

Модификации метода анализа сингулярного спектра для анализа временных рядов: Circulant SSA и Generalized SSA

Погребников Н. В., гр. 21.Б04-мм

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

Научный руководитель: д. ф.-м. н., проф. Голяндина Н. Э.

Санкт-Петербург, 2025

План доклада:

- ❶ Введение — методы, постановка задачи и цели.
- ❷ Критерии сравнения методов
- ❸ Сравнение SSA и GSSA
- ❹ Сравнение SSA, разложение Фурье и CiSSA
- ❺ Итоги и выводы.

Пусть $X = (x_1, \dots, x_N)$ – временной ряд длины N , $x_i \in \mathbb{R}$ – наблюдение в момент времени i .

$X = X_{\text{Trend}} + X_{\text{Periodics}} + X_{\text{Noise}}$, где:

- X_{Trend} – тренд, медленно меняющаяся компонента;
- $X_{\text{Periodics}}$ – сумма периодических компонент;
- X_{Noise} – шум, случайная составляющая.

Методы: **SSA** – метод, позволяющий раскладывать временной ряда в сумму интерпретируемых компонент (Golyandina, Nekrutkin и Zhigljavsky 2001); **GSSA** – модификация **SSA** на основе добавления весов (Gu и др. 2024); **CiSSA** – модификация **SSA** на основе циркулярной матрицы (Bogalo, Poncela и Senra 2020).

Задача: Описание модификаций в контексте теории **SSA**, сравнение алгоритмов, реализация их на языке R.

Пример

$$X = S + X_{\text{Noise}} = S^{(1)} + S^{(2)} + X_{\text{Noise}} = e^{An} \sin(2\pi\omega_1 n) + \cos(2\pi\omega_2 n) + \varepsilon_n.$$

ω_1, ω_2 – частоты; $\varepsilon_n \sim N(0, \sigma^2)$ – шум;

S – сигнал.

\hat{S} – оценка выделения сигнала методом.

$\hat{S}^{(1)}, \hat{S}^{(2)}$ – оценки разделения компонент $S^{(1)}, S^{(2)}$.

Критерии сравнения методов:

- Выделение сигнала;
- Разделимость;
- Постановка задачи (для CiSSA частоты предполагаются известными).

$X_N = X_N^{(1)} + X_N^{(2)}$. М – метод разделения ряда на компоненты с параметрами Θ . $\hat{X}_N^{(1)}$ – оценка $X_N^{(1)}$, восстановленная М.

Определение 1

Ряды $X_N^{(1)}$ и $X_N^{(2)}$ точно разделимы методом М, если существует такое Θ , что $\text{MSE} \left(X_N^{(1)}, \hat{X}_N^{(1)} \right) = 0$.

Определение 2

Ряды $X_N^{(1)}$ и $X_N^{(2)}$ асимптотически разделимы методом М, если существует последовательность $\Theta(N)$, $N \rightarrow \infty$, что $\text{MSE} \left(X_N^{(1)}, \hat{X}_N^{(1)} \right) \rightarrow 0$.

$X = (x_1, \dots, x_N)$ — временной ряд. $1 < L < N$ — длина окна.

Алгоритм SSA:

- 1 **Построение траекторной матрицы:**

$$X = \mathcal{T}_L(X) = [X_1 : \dots : X_K], \quad X_i = (x_i, \dots, x_{i+L-1})^T, \\ 1 \leq i \leq K, \quad K = N - L + 1.$$

- 2 **Сингулярное разложение (SVD) траекторной матрицы.**
- 3 **Группировка элементарных матриц SVD.**
- 4 **Восстановление временного ряда по матрицам SVD:**
 $X = \tilde{X}_1 + \dots + \tilde{X}_m.$

$$X = S + X_{\text{Noise}} = S^{(1)} + S^{(2)} + X_{\text{Noise}}$$

Определение 3 (Golyandina и Shlemov 2015)

Вложенный вариант SSA — двухэтапный метод:

1. *Задается r . \tilde{S} – сумма первых r слагаемых SVD разложения траекторной матрицы сигнала S с помощью базового **SSA**.*
2. *Применение другого метода к \tilde{S} для улучшения делимости: $\tilde{S} = \tilde{S}_1 + \tilde{S}_2$.*

SSA EOSSA (Golyandina, Dudnik и Shlemov 2023) является вложенным вариантом **SSA**.

Метод GSSA. Алгоритм

$X = (x_1, \dots, x_N)$ — временной ряд, параметры L и $\alpha \geq 0$.

$$w^{(a)} = (w_1, w_2, \dots, w_L) = \left(\left| \sin \left(\frac{\pi n}{L+1} \right) \right|^\alpha \right), \quad n = 1, 2, \dots, L.$$

Шаг 1 алгоритма GSSA:

$$X^{(\alpha)} = \mathcal{T}_L^{(\alpha)}(X) = [X_1^{(\alpha)} : \dots : X_K^{(\alpha)}],$$

$$X_i^{(\alpha)} = (w_1 x_{i-1}, \dots, w_L x_{i+L-2})^T, \quad 1 \leq i \leq K.$$

Шаги 2-4: аналогичны **SSA**.

