Санкт-Петербургский государственный университет
Прикладная математика и информатика

Учебная практика 3 (научно-исследовательская работа)

# «Tensor SSA для анализа временного ряда»

Выполнил:
Хромов Никита Андреевич
20.Б04-мм

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

Санкт-Петербург
2023

# Оглавление

# 1. Введение

Здесь должно быть введение.

## 2. Построение тензора и его разложение

Дан временной ряд $f$ длины $N$

$$f = (f_1, f_2, \dots, f_N).$$

Выбираются два натуральных параметра $I, L : I + L - 1 \leqslant N$, по ним высчитывается третий параметр $J = N - I - L + 2$. С учётом этих параметров строится траекторный тензор $\mathcal{X}$ размерности $I \times L \times J$ следующим образом

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

Слои тензора будут иметь следующий вид

$$\mathcal{X}_{,,j} = \begin{pmatrix} f_j & f_{j+1} & \cdots & f_{j+L-1} \\ f_{j+1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ f_{j+I-1} & \cdots & \cdots & f_{j+I+L-2} \end{pmatrix},$$

$$\mathcal{X}_{,l,} = \begin{pmatrix} f_l & f_{l+1} & \cdots & f_{l+J-1} \\ f_{l+1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ f_{l+I-1} & \cdots & \cdots & f_{l+I+J-2} \end{pmatrix},$$

$$\mathcal{X}_{i,,} = \begin{pmatrix} f_i & f_{i+1} & \cdots & f_{i+J-1} \\ f_{i+1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ f_{i+L-1} & \cdots & \cdots & f_{i+L+J-2} \end{pmatrix}.$$

К полученному тензору применяется HOSVD [1] — тензорное разложение, являющееся продолжением SVD на большие размерности. Результатом разложения является набор из одного тензора $\mathcal{Z}$ размерности $I \times L \times J$ и трёх ортогональных матриц $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, $\mathbf{U}^{(3)}$ размерностей $I \times I$, $L \times L$, $J \times J$ соответственно.

Этот набор удовлетворяет равенству

$$\mathcal{X} = \mathcal{Z} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)},$$

где $\times_n$ — произведение тензора на матрицу по $n$-му измерению. Оно определяется следующим образом: пусть $\mathcal{A}$ — тензор размерности $I_1 \times I_2 \times \ldots \times I_K$, $\mathbf{U}$ — матрица размерности

$J_n \times I_n$, тогда $\mathcal{A} \times_n \mathbf{U}$ — тензор размерности $I_1 \times I_2 \times \ldots \times I_{n-1} \times J_n \times I_{n+1} \times \ldots \times I_K$, который считается по формуле

$$(\mathcal{A} \times_n \mathbf{U})_{i_1 i_2 \ldots i_{n-1} j_n i_{n+1} \ldots i_K} = \sum_{i_n=1}^{I_n} a_{i_1 i_2 \ldots i_{n-1} i_n i_{n+1} \ldots i_K} u_{j_n i_n}.$$

Обозначим за $\mathcal{Z}_{i_n=\alpha}$ подтензор тензора $\mathcal{Z}$, полученный фиксированием индекса $i_n = \alpha$. Тензор $\mathcal{Z}$ удовлетворяет следующим свойствам:

1. подтензоры $\mathcal{Z}_{i_n=\alpha}$ и $\mathcal{Z}_{i_n=\beta}$ ортогональны для всех возможных значений $n$, $\alpha$, $\beta$ : $\alpha \neq \beta$:
$$\langle \mathcal{Z}_{i_n=\alpha}, \mathcal{Z}_{i_n=\beta} \rangle = 0 \qquad \alpha \neq \beta,$$
2. подтензоры расположены в порядке убывания их евклидовой нормы:
$$\|\mathcal{Z}_{i_n=1}\| \geqslant \|\mathcal{Z}_{i_n=2}\| \geqslant \ldots \geqslant \|\mathcal{Z}_{i_n=I_n}\|$$
для всех возможных значений $n$.

**Определение 2.1.** *Обозначим* $\sigma_i^{(n)} = \|\mathcal{Z}_{i_n=i}\|$ *и будем называть* $\sigma_i^{(n)}$ $i$*-м сингулярным числом тензора* $\mathcal{X}$ *по измерению* $n$.

**Определение 2.2.** *Векторы* $\mathbf{U}_i^{(n)}$ *будем называть* $i$*-м сингулярным вектором тензора* $\mathcal{X}$ *по измерению* $n$.

