

# Monte Carlo SSA for extracting weak signals

I.I. IVANOV<sup>1</sup>, P.P. PETROV<sup>2</sup>

<sup>1</sup>*University ...*

<sup>2</sup>*Institute ...*

*City, STATE*

e-mail: <sup>1</sup>ivanov@yandex.ru, <sup>2</sup>petrov@google.com

Abstract text.

**Keywords:** comma, separated, keywords, minimum 3, maximum 5

## 1 Введение

Рассмотрим следующую модель:  $\mathbf{X} = \mathbf{S} + \mathbf{R}$ , где  $\mathbf{X}$  — наблюдаемый временной ряд,  $\mathbf{S}$  — сигнал,  $\mathbf{R}$  — шум, т.е. реализация некоторого стационарного процесса. В работе рассматриваются две проблемы: проблема обнаружения сигнала  $\mathbf{S}$  и проблема выделения сигнала при его наличии.

Для решения первой проблемы используется метод Monte Carlo SSA (MC-SSA) [1], проверяющий гипотезу  $H_0 : \mathbf{S} = 0$ , а для решения второй — метод анализа сингулярного спектра (singular spectrum analysis, SSA) [2, 3]. Один из шагов SSA подразумевает визуальный анализ для определения компонент сигнала, поэтому возникает потребность в автоматизации этого шага, этой проблеме посвящены, например, работы [4–7]. Целью работы является определение подхода к автоматическому выделению слабых сигналов, обнаруживаемых критерием MC-SSA.

## 2 Метод autoMCSSA

Для описания разработанного метода необходимо ввести некоторые обозначения и предположения.

Для ряда  $\mathbf{X}$  длины  $N$  и  $0 \leq \omega_1 \leq \omega_2 \leq 0.5$  определим меру, следуя [4]

$$T(\mathbf{X}; \omega_1, \omega_2) = \frac{1}{\|\mathbf{X}\|^2} \sum_{k: \omega_1 \leq k/N \leq \omega_2} I_N(k/N),$$

где  $I_N$  — периодограмма ряда  $\mathbf{X}$ . Величину  $T(\mathbf{X}, \omega_1, \omega_2)$  можно рассматривать как долю вклада частот, содержащегося в интервале  $[\omega_1, \omega_2]$ .

В данной работе рассматривается модификация MC-SSA с поправкой на множественные сравнения [8]. В качестве векторов для проекции, необходимых для построения статистики критерия, были выбраны косинусы с равноотстоящими частотами  $\omega_k = k/(2L)$ ,  $k = 1, \dots, L$ . В этом случае можно говорить о значимых частотах, присутствующих в сигнале.

На рис. 1 изображена блок-схема алгоритма autoMCSSA. Заключается он в последовательном применении критерия MC-SSA до тех пор, пока гипотеза не

