

Singular Spectrum Analysis for time series: Introduction to this special issue

ANATOLY ZHIGLJAVSKY

In this introduction we briefly describe the methodology of the Singular Spectrum Analysis (SSA), some versions and extensions of the basic version of SSA as well as connections between SSA and subspace-based methods in signal processing. We also briefly touch upon some history of SSA and mention some areas of application of SSA.

AMS 2000 SUBJECT CLASSIFICATIONS: Primary 62M20, 60G35.

KEYWORDS AND PHRASES: Singular Spectrum Analysis, Time series analysis, Subspace-based methods.

INTRODUCTION

General. 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. From the algorithmic point of view, SSA can be considered as a typical subspace-based method of 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 (SVD) of a specific matrix constructed upon the 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 and hence enables SSA to have a very wide range of applicability.

History. The first publication, where the key ideas of SSA and more generally of the subspace-based methods of signal processing have been formulated, can be traced back to the eighteenth century [1]. However, the commencement of SSA is usually associated with publication in 1986 of the papers [2, 3] by Broomhead and King. Since then SSA has received much attention in literature. Arguably, the most influential papers on SSA published in the nineteen eighties and nineties, in addition to [2, 3], are [4–7].

The three books [8–10] are fully devoted to SSA. [8] is the first book on SSA; it is well-written but provides only an elementary introduction to SSA. [9] is a collection of papers

written entirely by St. Petersburg statisticians. All these papers are devoted to the so-called ‘Caterpillar’ methodology. This methodology is a version of SSA developed in the former Soviet Union independently of the main-stream SSA (remember ‘the iron curtain’!). The work on the ‘Caterpillar’ methodology has started long after the year of publication of [1] but well before 1986, the year of publication of [2, 3].

The main difference between the main-stream SSA of [2–7] and the ‘Caterpillar-SSA’ is not that much in the algorithmic details but rather in the assumptions and in the emphasis in the study of the SSA properties. To apply the main-stream SSA, one often needed to assume some kind of stationarity of the time series and think in terms of the ‘signal plus noise’ model (where the noise is often assumed to be ‘red’ rather than simply ‘white’). In the ‘Caterpillar-SSA’, the main methodological concept is the separability (of one subseries from another one) and neither the assumption of stationarity nor the model in the form ‘signal plus noise’ are required.

The main methodological principles described in [9] have been further developed in the monograph [10]. To some extent, this special issue follows the style of [9] and [10]. On the other hand, the recent developments, which are well represented in the present issue, are made in the style of classical time series analysis, classical signal processing and classical statistics. As a result, at present we do not only have many different versions of SSA but also a general methodology called ‘singular spectrum analysis’. This methodology unifies all these versions of SSA into a very powerful tool of time series analysis and forecasting.

Application areas. 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, 5] (climatology), [12] (meteorology), [13] (marine science), [14] (geophysics); for more references see [4–10, 14] and the papers in the present issue. More recent areas of application of SSA include engineering, medicine, econometrics and many other fields; for references see e.g. papers in the present issue, especially [23, 26, 29, 31–35].

BASIC SSA

The most common version of SSA is called ‘Basic SSA’. Below we give a short description of it. For details, see [10, Chapter 1] and [23].

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 $i + j = \text{const}$ are equal.

The columns X_j of \mathbf{X} can be considered as vectors 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 of these eigenvectors determines an l -dimensional subspace \mathbb{L} in \mathbb{R}^L , $l < L$. The L -dimensional data $\{X_1, \dots, X_K\}$ is then projected onto the subspace \mathbb{L} and the subsequent averaging over the diagonals yields some Hankel matrix $\tilde{\mathbf{X}}$ which is considered as an approximation to \mathbf{X} . The time series $\tilde{x}_1, \dots, \tilde{x}_N$, which is in the one-to-one correspondence with the matrix $\tilde{\mathbf{X}}$, provides an approximation to the original series x_1, \dots, x_N .

