МЕТОДИ КЕРУВАННЯ ТА ОЦІНЮВАННЯ В УМОВАХ НЕВИЗНАЧЕНОСТІ

UDC 681.5.015; 519.233

V. Gubarev, S. Melnychuk, N. Salnikov*

PARADIGM OF NONSTOCHASTIC APPROACH TO SYSTEM IDENTIFICATION

Vyacheslav Gubarev

Space Research Institute of NASU and SSAU, Kyiv,
orcid: 0000-0001-6284-1866

v.f.gubarev@gmail.com

Serhii Melnychuk

Space Research Institute of NASU and SSAU, Kyiv,
orcid: 0000-0002-6027-7613

serg.vik@ukr.net

Nikolay Salnikov

Space Research Institute of NASU and SSAU, Kyiv,
orcid: 0000-0001-9810-0963

salnikov.nikolai@gmail.com

The concept of a complex system in this work is understood as a large set of dynamic interconnected systems, the exact mathematical model of which is not known or has a very large dimension. In such situation the use of standard methods for synthesizing feedback becomes difficult or even impossible due to the degeneracy of the corresponding mathematical problems. One way out of this situation is to build an approximation model of reduced dimension. This can be done using a system of initial equations, if they are available, or using identification methods based on measurements of output and input variables acting on the system. In this case, the process of constructing a mathematical model is reduced to a sequential enumeration of possible models of increasing complexity. As a criterion for the adequacy of the model, the norm of deviation of the output of the adjusted model from the measured value of the output of the system under study is considered. The article deals with the construction of linear models, the complexity of which is determined by their dimension. In the framework of nonstochastic approach it is developed the methodological and mathematical basis for model reconstruction which describes processes in complex systems. Asymptotic modelling allows for such system to form model classes appropriate to solve identification problem. Precise description corresponds to infinite expansion so the model quality is improved when its dimension is in-

* S. Melnychuk is grateful to the National Research Fund of Ukraine for support (project 2020.02/0015 «Theoretical and experimental research of global disturbances of natural and man-made origin in the Earth-Atmosphere-Ionosphere system»).

creased. However errors in available data do not allow increase their dimension limitlessly due to ill-conditionality of the identification problem beginning from some dimension. Regularization procedure permits to determine the effective approximate solution of identification problem which for nonstochastic case is in agreement with errors in data. Properties and peculiarities of the proposed approach are illustrated by simulation results.

Keywords: system identification, linear regression, regularization, asymptotic modelling, approximate solution

Introduction

System identification, that is, the construction of a mathematical model of a system based on data obtained as a result of experiments, is primarily focused on complex dynamic processes for which it is impossible to establish the laws of their behaviour in other ways. This means that even in a general form it is impossible to indicate a class of mathematical models containing an exact description of the system under study.

Here we consider a new concept or paradigm of system identification within a nonstochastic approach for linear time invariant (LTI) systems. We shell assume the existence of abstract transfer function matrix $G(z)$ including nonrational cases, which connects the input and the measured output. According to the concept of asymptotic modelling, widely used in computational mathematics, an unknown transfer function can be written as finite or infinite expansions for some basis functions, which makes it possible to write a class of models in a form convenient for identification.

At such approach, we are talking about the construction of finite-dimensional approximating models. At the same time, with an increase in the dimension, the accuracy of the description of the system increases if certain requirements for the system under study are met. For example, truncated rational approximation of infinite-dimensional LTI system with finite-dimensional input and output that induced nuclear type Hankel operator with distinct singular values convergences to precise model as it dimension tends to infinity [1].

Asymptotically stable system can be written as an infinity expansion

$$G(z) = \sum_{k=0}^{\infty} G_k f_k(z) \tag{1}$$

where $\{f_k(z)\}_{k=0,1,2,\ldots}$ is a sequence of orthonormal functions and $\{G_k\}_{k=0,1,2,\ldots}$ is a sequence of parameters. Then task of identification is to find a finite-dimensional model

$$\tilde{G}(z) = \sum_{k=0}^{n-1} G_k f_k(z). \tag{2}$$

The accuracy of the model depend on the choice of basis functions and dimension *n.* The case of $f_k = z^{-k}$ corresponds to Infinite Impulse Response (IIR) modeling [2]. In some cases the use of Laguerre or Kautz polynomials gives a better approximation [3, 4]. In general, the advantage in identification and approximation can be achieved by using such orthonormal basis functions that correspond to dynamics close to the dynamics of the object under study [5].

A very popular form for rational approximation is the Linear Regression which directly links input and output variables [2]. Here, for ease of paradigm presentation, we consider an ARX model with a scalar input $u(t)$ and output $y(t)$

$$y(t) + a_1 y(t-1) + \ldots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \ldots + b_{n_b} u(t-n_b) \tag{3}$$

in discrete time $t = 0, 1, \ldots$. Let us denote vectors $\theta_a = [a_1, \ldots, a_{n_a}]^{\mathrm{T}}$, $\theta_b = [b_1, \ldots, b_{n_b}]^{\mathrm{T}}$, $\varphi_a(t) = [-y(t-1), \ldots, -y(t-n_a)]^{\mathrm{T}}$ and $\varphi_b(t) = [u(t-1), \ldots, u(t-n_b)]^{\mathrm{T}}$, so (3) can be written as follows

$$y(t) = \varphi_a^{\mathrm{T}}(t) \cdot \theta_a + \varphi_b^{\mathrm{T}}(t) \cdot \theta_b. \tag{4}$$

The special case $n_a = 0$ gives us Finite Impulse Response (FIR) approximate model. Any ARX model has an equivalent state-space model representation, so it is easy to pass from one description to another using linear transformation. It is important because the truncated state-space models are approximating for nuclear type nonrational system [1].

1. Paradigm of stochastic identification

Let us briefly consider identification paradigm in stochastic case according to [6]. Mathematical difficulties in system identification are associated with the presence of uncertainty in the data. The generally accepted is the stochastic interpretation of uncertainty, which assumes that measurement errors are independent and identically distributed (i.i.d.) random variables. This paradigm underlies all classical statistical methods in system identification. Model parameters of a given order are estimated by maximum likelihood methods (prediction errors), which in most cases are formulated as extremal problems

$$\hat{\theta} = \arg\min_{\theta \in D} \sum_{t=1}^{N} \| y(t) - \hat{y}(t/\theta) \|^2, \tag{5}$$

where $y(t)$ is a measured output, $\hat{y}(t/\theta)$ is a prediction by the model, D is a set of values θ. There are many publications devoted to the identification problem in such statement, see, for example [2, 7, 8].

Under concept of asymptotic modeling the key question in identification is a model order determination. It is known that a higher order model can better approximate the measured output of the system, i.e. reduce the error called «bias». On the other hand, a higher-order model is more sensitive to errors in the data, which corresponds to a larger variance in model parameter estimates. The mismatch between system and model includes both of these components. Traditional system identification often uses the bias-variance trade-off to minimize the total mean square error (MSE). Various procedures have been proposed for this, among which are the Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC) order criteria described in the books mentioned above.