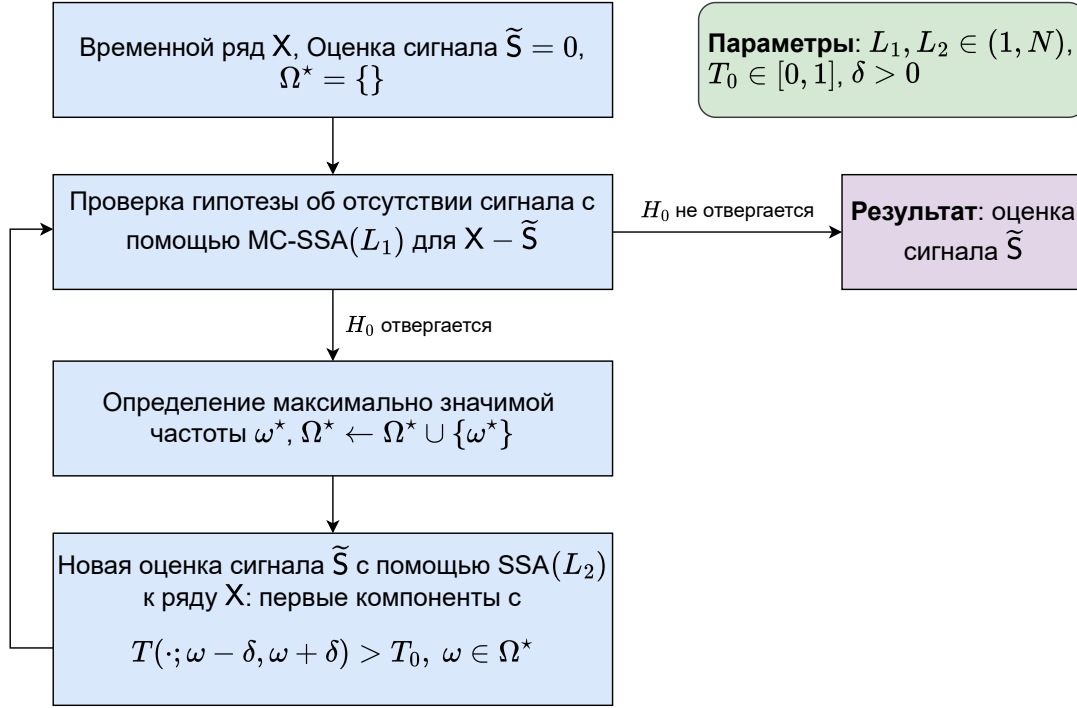


Рис. 1: Алгоритм autoMCSSA

перестанет отвергаться. Если на очередной итерации алгоритма гипотеза отвергается, определяется максимально значимая частота  $\omega^*$  и вычисляется новая оценка сигнала с помощью SSA к исходному ряду: для каждой частоты  $\omega$  из множества значимых частот  $\Omega^*$  выбираются первые компоненты с мерой  $T$  превышающей порог  $T_0$ . Как только гипотеза перестает отвергаться, алгоритм завершает свою работу и тогда  $\tilde{S}$  — итоговая оценка сигнала.

Оценивать частоту  $\omega^*$  будем следующим образом:

1. Найти индекс наиболее значимой частоты, т.е.  $k = \operatorname{argmax}_i (\hat{p}_i - c_i)$ , где  $c_i$  — верхняя граница доверительного интервала для  $\hat{p}_i$ ;
2. Вычислить значение  $\hat{\omega}$  как взвешенное среднее частот  $\omega_{k-1}, \omega_k, \omega_{k+1}$  с весами  $w_i = \max(0, \hat{p}_i - c_i)$ ;

Такой способ оценки позволяет получить более точную оценку  $\omega$  в случае, когда она не попадает в решетку  $k/(2L)$ .

Рассмотрим пример работы рассмотренного алгоритма. Пусть  $X = S + \xi$ , где  $\xi$  — модель AR(1) с параметрами  $\phi = 0.7$  и  $\sigma^2 = 1$ ,  $N = 200$ ,  $S = (s_1, \dots, s_N)$ ,

$$s_n = 0.075 e^{0.02n} \cos(2\pi n/8) + 2 \cos(2\pi n/4) + 0.2 \cdot (-1)^n.$$

На рис. 2 представлены первые 15 элементарно восстановленных с помощью SSA компонент. Сигналу соответствуют компоненты с индексами 1, 2, 5, 6 и 13. Не видя

