ON NEW MATHEMATICAL RESULTS IN RESEARCHING OF LINEAR TRANSFORMATIVE SYSTEMS

Stoyan V.A., Palamarchuk K.M.
Kyiv National Taras Shevchenko University

The problem of construction and research of solutions’ multitude (if they are) or root-mean-square approximations to them for the system of linear algebraic equations is set and solved. Obtained mathematical results on pseudo-inversion of these classically known algebraic systems are spread on linear integral and functionally transformed systems. The evaluation of accuracy of the obtained in this paper pseudo-inversions’ analytical dependencies of these three practically important discrete and discretely-continuous transformers is conducted. Conditions of uniqueness of pseudo-solutions, constructed in this way, are formulated. The form of the considered problems’ solutions is quite simple and convenient for practical implementation.

Keywords: Pseudo-inversions, linear algebraic systems, linear integral systems, linear functional systems, root-mean-square approximation.

Introduction. It is well known that functioning of many classically known processes and phenomena is described by mathematical models, the dependence of output characteristics from input ones is linear. Such models are constructed without difficulties and are easier to be researched. However, being mathematically correct in relation to the physical and technical merits of specific processes and phenomena, such models can’t be definitely solved or do not have solutions at all. This means, that to set and solve math problems by investigating the status of this process (the phenomenon), by classically known methods of analytic or computational mathematics is impossible. Here you need non-standard approaches to the solution of the problem, especially in cases, when the formulations of mathematical problems are practically oriented.

The given research is based on three classes of transformers, which by linear mathematical model combine discrete-continuous inputs-outputs and are described by systems of linear algebraic, integrated and functional equations. For each of them for previously published [1; 2] authors' scientific achievements will be constructed and researched for accuracy and uniqueness sets of root-mean-square approximations to the exact solution. The peculiarities of the mentioned above approach to pseudo-inverse approximation problems of these three common mathematical models are their simplicity in computer - practical implementation.

1. PSEUDO-INVERSIONS OF LINEAR ALGEBRAIC SYSTEMS

1.1. Optimization definition of pseudo-inverse matrix. Consider linear algebraic system

\[ Cx = y, \]  

(1)

in which \( C \in \mathbb{R}^{m \times n} \) and \( y \in \mathbb{R}^n \) – are known matrix and vector, and \( x \in \mathbb{R}^n \) is searched vector.

There are many approaches to solving system (1), which may have a solution (one or multiple) and may not have. In the study of system (1) we will consider the proposed and developed research methods in M. F. Kryyenko’s works [2; 3]. Thus, let’s construct and research solution of system (1), if it exists. In the absence of the latter, let’s construct the best root-mean-square (single or multiple) approximation to it. Found in this way \( x \) let’s call the general solution of system (1).

For construction of general solution of system (1) we introduce for consideration matrix \( C^* \), pseudo-inverse to \( C \), such as

\[ C^* y = \arg \min_{x \in \mathbb{R}^n} \| x \|, \]  

(2)

where

\[ \Omega_x = \arg \min_{x \in \mathbb{R}^n} \| Cx - y \|. \]  

(3)

Let’s research the properties of vector \( \hat{x} = C^* y \), which we name later pseudo-solution of system (1). For the beginning let’s consider, that matrix \( C^* \) exists and can be constructed singularly.

For this purpose define in matrix \( C \) \( r \) linearly independent columns. Determine via \( C_i = (c_{1,i},\ldots,c_{n,i}) \) matrix, formed by these columns. Taking into consideration the fact, that each column of matrix \( C \) can be decomposed via vectors \( c_{1,i},\ldots,c_{n,i} \), as a basis, matrix \( C \) is represented as follows:

\[ C = (C_1, \ldots, C_r), \]  

where \( a_i \in \mathbb{R} \cdot (i=1, m) \) - is vector of coefficients of decomposition \( i \)-th column of matrix \( C \) for basis \( c_{1,i},\ldots,c_{n,i} \). Thus \( C = C_1 C_2 \), where

\[ C_2 = (a_1, \ldots, a_n). \]  

(4)

Taking into account this, let’s construct solution of problem (1)-(3) in two stages:

1) we solve the problem of finding vector \( \hat{z} \in \mathbb{R}^r \), such as

\[ \hat{z} = \arg \min_{x \in \mathbb{R}^r} \| C_2 x - y \|; \]  

(5)

2) we find the minimum according to the norm of vector \( x \), such as

\[ C_1 x = \hat{z}. \]  

(6)

These problems have a single solution: the first – as problem of decomposition of vector \( y \in \mathbb{R}^n \) according to the system of vectors \( c_{1,i},\ldots,c_{n,i} \);
the second is that for system (6) rank of the major matrix equals to the rank of expanded one.

