

# Тензорный анализ сингулярного спектра

Хромов Никита Андреевич, гр.20.Б04-мм

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

Отчет по производственной практике (научно-исследовательская  
работа) (6 семестр)

Санкт-Петербург, 2023

# Введение

Временной ряд длины  $N$ :  $X = (x_1, x_2, \dots, x_N)$ .

$$X = T + P + R.$$

Возможные задачи:

- 1 Выделение сигнала из ряда: нахождение  $T + P$ ;
- 2 Отделение компонент сигнала: нахождение  $T$  и  $P$ .

Одним из методов решения этих задач является метод Singular Spectrum Analysis ([SSA](#)) Golyandina et al. (2001), Analysis of time series structure: SSA and related techniques.

# Введение

**SSA:** ряд  $X \Rightarrow$  матрица  $\mathbf{X} \Rightarrow$  SVD  $\mathbf{X}$ ,

**Tensor SSA:** ряд  $X \Rightarrow$  тензор  $\mathcal{X} \Rightarrow$  тензорное разложение  $\mathcal{X}$ .

- Примеры тензорных разложений:
  - ① High-order singular value decomposition (**HOSVD**);
  - ② Canonical polyadic decomposition (CPD).
- Для выделения сигнала и оценки его параметров в работе Papu et al. (2005) применялось HOSVD, и в работе De Lathauwer et al. (2011) применялось CPD.

**Задача:** изучить различные варианты тензорных разложений, применить их в задаче анализа временного ряда и сравнить результаты с SSA.

Из тензорных разложений первым было выбрано HOSVD, так как оно имеет наибольшее сходство с SVD.

# Описание HOSDV

Пусть имеется тензор  $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_M}$ , тогда HOSVD  $\mathcal{A}$ :

$$\mathcal{A} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_M=1}^{I_M} \mathcal{Z}_{i_1, i_2, \dots, i_M} U_{i_1}^{(1)} \circ U_{i_2}^{(2)} \circ \dots \circ U_{i_M}^{(M)},$$

где

- $\mathbf{U}^{(n)} = [U_1^{(n)}, \dots, U_{I_n}^{(n)}]$  — унитарные матрицы;
- Тензор  $\mathcal{Z} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_M}$  удовлетворяет свойствам
  - 1 полная ортогональность:

$$\langle \mathcal{Z}_{i_n=\alpha}, \mathcal{Z}_{i_n=\beta} \rangle = 0 \quad \alpha \neq \beta,$$

- 2 упорядоченность:

$$\|\mathcal{Z}_{i_n=1}\| \geq \|\mathcal{Z}_{i_n=2}\| \geq \dots \geq \|\mathcal{Z}_{i_n=I_n}\|.$$

# Свойства HOSVD

Все свойства представлены в работе De Lathauwer et al. (2000).

- HOSVD — единственное  $M$ -ортогональное разложение.
- При  $M = 2$  HOSVD совпадает с SVD.
- Пусть  $\text{rank}_n(\mathcal{A})$  — размерность пространства векторов измерения  $n$  тензора. Если в HOSVD тензора  $\mathcal{A}$   $r_n$  — наибольший индекс такой, что  $\|\mathcal{Z}_{i_n=r_n}\| > 0$ , то  $r_n = \text{rank}_n(\mathcal{A})$ .

- 

$$\|\mathcal{A}\|^2 = \sum_{i=1}^{R_1} \left(\sigma_i^{(1)}\right)^2 = \sum_{i=1}^{R_2} \left(\sigma_i^{(2)}\right)^2 = \dots = \sum_{i=1}^{R_M} \left(\sigma_i^{(M)}\right)^2 = \|\mathcal{Z}\|^2$$

$$\sigma_i^{(n)} = \|\mathcal{Z}_{i_n=i}\|, \quad R_n = \text{rank}_n(\mathcal{A}).$$

