

УДК 681.3, 681.5

ПОВЫШЕНИЕ ОПЕРАТИВНОСТИ АНАЛИЗА МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

А. Н. Кабанов, к.т.н., доцент кафедры АСУ РГРТУ, Рязань, Россия;

orcid.org/0000-0002-7248-0918, e-mail: kbnnik@yandex.ru

Д. Н. Фоломкин, старший преподаватель кафедры АСУ РГРТУ, Рязань, Россия;

orcid.org/0000-0002-1648-0374, e-mail: d.folomkin@yandex.ru

Целью работы является поиск решения для повышения оперативности анализа многомерных временных рядов. Предлагается технология обработки, которая состоит из нескольких этапов: фильтрации множества сигналов с использованием оперативного алгоритма поиска оптимального временного масштабного коэффициента спектрального разложения, кластеризации многомерных сигналов с помощью алгоритма Прима, обработки однородных групп сигналов на скользящем интервале с использованием многомерно-матричных преобразований.

Ключевые слова: многомерно-матричное представление, многомерные временные ряды, белый шум, цветной шум, кластеризация, алгоритм Прима, рекуррентные оценки, скользящий интервал, аппроксимация, коэффициенты разложения, временной масштабный коэффициент.

DOI: 10.21667/1995-4565-2020-71-172-179

Введение

Анализ многомерных временных сигналов представляет собой довольно сложный процесс, требующий больших затрат временных и вычислительных ресурсов. Поиск способов оптимизации анализа является важной и актуальной задачей.

В статье предлагается технология обработки многомерных сигналов, которая позволяет повысить оперативность и точность их анализа. Последовательность обработки можно разделить на три этапа.

На 1-м этапе необходимо осуществить фильтрацию сигналов для общего случая помех типа небелого шума. Подобное исследование проводилось авторами в работах [1, 2].

На 2-м этапе после получения отфильтрованных сигналов необходимо осуществить их разбиение на кластеры для повышения точности анализа по каждому кластеру. Распространение алгоритма Прима на многомерный случай позволяет расширить возможности анализа многомерных сигналов для повышения точности. Кластеризация многомерных временных рядов с использованием алгоритма Прима была рассмотрена в [3].

На 3-м этапе обработку сигналов предлагается проводить на скользящем интервале. Для решения линейных уравнений требуется обращение многомерных матриц. Предлагается метод обращения на основе многомерно-матричных преобразований, что также было рассмотрено авторами в [5].

В целом предлагаемая технология обработки позволяет существенно повысить оперативность и точность анализа многомерных временных сигналов. Результаты исследований отдельных этапов технологии представлены в ряде работ авторов [1-3, 5].

Способы представления многомерных матриц и правила действия над многомерными матрицами подробно изложены в работе [5].

Этап 1. Фильтрация сигнала с использованием оперативного алгоритма поиска оптимального временного масштабного коэффициента спектрального разложения в условиях небелого шума

Исходный многомерный процесс описывается уравнением:

$$y(1,0;n) = \varphi(0,1;n) \cdot c(2,0) + e(1,0;n), \quad (1)$$

где $y(1,0;n) = \{y_1(n), y_2(n), \dots, y_k(n)\}$ – наблюдаемые временные ряды; $\varphi(0,1;n) = \{\varphi_j[n]\}$, $j = \{1..L\}$ – известные функции разложения временных рядов; $c(2,0) = \{c_{ij}\}$, $i = \{1..k\}$, $j = \{1..L\}$ – оцениваемые коэффициенты разложения временных рядов; $e(1,0;n) = \{e_1(n), e_2(n), \dots, e_k(n)\}$ – помехи в наблюдаемых рядах типа белого шума. Допускаем, что помехи $e(1,0;n)$ – имеют одинаковые корреляционные функции для всех процессов. В случае небелого шума (цветного) возможен оперативный переход к белому [1]. Воздействуя на уравнение (1) отбеливающим оператором $K(1,1)$, получаем:

$$K(1,1) \cdot y(1,0;n) = K(1,1) \cdot \varphi(0,1;n) \cdot c(2,0) + K(1,1) \cdot e(1,0;n).$$

Подберем $K(1,1)$ так, чтобы произведение $K(1,1) \cdot e(1,0)$ являлось белым шумом.

