A Technique of Incorporating Spatial Dependence in MSSA Forecasts

Awichi, R.O.(ichbinrao@gmail.com);
Johannes Kepler University and
Uganda Martyrs University,
Kampala, Uganda.

Abstract

In this discussion, I present three major issues of time series analysis, namely: Singular Spectrum Analysis, SSA; Multivariate Singular Spectrum Analysis, MSSA and Spatial Dependence.
SSA is a recently developed tool for time series analysis. It is a model free approach to time series analysis and literally any time series with a notable structure can provide an application of SSA. Indeed it has a wide area of application including social sciences, medical sciences, finance, environmental sciences, mathematics, dynamical systems and economics. It is implemented under the platform of a software called CaterpilarSSA, although other packages, exist.
The aim of SSA is twofold: i) to make a decomposition of the original series into a sum of a small number of independent and interpretable components such as a slowly varying trend, oscillatory components and a structure less noise; ii) to reconstruct the decomposed series to make a forecast in the absence of the noise component.
MSSA is an extension of SSA to multivariate statistics and takes advantage of the delay procedure to obtain a similar formulation as SSA though with larger matrices for multivariate data. In environmental sciences and other areas where spatial data is an important focus of investigation, it is not uncommon to have attributes whose values change with space and time and quite often, due to spillovers or unobservable variables or omitted factors. This leads to spatial dependence that subsequently influence data analysis.
In light of spatial dependence, Kriging and inverse distance weighting techniques may be used to improve SSA and MSSA predictions. Here, I explore the inverse distance approach to improve on the forecast of the values. This technique is applied to climate data recordings from Upper Austria.

Keywords: *Time Series Analysis, Kriging, Inverse Distance Weighting and Spatial Dependence.*

1 Introduction

SSA, a recently developed tool for time series analysis, is a model free approach to the analysis of time series . The beginning of SSA is usually attributed to Broomhead et al. (1986). Literally, any time series with a notable structure can provide an application of SSA. It can be applied to many areas, see Golyandina et al. (2001), Hassani et al. (2010) and Kapl et al. (2010) for details. It is implemented under the platform of a software called CaterpilarSSA, Golyandina et al (2001) though other platforms exist.
SSA has two broad aims:

i) to make a decomposition of the original series into a sum of a small number of independent and interpretable components such as a slowly varying trend, oscillatory components and a structure less noise;
ii) to reconstruct the decomposed series so as to make a forecast in the absence of the noise component.
MSSA is a direct extension of SSA to multivariate analysis and has been applied before to climate studies, S. Raynaud et al. (2005).
Kriging and inverse distance techniques are ways in which spatial information can be included in the predictions of SSA and subsequently MSSA. These interpolation techniques have been used before to climate studies, see Hopkins et al. (1999) for details.

Section 2 contains a brief review of the basics of SSA. Section 3 discusses MSSA as an extension of the SSA techniques to multivariate data and presents a method of describing spatial dependence in MSSA. Results are in Section 4 while Conclusions appear in Section 5.

2 Singular Spectrum Analysis, SSA

The Basic SSA, as it is commonly referred to, has two main stages: Decomposition and Reconstruction; each of these stages consists of two steps as described below.
Let $T_N = \{t_0, t_2, \cdots, t_{N-1}\}$ be a real valued nonzero time series data of sufficient length N without missing values.
The first step is the Embedding step in which the (one) dimensional time series is transformed into a multidimensional lagged vector matrix $X = X_1 : \cdots : X_K$ of dimension $L \times K$, L being the window length and $K = N - L + 1$ whereby the rows and columns are subseries of the original series. This (Hankel) matrix is referred to as the trajectory matrix.
The second step is the Singular Value Decomposition, SVD. In this step, the trajectory matrix is factorized into a sum of elementary matrices by calculating and ordering eigenvalues λ_i of the matrix $S = XX^T$, calculating left and right singular vectors U_i and V_i and representing X as, $X = \sum_{i=1}^{d} X_i$ where d is the number of nonzero eigenvalues of S.
The third step is the Grouping step, the first of the Reconstruction stage. Here, the elementary matrices are split further into what is called resultant matrices through the eigentriple grouping.
Finally the last step of Diagonal Averaging transfers each resultant matrix into a time series through a process known as Hankelization. These steps can be summarized as in the algorithm below:
Step 1. (Computing trajectory matrix)
This is embedding where the one dimensional time series is converted into a multidimensional data of the trajectory matrix.
Step 2. (Constructing a matrix for SVD)
The matrix $S = XX^T$ is computed.
Step 3. (SVD of S)
The eigenvalues and eigenvectors of S are computed such that S is represented

as $S = PDP^T$. Here $D = diag(\lambda_1, \cdots, \lambda_L)$ is the diagonal matrix of the eigenvalues of S ordered so that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_L \geqslant 0$ and $P = (P_1, P_2, \cdots, P_L)$ is the corresponding orthogonal matrix of the eigenvectors of S.
Step 4. (Grouping)
Selection of m eigenvectors $P_{i_1}, \cdots, P_{i_m}$ is done. This complements the splitting of the elementary matrices X_i into several groups and summing as discussed earlier.
Step 5. (Reconstruction of the 1D series)
The one dimensional time series is reconstructed through diagonal averaging to complete the procedure of the software.
The details of these steps can be found in Golyandina et al. (2001).

