

# Тензорный анализ сингулярного спектра

Хромов Никита Андреевич, гр.20.Б04-мм

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

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

Рецензент: к.ф.-м.н. Усевич К.Д.

Санкт-Петербург,  
2024

# Введение

$X_q = \left( x_1^{(q)}, x_2^{(q)}, \dots, x_N^{(q)} \right)^T$ ,  $x_i \in \mathbb{R}$  — одномерный временной ряд

$X = \left( X^{(1)} : X^{(2)} : \dots : X^{(Q)} \right)$  — многомерный временной ряд

$$X = T + P + E$$

$T$  — медленно меняющаяся компонента (тренд)

$P$  — периодическая компонента (сезонность)

$E$  — случайная компонента (шум)

## Возможные задачи:

- 1 Выделение сигнала из ряда: нахождение  $S = T + P$
- 2 Разделение сигнала: нахождение компонент  $T$  и  $P$

Возможный метод решения: Singular Spectrum Analysis ([SSA](#))

[Broomhead, King (1986a)], [Golyandina, Nekrutkin, Zhigljavsky (2001)],

и его многомерное расширение Multivariate SSA ([MSSA](#))

[Broomhead, King (1986b)]

**Цель:** реализация тензорных расширений методов SSA и MSSA, исследование их свойств с точки зрения точности выделения сигнала и разделения компонент, сравнение расширений с базовыми методами.

## HO-SSA: алгоритм

$X = \sum_{m=1}^M S_m + E$  — одномерный временной ряд,

$S \stackrel{\text{def}}{=} \sum_{m=1}^M S_m$  — сигнал.

**Параметры:**  $I, L < N$  — длины окна,  $I + L \leq N + 1$ ,  $J = N - I - L + 2$ ,  
 $R$  — число элементов разложения, относимых к сигналу,  
 $\mathcal{I}_1, \dots, \mathcal{I}_M \subseteq \{1, 2, \dots, R\}$ ,  $\mathcal{I}_i \cap \mathcal{I}_j = \emptyset$  — наборы индексов, относимых к компонентам сигнала.

**Схема алгоритма HO-SSA для разделения компонент сигнала**

- ① Вложение  $X \xrightarrow{I, L} \mathcal{X} \in \mathbb{R}^{I \times L \times J}$  — траекторный тензор,
- ② Разложение

$$\mathcal{X} = \sum_{i=1}^{d_1} \sum_{l=1}^{d_2} \sum_{j=1}^{d_3} \mathcal{Z}_{ilj} U_i^{(1)} \circ U_l^{(2)} \circ U_j^{(3)}, \quad d_1 \leq I, d_2 \leq L, d_3 \leq J$$

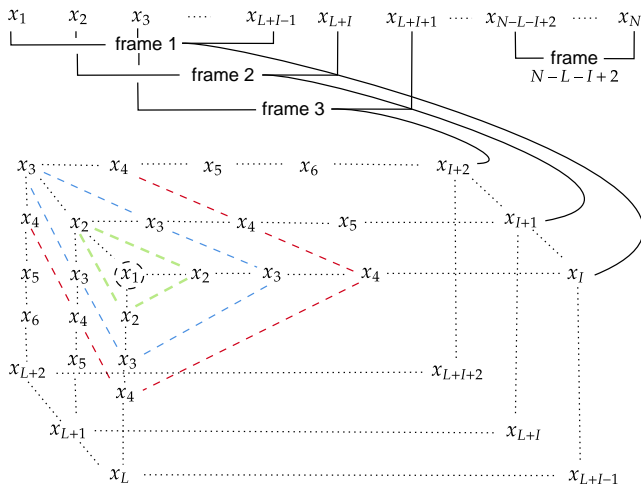
- ③ Группировка

$$\tilde{S}_m = \sum_{i \in \mathcal{I}_m} \sum_{l \in \mathcal{I}_m} \sum_{j \in \mathcal{I}_m} \mathcal{Z}_{ilj} U_i^{(1)} \circ U_l^{(2)} \circ U_j^{(3)}, \quad R \leq \min(d_1, d_2, d_3)$$

- ④ Восстановление усреднение  $\tilde{S}_m$  вдоль обобщённых антидиагоналей  $i + l + j = \text{const}$ .