2. Nonstochastic approach to EIV-identification

In many cases on practice, errors occur at output and input variables measurement. Such situation in system identification is called «errors in variables» (EIV) and was considered by many authors. Results obtained in the framework of stochastic paradigm are described in [9–13]. Recently EIV identification in frequency domain was considered in [14]. New results also presented in [15–17].

This article develops an alternative approach based on the non-stochastic paradigm. We assume that data errors are random variables belonging to known bounded sets with arbitrary distributions, including worst cases. With this formulation, the maximum likelihood method is not appropriate. Early the similar approach was implemented in so-called set membership identification (SMI) in which the main goal was to determine a guaranteed set of models, including the exact one. The description of these methods is given in [18–21].

The approach developing in this article is an alternative to SMI and aims to find a single approximate solution of identification problem consistent with errors in available data.

We have the following measurements of true values $u(t)$ and $y(t)$

$$\tilde{u}(t) = u(t) + \xi_u(t), \quad \tilde{y}(t) = y(t) + \xi_y(t), \tag{6}$$

where $\xi_u(t)$ and $\xi_y(t)$ are errors. According to the mentioned assumption $\xi_u(t)$ and $\xi_y(t)$ are unknown arbitrary random sequences satisfying conditions

$$\|\xi_u(t)\| \le \varepsilon_u, \ \ \|\xi_y(t)\| \le \varepsilon_y. \tag{7}$$

An important issue in identification is the informative input that excites all modes of the system. In stochastic methods, a persistent excitation signal of a given order is usually used for this. In the case of non-stochastic identification, we propose an active experiment in which each mode of the system at a certain moment of observation makes the maximum possible contribution to the output. For an asymptotically stable system, this can be achieved by two ways. The first is to collect data from separate experiments consisting of the intervals of excitation and following relaxation. An alternative is a single continuous experiment where excitation intervals alternate with relaxations. In both cases, the duration of relaxation should exceed the transient time.

Consider an observable and controllable discrete LTI system. Such a system can be represented by linear regression or its state-space equivalent

$$x(t+1) = Ax(t) + bu(t), \ \ y(t) = c^{\mathrm{T}} x(t) + du(t) \tag{8}$$

where t is a discrete time and «T» denotes a transposition. Assume that this system does not have multiple eigenvalues. Then we can choose the Jordan block realization: $A = \mathrm{diag}(A_p), \ \ b = \mathrm{col}(b_p), \ \ c = \mathrm{col}(c_p),$ where real eigenvalues $\lambda_p = \alpha_p$ correspond to blocks

$$A_p = \alpha_p, \ \ b_p = b_p^c, \ \ c_p = c_p^c, \tag{9}$$

and complex eigenvalues $\lambda_p = \alpha_p \pm i\beta_p$ correspond to blocks

$$A_p = \begin{pmatrix} \alpha_p & -\beta_p \\ \beta_p & \alpha_p \end{pmatrix}, \ b_p = \begin{pmatrix} b_p^c \\ b_p^s \end{pmatrix}, \ c_p = \begin{pmatrix} c_p^c \\ c_p^s \end{pmatrix}. \tag{10}$$

In case of finite-dimensional system with $y(t) = 0$ at $t \le 0$, $u(t) = 0$ at $t < 0$ and $u(t) \ne 0$ at $t \ge 0$ input-output ration (8)–(10) is equivalent to

$$y(t) = \sum_{j=0}^{t-1} \sum_{p=1}^{P} h_{0p}(t-j) \cdot u(j), \tag{11}$$

where $h_{0p}(k) = \rho_p^k [f_p^c \cos k\omega_p + f_p^s \sin k\omega_p], \quad \rho_p = |\lambda_p|, \quad \omega_p = \arg\lambda_p, \quad f_p^c = c_p^c b_p^c + c_p^s b_p^s, \ \ f_p^s = c_p^c b_p^s - c_p^s b_p^c.$ Parameters $\rho_p, \ \omega_p, \ f_p^c, \ f_p^s$ completely determine the system dynamics. In case of real eigenvalues we have $\rho_p = \alpha_p, \ \ \omega_p \equiv 0,$ $c_p^s = b_p^s = 0.$

In EIV identification the informativeness of the input is defined by the signal-to-noise ratio (SNR). The larger SNR of the output for each mode (10), the more informative the input is. Therefore, we will shape the input so that at the end of the excitation interval, a certain mode has the largest output signal. Different $u(t)$ across experiments should effectively excite each of the modes. Two types of input signal will be used for this. The first is a rectangular pulse

$$u(t) = \begin{cases} 1, & t = \overrightarrow{1, \tau_i} \\ 0, & \text{else} \end{cases} \tag{12}$$

where τ_i is of a varied duration across experiments: $\tau_i = i$, $i = 1, 2, \ldots, i_{\max}$. Short rectangular pulses provide stronger excitation of fast modes. Increasing the duration will add a contribution from slow modes. Therefore a set of rectangular pulses of different duration should provide separate excitation of individual modes corresponding to real eigenvalues (9).

The second type of input is a single harmonic

$$u(t) = \begin{cases} \sin \omega_j t, & t = \overrightarrow{1, \tau} \\ 0, & \text{else} \end{cases} \tag{13}$$

where frequency ω_j change across experiments $\omega_j = \Delta \cdot j$, $j = 1, 2, \ldots, j_{\max}$ and cover the range $(0, \pi/2)$. For modes corresponding to complex eigenvalues (10), we aim to catch resonant excitation when ω_j happens to be close to the natural frequency ω_p. When ω_j are varied with a small step, it is possible to provide resonant excitation of all oscillating modes of the system. In this case the duration τ is chosen to be sufficient to establish steady-state forced oscillations. With an appropriate choice of parameters $i_{\max}$, $j_{\max}$, Δ and τ it is possible to collect informative data with acceptable SNR with a fairly general knowledge of the system under study. More detailed information on this can be found in [22].

3. Method of identification

Here will be developed the identification method in the framework of nonstohastic approach. The goal is to obtaine a regularized solution that gives an approximation with an accuracy corresponding to the noise in data.

As mentioned, the asymptotic class of linear autoregressive models is considered to identify complex systems that in this class have a large or infinite dimension. Conception of approximate regularized solution means that the model dimension may be less then the order of the true system. Therefore, identification must include the choice of model order.