Для оперативного определения корреляционной матрицы целесообразно использовать метод аппроксимации корреляционной функции по системе ортогональных функций с подстраиваемым масштабом [2]. Рассмотрим алгоритм определения временного масштабного коэффициента [2] при аппроксимации корреляционной функции

$$K(t) = \sum_{i=1}^N C_i \cdot \varphi_i\left(\frac{t}{T1}\right) = \vec{\varphi}^T\left(\frac{t}{T1}\right) \cdot \vec{C},$$

где C_i , $T1$ – параметры, подлежащие определению. $T1$ выступает в роли масштабного коэффициента и оказывает существенное влияние на точность аппроксимации при ограниченном числе коэффициентов C_i . Априорный выбор $T1$ не всегда обеспечивает достаточную точность аппроксимации. За критерий близости исходного процесса и аппроксимирующей функции примем квадратичный критерий.

Существует несколько способов определить значения параметров C_i , $i = \overline{1, N}$, обеспечивающие минимум выбранного критерия. Однако большинство из них либо требуют сложных вычислений, либо занимают много времени.

Предлагается на основе оценок $\hat{C}_i$, определённых из условия минимума критерия только при двух различных значениях $\hat{T}1$, а именно $\hat{T}1 = T21$ и $\hat{T}1 = T22$, определить «истинные» значения параметров процесса $T1$ и C_i .

При такой постановке задачи в качестве измеряемых величин выступают уже не значения наблюдаемого процесса, а оценки коэффициентов

$$\hat{C}1^T = \{\hat{C}1_1(T21), \hat{C}1_2(T21), \dots, \hat{C}1_n(T21)\},$$

$$\hat{C}2^T = \{\hat{C}2_1(T22), \hat{C}2_2(T22), \dots, \hat{C}2_n(T22)\},$$

определённые из условия минимума критерия при двух различных временных масштабных коэффициентах $\hat{T}1 = T21$ и $\hat{T}1 = T22$, т.е. «измеряемые» величины могут быть представлены в виде

$$K1 = \bar{A}^{-1}(T21) \cdot \int_0^{\infty} \bar{\varphi}\left(\frac{t}{T21}\right) \cdot K(t) dt,$$

$$K2 = \bar{A}^{-1}(T22) \cdot \int_0^{\infty} \bar{\varphi}\left(\frac{t}{T22}\right) \cdot K(t) dt,$$

$$\bar{A}(T) = \int_0^{\infty} \bar{\varphi}\left(\frac{t}{T}\right) \cdot \bar{\varphi}^T\left(\frac{t}{T}\right) dt.$$

Подставляя в уравнения вместо $K(t)$ его аппроксимирующее выражение получаем

$$K1 = \bar{F}1 \cdot \bar{C}, \quad K2 = \bar{F}2 \cdot \bar{C},$$

где $\bar{F}1 = \bar{A}^{-1}(T21) \cdot \bar{B}(T21, T1)$, $\bar{F}2 = \bar{A}^{-1}(T22) \cdot \bar{B}(T22, T1)$, $\bar{B}(T_i, T_j) = \int_0^\infty \vec{\varphi}(\frac{t}{T_i}) \cdot \vec{\varphi}^T(\frac{t}{T_j}) dt$.