Замечание 1

При $\alpha = 0$, **GSSA** — в точности базовый алгоритм **SSA**.

Замечание 2

$w^{(a)}$ называются степенными синусными весами. Они могут иметь другой вид.

Определение 4

Пусть $X = (\dots, x_{-1}, x_0, x_1, \dots)$ — бесконечный временной ряд. **Линейный конечный фильтр** — оператор Φ , преобразующий X в $Y = (\dots, y_{-1}, y_0, y_1, \dots)$ по правилу:

$$y_j = \sum_{i=-r_1}^{r_2} h_i x_{j-i}, \quad j \in \mathbb{Z},$$

где $r_1 + r_2 + 1$ — ширина фильтра, $h_i \in \mathbb{R}$ — коэффициенты.

Пример. При применении фильтра Φ к $x_j = \cos 2\pi\omega j$, получается ряд $y_j = A_\Phi(\omega) \cos(2\pi\omega j + \phi_\Phi(\omega))$.

$\phi_\Phi(\omega)$ — фазово-частотная характеристика (ФЧХ).

$A_\Phi(\omega)$ — амплитудно-частотная характеристика (АЧХ).

Сравнение SSA и GSSA. Линейные фильтры 2

$X = (x_1, \dots, x_N)$, $(\sqrt{\lambda}, U, V)$ – собственная тройка **SSA**.

$U = (u_1, \dots, u_L)$. $\tilde{X} = \mathcal{T}_L \circ \mathcal{H}(\sqrt{\lambda}UV^T)$.

Запись SSA через линейный фильтр для средних точек:

$$\tilde{x}_s = \sum_{j=-(L-1)}^{L-1} \left(\sum_{k=1}^{L-|j|} u_k u_{k+|j|} / L \right) x_{s-j}, \quad L \leq s \leq K.$$

Аналогичное представление для GSSA:

$$\tilde{x}_s = \sum_{j=-(L-1)}^{L-1} \left(\sum_{k=1}^{L-|j|} u_k^{(\alpha)} u_{k+|j|}^{(\alpha)} w_k / \sum_{i=1}^L w_i \right) x_{s-j}, \quad L \leq s \leq K.$$

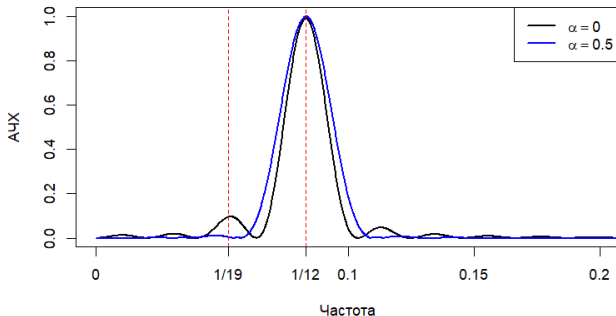
Замечание 1

Представление через линейные фильтры можно получить и для остальных точек ряда.

Сравнение SSA и GSSA. Пример

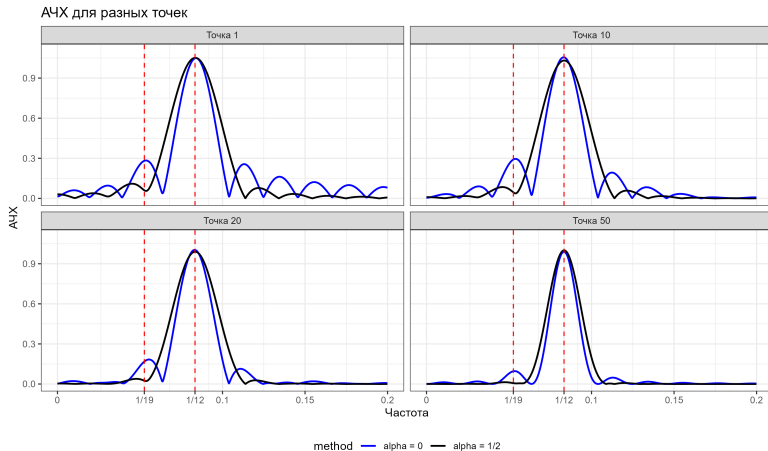
$X = X_{\sin} + X_{\cos} = \sin\left(\frac{2\pi}{12}n\right) + \frac{1}{2} \cos\left(\frac{2\pi}{19}n\right)$. $N = 96 \cdot 2 - 1$, $L = 48$.
Группировка: для X_{\sin} 1-2 SVD, для X_{\cos} 3-4 SVD.

