UDC 681.5.015(045)
DOI:10.18372/1990-5548.62.14391

[1]A. M. Silvestrov,
[2]O. M. M'yakshylo,
[3]G. I. Kryvoboka

MODIFICATION OF THE METHOD OF LEAST SQUARES

[1]National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
[2,3]Department of Information Systems, National University of Food Technologies
E-mails: [1]silvestrovanton@gmail.com ORCID 0000-0003-4554-6528, [2]mem2004@ukr.net,
[3]galinakryvoboka@gmail.com orcid.org/0000-0002-5011-0860

***Abstract*—*The article deals with the problems of using a focused multilevel information system in optimal management. In order to make such a system more efficient, it is proposed to modify the least-squares method, which provides unbiased estimates of the parameters in the real situation of noisy measurements of input and output signals of the primary converters. Parameters are estimated using the proposed method, which ensures smoothing of external influences of the model under study on the results. The efficiency of the considered method is confirmed by comparison with the method of least squares.*

Index Terms—Parametric identification; ordinary least squares; unbiasedness; estimation efficiency; modification of ordinary least squares; weight function.

I. INTRODUCTION

Recently, problems of identification and control of stochastic systems, which occupy a significant place in the theory and practice of management, are of particular interest. The behavior of a real object, which operates in the conditions of natural, industrial and other noises, is characterized by some uncertainty. In addition, automated systems for managing complex objects are usually attended by people who are characterized by some uncertainty about behavior. Describing such systems using well-known deterministic approaches is not always effective and does not reflect the true picture of the operation of the object. Methods for deterministic systems have not been developed to solve the problem of optimal management of this class of objects. Thus, the need to develop a mathematical model of a universal multilevel system is caused by the urgent needs of management practice, and this theory maximally approximates a formalized representation to the actual conditions of operation.

Finding optimal control in dynamic systems requires a solution in the process of managing a fairly complex mathematical problem. The mathematical model of the system is important. In order to ensure good control of a dynamic system, it is necessary that its mathematical model is known with sufficient accuracy. In this case, the identification subsystem plays an important role. The application of identification methods enables quantification of the best decisions and enables optimization of management.

II. PROBLEM STATEMENT

Due to the wide automation of production processes, there is a growing interest in the methods of constructing mathematical models of real dynamic systems, which are prone to uncontrolled random influences. As these models are approximate, their synthesis should state the requirements (optimality criteria) that they must satisfy. These criteria can be traditionally applied in the theory of experiment planning to optimality criteria, or special criteria that take into account the ultimate goal of using synthesized models in specific application problems solved on their basis. Such tasks include parametric optimization of dynamic systems, optimal control and other tasks.

Building a mathematical model of a real dynamic system is possible based on the results of either a passive or active experiment. Passive identification methods involve processing information collected by observing object entry and exit. Active identification methods involve the submission of the test object of the test signal to the input, the synthesis of which is carried out on the basis of the theory of optimal experiment and processing of "input-output" implementations.

In cybernetics, there is a principle of multiplicity of the mathematical model, but this does not facilitate the task of choosing the optimal model.

The optimal mathematical model should be not by itself, but by which the main task on the object (management, diagnostics, forecast, etc.) is best solved. Therefore, the construction of MM based on a priori data and the result of the experiment on a real object must be consistent with the main task. Then, from an unlimited number of approaches to the identification problem, it is necessary to determine the structural and parametric identification method and mathematical model that is the most optimal in the main indicator.

The lack of a clear classification and quantitative comparison of methods and models makes the task of finding the optimal method-model pair the most important task, which in order to effectively solve this classification (systematization) within information technologies. Namely, creation of "knowledge base" on the basis of qualitative and quantitative ordering of methods and mathematical models.

One of the ways to solve the actual problem of today is to use a focused multilevel information system. Such a system can be a system "MASS" [3].

Consider the real problem of any control system that requires an optimal result within the task. Three multilevel subsystems can provide the solution to this problem (Fig. 1):

– the optimized object management subsystem;
– the object identification subsystem;
– signal identification subsystem.

The effectiveness of such a system depends on its internal structure, that is, the fullness of methods and models that ensure the solution of the task.