**Замечание 1.** *Вычисление* HOSVD *тензора* $\mathcal{A}$ *с* $N$ *размерностями сводится к вычислению* SVD *на* $N$ *матрицах* $\mathbf{A}_{(n)}$, *которые вычисляются развёрткой тензора по* $n$*-му измерению [1].*

Другими словами, если $\mathcal{A}$ — тензор размерности $I_1 \times I_2 \times \ldots \times I_N$, то его развёртка по $n$-му измерению — это матрица $\mathbf{A}_{(n)}$ размерности $I_n \times I_{n+1} I_{n+2} \ldots I_N I_1 I_2 \ldots I_{n-1}$, в которой элемент $a_{i_1 i_2 \ldots i_N}$ тензора содержится в строке $i_n$ и столбце с номером равным

$$(i_{n+1}-1) I_{n+2} I_{n+3} \ldots I_N I_1 I_2 \ldots I_{n-1} + (i_{n+2}-1) I_{n+3} I_{n+4} \ldots I_N I_1 I_2 \ldots I_{n-1} + \ldots$$
$$+ (i_N - 1) I_1 I_2 \ldots I_{n-1} + (i_1 - 1) I_2 I_3 \ldots I_{n-1} + (i_2 - 1) I_3 I_4 \ldots I_{n-1} + \cdots + i_{n-1}.$$

К каждой из полученных матриц применяется SVD, в результате чего получаются $N$ матриц $\mathbf{U}^{(n)}$, составленных из левых сингулярных векторов соответствующих развёрток. Затем находится тензор сингулярных чисел

$$\mathcal{Z} = \mathcal{A} \times_1 \mathbf{U}^{(1)^{\mathrm{H}}} \times_2 \mathbf{U}^{(2)^{\mathrm{H}}} \ldots \times_N \mathbf{U}^{(N)^{\mathrm{H}}}.$$

В результате получается искомое разложение

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \ldots \times_N \mathbf{U}^{(N)}.$$