АЧХ для суммы фильтров собственных троек синуса



$\alpha = 0.5$: шире полоса пропускания фильтра, чем при $\alpha = 0$, но нет волнообразного поведения на краях.

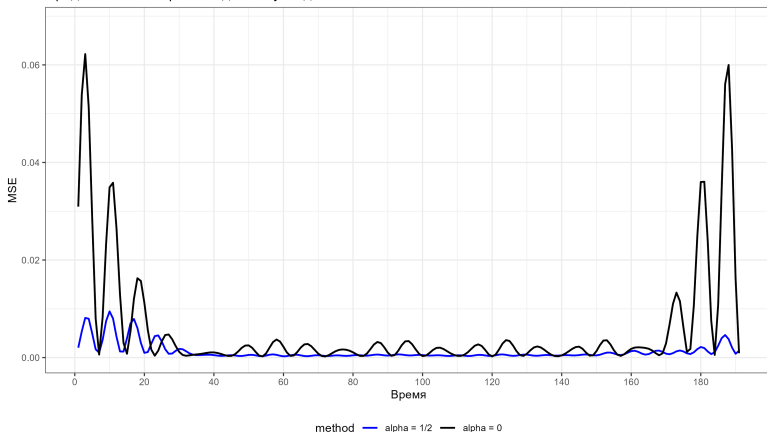
Сравнение SSA и GSSA. Пример продолжение



Таким образом, АЧХ фильтра также зависит от точки, для которой этот фильтр построен.

Сравнение SSA и GSSA. Пример продолжение 2

Средний MSE по времени для синусоиды



В начальных и конечных значениях ошибки больше.

Вывод. Вложенный вариант SSA + GSSA

Таблица 1: $X_{\sin} + X_{\cos} + \varepsilon_n$, $\varepsilon_n \sim N(0, 0.1^2)$, MSE оценок

Метод/Ошибка	X_{\sin}	X_{\cos}	X
SSA	5.68e-03	5.44e-03	7.48e-04
GSSA , $\alpha = 0.5$	1.21e-03	1.25e-03	1.04e-03
SSA + GSSA , $\alpha = 0.5$	1.06e-03	1.12e-03	7.15e-04

Получается вложенный вариант **SSA**.

Метод CiSSA. Алгоритм

$\mathbf{X} = (x_1, \dots, x_N)$ — временной ряд. $1 < L < N$ — длина окна.

Алгоритм CiSSA:

❶ Построение траекторной матрицы: как в SSA.

❷ $l = 1 : L$,

$$U_l = L^{-1/2}(u_{l,1}, \dots, u_{l,L}), \quad u_{l,j} = \exp(-i2\pi(j-1)\frac{l-1}{L}).$$

Элементарное разложение: $\omega_k = \frac{k-1}{L}$, $k = 1 : \lfloor \frac{L+1}{2} \rfloor$

$$\mathbf{X}_{\omega_k} = U_k U_k^H \mathbf{X} + U_{L+2-k} U_{L+2-k}^H \mathbf{X};$$

$$\mathbf{X}_{\omega_{\frac{L}{2}+1}} = U_{\frac{L}{2}+1} U_{\frac{L}{2}+1}^H \mathbf{X}, \text{ если } L \bmod 2 = 0,$$

Разложение: $\mathbf{X} = \sum_{k=1}^d \mathbf{X}_{\omega_k}$, $d = \lfloor \frac{L+1}{2} \rfloor$ (или $\frac{L}{2} + 1$).

❸ Группировка по частотам:

$$\bigsqcup_{j=1}^m \Omega_j = \bigsqcup_{j=1}^m [\omega_j^{(l)}, \omega_j^{(r)}] = [0, 0.5]. \quad \mathbf{X}_{\Omega_j} = \sum_{\omega_k \in \Omega_j} \mathbf{X}_{\omega_k}.$$

❹ Диагональное усреднение: как в SSA.

- ① **SSA**: базис адаптивный (зависит от X, L, N).
CiSSA: базис фиксированный (зависит от L, N).
- ② **CiSSA** – разложения Фурье для K векторов матрицы X с последующим диагональным усреднением слагаемых.
- ③ В **CiSSA** группировка по диапазонам частот. Алгоритм применим только, когда заранее известны частоты интересующих компонент.

