

8. The approach for complexity analysis of multivariate time series

Kazimieras Pukenas

Lithuanian Sports University, Kaunas, Lithuania

E-mail: kazimieras.pukenas@lsu.lt

(Received 15 June 2015; received in revised form 1 August 2015; accepted 11 August 2015)

Abstract. This paper proposes to estimate the complexity of a multivariate time series by the spatio-temporal entropy based on multivariate singular spectrum analysis (M-SSA). In order to account for both within- and cross-component dependencies in multiple data channels the high dimensional block Toeplitz covariance matrix is decomposed as a Kronecker product of a spatial and a temporal covariance matrix and the multivariate spatio-temporal entropy is defined in terms of modulus and angle of the complex quantity constructed from the spatial and temporal components of the multivariate entropy. The benefits of the proposed approach are illustrated by simulations on complexity analysis of multivariate deterministic and stochastic processes.

Keywords: multivariate time series analysis, multivariate entropy, multivariate singular spectrum analysis, Kronecker product expansion.

1. Introduction

Spatially extended complex dynamical systems, common in physics, biology and the social sciences such as economics, contains multiple interacting components and is nonlinear, non-stationary, and noisy, and one goal of data analysis is to detect, characterize, and possibly predict any events that can significantly affect the normal functioning of the system [1]. A lot of analysis techniques, originated from synchronization theory, nonlinear dynamics, information theory, statistical physics, and from the theory of stochastic processes, is available to estimate strength and direction of interactions from time series [2]. Among other techniques, the entropy measures are a powerful tool for the analysis of time series, as they allow describing the probability distributions of the possible state of a system, and therefore are capable to map changes of spatio-temporal patterns over time and detect early signs of transition from one to another regime. Various practical definitions and calculations of entropy exist [3-6], but most of them have been designed predominantly for the scalar case and are not suited for multivariate time series that are routinely measured in experimental and biological systems. Hence, suitable tools for the detection and quantitative description of the rich variety of modes that can be excited due to the interplay between spatial and temporal behavior are needed. Multivariate multiscale entropy (MMSE) has been recently suggested as a novel measure to characterize the complexity of multivariate time series [4]. The MMSE method involves multivariate sample entropy (MSampEn) and allows to account for both within- and cross-channel dependencies in multiple data channels. Furthermore, it operates over multiple temporal scales. The MMSE method is shown to provide an assessment of the underlying dynamical richness of multivariate observations and has more degrees of freedom in the analysis than standard MSE [4]. However, this MSampEn approach has limited application when the number of variables increases. Differently to the univariate sample entropy (SampEn) introduced by Richman and Moorman [5], which operates on delay vectors of a univariate time series, MSampEn calculates, for a fixed embedding dimension m , the average number of composite delay vector pairs that are within a fixed threshold r and repeats the same calculation for the dimension $(m + 1)$ [4]. Because the composite delay vectors are high-dimensional (the dimension is equal to the product of the number of variables and the embedding dimension of the reconstructed phase space), the probability that at least two composite delay vectors will match with respect to tolerance r quickly decreases as the number of variables increases. Among the other multivariate entropy approaches described is Multivariate Principal Subspace Entropy (MPSE) [6], introduced as a measure to estimate spatio-temporal

complexity of functional magnetic resonance imaging (fMRI) time series. MPSE is based on the multivariate Gaussian density assumption of the fMRI data. Multivariate singular spectrum analysis (M-SSA) [7-10] provides insight into the dynamics of the underlying system by decomposing the delay-coordinate phase space of a given multivariate time series into a set of data-adaptive orthonormal components and is a powerful tool for spatiotemporal analysis of the vector-valued processes. M-SSA combines two useful approaches of statistical analysis: (1) it determines – with the help of principal component analysis – major directions in the system’s phase space that are populated by the multivariate time series; and (2) it extracts major spectral components by means of singular spectrum analysis (SSA). However, the Shannon entropy constructed from the M-SSA eigenvalue spectrum has weak distinguishing ability and does not provide information about the degree of complexity for both temporal and spatial components of multivariate time series.

