A new approach to identifying the local structure of multidimensional chaotic time series

A V Makshanov, A E Zhuravlev and L N Tyndykar
Admiral Makarov State University of Maritime and Inland Shipping, 5/7, Dvinskaya str., 198035, Saint-Petersburg, Russia
E-mail: zhuravlevae@gumrf.ru

Abstract. The paper is devoted to the solution of the problems of mathematical supply of decision making during multichannel monitoring of large-scaled systems. The work also deals with space-time dynamics of multidimensional time series of different origins. Highly dynamical chaotic processes whose fine structure cannot be revealed by standard spectral methods are regarded. Technologies for dimension reduction based on data matrix representation on the first singular basis and multiple regression in projections’ space are developed.

1. Introduction
Let data matrix \(X_{ij}, i=1,\ldots,N, j=1,\ldots,m\) represent a stretch of \(m\)-dimensional chaotic discrete time series at a relative scale, so \(x_{ij}\) is out of scale random values. Chaotic nature implies fast local changes; therefore, methods based on averaging over extended sliding windows such as parametric or non-parametric spectra analysis are hardly to be used. Such series evidently contains local trends and alternations in the correlation structure, specifically turbulence stretches and segments with periodic constituents [2, 12, 22].

2. Problem statement
The following problems are to be regarded:
• prediction for a possibly large number of steps;
• revealing segments of an abnormal structure;
• exposing segments containing local periodics.

The basic approach here is founded on stationary stiffness assumption of \(X\)-series itself or of series of its finite differences [2, 3, 4, 10]. That makes it possible to achieve correct estimates of components’ averages and the covariance matrix \(\hat{\Sigma} (m \times m)\). Projections of \(X\) columns onto latent vectors of \(\hat{\Sigma}\) form \(m\) uncorrelated one-dimensional time series, whose contributions to initial series are estimated by their covariances. These series are treated by standard methods for one-dimensional series [4, 13]. If the assumption of stationary stiffness is invalid, the analysis is carried out in the system of sliding windows \(L\) steps wide that supposes the window width and the mutual recovering of neighboring windows.

Special interest represents the case of data of high dimension \(m\) and high dynamics where the window cannot be taken with a sufficient width to supply correct estimates of the covariance matrix. In this case, special methods of dimensionality reduction not founded on spectra’s estimates of covariances are available [4, 24].
For example, Fig.1 represents data on varying 5 dynamic parameters of the turbulent gas flow. Measurements are represented by the $X$ matrix $<N \times m>$, $m=5$, $N=144000$.

**Figure 1.** Data matrix $X$ with 5 components, 144000 time readings

To reveal possible fine structures, let us compare the average $\hat{x}$ and the covariance matrix $\hat{\Sigma}$, being estimated by the series of first finite differences of the whole series $X$ as a base and similar characteristics $\hat{y}(k)$, $\hat{S}(k)$. They were obtained in sliding windows of width $L$ with a successive shift on $d$ steps ($k$ is the number of the successive sliding window). The comparison is to be based on standard MANOVA statistics [4, 13]:

- Wilks’ statistics: $W(k) = \log \left( \frac{\hat{S}(k)}{\hat{\Sigma}} \right)$;
- Hotelling’s statistics: $H(k) = \text{trace}(\hat{S}(k) \cdot \hat{\Sigma}^{-1})$;
- Mahalanobis’ distance: $M(k) = \text{trace}(\hat{\Sigma}^{-1} \cdot (\hat{x} - \hat{y}(k))^T \cdot (\hat{x} - \hat{y}(k)))$;
- Kulbac-Leibler divergence: $J(k) = \{W(k) + [H(k) - m] + M(k)\}$.

The dynamics of these statistics in the sliding window of width $L=200$ steps with a subsequent shift of $d=200$ steps is displayed in Fig. 2a (the figure shows first 40000 readings, 400 windows with shift $d$). Fig. 2b shows similar results, but for the base, one of the windows $Y$ (No. 401) is taken. It makes it possible to single out the segments that are at some sense close to $Y$ and get predictions “by analogy” (precedent analysis [3, 4, 10, 11]).
The short-time predictors may be based on multiple regressions techniques [8,10,11]. Let $Y$ be the sliding window of width $L$ – the $L \times m$ ($m=5$) matrix. A column of learning data $y$ expands it – they are readings of the component to be predicted with a given lag $r$ – the depth of the prediction. The expanded $L \times (m+1)$ matrix supplies the block average vector $\hat{a}$ and the block covariance matrix $\hat{\Sigma}$:

$$
\hat{a} = \begin{bmatrix} \hat{a}_x \\ \hat{a}_y \end{bmatrix}, \quad \hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_{xx} & \hat{\Sigma}_{xy} \\ \hat{\Sigma}_{yx} & \hat{\Sigma}_{yy} \end{bmatrix}.
$$

The prediction $\hat{Y}$ has the form:

$$
\hat{Y} = \hat{a}_y + \hat{\Sigma}_{xy} \hat{\Sigma}_{yy}^{-1} (X - \hat{a}_x),
$$

where $X$ is measurements at the least available data stretch. The STD of the prediction has to be regarded with respect to local STD of the least data segment.