3.1. Model order selection. The type of control signals proposed above makes it possible to divide the identification process into independent tasks. Measurements on relaxation intervals are used to determine the vector a and its dimension. For this we will use the equation

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) - \ldots - a_n y(t-n), \tag{14}$$

which describe free motion on relaxation intervals. Applying controls (12) and (13), we collect the corresponding outputs into the following matrix:

$$\tilde{Y} = \begin{bmatrix} \tilde{Y}_{\text{relax}}^{\text{pulse}} \\ \tilde{Y}_{\text{relax}}^{\text{harm}} \end{bmatrix},$$

$$\tilde{Y}_{\text{relax}}^{\text{pulse}} = \text{col}([\tilde{y}_i^{\text{pulse}}(\tau_i + 1) \ \ldots \ \tilde{y}_i^{\text{pulse}}(\tau_i + l)), \ i = \overrightarrow{1, k_1}, \tag{15}$$

$$\tilde{Y}_{\text{relax}}^{\text{harm}} = \text{col}([\tilde{y}_j^{\text{harm}}(\tau + 1) \ \ldots \ \tilde{y}_j^{\text{harm}}(\tau + l)]), \ j = \overrightarrow{1, k_2},$$

where $\tilde{y}_i^{\text{pulse}}$ is a measured response on the i-th rectangular pulse (12), k_1 is a total number of rectangular pulses, $\tilde{y}_j^{\text{harm}}$ is a measured response on the j-th harmonic in-

put (13), k_2 is a total number of input harmonics and l is the relaxation duration, the same for all experiments. We denote the total number of rows of $\tilde{Y}$ by $k = k_1 + k_2$. Value l should be more then transient time and both k and l must exceed supposed model order.

Let us produce singular value decomposition (SVD) of the matrix $\tilde{Y}$

$$\tilde{Y} = U\Sigma V^{\mathrm{T}}, \tag{16}$$

where U, Σ, V are matrices of dimensions $k \times k$, $k \times l$ and $l \times l$ respectively.

In the deterministic case for a finite dimensional system the rank of $\tilde{Y}$ is equal to true dimension n and only the first n singular numbers $\sigma_1, \ldots, \sigma_n$ are non-zero. Then the dimension of the system can be precisely determined by the number of non-zero diagonal elements of Σ.

In the non-deterministic case the matrix $\tilde{Y}$ almost always will be of full rank and all singular numbers will be positive. Even then the true dimension n can sometimes be determined by singular values, if there is a gap between the values σ_n and σ_{n+1}. However there is often no clear gap. Result depends on the SNR of individual modes. Modes of the system with a low SNR becomes indistinguishable from background noise, making impossible in the non-stochastic case to establish the true dimension. In such cases, we can find a model of reduced dimension applying the known in regularization principle that approximate solution should be consistent in accuracy with the data errors.

For any chosen value $\hat{n}$ we can divide the measured output into «signal» and «noise» part using a low-rank approximation in Frobenius norm. For this we split (16)

$$U\Sigma V^{\mathrm{T}} = [U^S \; U^N] \begin{bmatrix} \Sigma^S & 0 \\ 0 & \Sigma^N \end{bmatrix} [V^S \; V^N]^{\mathrm{T}}, \tag{17}$$

where Σ^S contains $\hat{n}$ largest singular values. Indices «*S*» and «*N*» denote «signal» and «noise» part. As a result we get decomposition

$$\tilde{Y} = Y^S_{\hat{n}} + Y^N_{\hat{n}}, \tag{18}$$

where $Y^S_{\hat{n}} = U^S \Sigma^S (V^S)^{\mathrm{T}}$ and $Y^N_{\hat{n}} = U^N \Sigma^N (V^N)^{\mathrm{T}}$. Expansion (18) allows one to find the matrix of the given rank closest to full-rank matrix $\tilde{Y}$ in Frobenius norm using ε-rank property [23]. So we have $\operatorname{rank}(Y^S_{\hat{n}}) = \hat{n}$ and

$$\sigma_{\hat{n}+1} = \left\| Y^N_{\hat{n}} \right\|_F, \tag{19}$$

where subscript «*F*» means Frobenius norm, defined as $\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}$ for any matrix $A = (a_{ij})$. Matrix $Y^S_{\hat{n}}$ will be considered as the output of the desired model.

Again we represent the matrix $\tilde{Y}$ in the form

$$\tilde{Y} = Y + \Delta Y, \tag{20}$$

where Y corresponds to noise-free output generated by true system, and ΔY corresponds to the noise, which, according to condition (7), is bounded in the infinity norm

$$\left\| \Delta Y \right\|_\infty = \varepsilon_y. \tag{21}$$

Since all available information about Y is limited only by condition (21), then, according to the guaranteed approach, any Y' that satisfies the inequality $\left\| Y' - \tilde{Y} \right\|_\infty \le \varepsilon$

can be considered as corresponding to a noise-free output of the exact model. Among all exact models, we are looking for a model of minimal order, which corresponds to the minimal rank of the matrix Y'. Using decomposition (18) we determine the model by its output as $Y_{\hat{n}}^S$, increasing the dimension $\hat{n}$ until condition $\left\| Y_{\hat{n}}^S - \tilde{Y} \right\|_\infty \le \varepsilon_y$ is satisfied. From here comes the criterion

$$\left\| Y_{\hat{n}}^N \right\|_\infty \le \varepsilon_y < \left\| Y_{\hat{n}-1}^N \right\|_\infty . \tag{22}$$

Using equivalence of norms $\|A\|_\infty \le \|A\|_F \le \|A\|_\infty \sqrt{k \cdot l}$, which is valid for an arbitrary matrix $A \in R^{k \times l}$, from (22) and (19) we derive inequality

$$\frac{\sigma_{\hat{n}+1}}{\sqrt{kl}} \le \varepsilon_y < \sigma_{\hat{n}},$$

or

$$\begin{cases} \sigma_{\hat{n}} > \varepsilon_y, \\ \sigma_{\hat{n}+1} \le \varepsilon_y \sqrt{kl}. \end{cases} \tag{23}$$

The condition (23) can be used to determine the dimension of the model that is consistent in accuracy with the errors.

3.2. Parametric identification. After determining the order $n_a = \hat{n}$ of vector a, it can be found from the basic equations (3), (4), where $\varphi_b(t) = u \equiv 0$. For this we construct an overdetermined linear system

$$W_A \cdot \theta_a = w_A, \tag{24}$$

where the matrix W_A and vector w_A are formed from the matrix $Y_{n_a}^S$ as follows:

$$W_A = \begin{pmatrix} W_A^{(1)} \\ W_A^{(2)} \\ \vdots \\ W_A^{(l-n_a)} \end{pmatrix}, \; w_A = \begin{pmatrix} w_A^{(1)} \\ w_A^{(2)} \\ \vdots \\ w_A^{(l-n_a)} \end{pmatrix}.$$

Here, $W_A^{(1)}$ contains the first n_a columns of the matrix $Y_{n_a}^S$, the matrix $W_A^{(2)}$ also contains n_a columns, starting from the second, and so on. The vector $w_A^{(1)}$ is the (n_a+1)–th column of $Y_{n_a}^S$, $w_A^{(2)}$ is the (n_a+2)–th column, etc. The last one vector $w_A^{(l-n_a)}$ is the first column of the matrix $Y_{n_a}^N$.