Forecasting [10, Sect. 1.2]. The l -dimensional subspace \mathbb{L} constructed by the Basic SSA yields a linear recurrent formula which may be used for forecasting. As an alternative to this method of forecasting we may use the so-called ‘vector forecasting’. The main idea of the vector forecasting algorithm is in the consecutive construction of the vectors $X_i = (x_i, \dots, x_{i+L-1})^T$, for $i = K + 1, K + 2, \dots$ so that they lie as close as possible to the subspace \mathbb{L} created by the chosen l eigenvectors.

Other tasks [10, Chapter 1]. 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, filling in missing values [11, 14, 15] and other tasks. Also, the Basic SSA can be modified and extended in many different ways; some of them are discussed below.

Choice of parameters in the Basic SSA. There are two parameters to choose in the Basic SSA: the window length L and the group of l indices which determine the subspace \mathbb{L} . A rational or even optimal choice of these parameters should depend on the task we are using SSA for. A detailed discussion on this topic can be found in [23] and [10, Section 1.6]. There are versions of the Basic SSA where given the window length L , the group of l indices determining the subspace \mathbb{L} (for extraction of either trend or periodic components) is chosen automatically; see e.g. [6] and <http://www.gistatgroup.com/gus/autossa2.pdf>.

EXTENSIONS OF THE BASIC SSA

SSA for stationary series [5, 6]. 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 $c_{ij} = \frac{1}{N-k} \sum_{t=1}^{N-k} x_t x_{t+k}$ with

$i, j = 1, \dots, L$ and $k = |i - j|$. In the terminology of [10], this is ‘Toeplitz SSA’. Comparison of the Basic and Toeplitz SSA is discussed in [23]. Unsurprisingly, if the original series is stationary then the Toeplitz SSA slightly outperforms the Basic SSA. However, if the series is not stationary then the use of Toeplitz SSA may yield wrong results.

Monte-Carlo SSA [7]. 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 cases, special tests based on the Monte Carlo simulations may be used to test the hypothesis of the presence of a signal. If the signal is weak then it can only be detected but not extracted (separated from noise). Note also that the sensitivity of the Basic SSA to the presence of correlations in the noise component is discussed in [23].

Improvement or replacement of the SVD procedure.

A description of techniques that could increase the efficiency of computation of the SVD for the matrices appearing in SSA can be found in [30].

There are two main reasons why it may be worthwhile to replace the SVD operation in the Basic SSA with some other operation. The first reason is simplicity: in problems where the dimension of the trajectory matrix is very large, SVD may be too costly to perform; substitutions of SVD are available, see [16, 17]. The second reason is the analysis of the accuracy of SSA procedures based on the perturbation theory [25]. For example, in the problems of separating signal from noise, some parts of noise are often found in the SVD components mostly related 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 [18] (see also [32]) and is sometimes called the minimum-variance SSA.

Low-rank matrix approximations, Cadzow iterations and connections with signal processing.

As an approximation to the trajectory matrix \mathbf{X} , the Basic SSA yields a 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. In signal processing applications, however, a parametric form of an approximation is typically of prime importance, and hence one may wish to find a Hankel matrix of size $L \times K$ and rank $\leq l$ which gives the best approximation to \mathbf{X} ; this is a problem of the structured low-rank approximation [27]. The simplest procedure of finding a solution to this problem (not necessarily the globally optimal one though) is the so-called Cadzow iterations [19], which are the repeated alternating projections of the matrices (starting at \mathbf{X}) to the set of matrices of rank $\leq l$ (by performing the SVDs) 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 (and more generally, the solution to the structured low-rank approximation problem) lead to more accurate forecasting formulas than the Basic SSA [28].

SSA for change-point detection and subspace tracking [20], [10, Sect. 1.2]. Assume that the observations x_1, x_2, \dots 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 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, $X = \{x_1, \dots, x_N\}$ and $Y = \{y_1, \dots, y_N\}$. 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 ones 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 the univariate SSA forecasts for X , gives an evidence of a possible causality of Y on X . This observation can be formalized in the form of certain SSA-causality tests, see [21, 32].

2D-SSA for image processing. 2D-SSA is a straightforward extension of the Basic SSA and MSSA for analyzing images. The only difference between these three versions of SSA is in the construction of the trajectory matrix, see [9, 22, 35].