Задавшись некоторым значением $T1 = \hat{T}1$ определим оценку $\hat{C}1$:

$$\begin{aligned}\hat{C}1(T21, \hat{T}1) &= \bar{C}1 \cdot \bar{K}1, \\ \bar{C}1 &= \bar{F}1^{-1} = \bar{B}^{-1}(T21, \hat{T}1) \cdot \bar{A}(T21).\end{aligned}$$

Аналогично

$$\begin{aligned}\hat{C}2(T22, \hat{T}1) &= \bar{C}2 \cdot \bar{K}2, \\ \bar{C}2 &= \bar{F}2^{-1} = \bar{B}^{-1}(T22, \hat{T}1) \cdot \bar{A}(T22).\end{aligned}$$

При $\hat{T}1 = T1$ оценки $\hat{C}1(T21, T1)$ и $\hat{C}2(T22, T1)$ совпадут (при отсутствии помех). На практике за истинное значение $T1$ принимается значение $\hat{T}1$, при котором величина квадратичного критерия $Q = \sum_{i=1}^N \Delta C_i^2$ принимает минимальное значение. Очень важно подчеркнуть, что значения матриц $\bar{C}1$ и $\bar{C}2$ зависят лишь от отношения $\frac{T1}{T21}$ и от отношения $\frac{T1}{T22}$. Матрицы могут быть протабулированы заранее для выбранной системы функций $\varphi\left(\frac{t}{T}\right)$ [2].

Пример.

Вышеописанный алгоритм подходит для обработки любой функции, которую можно представить в виде (1). Для примера рассмотрим спектральное разложение переходной функции колебательного звена 2-го порядка (рисунок 1):

$$y(t) = k \left(1 - A \cdot e^{-\zeta \frac{t}{T}} \cdot \sin(\omega_0 \frac{t}{T} + \varphi) \right), \quad (1)$$

где передаточный коэффициент $k = 1$; постоянная времени в секундах $T = 5$; коэффициент затухания $\zeta = 0,707$; $A = \frac{1}{\sqrt{1-\zeta^2}}$, $\omega_0 = \sqrt{1-\zeta^2}$, $\varphi = \arctg\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right)$.

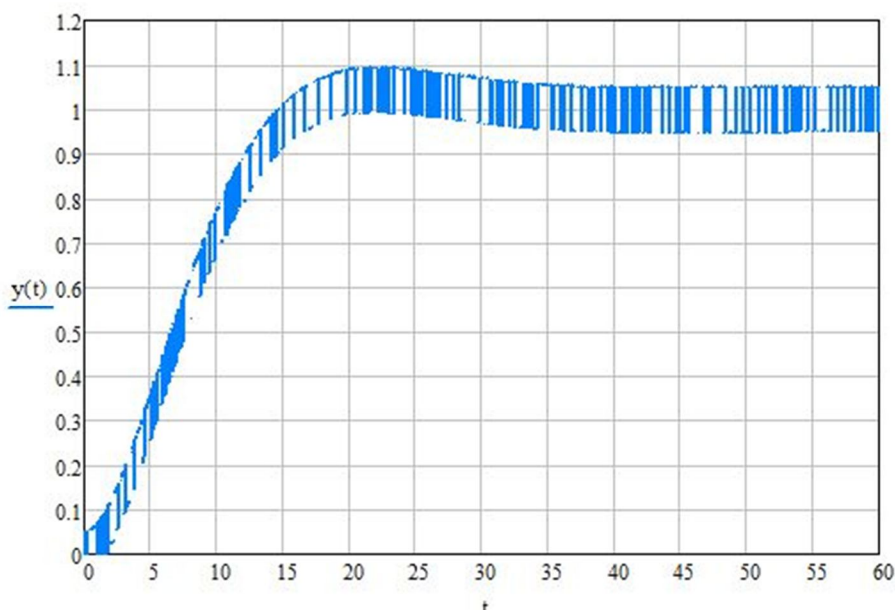


Рисунок 1 – Исходная функция – переходная функция колебательного звена 2-го порядка с помехой

Figure 1 – Original function - 2nd order link transition function with a noise

В качестве базиса для разложения выбираем первые пять функций Лягерра (ортонормированных), общее выражение для которых имеет вид

$$L_n(t) = \sqrt{2m} \cdot e^{-mt} \cdot \sum_{k=1}^n \frac{(n-1)!}{(n-k)! [(k-1)!]^2} (-2mt)^{k-1}, \quad (2)$$

где $m = \frac{1}{T}$.

Графики первых пяти функций Лягерра с МК $T = 1$ показаны на рисунке 2.

