Comparison of low rank tensorial approximation to discrete wavelet approximation of given nonstationary time series

Davit DATUASHVILI*
Cihan MERT**

Abstract

Approximation and filtration of time series belongs to one of most important problems in real word scientific application. Despite of plenty of existed methods, they are effective in very specific situations, most of them assume that time series is stationary or can be stationary by finite amount of differencing. In this article we will consider several time series and compare Singular spectrum analysis and its low rank tensorial approximation method to classical wavelet decomposition approach.

Keywords: DWT (discrete wavelet transform), SSA (singular spectrum analysis), Low Rank Tensorial approximation, Hankel data matrix, SNR (signal to noise ratio)

Introduction

Filtering and approximation of time series represents the classical problem in signal processing, Economic fields or in any modern scientific fields. Existence of noise and nonstationarity property of time series makes time series analysis a complex topic and sometimes effectiveness of any method used and its reliability represents subjective decision made by the researcher.

Literature Review

Analysis of the nonstationary time series requires specific methods because of the complexity of randomness and mathematical complex structure of time series. Two classical approaches for the analysis of nonstationary time series are Short Time Fourier Transform and Wavelet transform. Short Time Fourier Transform uses window function or in other words,it represents windowed Fourier transform and width of window is fixed during the analysis. The idea is that hopefully widnow width would be enough to catch stationary part of the series. On the other hand, wavelet transform uses special function called mother wavelet and it is scaled and delayed version functions to adjust the transformation according to the dynamic structure of series. Besides these two approaches, the newest methodology for time series analysis represents singular spectrum method and it is considered as a new tool in Time Series Analysis (Elsner, 1996). SSA constructs Hankel type of data matrix from the given time series. Using Singular value decompose, it computes left and right singular vectors. The idea of singular spectrum analysis is that it contains temporally similar structure of original time series (Milnikov, 2013; 2014) for time series that contain periodic deterministic components, left and right singular vectors have the same pseudospectral structure as original time series (Datuashvili. Mert & Milnikov, 2014) and then chooses appropiate singular values and singular vectors using algorithm of low rank tensorial approximation (Milnikov & Prangishvili, 2013). We can reconstruct the original time series with good approximation with significantly increased signal to the noise ratio.

Problem Statement

In this article we will consider the signal which consists of two parts: pure signal part and noise

$$s(t) = x(t) + noise \tag{1}$$

Here $x(t)$ is a pure part of the time series, while $s(t)$ is the result of the combination of noise and pure signal. By noise we mean uniformly distributed white noise. The goal is to remove noise and therefore, increase a signal to the noise ratio to minimize the distortion of the original signal. As mentioned above, we will compare wavelet and SSA method for denotation and approximation purpose.

*PDc. Study process Administrator, Faculty of Computer Technology and Engineering, International Black Sea University, Tbilisi, Georgia. E-mail:ddatuashvili@ibsu.edu.ge

** Assoc.Prof.Dr .Dean. , Faculty of Computer Technology and Engineering, International Black Sea University, Tbilisi, Georgia. E-mail:cmert@ibsu.edu.ge

Methodology

Singular Spectrum Analysis

A procedure for singular spectrum analysis is to construct the following matrix from the given time series

$$X_d = \begin{vmatrix} x[1] & x[2] & \dots & x[p] \\ x[2] & x[3] & \dots & x[p+1] \\ \dots & \dots & \dots & \dots \\ x[N-p] & x[N-p+1] & \dots & x[N] \end{vmatrix}$$

Where N samples of the time series x[1], x[2],…,x[N] are given. It should be noted that the main task of SSA (singular spectrum analysis) is to choose window length correctly (Elsner, 1996) meaning that we should choose the window length that is enough to hold all principal components and at the same time its singular vectors should be long enough to separate the noise and the signal from each other. After applying SVD decomposition, which sorts singular values and their corresponding singular vectors, leading singular vectors contribute to the main variance of the given signal.

The main benefit of SSA is that it can help us to increase the signal to noise ratio more than it was at the beginning.

Finally, a reconstruction of the original matrix using low rank tensorial approximation can be done by algorithm given in the article by Milnikov, & Prangishvili (2013).

Wavelet analysis and its discrete version

Discrete wavelet transform represents the special case of wavelet transform method for the given function $f(t)$

$$\psi(t) = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)$$

(2)

Where for discrete wavelet transform coefficients are the following ($(a^m, na^m b)$ for all m, n $\in Z$ (D.-X., 2002), summary for scaling parameter a is given below:

• Low scale *a* ➞ Compressed wavelet ➞ rapidly changing details ➞ High frequency ω.