While scanning the entire sample (Fig. 4, 144000 readings, $L=200$, $d=100$), it appears that the STD of the prediction monotonically increases with $r$, whereas bias firstly decreases and then begins to increase after $r=8$. It means that for underlying data, the approach with regard to the prediction horizon [10] is confined to about 7 steps. In contrast, standard polynomial predictors [8] show similar properties for only 1-2 steps.

The principal shortage of similar methods for chaotic multidimensional time series is that they require estimation in sufficiently representative windows that reduce a possibility of exposing fine peculiarities that show themselves only at smaller segments.
3. New mathematical technologies

At the end of the XX century, there were investigations on the analysis based on singular decompositions of data matrices going back to paper [15]. Nowadays they are regarded as approaches to the global concept of evolutionary programming [4, 5, 6, 14, 16, 23].
Let initial data matrix \( X = [x_{ij}] \), \( <n \times m> \) have the rank \( p \leq m \). The problem is to get its approximation by the matrix \( Y \) of a lower rank \( k < p \). Therefore, we have to find the \( <n \times m> \) matrix \( Y \) as a solution of the optimization problem:

\[
\sum_{i,j} (x_{ij} - y_{ij})^2 \rightarrow \min
\]

under restriction \( \text{rank}(Y) = k < \min(n,p) \). For data analysis it means that we have to explain the underlying structure of \( m \)-dimensional data by a smaller number \( k \) of generalized indices, \( k < p < m \).

An arbitrary real-valued matrix \( X <n \times m>, n>m \) may be represented in the form of a singular decomposition (SVD):

\[
X = L * S * R^T,
\]

where

- \( S = \text{diag}(s_1,...,s_n) \) is diagonal; numbers \( s_1 \geq s_2 \geq ... \geq s_n \geq 0 \) are singular values of \( X \);
- \( L \) is the \( <n \times n> \) matrix; its columns \( L_1,...,L_n \) are orthogonal unit vectors named as left singular vectors of \( X \); \( L^T L = L L^T = E \);
- \( R \) is an \( <m \times m> \) matrix; its columns \( R_1,...,R_m \) are orthogonal unit vectors named as right singular vectors of \( X \); \( R^T R = R R^T = E \).

Such form of (3) is referred to as a “short” one. Orthogonality here has to be understood in Euclidean sense, data remain statistically correlated. If the rang of \( X \) \( \text{rank}(X) = p < m \), only \( p \) of singular values is non-nulls.

Decomposition (3) may be rewritten as a sum of elementary matrices of a unit rang:

\[
X = \sum_{i=1}^{p} s_i L_i R_i^T = s_1 L_1 R_1^T + ... + s_p L_p R_p^T.
\]

**Theorem** (Eckart-Young [4, 15]). The solution of the extreme problem (1) is given by the sum of \( k \) initial summands if (3):

\[
X \equiv Y = \sum_{i=1}^{k} s_i L_i R_i^T = s_1 L_1 R_1^T + ... + s_k L_k R_k^T.
\]

Particularly with \( k = 1 \) the solution is based on the first (maximal) singular value with corresponding singular vectors:

\[
X \equiv Y = s_1 L_1 R_1^T.
\]

If \( X \) is a data matrix, it is converted here into a sum of the reduced number (\( k < m \)) of layers having the same dimensions, but of a simplified structure: every layer is a \( n \times m \) matrix of a unit rang.

The singular decomposition is robust to small perturbations of \( X \), so it is a well-conditioned procedure opposite to spectral expansions forming the base of multidimensional statistical analysis (MANOVA).

*Nowadays such approach has five basic lines of development.*

*When solving problems of recognition, classifying and clustering projections of data onto spaces that are generated by several singular components are used. These projections particularly form specific pseudo metrics for solving similar problems [4, 9, 23].*

*This way is shown to be plausible in situation analysis: the observed situation \( x_0 \) is associated with one of virtual situations \( x_1,\ldots,x_k \) that turns to be the nearest with respect to some pseudo metrics [18-20, 23].*

*In interpolating random fields, the value \( f(x_0) \) is estimated as a result of linear interpolation at \( k \) nearest points \( x_1,\ldots,x_k \):

\[
\hat{f}(x_0) = c_1 f(x_1) + ... + c_k f(x_k),
\]

where

\[
 c_j = 1 / (1 + d_j \sum_{i \neq j} d_i),
\]

\( d_j = (x_j - x_0)^2 \).
The point here is that the proximity measure \( d_j \) between \( x_0 \) and \( x_j \) is determined as Euclidean metrics in one of projection spaces.

