.
i) $A^{-*} = (A^{-1})^*$.

Remark 1. One can easily show that the subspaces $T_A = K_A^{\perp}$ and S_A are isomorphic. It is also known that $(\operatorname{Ker} A)^{\perp} = \operatorname{Im} A^*$, i.e., $K_A^{\perp} = S_{A^*}$.

The following important statement follows from Lemma 1.

Theorem 1. *Two nonsingular systems of vectors are equivalent (backward–equivalent) if and only if they are related via a non singular upper–triangular (lower–triangular) transformation.*

By the above, such transformations will be called equivalent (backward equivalent) transformations. This statement is well known, at least for direct chains $C_k = \{S_i\}_0^k$. Its generalization for backward chains $\widetilde{C}_k = \{S_{k-i,k}\}_0^k = \{\widetilde{S}_{i/k}\}_0^k$ is quite obvious.

Definition 1 describes two counter processes of a sequential orthogonalization. For any k, k steps of a backward orthogonalization starting from element x_k correspond to the single kth step of a direct one (for this element x_k). As result, for every k, $k = \overline{0, N}$, we obtain two orthogonal systems

$$\Phi_k = \{f_i^*\}_0^{k*} = |f_0 = x_0, \cdots, f_{k-1}, f_k|, \qquad \text{and}$$

$$\widetilde{\Phi}_k = \{f_{i/k}^*\}_{i=0}^{k*} = |f_{0/k} = \widetilde{f}_k, \cdots, f_{k/k} = x_k|$$

equivalent, in their own way, to system X. Here

$$\Phi_k = |\Phi_{k-1}, f_k|,$$

whereas the system

$$\widetilde{\Phi}_k = |\widetilde{\Phi}_{k-1/k}, x_k| = |\widetilde{f}_k, \widetilde{\Phi}_{1,k-1/k}, x_k|,$$

in general case could be completely restored at every step. Any of "utmost" of k–step orthogonalization vectors, f_k or $\widetilde{f}_k$, determine projectors P_k and Π_k on S_k and $S_k^{\perp}$ by recurrent equations of Lemma 1.

Before we prove an intermediate result (Theorem 2), consider some auxiliary propositions.

Lemma 2. *Let*

1. a linear operator $\boldsymbol{X} : H_2 \to H$, where H and H_2 are separable unitary spaces, be completely continuous;

2. H be a separable unitary space which is isomorphic to each of the following subspaces

$$T = K^{\perp} = (\operatorname{Ker} \boldsymbol{X})^{\perp} \subset H_2 \quad \text{and} \quad S = \operatorname{Im} \boldsymbol{X} \subset H;$$

3. V and W be linear invertible (i.e. $K_V = \emptyset$, and $K_W = \emptyset$) operators such that

$$V : H_1 \to \operatorname{Im} V = T \quad \text{and} \quad W : H_1 \to \operatorname{Im} W = S;$$

4. the domains of operators V and W are everywhere dense in H_1.

Then

1. the operator $W^* \boldsymbol{X} V$ is one–to–one invertible and

2. the Moore–Penrose generalized inverse operator can be written in the form

$$\boldsymbol{X}^{(-1)} = V (W^* \boldsymbol{X} V)^{-1} W^*.$$

Lemma 3. *Let $\boldsymbol{X} : E \to H$. Denote a scalar product in H by $(\,\cdot\,,\,\cdot\,)_H = (\,\cdot\,,\,\cdot\,)$. Let a scalar product in E be defined by means of a positive matrix R_E:*

$$(\lambda, \mu)_E = \overline{\mu}^T\, R_E\, \lambda.$$

Then the operator $\boldsymbol{X}^*$ conjugated to the row operator $\boldsymbol{X}$ has the form

$$\boldsymbol{X}^* = R_E^{-1}(\,\cdot\,, X).$$

Now we can formulate the following
Corollaries (of Lemmas 2 and 3).

1. $\Gamma(X) = R_E^{-1}(X, X) = \boldsymbol{X}^*\boldsymbol{X}$;
2. $\boldsymbol{X}^{(-1)} = (X, X)^{-1}(\,\cdot\,, X) = (\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*$;
3. $P(S) = X(X, X)(\,\cdot\,, X) = \boldsymbol{X}(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*$.

Prove (3). First recall that $S = \operatorname{Im}\boldsymbol{X}$.

Here Lemma 2 (where $H_1 = H_2 = E$) and auxiliary operators are applied with $W = \boldsymbol{X}$, $V = I$. We also take into account that $P(\operatorname{Im}\boldsymbol{X}) = \boldsymbol{X}\boldsymbol{X}^{(-1)}$. If $R_E = I$ then $\boldsymbol{X}^* = (\,\cdot\, X)$, $\Gamma = (X, X)$. Thereafter $R_E = I$.

Lemma 4. *Let $F \in \mathbb{R}$, ($\widetilde{F} \in \mathbb{L}$) and $Q \in \mathbb{Q}$ be a unitary operator in E. Then*

$$\{FQ, QF\} \in \mathbb{R} \quad (\{\widetilde{F}Q, Q\widetilde{F}\} \in \mathbb{L}) \textit{ if and only if } Q = I.$$

In what follows, we discuss corollaries from Lemma 1.

4. Orthogonal decompositions.

First we summarize corollaries to Definition 1, Lemma 1, and auxiliary Lemmas 2,3,4. This will supplement and specify the results of Theorem 1.

Theorem 2. *Let X be a linearly independent system of vectors in a unitary space H and let X_k, $k = \overline{0, N}$ be its subsystems, initial segments of system X, which are simultaneously the base of line C_k and of counter $\widetilde{C}_k$ of chains of subspaces S_i and $S_{i/k} = S_{k-i,k}$, $i = \overline{0, k}$.*

Then

1. systems of vectors

$$\begin{aligned}
\Phi_k &= |f_0, \cdots, f_k| = |\Phi_k, f_k| = X_k F_k, & \Phi_N &= \Phi; \\
\widetilde{\Phi}_k &= |f_{0/k}, f_{1/k}, \cdots, f_{k/k}| & & \\
&= |\widetilde{f}_k, f_{1/k}, \cdots, f_{k/k}| = X_k \widetilde{F}_k, & \widetilde{\Phi}_N &= \widetilde{\Phi};
\end{aligned}$$

$k = \overline{0, N}$, are orthogonal equivalents (direct and backward) of systems X_k, i.e., they are orthogonal bases of chains C_k and $\widetilde{C}_k$, and elements f_i, $f_{i/k}$, $i = \overline{0, k}$, are bases of orthogonal chain elements

$$\Delta S_k = S_k \ominus S_{k-1}, \quad \Delta S_{i/k} = S_{i,k} \ominus S_{i+1,k}$$

of these chains respectively;

2. The corresponding equivalent transformations are upper– ($F \in \mathbb{R}$) *and lower–* ($\widetilde{F} \in \mathbb{L}$) *triangular:*

$$F_k = |F_{k-1/k}, F_{(k)}| = (X_k, \Phi_k)^{-1} A_k^{-1} \in \mathbb{R},$$

$$\widetilde{F}_k = |\widetilde{F}_{[k]}, \widetilde{F}_{1,k/k}| = (X_k, \widetilde{\Phi}_k)^{-1} \widetilde{A}_k^{-1} \in \mathbb{L},$$

$$A_k = (\Phi_k, \Phi_k)^{-1}, \qquad \widetilde{A}_k = (\widetilde{\Phi}_k, \widetilde{\Phi}_k)^{-1}.$$

