Singular spectrum analysis for time series

Anatoly Zhigljavsky,
School of Mathematics, Cardiff University, Cardiff, U.K.

Singular spectrum analysis (SSA) is a technique of time series analysis and forecasting. It combines elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing. SSA aims at decomposing the original series into a sum of a small number of interpretable components such as a slowly varying trend, oscillatory components and a 'structureless' noise. It is based on the singular-value decomposition of a specific matrix constructed upon time series. Neither a parametric model nor stationarity-type conditions have to be assumed for the time series; this makes SSA a model-free technique.

The commencement of SSA is usually associated with publication of the papers [1,2] by Broomhead and King. Nowadays SSA is becoming more and more popular, especially in applications. There are several hundred papers published on methodological aspects and applications of SSA, see [3–7] and references therein. SSA has proved to be very successful, and has already become a standard tool in the analysis of climatic, meteorological and geophysical time series; see, for example, [4-6]. More recent areas of application of SSA include engineering, medicine, econometrics and many other fields. Most recent developments in the theory and methodology of SSA can be found in [7]. We start with 'Basic SSA', which is the most common version of SSA.

Basic SSA.

Let $x_1,\dots,x_N$ be a time series of length N. Given a window length L $(1<L<N)$, we construct the L-lagged vectors $X_i=(x_i,\dots,x_{i+L-1})^T$, $i=1,2,\dots,K=N-L+1$, and compose these vectors into the matrix $\mathbf{X}=(x_{i+j-1})_{i,j=1}^{L,K}=[X_1:\dots:X_K]$. This matrix has size $L\times K$ and is often called 'trajectory matrix'. It is a Hankel matrix, which means that all the elements along the diagonal $i+j=$const are equal.

The columns X_j of $\mathbf{X}$, considered as vectors, lie in the L-dimensional space $\mathbb{R}^L$. The singular-value decomposition of the matrix $\mathbf{X}\mathbf{X}^T$ yields a collection of L eigenvalues and eigenvectors. A particular combination of a certain number $l<L$ of these eigenvectors determines an l-dimensional subspace in $\mathbb{R}^L$. The L-dimensional data $\{X_1,\dots,X_K\}$ is then projected onto this l-dimensional subspace and the subsequent averaging over the diagonals gives us some Hankel matrix $\tilde{\mathbf{X}}$ which is considered as an approximation to $\mathbf{X}$. The series reconstructed from $\tilde{\mathbf{X}}$ satisfies some linear recurrent formula which may be used for forecasting.

In addition to forecasting, the Basic SSA can be used for smoothing, filtration, noise reduction, extraction of trends of different resolution, extraction of periodicities in the form of modulated harmonics, gap-filling [8,9] and other tasks, see [3]. Also, the Basic SSA can be modified and extended in many different ways some of which are discussed below.

Extensions of the Basic SSA.

SSA for analyzing stationary series [5]. Under the assumption that the series $x_1,\dots,x_N$ is stationary, the matrix $\mathbf{X}\mathbf{X}^T$ of the Basic SSA is replaced with the so-called lag-covariance matrix $\mathbf{C}$ whose elements are $c_{ij}=\frac{1}{N-k}\sum_{t=1}^{N-k}x_tx_{t+k}$ with $i,j=1,\dots,L$ and $k=|i-j|$. In the terminology of [3], this is 'Toeplitz SSA'.

Monte-Carlo SSA [6]. In the Basic SSA we implicitly associate the 'structureless' component of the resulting SSA decomposition with 'white noise' (this noise may not necessarily be random). In some applications, however, it is more natural to assume that the noise is 'coloured'.

In this case, special tests based on Monte Carlo simulations may be used to improve the quality of the separation of signal from noise.

Improvement or replacement of the singular-value decomposition (SVD) procedure. There are two main reasons why it may be worthwhile to replace the SVD operation in the Basic SSA with some another operation. The first reason is simplicity: in problems where the dimensions of the trajectory matrices is large, SVD may simply be too costly to perform; substitutions of SVD are available, see [10,11]. The second reason is the analysis of the accuracy of SSA procedures based on the perturbation theory [7]. For example, in the problems of separating signal from noise, some parts of the noise are often found in SVD components corresponding to the signal. As a result, a small adjustment of the eigenvalues and eigenvectors is advisable to diminish this effect. The simplest version of the Basic SSA with a constant adjustment in all eigenvalues was suggested in [12] and is sometimes called the minimum-variance SSA.