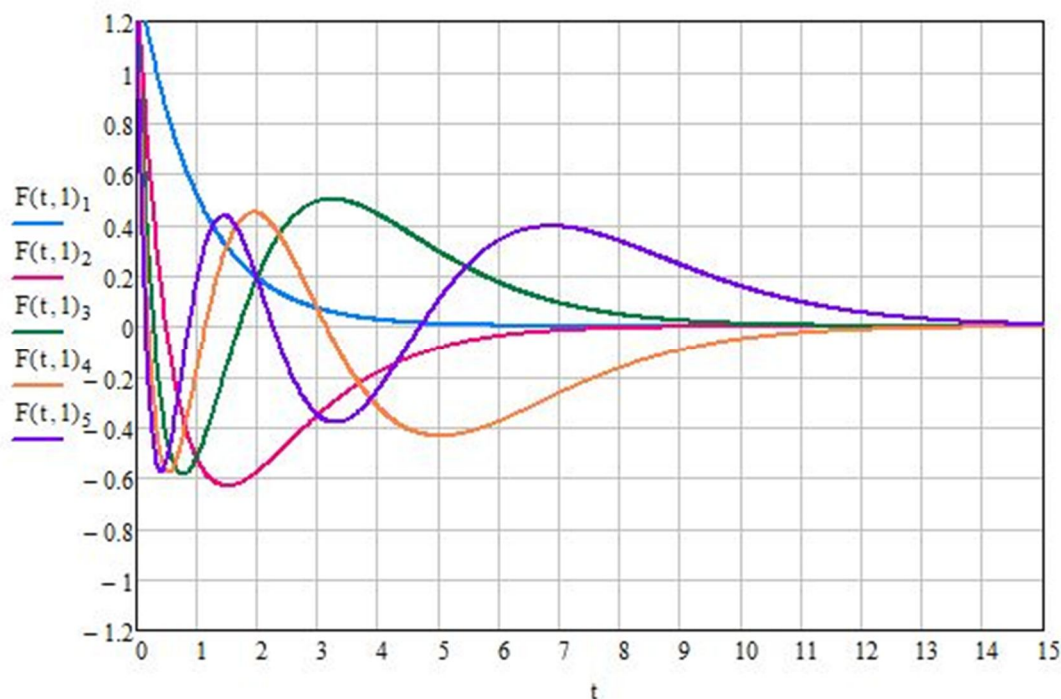


Рисунок 2 – Графики первых пяти функций Лягерра
Figure 2 – Laguerre functions (first five)

По графикам функций видно, что они затухают примерно при $t = 10$. Задаем два МК $T_{21} = 1$ и $T_{22} = 10$.

При этом векторы оценок коэффициентов разложения соответственно равны:

$$\begin{aligned} \hat{C}_1 &= [-1.371 \quad 1.224 \quad -0.991 \quad 0.738 \quad -0.504]^T, \\ \hat{C}_2 &= [-2.187 \quad -0.634 \quad 0.133 \quad 0.238 \quad 0.227]^T. \end{aligned}$$

Оптимальный МК равен $T_{opt} = 2$.

Оптимальные коэффициенты разложения равны:

$$C_{opt} = [-1.815 \quad 1.263 \quad -0.605 \quad 0.154 \quad 0.059]^T.$$

График аппроксимации выходного сигнала показан на рисунке 3.

Относительная погрешность аппроксимации $\sigma = 3,5 \cdot 10^{-3}$.

Таким образом, вышеописанный алгоритм обеспечивает довольно высокое качество аппроксимации.

