

Пространственно-временная реконструкция временных рядов на основе тензорного разложения

Горишний Максим Надежда Алсаханова

МФТИ

4 мая 2023 г.

Цель

Сравнить Singular Spectrum Analysis (SSA) и его тензорную версию (TSSA) в пространственно-временном разложении и реконструкции временных рядов при сравнимой вычислительной затратности

Идея

TSSA за счет использования пространственной информации должен лучше справляться с задачей.

$$L(\hat{x}) = MSE(x, \hat{x})$$

x — исходный временной ряд длины n .

l — длина окна для стадии эмбединга SSA.

m — длина окна для стадии эмбединга TSSA.

\hat{x} — реконструированный временной ряд.

$$X = [X_1, X_2, \dots, X_l] = \mathcal{H}X = \begin{pmatrix} x_1 & x_2 & \dots & x_l \\ x_2 & x_3 & \dots & x_{l+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{l_x} & x_{l_x+1} & \dots & x_{l_x+l-1} \end{pmatrix},$$
$$l_x = n - l + 1;$$

$$X = \sum_{i=1}^d X_i = \sum_{i=1}^d \sqrt{\lambda_i} u_i v_i^T;$$

$$I = \{l_1, l_2, \dots, l_q\}, \bigcup_{i=1}^q I = \{1, 2, \dots, d\}$$

$$\hat{X} = \sum_{i=1}^q \hat{X}_i = \sum_{i=1}^q \sum_{j \in I_i} X_j;$$

$$X = [X_1, X_2, \dots, X_m] = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ x_{m+1} & x_{m+2} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k_x+1} & x_{k_x+2} & \dots & x_{k_x+m} \end{pmatrix},$$
$$k_x = \lfloor \frac{n}{m} \rfloor$$

$$\hat{X} = \text{PARAFAC}(X)|_{\text{first}_k}$$



[Nina Golyandina \(2020\)](#)

Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing



[Sanei Kouchaki \(2013\)](#)

Tensor based singular spectrum analysis for nonstationary source separation



[DerkJan Dijk Samaneh Kouchaki, Saeid Sane \(2015\)](#)

Tensor based singular spectrum analysis for automatic scoring of sleep eeg



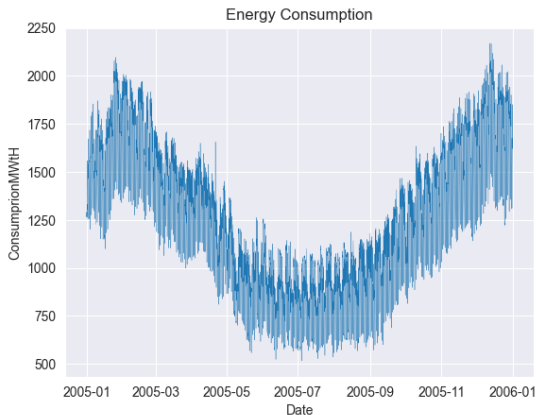
[V.V. Strijov K.R. Usmanova. \(2018\)](#)

Models of detection relationship between time series in forecasting problems

Синтетические квази-периодические ряды длины $n = 1000$

Продажи топлива в день на протяжении года $n = 365$

Потребление электричества в час на протяжении года $n = 8760$

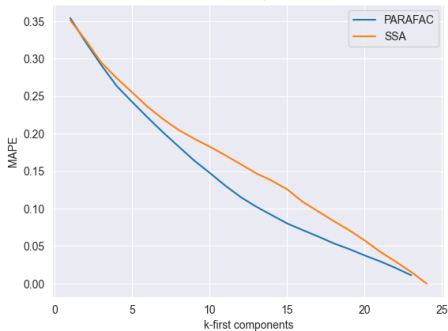


Для SSA — перебор размера окна l и выбор первых компонент k по сетке с шагом 1.

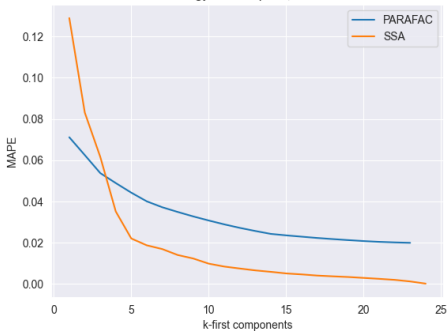
Для TSSA — перебор размера окна m , ганкелизация с размером l и выбор первых компонент k по сетке с шагом 1.

Результаты

Petrol sales, $l = 24$



Energy consumption, $l = 24$



На малых k TSSA показывает лучшие результаты, чем SSA. На больших k соотношение их точности зависит от данных и гиперпараметров.

- SSA практически не требует подбора гиперпараметров
- TSSA более устойчив и точен, но требует более тонкого подбора гиперпараметров
- В большинстве случаев TSSA работает медленнее SSA
- Алгоритмы не заменяют, а дополняют друг друга