3. If diagonal blocks of matrices F_k and $\widetilde{F}_k$ of these QR– and QL– representations of the system X are unity blocks then these representations are unique. The same holds for the equivalent orthonormal systems

$$\Phi_{\mathbf{n}} = XG, \qquad \widetilde{\Phi}_{\mathbf{n}} = X\widetilde{G}, \qquad \textit{where}$$

$$G = FA^{1/2} \in R, \qquad \widetilde{G} = \widetilde{F}\widetilde{A}^{1/2} \in L; \qquad \textit{also}$$

$$\widetilde{\Phi}_{\mathbf{n}} = \Phi_{\mathbf{n}} Q, \qquad \textit{where}$$

$$Q = (\widetilde{\Phi}_{\mathbf{n}}, \Phi_{\mathbf{n}}) = G^{-1}\widetilde{G} \in \mathbb{Q}$$

is a unitary operator in E.

Proof. The first and the second statements follow from the recursive equation for projectors from Lemma 1. Let us elucidate this for elements $f_i = x_i - P_{i-1}x_i$ (analogously for elements $f_{i/k} = x_i - P_{i+1,k}x_i$). From Lemmas 2,3 and their corollaries we deduce

$$P(\operatorname{Im} \boldsymbol{X}_k)(\cdot) = X_{k/N}\Gamma_{k/N}^{(-1)}(\,\cdot\,, X_{k/N}) = X_k\Gamma_k^{-1}(\,\cdot\,, X_k).$$

The generalized inverse operators $\boldsymbol{X}_k^{(-1)}$ are obtained from Lemma 2 by letting $W = X_k$, $V = I_k$. Then $\boldsymbol{X}_k^{(-1)} = I_k(X_k^*\boldsymbol{X}_k I_k)^{-1}X_k^*$ and Theorem 2 follows from Lemma 4.

Now, applying Theorem 2 to corollaries based on Lemma 1, one can deduce other important corollaries of this Lemma.

Corollary 1. *Counter orthogonalization processes of bases X_k for subspace chains S_k generate one–valued QR– and QL– decompositions of the corresponding matrices X_k and operators $\boldsymbol{X}_k$. They are unique.*

While calculating operators $\boldsymbol{X}_k^{(-1)} = X_{k/N}^{(-1)}$ in the proof of Theorem 2, we have applied auxiliary operators $W = X_k$ (Lemma 2). Row operators of Φ_k and $\widetilde{\Phi}_k$ satisfy the conditions of Lemma 2 for the operator W in accordance with (1) of Theorem 2. Using these operators, we obtain the other formulas for generalized inverse operator $\boldsymbol{X}_k^{(-1)}$: with a right– (upper–) and left– (under–) triangular invertible matrices. Thus we have

Corollary 2.

$$a. \quad \boldsymbol{X}_k^{(-1)} = (\boldsymbol{\Phi}_k^*\boldsymbol{X}_k)^{(-1)}\boldsymbol{\Phi}_k^*,$$

$$b. \quad \boldsymbol{X}_k^{(-1)} = (\widetilde{\boldsymbol{\Phi}}_k^*\boldsymbol{X}_k)^{(-1)}\widetilde{\boldsymbol{\Phi}}_k^*,$$

where

$$\boldsymbol{\Phi}_k^*\boldsymbol{X}_k = \mathbf{A}_k^{(-1)}\mathbf{F}_k^{(-1)} \in \mathbb{DR}, \qquad \widetilde{\boldsymbol{\Phi}}_k^*\boldsymbol{X}_k = \widetilde{\mathbf{A}}_k^{(-1)}\widetilde{\mathbf{F}}_k^{(-1)} \in \mathbb{DL},$$

and $\boldsymbol{\Phi}$, $\widetilde{\boldsymbol{\Phi}}$ are orthogonal system of vectors.

Here we use notations (g) and (h) for operators $\boldsymbol{A}_k$ and the result of Lemma 3 for the operator $\boldsymbol{X}^* = (\,\cdot\,, X)$. A generalized inversion of the $\boldsymbol{\Phi}_k^* \boldsymbol{X}_k$ and $\widetilde{\boldsymbol{\Phi}}_k^* \boldsymbol{X}_k$ in this corollary is handled by Lemma 2 with $W = V = I_k$.

5. Factorizations and inversions. Here we give some more corollaries from Lemma 1.

Corollary 3.

(a) *Invertible self–conjugate operators in E have unique $\mathbb{RDL}$– and $\mathbb{LDR}$– decompositions.*
(b) *They are defined by the processes of direct and backward orthogonalization of the initial system.*

PROOF. From Theorem 2 we have:

$$\begin{aligned}
\Gamma &= (X, X) = F^{-*}(\Phi, \Phi)F^{-1} = \widetilde{F}^{-*}(\widetilde{\Phi}, \widetilde{\Phi})F^{-1} \\
&= F^{-*}A^{-1}F^{-1} \qquad\qquad (\in \mathbb{LDR}) \\
&= \widetilde{F}^{-*}\widetilde{A}^{-1}\widetilde{F}^{-1} \qquad\qquad (\in \mathbb{RDL}); \\
\Gamma^{-1} &= (X, X)^{-1} = FAF^* \qquad (\in \mathbb{RDL}) \\
&= \widetilde{F}\widetilde{A}\widetilde{F}^* \qquad\qquad\qquad (\in \mathbb{LDR}),
\end{aligned}$$

where

$$F = (X, \Phi)^{-1}A^{-1} \in \mathbb{RD}, \qquad \widetilde{F} = (X, \widetilde{\Phi})^{-1}\widetilde{A}^{-1} \in \mathbb{LD}.$$

Remark 2. It is known that, if a matrix in some orthonormal basis defines a positive operator then it is a Gram matrix of a certain basis (i.e., of a nonsingular systems of vectors). In Corollary 4, we show that one can calculate decompositions from Corollary 3 without knowing the mentioned basis. Then the statement (a) of Corollary 3 follows from Lemma 1. Corollaries 6 and 8 show how to construct such bases. Therefore, the mentioned proposition on the “origins” of positive operators also follows from Lemma 1.

