

РЕАЛИЗАЦИЯ МЕТОДА СИНГУЛЯРНОГО СПЕКТРАЛЬНОГО АНАЛИЗА ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ МАТЕМАТИЧЕСКОГО ПАКЕТА SCILAB¹

Г. Ю. Юрченко, А. И. Якимов

Аннотация. В статье представлены этапы реализации метода сингулярного спектрального анализа временных рядов в математическом пакете SciLab.

Ключевые слова: сингулярный спектральный анализ, программный модуль, временной ряд.

1. ВВЕДЕНИЕ

Для исследования метода сингулярного спектрального анализа (SSA-метода) применяется математический пакет SciLab, который является открытым окружением и одновременно языком программирования, где и реализована программа, реализующая базовый метод SSA [1]. До применения среды SciLab исследования проводились на основе комплекса информационных технологий, представленных табличным процессором MS Excel, математическим пакетом Mathcad и пакетом статистической обработки данных Statistica [2], а недавно - на основе программного обеспечения (ПО) BelSim2#.SSA [3, 4]. Такая методика является достаточно трудоемкой из-за значительного числа ручных операций [5, 6], ПО BelSim2#.SSA же вызывал определенные вопросы в промежуточных результатах и также требовал достаточно большое количество ручных операций.

2. ЭТАП ВЛОЖЕНИЯ СИНГУЛЯРНОГО СПЕКТРАЛЬНОГО АНАЛИЗА

Для экспериментальных исследований исходный ряд $G = (g_0, g_1, \dots, g_{N-1})$ может быть задан по известным функциям либо определен из входного файла. По известным функциям (или из входного файла) ряд формируется в программном коде на языке SciLab:

```
t = (1:N)';  
X = -1 + t  
X = X + 10 * sin(2 * %pi * t / 30); // sine function with period length 30  
rand('seed', 0.123456789);  
rand('normal');  
noise = rand(X); // Gaussian noise  
X = X + noise;
```

Если же происходит ввод из файла, ряд загружается следующим образом:

```
X = read("C:\Users\user\Desktop\input.txt", N, 1);
```

После этого на экран будет выведен исходный ряд. Процедура вложения является преобразованием исходного одномерного ряда $G = (g_0, g_1, \dots, g_{N-1})$ в последователь-

¹ Работа выполнена в порядке личной инициативы по НИРС

ность L - мерных векторов, число которых равно $K = N - L + 1$.

$$\mathbf{A}_i = (g_{i-1}, \dots, g_{i+L-2})^T, \quad 1 \leq i \leq K. \quad (1)$$

Блок кода SciLab формирования матрицы A :

```

N = length(y) // y = G, т.е. исходный ряд
if L > N / 2 then
    L = N - L
end
K = N - L + 1
X = zeros(L, K)
for i = 1 : K
    X(:, i) = y(i : i + L - 1) // Заполнение матрицы A, X в коде
end

```

Эти вектора образуют траекторную матрицу $A = [\mathbf{A}_1; \dots; \mathbf{A}_K]$ ряда G , в которой $a_{ij} = g_{i+j-2}$, т. е. матрица A имеет одинаковые элементы на диагонали $i + j = \text{const}$.

3. РЕАЛИЗАЦИЯ ЭТАПА СИНГУЛЯРНОГО РАЗЛОЖЕНИЯ В SCILAB

Обозначим $S = A \cdot A^T \in R^{L \times L}$. Матрица $A \cdot A^T$ симметричная и неотрицательно определенная, а значит ее собственные числа $\{\mu_k\}_{k=1}^L$ вещественны и неотрицательны. Представленные в виде $\mu_1 \geq \dots \geq \mu_L \geq 0$ собственные числа называют сингулярными значениями матрицы S . Пусть $\mathbf{U}_1, \dots, \mathbf{U}_L$ – соответствующие им ортонормированные собственные вектора. Будем называть $p = \max\{k \mid \mu_k > 0\}$ порядком сингулярного разложения. Обозначим

$$\mathbf{V}_k = \frac{1}{\sqrt{\mu_k}} A^T \mathbf{U}_k, \quad k = 1, \dots, p. \quad (2)$$

Тогда сингулярным разложением матрицы A называется ее представление в виде суммы элементарных матриц

$$A = A_1 + A_2 + \dots + A_p, \quad A_k = \sqrt{\mu_k} \mathbf{U}_k \mathbf{V}_k^T. \quad (3)$$

Для проведения сингулярного разложения использована функция `svd()`:

$$[U, LAMBDA] = \text{svd}(C)$$

Функция `svd` осуществляет *SVD*-разложение прямоугольной матрицы. На выходе подпрограмма возвращает массив сингулярных значений *LAMBDA*, упорядоченных по убыванию, и, по требованию, матрицы U , причем возможно как возвращение только левых и правых сингулярных векторов, так и полных матриц размером $m \times n$ и $n \times n$ (в зависимости от выбранных параметров).