One of the main concepts in SSA is the aspect of separability of the original time series into signal and noise so that the analysis and forecasting can be done on signal in the absence of noise. This implies that if the original time series T_N can be written as $T_N = T_N^{(1)} + T_N^{(2)}$, then the trajectory matrix can be written as $X = X_N^{(1)} + X_N^{(2)}$ and the series is said to be (weakly) separable. Any further analysis will depend on the purpose for which the time series data was obtained, Hassani (2007). Once the time series is separable into signal and noise, then forecasting can be done on the signal without the noise component. Forecasting in SSA is done using the linear recurrence relations. A time series T_N is governed by a linear recurrence formula if it can be expressed as a linear combination of products of exponential, polynomial and harmonic series. For a separable time series, the signal component can always be expressed as a linear combination of these series.

3 Multivariate Singular Spectrum Analysis, MSSA

This is an extension of SSA to multidimensional data in which the trajectory of the m variate time series is given as; $X = (X^{(1)}, \cdots, X^{(r)}, \cdots, X^{(m)})$, X is of order $L \times mK$; K. Patterson et al. (2011), A. Zhigljavsky et al. (2009), Kapl et al. (2010).
The algorithms of MSSA are an extension of those of SSA. For emphasis, we present a brief discussion of the MSSA algorithms below.
Let $F^{(i)} = f_j^{(i)}; j = 1, \cdots, N$ and $i = 1, \cdots, m$ denote the multivariate time series.

Step 1: Embedding
For L $(1 < L < N)$ the window length, the embedding procedure produces K (K= N - L +1) lagged vectors, $X_j^{(i)}$ of the trajectory matrix X $(L \times mK)$ given as,
$X = [X^{(1)} : \cdots X^{(m)}]$; this is a block Hankel matrix.
Each of the m blocks of K columns corresponds to the trajectory matrix for a particular vintage. For simplicity, we have assumed here that N and K are uniform, otherwise if they different for each block, then the individual trajectory matrices are stacked horizontally and may have different column dimensions.
The trajectory space is the linear space spanned by the lagged vectors.

Step 2: SVD
This is basically similar to step 2 of SSA, though it is also possible to use the Principal Component Analysis, PCA instead of the SVD at this step. The PCA extracts the orthogonal components of the initial series to achieve a reduced dimensionality of the trajectory matrix, K. Patterson et al. (2011).
The major aim of the first two steps is to achieve separability of the components in the decomposition of the series. This makes the selection of L critical: it should not be too small since then not all the components will be captured, neither should it be too big because it becomes rather difficult to trace the behaviour of the series with only the very few "windows" in use. It should just be large enough to capture the essential behaviour of the time series. The degree of separability can be assessed empirically by means of the corresponding w- correlation coefficients. This step represents the trajectory matrix X as, $X = X_1 + \cdots + X_d$

Step 3: Grouping
Once the expansion, $X = X_1 + \cdots + X_d$ is achieved, the grouping procedure partitions the set of indices $\{1, \cdots, d\}$ into say s disjoint subsets $I_1, \cdots, I_s$ so that the resultant matrix X is represented in terms of these partitioned indices as in the SSA grouping step, i.e. $X = X_{I_1} + \cdots + X_{I_s}$.

Step 4: Diagonal Averaging
This last step is in a way opposite to step 1 and transforms each matrix of the grouped decomposition into a system of new (reconstructed) series of length N by the block Hankelization procedure below, Golyandina et al. (2001).
Forecasting and Separability of the MSSA technique are direct extension of the analysis given for SSA.
Environmental Science is an area where space-time behaviour is an important focus of study. Many times several unobservable factors affect the analysis of the data so collected. This leads to spatial dependence. To include factors due to spatial dependence into the analysis, one may use Kriging or a function of the distance (Euclidean) between the points of analysis – the inverse distance weighting technique.
The inverse distance weighting is a procedure that incorporates the spatial variation in the data set y_i while computing the forecasts. The weight w_i assigned to each data diminishes with the increase in distance d_i between the data points, Hopkins et al. (1999). The equation used in this study to calculate a new data set f_i for the MSSA prediction is as follows, Awichi et al. (2013):

$$f_i = \sum_{i=1}^{n} w_i y_i \quad \text{where} \quad w_i = \frac{1/d_i}{\sum_{i=1}^{n} (1/d_i)}$$

Kriging uses spatial correlation structure – the empirical variogram as its weights. The weights are the best linear unbiased estimator of the observed spatial correlation which is always modelled mathematically, see Hopkins et al. (1999) for details.

