IBSU Scientific Journal 2010, 4(2), 149-155

Heuristic Analysis of Time Series Internal Structure

Cihan MERT
Alexander MILNIKOV

Abstract

A method of analysis of Time Series Internal Structures based on Singular Spectrum Analysis is discussed. It has been shown that in the case when the Time Series contains deterministic additive components rank of the trajectory matrices equal to number of parameters of the components. Also it was proved that both eigen and factor vectors repeat shapes of the additive components and both eigen values and eigen vectors can be divided into additive groups. Some useful patterns of deterministic components were identified, which permit to provide graphical analysis of times series Internal Structures.

Keywords: *Singular spectrum Analysis, Time series decomposition, Singular vectors, singular values, deterministic additive components, patterns.*

Cihan Mert *is a PhD candidate at Faculty of Computer Technologies and Engineering at International Black Sea University, Tbilisi, Georgia, cmert@hotmail.com*
Alexander Milnikov *is a professor at Faculty of Computer Technologies and Engineering at International Black Sea University, Tbilisi, Georgia,alexander_milnikov@yahoo.com*

Introduction

The main results of Singular Spectrum Analysis (SSA) method are [1, 2, 3]: 1. an initial time series $S(s_1,..,s_N)$ of length N is transformed into a sequence of multidimensional vectors which is represented as a trajectory matrix X of order L×K , where L – arbitrary integer represented the dimension of the vectors and K=N-L+1; 2. The matrix X can be decomposed as

$$X = \sum_{i=1}^{d} X_i \, , \qquad (1)$$

where d- rank of the matrix X (number of its singular values not equal to zero[1]);

$$X_i = \sqrt{\lambda_i}\, U_i \otimes Vi \, ,$$

where U_i- eigen vectors of a matrix $S=XX^T$, corresponded to a value of λ_i;

V_i - eigen vectors (they also are referred as factor vectors) of a matrix $S^T=X^TX$, corresponded to a non zero value of λ_i;

$\otimes$ -sign of operation of tensor production.

Relationships between vectors U_i and V_i exist:

$$V_i = X^T U_i \sqrt{\lambda_i} \qquad (2)$$

and

$$U_i = X V_i \sqrt{\lambda_i} \qquad (2')$$

We have to notice, that the matrices Xi are the first rank matrices.

Decomposition (1) reflects the internal structure of the time series under consideration [2]. Identification of the mathematical type of the components of (1) is very important to understand the nature of the time series and to separate valuable (from information point of view) components from noise components.

In the present article one heuristic approach for such kind of analysis of time series internal structures is discussed.

---

[1] Singular values of matrix X coincide with eigen values of matrix $S=XX^T$.

Basic part

In general the initial time series is a sum of the following additive factors[4-9]: deterministic functions (for example, n-rank polynomials, exponential, periodical and their combinations) and noisy components. The deterministic functions, beside they trivially depend on time, may also contain various parameters, such as: coefficients - in the case of polynomials, frequencies, phases, magnitude – in the case of periodic components, exponents of powers etc. It means that each element of the time series depends on the certain fixed number k of such kind of parameters.

Assume now that initial time series comprises only certain additive combination of above mentioned deterministic functions. In this case the following proposition can proved.

Proposition 1

Number of nonzero singular values of the matrix X (rank of the matrix S) equals to number of parameters of deterministic components of initial time series.

Proof

Each element of initial time series can be represented as $s_i(t,\theta)$, where θ is k-dimensional vector of real parameters. Assume that $k< \min(L,K)$. Clear those elements of the matrix X depend on the vector θ and on time variable t that is on $k + 1$ parameters. Each column x_i of the matrix X belongs to L-dimensional space L^L. Analogically, vector-row xi (it has K coordinates) belongs to K-dimensional space L_K but completely located in the $k + 1$-dimensional subspace of L_K. It leads that the vector-columns (vector-rows) of matrix $S=XX^T$ depend only on k parameters (time parameter t has been convolved in the issue of operation of product S=XXT) and both type of vectors (columns and rows) located in the k-dimensional subspace. But in turn it follows that only k vectors among them can be linearly independent and that rank of matrix S equals to k. Q.E .D.Assume now that the time series contain only one deterministic component, depended on k parameters. In that case the following proposition hold true.

If the rank of the matrix S is k and it is independent of Values of L and K, then the given time series has rank k.

Proposition 2

In the above mentioned conditions: 1. number of independent Eigen and factor vectors is equal to k; and 2. Both Eigen and factor vectors repeat shapes of the additive component.

Proof

Proposition 1 directly implies the first assertion.

The following consideration implies the second one. Relationship (2) between Eigen and factor vectors shows that Eigen vectors are linear combinations of basis vectors of the space L^K. To see it represents matrix X as block-matrix of column-vectors $|x_1, x_2,...,x_k|$[2], then vector-columns can be considered as basis vectors of L^L. Performing multiplication we immediately get

$$U_i = \sqrt{\lambda_i} \sum_{j=1}^{K} x_i v_i^j \quad \text{(i=1,2,…,L)} \tag{3}$$

or in coordinate form

$$U_i^l = \sqrt{\lambda_i} \sum_{j=1}^{K} x_j^l v_i^j \ . \ \text{(i=1,2,…,K; l=1,2,…,L)}, \tag{4}$$

where U_i^l - 1coordinate of i vector U; x_j^l - 1 coordinate of i vector x and v_i^j − j coordinate of i vectors of the space L^K with constant decomposition coefficients $\sqrt{\lambda_i}\, v_i^j$ at j coordinate of U or, in other words, each component of eigen vector U_i is weighted sum of corresponding components of vectors x_i represented additive components. But later proves the second assertion of the proposition as decomposition coefficients $\sqrt{\lambda_i}\, v_i^j$ change only scale of the deterministic component. The same can repeated for factor vectors. Q.E .D.