By virtue of notations (d), (f) and Lemma 3, the equalities of Lemma 1 can be written in the following form:

$$\begin{aligned}
f_{k+1} &= \Pi_k x_{k+1} = X_{k+1}F_{(k+1)} \\
&= x_{k+1} - X_k \Gamma_k^{-1}(x_{k+1}, X_k), \\
\widetilde{f}_{k+1} &= \Pi_{1,k+1}x_0 = X_{k+1}\widetilde{F}_{[k+1]} \\
&= x_0 - X_{1,k+1}\Gamma_k^{-1}(x_0, X_{1,k+1}),
\end{aligned}$$

where $\qquad (x_{k+1}, X_k) = \Gamma_{(k+1/k)}, \qquad (x_0, X_{1,k+1}) = \Gamma_{[k+1/k]}.$

Corollary 4. *Sequential inversion of bordered invertible matrices Γ_k, $k = \overline{0, N}$ (essentially extending the invertible operators $E_k \to E_{k+1}$), can be realized by the* *following recurrent formulas:*

$$\begin{aligned}
\Gamma_{k+1}^{-1} &= \Gamma_{k/k+1}^{-1} + F_{(k+1)} a_{k+1} F_{(k+1)}^*, \\
F_{(k+1)} &= e_{k+1/k+1} - \Gamma_{k/k+1}^{(-1)} \Gamma_{(k+1)}; \\
\Gamma_{k+1}^{-1} &= \Gamma_{1,k+1/k+1}^{-1} + \widetilde{F}_{[k+1]} \widetilde{a}_{k+1} \widetilde{F}_{[k+1]}^*, \\
\widetilde{F}_{[k+1]} &= e_{0/k+1} - \Gamma_{1,k+1/k+1}^{(-1)} \Gamma_{[k+1]}; \\
a_{k+1} &= (x_{k+1}, f_{k+1})^{-1} = (f_{k+1}, f_{k+1})^{-1} = F_{(k+1)}^* \Gamma_{(k+1)}, \\
\widetilde{a}_{k+1} &= (x_{k+1}, \widetilde{f}_{k+1})^{-1} = (\widetilde{f}_{k+1}, \widetilde{f}_{k+1})^{-1} = \widetilde{F}_{(k+1)}^* \Gamma_{(k+1)}.
\end{aligned}$$

They are two well–known Frobenius formulas for inverting bordered block matrices.

Let as unite these formulas by equating their expressions for a single block of inverse matrices. We obtain a qualitatively new result. From formulas of the direct orthogonalization of Corollary 4 we have

$$\begin{aligned}
\Gamma_{k+1/k}^{-1} &= \Gamma_k^{-1} + F_{(k+1/k)} a_{k+1} F_{(k+1/k)}^*, \\
F_{(k+1/k)} &= -\Gamma_k^{-1} \Gamma_{(k+1/k)}, \\
a_{k+1} &= \left[\gamma_{k+1,k+1} - \Gamma_{(k+1/k)}^* \Gamma_k^{-1} \Gamma_{(k+1/k)}\right]^{-1} \\
&= \left[(x_{k+1}, x_{k+1}) - (X_k, x_{k+1}) \Gamma_k^{-1} (x_{k+1}, X_k)\right]^{-1}.
\end{aligned}$$

To calculate this block ($\Gamma_{k+1/k}^{-1}$) from the backward orthogonalization process, we apply it for the system $X_{k+1} = |X_k, x_{k+1}|$ consisting of two elements:

$$\begin{aligned}
\Gamma_{k+1/k}^{-1} &= \widetilde{a}_1 = (X_k, \Pi(x_{k+1}) X_k)^{-1} \\
&= \left[(X_k, X_k) - (x_{k+1}, X_k)(x_{k+1}, x_{k+1})^{-1}(X_k, x_{k+1})\right]^{-1}.
\end{aligned}$$

Corollary 5. *The following formula for inverting matrices with factorized additive increment is valid:*

$$\left[\Gamma_k - \Gamma_{(k+1/k)} \gamma_{k+1,k+1}^{-1} \Gamma_{(k/k+1)}^*\right]^{-1} = \Gamma_k^{-1} + F_{(k+1/k)} a_{k+1} F_{(k+1/k)}^*.$$

This formula (matrix type Riccati discrete equation) is widely used as well as its continuous analogs; in particular, it is used in the least–square and in the Kalman filtration methods.

Thus, the Frobenius formulas from Corollary 4 are formulas of $\mathbb{RDL}$–factorization of the matrix $\Gamma^{-1} = FAF$ (Frobenius factorization). This matrix having a triangular inverse could be calculated quite easily:

$$F_{(k+1)}^{-1} = e_{k+1/k+1} - F_{k/k+1}^{(-1)} F_{(k+1)}.$$

By applying this formula, we can obtain a $\mathbb{LDR}$–factorization $\Gamma = F^{-*} A^{-1} F^{-1}$ from Frobenius factorization. Thus, we arrive at the well–known Cholessky factorization algorithms of this type.

Corollary 6. *The Cholessky formulas of $\mathbb{LDR}$–factorization of a nonsingular positive matrix Γ could be written in the form:*

$$F_{k+1}^{-1} = e_{k+1/k+1} + A_{k/k+1}F^{*}_{k/k+1}\Gamma_{(k+1)},$$
$$a_{k+1}^{-1} = \gamma_{k+1,k+1} - F^{-*}_{(k+1/k)}A^{-1}_{k+1}F^{-1}_{(k+1/k)}.$$

Here we take into consideration the following formulas derived from those mentioned above:

$$F_{(k+1)} = e_{k+1/k+1} - \Gamma^{-1}_{k/k+1}\Gamma_{(k+1)} \quad \text{and}$$
$$\Gamma_k^{-1} = F_kA_kF_k^{*}.$$

Now we apply this reasoning to the backward factorization. We obtain

Corollary 7.

a) *The Frobenius ($\mathbb{RDL}$) and the Cholessky ($\mathbb{LDR}$) factorization algorithms generated by one of the orthogonalization processes are mutually dual and mutually inverse.*

b) *Processes of counter factorization correspond to counter orthogonalization processes.*

Remark 3. Inversion and factorization equations for arbitrary (nonsingular) matrices can be obtained from the analysis of biorthogonalization of two systems X and Y. Let $P(X^{\perp}\|Y)$ denote a projector onto the orthogonal supplement of the linear span of X parallel to the linear span of Y. (The orthoprojector onto $X^{\perp}$ is $P(X^{\perp}\|X)$.) An aim of biorthogonalization is to construct systems $\Phi_{\mathbf{n}} = XG$ and $\widehat{\Phi}_{\mathbf{n}} = Y\widehat{G}$ so that $(\Phi_{\mathbf{n}}, \widehat{\Phi}_{\mathbf{n}}) = \widehat{\boldsymbol{\Phi}}^{*}_{\mathbf{n}}\boldsymbol{\Phi}_{\mathbf{n}} = I$.

Let $X = |X_k, x_{k+1}|$, $Y = |Y_k, y_{k+1}|$. Then for a step of the direct biorthogonalization we have

$$P_{k+1}(\cdot) = P_k(\cdot) - f_{k+1}(f_{k+1}, g_{k+1})^{-1}(\cdot, g_{k+1}),$$
$$f_{k+1} = P_k\, x_{k+1}, \qquad g_{k+1} = P_k^{*}\, y_{k+1},$$

where
$$f_0 = x_0, \quad g_0 = y_0, \quad P_{-1} = I.$$

Thus, the method above yields us the general Frobenius and Cholessky equations for nonsingular matrices.

6. Some statements for Gram matrices. To formulate Corollary 8, we will need two auxiliary statements.

Lemma 5. *Nonsingular systems $X \in H_1$ and $Y \in H_2$ have the same Gram matrix if and only if there exists an isometric operator $U : H_1 \to H_2$ such that*

a) $Y = UX$;

b) $S \subset T_U$ *(Here $S = S_X = \operatorname{Im}\mathbf{X}$; $T = K^{\perp}$, $K = \operatorname{Ker} U$, see notation (a)).*

To prove the necessity, one can use the following simple fact.