**Результат алгоритма**  $\tilde{S}_m$  — оценки компонент  $S_m$

## HO-SSA: траекторный тензор



$I, L$  — параметры длины окна

## HO-SSA: разложение и группировка

## Определение

$n$ -ранг тензора  $\mathcal{X}$  ( $\text{rank}_n(\mathcal{X})$ ) — размерность пространства  $n$ -столбцов  $\mathcal{X}$ .

Идея выделения сигнала: приближение  $\mathcal{X}$  тензором  $\tilde{\mathcal{X}}$  меньших  $n$ -рангов.

Способы приближения меньшими рангами:

① Усечение HOSVD:

$$\mathcal{X} = \sum_{i=1}^{d_1} \sum_{l=1}^{d_2} \sum_{j=1}^{d_3} \mathcal{Z}_{ilj} U_i^{(1)} \circ U_l^{(2)} \circ U_j^{(3)} \mapsto \sum_{i=1}^{R_1} \sum_{l=1}^{R_2} \sum_{j=1}^{R_3} \mathcal{Z}_{ilj} U_i^{(1)} \circ U_l^{(2)} \circ U_j^{(3)} = \tilde{\mathcal{X}}$$

② Higher-Order Orthogonal Iteration (HOOI):

$$\tilde{\mathcal{X}} = \text{HOSVD} \left( \arg \min_{\mathcal{Y}} \|\mathcal{X} - \mathcal{Y}\| \right), \quad \text{rank}_n(\mathcal{Y}) = R_n$$

# HO-SSA: разделимость и ранги рядов в терминах SSA

- **Разделимость:** пусть  $S = S_1 + S_2$ ,  $S$ ,  $S_1$ ,  $S_2$  — траекторные матрицы этих сигналов с длиной окна  $L$  (траекторные тензоры с  $I=1$ )

## Определение (Разделимость в терминах SSA)

Сигналы  $S_1$  и  $S_2$   $L$ -разделимы, если существуют такие  $\mathcal{I}_1, \mathcal{I}_2$ , что

$$S = \sum_{j=1}^R \sqrt{\lambda_j} U_j V_j^T = \overbrace{\sum_{j \in \mathcal{I}_1} \sqrt{\lambda_j} U_j V_j^T}^{S_1} + \overbrace{\sum_{j \in \mathcal{I}_2} \sqrt{\lambda_j} U_j V_j^T}^{S_2}$$

- **Ранг ряда:** нахождение ранга аппроксимации  $R$ ?  
 $\Lambda^{(L)}(S)$  — пространство столбцов  $S$

## Определение (Ранг сигнала в терминах SSA)

$S$  имеет ранг  $R$ , если  $\forall L: R \leq \min(L, N - L + 1) \quad \dim \Lambda^{(L)}(S) = R$

# HO-SSA: разделимость

Пусть  $S = S_1 + S_2$ ,

$S, S_1, S_2$  — траекторные тензоры этих сигналов с длинами окна  $I$  и  $L$

## Теорема (О связи разделимости в SSA и HO-SSA)

*HOSVD тензора  $S$  можно представить в виде суммы HOSVD тензоров  $S_1$  и  $S_2$  тогда и только тогда, когда сигналы  $S_1$  и  $S_2$  слабо  $I$ - и  $L$ -разделимы в терминах SSA.*

## Замечание

*Теорема позволяет выделить класс сигналов, которые можно разделить методом HO-SSA, а также даёт рекомендации к выбору параметров  $I$  и  $L$ .*

HO-SSA:  $n$ -ранги траекторного тензора

## Теорема (О связи рангов рядов в SSA и HO-SSA)

Пусть сигнал  $S$  имеет ранг  $R$  в терминах SSA.

Тогда для любых значений параметров  $I$  и  $L$  таких, что

$$R \leq \min(I, L, N - I - L + 2),$$

$\text{rank}_1(\mathcal{S}) = \text{rank}_2(\mathcal{S}) = \text{rank}_3(\mathcal{S}) = R$ , где  $\mathcal{S}$  — траекторный тензор  $S$ , построенный по длинам окна  $I, L$ .

## Замечание

Теорема позволяет использовать известные результаты о рангах сигналов из теории SSA для аппроксимации траекторного тензора на шагах разложения и группировки в методе HO-SSA.