• High scale *a* ➞ Stretched wavelet ➞ slowly changing, coarse features ➞ Low frequency ω.

But it should be noted that there does not exist a precise relationship between scale and frequency, for different choice of mother wavelet, there is different relationship.

Approximation is done by selecting significant coefficients after wavelet transformation for discrete wavelet transform which represents a series of filter banks or a series of low pass and high pass filter. The approximation is done by selecting the level of approximation usually at most 5 level of the approximation and decomposition is considered as enough. Discrete wavelet decomposition in matlab can be easily computed by function wavedec. The summary of wavedec can be seen in matlab using help wavedec in matlab software. Also, for wavelet transformation, it highly depends on the scientist's experience and a subjective decision which mother basis will be chosen, because there exists a lot of wavelet functions and all those are characterized by different properties and therefore, have different applications. It should also be noted that wavelet transformation is sensitive to noise.

Experimental results

Analysis of electrical signal

In the first experiment we will consider time series consisting of electricity usage data which in matlab can be easily accessed by command

```
load leleccum;
```

For analysis purpose, let us consider first 3920 samples, fully code for this in matlab will be

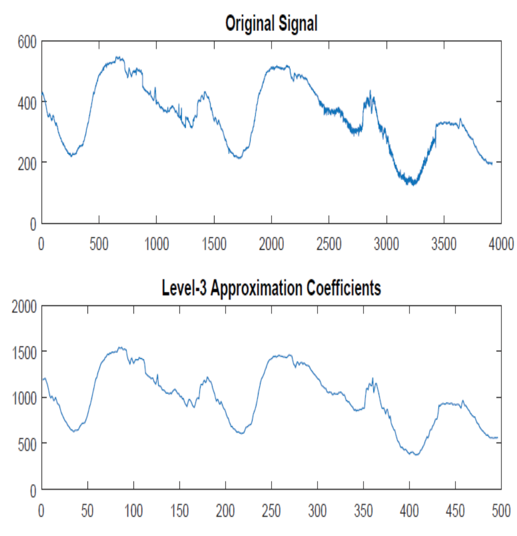
```
load leleccum;
sig = leleccum(1:3920);
```

Let us obtain discrete wavelet decomposition using level 5 with the 'sym4' wavelet.

```
[C, L] = wavedec (sig, 5,'sym4');
```

And finally the picture of original and approximated signal can be easily obtained by following set of commands

```
Lev = 3;
a3 = appcoef(C,L,'sym4',Lev);
subplot(2,1,1)
plot(sig); title('Original Signal');
subplot(2,1,2)
plot(a3);
 title('Level-3 Approximation Coefficients');
```



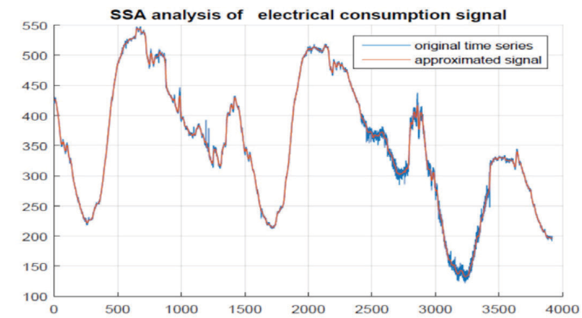
Discrete wavelet transform

One of the big drawbacks of this method is that while approximation seems ok, the length of reconstructed series is

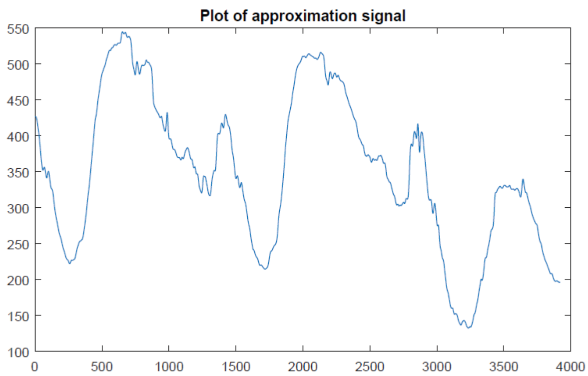
Journal of Technical Science and Technologies; ISSN 2298-0032; Volume 4, Issue 1, 2015

equal to 496, while the length of original time series is 3920. Note that increasing or decreasing level of approximation decreases the length of coefficients even more or increases the noise level in the original signal, for instance, for level 4, the length of new approximation coefficients was 251.