System (24) is a strongly overdetermined for large r and l. Then we can discard non-informative equations with small SNR, i.e. having $y_{\max} / \varepsilon_y$ less than some threshold, where $y_{\max}$ is the maximum modulus of regressors. As a result, we obtain a truncated but still overdetermined system

$$\bar{W}_A \cdot \theta_a = \bar{w}_A. \tag{25}$$

The solution of (25) can be found by ordinary least squares (OLS) or by total least squares (TLS) [23]. However, in practice, very often problem (25) turns out to be ill-posed, having an ill-conditioned matrix $\bar{W}_A$. Therefore, it is proposed to implement the LS based on the SVD decomposition of the matrix $\bar{Y}_{n_a}$

$$\bar{W}_A = U_1 \Sigma_1 V_1^{\mathrm{T}}. \tag{26}$$

 ISSN 2786-6491

Substituting (26) into (25), we obtain

$$\Sigma_1 \cdot \theta_1 = \bar{\bar{w}}_A, \tag{27}$$

where $\theta_1 = V_1^{\mathrm{T}} \theta_a$, $\bar{\bar{w}}_A = U_1^{\mathrm{T}} \cdot \bar{w}_A$. The ratio of the first singular value of the matrix Σ_1 to the last one determines the conditionality of the problem. Under good conditionality, the least squares solution of (26) is

$$\theta_a = V_1 \Sigma_1^{-1} U_1^{\mathrm{T}} \bar{w}_A. \tag{28}$$

When the matrix $\bar{W}_A$ is ill-conditioned, one should find a regularized solution, the construction of which is given below.

At the last stage of identification, the dimension n_b and coefficients of the vector b are determined. It is preceded by extracting data corresponding to a purely forced part of model. With an estimated vector a, such a signal $\tilde{\tilde{y}}(t)$ for the moment t is determined by the relation

$$\tilde{\tilde{y}}(t) = \tilde{y}(t) + a_1 \tilde{y}(t-1) + a_2 \tilde{y}(t-2) + \ldots + a_n \tilde{y}(t-n_a). \tag{29}$$

Then the basic equation for finding the vector b is

$$\tilde{\tilde{y}}(t) = b_1 u(t-1) + \ldots + b_m u(t-n_b), \tag{30}$$

where dimension n_b of the vector b in the main not exceed dimension of vector a. In principle, case $n_b > n_a$ is possible if there is a static connection between input and output, caused by very fast modes having a transient process shorter than the sampling step.

After obtaining $\tilde{\tilde{y}}(t)$ from (29), we select samples corresponding to excitation intervals defined by (12) and (13) and form an overdetermined system

$$W_B \cdot \theta_b = w_B, \tag{31}$$

similar to the construction of W_A and using $n_b = n_a$. Elements of the matrix W_B are measured control signals $u(t)$. The chosen method of excitation (12), (13) provides good conditionality of the matrix W_B, so system (31) can be solved by TLS.

This completes the solution of the identification problem. As a result we obtain an approximate mathematical model, consistent in accuracy with the errors in the data determined by conditions (6), (7).

4. Ill-conditioning

The solvability of the identification problem in the formulation under consideration is determined by the properties of the matrices W_A in (24) and W_B in (31), and, first of all, by their condition number. It was proposed to solve (24) and (31) using instead of OLS, a completely equivalent method based on the SVD decomposition of these matrices. Then the condition number of these matrices with respect to the norm $\|\cdot\|_2$ can be found from the relation

$$\kappa = \frac{\sigma_1}{\sigma_n}, \tag{32}$$

where $n = n_a, n_b$ correspond to the dimensions of the allocated blocks when choosing the order of the model. Due to the properties of the SVD decomposition, as the

number n increases, the number κ can only grow. The growth nature of $\kappa(n)$ of matrices W_A and W_B is usually different, and each of them depends on its own factors. Since the matrix W_B is formed from input signals, it is possible to control and influence behavior of $\kappa(n_b)$ in active experiments. The method of excitation with alternating intervals of excitation and relaxation makes it possible to provide a rather weak change of $\kappa(n_b)$ with increasing n_b. Moreover, there is every reason to believe, that with values of $\kappa(n_b)$ remaining close to unity, we will have the most informative signal.

The behavior of condition number of the matrix W_A depends on the following factors. What remains is its dependence on the informative input action, i.e. its ability to excite all modes of the system with an acceptable SNR. The SNR indicator, which is a relative value, naturally includes all errors acting on the system. This equally applies to both matrices W_A and W_B. The value of $\kappa(n_a)$ significantly depends on the dynamic features of the system, determined by the values of the invariants ω_p and ρ_p, as well as f_p^c and f_p^s. There is a rapid increase in the condition number with increase of n when the eigenvalues of the system are clustered or when there are very fast modes in the system. Small values of f_p^c and f_p^s lead to small output signals of the corresponding modes, so that for them the SNR becomes of the order of unity. Nevertheless, even for systems with the most favorable values of these invariants and with an informative input, the condition number of matrix W_A grows rather quickly as n_a increases, which is clearly seen from the Table.

Table

n_a	10	12	14
$\kappa(n_a)$	10^3	10^5	10^7

Table shows only the order of the condition number and it is calculated here only with exact data. With more unfavourable dynamic properties of the system, $\kappa(n_a)$ grows only faster, which makes the problem of parametric identification with increasing n_a more sensitive to errors in the data, i.e. we have a manifestation of bad conditionality of W_A. This property is fundamental and does not essentially depend on the dynamic properties of systems and methods that solve the problem of parametric identification. Therefore, at larger dimensions, when the problem becomes ill-posed, it is necessary to use regularization procedures, which can be used to ensure the stability of the obtained solutions and improve the quality of the models. Poor conditionality is also inherent to stochastic identification problems, which led to a shift in a paradigm for system identification [6].

The main attention in stochastic case was paid to the choice of various stabilizers that make it possible to ensure the consistency of the estimation. The most widely used is the Kernel structure for stabilizator, which contains parameters that are subject to tuning in order to provide optimal regularization. Various types of Kernel structure can be found in works [24–32]. Here in the framework of nonstohastic approach will be considered regularization more familiar to classical one [33].