## HOSVD-MSSA: алгоритм

Пусть  $X = \sum_{m=1}^M S_m + E$  —  $Q$ -мерный временной ряд

**Параметры:**  $L$  — длина окна,  $K = N - L + 1$ ,

$R, \mathfrak{I}_1, \dots, \mathfrak{I}_M$  — как в HO-SSA,

$R_3$  — число элементов разложения по третьему направлению,

$\mathfrak{Q}_1, \dots, \mathfrak{Q}_M \subseteq \{1, 2, \dots, R_3\}$ ,  $\mathfrak{Q}_i \cap \mathfrak{Q}_j = \emptyset$  — индексы группировки по третьему направлению

- **Вложение**  $X \xrightarrow{L} \mathcal{X} \in \mathbb{R}^{L \times K \times Q}$  — траекторный тензор
- **Разложение**

$$\mathcal{X} = \sum_{l=1}^{d_1} \sum_{k=1}^{d_2} \sum_{q=1}^{d_3} \mathcal{Z}_{l k p} U_l^{(1)} \circ U_k^{(2)} \circ U_q^{(3)}$$

- **Группировка**

$$\tilde{S}_m = \sum_{l \in \mathfrak{I}_m} \sum_{k \in \mathfrak{I}_m} \sum_{q \in \mathfrak{Q}_m} \mathcal{Z}_{l k p} U_l^{(1)} \circ U_k^{(2)} \circ U_q^{(3)}$$

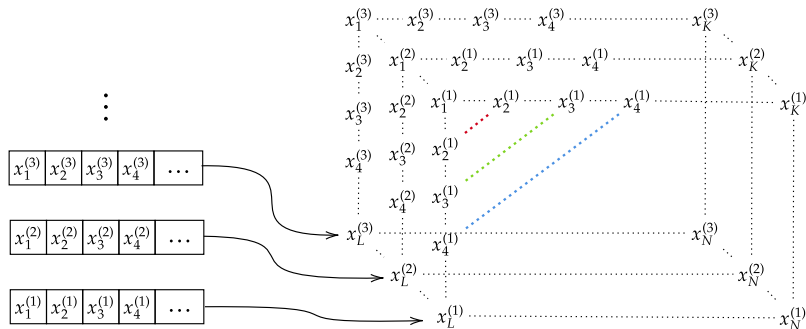
- **Восстановление** усреднение сечений третьего направления  $\tilde{S}_m$  по побочным диагоналям

**Результат алгоритма:**  $\tilde{S}_m$  — оценка  $S_m$

## HOSVD-MSSA: траекторный тензор многомерного ряда

$X$  — многомерный временной ряд длины  $N$

$L$  — длина окна,  $K = N - L + 1$



## HOSVD-MSSA: разделимость

Пусть  $\hat{S} = S_1 + S_2$ ,

$S$ ,  $S_1$ ,  $S_2$  — траекторные тензоры этих сигналов с длиной окна  $L$

$$\Lambda^{(I)}(S) = \text{span} \left\{ \left( s_i^{(q)}, s_{i+1}^{(q)}, \dots, s_{i+I-1}^{(q)} \right) \right\}$$

## Теорема

HOSVD тензора  $S$  можно представить в виде суммы HOSVD тензоров  $S_1$  и  $S_2$  тогда и только тогда, когда  $\Lambda^{(L)}(S_1) \perp \Lambda^{(L)}(S_2)$  и  $\Lambda^{(K)}(S_1) \perp \Lambda^{(K)}(S_2)$

## Замечание

Теорема позволяет выделить класс сигналов, которые можно разделить методом HOSVD-MSSA, а также даёт рекомендации к выбору параметра  $L$ .

HOSVD-MSSA:  $n$ -ранги траекторного тензора

## Теорема

Пусть  $S = (S^{(1)} : \dots : S^{(Q)})$ , тогда справедливы следующие утверждения.

- ①  $S$  имеет ранг  $R$  в терминах теории MSSA тогда и только тогда, когда для траекторного тензора  $\mathcal{S}$ , построенного по любой длине окна  $L < N$  такой, что  $R \leq \min(L, K)$  выполняется

$$\text{rank}_1(\mathcal{S}) = \text{rank}_2(\mathcal{S}) = R.$$

- ②  $\text{rank}_3(\mathcal{S})$  равен рангу матрицы, в строках которой содержатся одномерные сигналы  $S^{(q)}$ .

## Замечание

Теорема позволяет использовать известные результаты о рангах сигналов из теории MSSA для аппроксимации траекторного тензора по первым двум направлениям, а также даёт рекомендации к выбору ранга аппроксимации по третьему направлению на шагах разложения и группировки в методе HOSVD-MSSA.