Lemma 6. *The systems $X \in H_1$ and $Y \in H_2$ have the same nonsingular Gram matrix if (and only if) $U = YX^{(-1)} : H_1 \to H_2$ is an isometric operator.*

Corollary 8 (from Lemma 1) *The systems X and Y have similar orthogonalization matrices (direct F and backward $\widetilde{F}$) if and only if they satisfy the conditions of Lemma 5.*

Let Γ_c be a positive operator, $\Gamma_c^{-1} = F_c A_c F_c^* = G_c G_c^*$, where $G_c = F_c A_c^{1/2}$ is a normalized matrix. It could be, for instance, an $\mathbb{RDL}$–factorization of Γ_c^{-1} calculated by Corollary 4. Let $\Phi_{\boldsymbol{n}} = \Phi A^{1/2} = XFA^{1/2} = XG$ be an orthonormal system equivalent to the given system X (see statement (3) of Theorem 2) and hence $\Gamma = \Gamma(X) = G^{-*}G^{-1}$. Then the system $Y = \Phi_{\boldsymbol{n}} G_c^{-1}$ equivalent to the system X will have the given Gram matrix $\Gamma_c = \Gamma(Y) = G_c^{-*}G_c^{-1}$. Note that $G^{-1}\widetilde{G} = Q$ is a unitary operator in E and that $\widetilde{\Phi}_{\boldsymbol{n}} = \Phi_{\boldsymbol{n}} Q$.

Corollary 9 (from Lemma 1) *In a chain of subspaces one can construct the unique basis with the given Gram matrix.*

The case of Toeplitz Gram matrices is of a particular interest.

3. Homogeneous Systems

7. The principal result.

Definition 5. *A system of vectors X is called homogeneous (U–homogeneous) if it is*

1) *one–sized, i.e., $m_i = m$, for all $i = \overline{0, N}$,*
2) *its elements are related isometrically, and hold $x_{i+1} = Ux_i$, $x_i = U^* x_{i+1}$, $i = \overline{0, N-1}$.*

Here U is a partially isometric operator: $U : (\operatorname{Ker} U)^{\perp} = T_U \longrightarrow S_U = \operatorname{Im} U$, $S(X_{N-1}) \subset T_U$, $S(X_{1,N}) \subset S_U$.

Lemma 7. *A system of vectors has a Toeplitz Gram matrix if and only if it is homogeneous.*

Now we can formulate the principal result:

Theorem 3. *Assume that systems X are U–homogeneous and nonsingular. Then for $k = \overline{0, N-1}$ the following conditions are satisfied:*

1. the orthogonalization elements f and $\widetilde{f}$ satisfy the following equations:

$$f_{k+1} = Uf_k - \widetilde{f}_k \widetilde{\theta}_{k+1},$$

$$\widetilde{f}_{k+1} = \widetilde{f}_k - Uf_k \theta_{k+1},$$

$$f_0 = \widetilde{f}_0 = x_0,$$

where $\quad \theta_{k+1} = a_k \mu_{k+1}, \quad \widetilde{\theta}_{k+1} = \widetilde{a}_k \mu_{k+1}^*, \quad \mu_{k+1} = (g_k, Uf_k),$

and the factors a and $\widetilde{a}$ are defined as follows:

$$a_l = h_l^{-1}, \quad \widetilde{a}_l = \widetilde{h}_l^{-1}, \quad l = \overline{0, N}.$$

2. The Gram matrices h and $\widetilde{h}$ are nonsingular and satisfy the following non-linear recurrent equations: $h_0 = \widetilde{h}_0 = (x_0, x_0)$ and

$$h_{k+1} = h_k - \mu_{k+1} \widetilde{h}_k^{-1} \mu_{k+1}^*,$$

$$\widetilde{h}_{k+1} = \widetilde{h}_k - \mu_{k+1}^* h_k^{-1} \mu_{k+1}.$$

3. The corresponding recurrent equations for normalizing factors a and $\widetilde{a}$ can be written in the form: $a_0 = \widetilde{a}_0 = h_0^{-1}$ and

$$a_{k+1} = (I - \theta_{k+1} \widetilde{\theta}_{k+1})^{-1} a_k,$$

$$\widetilde{a}_{k+1} = (I - \widetilde{\theta}_{k+1} \theta_{k+1})^{-1} \widetilde{a}_k,$$

where all the inverse matrices exist.

4. Moreover, if the system of vectors X is one–dimensional $(m = 1)$ *then* $\theta = \widetilde{\theta}$ *are scalars less than 1 and factors* $a_k = \widetilde{a}_k$, $k = \overline{0, N}$, *are positive quantities. This means that* $a_0 = \widetilde{a}_0 = (x_0, x_0)^{-1}$ *and* $a_{k+1} = (1 - \theta_{k+1}^2)^{-1} a_k$, $1 - \theta_{k+1}^2 > 0$.

Remark 4. Formulas of types (2) and (3) are encountered in various applications related to the inverse problems for homogeneous media and dynamic systems. Examples of them can be inversion and block-triangular factorization of block–Toeplitz matrices, as well as solving the corresponding equation systems [21] and recurrent calculations of polynomials which are orthogonal on the unit circumference [18].

Let us mention some related applied problems: theories of propagation in isotropic space (astrophysics) [7,8], in homogeneous layer structures [5], the P. Gupillo model (geophysics) [9], and estimation and control problems in constant information systems (stationary process analysis): their restoration, modeling, identification [13,12,28].

Theorem 3 gives us the basis for these and other similar results. Lemma 1 is also the basis for many important results like the Frobenius formulas, matrix inversion lemmas, their block-triangular factorizations, $\mathbb{QDR}$– and $\mathbb{QDL}$– decompositions, etc.

These results could be formulated for two systems of vectors biorthogonalization processes similar to those discussed in this paper. For some problems mentioned above, there are solutions of their continual analogs (see a fundamental M. Krein's paper [20]).

8. Proof of Theorem 3. To prove it, we need some auxiliary facts. The essential one is the following

Theorem 4. *Elements* $f_{i/k}$, $i = \overline{0, k}$ *of the orthogonal system*

$$\widetilde{\Phi}_k = |f_{0/k}, f_{1/k}, \cdots, f_{k/k}|, \quad k = \overline{0, N}$$

(*see assertion 1 of Theorem 2*) *backward equivalent to U–homogeneous system X are determined by the elements*

$$\widetilde{f}_j = f_{0/j}, \qquad j = \overline{0, k}$$

via the formulas

$$f_{i/k} = U^i \widetilde{f}_{k-i}, \qquad i = \overline{0, k}.$$

Theorem 4 follows from

Lemma 8. *In a homogeneous system, the following equations hold for* $k = \overline{0, N}$, $i = \overline{0, N - k}$:

(1) $P_{i,k+i} x = U^i P_k U^{-i} x, \quad \forall x \in S_{i,N}$;
(2) $\Pi_{i,k+i} = U^i \Pi_k U^{-i} x, \quad \forall x \in S_{i,N}$.