When solving problem (5) let’s come from the fact, that
\[
|Cz - y|² = |Cz - y|²|Cz - y| = z^T C_1 C z - 2z^T C_1 y + y^T y.
\]
From here
\[
\text{grad}_z |Cz - y|² = 2C_1 C z - 2C_1 y = 0;
\]
\[
\hat{z} = (C_1^T C_1)^{-1} C_1 y.
\] (7)

We get the solution to the second problem by minimizing of Lagrange’s function
\[
\Phi = x^T x + \lambda^T (C z - \hat{z})
\]
Taking into consideration, that for
\[
\Phi = x^T x + \lambda^T (C z - \hat{z}) = 0
\]
under
\[
\text{grad}_z \Phi = 2x + \lambda^T C z = 0
\]
from (6) we have:
\[
C z C z^T \lambda + 2\hat{z} = 0.
\]
From here
\[
\lambda = -2(C z C z^T)^{-1} \hat{z},
\]
and searched
\[
x = C z^T \lambda
\]
Taking into consideration (7), we get:
\[
x = C y,
\]
where
\[
C^* = C z^T (C z C z^T)^{-1} (C z C z^T)^{-1} C^T
\] (8)
is searched, according to (2), (3) pseudo-inverse matrix.

1.2. The singular image of direct and pseudo-inverse rectangular matrices. Let’s come from the fact, that every orthogonal matrix \( C \) dimension \( p \times m \) we can apply as the product of two matrices \( C_1 \) and \( C_2 \) dimensions \( p \times r \) and \( r \times m \) accordingly, where \( r \) is the rank of matrix \( C \). Consequently
\[
C = C_1 C_2 = (c_{i1}; \ldots; c_{in}) \ldots (c_{m1}; \ldots; c_{mr})
\]
where \( c_{ij} \cdot (i = 1, r) \) – linearly independent columns of matrix \( C \) and \( c_{ij}^T \cdot (i = 1, r) \) – rows of matrix \( C^T \), defined in (4). This means, that
\[
C = \sum_{i=1}^{r} c_{ij} c_{ij}^T.
\] (9)

If vectors \( c_{ij} \) and \( c_{ij}^T \) we factor out according to the system of orthonormal vectors \( y \in R^p \cdot (i = 1, r) \) and \( x_0 \in R^m \cdot (i = 1, r) \) accordingly, then presentation (9) will look as:
\[
C = \sum_{j=1}^{m} x_0 j \lambda^j
\] (10)

To determine the properties of the system of orthonormal vectors \( y \), \( x \), numbers \( \lambda \), and their connection with matrix \( C \) let’s consider expression
\[
CC^T y \forall s \in \{1, \ldots, r\},
\]
Taking into consideration orthonormalist of vectors \( x \) and \( y \), we have:
\[
CC^T y_s = \sum_{j=1}^{m} x_j \lambda^j \sum_{i=1}^{r} y_i \lambda^j y_s = \sum_{i=1}^{r} y_i \lambda^j y_s = y_s \lambda^j.
\] (11)

Similarly we find, that
\[
C^T C x_s = \lambda^j x_s \forall s \in \{1, \ldots, r\}.
\] (12)

This means that the system of vectors \( x, y \) and numbers \( \lambda: (i = 1, r) \) exists: vectors \( x \) and \( y \) are proper vectors for matrices \( C^T C \) and \( C C^T \) with their own values, which are equal to \( \lambda^j \). Representation of matrix \( C \) in the form of (10) is possible, though practically it is difficult to construct.

Taking into consideration (10) of matrix \( C \), we construct a similar presentation for matrix \( C^T \). According to (8), where
\[
C = (y_1, \ldots; y_r) \begin{pmatrix} x_1^T \\ \vdots \\ x_r^T \end{pmatrix}
\]
we find:
\[
C^T = \begin{pmatrix} x_1 & \cdots & x_r \end{pmatrix} \begin{pmatrix} y_1^T \\ \vdots \\ y_r^T \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 \cdots \lambda_r \end{pmatrix}
\]
\[
= \begin{pmatrix} x_1 & \cdots & x_r \end{pmatrix} \begin{pmatrix} \lambda_1^{-2} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_r^{-2} \end{pmatrix} \begin{pmatrix} y_1^T \\ \vdots \\ y_r^T \end{pmatrix}^{-1} = \sum_{j=1}^{r} x_j y_j^T \lambda_j^{-1}.
\] (13)

Singualar representation of matrices \( C, C^T \) in the form of (10), (13) allows to prove, that
\[
C^T = (C^T C)^{-1} \cdot (C C^T)^{-1} C^T.
\]

The authenticity of the last representation of matrix \( C^T \) we check, coming from (10), (13).

