ЭЛЕКТРОНИКА, ИЗМЕРИТЕЛЬНАЯ ТЕХНИКА И РАДИОТЕХНИКА

УДК 681.5.015.4
DOI 10.21685/2072-3059-2020-1-2

В. С. Безяев, П. П. Макарычев

ИДЕНТИФИКАЦИЯ ПАРАМЕТРОВ МОДЕЛЕЙ ОБЪЕКТОВ МЕТОДОМ РЕГРЕССИОННОГО АНАЛИЗА

Аннотация.

Актуальность и цели. Рассматривается процедура идентификации параметров моделей динамических объектов методом регрессионного анализа. Приводится обоснование и выбор структуры, типов компонентов наилучшей дискретной модели в виде разностных уравнений *n*-го порядка. Обсуждается последовательность оценки численных значений параметров дискретной модели объекта, соответствия этих параметров экспериментальным данным. Предлагается интегральный квадратичный критерий оценки степени адекватности модели с использованием измерений в дискретные моменты времени. Используется базовый подход параметрической идентификации – метод наименьших квадратов, который при соблюдении линейности и дискретности обеспечивает простое и универсальное решение. Рассматриваются вопросы оценки параметров непрерывных моделей на основе значений параметров дискретной модели.

Результаты. Разработана процедура оценки параметров дискретной и непрерывной моделей динамического объекта на основе результатов наблюдения за входной и выходной переменной на заданном интервале времени.

Выводы. Структура регрессионной модели должна быть согласована со структурой непрерывной и дискретной модели исходя из предполагаемого состава полюсов и нулей. Количество нулей определяется из условия минимума среднеквадратичного отклонения рассчитанных значений от наблюдаемых значений выходной переменной. Оптимальное значение полюсов и нулей определяется методом полного перебора возможных вариантов.

Ключевые слова: регрессионный анализ, математическая модель динамического объекта, идентификация параметров модели, непрерывная модель, дискретная модель.

V. S. Bezyaev, P. P. Makarychev

IDENTIFICATION OF OBJECT MODEL PARAMETERS BY THE METHOD OF REGRESSION ANALYSIS

Abstract.

Background. The procedure of identification of parameters of models of dynamic objects by the method of regression analysis is considered. The substantiation and choice of structure, types of components of the best discrete model, in the form of difference equations of the order are given. The sequence of estimation of numerical values of parameters of discrete model of object, correspondence of these parameters to experimental data is discussed. We propose an integral quadratic criterion for assessing the adequacy of the model using measurements at discrete times. The basic approach of parametric identification is used-the least squares method, which, while respecting linearity and discreteness, provides a simple and universal solution. The questions of estimation of parameters of continuous models on the basis of values of parameters of discrete model are considered

Results. The procedure of estimation of parameters of discrete and continuous models of dynamic object on the basis of results of observation of input and output variable on the set interval of time is developed.

Conclusions. The structure of the regression model must be consistent with the structure of the continuous and discrete models based on the expected composition of the poles and zeros. The number of zeros and zeros is determined from the condition of the minimum standard deviation of the calculated values from the observed values of the output variable. The optimal value of the poles and zeros is determined by a complete search of possible options.

Keywords: regression analysis, mathematical model of dynamic object, identification of model parameters, continuous model, discrete model.

Введение

Задача об идентификации параметров моделей динамических объектов может быть решена методами регрессионного анализа или воздействия на вход сигнала [1–3]. При регрессионном анализе данные состоят из значений зависимой выходной переменной и независимой входной переменной. Регрессионная модель есть функция независимой переменной и параметров с добавленной случайной переменной. Параметры модели настраиваются таким образом, что модель наилучшим образом аппроксимирует выходные данные. Критерием качества оценки параметров модели объекта, как правило, служит среднеквадратичное отклонение рассчитанных значений от наблюдаемых значений выходной переменной. Предполагается, что значение выходной переменной в каждый момент времени есть сумма значений входной переменной, значений входной и выходной переменных, задержанных на интервалы времени, кратные шагу дискретизации Δt и случайной величины. Регрессионный анализ может использоваться для параметрической идентификации непрерывных и дискретных моделей динамических объектов [4, 5].

Постановка и решение задачи

Предположим, что по результатам наблюдения за функционированием объекта зафиксированы значения входной $\tilde{x}_k = \tilde{x}(k\Delta t)$ и выходной $\tilde{y}_k = \tilde{y}(k\Delta t) + \tilde{\varepsilon}(k\Delta t)$ переменных в дискретные моменты времени $k\Delta t$, которые приведены в табл. 1. Наблюдения проведены при наличии шума, характеризуемого случайной переменной $\varepsilon(k\Delta t)$.

Графики наблюдаемых переменных $\tilde{x}_k, \tilde{y}_k + \tilde{\varepsilon}_k$ приведены на рис. 1. На основе анализа значений и графиков переменных практически невозможно определить структуру и параметры модели: количество полюсов и нулей, вид корней характеристического уравнения. В связи с этим целесообразно определить максимально допустимое количество полюсов n и нулей m. Для практической реализации параметрической идентификации можно принять $n_{\max} = 4$, $m_{\max} = 3$ и установить требование $n > m$ [6, 7]. В этом случае потребуется на основе наблюдаемых значений входной и выходной переменных решить $n_{\max}(m_{\max} + 1)$ задач оценки параметров дискретной (непрерывной) модели и выбрать лучший вариант по заданному критерию.

Таблица 1

Значения наблюдаемых входной и выходной переменных

k	4	5	6	7	8	9	…	23	24	25
$\tilde{x}_k$	…	2,025	1,695	0,302	0,872	0,812	…	1,118	1,458	1,930
$\tilde{y}_k$	…	3,508	4,044	3,905	3,268	2,376	…	1,662	1,166	1,195

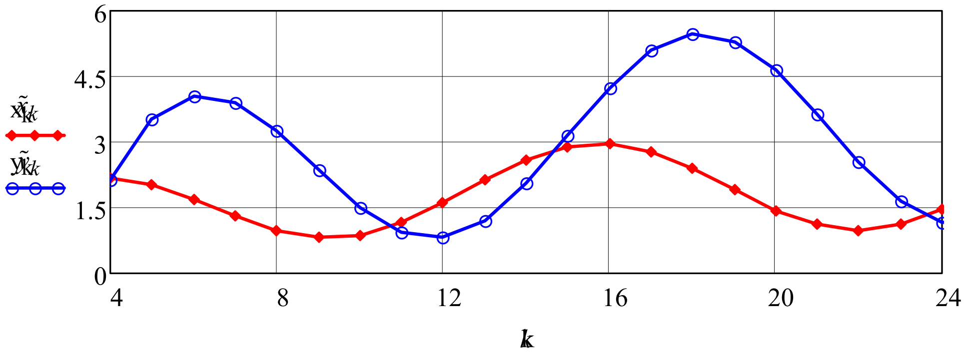

Рис. 1. Наблюдаемые переменные: $\tilde{x}_k$ – входная независимая переменная; $\tilde{y}_k$ – выходная зависимая переменная

Передаточная функция линейного объекта в изображении по Лапласу при выполнении условий $n \le n_{\max}$, $m \le m_{\max}$, $m < n$ имеет вид [6, 7]:

$$y(s) = \left[\prod_{i=1}^{m}(s+\beta_i) \Big/ \prod_{j=1}^{n}(s+\alpha_j)\right] Kx(s),\ n \ge m, \tag{1}$$

где α_j, β_i, K – полюса, нули и коэффициент передаточной функции исследуемого динамического объекта соответственно.

В соответствии с (1) конечно-разностную модель объекта, соответствующую передаточной функции (1), можно определить в виде

$$y_k = \sum_{i=1}^{n} a_i y_{k-i} + \sum_{j=0}^{m} b_i x_{k-j} + \varepsilon_k, \tag{2}$$

где a_i, b_j – весовые коэффициенты модели; y_{k-i}, x_{k-j} – наблюдаемые входная и выходная переменные; ε_k – случайная нерегистрируемая переменная, характеризующая шум на входе объекта.

Рассмотрим решение задачи на примере данных наблюдения, приведенных в табл. 1. Предположим, что размерность передаточной функции объекта определяется значениями $n = 2, m = 1$ и имеет вид

$$y(s) = \left[(s + \beta_1) \Big/ \prod_{j=1}^{2} (s + \alpha_j) \right] Kx(s) . \tag{3}$$

В этом случае регрессионную модель линейного объекта следует представить в виде конечно-разностного уравнения [7–9]:

$$\tilde{y}_k = B_1 \tilde{y}_{k-1} + B_2 \tilde{y}_{k-2} + B_3 \tilde{x}_k + B_4 \tilde{x}_{k-1} , \tag{4}$$

где B_1, B_2, B_3, B_4 – неизвестные весовые коэффициенты.

С использованием данных из табл. 1 и выбранной структуры регрессионной модели для решения задачи параметрической идентификации методом наименьших квадратов формируем матрицы **Y, W** [3, 5]:

W =

	1	2	3	4
8	3.90482	4.04451	0.97253	1.30167
9	3.26757	3.90482	0.81167	0.97253
10	2.37627	3.26757	0.87691	0.81167
11	1.51024	2.37627	1.1666	0.87691
12	0.92909	1.51024	1.62101	1.1666
13	0.80974	0.92909	2.13757	1.62101
14	1.20998	0.80974	2.59659	2.13757
15	2.05383	1.20998	2.89096	2.59659
16	3.15153	2.05383	2.95273	2.89096
17	4.24711	3.15153	2.76996	2.95273
18	5.08262	4.24711	2.3899	2.76996
19	5.46178	5.08262	1.90755	2.3899
20	5.29806	5.46178	1.4425	1.90755
21	4.63569	5.29806	1.10981	1.4425
22	3.64182	4.63569	0.99184	1.10981
23	2.56162	3.64182	1.11818	0.99184
24	1.66218	2.56162	1.45847	1.11818
25	1.16606	1.66218	1.92982	...

Y =

	1
8	3.26771
9	2.37637
10	1.51071
11	0.92939
12	0.81014
13	1.21061
14	2.05435
15	3.15174
16	4.24721
17	5.08257
18	5.46189
19	5.29794
20	4.6362
21	3.64191
22	2.56168
23	1.66236
24	1.1659
25	...

Матрица-столбец **Y** содержит значения наблюдаемой выходной переменной $\tilde{y}_k$ на временном интервале $[8\Delta t, 24\Delta t]$. Матрица **W** содержит 4 столбца. Два первых столбца содержат значения наблюдаемой выходной переменной $\tilde{y}_{k-1}, \tilde{y}_{k-2}$. Столбцы 3 и 4 матрицы **W** содержат значения входной наблюдаемой переменной $\tilde{x}_k, \tilde{x}_{k-1}$.

Между матрицами **Y**, **W** и вектором неизвестных весовых коэффициентов **B** регрессионной модели существует функциональная зависимость, которая определяется векторно-матричным уравнением:

$$\mathbf{Y} = \mathbf{WB} . \tag{5}$$

Решение уравнения (5) находится методом наименьших квадратов и для рассматриваемой постановки задачи регрессионного анализа определяется следующим образом [1–3]:

$$\mathbf{B} = (\mathbf{W}^T \cdot \mathbf{W})^{-1} \mathbf{W}^T \cdot \mathbf{Y}. \qquad (6)$$

Результаты решения задачи. Для наблюдаемых значений входной и выходной переменных, приведенных в табл. 1, выбранной структуры передаточной функции (3) вектор весовых коэффициентов регрессионной конечно-разностной модели имеет вид

$$\mathbf{B} = \begin{pmatrix} B_1 & B_2 & B_3 & B_4 \end{pmatrix} = \begin{pmatrix} 0,305 & -0,020 & -1,828 & 3,015 \end{pmatrix}.$$

Для оценки качества регрессионной модели значения выходной переменной рассчитывались по формуле

$$z_k = B_1 z_{k-1} + B_2 z_{k-2} + B_3 \tilde{x}_k + B_4 \tilde{x}_{k-1},\ z_2 = \tilde{y}_2,\ z_1 = \tilde{y}_1,\ k = 3, 4, ..., 25, \qquad (7)$$

на интервале времени $[8\Delta t, 24\Delta t]$.

Графики рассчитанной z_k и наблюдаемой переменной y_k приведены на рис. 2. Для лучшего восприятия рисунка график выходной наблюдаемой переменной y_k смещен на шаг дискретизации.

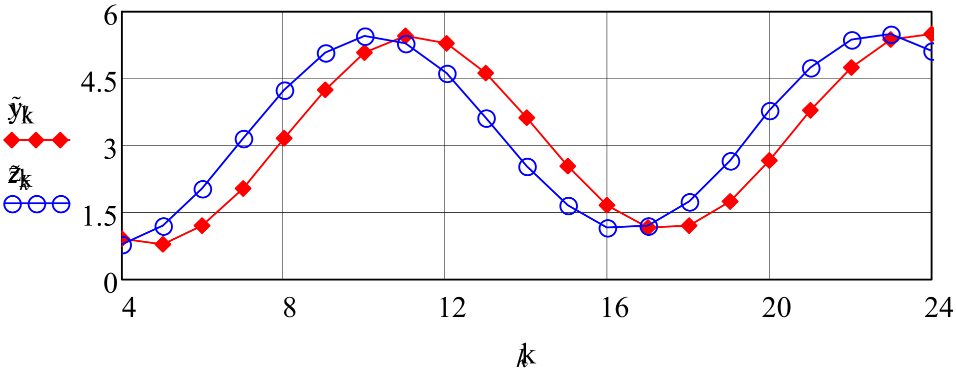

Рис. 2. Графики наблюдаемой $\tilde{y}_k$ и рассчитанной z_k выходной переменной

Количественно качество регрессионной модели экспериментально оценивалось путем расчета среднеквадратичного отклонения рассчитанных значений от наблюдаемых значений выходной переменной [10, 11]. Для регрессионной модели (7) расчеты выполнялись на ограниченном интервале по формуле

$$S_r = \sqrt{(1/22)\sum_{k=3}^{24}(z_k - \tilde{y}_k)^2}. \qquad (8)$$

В вычислительном эксперименте наблюдалось, что при изменении среднеквадратического отклонения случайной переменной ε_k от среднего значения на интервале $[0,1;\ 0,001]$ изменение среднеквадратического отклонения переменной z_k наблюдалось на интервале $[0,07;\ 0,001]$.

Рассмотрим построение дискретных и непрерывных моделей по результатам регрессионного анализа данных о наблюдаемых входной и выходной переменных. На основе заданной передаточной функции объекта (1, 3) можно составить систему уравнений [7, 8]:

$$\begin{cases} u(s) = \sum_{i=1}^{m} (s+\beta_i)\,x(s); \\ y(s) = Ku(s) \Big/ \sum_{j=1}^{n} (s+\alpha_j). \end{cases}$$

Примем $n = 2,\ m = 1$. Эта размерность модели достаточна, чтобы полно рассмотреть основные правила построения дискретной модели. В предположении, что корни вещественные, можно осуществить следующие обратные преобразования по Лапласу [4, 5]:

$$\frac{u_{k-1}^{2}}{(s+\alpha_1)} + \frac{Ku_k^1}{(s+\alpha_1)s} \to u_{k-1}^{2}\exp(\alpha_1\Delta t) + (\exp(\alpha_1\Delta t)-1)Ku_k^1 \big/ \alpha_1\,; \tag{9}$$

$$\frac{y_{k-1}}{(s+\alpha_2)} + \frac{u_k^2}{(s+\alpha_2)s} \to y_{k-1}\exp(\alpha_1\Delta t) + (\exp(\alpha_2\Delta t)-1)u_k^2 \big/ \alpha_2\,; \tag{10}$$

$$\frac{x_{k-1}}{(s+\beta)} + \frac{u_k^1}{(s+\beta)s} \to x_{k-1}\exp(\beta\Delta t) + (\exp(\beta\Delta t)-1)u_k^1 \big/ \beta\,. \tag{11}$$

На основе результатов обратного преобразования по Лапласу и выбранной формы представления дискретных моделей (2) выражение для дискретной модели, соответствующей (7)–(9), имеет вид

$$y_k = (a_1 + a_2)y_{k-1} - a_1 a_2 y_{k-2} + b_1 b_2 d_1 x_k - b_1 b_2 d_0 x_{k-1}\,,$$

где

$$a_1 = \exp(-\alpha_1\Delta t)\,,\quad a_2 = \exp(-\alpha_2\Delta t)\,;$$

$$b_1 = K(\exp(-\alpha_1\Delta t)-1)/\alpha_1\,,\ b_2 = (\exp(-\alpha_2\Delta t)-1)/\alpha_2\,;$$

$$d_0 = \beta\exp(-\beta_1\Delta t)\,/(\exp(-\beta_1\Delta t)-1)\,,\ d_1 = \beta/(\exp(-\beta_1\Delta t)-1)\,. \tag{12}$$

Из выражений (7), (12) следует, что между весовыми коэффициентами регрессионной модели (5) и параметрами передаточной функции (3) существуют функциональные связи [9], которые определяются системой уравнений:

$$\begin{cases} B_1 = \exp(\alpha_1\Delta t) + \exp(\alpha_2\Delta t), \\ B_2 = -\exp(\alpha_1\Delta t)\exp(\alpha_2\Delta t), \\ B_3 = \dfrac{K(\exp(\alpha_1\Delta t-1)}{\alpha_1}\cdot\dfrac{(\exp(\alpha_2\Delta t-1)}{\alpha_2}\cdot\dfrac{\beta}{\exp(-\beta\Delta t)-1}, \\ B_4 = -\dfrac{K(\exp(\alpha_1\Delta t-1)}{\alpha_1}\cdot\dfrac{(\exp(\alpha_2\Delta t-1)}{\alpha_2}\cdot\dfrac{\exp(-\beta\Delta t)\beta}{\exp(-\beta\Delta t)-1}. \end{cases} \tag{13}$$

Система уравнений (13) содержит 4 неизвестных параметра передаточной функции: $\alpha_1, \alpha_2, \beta, K$. Решая эту систему уравнений, получим значения параметров, определяющих весовые коэффициенты непрерывной модели и дискретных моделей с различными шагами дискретизации:

$$\alpha_1 = -5{,}001,\ \alpha_1 = -3{,}001,\ \beta = -1{,}001,\ K = 25{,}001.$$

Обсуждение

В соответствии с разработанной процедурой оценка параметров модели динамического объекта должна выполняться в следующей последовательности.

Анализируются данные. По результатам анализа определяются структура регрессионной модели и временной интервал для расчета параметров этой модели.

Выполняется оценка значений параметров регрессионных моделей. В предположении $1 \le n \le 4,\ 0 \le m \le n-1$ определяются параметры группы регрессионных моделей различной размерности (n, m):

$$\tilde{y}_k = a_1\tilde{y}_{k-1} + \ldots + a_n\tilde{y}_{k-n} + b_0\tilde{x}_k + \ldots + b_n\tilde{x}_{k-m}.$$

Для каждой регрессионной модели размерности n, m рассчитывается среднеквадратическое отклонение рассчитанных значений выходной переменной z_k от наблюдаемых значений $\tilde{y}_k$:

$$S_{n,m} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(z_k - \tilde{y}_k)^2}.$$

Значения выходной переменной рассчитывались по формуле

$$z_k = a_1 z_{k-1} + \ldots + a_n z_{k-n} + b_0 x_k + \ldots + b_n x_{k-m},\ k = 2,3,\ldots,N; z_0 = \tilde{y}_0, z_1 = \tilde{y}_1.$$

Из множества значений $S_{n,m}$ размерности $n \times m$ находится минимальное значение среднеквадратичного отклонения и определяется количество полюсов и нулей передаточной функции.

Решается система нелинейных уравнений (13) и определяются значения весовых коэффициентов ($\alpha_1, \alpha_2, \beta, K$), необходимых для построения непрерывных моделей и дискретных моделей с различным шагом дискретизации.

Заключение

Разработана процедура идентификации дискретной и непрерывной моделей динамического объекта по результатам наблюдения за входной и выходной переменной объекта в течение интервала времени. Процедура может быть использована для построения адекватных регрессионных, дискретных и непрерывных моделей. Для обеспечения адекватности моделей процедура предполагает согласование структуры регрессионной модели со структурой непрерывной (дискретной) модели исходя из предполагаемого состава полюсов и нулей. При отсутствии информации о структуре передаточной функции объекта аналитик вынужден осуществить построение группы моделей при различных составах полюсов и нулей передаточной функции. Для выбора модели с наибольшей адекватностью можно использовать методы дисперсионного и корреляционного анализа. Количество нулей и полюсов определяется из условия минимума дисперсии (среднеквадратичного отклонения) значений рассчитанных значений от наблюдаемых значений выходной переменной. Оптимальное количество полюсов и нулей передаточной функции объекта возможно определить методом полного перебора допустимых вариантов.

Библиографический список

1. **Montgomery, D. C.** Introduction to Linear Regression Analysis, 5th Edition / D. C. Montgomery, E. A. Peck, G. G. Vining. – John Wiley & Sons, Inc, 2012. – 672 p.
2. **Giri, I.** Procedure and interpretation of linear regression analysis using STATA / I. Giri, P. Chetty. – 2017. – URL: www.projectguru.in
3. **Isermann, R.** Identification of Dynamic Systems. An Introduction with Applications / R. Isermann, M. Műnchhof. – Berlin : Springer, 2010. – 710 p.
4. **Mangan, N. M.** Model selection for dynamical systems via sparse regression and information criteria / N. M. Mangan, J. N. Kutz, S. L. Brunton, J. L. Proctor // Proceedings of the royal society a. Mathematical, physical and engineering sciences. – 2017. – Vol. 473 (2204).
5. **Макарычев, П. П.** Построение модели системной динамики по результатам регрессионного анализа / П. П. Макарычев // Аналитические и численные методы моделирования естественно-научных и социальных проблем : материалы XII Междунар. науч.-техн. конф. – Пенза, 2017. – С. 113–117.
6. **Дилигенская, А. Н.** Идентификация объектов управления / А. Н. Дилигенская. – Самара : Самар. гос. техн. ун-т, 2009. – 136 с.
7. **Волгина, М. А.** Моделирование многокомпонентных систем на основе маркированных графов : монография / М. А. Волгина, П. П. Макарычев. – Пенза : Изд-во ПГУ, 2011. – 156 с.
8. **Макарычев, П. П.** Моделирование непрерывных и дискретных динамических систем : учеб. пособие / П. П. Макарычев ; под ред. Н. Н. Вашкевича. – Пенза : Изд-во Пенз. политех. ин-та, 1988. – 76 с.
9. **Макарычев, П. П.** Анализ функциональных зависимостей методами системной динамики и регрессионного анализа / П. П. Макарычев, А. Ю. Афонин, С. В. Шибанов // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2019. – № 2 (50). – С. 11–22.
10. **Дрейпер, Н.** Прикладной регрессионный анализ. Множественная регрессия / Н. Дрейпер, Г. Смит // Applied Regression Analysis. – 3-е изд. – Москва : Диалектика, 2007. – С. 912.
11. **Семенычев, В. К.** Параметрическая идентификация рядов динамики: структуры, модели, эволюция : монография / В. К. Семенычев, Е. В. Семенычев. – Самара : Изд-во СпмНЦ РАН, 2011. – 364 с.

References

1. Montgomery D. C., Peck E. A., Vining G. G. *Introduction to Linear Regression Analysis, 5th Edition*. John Wiley & Sons, Inc, 2012, 672 p.
2. Giri I., Chetty P. *Procedure and interpretation of linear regression analysis using STATA*. 2017. Available at: www.projectguru.in
3. Isermann R., Műnchhof M. *Identification of Dynamic Systems. An Introduction with Applications*. Berlin: Springer, 2010, 710 p.

4. Mangan N. M., Kutz J. N., Brunton S. L., Proctor J. L. *Proceedings of the royal society a. Mathematical, physical and engineering sciences*. 2017, vol. 473 (2204).
5. Makarychev P. P. *Analiticheskie i chislennye metody modelirovaniya estestvenno-nauchnykh i sotsial'nykh problem: materialy XII Mezhdunar. nauch.-tekhn. konf.* [Analytical and numerical methods for modeling natural-scientific and social issues: proceedings of XII International scientific and practical conference]. Penza, 2017, pp. 113–117. [In Russian]
6. Diligenskaya A. N. *Identifikatsiya ob"ektov upravleniya* [Identification of management objects]. Samara: Samar. gos. tekhn. un-t, 2009, 136 p. [In Russian]
7. Volgina M. A., Makarychev P. P. *Modelirovanie mnogokomponentnykh sistem na osnove markirovannykh grafov: monografiya* [Simulation of multicomponent systems based on marked graphs: monograph]. Penza: Izd-vo PGU, 2011, 156 p. [In Russian]
8. Makarychev P. P. *Modelirovanie nepreryvnykh i diskretnykh dinamicheskikh sistem: ucheb. posobie* [Modeling of continuous and discrete dynamical systems: teaching aid]. Penza: Izd-vo Penz. politekh. in-ta, 1988, 76 p. [In Russian]
9. Makarychev P. P., Afonin A. Yu., Shibanov S. V. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* [University proceedings. Volga region. Engineering sciences]. 2019, no. 2 (50), pp. 11–22. [In Russian]
10. Dreyper N., Smit G. *Applied Regression Analysis*. 3d ed. Moscow: Dialektika, 2007, pp. 912.
11. Semenychev V. K., Semenychev E. V. *Parametricheskaya identifikatsiya ryadov dinamiki: struktury, modeli, evolyutsiya: monografiya* [Parametric identification of the dynamics' series: structures, models, evolution: monograph]. Samara: Izd-vo SpmNTs RAN, 2011, 364 p. [In Russian]

Безяев Виктор Степанович
кандидат технических наук, доцент, кафедра математического обеспечения и применения ЭВМ, Пензенский государственный университет (Россия, г. Пенза, ул. Красная, 40)

E-mail: pm@pnzgu.ru

Bezjaev Viktor Stepanovich
Candidate of engineering sciences, associate professor, sub-department of software and computer applications, Penza State University (40, Krasnaya street, Penza, Russia)

Макарычев Петр Петрович
доктор технических наук, профессор, заведующий кафедрой математического обеспечения и применения ЭВМ, Пензенский государственный университет (Россия, г. Пенза, ул. Красная, 40)

E-mail: mpp@pnzgu.ru

Makarychev Pjotr Petrovich
Doctor of engineering sciences, professor, head of the sub-department of software and computer applications, Penza State University (40, Krasnaya street, Penza, Russia)

Образец цитирования:

Безяев, В. С. Идентификация параметров моделей объектов методом регрессионного анализа / В. С. Безяев, П. П. Макарычев // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2020. – № 1 (53). – С. 19–27. – DOI DOI 10.21685/2072-3059-2020-1-2.