[2] Here we use upper indices as a vector components numbers

If there are several components Proposition 2 leads to the

Corollary

If the time series is of rank k and contains only m deterministic components and each of them depends on k_i parameters ($k_1+k_2+...+k_m=k$) then:

1. both Eigen values and Eigen vectors can be divided into m groups;

2. number of members of each group equal to number of parameters of the current component k_i;

To prove Corollary it is enough to apply Proposition 2 to each component of the time series.

The Propositions 1 and 2 permit to discuss and to identify effects of presence in the time series of some useful, from practical point of view, deterministic components

1. Polynomial Patterns

Polynomial patterns used to identify trends. Consider the general polynomial series $f_n=P_m(n)$, where $P_m(t)$ – is a polynomial function of order m. Since it completely defined by m+1 coefficients and if L>m and N enough large, one can say that rank of the time series is m+1 and singular vectors have polynomial structure.

1.1. The linear series $f_n=a+bn$ is a series of rank 2. If $L\geq 2$, then one can define two span vectors[3] of space : $(1,1,\ldots,1)^T$ and $(0,1,2,\ldots,L\text{-}1)^T$.

1.2. The quadratic sequence $f_n= a+bn+cn^2$. For $3\leq L\leq N\text{-}2$ rank of the time series is 3: space L^L is spanned by two vectors from previous example and the third one – $(0,1^2, 2^2, \ldots, (L\text{-}1)^2)$.

2.Exponential-periodic Patterns

The patterns are important because they depicted vibration processes (they represent solution solutions of linear differential equations). General shape of the patterns is

[3] Both are coefficients at $f_n=a+bn$ $(n=0,1,2,\ldots L\text{-}1)$

$$f_n = Ae^{\alpha n}\cos(2\pi f n + \varphi), \quad (5)$$

where f- is a frequency and φ - is a phase. Because of Nyquist frequency we assume that $f \in [0,1/2]$.

2.1. f=0 and $\cos\varphi \neq 0$. Then we have only one parameter α and the time series has rank1. The bases vector is $(1, e^{\alpha}, e^{2\alpha}, ..., e^{\alpha(L-1)})$;

2.2. f=1/2 and $\cos\varphi \neq 0$. In this case $f_n = -Ae^{\alpha n}\cos(\pi n + \varphi)$, the time series again has rank=1 and span vector is $(1, -e^{\alpha}, e^{2\alpha}, ..., (-1)^{L-1}e^{\alpha(L-1)})$.

2.3. $f \in (0,1/2)$. The sequence (5) has rank 2, because the sequence depends on 2 parameters: α and f. Components of two basis vectors of L^2 are: $x_k = e^{\alpha(k-1)}\cos(2\pi f(k-1))$ and $y_k = e^{\alpha(k-1)}\sin(2\pi f(k-1))$ (1<k≤L-1).

2.4. $\alpha = 0$ and $f \in (0,1/2)$. This is pure periodical pattern. The time series has rank=2 and two basis vectors: $x_k = \cos(2\pi f(k-1))$ and $y_k = \sin(2\pi f(k-1))$ (1<k≤L-1).

Listed items permit to provide graphical analysis of times series decomposed according to (1).

Conclusion

It was showed that both Eigen and factor vectors repeat shapes of the additive components and both Eigen values and Eigen vectors can be divided into additive groups. The latter permitted to identify several useful practically deterministic patterns: linear and quadratic trends, pure periodical trends, periodic exponential trends etc. The patterns can be efficiently used for practical problems of time series analysis.

References

Fowler A.C. and Kember G., 1998, Singular systems analysis as a mowing-window spectral method. European Journal of Applied Mathematics, 9:55–79.

Nekrutkin V.V., 1997, Decomposition of time series. In D. Danilov and A. Zhigljavsky, editors, Principal Components of Time Series: the 'Caterpillar' method, pages 194–227. University of St.Petersburg, St.Petersburg.

Bouvet M. and Clergeot H., 1988, Eigen and singular value decomposition technique for the solution of harmonic retrieval problems. In E.F. Deprettere, editor, SVD and Signal

Processing: Algorithms, Applications and Architectures, pages 93–114. North-Holland, Amsterdam.

Anderson T.W., 1994, Statistical Analysis of Time Series. Wiley, New York.

Brillinger D., 1975, Time Series. Data Analysis and Theory. Holt, Rinehart and Winston, Inc., New York.

Worsley K.J. and Friston K.J. ., 1995, Analysis of fMRI Time-Series Revisited – Again. NeuroImage, 2:173-181.

Goertzel B., 1998, Delay Vectors as Perceptual Chunks: Understanding Nonlinear Time Series Analysis as Self-Organizing Cognition. Journal of Dynamical Psychology.

Goertzel B., 1996, Subself Dynamics in Human and Computational Intelligence. CC-AI: The Journal for the Integrated Study of Artificial Intelligence, Cognitive Science and Applied Epistemology, Vol. 13, No. 2-3, pp. 115-140.

Howes R., Eccleston L., Gonçalves J., Stan G.-B., Warnick S., 2008, Dynamical structure analysis of sparsity and minimality heuristics for reconstruction of biochemical networks. Proceedings of the 47th IEEE Conference on Decision and Control (CDC 2008), Cancun, Mexico, 9-11 December.