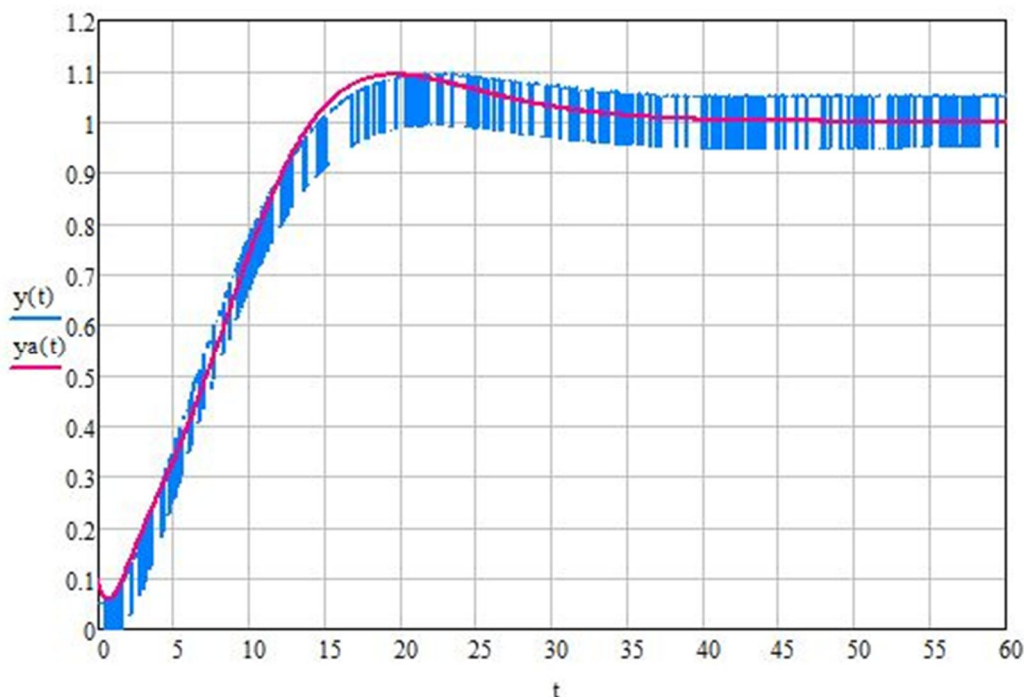


Рисунок 3 – Аппроксимация переходной функции 2-го порядка функциями Лягерра
Figure 3 – Approximation of the 2nd-order transition function by the Laguerre functions

Этап 2. Кластеризация многомерных временных рядов на основе алгоритма Прима

Для повышения оперативности обработки множество наблюдаемых рядов необходимо разбить на однородные группы по наблюдаемым данным $\{y_1(n), y_2(n) \dots y_k(n)\}$. Осуществляем построение кратчайшего остовного дерева путем распространения алгоритма Прима на многомерный случай.

Считаем, что единицы измерения всех рядов одинаковы. В противном случае необходимо привести ряды к безразмерным величинам, используя формулу $y(t) = \frac{y(t) - \bar{y}(t)}{\delta \cdot y(t)}$.

Адаптивная кластеризация множества элементов производится путем удаления части ребер графа по критерию минимальной суммарной дисперсии классов. Для разбиения множества элементов на K классов удаляются $K-1$ ребер [3].

Пример.

Построим многомерный временной ряд (рисунок 4), используя следующие функции:

– $y1, y2$ – переходные функции колебательного звена 2-го порядка $y(t) = k \left(1 - A \cdot e^{-\zeta \frac{t}{T}} \times \sin(\omega_0 \frac{t}{T} + \varphi) \right)$ с параметрами $k = 1, T = 5$ для $y1$ и $T = 3$ для $y2$, $\zeta = 0.707$ для $y1$ и $\zeta = 0.5$ для $y2$, $A = \frac{1}{\sqrt{1-\zeta^2}}$, $\omega_0 = \sqrt{1-\zeta^2}$, $\varphi = \arctg\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right)$;

– $y3, y4, y5$ – экспоненты $y(t) = e^{-\frac{t}{T}}$ с параметрами $T = 1, T = 3$ и $T = 5$ соответственно.