Here we take into account that $U^i U^{-i} = I$ in $S_{i,N}$.

Proof of Theorem 4. By Lemma 8, for $k = \overline{0, N}$, $i = \overline{0, k}$, we have:

$$f_{i/k} = \Pi_{i+1/k} x_i = (U_i U_i^* - U^i P_{1,k-i} U^{-i}) x_i =$$

$$= U^i \Pi_{1,k-i} U^{-i} x_i = U^i \Pi_{1,k-i} x_0 = U^i \widetilde{f}_{k-i}.$$

Theorem 4 is complete.

Here we turn back to Theorem 3. To prove it, we need two more statements. We obtain the first one from Lemma 8 (2). This is a key fact for homogeneous systems:

Corollary (from Lemma 8) *For $k = \overline{0, N-1}$, the following equality holds:*

$$\Pi_{1,k} x_{k+1} = \Pi_{1,k} U x_k = U f_k.$$

The second statement follows from Lemma 8 (1) and the first equation of Lemma 1.

Lemma 9. *For $k = \overline{0, N-1}$, the following equality holds:*

$$\Pi_{1,k+1} = U \Pi_k U^* = \Pi_{1,k} - P(U f_k).$$

It follows from the corollary above that

$$\Pi_{1,k+1} = I - U P_k U = I - U \left(P_{k-1} + P(f_k)\right) U^*.$$

Proof of Theorem 3. We calculate the element

$$f_{k+1} = \Pi_k x_{k+1}$$

by applying the second equation for the projector from Lemma 1. As a result, we obtain the first equation of Theorem 3. We assume that the projector of Lemma 9 affects the element x_0. Then we obtain the second equation of Theorem 3.

Formulas of the second statement are obtained by applying the equations of the first statement to the equations $h_k = (f_k, x_k)$ and $\widetilde{h}_k = (\widetilde{f}_k, x_0)$.

The equations of the third statement follow from the definitions of coefficients μ and parameters θ. The formulas of the fourth statement are obvious. The nonsingularity of the invertible matrix follows from the nonsingularity of X, the first statement of Lemma 1, and $\mathbb{RDL}$– and $\mathbb{LDR}$– Gram matrix factorizations of X obtained above. From here we obtain that

$$\det \Gamma(X) = \mathbf{\Pi}_0^N h_i = \mathbf{\Pi}_0^N \widetilde{h}_i.$$

8. Examples of homogeneous systems.

From Theorem 3 we obtain, for instance, some procedures of block Toeplitz matrices inversions [21], synthesis of orthogonal polynomials on the unit circumference [18] and a Hardy space [1,2], fast algorithms for signal filtration [13] and variational identification of difference and differential equations [29]. These applications can be obtained by the construction in H described below as well as by the similar construction in a Hardy space [1,2], which is considered as a particular case.

Example 1. Let $Y = Y_{N+n} = \{U^i y_0\}_0^{N+n}$ be a homogeneous and orthonormal system of vectors in H (i.e., $(Y, Y) = I_{N+n}$), and let the polynomial

$$x_0 = p = \sum_0^n y_i p_i = \sum_0^n U^i y_0 p_i = p(U, y_0) = p(y_0)$$

be an initial vector of the homogeneous system

$$X = \{U^j x_0\}_0^N = X_N = X(p) = p Y_N.$$

Then $\Gamma = C = (X, X) = G^{-*}G^{-1}$, $G = FA^{1/2} \in \mathbb{R}$ is a Toeplitz matrix. Theorem 3 and Frobenius algorithms in Corollary 4 yield us the known effective technique of its inversion (see [21]).

Example 2. Let the columns of the matrix

$$G = FA^{1/2} = \{g_{ij}\}_0^N \in R$$

be the coefficients of orthonormal polynomials [2]

$$p_k(U) = \sum_0^k U^i x_0 g_{ik} = f_k a_k^{1/2}.$$

From the orthonormal system $\Phi_{\boldsymbol{n}}{=}XG$ we easily obtain the algorithms calculating these polynomials (i.e., g_{ik}, see [18]), which are similar to the algorithms of Toeplitz matrices inversion [21].

Example 3. Block–vector $F_{(k+1/k)}$ from Frobenius procedure of the inversion of the self–conjugate nonsingular matrix (Corollary 4) is also known to be an optimal linear predictor of length k for $\breve{x}_{k+1}$ of the random process X_{k+1} with covariational matrix Γ_{k+1} according to the visualized value $\breve{X}_k$.

It is easily seen that the vector components $F_{[k+1/k]}$ are coefficients of the revertible (backward) predictor, i.e., of the restorer of the former (initial) value $\breve{x}_0$ of this process due to he fact that we know covariational matrix Γ_{k+1} and the observable values $\breve{X}_{1,k+1}$.

Example 4. Now we give a pure mathematical example. Let H be a Hilbert space of Gauss random values. Then the norm is the dispersion of random values and the scalar product is the covariational of two random values. The ordered system of vectors X in H is a process consisting of random elements $x_i, i = \overline{1, N}$. A covariational matrix of this process is the matrix $\Gamma = (X, X) = \mathbb{E}(\breve{X}, \breve{X})$, where $\mathbb{E}$ is the symbol for the mathematical expectation.

The minimized value for the least–mean–square predictor for $\breve{x}_{k+1}$ equals to $J = \mathbb{E}\|\breve{x}_{k+1} - \widehat{x}_{k+1/k}\|^2$.

The linear prediction value is computed as the element $\widehat{x}_{k+1/k} = \breve{X}_k\lambda_{k+1}$. Let us now differentiate J with respect to λ. The derivative vector can be written in the form $\mathbb{E}((\breve{X}_k\lambda_{k+1} - \breve{x}_{k+1}), \breve{X}_k)$. After equating it to zero we obtain $\lambda_{k+1} = (X_k, X_k)^{-1}(x_{k+1}, X_k)$. Comparing the latter with the second equation of Corollary 4, we see that $\lambda_{k+1} = -F_{(k+1/k)}$ (see notation (f)).

It follows from Corollary 4 that the complete vector $F_{(k+1)}$ is a prediction vector for the error coefficient $\breve{x}_{k+1} - \widehat{x}_{k+1/k}$.

Thus, optimal predictors (direct and backward) are calculated by means of equalities from Corollary 4.

Example 5. This example is associated with the homogeneous construction under consideration. If H is the space of random values then any orthonormal homogeneous system Y in H is a process of independent random values with a unit dispersion (basis in $S(Y) \subset H$).

The system X is a random process formed as the moving average of process Y with coefficients α. A covariation matrix of this process is a Toeplitz matrix $\Gamma = C = (X, X)$. The optimal predictors are calculated like in Examples 3,4 with the use of Frobenius formulas in Corollary 4 and Theorem 3.

Example 6. Now we give two more examples associated with signal estimation and their modeling. Whereas in the previous case, there is used a signal model in the regressive form of moving average $X = pY$ (model subspace $S(X) = S(pY)$), in the ensuing cases a model conjugated to X is applied. This is an auto–regression model $X^{\perp}$. Consider a model subspace

$$M = S(X^{\perp}) = S^{\perp}(X) = S \ominus S(X), \qquad S = S(Y).$$

This subspace is a set of its transient processes. It is the orthogonal complement $S^{\perp}(X)$ to the subspace $S(X)$ of the regression–moving–average model. The Gram matrix of the basis $X = pY = S \ominus M$ in $S(X)$ is $\Gamma(X) = C$.