In the present paper, it is shown that high-dimensional block Toeplitz covariance matrices (on which M-SSA operates) estimation via Kronecker product expansions makes it possible to define the complexity of multivariate time series by means of the spatial and temporal components and to obtain more information about the underlying process.

2. Description of the algorithm

Let $\mathbf{x} = \{x_d(n): d = 1, \dots, D, n = 1, \dots, N\}$ be a multivariate time series with D components (variables) of length N . In practical implementation (neuroimaging with EEG, geosciences), this denotes samples of a space–time random process defined over a D -grid of space samples and an N -grid of time samples. It is assumed that each component has been centred and normalized. Each time series component of the multivariate system can be reconstructed in M -dimensional phase space by selecting the embedding dimension M and time delay τ . Each phase point in the phase space is thus defined by [11]:

$$\mathbf{x}_d(n) = [x_d(n), x_d(n + \tau), \dots, x_d(n + (M - 1)\tau)]^T, \tag{1}$$

where $n = 1, 2, \dots, N - (M - 1)\tau$, and $(\cdot)^T$ denotes the transpose of a real matrix. At $\tau = 1$, the reconstructed phase space matrix \mathbf{X}_d with M rows and $L = N - M + 1$ columns (called a trajectory matrix) is defined by:

$$\mathbf{X}_d = \begin{bmatrix} x_d(1) & \cdots & x_d(N - M + 1) \\ \vdots & \ddots & \vdots \\ x_d(M) & \cdots & x_d(N) \end{bmatrix}, \tag{2}$$

and encompasses M delayed versions of each component. The total trajectory matrix of the multivariate system \mathbf{X} will be a concatenation of the component trajectory matrices \mathbf{X}_d computed for each component, that is, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_D]^T$, and has DM rows of length $L = N - M + 1$. The classical M-SSA algorithm [8, 9] then computes the covariance matrix $\mathbf{\Sigma} = 1/L \cdot \mathbf{X}\mathbf{X}^T \in \mathbb{R}^{DM \times DM}$ of \mathbf{X} and its eigendecomposition. The total trajectory matrix \mathbf{X} can be rearranged such that the delayed versions of all D components follow one another, that is, $\mathbf{X} = [\mathbf{X}_{D1}, \mathbf{X}_{D2}, \dots, \mathbf{X}_{DM}]^T$. In this case, and for $L \gg M$, the covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{MD \times MD}$ takes the block Toeplitz form:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_0 & \mathbf{\Sigma}_1 & \cdots & \mathbf{\Sigma}_{M-2} & \mathbf{\Sigma}_{M-1} \\ \mathbf{\Sigma}_{(-1)} & \mathbf{\Sigma}_0 & \cdots & \mathbf{\Sigma}_{M-3} & \mathbf{\Sigma}_{M-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{\Sigma}_{(-(M-2))} & \mathbf{\Sigma}_{(-(M-3))} & \cdots & \mathbf{\Sigma}_0 & \mathbf{\Sigma}_2 \\ \mathbf{\Sigma}_{(-(M-1))} & \mathbf{\Sigma}_{(-(M-2))} & \cdots & \mathbf{\Sigma}_{(-1)} & \mathbf{\Sigma}_0 \end{bmatrix}, \tag{3}$$

with equal block submatrices $\Sigma_Q \in \mathbb{R}^{D \times D}$ along each (top left to lower right) diagonal. The covariance matrix Σ combines all auto- as well as cross-covariances up to a time lag equal to $M - 1$. The spatio-temporal covariance matrix $\Sigma \in \mathbb{R}^{MD \times MD}$ can be represented as a Kronecker product (KP) [12-16]:

$$\Sigma = \sum_{i=1}^r \mathbf{T}_i \otimes \mathbf{S}_i, \quad (4)$$

for some separation rank r , some sequence of $\mathbf{T}_i \in \mathbb{R}^{M \times M}$ and $\mathbf{S}_i \in \mathbb{R}^{D \times D}$. The KP parametrization assumes that an arbitrary spatio-temporal correlation can be modelled as a product of a spatial and a temporal factor. These two factors are independent of each other; hence, the spatial and temporal correlations are separated from each other in the KP model. The spatial covariance matrix captures the spatial correlations between the components of multivariate time series (in real-world application, signals registered at different sensors) and the temporal covariance matrix captures the temporal correlations as lag dependent. The meaning of Eq. (4) is that the temporal covariance matrix \mathbf{T}_i is fixed in space and the spatial covariance matrix \mathbf{S}_i does not vary over time. In other words, space and time are not correlated. In order to effectively detect the structure of the spatio-temporal dependencies in high-dimensional multivariate systems, we use the full rank approximation model, i.e., define $r = M$. The solution to the nearest Kronecker product for a space-time covariance matrix (NKPST) problem is given by the singular value decomposition (SVD) of a permuted version of the space-time covariance matrix $\Sigma \in \mathbb{R}^{MD \times MD}$ in the following:

1) The matrix $\Sigma \in \mathbb{R}^{MD \times MD}$ is rearranged into another matrix $\Xi(\Sigma) \in \mathbb{R}^{M^2 \times D^2}$ such that the sum of squares that arise in $\|\Sigma - \sum_{i=1}^r \mathbf{T}_i \otimes \mathbf{S}_i\|_F$ is exactly the same as the sum of squares that arise in $\|\Xi(\Sigma) - \sum_{i=1}^r \text{vec}(\mathbf{T}_i) \otimes \text{vec}(\mathbf{S}_i)^T\|_F$, where $\text{vec}(\mathbf{W})$ denotes the vectorized form of the square matrix \mathbf{W} (concatenation of columns of \mathbf{W} into a column vector). The M^2 rows of Ξ are defined as the transpose of the vectorized $D \times D$ block submatrices of Σ , where each block submatrix is $\Sigma_Q(i, j) = \Sigma_{(i-1)D+1:iD, (j-1)D+1:jD}$, $i, j = 1, \dots, M$. That is, the permutation operator $\mathcal{R}: \mathbb{R}^{MD \times MD} \rightarrow \mathbb{R}^{M^2 \times D^2}$ is defined by setting the $((i-1)M + j)$ th row of Ξ equal to $\text{vec}(\Sigma_Q(i, j))^T$, $i, j = 1, \dots, M$.

2) The SVD is performed on the matrix Ξ :

$$\Xi = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T = \sum_{i=1}^r \mathbf{u}_i \lambda_i \mathbf{v}_i^T, \quad (5)$$

and the matrices \mathbf{T}_i and \mathbf{S}_i are defined for the selected Kronecker product (full rank approximation) model by $\text{vec}(\mathbf{T}_i) = \lambda_i \mathbf{u}_i$ and $\text{vec}(\mathbf{S}_i) = \lambda_i \mathbf{v}_i$, where \mathbf{u}_i and \mathbf{v}_i are left and right singular vectors corresponding to the singular value λ_i [13, 16]. This is converted back to a square $\mathbf{T}_i \in \mathbb{R}^{M \times M}$ and the $\mathbf{S}_i \in \mathbb{R}^{D \times D}$ matrix by applying the inverse permutation operator.