Сравнение SSA, Фурье, CiSSA. Точная делимость

Фиксируем временной ряд $X = X_1 + X_2 =$
 $= A_1 \cos(2\pi\omega_1 n + \varphi_1) + A_2 \cos(2\pi\omega_2 n + \varphi_2).$

Метод	Условия точной делимости
SSA	$L\omega_1, L\omega_2, K\omega_1, K\omega_2 \in \mathbb{N}, \omega_1 \neq \omega_2, A_1 \neq A_2$
SSA EOSSA	$\omega_1 \neq \omega_2$
Фурье	$N\omega_1, N\omega_2 \in \mathbb{N}, \omega_1 \neq \omega_2$
CiSSA	$L\omega_1, L\omega_2 \in \mathbb{N}, \omega_1 \neq \omega_2$

Таким образом, условия на разделение косинусов, слабее у методов **CiSSA** и **Фурье**, чем у **SSA**.

Сравнение SSA, Фурье, CiSSA. Асимптотическая делимость

Метод	Полиномы	Гармоники	Эксп.-мод. функции
SSA	✓	✓	✓
SSA EOSSA	✓	✓	✓
Фурье	✗	✓	✗
CiSSA	✗	✓	✓

✓ — класс функций асимптотически делим методом.

Когнитивная нагрузка (Yedukondalu и др. 2025)

- Разложили сигналы ЭЭГ (наборы MAT, STEW) с помощью CiSSA на частотно-временные компоненты для отслеживания мозговой активности.
- Создали новые признаки из компонент.
- Классифицировали когнитивную нагрузку (низкая/высокая или лёгкая/средняя/высокая) с KNN, SVM.

Таяние ледников (Dey и др. 2023)

- Рассматривается таяние ледников. Цель работы – отделить долгосрочную тенденцию от сезонных сигналов.
- Применили CiSSA ($L = 10$) к стратиграфии кернов для разделения долгосрочных трендов и сезонных сигналов (пыль, соль).

По полученным результатам, можно следующие выводы:

- ❶ **CiSSA** показывает себя лучше **Фурье**;
- ❷ На разделение периодических компонент для базового **SSA** накладываются более строгие ограничения относительно **CiSSA**. В остальных случаях **SSA** работает лучше;
- ❸ **SSA EOSSA** исправляет недостатки базового **SSA**.
- ❹ Имеет смысл вложенный вариант с **CiSSA**.

FSSA – метод разложения функциональных временных рядов, совмещающий подходы функционального PCA, SSA.

Вход:

- $\{y_t(s)\}_{t=1}^N$, $y_t(s) \in \mathcal{L}^2([0, 1])$.
- Длина окна L , базис.

Сравним с **2d-SSA**, **MSSA**.




Результаты данного исследования:

- Выявлены сильные и слабые стороны методов;
- Предложены собственные вложенные модификации;
- Методы реализованы на языке R.

Последующие действия:

- Рассмотрение FSSA;
- Реализация вложенного варианта с **CiSSA**.

Спасибо за внимание!

-  Bogalo, Juan, Pilar Poncela и Eva Senra (2020). «Circulant singular spectrum analysis: A new automated procedure for signal extraction». В: *Signal Processing* 177. ISSN: 0165-1684. DOI: 10.1016/j.sigpro.2020.107750. URL: <http://www.sciencedirect.com/science/article/pii/S0165168420303264>.
-  Dey, Rahul и др. (2023). «Application of visual stratigraphy from line-scan images to constrain chronology and melt features of a firn core from coastal Antarctica». В: *Journal of Glaciology* 69.273, с. 179—190. DOI: 10.1017/jog.2022.59.
-  Golyandina, Nina, Pavel Dudnik и Alex Shlemov (2023). «Intelligent Identification of Trend Components in Singular Spectrum Analysis». В: *Algorithms* 16.7, с. 353. DOI: 10.3390/a16070353. URL: <https://doi.org/10.3390/a16070353>.



Golyandina, Nina, Vladimir Nekrutkin и Anatoly Zhigljavsky (2001). *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman и Hall/CRC. URL:

https://www.academia.edu/34626051/Analysis_of_Time_Series_Structure_-_SSA_and_Related_Techniques.



Golyandina, Nina и Alex Shlemov (2015). «Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series». В: *Statistics and Its Interface* 8.3, с. 277—294. ISSN: 1938-7997. DOI: 10.4310/sii.2015.v8.n3.a3. URL: <http://dx.doi.org/10.4310/SII.2015.v8.n3.a3>.

Список литературы III



Gu, Jialiang и др. (2024). «Generalized singular spectrum analysis for the decomposition and analysis of non-stationary signals». В: *Journal of the Franklin Institute Accepted/In Press*. ISSN: 0016-0032. DOI: 10.1016/j.jfranklin.2024.106696. URL: <https://doi.org/10.1016/j.jfranklin.2024.106696>.



Yedukondalu, Jammisetty и др. (янв. 2025). «Cognitive load detection through EEG lead wise feature optimization and ensemble classification». В: *Scientific Reports* 15. DOI: 10.1038/s41598-024-84429-6.