OVERVIEW OF THE PRESENT ISSUE

This issue contains one literature review [27] and twelve research papers [23–26, 28–35]. These papers are devoted to the SSA theory, the SSA methodology, computational aspects of SSA and applications of SSA.

Theory of SSA. The main theoretical paper in this issue is [25]; this paper is fully devoted to the perturbation analysis in SSA and subspace-based methods of signal processing. Another important theoretical paper is [24], where the concept of SSA separability is further developed (relative to [10]) and studied through the apparatus of the roots of characteristic polynomials of the linear recurrent formulas of the SSA approximation of the signal (in the ‘signal plus noise’ model).

Methodology of SSA. The key paper on the methodology of SSA is [23]. This is an essential read for anyone who is interested in applying SSA beyond trivial. The main question which is discussed in [23] is the choice of the window length, but the discussions in the paper go far further than this. Elements of SSA methodology are also discussed in [24, 26, 28, 29] and [32].

SSA and signal processing. Many discussions in the papers on the theory and methodology of SSA touch the important question of the similarities and dissimilarities between SSA and the subspace-based methods of signal processing. Although there are many similarities between them, there is a fundamental difference too. This difference lies in fact that the model of the form ‘signal plus noise’ is obligatory in the signal processing; consequently, the main aim of the signal processing methods is the estimation of the parameters of the model (which in our case is a linear recurrent formula). In SSA the aims of the analysis may be different (for example, forecasting) and the parameters of the approximatory time series are of secondary importance. The similarities and dissimilarities between SSA and the subspace-based methods of signal processing are discussed in [23–25], see also [27–29].

Computational aspects of SSA. In [30], a thorough study of computational aspects of SSA procedures is performed. Also, [30] provides a description of techniques that could be used to increase the efficiency of these procedures. Using the techniques suggested in [30], a standard laptop can routinely perform SSA with the window lengths up to several thousand.

Comparison of SSA with other techniques. From the methodological point of view, SSA is compared with subspace-based techniques of signal processing in [23, 25] and [28]. Numerical comparison of SSA with ARIMA and other classical methods of time series analysis can be found in [29, 31–34].

Applications of SSA. The papers [31, 33–35] are devoted to applications of SSA in new areas. These areas are: prediction of steel prices [31], extraction of a weak fetal heart signal from noisy maternal ECG [33], DNA microarray gene expressions [34] and image processing [35]. Applications of SSA are also a significant part of the papers [26, 29, 32] where methodological aspects of SSA are considered.