Finally, the SVD is performed on the estimated matrices \mathbf{T}_i and \mathbf{S}_i , and the spatial and temporal components of the multivariate entropy (MVE) can be defined as the Shannon entropy of the corresponding normalized eigenvalue spectrum:

$$E_S = -\frac{1}{\ln D} \sum_{j=1}^D \sigma_{S,j} \ln \sigma_{S,j}, \quad (6)$$

$$E_T = -\frac{1}{\ln M} \sum_{j=1}^M \sigma_{T,j} \ln \sigma_{T,j}, \quad (7)$$

where $\sigma_{S,j} = \lambda_{S,j} / \sum \lambda_{S,j}$ and $\sigma_{T,j} = \lambda_{T,j} / \sum \lambda_{T,j}$, are the summarized $\lambda_{S,j} = \sum_{i=1}^r \lambda_{S,j}^{(i)}$, $\lambda_{T,j} = \sum_{i=1}^r \lambda_{T,j}^{(i)}$ and normalized eigenvalues of the decomposed matrices \mathbf{S}_i and \mathbf{T}_i respectively. The Shannon entropy quantifies the degree of uniformity in the distribution of the eigenvalues: it tends to unity at uniform distribution of the eigenvalues and tends to zero if the eigenvalue spectrum is concentrated into a single first largest eigenvalue. The entropy features E_S and E_T can be expressed in complex form $E_\Sigma = E_S + iE_T$, and the multivariate spatio-temporal entropy can be defined by modulus $|E_\Sigma|$ and angle φ_E of this complex quantity. The angle φ_E reflects the difference between E_S and E_T and ranges from 0 (at $E_T = 0$) to $\pi/2$ (at $E_S = 0$).

3. Simulation results

To validate the ability of the proposed algorithm to reveal both within- and cross-component properties in multivariate data, we generated a multivariate time series with defined changes of the complexity. The overembedding ($M = 16$) for all trials was chosen in order to capture the temporal properties of the time series. All data processing and analyses were performed using Matlab software (MathWorks, Natick, MA).

First, to illustrate the ability of the proposed algorithm to model cross-component properties in multivariate data, we considered a multivariate deterministic system – a chain of diffusively coupled Rössler oscillators [9, 17]:

$$\begin{aligned} \dot{x}_j &= -\omega_j y_j - z_j, \\ \dot{y}_j &= \omega_j x_j + \alpha y_j + c(y_{j+1} - 2y_j + y_{j-1}), \\ \dot{z}_j &= 0.1 + z_j(x_j - 8.5). \end{aligned} \quad (8)$$

The position in the chain is given by the index $j = 1, \dots, J$; $\omega_j = \omega_1 + 0.02(j - 1)$ are the associated natural frequencies, with $\omega_1 = 1$, and we assume that free boundary conditions $x_0(n) = x_1(n)$ and $x_{J+1}(n) = x_J(n)$ exist. The frequency mismatch $\Delta_{jk}\omega$ between oscillators j and k is given by $\Delta_{jk}\omega = 0.02|j - k|$. The parameter $\alpha = 0.15$, and $c \geq 0$ is the coupling strength. We considered a chain of $J = 10$ Rössler oscillators and generated a time series using the ODE45 integrator of Matlab. The solution was sampled at time intervals $t = 0.1$, and the proposed algorithm was applied to x -components of length $N = 2048$ of the Rössler systems. Fig. 1(a) shows the $|E_\Sigma|$ and φ_E curves when the coupling strength c is varied. As c increases, the oscillators form synchronous clusters within which the observed frequency is the same [9]. With a further increase of c the number of clusters decreases as clusters merge, and at sufficiently high c all oscillators are synchronized, i.e., all the mean frequencies coincide, although the amplitudes of the oscillators remain chaotic [17]. MVE is characterized by two components – modulus $|E_\Sigma|$ and angle φ_E – that provide comprehensive information about the inherent structure of the multivariate data. The increasing value of φ_E indicates that $|E_\Sigma|$ decreases due to the growing strength of dependence between components of a multivariate time series.

To illustrate the ability of the proposed algorithm to model within-component properties in multivariate data, we considered a chain of above-mentioned Rössler oscillators at $c = 0$ and gradually increased the parameter α from 0.10 to 0.35. The larger α , the more complex the resulting time series will be. The results are displayed in Fig. 1(b). In this case, the increased value of φ_E indicates that $|E_\Sigma|$ slightly increases because of the degree of randomness of a time series.