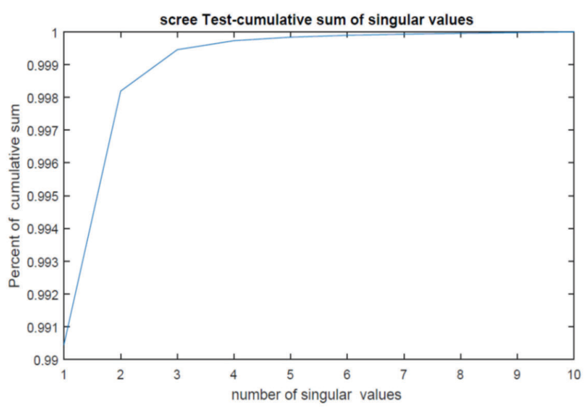
Now let us consider the analysis of the same time series using singular spectrum analysis. Window length was chosen 10 and for low rank tensorial approximation first matrix of tensorial product was used.



Here is also a separate plot of approximated signal. From the picture we can see that the noise level is significantly removed while keeping the original structure.



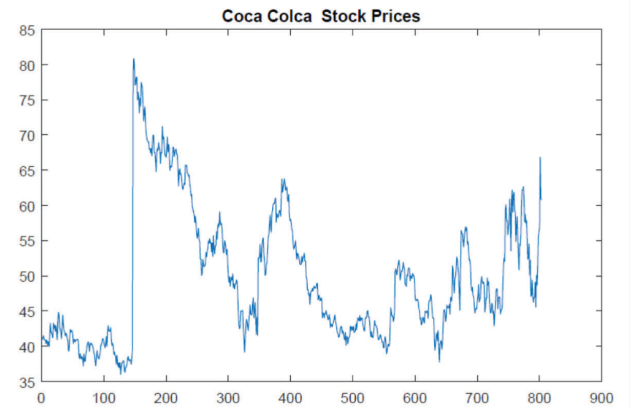
While it repeats the main features and structure of the original signal, the length of the approximated sample is again 3920. Here is also a plot of singular values cumulative plots or scree test.



As we see the first singular value contains 99% of total variation. To reconstruct the original signal, we should include all singular values during the low rank tensorial approximation algorithms.

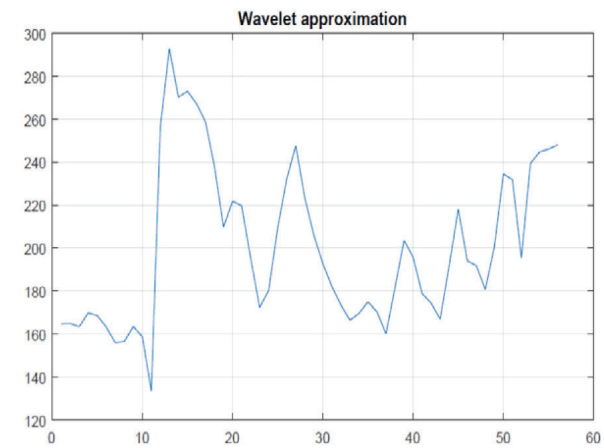
Analysis of weekly taken Coca-Cola prices

In the second experiment we will consider Historical stock prices of Coca-Cola Company from 2/01/2000 to 06/05/2015. Data was taken weekly, total sample size equals 803. Original signals plot and its discrete wavelet analysis result is given below:



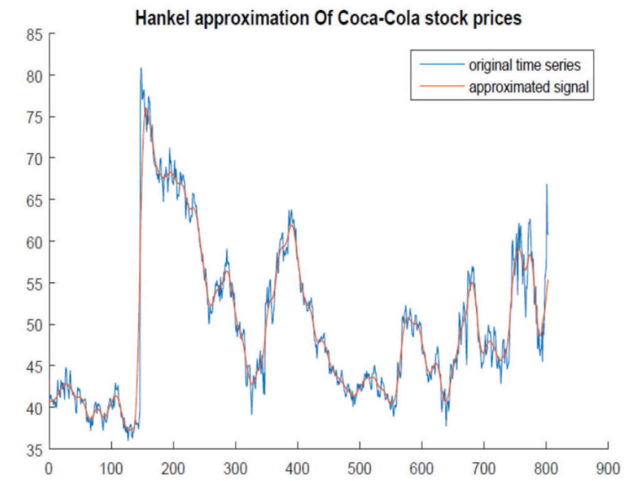
Original time series

Discrete wavelet transform coefficients plot is given below:



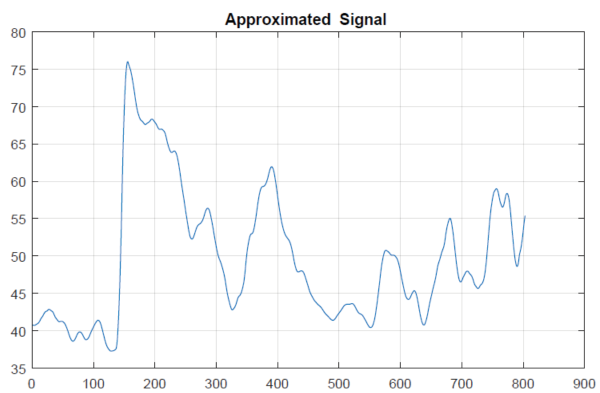
Wavelet approximation of Coca-Cola

Again in our case the length of coefficients is 56, more than 10 times short compared to the original series. We will apply the low rank tensorial approximation to this series.



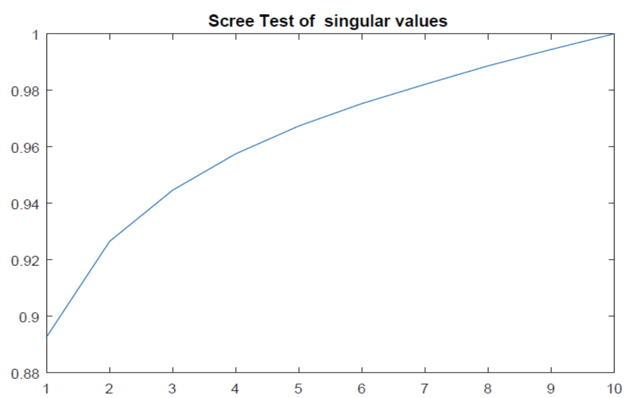
Hankel approximation

We can also see the plot of approximation result separately. From the picture we can see easily that the noise level is significantly reduced. We kept samples size equal to original and also we obtained the most significant parts from the noise signal. As it was mentioned above, if we want to reconstruct original time series, we should use all singular values during the approximation procedure. Generally the idea of low rank tensorial approximation is not fully reconstruction of time series, but partially reconstructed time series with the properties that are mostly similar to the original one and at the same time the noise level in returned signal is significantly reduced.



Approximated signal

Finally, we will show the scree test of singular values, from which total 10 singular values contain total 100% of variance of the original time series, therefore for fully reconstruction use all these singular values will work fine.



Conclusion and Recommendation

Comparison showed that while wavelet transform worked fine for given time series models, its main drawbacks are that they are sensitive to the noise and also choosing correct wavelet basis highly depends on researcher's objectives and practical experience.

Resulted approximated coefficients are too small compared to the original series, which may have a great impact for further analysis of signal, starting from its forecasting continued to spectral analysis, because spectral analysis of time series highly depends on its total observation duration.

Different from Wavelet transform, SSA analysis showed that it not only keeps original time series structure, but also significantly increases the signal to the noise ratio, which increases statistical stability of the further analysis and forecasting reliability.

References

Milnikov, A. & Prangishvili, A. (2013). A Low Rank Tensorial Approximations method of computation of Singular Values and Singular Vectors for SVD problem . *Journal of Mathematical Sci*. (pp. 512-517). New-Yotk: Springer.
.
Dautashvili, D.; Mert, C. & Milnikov, A. (2014). A New Approach to Detecting Deterministic Periodic Components in Noise. *5th International Conference on Circuits, Systems, Control, Signals* . Salerno.

D.-X., Z. (2002). Wavelet analysis: twenty years' developments. *Wavelet analysis: twenty years' developments: proceedings of the International Conference of Computational Harmonic Analysis* (p. 318). Hong Kong: World Scientific Publishing Co.Pte.Ltd.

Elsner, J.B. (1996). *Singular Spectrum Analysis*. Milwaukee: Springer Science.

Milnikov, A. (2013- 2014). Pseudospectral Structure of the Singular Vectors of Nonstationary Time. *Latest Trends in Circuits, Systems, Signal Processing and Automatic Control* , 2-3.