## Свойства HOSVD

- Векторы тензора  $\mathcal{A}$  по измерению  $n$  в основном содержат вклады в направлении  $U_1^{(n)}$ , величина этого вклада равна  $\sigma_1^{(n)^2}$ . Следующий по величине вклад по измерению  $n$  достигается в направлении  $U_2^{(n)}$ , перпендикулярном  $U_1^{(n)}$ , с величиной  $\sigma_2^{(n)^2}$ , и т.д.
- Определим тензор  $\hat{\mathcal{A}}$  отбрасыванием наименьших сингулярных значений  $\sigma_{I'_n+1}^{(n)}, \sigma_{I'_n+2}^{(n)}, \dots, \sigma_{R_n}^{(n)}$ , тогда

$$\|\mathcal{A} - \hat{\mathcal{A}}\|^2 \leq \sum_{i_1=I'_1+1}^{R_1} \left(\sigma_{i_1}^{(1)}\right)^2 + \dots + \sum_{i_M=I'_M+1}^{R_M} \left(\sigma_{i_M}^{(M)}\right)^2.$$

# Описание метода HOSVD SSA

Имеется временной ряд  $X = (x_1, x_2, \dots, x_N)$ . Приведём формулировки алгоритма HOSVD SSA для решения различных задач.

**Входные данные алгоритма:**  $I, L : 1 \leq I, L \leq N, I + L \leq N + 1$ .

**Траекторный тензор:** тензор размерности

$I \times L \times J = N - I - J + 2$ , строится по ряду:

$$\mathcal{X}_{i,l,j} = x_{i+l+j-2} \quad i \in \overline{1:I}, l \in \overline{1:L}, j \in \overline{1:J}.$$

Слои траекторного тензора:

$$\mathcal{X}_{:,j} = \begin{pmatrix} x_j & x_{j+1} & \dots & x_{j+L-1} \\ x_{j+1} & x_{j+2} & & \vdots \\ \vdots & & \ddots & \vdots \\ x_{j+I-1} & \dots & \dots & x_{j+I+L-2} \end{pmatrix}.$$

# HOSVD SSA: отделение компонент сигнала

**Задача:** отделение компонент сигнала в ряде.

- 1 **Вложение:** выбор параметров  $I, L$  и построение по ним траекторного тензора  $\mathcal{X}$ ;
- 2 **Разложение:** Проведение HOSVD траекторного тензора  $\mathcal{X}$ , получение его представления в виде

$$\mathcal{X} = \sum_{i=1}^I \sum_{l=1}^L \sum_{j=1}^J z_{i,l,j} \mathbf{U}_i^{(1)} \circ \mathbf{U}_l^{(2)} \circ \mathbf{U}_j^{(3)};$$



# HOSVD SSA: отделение компонент сигнала

- ③ **Группировка:** разбиение множества индексов  $\mathfrak{S} = \{1, 2, \dots, \min(I, L, J)\}$  по смыслу на непересекающиеся множества  $\mathfrak{S}_k$ ,  $k \in \overline{1:m}$  и построение по этому разбиению тензоров

$$\mathcal{X}^{(\mathfrak{S}_k)} = \sum_{i \in \mathfrak{S}_k} \sum_{l \in \mathfrak{S}_k} \sum_{j \in \mathfrak{S}_k} \mathcal{Z}_{i,l,j} \mathbf{U}_i^{(1)} \circ \mathbf{U}_l^{(2)} \circ \mathbf{U}_j^{(3)}.$$

- ④ **Восстановление:** получение рядов  $\mathbf{X}^{(k)} = \mathbf{X}^{(\mathfrak{S}_k)}$  по тензорам  $\mathcal{X}^{(\mathfrak{S}_k)}$  посредством их усреднения вдоль плоскостей  $i + l + j = \text{const}$ .

**Результат алгоритма:** набор рядов  $\mathbf{X}^{(k)}$  таких, что

$$\mathbf{X} = \sum_{k=1}^m \mathbf{X}^{(k)}.$$

# HOSVD SSA: отделение сигнала усечением HOSVD

**Задача:** выделение сигнала из ряда.

Способы реализации HOSVD SSA для решения этой задачи: **усечение HOSVD** и High-order orthogonal iteration (**HOOI**).  
Начнём с усечения HOSVD.

Первые два шага в этом алгоритме — вложение и разложение, были описаны выше.