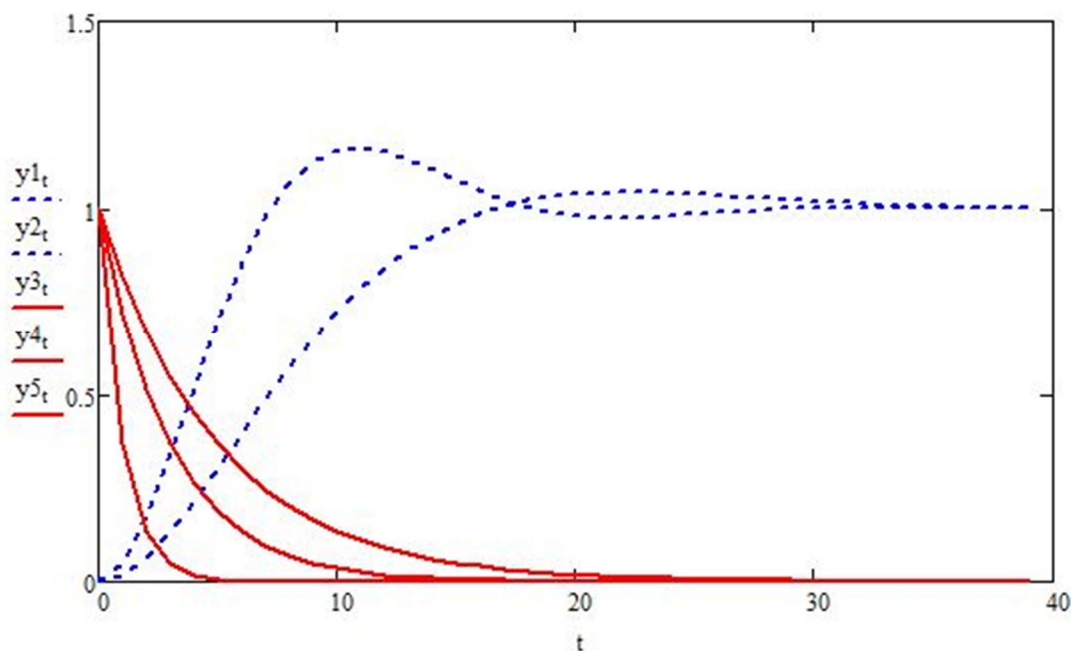


Рисунок 4 – Набор функций, представляющих многомерный временной ряд
Figure 4 – Set of functions that represents multidimensional time series

Строим начальную матрицу расстояний по формуле $D_{i,j} = \sum_{t=0}^{T_{n-1}} \sqrt{(y_i^t - y_j^t)^2}$

	y1	y2	y3	y4	y5
y1	0	5,312	33,707	33,776	33,167
y2	5,312	0	37,559	37,045	35,938
y3	33,707	37,559	0	1,946	3,933
y4	33,776	37,045	1,946	0	1,987
y5	33,167	35,938	3,933	1,987	0

После применения алгоритма Прима, получаем следующий граф (рисунок 5).

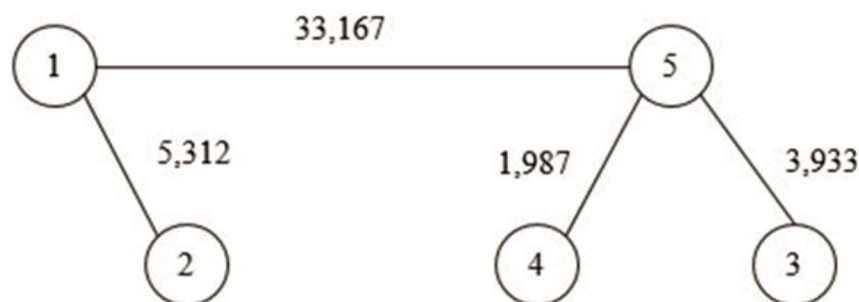


Рисунок 5 – Граф после применения алгоритма Прима
Figure 5 – Rest graph after applying of Prim's algorithm

Визуально наблюдаем два кластера: в первый входят ряды $y1$ и $y2$, во второй – ряды $y3$, $y4$, $y5$.

Этап 3. Обработка однородных групп временных рядов на скользящем постоянном интервале

Исходный процесс $y(n)$ представлен в матричном виде (1).