5. Regularization

Since an ill-posed problem arises when solving (24), it is proposed to find a solution from (27) using the stabilizer Σ_{stab} in the following form:

$$\Sigma_{\text{stab}} = \begin{bmatrix} \sigma_n / \sigma_1 & 0 & \dots & 0 & 0 \\ 0 & \sigma_n / \sigma_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n / \sigma_{n-1} & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

where $\sigma_1, \sigma_2, \dots, \sigma_{n_a}$ are singular values of Σ_1. Instead of (27) we get the system

$$(\Sigma_1 + \alpha\Sigma_{\text{stab}}) \cdot \theta_1 = \bar{\bar{y}}^1, \tag{33}$$

where α — is a regularization parameter, which is chosen from the residual principle by [33]. Parameter α changes in a geometric progression

$$\{\alpha_k\} = \alpha_0\{\gamma^k\},\ 0 < \gamma < 1,\ k = 0, 1, 2, \dots,$$

where α_0 is chosen to be of order 10-2. With $k = 0$ we have an regularized solution $\theta_a^{(0)}$

$$\theta_a^{(0)} = V_1 \cdot (\Sigma_1 + \alpha_0\Sigma_{\text{stab}})^{-1} U_1^{\text{T}} \bar{y}. \tag{34}$$

Without regularization, the vector $\theta_b^{(0)}$ is found and the impulse response function of the model is calculated. It is compared with the impulse response function of the system, which can be taken as the free motion on the relaxation interval after excitation by a long rectangular pulse. Let $\hat{h}^0(t)$ be the impulse response function of the model, and $\tilde{h}(t)$ — impulse response function of the systems. The accuracy of the solution is defined as

$$\left\| \tilde{h}(t) - \hat{h}^0(t) \right\|_\infty = \max_t \left| \tilde{h}(t) - \hat{h}^0(t) \right|.$$

Next, we check the fulfillment of the condition

$$\left\| \tilde{h}(t) - \hat{h}^0(t) \right\|_\infty \le 2\varepsilon_y. \tag{35}$$

When the condition (35) is met, we obtain a regularized solution that is consistent in accuracy with the data error. Otherwise, we take $\alpha = \alpha_2$ and repeat the above steps. The decrease of α continues until (35) is fulfilled.

If, starting from some k, the monotonic character of the residual in (28) is violated, we increase n_a and n_b by one and continue the search for a regularized solution. The increase in dimensions is interrupted if there is no improvement in the quality of the model. In this case, the model of the smallest dimension is taken, among those that give a suitable approximation.

6. Simulation results

This section presents experimental studies of the fundamental features of identification for systems with different dynamic properties and structures using numerical simulation. They are mainly determined by the system dimension and invariants, including eigenvalues and identifiability parameters characterizing the contribution of each individual mode to the output.

To do this, we use the model representation (10), (11), in which the specified invariants are represented by parameters ρ_p, ω_p and f_p^c, f_p^s. The complexity of the system is characterized by dimension n. Errors in measured output and input were assumed to be uniformly distributed over the membership intervals (7).

For a given model dimension, the main characteristics that determine the quality of parameter estimation are the conditionality of the matrix $\bar{W}_A$ in the overdetermined system (25) and the matrix W_B in a similar system (31). The proposed excitation of the system by a large number of signals of types (12), (13) provides good conditionality of the matrices W_B in a wide range of values for any system under study.

The conditionality of the matrix $\bar{W}_A$ essentially depends on the dynamic properties of the system, determined by its dimension and the values of the above invariants. As the dimension of the system and the approximating model increase, the conditionality grows rapidly. This is clearly seen from Fig. 1 and Fig. 2, where dependence of conditionality of the matrix $\bar{W}_A$ on model order is shown for the systems of order 6 (Fig. 1) and of order 15 (Fig. 2). The dashed line shows the case with an uncertainty level of 10^{-6}.

On Fig. 1, the data generating system has dimension $n = 6$, with two real and two complex conjugate eigenvalues. The conditionality of the system with $n = 15$, which has three real and six complex conjugate eigenvalues, is shown in Fig. 2.

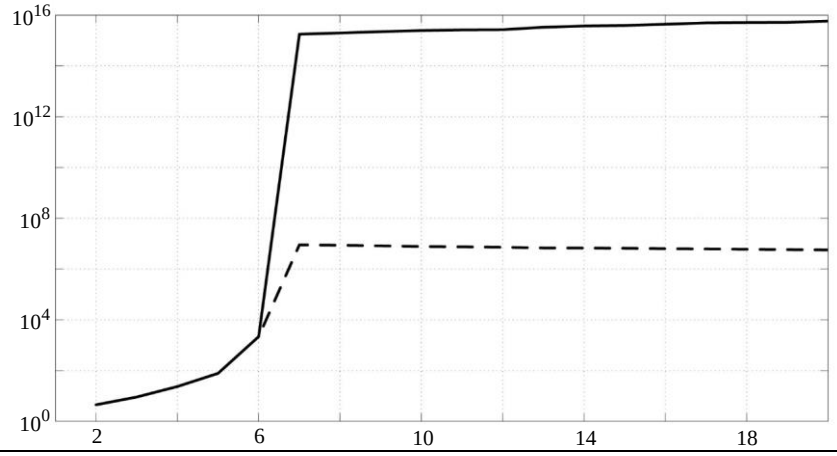

Fig. 1

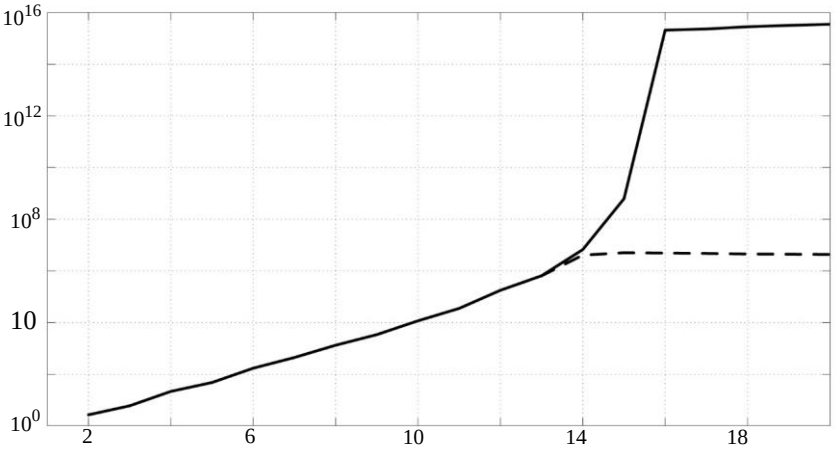

Fig. 2

 ISSN 2786-6491

The conditionality of the matrix $\bar{W}_A$ for approximate models of different dimensions is shown to the left of $n = n_a$. The solid line shows the deterministic case, when there is no uncertainty in the data. The dashed line shows the case with an uncertainty level of 10^{-6}.

As can be seen from the figures, starting from a certain dimension, the condition number reached a constant level. In the deterministic case, the growth of conditionality stops when $n_a = n+1,$ and is set at the level of the computational error (10^{-16}). In the non-deterministic case, the increase in conditionality may stop earlier, as soon as it reaches a level corresponding to the uncertainty in the data (10^{-16}). Then the dimension is determined to be smaller than the true one, since some of the system modes have a signal with too low SNR.