- ③ **Усечение:** выбор ранга сигнала  $r$  и обнуление матриц-слоёв тензора  $\mathcal{Z}$  с номерами  $k > r$  по каждому измерению. Построение по этому усечению тензора  $\hat{\mathcal{X}}$ .
- ④ **Восстановление:** усреднение тензора  $\hat{\mathcal{X}}$  вдоль плоскостей  $i + l + j = \text{const}$ .

**Результат алгоритма:** полученный усреднением ряд  $\hat{X}$  будем считать сигналом.

# HOSVD SSA: отделение сигнала с помощью HOOI

**Задача:** выделение сигнала из ряда.

Второй способ использует HOOI — метод приближения тензора другим тензором с меньшими значениями  $n$ -рангов. В отличие от усечения, этот метод является оптимальным.

Первый шаг алгоритма — вложение, совпадает с предыдущими алгоритмами, поэтому опишем его начиная со второго шага.

- ② **HOOI:** Выбор ранга сигнала  $r$  и применение к  $\mathcal{X}$  HOOI с набором  $n$ -рангов  $(r, r, r)$ . Результат — оптимальное приближение тензором  $\hat{\mathcal{X}}$  с  $n$ -рангами  $r$ .
- ③ **Восстановление:** усреднение тензора  $\hat{\mathcal{X}}$  аналогично восстановлению в варианте с усечением.

**Результат алгоритма:** полученный усреднением ряд  $\hat{X}$  будем считать сигналом.

# Свойства HOSVD SSA, делимость рядов

**Делимость** и **ранг** рядов являются важными понятиями в теории SSA. Рассмотрим эти понятия для алгоритма HOSVD SSA

## 1 Делимость рядов

### Утверждение

*Если временные ряды  $\tilde{X}$  и  $\hat{X}$  длины  $N$  слабо  $I$ - и  $L$ -делимы в смысле теории SSA, то существует такое HOSVD траекторного тензора  $\mathcal{X}$  ряда  $X = \tilde{X} + \hat{X}$ , что его можно в виде суммы HOSVD траекторных тензоров рядов  $\tilde{X}$  и  $\hat{X}$ .*

Понятие слабой делимости рядов из SSA применимо к тензорному случаю

# Свойства HOSVD SSA, ранг ряда

## 2 Ранг ряда

### Утверждение

*Пусть временной ряд  $X$  имеет конечный ранг  $d$  в терминах SSA. Тогда для любых значений параметров  $I$  и  $L$  таких, что*

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

*количество ненулевых сингулярных чисел по каждому измерению в HOSVD траекторного тензора  $\mathcal{X}$  этого ряда с параметрами  $I$  и  $L$  будет равно  $d$ .*

Понятие ранга ряда имеет тот же смысл в терминах HOSVD SSA, что и в стандартной теории SSA, причём ряды конечного ранга имеют одинаковые ранги в тензорном и стандартном случаях.

# Примеры слабой разделимости

- Экспонента и косинус с экспоненциально-модулированной амплитудой:  $\tilde{x}_n = e^{-\alpha n} \cos(2\pi n/T + \varphi)$ ,  $\hat{x}_n = e^{\alpha n}$ ,  $n \in \overline{1:N}$ .

Если  $(N+2) \vdots T$ , то при выборе  $I, L: I+L < N+1$ , делящихся нацело на  $T$   $\tilde{X}$  и  $\hat{X}$  слабо разделимы.

- Два косинуса:  $\tilde{x}_n = \cos(2\pi\omega n + \varphi)$ ,  $\hat{x}_n = \cos(2\pi\omega' n + \varphi')$ ,  $0 < \omega < 1/2$ . тогда ряд  $\tilde{X}$  отделим от ряда  $\hat{X}$  в смысле Tensor SSA тогда и только тогда, когда  $\tilde{X}$  и  $\hat{X}$  слабо разделимы тогда и только тогда, когда  $\omega \neq \omega'$ ,  $I, L, J > 2$  и  $I\omega, I\omega', L\omega, L\omega', J\omega, J\omega'$  — целые числа.

В последнем примере при одинаковых амплитудах происходит смешение компонент и в теории SSA говорят, что наблюдается приближённая разделимость.

Как формализовать приближённую и точную разделимость в теории HOSVD SSA пока неясно.