Here we take into account
\[
CC^T = \sum_{j=1}^{r} x_j \lambda^j \sum_{k=1}^{r} y_k \lambda^j y_s = \sum_{j=1}^{r} y_j \lambda^j y_s
\]
\[
C^T C = \sum_{j=1}^{r} x_j \lambda^j \sum_{k=1}^{r} x_k \lambda^j x_s = \sum_{j=1}^{r} x_j x_j^T \lambda^j
\]
From here
\[
C^T (C^T)^{-1} = \sum_{j=1}^{r} x_j y_j^T \lambda^j \sum_{k=1}^{r} y_k \lambda^j y_s = \sum_{j=1}^{r} x_j y_j^T \lambda^{-1} = C^T
\]
\[
(C C^T)^{-1} = \sum_{j=1}^{r} x_j x_j^T \sum_{k=1}^{r} x_k x_k^T \lambda^{-1} = \sum_{j=1}^{r} x_j x_j^T \lambda^{-1} = C.
\]

1.3. Projection properties of pseudo-inversed matrices. Presentation (10), (13) of matrices \( C \) and \( C^T \) allow in the range within the above mentioned notations, write and illustrate some of the properties of the pseudo-inversed matrices. They are quite interesting in themselves and we need them for construction of defined above general solution of system (1).

To determine these properties we introduce for consideration matrices \( CC^T, C^T C \), \( Z(C) = I_m - C C^T \), \( Z(C)^T = I_p - C^T C \),
\[
Z(C) = I_m - C C^T
\]
System of vectors \( y_1, \ldots, y_m \) we supplement with the system of orthonormal vectors \( y_{r+1}, \ldots, y_m \), orthogonal to \( y_1, \ldots, y_m \).

Then
\[
\sum_{j=1}^{m} y_j y_j^T = I_p,
\]
CC' = \sum_{j=1}^{n} y_j y_j';
\tag{14}
\begin{align*}
Z(C') &= I_p - CC' = \sum_{j=1}^{n} y_j y_j'.
\end{align*}

If consider
CC'y = \sum_{j=1}^{n} y_j y_j'y,
\tag{15}
and
\begin{align*}
(I_p - CC')y &= \sum_{j=1}^{n} y_j y_j'y,
\end{align*}

where y \in R^n, then y'y is nothing else than the projection of vector y on orthogonal complement to linear shell, stretched on vector-columns of matrix C.

For vector y \in R^n and (p=\infty) - dimensional matrix C, such that for orthonormal vectors x_j', x_j, which are basic for a vector of rows c_j', c_j of matrix C and for which run ratio (11), (12), matrices C', Z(C) and Z(C') = I_p - C'C are projection on a linear shell, stretched on a vector-columns of matrix and orthogonal complement to this shell correspondingly.

1.4. General solution of system of linear algebraic equations. Projection properties of matrices CC', C'C, Z(C) and Z(C') allow to construct and research the general solution of system of linear algebraic equations (1).

For the construction of the latter we come from the fact, that according to the considered above, the solution of equation (1), if it exists, is written via pseudo-inverse of matrix C' by ratio
\begin{align*}
x = C' y.
\end{align*}

In the general case
\begin{align*}
x = C' y + w,
\end{align*}

where w is a random vector of dimension m, such as
\begin{align*}
Cw = 0.
\end{align*}

The latter means, that vector w must be an orthogonal to vector of rows of matrix C, i.e. w must belong to the orthogonal compliment to linear shell, stretched on vector-columns of matrix C, and therefore
\begin{align*}
w = Z(C)w \quad \forall v \in R^m,
\end{align*}

correspondingly
\begin{align*}
x = C' y + Z(C)w
\end{align*}
\tag{17}

This means, that in general case, if we take into account that det(C'C) > 0 is the condition of non-degeneracy of matrix C and
\begin{align*}
\|Z(C)y\|^2 = y'Z(C)y = 0
\end{align*}
is a condition of equality to zero of projection vector y on orthogonal complement to linear shell, stretched on vector-columns of matrix C, which describes the condition under which vector y can be spread according by vector-columns of matrix C, i.e. the condition, when equation (1) is exactly solved, then it is possible to conclude, that solution (17) of system (1) will be:

1) single and accurate, when
\begin{align*}
y'Z(C)y = 0;
\end{align*}
\begin{align*}
det(C'C) > 0;
\end{align*}

2) give a set of solutions under
\begin{align*}
y'Z(C)y = 0;
\end{align*}
\begin{align*}
det(C'C) = 0;
\end{align*}

3) single pseudo-solution, such that
\begin{align*}
e^2 = \min_{\forall \in \Omega} \|Cx - y\|^2 = y'Z(C'y
\end{align*}

when
\begin{align*}
y'Z(C)y > 0;
\end{align*}
\begin{align*}
det(C'C) > 0;
\end{align*}

4) give a set
\begin{align*}
\Omega = \text{Arg} \min_{\forall \in \Omega} \|Cx - y\|^2
\end{align*}
of pseudo-solutions with inaccurate \( e^2 \), defined by ratio (20), when
\begin{align*}
y'Z(C)y > 0;
\end{align*}
\begin{align*}
det(C'C) = 0.
\end{align*}

2. PSEUDO-INVERSIONS OF LINEAR INTEGRAL SYSTEMS.

Let’s consider the question of constructing a general solution of a linear integral system
\begin{align*}
\int_0^1 A(t)x(t)dt = y;
\end{align*}
\begin{align*}
\tag{21}
\end{align*}

under the known matrix function \( A(t) \in R^{m \times n} \) \( \forall \in [0,T] \) and vector \( y \in R^p \).

For construction of vector-function
\begin{align*}
x(t) = \text{arg} \min_{\forall \in \Omega} \| A(t)x(t)dt - y \|^2,
\end{align*}
\begin{align*}
\tag{22}
\end{align*}

let’s solve for the beginning the problem of constructing values \( x(t) \ (i = 1, N) \) of this vector, provided that \( x_i \ (i = 1, N) - \) points of the interval’s \([0,T]\) discretization. Under these conditions, system (21) will be written in the form
\begin{align*}
\sum_{i=1}^{N} A(t_i)x(t_i)\Delta t_i = y,
\tag{23}
\end{align*}

where \( \Delta t_i \) is a step of discretization of interval \([0,T]\). The problem of finding values \( x(t) \ (i = 1, N) \) comes to root-mean-square inversion of the system
\begin{align*}
\overline{A}(t)\overline{x}(t) = y,
\tag{24}
\end{align*}
in which \( \overline{A}(t), \overline{x}(t) \) are matrix and vector functions of discrete argument \( i = 1, N \) are such, that
\begin{align*}
\overline{A}(t) = (A(t_1), ..., A(t_n))\sqrt{\Delta t_n},
\tag{25}
\end{align*}
\begin{align*}
\overline{x}(i) = \text{col}(x(t_1), ..., x(t_n))\sqrt{\Delta t_n},
\end{align*}
\begin{align*}
\overline{A}(t) = A(t)\sqrt{\Delta t_n},
\tag{26}
\end{align*}
\begin{align*}
\overline{x}(i) = x(t)\sqrt{\Delta t_n}.
\end{align*}
According to (17) solution of (24) is such that
\[ |A(t)x(t) - y|^2 \rightarrow \min, \]
with
\[ x(t) = A^*(t)P^*y + v(t) - A^*(t)P^*p, \]
where
\[ p = A(t)x(t), \]
when
\[ v(t) = v(t_1), \ldots, v(t_N), \]
and random \( v(t) \in R^m \cdot (t \in [0, T]) \) and vector \( x(t) \).

With regard to definition (25) of matrix and vector functions \( A(t) \) and \( x(t) \) from (26) we find
\[ x(t) = A^*(t)P^*y + v(t) - A^*(t)P^*p, \]
where now
\[ P = \sum_{i=1}^{N} A(t_i)A^*(t_i)\Delta t, \]
\[ p = \sum_{i=1}^{N} A(t_i)v(t_i)\Delta t. \]

Solutions (26) and (27) will be accurate regarding to (24) and (23), if (see, (18), (19))
\[ \varepsilon^2 = y^2 - y^2P^*y = 0. \]

When \( \varepsilon^2 > 0 \) ratios (26) and (27) will be determined the best according to the root-mean-square criterion of approximation to the solution. According to (18), (21) these solutions will be unique (\( v(t) = 0, v(t) = 0 \)), if
\[ \det(A(t)A(t)) = \det(A(t)A(t)|_{i=1}^{N}) > 0. \]