To demonstrate the performance of the proposed algorithm for quantifying changes in the dynamics underlying the experimental data set, we have applied the algorithm to multi-channel EEG recordings from epilepsy patients. It is well known that during epileptic seizures EEG time series display a marked change of the dynamics (mainly oscillations of comparatively large amplitude and well defined frequency). The EEG data with the clinically marked start and end of seizure are taken from the CHB-MIT Scalp EEG Database

(<http://www.physionet.org/physiobank/database/chbmit/>). We analyzed 23 contacts of one-minute raw (unfiltered) scalp EEG recordings (from record, numbered chb01/chb01_04.edf) using the block Toeplitz covariance matrix, which was constructed over a moving window of length 3.9 s, corresponding to 1000 data points. This window was shifted with 90 % overlap over the whole EEG recording, and the time evolution of the $|E_{\Sigma}|$ and φ_E was computed. As we can see in Fig. 2, both $|E_{\Sigma}|$ and φ_E indicate sudden changes of the brain dynamics in preictal and ictal states. Furthermore, the angle φ_E of the multivariate entropy provides essential information about factors that cause this complicated dynamics. The considerable positive deviation of φ_E from its mean value denotes that $|E_{\Sigma}|$ changes due to the growing strength of dependence between components of a multivariate time series, and also because of the degree of randomness of a time series.

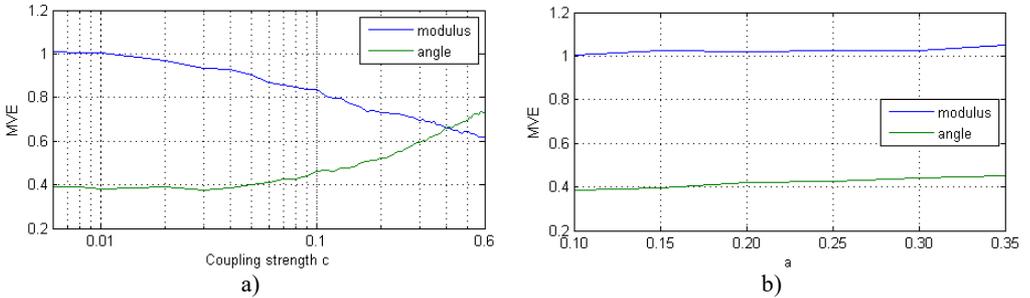


Fig. 1. Multivariate entropy (modulus $|E_{\Sigma}|$ and angle φ_E) for: a) a chain of coupled Rössler oscillators (x components) at various coupling strengths c ; b) a chain of Rössler oscillators (x components) at $c = 0$ and various values of parameter α

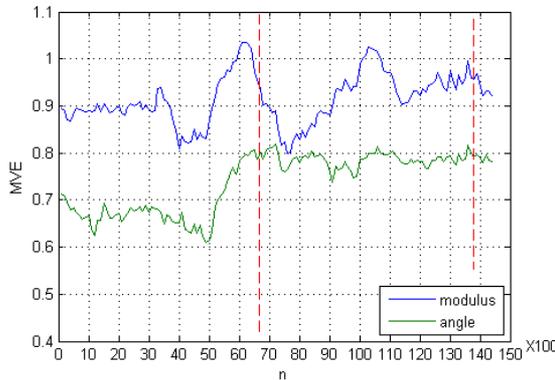


Fig. 2. Multivariate entropy (modulus $|E_{\Sigma}|$ and angle φ_E) applied on epileptic EEG. The period between two vertical dashed lines is visually identified by the neurologists as the seizure period

4. Conclusion

A simulation involving a multivariate deterministic and stochastic time series shows that the approach based on a full rank Kronecker product expansion of the block-Toeplitz covariance matrix, calculated from the full augmented trajectory matrix, can reveal temporal dynamics and relationship between the signals; furthermore, it is a suitable tool for the assessment of the underlying dynamic complexity of multivariate observations. The proposed approach defined in terms of modulus and angle of the complex quantity constructed from the spatial and temporal components of the multivariate entropy reflects both the integrative properties of the systems and the ratio between the complexities in the spatial and temporal directions. The algorithm can be applied to the estimation of the complexity of multivariate time series with a large number (tens) of variates (e.g. EEG), when the application of the other approaches becomes problematic.