The results presented in Fig. 1, 2 are obtained for a well-identified system, when ω_p is uniformly distributed over the interval $\left[0, \frac{\pi}{2}\right]$, all ρ_p are close to unity, and all values of f_p^s and f_p^c are the same. If the eigenvalues are clustered or there are modes with relatively small coefficients f_p^s and $f_p^c,$ then the growth of the conditionality $\bar{W}_A$ with increasing dimension accelerates even more.

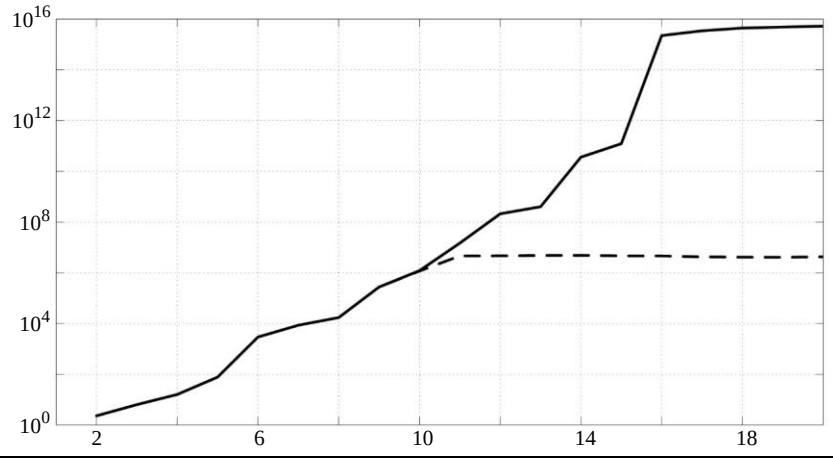

Fig. 3

Fig. 3 shows the conditionality of $\bar{W}_A$ for a system with $n = 15,$ whose eigenvalues form clusters near the stability boundary. It can be seen that for a well-identified system, the condition number for model of order $n_a = 10$ has a value of about $10^4,$ and for a system with clustering of eigenvalues, it has a value of about 10^6.

The property of asymptotic convergence for the described identification method was investigated. The same system that was used to build Fig. 2 was chosen. The responses of the system and asymptotic models for different n_a to a rectangular pulse are shown in Fig. 4–6. On Fig. 4 the model has a dimension of $n_a = 10,$ in Fig. 5 — $n_a = 13,$ in Fig. 6 — $n_a = 15,$ i.e. coincided with the dimension of the original system. The plots to the left at the figures corresponds to the deterministic case and the right plots correspond to the non-deterministic case with input and output uncertainty of value 10^{-6}. True system response is shown as solid line, and model response is shown as dashed line.

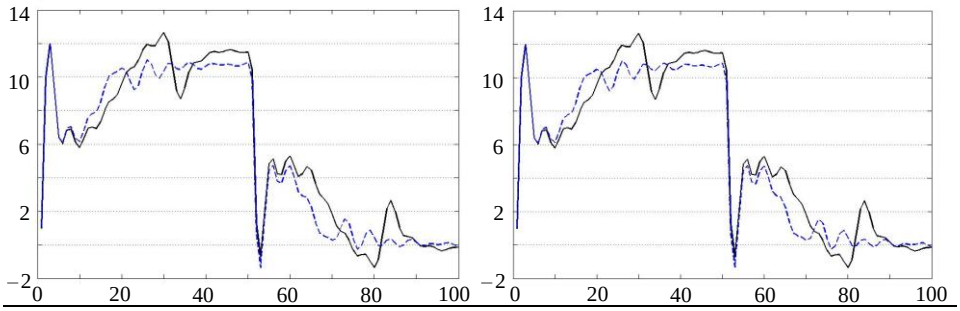

Fig. 4

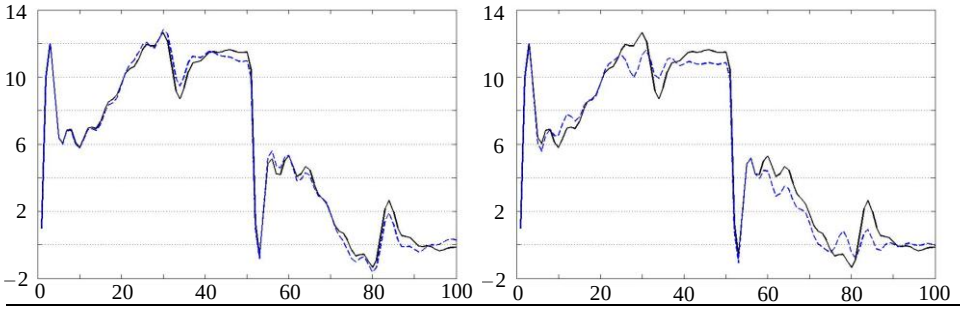

Fig. 5

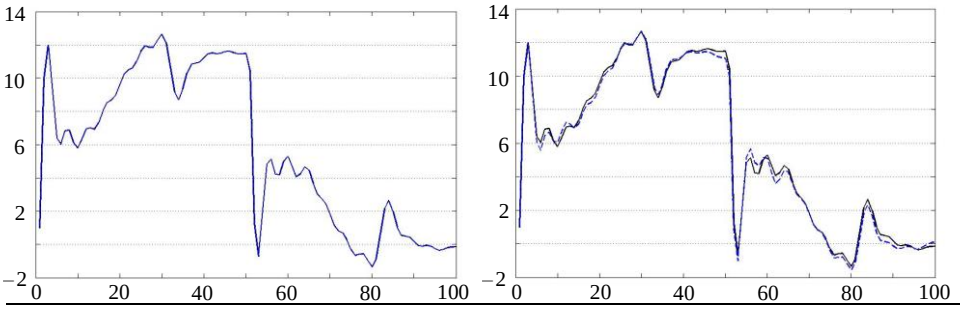

Fig. 6

Since the conditioning is bad in all cases, the regularization procedure (34), (35) was used. When approaching the exact dimension, the quality of the model improved.

When n_a exceed n, in the deterministic case, the coincidence of impulse response functions was obtained, which indicates the construction of a non-minimal model. For such models, the rank of the product of the controllability and observability matrices remained equal to n. In the non-deterministic case, the rank of the identifiability matrix increased with a probability of almost one, but the quality of the model did not improve significantly.

At a higher level of uncertainty, this effect appears for asymptotic models starting from a certain dimension n_a less than n, so further increase in dimension did not significantly improve the quality of the model. This value can be considered as the most appropriate dimension of the asymptotic model. It corresponds as a rule, to the stop of the growth of the conditionality of the matrix Y_{n_a}, according to Fig. 3. Thus, there is reason to believe that the dimension n_a, at which the condition number reaches saturation, is the most appropriate for the approximating model.

The dependence of the singular values of the matrices $\tilde{Y}$ (15) and $\bar{W}_A$ (25) on dynamic invariants and the magnitude of the uncertainty was studied. It has been found that the singular values can be used to judge the dynamic properties of the system, the number of essentially excited modes, and the level of noise.