## Численные результаты: сравнение HO-SSA с SSA

Пусть временной ряд имеет вид

$$X = (s_1 + \varepsilon_1, s_2 + \varepsilon_2, \dots, s_N + \varepsilon_N),$$

где  $N = 71$ ,  $s_n = 30 \cos(2\pi n/12)$ ,  $\varepsilon_n$  — шум.

Table: RMSE оценки сигнала: SSA.

Вид шума	$L = 12$	$L = 24$	$L = 30$	$L = 36$
Белый, $\sigma^2 = 25$	1.82	1.42	<b>1.40</b>	1.42
Красный, $\delta^2 = 5$ , $\varphi = 0.5$	1.31	1.03	<b>1.01</b>	1.03
Красный, $\delta^2 = 5$ , $\varphi = 0.9$	1.88	1.37	<b>1.34</b>	1.36

Table: RMSE оценки сигнала: HO-SSA.

Вид шума \ $I \times L$	$19 \times 30$	$12 \times 31$	$7 \times 36$	$12 \times 37$	$12 \times 49$
Белый, $\sigma^2 = 25$	1.62	1.56	<b>1.49</b>	1.53	1.63
Красный, $\delta^2 = 5$ , $\varphi = 0.5$	1.19	1.14	<b>1.08</b>	1.12	1.17
Красный, $\delta^2 = 5$ , $\varphi = 0.9$	1.51	1.44	<b>1.39</b>	1.42	1.56

## Численные результаты: сравнение HOSVD-MSSA с MSSA и 2D-SSA

$$X = (X_1 : \dots : X_Q), \quad X_q = \left( x_1^{(q)}, \dots, x_N^{(q)} \right)^T, \quad x_n^{(q)} = \hat{s}_n^{(q)} + \tilde{s}_n^{(q)} + \varepsilon_n^{(q)},$$

где  $\varepsilon_n^{(q)} \sim N(0, 0.01)$  и независимы, модель:  $s_n^{(q)} = C_q \cos(2\pi n \omega_q + \varphi_q)$

Вид сигнала	MSSA	HOSVD-MSSA	2D-SSA
Равные сигналы	0.026	0.019	<b>0.014</b>
	0.025	0.016	<b>0.014</b>
Различие амплитуд	0.029	<b>0.019</b>	0.086
	0.029	<b>0.019</b>	0.083
Линейные фазы	<b>0.026</b>	<b>0.025</b>	0.117
	<b>0.025</b>	<b>0.025</b>	0.114
Произвольные фазы	<b>0.026</b>	<b>0.025</b>	0.034
	<b>0.025</b>	<b>0.025</b>	0.033
Разделимость с const	<b>0.017</b>	<b>0.017</b>	0.023
	0.025	<b>0.019</b>	0.033
Различие частот	0.024	0.018	<b>0.012</b>
	0.024	<b>0.018</b>	0.031
	0.024	<b>0.014</b>	0.026
Ортогональность по каналам	0.031	<b>0.023</b>	0.025
	0.030	<b>0.022</b>	0.024

## Численные результаты: смещение и дисперсия

$X = \hat{S} + \tilde{S} + E$ , где  $\hat{s}_n^{(1)} = 3$ ,  $\hat{s}_n^{(2)} = -1.5$ ,  $\tilde{s}_n^{(1)} = \cos(2\pi n/20)$  и  $\tilde{s}_n^{(2)} = 2 \cos(2\pi n/20)$ , а  $\varepsilon_n^{(p)} \sim N(0, \sigma^2)$  и независимы.