Каждая из матриц A_k имеет ранг, равный единице. Поэтому их можно назвать элементарными матрицами. Вектор \mathbf{U}_k называют k -м левым сингулярным вектором

или просто k -м собственным вектором, вектор \mathbf{V}_k – правым сингулярным вектором. Набор $\langle \sqrt{\mu_k}, \mathbf{U}_k, \mathbf{V}_k \rangle$ называют k -ой собственной тройкой. Обозначим корень собственного числа матрицы S через $\lambda_k = \sqrt{\mu_k}$ и будем использовать это обозначение в дальнейших исследованиях.

4 Этап группировки в SciLab

Вид левых и правых сингулярных векторов, трактуемых в SSA как временные ряды, является очень важным для следующего шага метода – группировки. При этом для одномерного SSA левые и правые сингулярные вектора обладают определенной симметрией, так как в этих случаях сингулярные разложения траекторных матриц с длиной окна L и $K = N - L + 1$ эквивалентны.

Процедура группировки формально одинакова для всех разновидностей SSA. На основе разложения (3) процедура группировки делит все множество индексов $\{1, \dots, p\}$ на m непересекающихся подмножеств I_1, \dots, I_m . Пусть $I = \{i_1, \dots, i_p\}$, тогда результирующая матрица A_I , соответствующая группе I , определяется как $A_I = A_{i_1} + \dots + A_{i_p}$. Такие матрицы вычисляются для $I = I_1, \dots, I_m$, тем самым разложение (2.3) может быть записано в сгруппированном виде:

$$A = A_{I_1} + \dots + A_{I_m}. \quad (4)$$

Процедура выбора множеств $I = I_1, \dots, I_m$ и называется группировкой собственных троек. Для определения $I = I_1, \dots, I_m$ в скрипте используются сингулярные вектора. По структуре сингулярного вектора определяется его принадлежность к тренду, колебательной составляющей или шуму.

В программном коде для этого процесса написан ряд инструментов:

```
I=[1,2]; //массив применяемых для группировки компонент
```

Оценив форму компонент, массив I заполняется номерами компонент, соответствующих собираемой в текущий момент составляющей. Ниже представлена универсальная процедура для группировки:

```
function [y] = ssa_reconstruct(LAMBDA, U, V, I)
    I = matrix(I, 1, -1)
    [K, L] = size(V)
    N = K + L - 1
    // Этап 2: реконструкция
    // Этап 3: группировка
    for i = 1
        LAMBDA_U(:, i) = LAMBDA(i) * U(:, i)
    end
    // реконструированные компоненты
    RC = LAMBDA_U(:, I) * V(:, I)'
```

5 ЭТАП ДИАГОНАЛЬНОГО УСРЕДНЕНИЯ В SciLAB

На последнем шаге базового алгоритма каждая матрица сгруппированного разложения переводится в новый ряд длины N . Для произвольной матрицы X процедуру приведения ее к ганкелевому виду и последующему преобразованию в ряд (обозначим его как G^B) выразим следующим образом. Пусть X – матрица размера $L \times K$ с элементами x_{ij} , $1 \leq i \leq L$, $1 \leq j \leq K$. Положим $L^* = \min(L, K)$, $K^* = \max(L, K)$ и $N = L + K - 1$. Пусть $z_{ij} = x_{ij}$, если $L < K$ и $z_{ij} = x_{ji}$ в остальных случаях. Тогда диагональное усреднение переводит матрицу X в ряд $(g_0^B, \dots, g_{N-1}^B)$ по формуле

$$g_k^B = \begin{cases} \frac{1}{k+1} \sum_{j=1}^{k+1} z_{j, k-j+2} & | 0 \leq k \leq L^* - 1; \\ \frac{1}{L^*} \sum_{j=1}^{L^*} z_{j, k-j+2} & | L^* - 1 \leq k \leq K^*; \\ \frac{1}{N-k} \sum_{j=k-K^*+2}^{N-K^*+1} z_{j, k-j+2} & | K^* \leq k \leq N. \end{cases} \quad (5)$$