Example 7. Let $p(U) : S(Y_N) \to S(Y)$ (see Example 1) and let

$$M = M(x_0) = S^{\perp}(X) = S(Y) \ominus S(X).$$

Then the following fact takes place.

Theorem 5. *An element φ is in M if and only if $\varphi \in \operatorname{Ker} p^*(U)$, i.e., $p(U)\varphi = 0$ or $\widetilde{p}\varphi = r_{(n-1)}$. (Here $\widetilde{p} = U^n p^*$ and $r_{(n-1)} = r_{(n-1)}(U)$ is an $(n-1)$–polynomial.)*

Corollary. *Let $\varphi = \sum_0^{N+n} y_k \varphi_k \in M$. Then the sequence of coordinates $\varphi_k = (\varphi, y_k)$ of this function in the basis $Y \in H$ is a solution to the difference equation $\sum_0^n \varphi_{i+j} p_i^* = 0, \quad j = \overline{0, N}$.*

The projection h from H onto $M(x_0)$, for a given x_0, is a smoothing and a filtration of a problem. Theorem 1 yields us symmetric fast algorithms for a numerical solution of such a problem.

Example 8. The search for an x_0 whose distance $\rho(h, M)$ between the function h and the subspace M is minimal, is a variational problem of approximated modeling by a difference equation for a function or a sequence. This method is also used for identifying differential equations with their corresponding discretization. For details and references see [29,30].

Remark 5. The last point to be stressed here is that, given a positive Toeplitz matrix C, its decomposition (X, X), where $X = X(\alpha)$ is a homogeneous systems of above–mentioned type, can be obtained by a vector–matrix analog of the well–known procedure of calculating square roots. Therefore, it follows from Corollary 9 that the subspace chain has a set of bases where a chain of its orthoprojectors satisfies equations from Theorem 3 with different initial conditions and operators U.

4. Variational Identification

Here we develop and explain the ideas of Examples 6,7,8 in p. 8 of the previous section.

1. Theory. Theorem 5 and its Corollary may be successfully adapted to the approximating identification problem in a finite interval for a complex dynamical object via a linear constant model. We call it a *piecewise linear (dynamical) approximation* [29].

According to this Corollary, projecting of an arbitrary function $\psi \in S(Y) = S$ onto the subspace $M = S \ominus S(pY_N)$ which was described in the Examples 1 and 6 of p. 8 of section 3 is equivalent to projecting an arbitrary element (column

$(N+n+1)$–vector) $[\psi] = \{\psi_i\}_0^{N+n} = |\psi_0^*, \cdots, \psi_{N+n}^*|^*$ in spaces $l^2[0, N+n] = l^2$ or $E = E^{N+n+1} = E[0, N+n]$ onto the kernel of the following difference operator

$$D = D_\alpha = \sum_0^n s^i \alpha_i \, : E \to E[0, N],$$

where s is a shift operator such that $s\psi_i = \psi_{i+1}$, and $\alpha_i = p_i^*$. This means that if

$$\psi = \sum_0^{N+n} U^i y_0 \psi_i \in S \Longleftrightarrow [\psi] \in E_N \text{ or } l_N^2 \quad \text{ then}$$

$$D\psi_k = \sum_0^{N+n} \psi_{k+i}\alpha_i \in E[0, N], \quad i.e. \quad k = \overline{0, N}.$$

We recall here that if $\psi \in S = S(Y) \subset H$ (Y is an orthonormal system from Example 1 in point 8 of section 3) then $\psi_i = (\psi, y_i) = (\psi, y_i)_H$. The kernel of operator D is a vector of coefficients for the function $\varphi \in M$, therefore $D\varphi_k = 0$. In E we have an important equation [31,28]:

$$A^* \cdot [\varphi] = 0, \quad A = A_\alpha,$$

where A is the matrix of the operator D^* in $E[0, N]$ and A^* is the matrix of the operator D in E.

We consider the special Toeplitz $(N+n+1) \times (N+1)$–matrix, a matrix of the sliding vector [16]:

$$A = \begin{vmatrix} \alpha_0^* & & 0 \\ \vdots & \ddots & \\ \alpha_n^* & \cdots & \alpha_0^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{vmatrix} = \begin{vmatrix} p_0 & & 0 \\ \vdots & \ddots & \\ p_n & \cdots & p_0 \\ & \ddots & \vdots \\ 0 & & p_n \end{vmatrix}.$$

Its first column is denoted by x_0.

One can easily see that, if a function ψ is in S_n then the vector of coefficients for the function $\xi = p\psi$ is $[\xi] = A \cdot [\psi]$.

This means that the matrix A in the basis Y is an operator matrix of multiplying by a polynomial [1,2].

Now it is not difficult to see that the columns of the matrix A form a homogeneous basis X in the subspace $S(X) = S(pY)$ described above (Examples 1,6,7 of p. 8, section 3). This means that the projector Π onto the orthogonal complement to $S(X)$ can be calculated via the formulas of Theorem 3, Lemma 1, and Definition 1.

In these common results, it is necessary to assume that

$$H = E, \quad x_0 = \eta_0, \quad X = A, \quad \text{and} \quad U = \{\delta(i, j+1)\}_0^{N+n},$$

where U is "down–shift" operator in E and $\delta(i, j)$ is the Kronecker symbol. The matrix of the partially isometric shift operator U in basis E is

$$U = \begin{vmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{vmatrix}.$$

Now we have $A = |x_0, x_1, \cdots, x_N|$, $x_i = U^i x_0$. We see that the equation $A^* \cdot [\varphi] = 0$ is equivalent to the difference equation with operator D for the components of the vector $[\varphi] \in \operatorname{Ker} D$. We can consider the variational identification problem for this equation. We will see that this is an approximating identification.

2. Application. Any estimation problem can be reduced to some projecting problem. Let $[\psi] \in E$ be a reference vector of some recorded or desired dynamical process. The problem is to find a linear constant difference equation, i.e., to estimate its coefficients whose transient process with suitable (optimal) initial conditions would describe the given process in the best way.

The possible formulation of this problem is to find the minimum of the functional

$$J = \|[\psi] - [\varphi]\|_E^2 \quad \text{if} \quad A_\alpha^*[\varphi] = 0$$

with respect to α (see [29,30]).

Special case. If α is given, then the formulated problem is a smoothing problem for the sequence of the components of the vector $[\psi]$ via a constant difference equation (model). This optimization step of the variational functional J is the search for the best initial conditions of the approximating transient process of the given model. The solution to this problem is

$$[\widehat{\varphi}] = \Pi[\psi], \quad \Pi = \Pi(A) = \Pi_N = I - P_N(A) = I - P,$$

where

$$\Pi = I - A(A^*A)^{-1}A^*, \quad P = A(A^*A)^{-1}A^*.$$

Lemma 1 and Theorem 3 yield us *fast* calculation algorithms without Riccati equation for projector Π and the projection $[\widehat{\varphi}]$.

