

Improving Spectral Resolution of a Stationary Signal Using Singular Value Decomposition

Davit DATUASHVILI *

Abstract

A new approach to detecting Periodic Components with random phases of stationary time series corrupted by random component is considered. The approach is based on usage of the new time series, which is the union of the principal singular vectors of the data matrix of the original time series. The approach allows significant improvement of resolving capacity for all existed method of pseudospectra estimation. The comparative examples, illustrated practical efficiency of the approach are considered.

Keywords: singular value decomposition, spectral resolution, SNR (signal to noise ratio), PSD (power spectral density), stationary signal

Introduction

Spectral resolution is one of the biggest problems in digital signal processing and signal analyzing, while in Business sciences, forecasting and modelling of given time series is very crucial among a large variety of random processes (having trend, stationary, ergodic, purely random, etc.). Processes containing periodical and random components are also very important (Milnikov, 2013). Correct estimation of parameters, as well as increasing the ability of resolving capacity or ability to separate components, which are close to each other in amplitudes and frequencies, are very important to model data correctly, and finally to get accurate result for analysis and forecasting.

There exist several approaches for spectral estimation of stationary signal in modern literature, including parametric (Auto Regressive (AR), Autoregressive Moving Average (ARMA) etc) and non-parametric models. While these methods work well in many cases, they all have basic limitations. Parametric method requires some prior knowledge about time series, for instance - selection of correct order of AR/ARMA model, at the same time these methods are sensitive to noise and in case of presence of noise, it can give us an incorrect result. Another variant of AR model is a Subspace method, which requires number of parameters to be known in advance (Hayes, 1996). On the other hand, non-parametric method lacks the possible information about signal, and also it is not statistically stable. For instance, periodogram has variance problem (Hayes, 1996). There exist several approaches to reduce variance in periodogram method, but it will also decrease its spectral resolution ability.

Objectives of Research

In this article we are considering the following model for analyzing given signal represented by some finite number of periodic components with additive white noise

$$x(t) = \sum_{j=1}^m A_j e^{(2\pi f_j t + \theta_j)i} + w(t) \quad (1)$$

where A_j –magnitude of j -th periodic component;
 f_j –frequency of j -th periodic component;
 θ_j – uniformly distributed random phase of j -th periodic component;
 $w(t)$ - random component (white noise in our case);
 m - is number of periodic components.

Due to the phases are uniformly distributed random variables, thus this model represents stationary signal (Marple, 1987). We also assume that the ratio of frequencies can be any number: so, the time series (1) is not harmonic.

The main problem of spectral estimation method is to effectively separate components from each other and to restore good spectral pictures. Various parametric and non-parametric methods fail in this task. Recently it has been proved that spectral resolution and resolving capacity can be improved by concatenating principal singular vectors of Hankel data matrix after its singular value decomposition. The latter allows artificial increasing of total observation time, which will help us to identify components that are very close to each other (Milnikov, 2013). It should be noted that for given series with length N and sampling period t_s (where f_s is sampling frequency, total observation time is given by ,

*Department of Computer Technologies and Engineering, International Black Sea University, Tbilisi, Georgia.
 E-mail: ddatuashvili@ibsu.edu.ge

and spectral resolution limit is $\frac{1}{N * T_s}$. Concatenating principal singular vectors allows introducing a time parameter $T_o^n = mNt_s$, which is not a physical time of observation, but is equal to m times, the physical observation time. It was called **the time of singular unfolding** of the initial time series. Obviously, value of the spectral resolution of the time series, consisting of the union of the principal singular vectors, increases m times. We emphasize that this result is independent of the spectral estimation method used. Another advantage of this approach consists in increasing of the statistical stability of spectral estimates (reduction ratio noise / signal) that is achieved by filtering properties of the low-rank tensor approximation: in all singular vectors ratio noise / signal ratio is less than in the original series (Milnikov, 2013)

In the recent article we tested this approach on stationary signal and also compared several existed methods to see that improving spectral estimation by concatenation of singular vectors will overcome those disadvantages that occur during usage of modern spectral estimation methods.

Methodology

Let us consider the following model:

$$x(t) = \sum_{j=1}^m A_j e^{(2\pi f_j t + \theta_k)i} + w(t) \quad (2)$$

For $x(t)$ to be stationary, it is necessary that phases be uniformly distributed random variables, let us take the following frequencies $f_1=15$, $f_2=15.3$, $f_3=15.7$ and $f_4=16$ and amplitudes be $A_1=15$, $A_2=16$, $A_3=16.4$, and $A_4=15.5$. Sampling frequency is $f_s=100$, which means that sampling

period is equal $t_s \frac{1}{f_s} = 0.01$, sample length is equal 294 and

therefore, total observation time is equal $294 * 0.01 = 2.94$

resolution limit will be $\frac{1}{2.94} = 0.34013605$, power spectrum

of original signal using periodogram estimator is

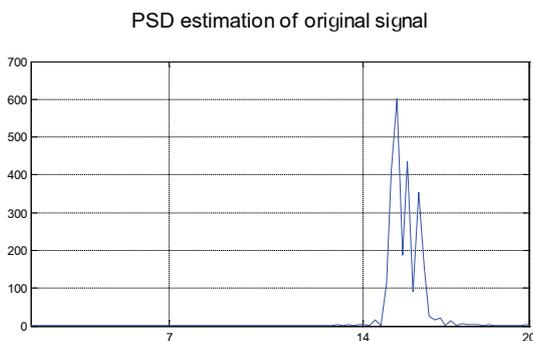


Figure 1. Spectral structure of original signal

From Figure 1, we see that given sample size is not enough to detect hidden periodicity. We can see just three peaks, but the fourth one is lost below. We will consider several approaches.

Discrete Fourier Transform method

We tested Discrete Fourier Transform (DFT) to estimate spectrum of the time series under consideration.

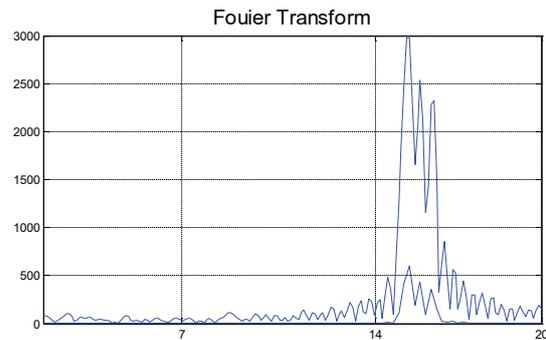


Figure 2. PSD estimation by Fourier method

The fig.2 shows that the DFT failed to detect 4 frequencies, and due to the noise false peaks have been appear peaks. Also due to small sample size, spectral window also effects on the resolving capacity, different technical difficulties are characters of Fourier method. First of all, it should be noted that frequencies do not satisfy Fourier harmonics criteria, also nonstationarity of signal makes Fourier analysis less useful.

Eigenanalysis based approach

Eigenanalysis method is based on eigenvalue decomposition of autocorrelation matrix of given data sample, and using its eigenvalue and eigenvectors. Eigen analysis method separates signal space from noise space and then estimate power spectrum from those spaced: those approaches are named as signal subspace and noise subspace methods. Two methods are considered – multiple signal classification method (MUSIC) and eigenvector method. PSD estimation using MUSIC will have the following form.

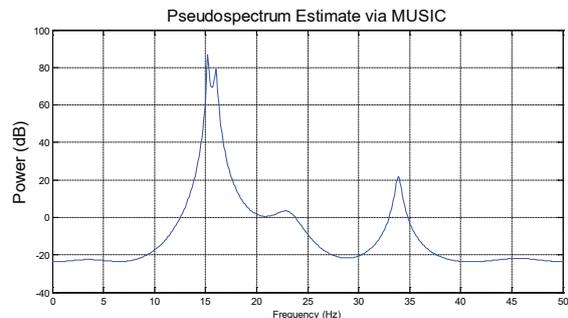


Figure 3. PSD estimation using music method

MUSIC method is characterized by the occurrence of spurious peaks. From the picture, it is easily checked that false peak occurred at frequencies near the 35 Hz that is because of noise as MUSIC method is sensitive to noise.

Similar is eigenvector method. Estimate of PSD using eigenvector method will be following:

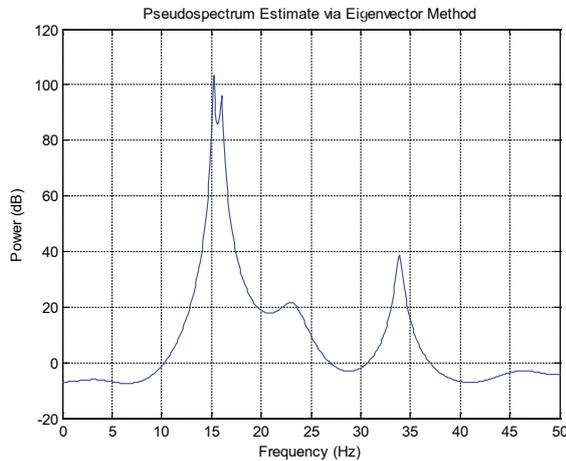


Figure 4. PSD estimation using Eigenvector method

The method again fails to detect close frequencies and again shows false peak near the 35 Hz. It should be noted that both methods require prior estimation of number of periodic components. It is clear that the latter constricts their abilities to solve problems connected with detecting of periodic components in corrupted signals.

Autoregressive approach for PSD estimation

Autoregressive model assume output of process linearly depends on its own previous values, it can be represented as

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t \tag{3}$$

where c is constant, ε_t is white noise and $\varphi_1, \dots, \varphi_p$ are parameters, PSD estimation for given signal with noise variance $Var(Z_t) = \sigma_Z^2$ will be

$$S(f) = \frac{\sigma_Z^2}{|1 - \sum_{k=1}^p \varphi_k e^{-2\pi i k f}|^2}$$

Despite the fact that AR spectral estimation is considered as high resolution methods, one of the basic problems of AR model is determining the order of equation (3), as choosing the correct order leads to incorrect estimation of PSD. Generally order is defined between from $N/3$ to $N/2$, where N is sample length. Matlab has several built in function for estimation of power spectrum using AR model, one of them is pyulear, which estimates power spectrum using Yule-Walker equations. It derives an all-pole model to represent the spectrum, so the correct choice of the model order

is crucial. Spectrum using pyulear method will have the following form:

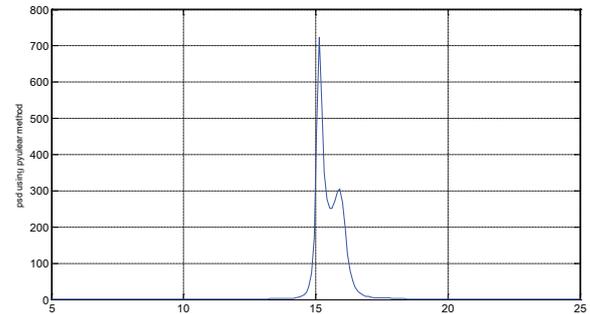


Figure 5. PSD estimation using pyulear method

From Fig.5 it is easy to see that PSD estimation using pyulear method is insufficient: it could not identify two peaks, besides the second peak has not been well separated from the first one. We also tested another approach - Burg method. This method fits an autoregressive (AR) model to the signal by minimizing the forward and backward prediction errors.

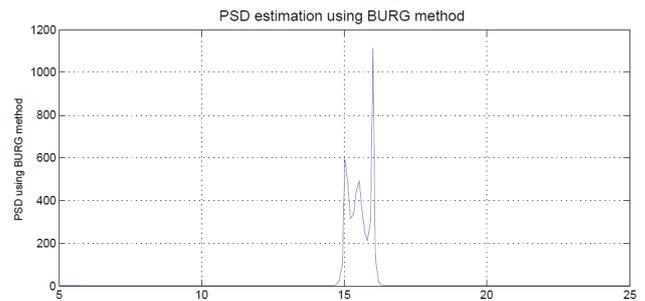


Figure 6. PSD estimation using BURG method

Despite the fact that BURG method showed better result than pyulear, it still cannot identify forth frequency.

During the analysis of all these procedures, we can see that sample size has very great effect on correct estimation of Power spectrum density, small data length basically causes incorrect and incomplete estimation and finally spectral resolution ability of above mentioned methods will be very low.

SVD approach for Hankel type of matrix

Let us consider following Data matrix

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

where $x[1], x[2], \dots, x[N]$ are samples of the time series and N – number of samples. SVD decomposition of the Data matrix allows obtaining both left and right singular vectors of matrix X_d . It is significant to note that sizes of the matrix were taken as 275×20 . After concatenation of the first and the

third left and right singular vectors value of time of singular unfolding will be equal to $T_s = 5.5$ sec ($T_s = 275 \cdot 2 \cdot 0.01$). The latter that means that the resolution limit now will be $f_r = 0.18$. It has decisive meaning for improving resolution capacity. In the Fig. 7 the result of PSD of new time series obtained by means of concatenation of the first and the third left and right singular vectors are shown.

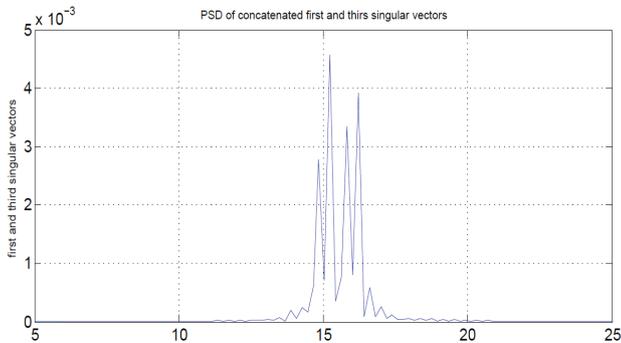


Figure 7. Concatenation of first and third singular vectors and their PSD estimation

Here we introduce a criterion for estimating the reliable separation of the two peaks of a pseudospectrum. It is assumed that the two peaks corresponding to two periodic deterministic components are separated, if a “dip” notch between them not less than 3 dB. Therefore as a criterion of separation of two peaks one should use the following coefficient of separability

$$S(f) = \frac{P(f_j)}{P(f_i)} \geq 10^{0.3} \quad (i \neq j)$$

where $P(f_j)$ - the value of a smaller peak (between two comparable ones) at the frequency f_j ;

$P(f_j)$ - value of notch at frequency f_i .

Usage of the coefficient of separability shows that minimal of them is equal

$$S(f) = \frac{11.50}{2.33} > 10^{0.3}$$

which proves that all frequencies were separated with sufficient reliability.

Results

Using the new approach for stationary signal analysis showed, that spectral resolution ability of periodic signal highly depends on its total observation time. For instance, if we compare total observation time of original signal to the new one, we can see that for original signal total observation time is 2.94 sec, while for new signal formed by concatenation of left singular vectors, we will get signal with observation time 5.5 sec, which is approximately 2 times of original

signals time. In other words, we have 100% increase of observation time, as long frequency resolution limit and total observation time are reciprocal to each other, we are getting stronger ability to resolve (separate) components that are too much close to each other. Furthermore, we can easily detect that methodology based on singular value decomposition plays role as a filter, which means that in new time series level of noise is significantly reduced, while power of signal is increased, all these results prove that new method based on singular value decomposition is statistically stable and reliable.

Conclusion

In this article, it was shown that increasing spectra resolution by concatenating signal eigenvectors is also applicable for stationary signal that will assign this procedure universal character, which means that it overcomes disadvantages of existed methods connected to spectral resolution ability. Therefore, the method presented in this article can be used both for nonstationary and stationary signals, to significantly increase resolving capacity.

References

- Hayes, M. H. (1996). Statistical Digital Signal Processing and Modeling.
- James B. Elsner, A. A. (1996). Singular Spectrum Analysis: A New Tool in Time Series Analysis. New York: Springer; 1996 edition.
- Marple, S. L. (1987). Digital Spectral Analysis: With Applications. Prentice Hall.
- Milnikov, A. (2013). Singular decomposition of the data matrix of time series with periodic components and the resolving capacity of their pseudospectral estimation. Circuits, Systems and Communications.
- Milnikov, A. (2012). A Tensorial Approximations Method as a Universal Filter, International Journal of Mathematical Models and Methods in Applied Sciences, ISSN: 1998-0140, Issue 4, Volume 6, 2012, Pp. 600-607.