Выражение (5) соответствует усреднению элементов матрицы вдоль побочных диагоналей $i + j = k + 2$: выбор $k = 0$ дает $g_0^B = x_{11}$, для $k = 1$ получаем $g_1^B = (x_{12} + x_{21})/2$ и т. д. Применяв диагональное усреднение к матрицам, полученным на этапе группировки, приходим к разложению исходного ряда в сумму m рядов [1]:

```
// диагональное усреднение
y = zeros(N, 1)
for i = 1 : K + L - 1
    v = adiaG(RC, i)
    y(i) = sum(v) / length(v)
end
y = real(y)
endfunction
```

6. ЗАКЛЮЧЕНИЕ

Программный модуль для исследования сингулярного спектрального анализа разработан на основе математического пакета SciLab с применением функционального программирования для простоты реализации и изменения другими пользователями.

Файл с исходным временным рядом представляет собой текстовый файл, содержащий последовательность вещественных чисел. Результатом работы программы является матрица восстановленной компоненты и ее график на фоне исходного ряда. Включена возможность сохранения в файл и задание исходного ряда в формульной форме.

Технические требования SciLab:

Операционная система MS Windows: Windows XP, Windows Vista, Windows 7, Windows 8; GNU/Linux: CentOS, Debian, Redhat, Fedora, Suse, Ubuntu; Mac OS X 10.6.x

и выше. Процессор: Pentium IV class (или эквивалентный) и выше, оперативная память: 2 Gb RAM (минимум 1 Gb), пространство на жестком диске: 600 Mb.

Литература

1. **Golyandina, N.** Analysis of Time Series Structure: SSA and Related Techniques / N. Golyandina, V. Nekrutkin, A. Zhigljavsky. – Boca Raton: Chapman & Hall/CRC, 2001. – 310 p.
2. **Якимов, Е. А.** Исследование SSA-метода на основе комплексного применения информационных технологий / Е. А. Якимов // Доклады БГУИР. – 2010. – № 2(48). – С. 77–83.
3. **Якимов, Е. А.** Программный модуль анализа сингулярного спектра числовых последовательностей данных «BelSim2#.SSA»: свидетельство о регистрации компьютерной программы № 518 / Е. А. Якимов, Д. М. Албкейрат, Н. М. Чапаров, А. И. Якимов. – Минск: НЦИС, 2013. – Заявка № С20130042. – Дата подачи: 23.05.2013.
4. **Чапаров, Н. М.** Разработка программного обеспечения для исследования временных рядов методом сингулярного спектрального анализа / Н. М. Чапаров; науч. рук.: А. И. Якимов, Е. А. Якимов // 49-я студенческая научно-техническая конференция Белорусско-Российского университета: материалы конф., редкол.: И. С. Сазонов (гл. ред.) [и др.]; 16-17 мая 2013 г. – Могилев: Белорус.-Рос. ун-т, 2013. – С. 224.
5. **Абдул Хак, М. А.** Исследование числовых последовательностей данных методом анализа сингулярного спектра / М. А. Абдул Хак; науч. рук.: А. И. Якимов, Е. А. Якимов // 50-я студенческая научно-техническая конференция Белорусско-Российского университета: материалы конф., редкол.: И. С. Сазонов (гл. ред.) [и др.]; 22 мая 2014 г. – Могилев: Белорус.-Рос. ун-т, 2014.
6. **Микулич, А. Н.** Особенности исследования метода сингулярного спектрального анализа временных рядов / А. Н. Микулич, В. А. Долгий, В. О. Батыщиков; науч. рук.: А. И. Якимов, Е. А. Якимов // 51-я студенческая научно-техническая конференция Белорусско-Российского университета: материалы конф., редкол.: И. С. Сазонов (гл. ред.) [и др.]; 21-22 мая 2015 г. – Могилев: Белорус.-Рос. ун-т, 2015. – С. 139.

Юрченко Глеб Юрьевич

Магистрант электротехнического факультета
Белорусско-Российский университет, г. Могилев
Тел.: +375(29) 104-43-96
E-mail: gedxblood@mail.ru

Якимов Анатолий Иванович

Доцент кафедры Автоматизированные системы управления, канд. техн. наук
Белорусско-Российский университет, г. Могилев
Тел.: +375(222) 25-24-47
E-mail: ykmt@tut.by