Research Article
Special Issue on Econometrics and Business Analytics

Andrey Makshanov, Alexander Musaev, and Dmitry Grigoriev*

Analyzing and forecasting financial series with singular spectral analysis

https://doi.org/10.1515/demo-2022-0112
received August 26, 2021; accepted April 22, 2022

Abstract: Modern techniques for managing multidimensional stochastic processes that reflect the dynamics of unstable environments are proactive, which refers to decision making based on forecasting the system’s state vector evolution. At the same time, the dynamics of open nonlinear systems are largely determined by their chaotic nature, which leads to a violation of stationarity and ergodicity of the series of observations and, as a result, to a catastrophic decrease in the efficiency of forecasting algorithms based on traditional methods of multivariate statistical data analysis. In this article, we make an attempt to reduce the instability influence by employing singular spectrum analysis (SSA) algorithms. This technique has been employed in a wide class of applied data analysis problems formulated in terms of singular decomposition of data matrices: technologies of immunocomputing and SSA.

Keywords: multidimensional chaotic processes, forecasting, singular spectrum analysis, immunocomputing, Forex

MSC 2020: 37M20, 37M10, 90C90

1 Introduction

One of the important problems of making management decisions based on monitoring is the low predictability of the resulting series. The reason for this is the high instability of some classes of open nonlinear systems categorized as dynamic chaos [8,10,11,21]. Examples of such unstable systems are gas-dynamic and hydrodynamic turbulent flows, high-temperature plasma, etc. [9,13,22]. In finance, chaotic behavior is particularly prominent in inertia-less environments such as electronic capital markets [1,15,17,26].

Investigating a multidimensional process is meaningful only if the parameters of the observation vector are correlated. Otherwise, the solution is limited to sequentially examining one-dimensional processes. Inter-dependencies make a regularization of chaotic observations possible (in other words, forecasting in general). However, the conventional assessment of inter-dependencies encounters serious problems due to the chaotic nature of the series [14,15]. Making correlation estimates over a large time interval is not feasible, and immediate estimates on a limited observation window are not stable. In addition, correlations larger than 0.9 can lead to degeneracy of the observation matrix. Thus, an alternative approach based on either of two branches of singular value decomposition (SVD) appears promising immunocomputing (IC) [4,23,24] and singular spectrum analysis (SSA) [2,6,7]. IC has found its use in the tasks of information...
processing for molecular protein complexes of the immune system. First of all, it concerns the issues of recognition and classification of foreign cells intended for stimulating the body’s defense mechanisms [4, 24]. As a result, the applied aspects of SVD analysis obtained the name of IC.

This article investigates the possibility of applying IC and SSA to forecasting multidimensional chaotic environments. We consider highly dynamic chaotic processes which are not suitable for multivariate statistical analysis [2, 19]. We use dimension reduction techniques based on the representation of data matrices in first singular basis in space.

2 Related work

The main directions of IC development are related to its practical applications, in particular, in classification and clustering problems. For example, recognition of multidimensional images uses vector projections on space generated by several singular components. This approach gives rise to a specific pseudometric (proximity measure) called the \(L_r\) distance [2, 12].

The problems of situational analysis are solved similarly: by mapping the observed situation, determined by state vector \(X_0, \ldots, X_k\), to a regular situation by the closest value of the pseudometric in \(L_r\).

In problems of approximation of random fields, the value of the sought hypersurface \(f(x_0)\) is estimated with linear interpolation over \(k\) nearest points \(x_0, \ldots, x_k\): \(f = c_0 f(x_0) + \ldots + c_k f(x_k)\), where \(c_j = \left(1 + d_j \sum_{i=0}^{k} d_i^{-1}\right)^{-1}\).

A variant of the \(L_r\) pseudometric is used as \(d_j, j = 1, \ldots, k\) from \(x_0\) to \(x_j\), \(j = 1, \ldots, k\).

If a segment of an \(m\)-dimensional series is represented as a transport matrix of size \(n \times m\), then it can be approximated by a sum of elementary matrices of unit rank, thus making it possible to analyze separately series terms of the simplified structure. This approach significantly reduces the dimension of the original problem.