4 Application

The data for this example is monthly recordings of temperature and rainfall at several locations in Upper Austria, for details see Mateu et al.(2012).
The one dimensional analysis for this example is the logarithm of the Linz rainfall recordings for the period January 2004 to December 2005.
The graph in Fig.1 gives a comparison of the initial and reconstructed series.

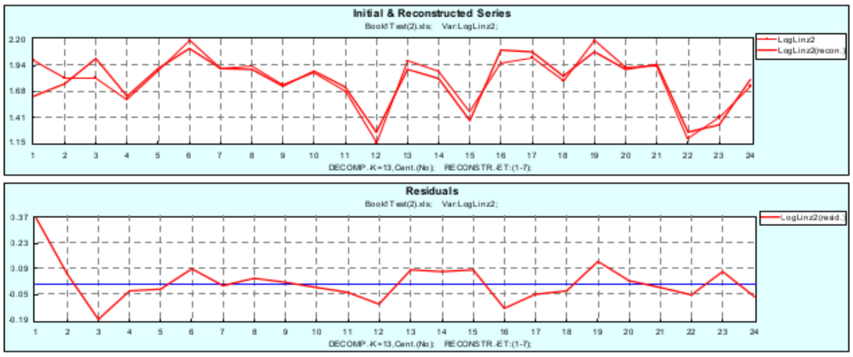

Fig.1: Initial and Reconstructed Series Compared.

For missing values in the calculation using the inverse weighting, a new weight is calculated by excluding the corresponding distance measure of that particular location from the w_is.
Fig.2 shows the comparison of the initial and reconstructed series for the 2 dimensional MSSA series.

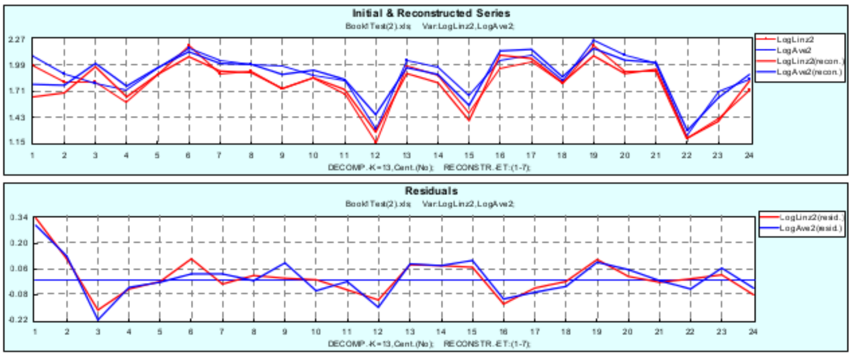

Fig 2: Initial and Reconstructed Series for Linz and Averages

5 Conclusion

This short presentation illustrates the basic capabilities of SSA in separating the components of time series without any assumptions for the time series data, i.e. the non model approach to time series analysis. The inverse distance weighting gives a slightly better prediction particularly for the in-sample forecast when compared with the actual data. A better picture is expected with a larger data set. Further studies will be done to refine the method for use with any environmental data set.

REFERENCES

Awichi R. O. and Müller, W.G. (2013). Improving SSA Predictions by Inverse Distance Weighting, *Revstat Statistics Journal*, **11(1)**, 105–119.

Broomhead, D. S. and King, G.P. (1986). Extracting Qualitative Dynamics from Experimental Data, *Physica D,* **20**, 217–236.

Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques*, Chapman & Hall; CRC, New York- London.

Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison, *Journal of Data Science*, **5**, 239–257.

Hassani, H. and Thomakos, D. (2010). A Review of Singular Spectrum Analysis for economic and financial time series, *Statistics and Its Interface*, **3**, 377–397.

Hopkins, L.P., Ensor, K.B. and Rifai, H.S. (1999). Empirical Evaluation of Ambient Ozone Interpolation Procedure to Support Exposure Models; *Journal of the Air and Waste Management Association*, **49**; 839–846.

Kapl, M. and Müller, W.G. (2010). Prediction of steel prices: A comparison between a conventional regression model and MSSA; *Statistics and Its Interface*, **3**, 369–275.

Mateu, J. and Müller, W.G. (2012). *Spatio-temporal Design: Advances in Efficient Data Acquisition;(Statistics in Practice)*, Wiley.

Patterson, K.; Hassani H., Saeed Heravi & Zhigljavsky, A. (2011). Multivariate Singular Spectrum Analysis for Forecasting Revisions to real-time data, *Journal of Applied Statistics*, **38:10**, 2183–2211.

Raynaud, S.; Yiou, P.; Kleeman, R. and Speich, S. (2005). Using MSSA to determine explicitly the oscillatory dynamics of weakly nonlinear climate systems; *Journal of Nonlinear Processes in Geophysics*, **12**, 807–815.

Zhigljavsky,A.; Hassani, H. and Saeed Heravi (2009). Forecasting European Industrial Production with Multivariate singular spectrum analysis, *International journal of Forecasting*, **25(1)**, 103–118.