Модель процесса по результатам наблюдения на интервале $[n_0, n_0 + N]$ имеет вид

$$y_a(1, 0; n) = \phi(0, 1; n) \cdot \hat{c}(2, 0; n_0 + N).$$

Для квадратичного критерия оптимальные оценки определяются соотношением

$$\hat{c}(2, 0; n_0 + N) = A^{-1}(1, 1; n_0 + N) \cdot Z(2, 0; n_0 + N),$$

где $A(1, 1; n_0 + N) = \sum_{n=n_0}^{n_0+N} x^T(0, 1; n) \cdot x(0, 1; n),$

$$Z(2, 0; n_0 + N) = \sum_{n=n_0}^{n_0+N} x^T(0, 1; n) \cdot y(1, 0; n).$$

При вычислении обратной матрицы $A^{-1}(1, 1; n_0 + N)$ можно использовать следующие матричные соотношения [4]:

– если $A(1, 1) = B(1, 1) + x(1, 0) \cdot x(0, 1)$, то

$$A^{-1} = B^{-1} - \frac{B^{-1} \cdot x(1, 0) \cdot x(0, 1) \cdot B^{-1}}{1 + x(0, 1) \cdot B^{-1}(1, 0) \cdot x(1, 0)}; \quad (3)$$

– если $A(1, 1) = B(1, 1) - x(1, 0) \cdot x(0, 1)$, то

$$A^{-1} = B^{-1} + \frac{B^{-1} \cdot x(1, 0) \cdot x(0, 1) \cdot B^{-1}}{1 - x(0, 1) \cdot B^{-1}(1, 0) \cdot x(1, 0)}. \quad (4)$$

При скольжении с постоянным интервалом составляющие процесса в начале интервала удаляются из матрицы A , последующие составляющие добавляются в матрицу A :

$$\begin{aligned} A(1, 1; n_0 + N + 1) &= A(1, 1; n_0 + N) - x^T(1, 0; n_0) \cdot x(0, 1; n_0) + \\ &+ x^T(1, 0; n_0 + N + 1) \cdot x(0, 1; n_0 + N + 1). \end{aligned}$$

При этом необходимо использовать выражения (3) и (4), что усложняет расчеты и снижает оперативность.

Аналогичное соотношение можно записать для Z :

$$\begin{aligned} Z(2, 0; n_0 + N + 1) &= Z(2, 0; n_0 + N) - x^T(1, 0; n_0) \cdot y(1, 0; n_0) + \\ &+ x^T(1, 0; n_0 + N + 1) \cdot y(1, 0; n_0 + N + 1). \end{aligned}$$

Для повышения оперативности анализа предлагается использовать следующее утверждение. При многомерно-матричном представлении структура обратной многомерной матрицы получается транспонированием матрицы, а значения ее элементов совпадают со значениями обратной квадратной матрицы [5].

Заключение

Таким образом, в статье предложена технология обработки многомерных сигналов, которая позволяет повысить оперативность и точность их анализа. Технология состоит в последовательном применении этапов: фильтрации множества сигналов с использованием оперативного алгоритма поиска оптимального временного масштабного коэффициента спектрального разложения, кластеризации многомерных сигналов с помощью алгоритма Прима, обработки однородных групп сигналов на скользящем интервале и использовании многомерно-матричных преобразований для оперативного нахождения обратных матриц. Данная последовательность позволяет решить поставленную в статье задачу: повысить оперативность обработки многомерных временных рядов.

Библиографический список

1. Повышение оперативности получения уравнения множественной линейной регрессии при помехе типа белого шума / Д. Н. Фоломкин, А. Н. Кабанов, В. В. Аверьянова, А. М. Гаушко // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А. Н. Пылькина. – М.: Горячая линия-Телеком, 2016. – С. 7-10.
2. Фоломкин Д. Н. Оперативный алгоритм спектрального анализа контролируемого процесса с помощью функций сложной формы / Д. Н. Фоломкин, А. Н. Кабанов, Г. И. Нечаев // Вестник РГРТУ. 2010. – № 33. – Рязань: РГРТУ, 2010. – С. 103-107.