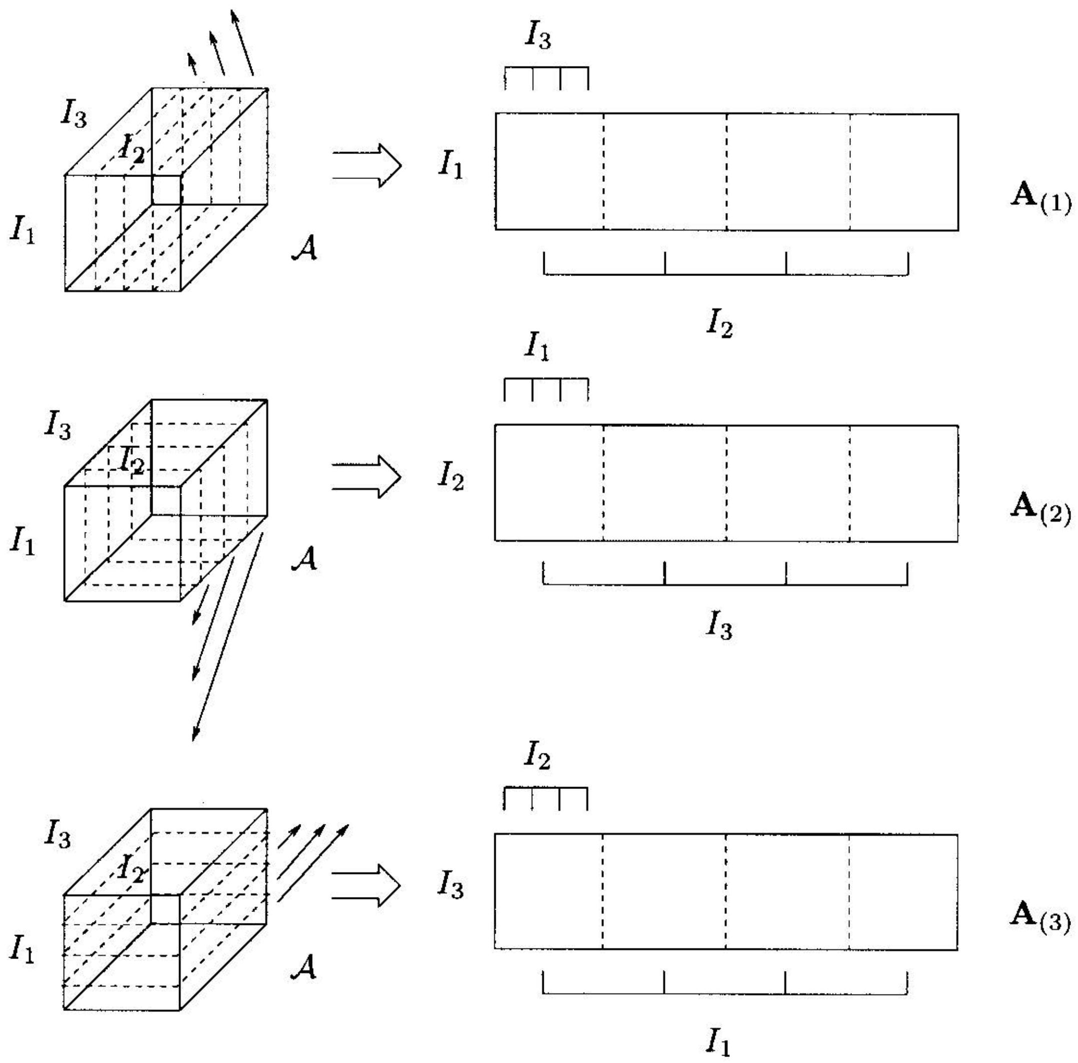


Рис. 1. Развёртка тензора $\mathcal{A}$ размерности $I_1 \times I_2 \times I_3$ в матрицы $\mathbf{A}_{(1)}$, $\mathbf{A}_{(2)}$, $\mathbf{A}_{(3)}$ размерностей $I_1 \times (I_2 I_3)$, $I_2 \times (I_3 I_1)$, $I_3 \times (I_1 I_2)$ соответственно

## 3. Свойства TSSA

В силу аналогичности свойств SVD и HOSVD, многие определения и свойства из теории SSA можно перенести на тензорный случай.

**Утверждение 3.1.** $\tilde{F} = (\tilde{f}_1, \ldots, \tilde{f}_N)$, $\hat{F} = (\hat{f}_1, \ldots, \hat{f}_N)$ *— временные ряды длины* $N$. *Пусть ряд* $F$ *является суммой этих рядов. Траекторные тензоры рядов равны соответственно:* $\tilde{\mathcal{X}}$, $\hat{\mathcal{X}}$, $\mathcal{X}$. *Тогда существует сингулярное разложение тензора* $\mathcal{X}$ *с параметрами* $I, L$, *которое можно представить в виде суммы сингулярных разложений тензоров* $\tilde{\mathcal{X}}$ *и* $\hat{\mathcal{X}}$ *с теми же параметрами в том и только том случае, когда взаимно ортогональны все подряды рядов* $\tilde{F}$ *и* $\hat{F}$ *длины* $I$, $L$, $J = N - I - L + 2$, *то есть*

1. $\tilde{f}_k \hat{f}_m + \ldots + \tilde{f}_{k+I-1} \hat{f}_{m+I-1} = 0 \qquad \forall k,\, m \in \overline{1 : N - I + 1}$,

*2.* $\tilde{f}_k\hat{f}_m + \ldots + \tilde{f}_{k+L-1}\hat{f}_{m+L-1} = 0 \qquad \forall k,\, m \in \overline{1:N-L+1}$,

*3.* $\tilde{f}_k\hat{f}_m + \ldots + \tilde{f}_{k+J-1}\hat{f}_{m+J-1} = 0 \qquad \forall k,\, m \in \overline{1:N-J+1}$.

*Доказательство.* Сингулярные разложения тензоров $\mathcal{X}, \tilde{\mathcal{X}}, \hat{\mathcal{X}}$ могут быть представлены в виде следующих сумм:

$$\mathcal{X} = \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J} \mathcal{Z}_{i,l,j}\mathbf{U}_i^{(1)} \circ \mathbf{U}_l^{(2)} \circ \mathbf{U}_j^{(3)},$$
$$\tilde{\mathcal{X}} = \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J} \tilde{\mathcal{Z}}_{i,l,j}\tilde{\mathbf{U}}_i^{(1)} \circ \tilde{\mathbf{U}}_l^{(2)} \circ \tilde{\mathbf{U}}_j^{(3)},$$
$$\hat{\mathcal{X}} = \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J} \hat{\mathcal{Z}}_{i,l,j}\hat{\mathbf{U}}_i^{(1)} \circ \hat{\mathbf{U}}_l^{(2)} \circ \hat{\mathbf{U}}_j^{(3)}.$$

Сумма $\mathcal{X} = \sum_i\sum_l\sum_j \tilde{\mathcal{Z}}_{i,l,j}\tilde{\mathbf{U}}_i^{(1)} \circ \tilde{\mathbf{U}}_l^{(2)} \circ \tilde{\mathbf{U}}_j^{(3)} + \sum_i\sum_l\sum_j \hat{\mathcal{Z}}_{i,l,j}\hat{\mathbf{U}}_i^{(1)} \circ \hat{\mathbf{U}}_l^{(2)} \circ \hat{\mathbf{U}}_j^{(3)}$ является сингулярным разложением $\mathcal{X}$ в том и только том случае, когда пары векторов $\tilde{\mathbf{U}}_k^{(\sigma)}$, $\hat{\mathbf{U}}_m^{(\sigma)}$ взаимно ортогональны при всех возможных значениях $\sigma, k, m$. Это равносильно ортогональности линейных пространств $\mathcal{L}_1^{(\sigma)}$, $\mathcal{L}_2^{(\sigma)}$, построенных на векторах $\tilde{\mathbf{U}}_k^{(\sigma)}$ и $\hat{\mathbf{U}}_m^{(\sigma)}$ соответственно.

Рассмотрим пространства $\mathcal{L}_1^{(1)}$, $\mathcal{L}_2^{(1)}$: это пространства первых измерений тензоров $\tilde{\mathcal{X}}$ и $\hat{\mathcal{X}}$, то есть пространства построенные на векторах вида $\tilde{\mathcal{X}}_{,l,j}$ и $\hat{\mathcal{X}}_{,l,j}$ соответственно. Вспоминая вид тензоров $\tilde{\mathcal{X}}$ и $\hat{\mathcal{X}}$ получаем, что условие ортогональности этих линейных пространств равносильно первому условию из формулировки утверждения.

Оставшиеся два условия получаются аналогично из условий ортогональности оставшихся двух пар линейных пространств. □

Из утверждения 3.1 следует, что понятие слабой разделимости ряда из теории SSA [2] применимо и к тензорному случаю.

**Следствие 1.** *Если временные ряды $\tilde{F}$ и $\hat{F}$ длины $N$ слабо $I$- и $L$-разделимы в смысле теории* SSA*, то существует такое* HOSVD *траекторного тензора $\mathcal{X}$ ряда $F = \tilde{F} + \hat{F}$, что его можно разбить на две части, являющиеся* HOSVD *траекторных тензоров, составленных по рядам $\tilde{F}$ и $\hat{F}$.*

**Замечание 2.** *Понятие сильной разделимости можно перенести со стандартного случая на тензорный непосредственно, с поправкой на определение 2.1 сингулярных чисел для тензора.*

### 3.1. Примеры разделимости рядов в тензорном случае

Рассмотрим условия разделимости рядов $\tilde{F} = (\tilde{f}_1,\, \tilde{f}_2, \ldots,\, \tilde{f}_N)$, $\hat{F} = (\hat{f}_1,\, \hat{f}_2, \ldots,\, \hat{f}_N)$ в некоторых частных случаях.

- Отделимость от константного ряда

  Пусть $\tilde{f}_n = c \neq 0$ для $n \in \overline{1:N}$. Тогда необходимые и достаточные условия отделимости от него ряда $\hat{F}$ в смысле TSSA следующие:

  1. Ряд $\hat{F}$ имеет целый период $T$, и $I/T,\ L/T,\ J/T$ — целые;
  2. $\hat{f}_1 + \hat{f}_2 + \ldots + \hat{f}_T = 0$.

- Отделимость от экспоненциального ряда

  Пусть $\tilde{f}_n = e^{\alpha n}$ для $n \in \overline{1:N}$. Тогда необходимые и достаточные условия отделимости от него ряда $\hat{F}$ в смысле TSSA следующие:

  1. Ряд $(\tilde{f}_1\hat{f}_1,\, \tilde{f}_2\hat{f}_2, \ldots,\, \tilde{f}_N\hat{f}_N$ имеет целый период $T$, и $I/T,\ L/T,\ J/T$ — целые;
  2. $\tilde{f}_1\hat{f}_1 + \tilde{f}_2\hat{f}_2 + \ldots + \tilde{f}_N\hat{f}_T = 0$.

- Отделимость от гармонического ряда

  Пусть $\tilde{f}_n = \cos(2\pi\omega n + \varphi)$, где $0 < \omega < 1/2$, и $I, L, J > 2$. Положим $\hat{f}_n = \cos(2\pi\omega' n + \varphi')$, тогда ряд $\tilde{F}$ отделим от ряда $\hat{F}$ в смысле TSSA тогда и только тогда, когда $\omega \neq \omega'$ и $I\omega,\ I\omega',\ L\omega,\ L\omega',\ J\omega,\ J\omega'$ — целые числа.

## 4. Другие разложения

Помимо HOSVD, существует ещё одно разложение тензора в сумму тензоров 1 ранга: CANDECOMP-PARAFAC (CP) [3, 4]. Это разложение нам не подходит, так как количество компонент в этом разложении равно тензорному рангу, который в общем случае не удовлетворяет одному из основных свойств SSA: ранг траекторного тензора, построенного по сумме двух рядов может оказаться больше, чем сумма рангов траекторных тензоров, построенных на каждом из этих рядов.

Кроме того можно рассмотреть следующий случай. В терминах CP ряды $f_1^{(i)} = 3$, $f_2^{(i)} = \sin(2\pi i/3)$ имеют тензорные ранги 1 и 3 соответственно, а у ряда $f^{(i)} = f_1^{(i)} + f_2^{(i)}$ ранг равен 3. Притом ни один из элементов разложения не даёт константный ряд, то есть разделения константного ряда от синуса не произошло, хотя в теории SSA эти ряды являются сильно разделимыми при условии делимости всех измерений на период синуса.

```
s <- rep(3, 132)
p <- tssa3(s, rank = 1, I = 42, L = 45)
```

```
print(p$cp$conv)
```

```
## [1] TRUE
```

```
rec <- t3.reconstruct(p, list(1))
mse(rec[[1]], s)
```

```
## [1] 5.438359e-30
```

```
s <- sin(2 * pi * 0:132 / 3)
p <- tssa3(s, 2, 45, 45)
```

```
print(p$cp$conv)
```

```
## [1] FALSE
```

```
rec <- t3.reconstruct(p, list(1:2))
mse(rec[[1]], s)
```

```
## [1] 0.03323903
```

```
s <- sin(2 * pi * 0:132 / 3)
p <- tssa3(s, 3, 45, 45)
```

```
print(p$cp$conv)
```

```
## [1] TRUE
```

```r
rec <- t3.reconstruct(p, list(1:3))
mse(rec[[1]], s)
```

```
## [1] 1.446833e-16
```

```r
s.const <- 3
s.sin <- sin(2 * pi * 0:132 / 3)
s <- s.const + s.sin
p <- tssa3(s, 3, 45, 45)
```

```r
print(p$cp$conv)
```

```
## [1] TRUE
```

```r
print(p$modes)
```

```
## $I
## [1] 45
## $L
## [1] 45
## $J
## [1] 45
```

```r
rec <- t3.reconstruct(p, list(1, 2, 3))
mse(rec[[1]], s.const)
```

```
## [1] 9.043692
```

```r
mse(rec[[2]], s.const)
```

```
## [1] 1.32778
```

```
mse(rec[[3]], s.const)
```

```
## [1] 3.431616
```

Другим недостатком CP разложения является то, что это итерационный метод, причём процесс итерации начинается с генерации случайной матрицы, в связи с чем на одних и тех же данных он может выдавать разные результаты, в том числе может как сойтись, так и нет. Возможно можно добиться лучших результатов, используя CP, если строить тензор по ряду другим образом, подбирать другие параметры разложения и использовать другие методы вычисления CP, однако изучение этого вопроса выходит за рамки этой работы.

## Заключение

Здесь должно быть заключение.

# Список литературы

1. De Lathauwer Lieven, De Moor Bart, Vandewalle Joos. A Multilinear Singular Value Decomposition // SIAM Journal on Matrix Analysis and Applications. — 2000. — Vol. 21, no. 4. — P. 1253–1278. — Access mode: https://doi.org/10.1137/S0895479896305696.
2. Golyandina Nina, Nekrutkin Vladimir, Zhigljavsky Anatoly. Analysis of time series structure: SSA and related techiques. — Chapman & Hall/CRC, 2001.
3. Harshman Richard A. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-model factor analysis. — 1970. — Vol. 16. — P. 1–84.
4. Carroll J. Douglas, Chang Jih Jie. Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition // Psychometrika. — 1970. — Vol. 35. — P. 283–319.