With ratio (7), when \( N \rightarrow \infty \) we find vector-function
\[ x(t) = A^*(t)P^*y + v(t) - A^*(t)P^*p, \]
where at random integrated on \([0, T]\) vector-function \( v(t) \in R^{m} \)
\[ P = \int_{0}^{T} A(t)A^*(t)dt, \quad p = \int_{0}^{T} A(t)v(t)dt. \]

This vector-function, according to (22) satisfies (21). When this
\[ \min_{x \in \Omega} \int_{0}^{T} |A(t)x(t) - y|^2 dt = \varepsilon^2, \]
and \( v(t) = 0 \), if
\[ \lim_{N \rightarrow \infty} |A(t)A(t)|_{i=1}^{N} > 0. \]

3. PSEUDO-INVERSION OF LINEAR FUNCTIONAL SYSTEMS.

Let’s consider the question of construction of general solution of a linear functional system of the form
\[ B(t)x = y(t) \cdot (t \in [0, T]) \]
when matrix and vector functions \( B(t) \in R^{m \times n} \) and \( y(t) \in R^{m} \) \( v(t) \in R^{n} \) are known.

For construction of vector
\[ x = \arg \min_{x \in \Omega} \int_{0}^{T} |B(t)x - y(t)|^2 dt \rightarrow \min \]
let’s consider for the beginning the problem of finding a set
\[ \Omega = \left\{ x \in R^m : \sum_{i=1}^{N} |B(t)x - y(t)|^2 \rightarrow \min_x \right\}, \]
which is the set of solutions (or root-mean-square approximations to them) for system (31), defined by points \( t, i = [1, N] \), selected on the interval \([0, T]\) with a step \( \Delta t. \)

Let’s consider from, that problem (33) is equivalent to the problem of root-mean-square system’s inversion
\[ B(t)x = \overline{y(t)}, \]
in which
\[ B(t) = \text{col}(B(t), i = [1, N]) \sqrt{\Delta t}, \]
\[ \overline{y(t)} = \text{col}(y(t), i = [1, N]) \sqrt{\Delta t}. \]

and matrix and vector functions of discrete argument \( i \) are such, that
\[ \overline{B(i)} = B(t) \sqrt{\Delta t}, \]
\[ \overline{y(t)} = y(t) \sqrt{\Delta t}. \]

According to (17), solution of (34) is such, that
\[ \|B(t)x - \overline{y(t)}\| \rightarrow \min_x, \]
will be
\[ x = P^*B_y + v - P^*P v \]
at random \( v \in R^{n} \) (if \( P^*P = 0 \)) and \( v = 0 \) (if \( P = 0 \)) and
\[ P = B^*(t)\overline{B(t)} = \sum_{i=1}^{N} B^*(t_i)B(t_i)\Delta t, \]
\[ B_y = B^*(t)\overline{y(t)} = \sum_{i=1}^{N} B^*(t_i)y(t_i)\Delta t. \]

Solution (38) will be accurate regarding to (34) and (33), if (see, (18), (19))
\[ \varepsilon^2 = y^2 - B^*P^*B_y = 0 \]
when
\[ y^2 = \sum_{i=1}^{N} y(t_i)y(t_i)\Delta t. \]

For the other conditions by ratio (37) is defined the best root-mean-square approximate to solution (34) such, that
\[ \min_{x \in \Omega} \int_{0}^{T} |B(t)x - y|^2 dt = \varepsilon^2. \]

Pointing \( N \rightarrow \infty \) from the system (34) we return to system (31), and from problem (37) to (32). It is easy to see, that the solution of the latter in this case is defined by ratio (38), in which now
\[ P = \int_{0}^{T} B^*(t)B(t)dt, \quad B_y = \int_{0}^{T} B^*(t)y(t)dt. \]

The accuracy and uniqueness of solution (38), as before, will be defined by values \( \varepsilon^2 \) and \( P \) and \( P \), written with consideration (40), when
\[ y^2 = \int_{0}^{T} y(t)y(t)dt. \]

Conclusions. Consequently, set and successfully resolved are complex mathematical problems of pseudo-inversion of linearly transformatory systems with vector inputs-outputs. These which systems: 1) discretely defined input signal turn into both static and dynamic output; 2) dynamic vector input integrate into static output.

Mathematical researches conducted in this paper permit according to the known output signal to renew input vector (vector-function), if transformer’s structure allows doing it.
Regardless of the transformer’s structure and its multiply-combining properties, in this paper’s representation of pseudo-inversions sets are constructed, considering each of transformers, which is accomplished by the convenient checking conditions of accuracy and uniqueness of the latter. It’s quite general and successfully used in engineering-constructive projects.

References: