

# ПОСТРОЕНИЕ МОДЕЛИ ЧИСЛОВЫХ ДАННЫХ МЕТОДОМ СИНГУЛЯРНОГО СПЕКТРАЛЬНОГО АНАЛИЗА В ИМИТАЦИОННОМ МОДЕЛИРОВАНИИ

© Е.А. Якимов<sup>1</sup>, Е.М. Борчик<sup>2</sup>, А.И. Якимов<sup>3</sup>

Белорусско-Российский университет,  
212030, Беларусь, г. Могилев, пр. Мира, 43.

Представлены методика и результаты исследования погрешности восстановления аддитивных составляющих временного ряда методом сингулярного спектрального анализа. Даны рекомендации по практическому применению сингулярного спектрального анализа в имитационном моделировании на этапе эксплуатации имитационных моделей. Для построения аналитической модели шумовой составляющей предложена методика построения функции плотности распределения семейства Пирсона. Представлена технология применения сингулярного спектрального анализа и программного средства для определения типа шумовой составляющей временного ряда.

Ил. 2. Библиогр. 9 назв.

*Ключевые слова:* сингулярный спектральный анализ; имитационная модель; комплексная информационная система; временной ряд.

## BUILDING NUMERICAL DATA MODEL BY SINGULAR SPECTRUM ANALYSIS METHOD IN SIMULATION

E.A. Yakimov, E.M. Borchik, A.I. Yakimov

Belarusian-Russian University,  
43 Mira Av., Mogilev, Belarus, 212030.

The paper presents the procedure and results of studying errors of time series additive component recovery by the method of singular spectrum analysis. The recommendations on the practical use of the singular spectrum analysis in simulation at the stage of simulation model operation are given. The methods of the Pearson family density distribution function are proposed for building the analytical model of a noise component. The application technology of singular spectrum analysis and software to determine the type of the noise component of the time series is presented.

2 figures. 9 sources.

*Key words:* singular spectrum analysis; simulation model; integrated information system; time series.

**Введение.** В настоящее время актуальным направлением совершенствования деятельности предприятия является использование имитационной модели (ИМ) в контуре управления. Построение имитационной модели реализуется последовательным выполнением следующих этапов [4]: разработка содержательного описания исследуемой системы; разработка концептуальной

модели; разработка формальной и построение имитационной модели системы; программирование и отладка имитационной модели; испытание и исследование модели в составе комплексной информационной системы (КИС) предприятия; эксплуатация имитационной модели, предполагающая длительное использование ИМ в контуре управления промышленным предприятием. Эксплуатация

<sup>1</sup>Якимов Евгений Анатольевич, ассистент кафедры автоматизированных систем управления, тел.: 375293124194, e-mail: e-soft@bk.ru

Yakimov Evgeny, Assistant Lecturer of the Department of Automated Control Systems, tel.: +375293124194, e-mail: e-soft@bk.ru

<sup>2</sup>Борчик Екатерина Михайловна, аспирант, тел.: 375297468310, e-mail: katrinb15@gmail.com

Borchik Ekaterina, Postgraduate, tel.: 375297468310, e-mail: katrinb15@gmail.com

<sup>3</sup>Якимов Анатолий Иванович, кандидат технических наук, доцент кафедры автоматизированных систем управления, тел.: 375447163816, e-mail: ykm@tut.by

Yakimov Anatoly, Candidate of technical sciences, Associate Professor of the Department of Automated Control Systems, tel.: 375447163816, e-mail: ykm@tut.by



имитационных моделей в производственных условиях затруднена ввиду отсутствия методов обработки накопленных данных и подготовки их для использования в моделях. Эффективная подготовка входных данных для имитационного моделирования становится возможной при использовании данных, накопленных в информационных системах организаций [8].

В имитационном моделировании в основном используются статистические методы оценки последовательностей данных. В [6] представлен анализ программных средств для определения закона распределения случайных величин, используемых для ввода в имитационную модель, например, программа StatFit имеет 32 различных распределения, BestFit имеет 37 распределений вероятности, ExpertFit [9] имеет 40 распределений. Для наиболее точного определения типа распределения данных используются обобщенные законы распределения Пирсона, Джонсона и др. В последнее время все более часто используются методы, основанные на искусственных нейронных сетях, с применением нечетких множеств и нечеткой логики, кластерном анализе и др. В то же время отсутствуют сведения об использовании в имитационном моделировании сингулярного спектрального анализа (другое название SSA – Singular Spectrum Analysis) последовательностей числовых данных, который позволяет получить дополнительную информацию о структуре данных разложением исходной последовательности на трендовую, периодическую и шумовую составляющие [7].

**Содержательное описание сингулярного спектрального анализа.** Данные в комплексной информационной системе, представленные вещественными числами и определенные последовательно во времени, именуют временной последовательностью данных. В общем случае такую последовательность числовых данных, обычно упорядоченных во времени, называют временным рядом. При исследовании SSA трендом полагают медленно меняющуюся составляющую, описывающую глобальное поведение ряда. Под шумовой составляющей ряда понимают реализацию случайной последовательности чисел. Сингулярный спектральный анализ временных рядов включает следующие этапы: вложение, сингулярное разложение, группировку, диагональное усреднение.

Чтобы выделить аддитивную составляющую ряда  $G$  длины  $n$  ( $G = G_T + G_H + G_N$ , причем  $G_T$ ,  $G_H$  и  $G_N$  – соответственно трендовая, гармоническая и шумовая составляющие) SSA-методом по ряду  $G$  строится траекторная матрица  $A$  заданного размера  $L \times K$ ,  $1 < L < n$ ,  $K = n - L + 1$  ( $L$  называется длиной окна), которая по правилам построения является ганкелевой.

Для матрицы  $S = A \cdot A^T$  вычисляются решением характеристического уравнения  $S - \mu \cdot E = 0$  собственные числа  $\{\mu_k\}_{k=1}^L$ , а также левые  $\{U_k\}_{k=1}^L$  и

правые ортонормированные собственные вектора  $\{V_k\}_{k=1}^L$ . Матрица  $A$  сингулярным разложением представляется в виде  $A = \sum_k A_k$ ,  $A_k = \sqrt{\mu_k} \cdot U_k \cdot V_k^T$ . Обозначим корень собственного числа матрицы  $S$  через  $\lambda_k = \sqrt{\mu_k}$ .

Наборы  $\langle \lambda_k, U_k, V_k \rangle$  группируются для получения трендовой, гармонической и шумовой составляющих. Соответственно определяются матрицы  $A_T = \sum_{k \in T} A_k$ ,  $A_H = \sum_{k \in H} A_k$ ,  $A_N = \sum_{k \in N} A_k$ .

После диагонального усреднения матриц  $A_T$ ,  $A_H$ ,  $A_N$  получают восстановленные аддитивные составляющие  $G_{T\theta}$ ,  $G_{H\theta}$ ,  $G_{N\theta}$ . Если  $G_T$ ,  $G_H$  и  $G_N$  заданы известными функциями соответственно  $F_T$ ,  $F_H$  и известен закон распределения  $G_N$ , то можно оценить показатели ошибки восстановления  $\Delta F_T = F_T - G_{T\theta}$ ,  $\Delta F_H = F_H - G_{H\theta}$  и тип распределения восстановленной шумовой составляющей  $G_{N\theta}$ . Показатели погрешности восстановления определяются характеристиками положения: математическое ожидание, оцениваемое средним (Mean), медиана (Med); характеристиками рассеяния: среднеквадратическое отклонение, оцениваемое стандартным отклонением (s), максимум (Max), минимум (Min), диапазон (Range) [1].

**Методика исследования временных рядов.** Исследование сингулярного спектрального анализа временных рядов с комплексным применением информационных технологий основано на использовании табличного процессора MS Excel, математического пакета Mathcad, пакета статистической обработки данных Statistica и отличается применением сдвиговых процедур при формировании ганкелевой матрицы на этапе вложения и при восстановлении составляющих исходного временного ряда на этапе диагонального усреднения [5]. На этапе группировки наборов  $\langle \lambda_k, U_k, V_k \rangle$  в MS Excel используется лепестковая диаграмма, которая является аналогом графика в полярной системе координат, отображая распределение значений относительно начала координат. По особенностям представления сингулярных векторов на лепестковой диаграмме принимается решение о принадлежности их одной группе.

Для оценивания результатов исследования каждой из составляющих временного ряда используется сумма модулей значений элементов временного ряда, которая именуется Ф-оценкой соответственно для трендовой, гармонической и шумовой составляющих:

$$\Phi_M = \sum_{i=1}^n |g_{Mi}|, M = T, H, N,$$



где  $g_{Mi}$  – значение  $i$ -го элемента  $M$ -ой составляющей временного ряда;  $n$  – длина временного ряда; индексы  $T, H, N$  – соответственно трендовая, гармоническая и шумовая составляющие.

При анализе временных рядов и сравнении результатов SSA-преобразований используются относительные  $\Phi$ -оценки, определяемые соответственно для трендовой, гармонической и шумовой составляющих:

$$\varphi_M = \Phi_M / \min(\Phi_T, \Phi_H, \Phi_N) \mid M = T, H, N.$$

Принятые  $\Phi$ -оценки характеризуют соотношение между значениями составляющих временного ряда. Аналогично  $\Phi$ -оценкам и относительным  $\Phi$ -оценкам составляющих временного ряда приняты оценки для сингулярной последовательности корней собственных чисел:

$$\Phi_{\lambda M} = \sum_{i=1}^k \lambda_i, \mid M = T, H, N;$$

$$\varphi_{\lambda M} = \Phi_{\lambda M} / \min(\Phi_{\lambda T}, \Phi_{\lambda H}, \Phi_{\lambda N}) \mid M = T, H, N,$$

где  $k$  – количество корней собственных чисел.

В теории сингулярного спектрального анализа показано, что любой ряд, являющийся суммой экспонент и полиномов, асимптотически отделим от периодического ряда; любой гармонический ряд с частотой  $\omega_1$  асимптотически отделим от другого гармонического ряда с частотой  $\omega_2 \neq \omega_1$ ; это же верно и для экспоненциально-модулированных гармоник. Эти два результата дают возможность, в частности, приближенно отделять тренд (понимаемый как медленно меняющаяся аддитивная составляющая ряда) от периодических компонент и от шума, а также разделять периодические компоненты на их гармонические составляющие. Однако для имитационного моделирования случайных процессов отсутствуют результаты по определению типа распределения шумовой составляющей при сингулярном спектральном анализе. Для восполнения этого пробела проведены исследования с наиболее распространенной полиномиальной и гармонической составляющими при различных типах шумовой составляющей.

В практике имитационного моделирования случайные последовательности определяются типами непрерывных распределений, относящихся к различным классам. Например, при нормальном распределении случайная величина принимает значения на положительной и отрицательной полуоси, экспоненциальное распределение относится к классу распределений с возможными значениями на положительной полуоси, равномерное – к классу распределений с возможными значениями на ограниченном интервале.

**Методика построения функции плотности распределения семейства Пирсона.** Пусть в ходе имитационных экспериментов получена последовательность данных, представленная выборкой:

$$X = \{x_i \mid x_i \in R, i = 1, \dots, n\}.$$

Необходимо построить функцию плотности

распределения (кривую), наилучшим образом описывающую выборку  $X$  на заданном интервале  $[a, b]$ .

Предлагается применение плотностей обобщенного распределения Пирсона с последующей проверкой комплексом из трех статистических критериев (Пирсона, Колмогорова-Смирнова, Мизеса) соответствия закону распределения с построенной плотностью. Семейство кривых Пирсона включает семь основных типов и три частных случая распределений: равномерное, нормальное, экспоненциальное.

Методика построения функции плотности распределения семейства Пирсона, соответствующей эмпирическому распределению выборки  $X$ , реализована в программном модуле «BelSim2#.random» [3] и состоит из следующих этапов:

**Этап 1.** Определение точечных оценок выборки  $X$ .

Рассчитываются следующие точечные оценки выборки  $X$ : минимальное и максимальное значения выборки  $Min, Max$ ; диапазон  $Range$ ; начальный момент первого порядка  $v_1$ ; центральные моменты  $\mu_2, \dots, \mu_4$ ; дисперсия  $D = \mu_2$ ; среднеквадратическое отклонение  $\sigma = \sqrt{D}$ ; коэффициенты асимметрии и эксцесса, введенные Пирсоном,  $\beta_1 = \mu_3^2 / \mu_2^3$ ,  $\beta_2 = \mu_4 / \mu_2^2$ , а также в соответствии со статистическим пакетом STATISTICA:  $\gamma_3 = \sqrt{\beta_1}$ ,  $\gamma_4 = \beta_2 + 3$ .

**Этап 2.** Построение функции  $f(x)$  плотности распределения Пирсона в соответствии с классификацией.

**Шаг 2.1.** Классификация типа кривой плотности распределения Пирсона.

Для определения типа кривой плотности распределения Пирсоном рассматривается показатель

$$\aleph = (\beta_1(\beta_2 + 3)^2) / (4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1)). \quad (1)$$

Значение показателя  $\aleph < 0$  соответствует типу I;  $\aleph = 0$  – типам II, VII;  $0 < \aleph < 1$  – определяет тип IV;  $\aleph = 1$  – тип V;  $\aleph > 1$  – тип VI;  $\aleph = \pm\infty$  тип III. Для типа II выполняются условия  $\beta_1 = 0$ ,  $\beta_2 \neq 3$ ; для типа VII –  $\beta_1 = 0$ ,  $\beta_2 = 3$ . При  $(\beta_1, \beta_2) = (0; 1,8)$  имеет место равномерное распределение ( $R(x)$ ); при  $(\beta_1, \beta_2) = (0, 3)$  – нормальное ( $N(x)$ ); при  $(\beta_1, \beta_2) = (4, 9)$  – экспоненциальное ( $E(x)$ ). Частный случай распределения для типа I – бета-распределение 1-го рода; для типа II – равномерное распределение; для типа III – гамма-распределение и распределение хи-квадрат; для распределений типа VI – бета-распределение 2-го рода и распределение Фишера-Снедекора; для типа VII – распределение Стьюдента.



Для оценки значений параметров функции плотности распределения  $f(x)$  применяется классический метод моментов. Все значения параметров функции  $f(x)$  выражаются через центральные моменты  $\mu_2, \dots, \mu_4$  и начальный момент первого порядка  $\nu_1$ , за исключением вырожденных распределений  $R(x)$ ,  $N(x)$ ,  $E(x)$ , задаваемых меньшим количеством моментов. Для построенной функции  $f(x)$  в соответствии с методикой, предложенной Пирсоном, определяется нормирующий множитель  $N$ , расчет которого выполняется исходя из условия

$$N \cdot \int_a^b f(x) = 1, \quad (2)$$

где  $[a, b]$  – интервал, в пределах которого производится построение кривой.

**Шаг 2.2.** Проверка гипотезы принадлежности выборки  $X$  закону распределения с плотностью  $f(x)$ .

Для проверки гипотезы выбраны следующие критерии согласия:  $\chi^2$  Пирсона,  $\lambda$  Колмогорова-Смирнова,  $\omega^2$  Мизеса. Вначале к  $(X, f(x))$  применяются критерии  $\chi^2$  и  $\lambda$ , сохраняются результаты их работы. Если логические значения результатов работы критериев  $\chi^2$  и  $\lambda$  эквивалентны, то на этом этап статистической проверки гипотезы заканчивается.

Иначе – дополнительно применяется критерий  $\omega^2$ , результат работы которого принимается в качестве заключения о проверке гипотезы.

Сохраняемые результаты работы критериев  $\chi^2$ ,  $\lambda$ ,  $\omega^2$ :

1) наблюдаемые значения критериев:

$$\Delta_{\chi^2} \in R_+ \cup \{0\},$$

$$\Delta_{\lambda} \in R_+ \cup \{0\},$$

$$\Delta_{\omega^2} \in R_+ \cup \{0\};$$

2) критические значения критериев:

$$\Delta_{\chi^2}^{kp} \in R_+, \Delta_{\lambda}^{kp} \in R_+, \Delta_{\omega^2}^{kp} \in R_+;$$

3) отношение наблюдаемых значений критериев к критическим:

$$dL_1 = \Delta_{\chi^2} / \Delta_{\chi^2}^{kp}, dL_2 = \Delta_{\lambda} / \Delta_{\lambda}^{kp}, dL_3 = \Delta_{\omega^2} / \Delta_{\omega^2}^{kp}; \quad (3)$$

4) логические результаты работы критериев:

$$bL_1 = \begin{cases} \text{true}, & dL_1 < 1; \\ \text{false}, & dL_1 \geq 1, \end{cases}$$

$$bL_2 = \begin{cases} \text{true}, & dL_2 < 1; \\ \text{false}, & dL_2 \geq 1, \end{cases}$$

$$bL_3 = \begin{cases} \text{true}, & dL_3 < 1; \\ \text{false}, & dL_3 \geq 1. \end{cases} \quad (4)$$

**Этап 3.** В случае отклонения на этапе 2 статистическими критериями гипотезы о принадлежности выборки  $X$  закону распределения с плотностью  $f(x)$ , производится перебор с последующей проверкой комплексом статистических критериев  $\chi^2$ ,  $\lambda$ ,  $\omega^2$  всех основных типов кривых семейства за исключением самой  $f(x)$ , проверенной на этапе 2.

**Этап 4.** По запросу пользователя выполняется построение функций распределений для трех частных случаев (равномерного, нормального, экспоненциального) с последующей проверкой комплексом статистических критериев  $\chi^2$ ,  $\lambda$ ,  $\omega^2$ .

**Этап 5.** По результатам работы комплекса статистических критериев производится выбор кривой плотности распределения, наилучшим образом описывающей выборку.

Выбор производится на множестве тех кривых  $\{f_i(x)\}$ ,  $i = 1, \dots, |f_i(x)|$ , которые не отклонены статистическими критериями, т.е. для результатов (4) выполняется одно из условий:

$$bL_1[i] \wedge bL_2[i] = \text{true}, \quad (5)$$

$$(bL_1[i] \vee bL_2[i]) \wedge bL_3[i] = \text{true}, \quad (6)$$

где  $[i]$  – номер функции  $f_i(x)$ ;  $bL_1, \dots, bL_3$  – результаты проверки  $f_i(x)$  по (4).

Функция плотности распределения  $f^*(x)$  наилучшим образом описывает на интервале  $[a, b]$  выборку  $X$  тогда и только тогда, когда выполняется условие

$$(f^*(x) = f_{i_0}(x) \in \{f_i(x)\}) [\rho(X, f_{i_0}(x)) = \min\{\rho(X, f_i(x)) \mid i = 1, \dots, |f_i(x)|\}], \quad (7)$$

где  $\rho(X, f_i(x))$  – евклидова метрика вида (8), если для всех  $i = 1, \dots, |f_i(x)|$  выполняется условие (5), т.е. применялись только 2 критерия:  $\chi^2$  и  $\lambda$ :

$$dL = \rho(X, f_i(x)) = \sqrt{(dL_1[i])^2 + (dL_2[i])^2}, \quad i = 1, \dots, |f_i(x)|, \quad (8)$$

либо  $\rho(X, f_i(x))$  – евклидова метрика вида (9), если существуют такие функции  $f_i(x)$ , для которых выполняется условие (6) и при этом не выполняется условие (5), т.е. применялись 3 критерия:  $\chi^2$ ,  $\lambda$ ,  $\omega^2$ :

$$dL = \rho(X, f_i(x)) = \sqrt{(dL_1[i])^2 + (dL_2[i])^2 + (dL_3[i])^2}, \quad i = 1, \dots, |f_i(x)|. \quad (9)$$

Для расчета значений евклидовой метрики (9) по результатам  $dL_1, \dots, dL_3$  вида (3) необходимо



произвести дополнительный расчет 3-го критерия ( $\omega^2$ ) для всех таких функций  $f_i(x)$ ,  $i = 1, \dots, |f_i(x)|$ , для которых выполняется условие (5).

В том случае, если все построенные кривые отклонены статистическими критериями, нет возможности построения кривой плотности распределения семейства Пирсона, описывающей данную выборку. В этом случае рекомендуется исключение выбросов и/или разделение исходной выборки на несколько однородных выборок с последующим построением на каждой из них своей функции плотности распределения.

Для построения функции плотности распределения, наилучшим образом описывающей выборку  $X$  на интервале  $[a, b]$ , проводится эксперимент с использованием программного модуля «BelSim2#.random», на вход которого подается файл model.xml с исходными данными, значениями параметров и списком откликов модели.

#### Результаты исследований и их обсуждение.

При исследовании методом сингулярного спектрального анализа временных рядов показано на примере для равномерного, экспоненциального и нормального распределений, что шумовая составляющая восстановленного ряда смещается в сторону нуля на величину математического ожидания шумовой составляющей исходного ряда с погрешностью до 6% и возрастает до 30% при росте дисперсии шума исходного ряда до 25 раз. Стандартное отклонение шумовой составляющей восстановленного ряда отличается от исходного на 7–9% во всех случаях.

Показано, что восстановление гармонической составляющей затруднено при наличии значительной шумовой составляющей. Например, в исследованиях рассматривался временной ряд с соотношением  $\Phi_{\Delta T} : \Phi_{\Delta H} : \Phi_{\Delta N} = 28,6 : 1,0 : 10,5$ , т.е. шум на порядок превышает гармоническую составляющую. При этом восстанавливается лишь количество периодов и максимальная абсолютная ошибка превышает амплитуду гармонической составляющей более чем в два раза.

Исследовано влияние процедуры сжатия/растяжения (пропорциональное уменьшение/увеличение значений элементов ряда путем умножения на коэффициент  $1/\alpha$  ( $\alpha > 1$ )) ( $0 < \alpha < 1$ ) временного ряда на качество восстановления его составляющих методом сингулярного спектрального анализа. Показано, что изменение элементов временного ряда  $G$  пропорционально величине  $\alpha \in \mathbb{R}$  изменяет показатели (характеристики положения: среднее (Mean), медиана (Med); характеристики рассеяния: стандартное отклонение (s); максимум (Max); минимум (Min), диапазон (Range)) погрешности восстановления аддитивных трендовой  $G_T$ , гармонической  $G_H$  и шумовой  $G_N$  составляющих,  $G = G_T + G_H + G_N$ , методом сингулярного спектрального анализа также пропорционально величине  $\alpha$ .

Таким образом, можно утверждать об исследова-

нии нормированного временного ряда, определяемого по формуле  $G_n = \beta \cdot [G - Mean(G)] / \alpha + \gamma$ ,  $Mean(G)$  – среднее элементов  $x_i$ ,  $i = 0, \dots, n-1$  временного ряда  $G$ ,  $1/\alpha$  – коэффициент сжатия/растяжения, при этом параметр  $\alpha$  определяется наибольшим по модулю значением элемента ряда  $G - Mean(G)$ ;  $\beta$  – нормирующий параметр масштаба, определяемый верхней границей исследуемых значений элементов нормированного временного ряда, как правило,  $\beta = 1$ ;  $\gamma$  – параметр сдвига элементов временного ряда в область положительных вещественных чисел,  $\gamma \geq \beta$  [2].

Показано, что шумовая составляющая сохраняет свой тип распределения при восстановлении, что позволяет использовать полученные результаты для построения модели входных данных при имитационном моделировании случайных процессов. Проверка соответствия полученных в ходе преобразования случайных чисел известному типу распределения выполнена по критерию Колмогорова – Смирнова ( $p = 0,01$ ). В то же время проведенные численные исследования с ограниченным классом функций не позволяют выработать общие рекомендации для всех классов детерминированных функций при определении типа распределения шумовой составляющей. Поэтому предложено проводить исследования числовых последовательностей данных с применением сингулярного спектрального анализа, приводя их к нормированному виду путем операций сжатия/растяжения и смещения, и с использованием обобщенных законов распределения случайных величин, что обеспечивает построение банка знаний по исследованию временных последовательностей данных методом сингулярного спектрального анализа.

**Технология применения SSA-преобразования и программного средства для определения типа шумовой составляющей.** Исходный временной ряд получен в ходе проведения имитационных экспериментов:

$G = \langle 95,0; 95,6; 97,1; 97,5; 98,0; 98,5; 99,0; 98,2; 99,2; 98,7; 99,3; 99,5; 100,0; 100,5; 101,0; 100,2; 99,4; 99,8; 101,2; 99,4; 102,0; 103,4; 104,4; 105,5; 104,0; 105,0; 104,2; 105,2; 104,3; 105,4; 105,6; 106,6; 107,6; 105,8; 106,8; 107,2; 107,1; 106,1; 105,6; 106,0; 107,6; 107,7; 109,0 \rangle$ .

Для решения задачи построения модели последовательности числовых данных применяется программный модуль BelSim2#.random, в списке параметров которого установлены флаги обязательной проверки всех функций семейства Пирсона и трех частных случаев: равномерного ( $R(x)$ ), нормального ( $N(x)$ ), экспоненциального ( $E(x)$ ).

С применением программы BelSim2#.random рассчитаны точечные оценки выборки  $X$ , в частности:  $Min = 95,00$ ,  $Max = 109,00$ ,  $v_1 = 102,54$ ,  $\mu_2 = 14,295$ ,  $\mu_3 = -7,698$ ,  $\mu_4 = 355,460$ ,  $\beta_1 = 0,020$ ,  $\beta_2 = 1,74$ ,  $\varkappa = -0,0064$  (1). Для вы-



борочных данных построены функции плотностей распределений  $f_5(x)$ ,  $f_7(x)$ ,  $R(x)$ ,  $E(x)$ ,  $N(x)$ .

С заданной точностью  $\varepsilon = 0,1$  критерий классификации типов кривых Пирсона  $\aleph = 0$ .

Результат работы программы BelSim2#.random:  
 $f_5(x) = 6,244E30 \cdot (x - 88,500)^{-20,020} \cdot \exp(-281,084 / (x - 88,500))$ ;  
 $f_7(x) = 0,114 \cdot \exp(-(x - 102,540)^2 / 28,590)$ ;  
 $R(x) = 0,071$ ;  
 $E(x) = 9,752E-3 \cdot \exp(-1 \cdot 9,752E-3 \cdot x)$ ;  
 $N(x) = 1 / (3,781 \cdot (2 \cdot \pi)^{1/2}) \cdot \exp(-1 \cdot (x - 102,54)^2 / 28,590)$ .

При проверке критериями  $\chi^2$ ,  $\lambda$ ,  $\omega^2$  гипотезы о принадлежности выборки  $X$  законам распределения с построенными плотностями  $f_5(x)$ ,  $N(x)$ ,  $E(x)$  отклонены, для отклонения гипотез о распределении выборок по законам с распределениями  $f_7(x)$ ,  $R(x)$  нет оснований. Так как для функции  $R(x)$  значение  $dL$ , рассчитанное по формуле (9), минимально, справедливо соотношение  $f^*(x) = R(x)$ .

Для исследования внутренней структуры временного ряда  $G$  проводится сингулярный спектральный анализ. Методика приведения временного ряда  $G$  к нормированному виду состоит из следующих шагов:

Шаг 1. Находится величина математического ожидания значений временного ряда:  $Mean = 102,54$ .

Шаг 2. Находится разность между значениями элементов ряда и его математического ожидания:  $G - Mean$ .

Шаг 3. Находится максимальное по модулю значение из ряда  $G - Mean$ :  $Max = 7,54$ .

Шаг 4. Делим полученные значения на шаге 2  $G - Mean$  на величину  $Max$ , найденную на шаге 3.

Шаг 5. Умножаем полученные значения на шаге 4 на нормирующий множитель  $\beta$ , равный единице.

Шаг 6. Смещаем полученные значения на шаге 5 на величину параметра сдвига  $\gamma$ , в данном примере равного единице. Применяя сингулярный спектральный анализ к нормированному ряду  $G_n$ , построим его модель разложением на трендовую, гармоническую и шумовую составляющие.

Восстановленная трендовая составляющая  $G_{ТВ}$  представлена на рис. 1.

Построение регрессионной модели для трендовой составляющей  $G_{ТВ}$  может быть выполнено в статистическом пакете STATISTICA.

Восстановленная гармоническая составляющая  $G_{НВ}$  представлена на рис. 2.

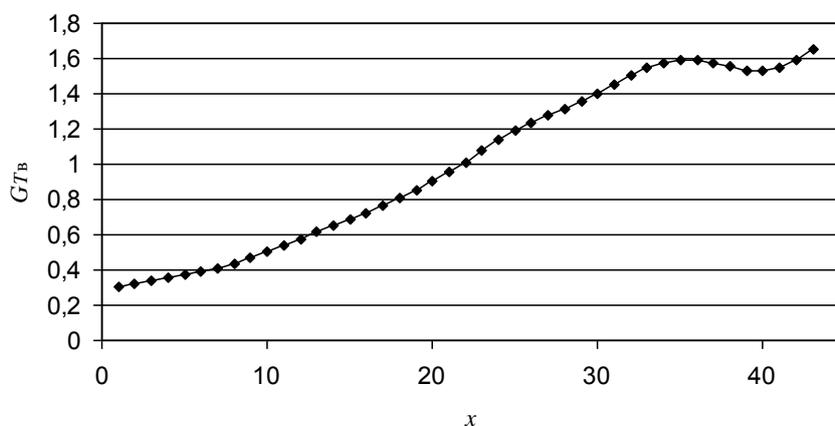


Рис. 1. Восстановленная трендовая составляющая

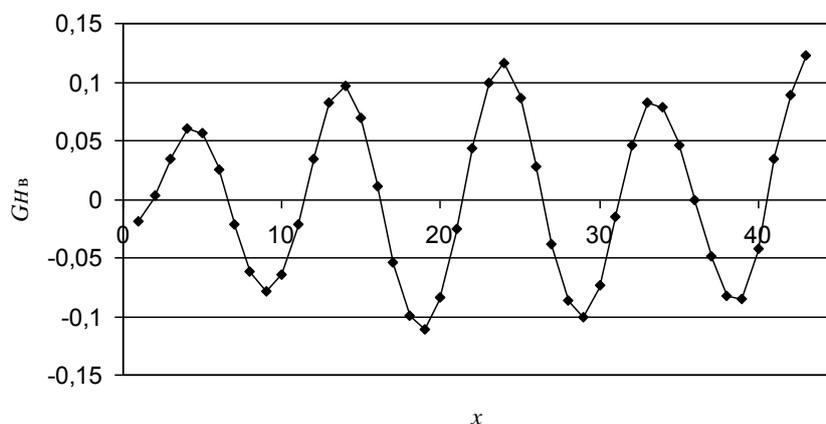


Рис. 2. Восстановленная гармоническая составляющая



При автоматизации построения аналитической модели  $G_{НВ}$  могут быть использованы функциональные возможности математического пакета MathCad, в частности, дискретное преобразование Фурье.

Соотношение  $\phi$ -оценок для аддитивных составляющих нормированного временного ряда  $\phi_T$ :  $\phi_H : \phi_N = 8,4 : 2,2 : 1,0$ .

Для определения закона распределения шумовой составляющей воспользуемся программным модулем BelSim2#.random. При этом рассчитаны точечные оценки выборки  $X$  с восстановленной шумовой составляющей, в частности:  $Min = -0,194$ ,  $Max = 0,136$ ,  $v_1 = -0,0104$ ,  $\mu_2 = 0,0056$ ,  $\mu_3 = -1,123$ ,  $\mu_4 = 8,743$ ,  $\beta_1 = 0,072$ ,  $\beta_2 = 2,796$ ,  $\aleph = -0,0884$  (1). Для выборочных данных построены функции плотностей распределений  $f_1(x)$ ,  $f_2(x)$ ,  $f_3(x)$ ,  $f_5(x)$ ,  $f_7(x)$ ,  $R(x)$ ,  $E(x)$ ,  $N(x)$ . С заданной точностью  $\varepsilon = 0,1$  критерий классификации типов кривых Пирсона  $\aleph = 0$ .

Результат работы программы BelSim2#.random:

$f1(x) = 5,077 * (1 + (x - 0,0104) / 0,321)^7,141 * (1 - (x - 0,0104) / 0,333)^7,426$ ;  
 $f2(x) = 5,346 * (1 - (x - 0,0104)^2 / 0,153)^{12,206}$ ;  
 $f3(x) = 5,524 * \exp(99,60 * (x - 0,0104)) * (1 + (x - 0,0104) / (-0,547))^{54,47}$ ;  
 $f5(x) = 0,024 * (x - 0,289)^{-20,072} * \exp(-5,584 / (x - 0,289))$ ;  
 $f7(x) = 5,512 * \exp(-(x - 0,0104)^2 / 0,0112)$ ;

$R(x) = 3,032$ ;

$E(x) = -96,19 * \exp(96,19 * x)$ ;

$N(x) = 1 / (0,075 * (2 * \pi)^{1/2}) * \exp(-1 * (x - 0,0104)^2 / 0,0112)$ .

При проверке критериями  $\chi^2$ ,  $\lambda$ ,  $\omega^2$  гипотезы о принадлежности выборки  $X$  законам распределения с построенными плотностями  $f_1(x)$ ,  $E(x)$  отклонены, для отклонения гипотез о распределении выборок по законам с распределениями  $f_2(x)$ ,  $f_3(x)$ ,  $f_5(x)$ ,  $f_7(x)$ ,  $R(x)$ ,  $N(x)$  нет оснований. Так как для функции  $f_7(x)$  значение  $dL$ , рассчитанное по формуле (9), минимально, справедливо соотношение  $f^*(x) = f_7(x)$ .

**Заключение.** Таким образом, определена внутренняя структура исследуемой равномерно распределенной последовательности числовых данных, включающая трендовую составляющую (см. рис. 1), гармоническую составляющую (см. рис. 2) и шумовую составляющую, соответствующую типу VII кривой плотности распределения семейства Пирсона. В отличие от применения статистических методов при анализе результатов имитационных экспериментов полученная модель числовых данных дает дополнительную информацию для принятия решения, чем достигается повышение эффективности эксплуатации имитационных моделей.

#### Библиографический список

1. Крамер Г. Математические методы статистики / пер. с англ.; 2-е изд. М.: Мир, 1975. 648 с.
2. Программный модуль расчета кривой плотности распределения случайной составляющей в последовательности данных «BelSim2#.random»: свидетельство о регистрации компьютерной программы № 306 / Е.А. Якимов [и др.]. Минск: НЦИС, 2011. Заявка № С20110024. Дата подачи: 04.04.2011.
3. Якимов А.И. Технология имитационного моделирования систем управления промышленных предприятий: монография. Могилев: Белорус.-Рос. ун-т, 2010. 304 с.
4. Якимов Е.А. Исследование SSA-метода на основе комплексного применения информационных технологий // Доклады БГУИР. 2010. № 2(48). С.77–83.
5. О приведении последовательностей данных к нормированному виду для преобразования методом сингулярного спектрального анализа / Е.А. Якимов [и др.] // Изв. Гомельс. гос. ун-та им. Ф. Скорины. 2011. № 6(69). С.204–214.
6. Biller B., Gunes C. Introduction to Simulation Input Modeling [Электронный ресурс] Proceedings of 2010 Winter Simulation Conference. 2010. P. 49–58. URL: <http://www.informs-sim.org/wsc10papers/006.pdf> (дата обращения: 10.04.2012).
7. Golyandina N., Nekrutkin V., Zhigljavsky A. Analysis of Time Series Structure: SSA and Related Techniques. Boca Raton: Chapman & Hall/CRC, 2001. 310 p.
8. Johansson B., Johansson J., Kinnander A. Information structure to support discrete event simulation in manufacturing systems // Proceedings of the 2003 Winter Simulation Conference, ed. S. Chick, P. J. Sanchez, D. Ferrin and D. J. Morrice. New Orleans, Louisiana, 2003. P.1290–1295.
9. Law A. M. ExpertFit Version 7 [Электронный ресурс] // User's Guide.pdf [2006. 144 p.] URL: <http://www.averill-law.com/> (дата обращения: 23.03.2011).