3. **Фоломкин Д. Н.** Отбор однородных групп экспертов на основе метода парного сравнения / Д. Н. Фоломкин, А. Н. Кабанов, А. М. Гаушко // Современные технологии в науке и образовании СТНО-2016: сб. тр. междуна. науч.-техн. и науч.-метод. конф.: в 4 т. / под общ. ред. О. В. Миловзорова. – Рязань: РГРТУ, 2016. Т. 3. – С. 71-73.

4. **Бронштейн И. Н.** Справочник по математике для инженеров и учащихся втузов : справочник / И. Н. Бронштейн, К. А. Семендяев. – СПб.: Лань, 2010. – 608 с.

5. Теоретико-множественные и структурно-математические основы описания дискретных систем: учеб. пособие / А. Н. Кабанов. Рязань. Рязан. гос. радиотехн. акад. 2004. – 52 с.

UDC 681.3, 681.5

IMPROVING THE EFFICIENCY OF MULTIVARIATE TIME SERIES ANALYSIS

A. N. Kabanov, Ph.D. (Tech.), associate Professor of MIS RSREU, Ryazan, Russia;
orcid.org/0000-0002-7248-0918, e-mail: kbnnik@yandex.ru

D. N. Folomkin, senior lecturer of MIS RSREU, Ryazan, Russia;
orcid.org/0000-0002-1648-0374, e-mail: d.folomkin@yandex.ru

The problem to increase the efficiency of multidimensional time series analysis is considered. The aim of the work is to find solutions to increase the efficiency of multidimensional time series analysis. Several approaches are proposed to optimize the analysis, namely: the use of an operational algorithm for determining the optimal temporal scale factor of spectral decomposition under color noise conditions, clustering of multidimensional time series based on Prim's algorithm, and obtaining recurrent estimates of the approximation of multidimensional time series on a moving interval.

Key words: multidimensional representation, multidimensional time series, white noise, color noise, clustering, Prim's algorithm, recurrence estimates, moving interval, approximation, decomposition coefficients, time scale factor.

DOI: 10.21667/1995-4565-2020-71-172-179

References

1. Povyshenie operativnosti polucheniya uravneniya mnozhestvennoj linejnoy regressii pri pomekhe tipa belogo shuma / D. N. Folomkin, A. N. Kabanov, V. V. Aver'yanova, A. M. Gaushko. *Matematicheskoe i programnoe obespechenie vychislitel'nyh sistem : mezhvuzov. sb. nauch. tr. pod red. A. N. Pyl'kina*. M.: Goryachaya liniya-Telekom, 2016, pp. 7-10.

2. **Folomkin D. N.** Operativnyj algoritm spektral'nogo analiza kontroliruемого processa s pomoshch'yu funkciy slozhnoy formy. D. N. Folomkin, A. N. Kabanov, G. I. Nechaev. *Vestnik RGRTU*. 2010, vol. 33. Ryazan': RGRTU, 2010, pp. 103-107.

3. **Folomkin D. N.** Otbor odnorodnyh grupp ekspertov na osnove metoda parnogo sravneniya. D. N. Folomkin, A. N. Kabanov, A. M. Gaushko. *Sovremennye tekhnologii v nauke i obrazovanii STNO-2016: sb. tr. mezhdunar. nauch.-tekhn. i nauch.-metod. konf.: vol. 4. pod obshch. red. O. V. Milovzorova*. Ryazan': RGRTU, 2016, vol. 3, pp. 71-73.

4. **Bronshtejn, I.N.** *Spravochnik po matematike dlya inzhenerov i uchashchihsya vtuzov: spravochnik*. I. N. Bronshtejn, K. A. Semendyaev. Sankt-Peterburg.: Lan', 2010. 608 p.

5. *Teoretiko-mnozhestvennye i strukturno-matematicheskie osnovy opisaniya diskretnykh sistem*: Ucheb. posobie. A.N. Kabanov. Ryazan'. Ryazan. gos. radiotekhn. akad. 2004. 52 s.