An interesting direction in analyzing one-dimensional random series is a group of methods based on embedding a time series in a multidimensional space followed by a singular decomposition of the resulting Hankel matrix. This approach is also based on SVD and is known in the literature as “Caterpillar” [16, 18]. The Caterpillar method identifies time series components and solves the problems of forecasting, parameter estimation and detecting various types of decomposition. The applications of this approach that use the projection into the space of principal singular components are based on the application of the Euclidean metric in the space of projections, which makes it much easier to analyze and interpret the results.

3 Methods

3.1 Basics of IC and SSA approaches

We have a traditional problem for data analysis of representing the structure of \(m\)-dimensional data with \(k\) generalized features, where \(k < p < m\). It is usually solved using the principal component method. However, its effectiveness significantly depends on the correlation properties of the observation matrix. The correlations between variables in a multidimensional chaotic process change rapidly and within very wide limits. Very strong correlation can lead to degeneracy or poor conditionality of the evaluation task. Therefore, we consider the singular decomposition technique (or IC) as an alternative approach to the problem of data dimension compression.

The problem consists in approximating a matrix of multidimensional observations \(X\) of size \(n \times m\), \(n > m\) and rank \(p \leq m\) by another matrix \(Y\) of a smaller rank \(k < p\). The corresponding approximation is carried out by minimizing the quadratic distance between the matrices

\[(X - Y)^T(X - Y) = \min,\]  

restricting \(\text{rank}(Y) = k < \min(n, p)\).
The solution of this problem was found in [5]. A real observation matrix \( X \) of dimension \( (n \times m) \) can be represented with an SVD (LR-decomposition):

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

(2)

where \( S = \text{diag}(s_1, s_2, \ldots, s_n) \) is a diagonal matrix whose elements \( s_1 \geq s_2 \geq \ldots \geq s_n \geq 0 \) are called singular values of the matrix \( X \).

\( L \) is a matrix of size \( (n \times m) \), the columns \( L_1, \ldots, L_n \) of which are orthogonal vectors of unit length, i.e., \( L_i^T L_i = E \), where \( E \) is the unit matrix. These columns are the left singular vectors of \( X \).

\( R \) is a matrix of size \( (m \times m) \), the columns \( R_1, \ldots, R_m \) of which are also orthogonal vectors of unit length \( R_i^T R_i = E \) which are called the right singular vectors of \( X \). These vectors are orthogonal in the Euclidean sense; from a probabilistic point of view, they are correlated.

If the rank of the observation matrix is \( \text{rank}(X) = p < m \), then among the singular numbers only \( p \) will be nonzero. In this case, the decomposition (2) can be rewritten as a sum of elementary matrices of unit rank:

\[
X = \sum_{i=1}^{p} s_iL_iR_i^T = s_1L_1R_1^T + \cdots + s_pL_pR_p^T.
\]

(3)

According to the Eckart-Young theorem [2,5], the solution of optimization problem (1) is the sum of the first \( k \) terms in (3), i.e., \( X = Y = \sum_{i=1}^{p} s_iL_iR_i^T = s_1L_1R_1^T + \cdots + s_kL_kR_k^T \).

With \( k = 1 \) (one-dimensional case), the best approximation is given by the first (maximum) singular value and the corresponding singular vectors \( A = s_1L_1R_1^T \). The matrix of observations \( X \) in this case turns into the sum of a small number of matrix segments of the same dimension, but of a very simple structure: each of them is a matrix of unit rank.

An important feature of SVD is its stability to small perturbations of the observation matrix. In other words, this representation of each matrix is a well-conditioned procedure. Such properties are not characteristic of the traditional spectral decomposition used in problems of multidimensional statistical analysis. As it has been already noted, this is important for processing multidimensional chaotic processes with a strongly pronounced dependence between the individual parameters of the observation vector.