- If a segment of \( m \)-dimensional time series is presented in the form of \( n \times m \) matrix, it may be approximated as a sum of elementary matrices of a unit rang. Any column of those summands produces new one-dimensional series and all of them may be treated separately or in some combinations. It permits to reduce and/or simplify treating a number of multidimensional problems.

- For one-dimensional series, a group of methods based on the deposition of the series into multidimensional space with analyzing the singular decomposition of the output Henkel matrix is proposed. One of most known methods here is Singular Spectrum Analysis (SSA) [5,6,14,16]. Such approaches are intended to single out trends and periodics, to reveal different kinds of dissensions, etc.

Based on those approaches, new algorithms of identification of the local structure of multidimensional time series are to be elaborated.

Now let us construct for a \( <n \times m>, n \times m \) data matrix \( X \) of the singular representation (3) and examine the dispersion criterion:

\[
D(k) = \frac{1}{\sum_{j=1}^{m} s_j} \sum_{j=1}^{m} s_j .
\]

For the example under consideration with \( k=3 \), \( D=0.98 \). Therefore, it is possible to be bounded by only three summands (projections) with negligible small loss of information. These summands are:

\[
X(i) = L_s R_i^t, \quad i=1,\ldots,k, \quad k=3.
\]

Every summand (projection) looks like the \( n \times m \) matrix of rank one, so it is sufficient to extract out of \( X(i) \) only the first (for example) column \( x_1(i) \) and a string of recount coefficients:

\[
C_j(i) = \frac{\sum_{j=1}^{m} x_j(i)}{\sum_{j=1}^{m} x_j(1)}, \quad j=1,\ldots,m.
\]

That will be necessary in restoring the estimate of the \( i \)-th layer and returning to originate coordinates. Now let us form the projections’ matrix \( Z=[x_1(1),\ldots,x_1(k)] \) (Fig.5) and build for the \( k \)-component \( (k=3) \) segment \( Z \) predictions of different widths \( r, r=1,2,\ldots, R \) using a multiple regression approach (1). We return to original variables on the base of reserved recount coefficients (5).

**Figure 5.** First three singular components – columns of the matrix \( Z \) for one data segment \( (L=200) \)
4. Results
The advantage of such approach is the possibility to reduce the length of the sliding window $L$. For $k=5$ correct estimate, the mean and covariance matrix (15 parameters) require about 200 measurements. When $k=3$ (9 parameters), the number of 100 measurements is sufficient. With a larger dimension $m$, the gain may be rather substantial. Fig. 6 shows errors with respect to the prediction width $r$ after averaging over all the length of $X$ and over all components with $L=200, d=100$ (in contrast to Fig. 4).

![Prediction errors](image1)

**Figure 6.** Averaged errors depending on the required prediction depth $r$ ($L=200, d=100$)

![Prediction compared to real process](image2)

**Figure 7.** An example of return to origin variables using estimated layers restored by recount coefficients (5). Prediction of component 1 to 10 steps ahead.
Investigation of a large number of predictions in different windows and their averages over $X$ show that dimension reduction ($k=3$) allows one to reduce the width of the sliding window about 2 times; the prediction horizon enlarges to about 2.2 times.

5. Conclusion
The class of chaotic multidimensional time series is highly diverse. Therefore, the proposed algorithms have to be regarded only by a way of a possibly useful instrument set. There are many settings where this set of tools may considerably reduce the dimension and, due to this, extend the prediction horizon or be able to scrutinize fine local details. The perspective is supposed to be in diverse combinations of singular analysis for reducing the dimension and more classic MANOVA methods for better understanding the data peculiarities.