REFERENCES

- [1] DE PRONY, G. S. (1795). Essai experimental et analytique: sur les lois de la dilatabilité de fluides elastique et sur celles de la force expansive de la vapeur de l'alkool, a differentes temperatures. *Journal de l'Ecole Polytechnique* **1**(22) 24–76.
- [2] BROOMHEAD, D. S. and KING, G. P. (1986a). Extracting qualitative dynamics from experimental data. *Physica D* **20** 217–236. [MR0859354](#)
- [3] 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. [MR0854700](#)
- [4] FRAEDRICH, K. (1986). Estimating the dimension of weather and climate attractors. *J. Atmos. Sci.* **43** 419–432. [MR0838629](#)
- [5] VAUTARD, R. and GHIL, M. (1989). Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D* **35**, 395–424. [MR1004204](#)
- [6] VAUTARD, R., YIOU, P., and GHIL, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signal. *Physica D* **58**, 95–126.
- [7] ALLEN, M. R. and SMITH, L. A. (1996) Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise, *Journal of Climate*, **9**, 3373–3404.
- [8] ELSNER, J. B. and TSONIS, A. A. (1996). *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. Plenum.
- [9] DANILOV, D. and ZHIGLJAVSKY, A. A. eds (1997) *Principal Components of Time Series: 'Caterpillar' Method*, St. Petersburg University (in Russian).
- [10] GOLYANDINA, N. E., NEKRUTKIN, V. V., and ZHIGLJAVSKY, A. (2001). *Analysis of Structure of Time Series: SSA and Related Techniques*, CRC / Chapman & Hall, Boca Raton. [MR1823012](#)
- [11] <http://www.gistatgroup.com/cat/>
- [12] WEARE, B. C. and NASSTROM, J. S. (1982). Examples of Extended Empirical Orthogonal Functions. *Monthly Weather Review* **110** 481–485.
- [13] COLEBROOK, J. M. (1978). Continuous plankton records – zooplankton and environment, northeast Atlantic and North Sea, 1948–1975. *Oceanol. Acta* **1** 9–23.
- [14] KONDRASHOV, D. and GHIL, M. (2006) Spatio-temporal filling of missing points in geophysical data sets, *Nonlin. Processes Geophys.*, **13**, 151–159.
- [15] 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. [MR2326115](#)
- [16] GOLUB, G. and VAN LOAN, C. (1996) *Matrix computations*, Johns Hopkins University Press. [MR1417720](#)
- [17] MOSKVIN, V. and SCHMIDT, K. M. (2003) Approximate projectors in singular spectrum analysis. *SIAM J. Matrix Anal. Appl.*, **24**, 932–942. [MR2003313](#)
- [18] VAN HUFFEL, S. (1993) Enhanced resolution based on minimum variance estimation and exponential data modeling, *Signal processing*, **33**, 333–355.
- [19] CADZOW, J. A. (1988). Signal Enhancement – A Composite Property Mapping Algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing* **36**(1) 49–62.
- [20] MOSKVIN, 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. [MR1983338](#)
- [21] 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 (in press).
- [22] GOLYANDINA, N. E. and USEVICH, K. D. (2010). 2D-extension of singular spectrum analysis: algorithm and elements of theory. In: *Matrix Methods: Theory, Algorithms, Applications* World Scientific, 449–473.
- [23] GOLYANDINA, N. (2010). On the choice of parameters in Singular Spectrum Analysis and related subspace-based methods. *Statistics and Its Interface* **3** 259–279.
- [24] USEVICH, K. (2010). On signal and extraneous roots in Singular Spectrum Analysis. *Statistics and Its Interface* **3** 281–295.
- [25] NEKRUTKIN, V. (2010). Perturbation expansions of signal subspaces for long signals. *Statistics and Its Interface* **3** 297–319.
- [26] PEPELYSHEV, A. and ZHIGLJAVSKY, A. (2010). Assessing the stability of long-horizon SSA forecasting. *Statistics and Its Interface* **3** 321–327.
- [27] MARKOVSKY, I. (2010). Bibliography on total least squares and related methods. *Statistics and Its Interface* **3** 329–334.
- [28] GILLARD, J. (2010). Cadzow's basic algorithm, alternating projections and singular spectrum analysis. *Statistics and Its Interface* **3** 335–343.
- [29] TOME, A. M., TEIXEIRA, A. R., FIGUEIREDO, N., SANTOS, I., GEORGIEVA, P., and LANG, E. (2010). SSA of biomedical signals: a linear invariant systems approach. *Statistics and Its Interface* **3** 345–355.
- [30] KOROBEYNIKOV, A. (2010) Computation- and space-efficient implementation of SSA. *Statistics and Its Interface* **3** 357–368.
- [31] MUELLER, W. G. and KAPL, M. (2010). Prediction of steel prices: a comparison between a conventional regression model and MSSA. *Statistics and Its Interface* **3** 369–375.
- [32] HASSANI, H. and THOMAKOS, D. (2010). A Review on Singular Spectrum Analysis for economic and financial time series. *Statistics and Its Interface* **3** 377–397.
- [33] GHODSI, M., HASSANI, H., and SANEI, S. (2010). Extracting fetal heart signal from noisy maternal ECG by multivariate Singular Spectrum Analysis. *Statistics and Its Interface* **3** 399–411.
- [34] TANG TSZ-YAN, LIEW WEE-CHUNG, and YAN HONG (2010). Periodicity analysis of DNA microarray gene expression time series profiles in mouse segmentation clock data. *Statistics and Its Interface* **3** 413–418.
- [35] RODRÍGUEZ-ARAGÓN, L. and ZHIGLJAVSKY, A. (2010). Singular Spectrum Analysis for image processing. *Statistics and Its Interface* **3** 419–426.

Anatoly Zhigljavsky
 School of Mathematics
 Cardiff University
 Cardiff CF24 4AG, UK
 E-mail address: ZhigljavskyAA@cf.ac.uk