Common case. Assume that the vector α is unknown. Then the value is calculated as

$$\rho^2 = \widehat{J} = \|[\psi] - [\widehat{\varphi}]\|^2.$$

i.e., $\rho^2 = \rho^2(\alpha) = [\psi]^* A(A^*A)^{-1}A^*[\psi] = [\psi]^* P[\psi]$, which is the square of the perpendicular from the given process vector ψ onto the kernel of the desired difference operator D_α.

The solution of the approximating identification is $\widehat{\alpha} = \arg\min \rho^2(\alpha)$. Lemma 1 and Theorem 3 provide us the *fast* calculation algorithms for value $\rho^2(\alpha)$. In addition, the author has found the effective iterative procedure with wide fast convergency domain for minimizing the value $\rho^2(\alpha)$ with respect to α.

3. Example 9. On Figures 1 and 2, we represent the identification problem described above in the trivariate case [2,29].

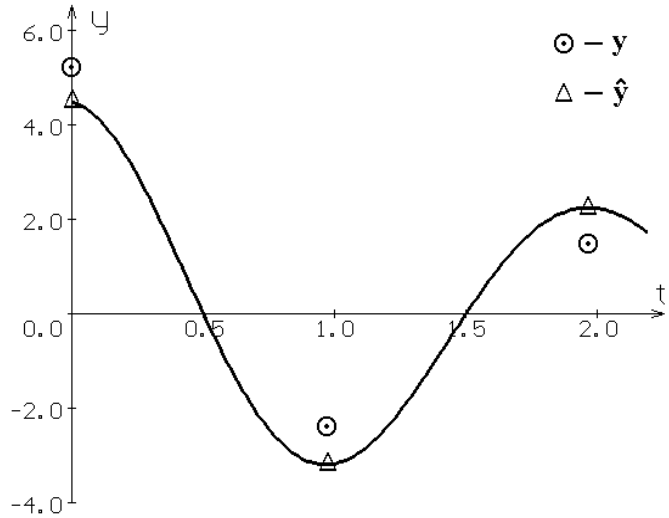

Figure 1: The approximation of the three references by solving the first order difference equation and by solving corresponding differential equation with complex–valued coefficients:

$$\dot{y} + a_0 y = 0, \quad a_0 = \ln\sqrt{2} + \mathbf{i}\cdot\pi.$$

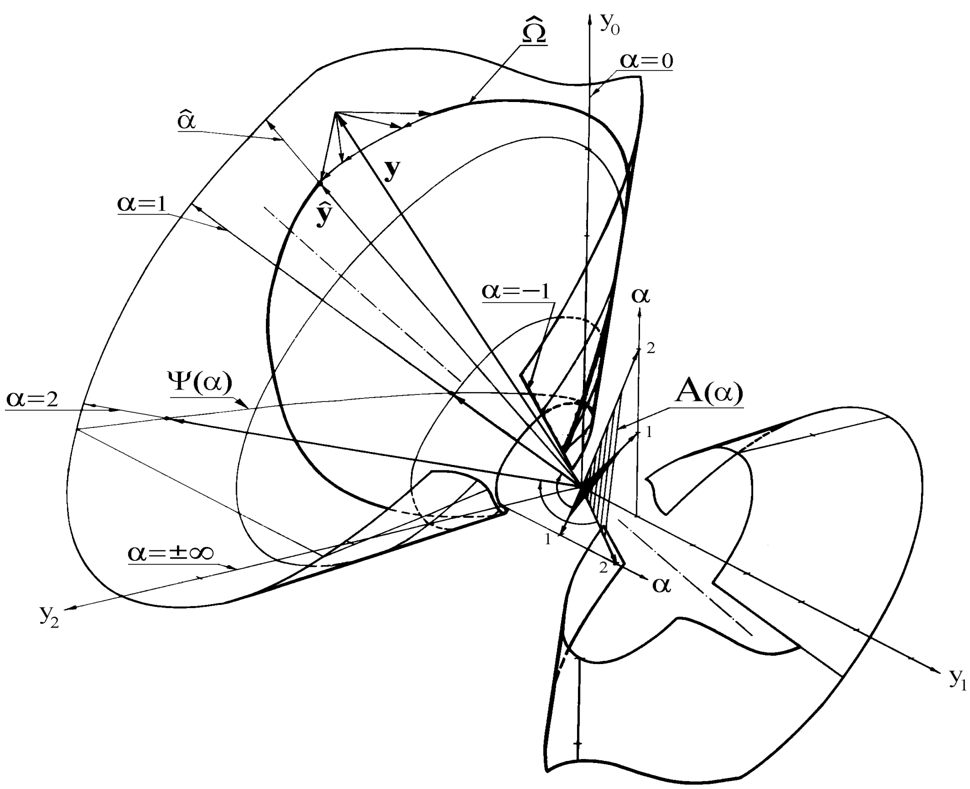

Figure 2: Variational projection in the trivariate case.

Figure 1 illustrates three references composing the trivariate vector $y = [\psi]$. We wish to approximate these three references via a suitable transient process of the first order model. It means that in this example we have $N+n = 3$, $n = 1$. The model equation has the form

$$y_{k+1}+\alpha\, y_k = 0.$$

On Figure 1, the transient process of the corresponding differential equation is represented. It is obtained by the “dediscretization” of the difference equation obtained via the identification method described above.

Figure 2 illustrates the property that the set of the subspaces $M_\alpha = \Psi_\alpha$, which are the kernels of the difference operators

$$D_\alpha = y_{k+1}+\alpha\, y_k = 0, \quad \alpha \in (-\infty, +\infty),$$

in this case is the cone surface denoted on Figure 2 by Ω. The set of all projection points $[\widehat{\varphi}] = \widehat{y}$ is denoted by $\widehat{\Omega}$.

5. Conclusion

One of the results of the theory discussed above is the variational approach to the function approximation consisting of solving of an ordinary linear differential equation with constant coefficients (Examples 6,7,8 of p. 8, section 3; see also [17,29]).

This approach to the approximation is a generalization of a classical problem of finding an approximation for some function y by means of a polynomial $\widehat{y}$ of a degree $(n-1)$, i.e., by the solution to the common stationary differential equation

$$\widehat{y}^{(n)} = 0$$

(see [30]). All the coefficients in this simplest approximating equation are known. It is a possible case when the coefficients of an arbitrary approximating equation are not specified; here we come to the problem of finding the best equation in this class which minimizes the least squares functional. Therein lies the essence of the variational method of identification proposed by the author of this paper in due time [31,32].

Variational approach to identification is highly error resistable both in the original data and in the model structure. Therefore it would be most appropriate to apply it for approximating complex objects by a simplified model of the described type on the short path segment [33]. In addition, due to the “fast” orthogonalization algorithms described here and due to the special iterative optimization procedures for equation parameters [2,17,30] proposed by the author of current paper, this approach can be also used in the real time feedback systems. This is essential for a number of problems of adaptive control with an identifier [29].

6. Acknowledgments

Author is grateful to G.V. Demidenko, G.V. Shevchenko, and A.A. Lomov for their attention to this research and for useful discussions.