References
[1] Hsieh I-C, Huang Y 2021 Sensitivity analysis and visualization for functional data Journal of Statistical Computation and Simulation 1–23. doi:10.1080/00949655.2020.1863405
[2] Lahmiri S, Bekiros S 2020 Nonlinear analysis of Casablanca Stock Exchange, Dow Jones and S&P500 industrial sectors with a comparison Physica A: Statistical Mechanics and Its Applications 539 122923. doi:10.1016/j.physa.2019.122923
[3] Sitarachu K, Zia R K P, Bachmann M 2020 Exact microcanonical statistical analysis of transition behavior in using chains and strips Journal of Statistical Mechanics: Theory and Experiment 2020(7) 073204. doi:10.1088/1742-5468/ab97bc
[4] Makshanov, Musaev 2014 Diachronic analysis of nonstationary random processes SPIIRAS Proceedings 1(1) 360. doi:10.15622/sp.1.26
[5] Man A X, Culpepper S A 2020 A Mode-Jumping Algorithm for Bayesian Factor Analysis Journal of the American Statistical Association 1–14. doi:10.1080/01621459.2020.1773833
[6] Xu C, Yang M, Zhang J 2020 Fast deflation sparse principal component analysis via subspace projections Journal of Statistical Computation and Simulation 90(8) 1399–1412. doi:10.1080/00949655.2020.1728761
[7] Makshanov, Sherstuk 2014 Automatic devices with stack memory application for evaluation combinations tasks SPIIRAS Proceedings 1(1) 282. doi:10.15622/sp.1.20
[8] Zhang Y, Zuo F, Guan X 2020 Integrating case-based analysis and fuzzy optimization for selecting project risk response actions Physica A: Statistical Mechanics and Its Applications 545 123578. doi:10.1016/j.physa.2019.123578
[9] Ermolaev V, Kozlovskiy S, Makshanov A 2009 IGIS Capabilities Application to Controlling Polystatic Detection Systems Ensuring Security of Marine Economic Activity Information Fusion and Geographic Information Systems 265–276. doi:10.1007/978-3-642-00304-2_18
[10] Ma C, Lu J, Liu H 2020 Inter-Subject Analysis: A Partial Gaussian Graphical Model Approach Journal of the American Statistical Association 1–57. doi:10.1080/01621459.2020.1841645
[11] Makshanov A, Zhuravlev A, Tyndykar L 2020 Elaboration of Multichannel Data Fusion Algorithms at Marine Monitoring Systems Advances in Intelligent Systems and Computing 909–923. doi:10.1007/978-3-030-37919-3_90
[12] Wang J, Hu J, Shen S, Zhuang J, Ni S 2020 Crime risk analysis through big data algorithm with urban metrics Physica A: Statistical Mechanics and Its Applications 545 123627. doi:10.1016/j.physa.2019.123627
[13] Mamunts D G, Marley V E, Kulakov L S, Pastushok E M, Makshanov A V 2018 The use of the authentication technology blockchain platform for the marine industry 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ECONRus). doi:10.1109/eiconrus.2018.8317032
[14] Wang Y, Shang P 2020 Complexity analysis of time series based on generalized fractional order cumulative residual distribution entropy Physica A: Statistical Mechanics and Its Applications 537 122582. doi:10.1016/j.physa.2019.122582
[15] Katorin Y F, Makshanov A V, Danilin G V, Yemelyanov V A, Ovcharenko I K 2020 Improving the QoS Multiservice Networks: New Methods, Impact on the Security of Transmitted Data 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). doi:10.1109/eiconrus49466.2020.9039503

[16] Liu H, Zhang X, Zhang X 2020 Multiscale complexity analysis on airport air traffic flow volume time series Physica A: Statistical Mechanics and Its Applications 548 124485. doi:10.1016/j.physa.2020.124485

[17] Ishihara M 2020 Derivation of density operators for generalized entropies with quantum analysis Physica A: Statistical Mechanics and Its Applications 543 123419. doi:10.1016/j.physa.2019.123419

[18] Ramesh G K 2020 Analysis of active and passive control of nanoparticles in viscoelastic nanomaterial inspired by activation energy and chemical reaction Physica A: Statistical Mechanics and Its Applications 550 123964. doi:10.1016/j.physa.2019.123964

[19] Lin Z, Flournoy N, Rosenberger W F 2020 Random norming aids analysis of non-linear regression models with sequential informative dose selection Journal of Statistical Planning and Inference 206 29–42. doi: http://dx.doi.org/10.1016/j.jspi.2019.09.003.

[20] Song K, Shi J, Yi X 2020 A time-discrete and zero-adjusted gamma process model with application to degradation analysis Physica A: Statistical Mechanics and Its Applications 560 125180. doi:10.1016/j.physa.2020.125180

[21] Ma Z, Chen G 2020 Bayesian joint analysis using a semiparametric latent variable model with non-ignorable missing covariates for CHNS data Statistical Modelling 1471082X1989668. doi:10.1177/1471082x19896688

[22] Tsompanaki E, Caroni C 2020 Local influence analysis of the 2PL IRT model for binary responses Journal of Statistical Computation and Simulation 1–23. doi:10.1080/00949655.2020.1858298

[23] Chen W 2020 Statistical Analysis of Coastal Port Competitiveness Factors Based on SEM Model Journal of Coastal Research 103(sp1) 190. doi:10.2112/si103-041.1

[24] Patone M, Zhang L 2020 On Two Existing Approaches to Statistical Analysis of Social Media Data International Statistical Review. doi:10.1111/insr.12404