For example, Fig. 7 shows the singular values of the matrix $\tilde{Y}$ (15) with different noise levels. Generating system has dimension 15 with 6 complex conjugate pairs and 3 real eigenvalues. The values of the invariants f_p^c and f_p^s for different modes differ up to 100 times. Solid line shows singular values in the deterministic case, dashed and dashed-dotted lines correspond to the noise $\varepsilon_u = \varepsilon_y = 10^{-6}$ and $\varepsilon_u = \varepsilon_y = 10^{-3}$, respectively.

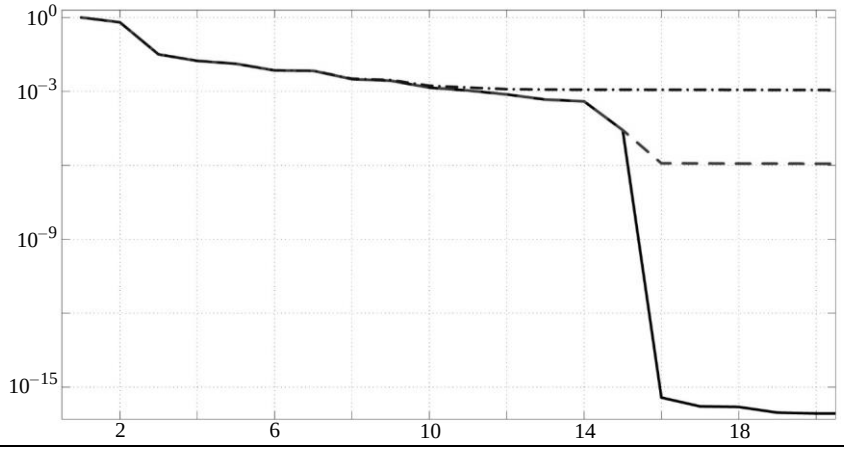

Fig. 7

Dealing specifically with this system in the deterministic case, one can hope to build a full-size model. In the case $\varepsilon_u = \varepsilon_y = 10^{-6}$, the successful construction of a full-size model is unlikely. In the case $\varepsilon_u = \varepsilon_y = 10^{-3}$, a model with $n_a \geq 10$ is non-minimal, because all further modes will have an SNR below one.

Conclusion

The non-stochastic approach to the identification of complex systems in the framework of asymptotic modeling is more realistic for practical use, since it uses less stringent restrictions on the uncertainty in the data, including finite-time observation. To ensure the effectiveness of this method, two main requirements must be met.

The first, it is necessary that the input signal be universally informative, i.e. for a system with arbitrary dynamic properties, the output of each mode of the system had an acceptable SNR. Otherwise, low SNR modes are poorly identified. The proposed design of active experiment with alternating intervals of excitation and relaxation partially solves this problem, due to the use of resonance effects. If the frequencies of the input signals vary with a small step, then the informativeness of the input becomes more universal. In principle, the most versatile informative input signals are white noise signals. However, their implementation in practice can cause certain problems. The proposed method of excitation allows you to control its spectrum, which allows you to use, for example, a priori knowledge about the dynamic properties of the system.

The second important aspect of non-stochastic identification is related to the choice of model dimension, which is the best in each particular case. To do this, within the framework of the described method, it is proposed to use the principle of choosing a model that is consistent in accuracy with the errors in the available data. In the first approximation, it is used in condition (20). Further, for the dimension established by (23), the problem of parametric identification is solved, after which the consistency of the

model and the system is checked by the difference in responses. If the principle is fulfilled, then the taken dimension of the model is final. If the model error is greater, we increase its dimension and perform parametric identification again. The principle of consistency is checked and a decision is made to stop or to continue increasing the dimension of the model. In this case, if the problem is ill-posed, regularization should be used. However, the case is not ruled out when the quality of the model reaches saturation and does not improve. Then, as the desired dimension, we take the smallest one corresponding to saturation.

In conclusion, we note that the process of identifying complex systems cannot be reduced to a formal solution of a strictly formulated mathematical problem. It is rather a research process of establishing patterns of system behavior based on the data of one or more experiments with the analysis of their results and the use of mathematical tools, the basics of which are described in this paper.

В.Ф. Губарев, С.В. Мельничук, М.М. Сальніков

ПАРАДИГМА НЕСТОХАСТИЧНОГО ПІДХОДУ ДО ІДЕНТИФІКАЦІЇ СИСТЕМ

Губарев В'ячеслав Федорович

Інститут космічних досліджень НАН України та ДКА України, м. Київ,

v.f.gubarev@gmail.com

Мельничук Сергій Вікторович

Інститут космічних досліджень НАН України та ДКА України, м. Київ,

serg.vik@ukr.net

Сальніков Микола Миколайович

Інститут космічних досліджень НАН України та ДКА України, м. Київ,

salnikov.nikolai@gmail.com

Під поняттям складна система у цій роботі розуміється велика сукупність динамічних взаємодіючих систем, точна математична модель яких невідома чи має дуже велику розмірність. Застосування стандартних методів синтезу зворотних зв'язків для таких систем стає складним і навіть неможливим із-за виродженості відповідних математичних задач. Один із виходів із такої ситуації полягає у побудові апроксимаційної моделі зниженої розмірності. Це може бути зроблено з використанням системи вихідних рівнянь, якщо вони є, або методів ідентифікації на основі вимірювань вихідних і вхідних змінних, що діють на систему. У цьому випадку процес побудови математичної моделі зводиться до послідовного перебору можливих моделей зі зростаючою складністю. Як критерій адекватності моделі розглядається норма відхилення виходу моделі від виміряного значення виходу досліджуваної системи. У статті розглядається побудова лінійних моделей, складність яких визначається розмірністю. У рамках нестохастичного підходу розроблено методологічну та математичну основу реконструкції моделей, що описують процеси у складних системах. Асимптотичне моделювання дозволяє для такої системи формувати класи моделей, які підходять для розв'язання задачі ідентифікації. Точний опис відповідає нескінченному розширенню, тому якість моделі покращується зі збільшенням її розмірності. Однак помилки в наявних даних не дозволяють безмежно

 ISSN 2786-6491

збільшувати їхню розмірність через погану обумовленість задачі ідентифікації, починаючи з деякого виміру. Процедура регуляризації дозволяє визначити ефективне наближене рішення задачі ідентифікації, яке для нестохастичного випадку узгоджується з помилками даних. Властивості та особливості пропонованого підходу ілюструються результатами моделювання.

Ключові слова: ідентифікація системи, лінійна регресія, регуляризація, асимптотичне моделювання, наближений розв'язок.