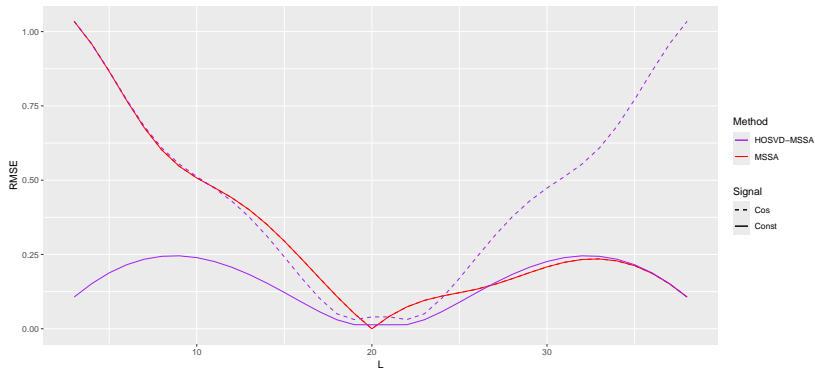


Figure: Зависимость RMSE оценки сигнала от длины окна:  $\sigma = 0$

## Численные результаты: смещение и дисперсия

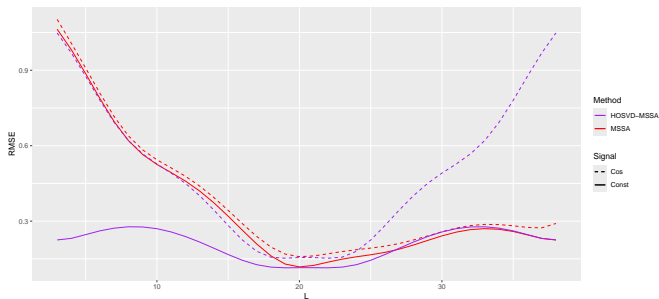


Figure:  
Зависимость  
RMSE оценки  
сигнала от  
длины окна:  
 $\sigma = 0.5$

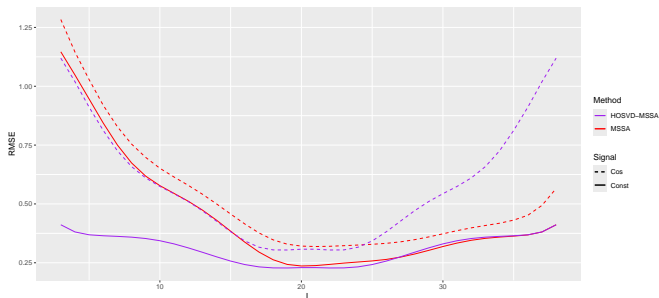
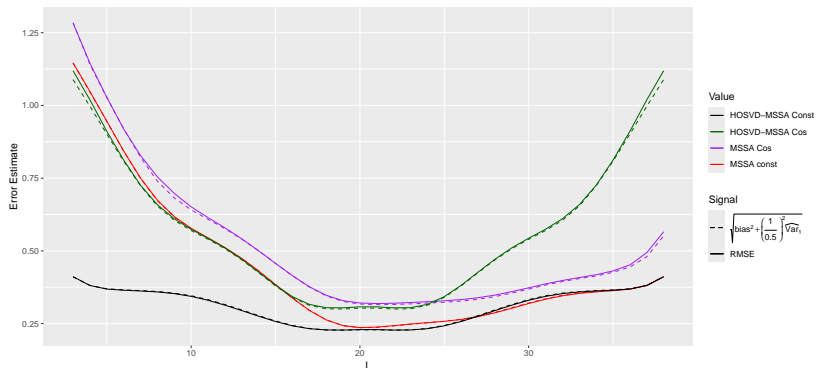


Figure:  
Зависимость  
RMSE оценки  
сигнала от  
длины окна:  
 $\sigma = 1$



## Численные результаты: смещение и дисперсия



**Figure:** Сравнение значений RMSE при  $\sigma = 1$  и аппроксимации  $\sqrt{\text{bias}^2 + \widehat{\text{Var}}_2}$  в зависимости от длины окна

# Результаты

## ④ Свойства HO-SSA

- Критерий разделимости HO-SSA, связь с разделимостью SSA
- Связь рангов траекторного тензора с SSA-рангом ряда
- Трудоёмкость алгоритма

## ② Свойства HOSVD-MSSA

- Критерий разделимости HOSVD-MSSA
- Теорема о рангах траекторного тензора, связь с MSSA-рангом и 3-ранг
- Утверждение о симметричности траекторного тензора относительно замены длины окна  $L$  на  $K$

## ③ Численные выводы

- Отсутствие преимущества HO-SSA над SSA
- Преимущество HOSVD-MSSA над MSSA в условиях разделимости, и над 2D-SSA для гармонических сигналов с равными амплитудами
- Возможное преимущество HOSVD-MSSA над MSSA в условиях приближительной разделимости для HOSVD-MSSA и точной для MSSA за счёт меньшей дисперсии оценки тензорным методом

## ④ Реализация методов HO-SSA и HOSVD-MSSA на языке R в стиле пакета Rssa (исходный код опубликован в репозитории Zenodo)