Singular matrix decomposition is stable to small matrix perturbations, i.e., it is a well-conditioned procedure. Such properties are not characteristic of spectral decomposition used in problems of multidimensional statistical analysis. Several main developments have arisen from this approach.

1. In problems of recognition, classification and clustering, vectors are projected on the space generated by several singular components (3), which generates a specific pseudometric [12, 24].

2. The tasks of situational analysis are solved in a similar fashion: the observed situation \( x_0 \) is associated with the closest by the pseudometric of the regular situations \( x_1, \ldots, x_k \) [24].

3. In random field interpolation, \( f(x_0) \) is estimated via linear interpolation on \( k \) nearest points \( x_1, \ldots, x_k \):

\[
f = c_1 f(x_1) + \ldots + c_k f(x_k),
\]

where

\[
G = \frac{1}{1 + d \sum_{i=1}^{k} \frac{1}{d_i}}.
\]

This approach stands out by employing the metric used in the projection space as a measure of proximity \( d_i \) from \( x_0 \) to \( x_i \).

4. If a segment of an \( m \)-dimensional series is represented as a matrix \( (n \times m) \), then it can be approximated by the sum of elementary matrices of unit rank, thus making it possible to analyze separately series-terms of the simplified structure. This significantly reduces the dimensionality of the original problem.

Based on these approaches as well as Caterpillar method, new algorithms for identifying local structure of multidimensional chaotic time series can be built.
3.2 Singular analysis algorithms for individual selected components

Let the selected component be a one-dimensional time series $y = (y_1, \ldots, y_N)$. We will match it with a Hankel matrix of size $K \times L$:

$$
Y = \begin{bmatrix}
  y_1 & y_2 & \cdots & y_L \\
  y_2 & y_3 & \cdots & y_{L+1} \\
  \vdots & \vdots & \ddots & \vdots \\
  y_K & y_{K+1} & \cdots & y_{L+K+1}
\end{bmatrix}, \quad L + K + 1 = N,
$$

where $L$ is the width of the sliding window. Let us construct a decomposition (3) for it. Then, to each real harmonic of the series $y_1, \ldots, y_N$ corresponds $m^*$, $m^* < p$ different singular numbers, and it is determined by the sum of $m^*$ terms in Eq. (3) that correspond to these numbers. To fully restore such a component, it is necessary to average over the diagonals of the matrix $X$ of the same name.

At this stage, the choice of the width of the sliding window $L$ is the most problematic issue. In the process of optimizing the algorithm, one can also vary the window shift parameter $d$.

One of the problems is that it is impossible to build a hierarchy of the corresponding components based on the values of singular numbers $s_i$, $i = 1, \ldots, p$: the periodic component that is important for analysis is not necessarily associated with one of the largest singular values. This problem is especially acute for the rapidly changing chaotic processes characteristic of the dynamics of quotations of financial instruments. In the process of tracking the local fluctuations of the components, it is necessary to consider the sequential differences of the singular values determined on each window. The appearance of zero in such a sequence indicates the presence of a quasi-periodic component, which can be visualized either by highlighting the corresponding singular components or by considering the change in a pair of components with equal (or sufficiently close) singular numbers on their phase plane [2,6].

3.3 Analyzing correlation properties of financial series

As an example of initial data, consider a segment of observations of centered quotation values of five currency instruments with the highest degree of correlation over an observation interval of 500 days (Figure 1).
The instruments (EURUSD, NZDUSD, AUDUSD, AUDJPY and NZDJPY) were selected according to the degree of correlation from 16 overall most common currency instruments. The reason for the interdependence of currency pairs is largely related to the nature of international trade and global financial flows. The currencies of countries with large trade deficits tend to have a negative correlation with countries that run surpluses. Similarly, the currencies of wealthy commodity exporters will often have a negative correlation with countries that rely heavily on imports. The color representation of their correlation matrix is shown in Figure 2. The color scheme for their correlation matrix is shown in Figure 3. The correlation estimates based on the data obtained over the entire observation interval range oscillate within 0.91–0.98, which indicates poor conditionality of the observation matrix.