REFERENCES

1. Glover K., Curtain R.F., Partington J.R. Realization and approximation of linear infinite-dimensional systems with error bounds. *SIAM Journal of Control and Optimization*. 1988. Vol. 26, N 4. P. 863–898. https://doi.org/10.1137/0326049.
2. Ljung L. System identification — theory for the user (2nd ed.). Upper Saddle River, NJ : Prentice-Hall, 1999. 631 p.
3. Wahlberg B. On the use of orthogonalized exponentials in system identification. *In Report LiTH-ISY-1099, Department of Electrical Engeneering*. Linkoping University, Sweden, 1990. 137 p. Wahlberg B. System identification using Laguerre models. *IEEE Transactions on Automatic Control*. 1991. Vol. 36, N 5. P. 551–562. DOI:10.1109/9.76361
4. Van den Hof P.M.J., Heuberger P.S.C., Bokor J. System identification with generalized orthonormal basis functions. *Automatica*. 1995. Vol. 32, N 12. P. 1821–1834. DOI:10.1016/0005-1098(95)00074-4
5. Ljung L., Chen T., Mu B. A shift in paradigm for system identification. *International Journal of Control*. 2020. Vol. 93, N 2. P. 173–180. DOI:10.1080/00207179.2019.1578407
6. Söderström T., Stoica P. System identification. U.K. : Prentice-Hall, 1989. 637 p.
7. Pintelon R., Schoukens J. System identification — a frequency domain approach. New York, NY : IEEE Pres., 2001. 648 p.
8. Söderström T. Accuracy analysis of the Frisch scheme for identifying errors-in-variables system. *IEEE Transaction on Automatic Control*. 2007. Vol. 52, N 6. P. 985–997. doi: 10.1109/TAC.-2007.899005.
9. Söderström T. Errors-in-variables methods in system identification. *Automatica*. 2007. Vol. 43, N 6. P. 939–958. DOI:10.1016/j.automatica.2006.11.025
10. Van Huffel S., Lemmerling P.(Eds.). Total least squares and errors-invariables modelling: Analysis, algorithms and applications. Springer, 2002. 388 p.
11. Guidorzi R.P., Beghlli S., Soverini U. The frisch scheme in dynamic system identification. *Automatica*. 1990. Vol. 26, N 1. P. 171–176. DOI:10.1016/0005-1098(90)90168-H
12. Hong M., Söderström T., Zheng W.X. Accuracy analysis of bias-eliminating least squares estimates for errors-in-variables systems. *Automatica*. 2007. Vol. 43, N 9. P. 1590–1596. DOI:10.1016/j.automatica.2007.02.002
13. Söderström T., Soverini U. Errors-in-variables identification using maximum likelihood estimation in the frequency domain. *Automatica*. 2017. Vol. 79. P. 131–143. DOI:10.1016/-j.automatica.2017.01.016
14. Zhang E., Pintelon R. Identification of multivariable dynamic errors-in-variables system with arbitrary inputs. *Automatica*. 2017. Vol. 82. P. 69–78. DOI:10.1016/j.automatica.2017.04.031
15. Zhang E., Pintelon R. Nonparametric identification of linear dynamic errors-in-variables systems. *Automatica*. 2018. Vol. 94. P. 416–425. DOI:10.1016/j.automatica.2018.04.039
16. Kang H., Gu G., Zheng W.X. A graph subspace approach to system identification based on errors-in-variables system models. *Automatica*. 2019. Vol. 109. DOI:10.1016/j.automatica.-2019.108535
17. Milanese M., Norton J., Piet-Lahanier H., Walter, E.(eds). Bounding Approaches to System Identification. New York : Plenum Press, 1996. 586 p.
18. Garulli A., Reinelt W. On model error modeling in set membership identification. *IFAC Proceedings Volumes*. 2000. Vol. 33, N 15. P. 169–174. DOI:10.1016/S1474-6670(17)-39745-8.
19. Giarre L., Kacewicz B.Z., Milanese M. Model quality evaluation in set membership identification. *Automatica*. 1997. Vol. 33, N 6. P. 1133–1139. DOI:10.1016/S0005-1098(97)00007-1.
20. Milanese M., Novara C. Set membership identification of nonlinear systems. *Automatica*. 2004. Vol. 40, N 6. P. 957–975. DOI:10.1016/j.automatica.2004.02.002

21. Gubarev V., Melnychuk S., Salnikov N. Identification of complex systems in the class of linear regression models. *In Recent Developments in Automatic Control Systems*. Edited by: Kondratenko Y.P., Kuntsevich V.M., Chikrii A.A., Gubarev V.F. River Publishers, 2022. P. 167–195.
22. Golub G.H., Van Loan C.F. Matrix computations. London : Johns Hopkins University Press, 2012. 694 p.
23. Pillonetto G., Chiuso A., De Nicolao G. Regularized estimation of sums of exponentials in spaces generated by stable spline kernels. *Proc. 2010 American Control Conference (Baltimor, MD: Marriott Waterfront)*. P. 498–503. DOI:10.1109/ACC.2010.5530862
24. Pillonetto G., Chiuso A., De Nicolao G. Prediction error identification of linear systems: A nonparametric Gaussian regression approach identification. *Automatica*. 2010. Vol. 47, N 2. P. 291–305. DOI:10.1016/j.automatica.2010.11.004
25. Chen T., Andersen M.S., Ljung L., Chiuso A., Pillonetto G. System identification via sparse multiple kernel-based regularization using sequential convex optimization techniques. *IEEE Transactions on Automatic Control*. 2014. Vol. 59, N 11. P. 2933–2945. DOI:10.1109/TAC.2014.-2351851.
26. Chen T., Ohlsson H., Ljung L. On the estimation of transfer functions, regularizations and Gaussian processes — revisited. *Automatica*. 2012. Vol. 48, N 8. P. 1525–1535. DOI:10.1016/-j.automatica.2012.05.026
27. Chen T., Ardeshiri T., Carli F.P., Chiuso A., Ljung L., Pillonetto G. Maximum entropy properties of discrete-time first-order stable spline kernel. *Automatica.* 2015. Vol. 66. P. 34–38. DOI:10.-1016/j.automatica.2015.12.009
28. Dinuzzo F. Kernel for linear time invariant system identification. *SIAM Journal on Control and Optimization.* 2015**.** Vol. 53, N 5. P. 3299–3317. DOI:10.1137/130920319
29. Carli F.P., Chen T., Ljung L. Maximum entropy kernels for system identification. *IEEE Transactions on Automatic Control*. 2017. Vol. 63, N 3. P. 1471–1477. DOI:10.1109/-TAC.2016.2582642.
30. Markonato A., Schoukens M., Schoukens J. Filter-based regularization for impulse response modelling. *IET Control Theory & Applications*. 2017. Vol. 11, N 2. P. 194–204. DOI:10.1049/iet-cta.2016.0908
31. Zorzi M., Chiuso A. The harmonic analysis of kernel functions. *Automatica*. 2018. Vol. 94. P. 125–137. DOI:10.1016/j.automatica.2018.04.015
32. Tikhonov A.N., Arsenin V.Y. Solution of ill-posed problems. Washington, DC : Winston/Wiley, 1977. 272 p.

Submitted 14.03.2023