# Сравнение HOSVD SSA и Basic SSA

- $x_n = 2e^{0.035n}$ , шум — белый гауссовский,  $\sigma^2 = 2.25$ ;
- $x_n = 2 + 0.1n$ , шум — белый гауссовский,  $\sigma_1^2 = 2.25$ ,  $\sigma_2^2 = 0.04$ ;
- $x_n = 30 \cos(2\pi n/12)$ , шум — белый гауссовский,  $\sigma^2 = 25$ , и красный,  $\delta = \sqrt{5}$ ,  $\varphi_1 = 0.5$ ,  $\varphi_2 = 0.9$ .

Оценка точности — RMSE по 500 реализациям шума:  $X$  — ряд длины  $N$ ,  $S_i$  — выделенный в  $i$ -м эксперименте сигнал, тогда

$$\widehat{\text{RMSE}}(m) = \sqrt{\frac{1}{m} \sum_{i=1}^m \widehat{\text{MSE}}(S_i, X)}, \quad \widehat{\text{MSE}}(S, X) = \frac{1}{N} \sum_{k=1}^N (s_i - x_i)^2.$$

Во всех этих случаях метод SSA показал точность отделения сигнала выше, чем HOSVD SSA.

## Сравнение HOSVD SSA и Basic SSA

**Таблица:** RMSE восстановленного с помощью SSA сигнала, порождённого косинусом.  $N = 71$

$L$ вид шума	12	24	30	36
белый, $\sigma^2 = 25$	1.82	1.42	<b>1.40</b>	1.42
красный, $\varphi = 0.5$	1.31	1.03	<b>1.01</b>	1.03
красный, $\varphi = 0.9$	1.88	1.37	<b>1.34</b>	1.36

**Таблица:** RMSE восстановленного с помощью HOSVD SSA с использованием HOOI сигнала, порождённого косинусом.  $N = 71$

$I \times L$ вид шума	12×12	12×24	12×30	24×24	24×30	30×36
белый, $\sigma^2 = 25$	1.63	1.53	1.56	1.65	1.62	<b>1.49</b>
красный, $\varphi = 0.5$	1.17	1.12	1.14	1.21	1.19	<b>1.08</b>
красный, $\varphi = 0.9$	1.56	1.42	1.44	1.54	1.51	<b>1.39</b>



# Особый случай

Приведём особый пример, в котором HOSVD SSA оказался точнее базового SSA.

$$x_n = \sin(2\pi n/3 + \pi/2), n \in \overline{1:9}.$$

Шум: красный с параметрами  $\delta = 0.1$ ,  $\varphi = 0.9$ .

**Таблица:** RMSE восстановленного с помощью различных методов короткого сигнала, порождённого синусом.

SSA	HOSVD SSA (HOSVD)	HOSVD SSA (HOOI)
0.116	0.110	<b>0.095</b>

# Заклучение





Таким образом:

- Описан и реализован алгоритм HOSVD SSA.
- Выведены аналоги важных свойств SSA для метода HOSVD SSA.
- Описаны примеры использования метода.
- Проведено численное сравнение точности методов HOSVD SSA и SSA, найден особый случай.

Остаётся для изучения:

- Изучение причин возникновения особого случая.
- Определение точной и приближённой разделимости, нахождение метода отделения компонент при отсутствии точной.
- Возможность применения других тензорных разложений (CPD).
- Подтверждение результатов, утверждающих о преимуществах Tensor SSA над SSA.

# Список литературы

-  Golyandina Nina, Nekrutkin Vladimir, Zhigljavsky Anatoly. Analysis of time series structure: SSA and related techniques. — Chapman & Hall/CRC, 2001.
-  Papy J. M., De Lathauwer L., Van Hu el S. Exponential data tting using multilinear algebra: the single-channel and multi-channel case // Numerical Linear Algebra with Applications. — 2005. — P. 809–826.
-  De Lathauwer Lieven. Blind Separation of Exponential Polynomials and the Decomposition of a Tensor in Rank- $(L_r, L_r, 1)$  Terms // SIAM Journal on Matrix Analysis and Applications. — 2011. — P. 1451–1474.
-  De Lathauwer Lieven, De Moor Bart, Vandewalle Joos. A Multilinear Singular Value Decomposition // SIAM Journal on Matrix Analysis and Applications. — 2000. — P. 1253–1278.