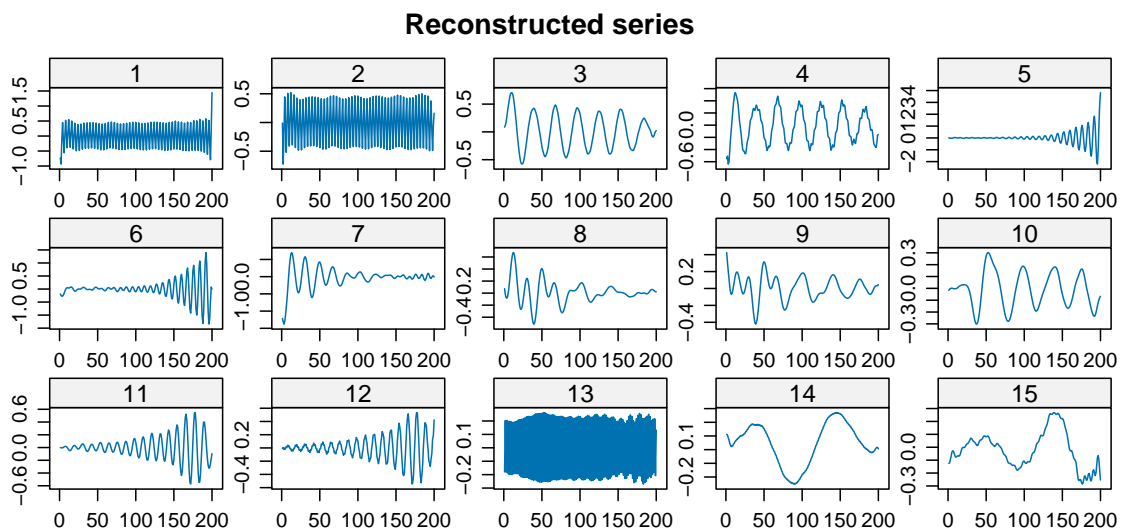


Рис. 2: Элементарные восстановленные компоненты

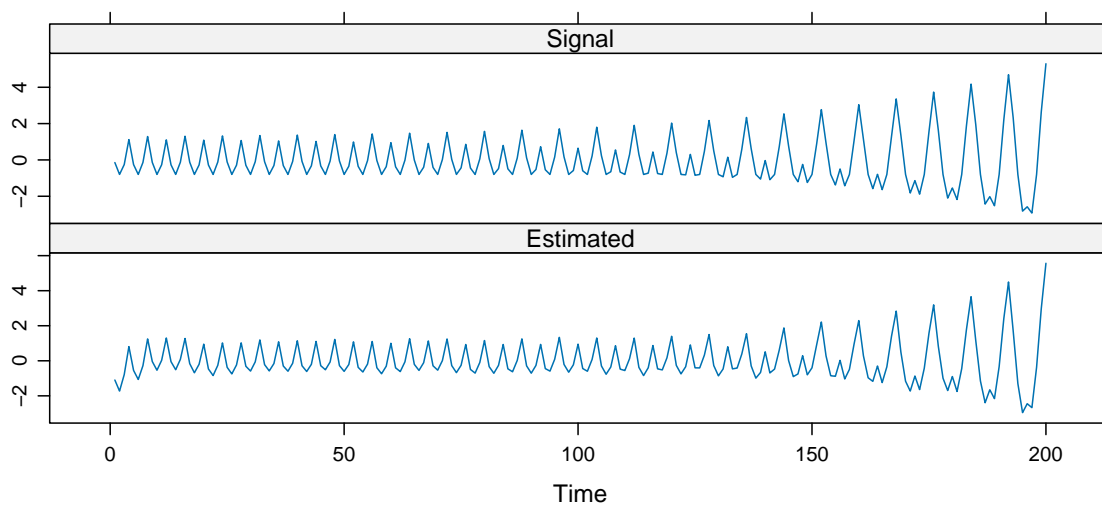


Рис. 3: Результат autoMCSSA ( $L_1 = 50$ ,  $L_2 = 100$ ,  $\delta = 1/80$ ,  $T_0 = 0.5$ )

формулы, по которой этот сигнал задается, сказать наверняка, какие компоненты неслучайные проблематично, поскольку компоненты 3, 4 и 11, 12 похожи на пары гармоник. Мы применили алгоритм autoMCSSA к этому ряду и получили, что разработанный метод правильно идентифицировал компоненты, соответствующие сигналу, на рис. 3 представлены истинная форма сигнала  $S$  и его оценка методом autoMCSSA.

## Список литературы

- [1] Allen M., Smith L. Monte Carlo SSA: detecting irregular oscillations in the presence of coloured noise // *Journal of Climate*. — 1996. — Vol. 9. — P. 3373–3404.
- [2] Broomhead D., King G. Extracting qualitative dynamics from experimental data // *Physica D: Nonlinear Phenomena*. — 1986. — Vol. 20, no. 2–3. — P. 217–236.
- [3] Golyandina N., Nekrutkin V., Zhigljavsky A. *Analysis of Time Series Structure: SSA and Related Techniques*. — Chapman and Hall/CRC, 2001.
- [4] Александров Ф., Голяндина Н. Автоматизация выделения трендовых и периодических составляющих временного ряда в рамках метода «Гусеница»-SSA // *Exponenta Pro (Математика в приложениях)*. — 2004. — Vol. 7-80. — P. 54–61.
- [5] Kalantari M., Hassani. H. Automatic grouping in singular spectrum analysis // *Forecasting*. — 2019. — Vol. 1, no 1. — P. 189–204.
- [6] Bogalo J., Poncela P., Senra E. Circulant singular spectrum analysis: A new automated procedure for signal extraction // *Signal Processing*. — 2021. — Vol. 179.
- [7] Golyandina N., Dudnik P., Shlemov A. Intelligent Identification of Trend Components in Singular Spectrum Analysis // *Algorithms*. — 2023. — Vol. 16. — ID 353.
- [8] Golyandina N. Detection of signals by Monte Carlo singular spectrum analysis: multiple testing // *Statistics and Its Interface*. — 2023. — Vol. 16. no 1. — P. 147–157.