Spatiotemporal forecasting with convolutions and tensor decomposition

A Preprint

David S. Hippocampus*
Department of Computer Science
Cranberry-Lemon University
Pittsburgh, PA 15213
hippo@cs.cranberry-lemon.edu

Elias D. Striatum
Department of Electrical Engineering
Mount-Sheikh University
Santa Narimana, Levand
stariate@ee.mount-sheikh.edu

Abstract

Tensor based time series decomposition methods based on singular spectrum analysis showed great results in both denoising and interpretability. Several forecasting techniques based on them were already explored, yet none provided simultaneously accurate, stable and computationally cheap inferring. After an in-depth study of well known models we facilitated a new one comprising all three requirements for non-stationary quasi-periodic time series. The model was then tested on real-life data of electricity consumption and other well-explored datasets.

Keywords First keyword · Second keyword · More

1 Introduction

Singular spectrum analysis (SSA) is a method widely used in the past decades in different areas, from economics to biology and social science [1]. One of main advantages is its ability to extract underlying frequencies from complex and multidimensional data, resulting in variable number of components. SSA consists of two main stages: decomposition and reconstruction, both adjustable in terms of methods and hyperparameters used.

Its tensor modification (TSSA) offers [2] a more robust and accurate results by converting a series into a tensor and using parallel factor analysis (PARAFAC) decomposition instead of the usual SVD. It optimizes usage of information initially available to a model in cost of working with more multidimensional data SSA does. The problem of reconstruction is left untouched however.

Empirical mode decomposition (EMD) can then be used to supervise TSSA, giving us TSSA-EMD[3]. It provides a better way to identify the number of frequency components within each subspace. With such enhancement algorithms achieves [4] a distinguishable growth in accuracy of signal reconstruction with denoising, leaving other methods far behind in particular tasks.

The problem of forecasting time series has not yet being mentioned. Basic SSA shows [5] adequate results when working with series of constant-limited variation function. However it becomes highly unstable in two basic cases. With several outliers already false frequencies are extracted at decomposition stage, what does not lead to reconstruction defects but enforces unacceptable error even at earlier points of prediction. A variation growth affects SSA the same way, usually creating frequencies of much higher amplitude than are expected. That makes the early predictions seem accurate, yet giving unrealistic forecast long-wise.

To conclude, SSA is a powerful but limited method.

*Use footnote for providing further information about author (webpage, alternative address)—not for acknowledging funding agencies.

At the same time SSA does not utilize spatial information. Given a set of parallel time series it is meant to decompose each separately. TSSA instead can show better performance by working with them as with a whole dataset. Knowing this we experiment to determine the difference between SSA and TSSA forecasting and try to create a robust model of signal decomposition, reconstruction and prediction.

2 Problem

2.1 Time series

Let $x = [x_1, x_2, ..., x_n]$, $x_i \in \mathbb{R}$ be a 1D time series, namely a vector. We suppose x has no trend, is quasi-periodic and it's phase trajectory is stationary.

[do we need to describe phase trajectory here?]

2.2 Hankelization

One of the SSA steps requires Hankelization operator $\mathcal{H}$ to be described. With 2D matrix $i \times j$ it works as follows:

$$\mathcal{H}M = \begin{pmatrix} \widetilde{m}_1 & \widetilde{m}_2 & \dots & m_i \\ \widetilde{m}_2 & \widetilde{m}_3 & \dots & m_{i+1} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{m}_j & \widetilde{m}_{j+1} & \dots & \widetilde{m}_{i+j-1} \end{pmatrix}$$
$$\widetilde{m}_k = \sum_{i,j \in D_k} m_{i,j}/numD_k \tag{1}$$
$$D_k = \{(\widehat{i}, \widehat{j}) : 1 \leq \widehat{i} \leq i, 1 \leq \widehat{j} \leq j, \widehat{i} + \widehat{j} = k + 1\}$$

For a higher dimensional tensor $\mathcal{H}$ performs first index wise.

2.3 SSA

The basic SSA consists of two stages. The first is decomposition, it maps time series to a number of components. The second is reconstruction, it combines components to frequencies of interest.

Decomposition: given a 1D vector x of size n and window size l it first creates a 2D embedding of it, a trajectory matrix X of size $l_x \times l$:

$$X = [X_1, X_2, ..., X_l] = \begin{pmatrix} x_1 & x_2 & \dots & x_l \\ x_2 & x_3 & \dots & x_{l+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{l_x} & x_{l_x+1} & \dots & x_{l_x+l-1} \end{pmatrix}$$
$$l_x = n - l + 1 \tag{2}$$

The actual value of l is chosen arbitrarily to include enough information about the signal variance.

Singular value decomposition (SVD) is then applied to X to obtain its singular values and vectors:

$$X = \sum_{i=1}^{d} X_i = \sum_{i=1}^{d} \sqrt{\lambda_i} u_i v_i^T \tag{3}$$

Where λ_i is the ith eigenvalue, u_i is the corresponding left singular vector and v_i is the ith right singular vector; d is the total number of eigenvalues.

Reconstruction: given a set of components X_i it first splits them into disjoint set and sums independently:

$$I = \{I_1, I_2, \dots, I_q\}, \bigsqcup_{i=1}^{q} I = \{1, 2, \dots, d\}$$
$$X = \sum_{i=1}^{q} \hat{X}_i = \sum_{i=1}^{q} \sum_{j \in I_i} X_j \tag{4}$$

The choice of I is one of the most important aspects of SSA in terms of retrieving the desired frequencies.

Hankelization is then applied to X and a time series is reconstructed from it (described later).

2.4 TSSA

TSSA is a generalization of SSA to work with a set of parallel time series. It is based on the idea that the time series are not independent and that the information contained in one time series can be used to improve the reconstruction of another.

2.5 Forecasting models based on SSA and TSSA

Copypaste from [Time series SSA forecasting] tutrial link from LinkReview could be found here.

No idea on how to predict with TSSA - we are to explore this I suppose

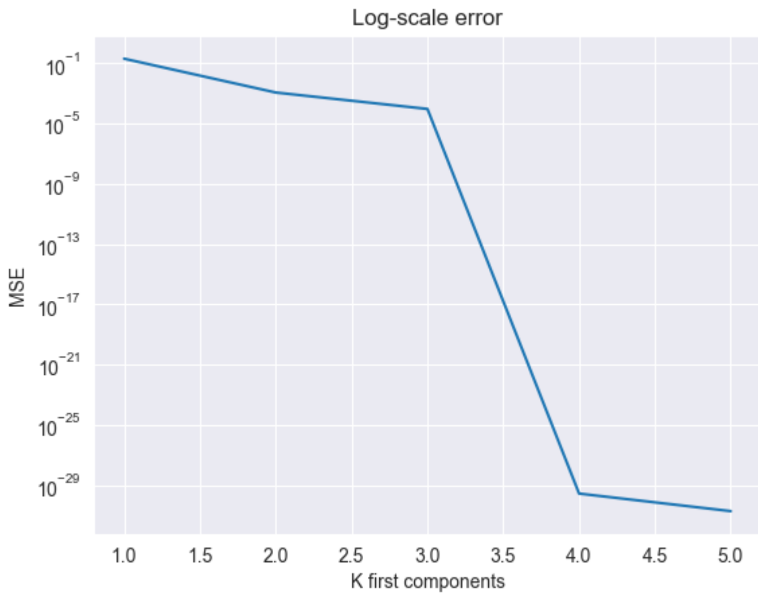
3 Experiment

Preliminary experimets are currently described here. The section will be updated when the real data will be utilized.

We use $f(t) = \sin(\pi t) + np.sin(et)$ for $t \in (0, 100)$, sampling it to 100 points to create a time series. We then applied SSA with different numbers of components taken at reconstruction step. The data generated looks as follows:

The actual results for log-MSE error for SSA with window size 5 are:

Рис. 1: Sample figure caption.



With first 4 components taken at reconstruction step the error is non-significant already. More complex data must be taken to see the difference when applying PARAFAC.

4 Trash

4.1 Citations

Citations use `natbib`. The documentation may be found at

`http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf`

Here is an example usage of the two main commands (`citet` and `citep`): Some people thought a thing [Kour and Saabne, 2014a, Hadash et al., 2018] but other people thought something else [Kour and Saabne, 2014b]. Many people have speculated that if we knew exactly why Kour and Saabne [2014b] thought this...

4.2 Figures

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi. See Figure 1. Here is how you add footnotes. [2] Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

4.3 Tables

See awesome Table 1.

The documentation for `booktabs` ('Publication quality tables in LaTeX') is available from:

[2] Sample of the first footnote.

Таблица 1: Sample table title

Part		
Name	Description	Size (μm)
Dendrite	Input terminal	$\sim$100
Axon	Output terminal	$\sim$10
Soma	Cell body	up to 10^6

`https://www.ctan.org/pkg/booktabs`

4.4 Lists

- Lorem ipsum dolor sit amet
- consectetur adipiscing elit.
- Aliquam dignissim blandit est, in dictum tortor gravida eget. In ac rutrum magna.

Список литературы

George Kour and Raid Saabne. Real-time segmentation of on-line handwritten arabic script. In Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on, pages 417–422. IEEE, 2014a.

Guy Hadash, Einat Kermany, Boaz Carmeli, Ofer Lavi, George Kour, and Alon Jacovi. Estimate and replace: A novel approach to integrating deep neural networks with existing applications. arXiv preprint arXiv:1804.09028, 2018.

George Kour and Raid Saabne. Fast classification of handwritten on-line arabic characters. In Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of, pages 312–318. IEEE, 2014b. doi:10.1109/SOCPAR.2014.7008025.