References

[1] A.O. Yegorshin, *On one projection problem in Hardy space*, In: International Conference honoring academician Sergei K. Godunov *Mathematics in Applications*, (August 25–28 1999, Novosibirsk), Abstract, Novosibirsk State University, IM SB RAN et al, Novosibirsk, 1999, 159–161.

[2] A.O. Yegorshin, *On one estimation method of modelling equation coefficients for sequences*, Sibirskyi Zhurnal Industrialnoy Matematiki (Siberian Zhournal of Industrial Mathematics), 2(6) (2000), 78–96 (In Russian).

[3] L. Ljung, T. Kailath, and B. Friedlander, *Scattering theory and linear least squares estimation. Part I: Continuous–time problem*, Proc. IEEE, 64(1) (1976), 131– 139.

[4] B. Friedlander, T. Kailath, and L. Ljung, *Scattering theory and linear least squares estimation. Part II: Discrete–time Problems*, J. Franklin Inst., 301(1–2) (1976), 71–82.

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

[6] R. Redheffer, *Difference equations and functional equation in transmission–line theory*, In: E.F. Beckenbach, editor, *Modern Mathematics for the Engineer, chapter 12*, McGraw–Hill, New York, 1961.

[7] V.A. Ambarzumian, *Diffuse reflection of light by foggy medium*, Dokl. Akad. Nauk SSSR, 38(8) (1943), 257–261.

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

[9] E.A. Robinson, *Spectral approach to geophisical inversion by Lorentz, Fourier, and Radon transforms*, Proc. IEEE, 70(9) (1976), 1039–1054.

[10] R.A. Wiggins, E.A. Robinson, *Recursive solution to the multichannel filtering problem*, J. Geophys. Res., 70 (1965), 1885–1891.

[11] A. Lindquist, *Optimal filtering of continuous–time stationary processes by means of the backward innovation process*, SIAM J. Control, 12(4) (1974), 747–754.

[12] A. Lindquist, *A new algorithm for optimal filtering of discrete–time stationary processes*, SIAM J. Control, 12(4) (1974), 736–746.

[13] T. Kailath, *Some new algorithms for reqursive estimation in constant linear system*, IEEE Trans. Inform. Theory, IT–19(6) (1973), 750–760.

[14] T. Kailath, *Some new result and insights in linear least–squares estimation*, In: Proc. First Joint IEEE–USSR Workshop *Information Theory*, (December 1975, Moscow), Moscow, 1975 [rept. with correction in T. Kailath, *Lectures in Linear–Least–Squares Estimation*, Springer–Verlag, Berlin and New York, 1978].

[15] T. Kailath, *Some alternatives in recurcive estimation*, Int. J. Control, 32(2) (1980), 311–328.

[16] A.O. Yegorshin, *Least square method and fast algorithms of identification and filtration (VI Method)*, Avtometriya, 1 (1988), 30–42 (In Russian).

[17] A.O. Yegorshin, *On one variational problem of the mathematical modelling and parametrical identification,* In: Proceeding of the IASTED International Conference *Automation, Control, and Information Technology* (ACIT–2002, Novosibirsk, Russia, June 10–13 2002), ACTA Press, Anaheim–Calgary–Zurich, 2002, 267–272.

[18] N.I. Ahiezer, *The Classical Moment Problem*, Hafner Publishing Company, New York, 1965.

[19] U. Grenander, G. Szegio, *Toeplitz Forms and Their Applications*, Univ. of California Press, Berkeley–California, 1958.

[20] I.C. Gohberg, M.G. Krein, *Theory of Volterra Operators in Hilbert Space*, In: *Transl. of Mathematical Monographs*, Providence, Amer. Math. Soc., Rhode Island, 28 (1970).

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

[22] M.G. Krein, *On integral equations leading to second–order differential equations*, Doklady Akad. Nauk SSSR, 97(1) (1955), 21–24.

[23] M.G. Krein, *The continuous analogues of theorems on polynomials orthogonal on the unit circle*, Dokl. Akad. Nauk SSSR, 105(4) (1955), 637–640.

[24] B. Friedlander, T. Kailath, M. Morf, L. Ljung, *Extended Levinson and Chandrasekhar equations for general discret–time linear estimation problem*, IEEE Trans. Automat. Control, AC–23(4) (1978), 653–659.

[25] T. Kailath, L. Ljung, M. Morf, *Generalized Krein–Levinson equations for efficient calculation of Fredholm resolvents of nondisplacement kernals*, In: Ed. I. Gohberg, and M. Kac, editors, *Topics in Functional Analysis*, Essays Deducated to M.G. Krein on the Occasion of His 70th Birthday, *Advances in Mathematics Supplementary Studies*, Academic Press, New York–San Francisco–London, 3 (1978) 169–184.

[26] B. Friedlander, *Lattice filter for adaptive processing*. Proc. IEEE 70(8) (1982), 829–867.

[27] D. Lee, B. Friedlander, M. Morf, Recursive ladder algorithms for ARMA modeling. In: Proc. 19th IEEE Conf. *Decision and Control*, (Albuquerce, NM, Dec. 1980), Albuquerce, NM, 1980, 1225–1241, [Also IEEE Trans.Automat. Contr., AC–27(Aug.) (1982), 753–754].

[28] A.O. Egorshin (Yegorshin), A.A. Lomov, *Variational identification and filtration via fast algorithms*, In: Prep. 8–IFAC/IFORS Symp. *Identification and System Parameter Estimation*, (Beijing, China, August 27–31 1988), Pergamon Press, Beijing, China and Oxford, New York, 2 (1988), 665–671.

[29] A.O. Yegorshin, *Parameters optimisation of the stationary models in unitary space*, Automatica and Remote Control, 65(12) (2004), 1885–1903.

[30] A.O. Yegorshin, *Variational discretization, and identification of linear stationary differential equations*, In: Proc. III International Conf. System *Identification and Control Problems* (SICPRO'04, January 28–30 2004, Moscow), Institute of Control Sciences, Moscow, 2004, 1824–1883 (In Russian).

[31] A.O. Yegorshin, *Numerical closed methods of linear objects identification*, In: V.M. Alexandrov, editor, *Optimal and Selfadjasting Systems* (Optimal'nye i Samonastraivayushiesia sistemy), IAE SB SSSR, Novosibirsk, 1971, 40–53 (In Russian).

[32] V.P. Budyanov, A.O. Yegorshin, N.P. Philippova, *On solving some problems of dynamic processes analysis based on experimental data with the help computers*, In: A.M. Iskol'dskyi, editor, *Problems of constracting scientific investigation automation* (Voprosy postroeniya sistem avtomatizatshii nauchnyh issledovaniy), IAE SB SSSR, Novosibirsk, 1974, 47–66 (In Russian).

[33] A.O. Egorshin (Yegorshin), *Identification, modelling and adaptation (On optimal identification and modelling problem)*. In: Proc. IV IFAC Symp. *Identification and System Parameter Estimation*, (Tbilisi, September 21–27 1976), North–Holland Publ. Co, Amsterdam–New York–Oxford, 3 (1978) 2143–2154.

Alexey O. Yegorshin
Sobolev Institute of Mathematics,
Koptyug ave., 4,
630090, Novosibirsk, 90, Russia
E-mail address: egorshin@math.nsc.ru