References

- [1] **Seely A. J. E., Maclem P. T.** Complex systems and the technology of variability analysis. *Critical Care*, Vol. 8, 2004, p. 367-384.
- [2] **Lehnertz K., Ansmanna G., Bialonski S., Dickten H., Geier Ch., Porz S.** Evolving networks in the human epileptic brain. *Physica D*, Vol. 267, 2014, p. 7-15.
- [3] **Bollt E. M., Skufca J. D., McGregor S. J.** Control entropy: a complexity measure for nonstationary signals. *Mathematical Biosciences and Engineering*, Vol. 6, Issue 1, 2009, p. 1-25.
- [4] **Ahmed M. U., Mandic D. P.** Multivariate multiscale entropy: a tool for complexity analysis of multichannel data. *Physical Review E*, Vol. 84, 2011, p. 061918.
- [5] **Richman J. S., Moorman J. R.** Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology Heart and Circulatory Physiology*, Vol. 278, 2000, p. H2039-H2049.
- [6] **Schütze H., Martinetz T., Anders S., Mamlouk A. M.** A multivariate approach to estimate complexity of fMRI time series. *Lecture Notes in Computer Science*, Vol. 7553, 2012, p. 540-547.
- [7] **Plaut G., Vautard R.** Spells of low-frequency oscillations and weather regimes in the northern hemisphere. *Journal of the Atmospheric Sciences*, Vol. 51, Issue 2, 1994, p. 210-236.
- [8] **Ghil M., Allen M. R., Dettinger M. D., Ide K., Kondrashov D., Mann M. E., et al.** Advanced spectral methods for climatic time series. *Reviews of Geophysics*, Vol. 40, 2002, p. 1-41.
- [9] **Groth A., Ghil M.** Multivariate singular spectrum analysis and the road to phase synchronization. *Physical Review E*, Vol. 84, 2011, p. 036206.
- [10] **Golyandina N., Korobeynikov A., Shlemov A., Usevich K.** Multivariate and 2D extensions of singular spectrum analysis with the Rssa package. *arXiv:1309.5050*, 2013, p. 1-74.
- [11] **Kantz H., Schreiber T.** *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge, 2003.
- [12] **Loan C. V., Pitsianis N.** Approximation with Kronecker Products, in *Linear Algebra for Large Scale and Real Time Applications*. Kluwer Publications, Amsterdam, 1993, p. 293-314.
- [13] **Genton M. G.** Separable approximations of space-time covariance matrices. *Environmetrics*, Vol. 18, 2007, p. 681-695.
- [14] **Tsiligkaridis T., Hero A. O.** Covariance estimation in high dimensions via Kronecker product expansions. *IEEE Transactions on Signal Processing*, Vol. 61, Issue 21, 2013, p. 5347-5360.
- [15] **Wirfält P., Jansson M.** On Kronecker and linearly structured covariance matrix estimation. *IEEE Transactions on Signal Processing*, Vol. 62, Issue 6, 2014, p. 1536-1547.
- [16] **Bijma F., de Munck J. C., Heethaar R. M.** The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products. *NeuroImage*, Vol. 27, 2005, p. 402-415.
- [17] **Osipov G. V., Pikovsky A. S., Rosenblum M. G., Kurths J.** Phase synchronization effects in a lattice of nonidentical Rössler oscillators. *Physical Review E*, Vol. 55, 1997, p. 2353-2361.