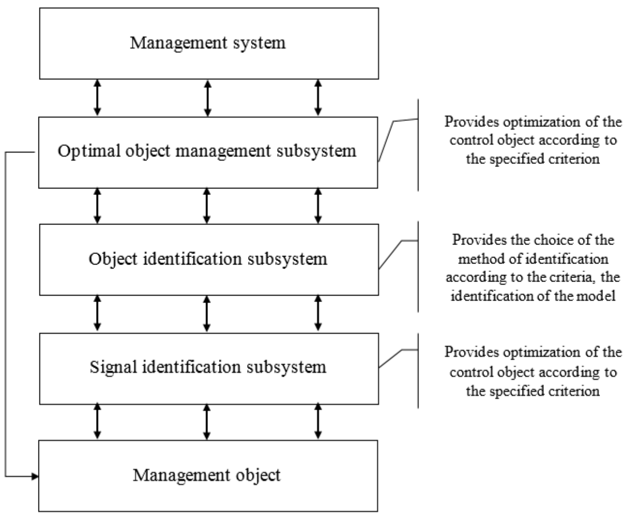

Fig. 1. Optimal multilevel information system

Taking into account the properties of real objects and the fundamental approximation of their models, we introduce the definition of optimal estimates β^* in the class of nontargeted (without taking into account the main indicator I_{qk}) [3]. For real systems Σ_∞ (limited power), any coordinate $y^*(t)$ is a completely continuous [7] argument function t, measured at discrete moments of time t, or simply k. According to the Weierstrass theorem [7], a continuum on a segment $[t_1,t_2]$ or $(k=\overline{1,M})$ function $y^*(t)$ or $y^*(k)$ can be approximated with any accuracy by some system of linearly independent functions $x_i^*(t)$ or $x_i^*(k)$, $i=1, 2,...$. For a given finite basis of functions $x_i^*(t)$ or $x_i^*(k)$ estimates $\hat{\beta}$ will be optimal estimates $\hat{\beta}^*$, if they minimize to some extent the distance between the exact signal $y^*(t)$ or $y^*(k)$ its model $X^*\beta$. Because we only know that $Y=Y^*+\varepsilon$, then as a physically realized measure of closeness Y^* to $\hat{Y}$ it is possible to take the difference functional Y and $\hat{Y}$. For example, in Euclidean space functional $\varepsilon^T\varepsilon$ leads to the estimation LSM [3], which are unbiased and effective under the condition that there is a Gaussian white noise. For a finite dimension n of a vector $\hat{\beta}$ infinite-dimensional space, stretched over a function $x_i^*(t)$, is reduced to a dimension subspace n, in which the basis forms a system of functions $x_i^*(t)$, $i=\overline{1,n}$, and the coefficients $\hat{\beta}_i$ represent the coordinates of the functions in the basis $\{x_i^*\}$ of the n-dimensional space. With a minimum of functional $\varepsilon^T\varepsilon$

$$\frac{\partial}{\partial\beta}(\varepsilon^T\varepsilon)=0, \tag{1}$$

which is necessary and sufficient (due to linearity ε through β and linear independence x_i^*), obtain a system of normal equations

$$A\hat{\beta}=B, \tag{2}$$

where (in the case of discrete time measurements) $A=\left[a_{ij}\right]$ – symmetric positively defined matrix of scalar products $a_{ij}=\sum_{k=1}^{M}x_i^*(k)x_j^*(k),\ i=\overline{1,n},\ j=\overline{1,n},$ $B=[b_i]$ is vector column $b_i=\sum_{k=1}^{M}y(k)x_i^*(k),$ $i=1,n;$ $\hat{\beta}=\left[\hat{\beta}_i\right]_1^n$ is the vector column estimates $\hat{\beta}_i$ of LSM.

From an infinite dimensional space of functions y, x moving into a n-dimensional vector space R^n, which the solution of the system (2) is searched. Then, by the best nontargeted estimate $\hat{\beta}^*$ for the model, we mean the estimate that minimizes the mean square of the difference ε^* between the exact values $y^*(k)$ and their model $\hat{y}(k)$:

$$\varepsilon^*(k) = y^*(k) - X^*(k)\hat{\beta}^*, k = \overline{1, M}. \quad (3)$$

In the general case $y(k)$ and $x(k)$ represent the coordinates of the real system Σ_∞, noisy obstructed and measured with errors. Consider an estimation algorithm that provides an opportunity for noisy $y(k)$ and $x(k)$ get estimates close $\hat{\beta}$ to $\hat{\beta}^*$.

III. THE INTEGRATED LEAST SQUARES METHOD

The method of least squares (MLS), due to its wide field of application, occupies an exceptional place among the methods of parametric identification, but its disadvantage is the variation of values of functional *I*. This disadvantage can be eliminated by further averaging over a set of quasi-statistically independent functionals close to the rms for accurate data. Such functionals can be shifted in time *t* by the interval θ of average products $\frac{1}{T}\int_0^T \varepsilon(t)\,\varepsilon(t+\theta)dt$. Averaging them over the interval $[-\tau_1, \tau_1]$, we get the functionality:

$$I = \frac{1}{2}\int_{-\tau_2}^{\tau_1} \eta(\theta)\int_0^T \varepsilon(t)\cdot\varepsilon(t+\theta)dtd\theta, \quad (4)$$

where $\eta(\theta)$ is the weight function.

From the necessary condition for a minimum with respect to β_k, $k = \overline{1,n}$ of the exponent (6):

$$\frac{\partial I}{\partial \beta_k} = \frac{1}{2}\int_{-\tau_1}^{\tau_1}\eta(\theta)\int_0^T\left[\frac{\partial\varepsilon(t)}{d\beta_k}\varepsilon(t+\theta) + \varepsilon(t)\frac{\partial\varepsilon(t+\theta)}{\partial\beta_k}\right]dtd\theta$$

$$= \int_{-\tau_1}^{\tau_1}\eta(\theta)\int_0^T\left(-x_k(t)\right)\left[y(t+\theta) - \sum_{i=1}^{n}\beta_i x_i(t+\theta)\right]$$

$$+\left(-x_k(t+\theta)\right)\left[y(t) - \sum_{i=1}^{n}\beta_i x_i(t)\right] = 0, \ k = \overline{1,n}, \quad (5)$$

Let's obtain a system of equations:

$$A\cdot\hat{\beta} = B, \quad (6)$$

where *A* is the matrix $n\times n$ with elements a_{ik}; *B* is matrix-column $n\times 1$ with the elements b_k:

$$a_{ik} = \sum_{1=-p}^{p}\eta(l)\sum_{j=1}^{M}\left[(x_i(j+l))x_k(j) + x_i(j)x_k(j+l)\right],$$

$$b_k = \sum_{1=-p}^{p}\eta(l)\sum_{j=1}^{M}\left[(y(j+l))x_k(j) + y(j)x_k(j+l)\right].$$

The solution of system (6) gives the required estimate $\hat{\beta}$:

$$\hat{\beta} = A^{-1}\cdot B. \quad (7)$$

The weight function $\eta(m)$ can be found in the class of finite functions that are symmetric with respect to *m* = 0 (such that $\eta(0) = \eta(\pm m_{kr})$. For example:

$$\eta(m) = \eta(m,\gamma,\theta) = \left(1+|m|\right)^\theta\left(1-\cos\frac{|\pi m|}{m_{kr}}\right)^\gamma, \quad (8)$$

where $\theta\in(\pm\infty), \gamma\in(0,\infty), m_{kr}$ is determined when:

$$\det\left[X^T\left(X_{m_{kr}} + X_{-m_{kr}}\right)\right] \cong 0. \quad (9)$$

The parameters θ and γ are optimized for the main (external) exponent *I* [2], [6]. The parameter γ affects the width of the pulse $\eta(m)$, and θ its asymmetry relative to the maximum.

The proposed method is called the Integrated Least Squares (IMLS) method.

IV. COMPARATIVE ANALYSIS OF IMLS AND MLS

The quality of parametric estimation is affected by the degree of interrelation of the variables $x_i(t)$, $i = \overline{1, n}$ and not by their number. Therefore, let's confine ourselves to a simple example. The equation describing the process is presented in the form (4), where:

$$y^*(k) = \beta_1^* x_1^*(k) + \beta_2^* x_2^*(k), \quad k = \overline{1{,}1000},$$

$$\beta_1^* = \beta_2^* = 1, \ x_1^* = \sin\frac{\pi k}{500}, \ x_2^* = \sin\left(\frac{\pi k}{500} + \frac{\pi}{6}\right).$$

On measurements $y(k)$, $x_1(k)$, $x_2(k)$ white noise is imposed – random numbers with uniform distribution in the range [±1]. For an objective estimation of the displacement and spread of estimates β_1, β_2 ten statistically independent realizations of noise are generated. The results of identifying the coefficients β_1, β_2 for the least squares method and the proposed method are given in Table I. Estimates of β_1 and β_2 by LSM (Table I) are underestimated by almost 50% (9). However, there is a regularization [6]: the spread $\sigma_{\beta i}$ of estimates β_i is 0.02 and 0.05. In the proposed method (Table I), the estimates are almost unchanged: 1.005 and 0,943, but the spread is greater than in a regularized LSM (0, 15, 0.16). Reducing the spread is possible due to a compromise between displacement and dispersion by changing the parameters θ and γ weight functions.

In the case of noise only in the original variable (Table II) (ideal situation for least squaresmethod), the estimates are unbiased, but the spread of the estimates for this method (0.07 and 0.09) is greater than the spread (0.05 and 0.08) optimization of parameters θ and γ of the function $\eta(m)$. In the case that there is an opportunity to optimize $\eta(m)$ [4], the gain of the proposed method in the sense of unbiasedness and the effectiveness of estimates relative to least squares method is much larger.

TABLE I. THE RESULTS OF ESTIMATING THE PARAMETERS FOR NOISINESS OF INPUT AND OUTPUT VARIABLES

N	Least squares method		Estimations by the proposed method	
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
1	0.4790	0.4981	1.0282	0.9094
2	0.4607	0.4493	1.0179	0.8844
3	0.4843	0.5663	1.0916	0.8435
4	0.5024	0.5401	1.0020	0.9290
5	0.5246	0.4659	1.0313	0.9798
6	0.4997	0.5058	1.2904	0.7437
7	0.4849	0.5255	0.7093	1.2307
8	0.4919	0.4431	0.8283	1.1910
9	0.4676	0.4856	1.0825	0.7197
10	0.4642	0.6015	0.9653	1.0112
$\bar{\beta}$	**0.4860**	**0.5082**	**1.0047**	**0.943**
$\sigma^2_{\beta i}$	**0.00038**	**0.0026**	**0.0241**	**0.0283**
σ_{β_i}	**0.0197**	**0.0511**	**0.1551**	**0.1682**

TABLE II. THE RESULTS OF ESTIMATING THE PARAMETERS WITH NOISINESS OF ONLY THE ORIGINAL VARIABLES

N	Least squares method		Estimations by the proposed method	
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
1	0.9781	0.9212	0.9497	1.1019
2	1.0541	0.9371	1.0234	0.8554
3	0.9329	1.0817	0.9825	1.0618
4	1.1181	0.8819	1.0111	0.9132
5	0.9847	1.0327	1.1907	0.9807
6	1.0009	1.0192	1.1018	0.9823
7	1.1549	0.8258	0.9866	1.1244
8	0.9407	1.0765	1.0216	0.9879
9	0.9578	1.0823	0.9639	1.0861
10	1.0007	0.9412	1.0961	0.9946
$\bar{\beta}$	**1.0123**	**0.9800**	**1.0280**	**1.0089**
$\sigma^2_{\beta i}$	**0.0055**	**0.0083**	**0.0027**	**0.0073**
σ_{β_i}	**0.0744**	**0.0911**	**0.0522**	**0.0854**

V. CONCLUSIONS

Theoretical [2], [3] and experimental calculations confirm that the proposed method of IMNC allows in a real situation noisy measurements of input and output signals of the primary converters, to obtain unbiased parameter estimates close to the MNC estimates for accurate measurements, as well as variance of estimates, smaller variance, MNCs. Choosing the weight function parameters can improve the accuracy of the obtained estimates, which will make it possible to use it effectively in targeted multilevel real-estate information management systems.

REFERENCES

[1] *Methods of classical and modern theory of automatic control*: Textbook in 3 volumes, under. ed. N. D. Egupova. Moscow: From the Moscow State Technical University. NOT. Bauman, 2000, 1000 p. (in Russian)

[2] M. Ya. Ostroverkhov and A. M. Silvestrov, *Methods of study of electrical systems and complexes*, Kyiv: TALKOM, 2019, 300 p.(in Ukrainian)

[3] A. N. Silvestrov and P. I. Chynaev, *Identification and optimization of automatic systems*, Moscow, Energoatomizdat, 1987, 200 p. (in Russian).

[4] P. Eikhoff, *Fundamentals of identification of control systems*, Moscow: Mir, 1975, 683 p. (in Russian)

[5] A. N. Tikhonov and V. Ya. Arsenin, *Methods for solving ill-posed problems*, Moscow: Nauka, 1979, 286 p. (in Russian).

[6] A. G. Ivakhnenko, *Long-term forecasting and management of complex systems*, Kyiv: Technica, 1975, 312 p. (in Russian).

[7] A. N. Kolmogorov and S. V. Fomin, *Elements of the theories of functions and functional analysis*. Moscow: Nauka, 1972 (in Russian).

Received September 27, 2019.

Silvestrov Anton. orcid.org/0000-0003-4554-6528
Doctor of Engineering Science. Professor.
National Technical University of Ukraine “Ihor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (1969).
Research area: filtration, identification of Complex Systems.
Publications: more than 200.
E-mail: silvestrovanton@gmail.com

M'yakshylo Olena. Candidate of Science (Engineering). Associate Professor.
Department of Information Systems, National University of Food Technologies, Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine (1974).
Research area: data mining in new decision-making information technologies.
Publications: more than 40
E-mail: mem2004@ukr.net

Kryvoboka Galina. orcid.org/0000-0002-5011-0860
Post-graduate student.
Department of Information Systems, National University of Food Technologies, Kyiv, Ukraine.
Education: Vinnitsa National Technical University, Vinnitsa, Ukraine, (2010).
Research interests: Information Technology.
Publications: 8.
E-mail: galinakryvoboka@gmail.com

А. М. Сільвестров, О. М. М’якшило, Г. І. Кривобока. Модифікація методу найменших квадратів

У статті розглянуто проблеми застосування цілеорієнтованої багаторівневої інформаційної системи в оптимальному управлінні. З метою ефективності застосування такої системи запропоновано модифікацію методу найменших квадратів, що забезпечує отримання незміщених оцінок параметрів в реальній ситуації зашумлених вимірів вхідних і вихідних сигналів первинних перетворювачів. Проведено оцінювання параметрів за допомогою запропонованого методу, що забезпечує згладжування зовнішніх впливів досліджуваної моделі на результати. Підтверджено ефективність розглянутого методу шляхом порівняння з методом найменших квадратів.

Ключові слова: ідентифікація, оптимальне управління; МНК-оцінювання; незміщеність; ефективність оцінок; інтегрований МНК; вагова функція.

Сільвестров Антон Миколайович. orcid.org/0000-0003-4554-6528
Доктор технічних наук. Професор.
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут. Київ, Україна (1968).
Напрям наукової діяльності: фільтрація, ідентифікація складних систем.
Кількість публікацій: більше 200.
E-mail: silvestrovanton@gmail.com

М’якшило Олена Михайлівна. Кандидат технічних наук. Доцент.
Кафедра інформаційних системи, Національний університет харчових технологій, Київ, Україна.
Освіта: Київський політехнічний інститут. Київ, Україна (1974).
Напрям наукової діяльності: інтелектуальний аналіз даних у нових інформаційних технологіях прийняття рішень.
Кількість публікацій: більше 40.
E-mail: mem2004@ukr.net

Кривобока Галина Іванівна. orcid.org/0000-0002-5011-0860
Аспірант.
Кафедра інформаційних системи, Національний університет харчових технологій, Київ, Україна.
Освіта: Вінницький національний технічний університет, Вінниця, Україна (2010).
Напрям наукової діяльності: інформаційні технології.
Кількість публікацій: 8.
E-mail: galinakryvoboka@gmail.com

А. Н. Сильвестров, Е. М. Мякшило, Г. И. Кривобока. Модификация метода наименьших квадратов

В статье рассмотрены проблемы применения целеориентированной многоуровневой информационной системы в оптимальном управлении. В целях эффективности применения такой системы предложено модификацию метода наименьших квадратов, что обеспечивает получение несмещенных оценок параметров в реальной ситуации зашумленных измерений входных и выходных сигналов первичных преобразователей. Проведена оценка параметров с помощью предложенного метода, обеспечивающего сглаживание внешних воздействий исследуемой модели на результаты. Подтверждена эффективность рассматриваемого метода путем сравнения с методом наименьших квадратов.

Ключевые слова: идентификация, оптимальное управление; МНК-оценивания; несмещенность; эффективность оценок; интегрированный МНК; весовая функция.

Сильвестров Антон Николаевич. orcid.org/0000-0003-4554-6528
Доктор технических наук. Профессор.
Национальный технический университет Украины «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина.
Образование: Киевский политехнический институт. Киев, Украина (1968).
Направление научной деятельности: фильтрация, идентификация сложных систем.
Количество публикаций: более 200.
E-mail: silvestrovanton@gmail.com

Мякшило Елена Михайловна. Кандидат технических наук. Доцент.
Кафедра информационных систем, Национальный университет пищевых технологий, Киев, Украина.
Образование: Киевский политехнический институт. Киев, Украина (1974).
Направление научной деятельности: интеллектуальный анализ данных в новых информационных технологиях принятия решений.
Количество публикаций: более 40.
E-mail: mem2004@ukr.net

Кривобока Галина Ивановна. orcid.org/0000-0002-5011-0860
Аспирант.
Кафедра информационных систем, Национальный университет пищевых технологий, Киев, Украина.
Образование: Винницкий национальный технический университет, Винница, Украина (2010).
Направление научной деятельности: информационные технологии.
Количество публикаций: 8.
E-mail: galinakryvoboka@gmail.com