Low-rank matrix approximations, Cadzow iterations, connections with signal processing. As an approximation to the trajectory matrix $\mathbf{X}$, the Basic SSA yields the Hankel matrix $\tilde{\mathbf{X}}$. This matrix is obtained as a result of the diagonal averaging of a matrix of rank l. Hence $\tilde{\mathbf{X}}$ is typically a matrix of full rank. However, in many signal processing applications, when a parametric form of an approximation is of prime importance, one may wish to find a Hankel matrix of size $L \times K$ and rank l which gives the best approximation to $\mathbf{X}$; this is a problem of the structured low-rank approximation [13]. The simplest procedure of finding a solution to this problem (not necessarily the globally optimal one though) is the so-called Cadzow iterations [14] which are the repeated alternating projections of the matrices (starting at $\mathbf{X}$) to the set of matrices of rank l (by performing the singular-value decompositions) and to the set of Hankel matrices (by making the diagonal averaging). That is, Cadzow iterations are simply the repeats of the Basic SSA. It is not guaranteed however that Cadzow iterations lead to more accurate forecasting formulas than the Basic SSA [7].

SSA for change-point detection and subspace tracking [15]. Assume that the observations $x_1, x_2, \ldots$ of the series arrive sequentially in time and we apply the Basic SSA to the observations at hand. Then we can monitor the distances from the sequence of the trajectory matrices to the l-dimensional subspaces we construct and also the distances between these l-dimensional subspaces. Significant changes in any of these distances may indicate on a change in the mechanism generating the time series. Note that this change in the mechanism does not have to affect the whole structure of the series but rather only a few of its components.

SSA for multivariate time series. Multivariate (or multichannel) SSA (shortly, MSSA) is a direct extension of the standard SSA for simultaneous analysis of several time series. Assume that we have two series, $X = \{x_1, \ldots, x_N\}$ and $Y = \{y_1, \ldots, y_N\}$. The (joint) trajectory matrix of the two-variate series (X, Y) can be defined as either $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ or $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})^T$, where $\mathbf{X}$ and $\mathbf{Y}$ are the trajectory matrices of the individual series X and Y. Matrix $\mathbf{Z}$ is block-Hankel rather than simply Hankel. Other stages of MSSA are identical to the stages of the univariate SSA except that we build a block-Hankel (rather than ordinary Hankel) approximation $\tilde{\mathbf{Z}}$ to the trajectory matrix $\mathbf{Z}$.

MSSA may be very useful for analyzing several series with common structure. MSSA may also be used for establishing a causality between two series. Indeed, the absence of causality of Y on X implies that the knowledge of Y does not improve the quality of forecasts of X. Hence an improvement in the quality of forecasts for X which we obtain using MSSA against univariate SSA forecasts for X gives us a family of SSA-causality tests, see [16].

References

[1] Broomhead D.S. and King G. P. (1986a) Extracting qualitative dynamics from experimental data. Physica D, **20**, 217-236.

[2] Broomhead D. S. and King G. P. (1986b) On the qualitative analysis of experimental dynamical systems. In: Nonlinear Phenomena and Chaos, Adam Hilger, Bristol, 113-144.

[3] Golyandina N., Nekrutkin V. and Zhigljavsky A. (2001). *Analysis of Time Series Structure: SSA and related techniques*, Chapman & Hall/CRC, N.Y. – London.

[4] Vautard R., Yiou P. and Ghil M. (1992). Singular spectrum analysis: a toolkit for short, noisy and chaotic series, Physica D, **58**, 95–126.

[5] Vautard R. and Ghil M. (1989) Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, Physica D, 35, 395-424.

[6] Allen M.R. and Smith L.A. (1996) Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise, Journal of Climate, **9**, 33733404.

[7] Zhigljavsky A., ed. (2010) *Statistics and Its Interface*, Special issue on the singular-spectrun analysis for time series, Springer.

[8] Kondrashov D. and Ghil M. (2006) Spatio-temporal filling of missing points in geophysical data sets, Nonlin. Processes Geophys., 13, 151-159.

[9] Golyandina N. and Osipov E. (2007) The Caterpillar-SSA method for analysis of time series with missing values, J. of Statist. Plann. Inference, **137**, 2642–2653.

[10] Golub G. and van Loan C. (1996) *Matrix computations*, Johns Hopkins University Press.

[11] Moskvina V. and Schmidt K. M. (2003) Approximate projectors in singular spectrum analysis. SIAM J. Matrix Anal. Appl. **24**, 932–942.

[12] Van Huffel S. (1993) Enhanced resolution based on minimum variance estimation and exponential data modeling, Signal processing, **33**, 333–355.

[13] Markovsky I., Willems J.C., Van Huffel S. and De Moor B. (2006) *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*, SIAM.

[14] Cadzow J. A. (1988) Signal enhancement a composite property mapping algorithm, IEEE Transactions on Acoustics, Speech, and Signal Processing, **36**, 49-62.

[15] Moskvina V. G. and Zhigljavsky A. (2003). An algorithm based on singular spectrum analysis for change-point detection, Commun. Statist. – Simul. and Comput., **32**, 319–352.

[16] Hassani H., Zhigljavsky A., Patterson K. and Soofi A. (2010). A comprehensive causality test based on the singular spectrum analysis, *Causality in Science*, Oxford University Press.