**Figure 2:** Correlation matrix estimates for 16 common currency instruments.

**Figure 3:** Correlation matrix estimates for the five currency instruments with the highest correlation.
The obtained conclusion indicates that in the problems of integral representation of the dynamics of market segments, transition from the traditional approach based on component data analysis to IC-based (i.e., SSA algorithm) representations would be practical.

3.4 Calculating singular components for multidimensional series of the currency exchange market state

Let a matrix of observations \( X(n \times m) \), \( n > m \) correspond to the example with five correlated financial instruments. We are going to construct an SVD (2) for this matrix. Limiting ourselves to three projections \( k = 3 \), we present the corresponding terms as

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

Each projection is a matrix \( (n \times m) \) of unit rank, so only the first column \( x_i(1) \), \( i = 1, \ldots, k \) and the following set of coefficients need to be extracted from each projection \( X(i) \):

\[
C_j(i) = \frac{\sum_{j=1}^{m} x_j(i)}{\sum_{j=1}^{m} x_j(i)}, \quad j = 1, \ldots, m, \tag{4}
\]

which will be needed to return to the original variables.

Figure 4 plots changes in the values of the first three singular components at the observation interval \( T = 200 \) minute counts. This observation interval is extracted for an example from the general dataset shown in Figure 1.

It is important to point out that for the example at \( k = 3 \), the variance criterion \( D(k) = \frac{\sum_{j=1}^{k} \sigma_j^2}{\sum_{j=1}^{m} \sigma_j^2} \) is equal to 0.98, so that the transition to decomposition (3) occurs almost without loss of information.

Figure 5 shows a plot of logarithms of individual singular components. We can infer that various singular components are unequally significant, and that it is possible to limit the analysis results to 2–3 components, which solves the problem of visualizing multidimensional data.

![Figure 4: Plots of changes in the values of the first three singular components on observation interval \( T = 200 \) minute counts.](image-url)
3.5 Forecasting using SSA

1. Fix a sliding window $X$, $L$ counts wide.
2. At each step, construct an SVD for matrix $X$. The ratio of the sum of several first singular values to the sum of all singular values is interpreted as the fraction of information explained by these first singular components.
3. Each selected component $X_i$ is a matrix of unit rank and the same dimensions as $X$, therefore, only their first columns (or rows) and coefficients (4) for recovering their estimates are necessary for further analysis.
4. Apply one of the SSA procedures [2,6] to the selected one-dimensional series to provide filtering, interpolation and forecasting.
5. Using the estimates, recover matrices $X_i$ and calculate their weighted sum, which is the estimate for the initial matrix $X$.
6. Correct the estimate of $X$ with respect to the bias and scale parameters.

4 Computational experiments

As an example of using the fusion of IC and SSA, we analyze the above-presented data of five highly correlated currency instruments. Using the Caterpillar method, we select a segment of series of $L$ minute counts and construct a forecast for each component using SSA. Figure 6 shows plots of the resulting forecast with the use of extrapolation of singular components, followed by recalculation into the initial dimension of the observed chaotic process.

The figure implies a fundamental feasibility of restoring a correct approximation of a chaotic process and the possibility of forecasting it using pre-aggregated data based on SSA.

This approach is promising: the aggregated form contains information about the mood of a given market segment, and the forecast takes into account its general trends. Another expected advantage of this technique is its increased resistance to strong correlation between individual financial instruments of the selected segment.

Scanning the entire series $r$ counts ahead (count step is 4 s, width of sliding window $L = 2$ min, window shift $d = 1$ min; Figure 7), we find out that the SD of the forecast increases monotonically with the increase
of \( r \), and bias decreases at first, but begins increasing with \( r = 8 \) and so on. We can conclude that this approach limits the forecast horizon \( r \) at around 7 steps or 30 s. At the same time, standard procedures of local polynomial forecasting [16,18] possess the same qualities only when forecasting 1 or 2 steps (not more than 10 s) ahead. However, the final conclusions about the feasibility of employing SSA in the tasks of proactive management require additional research. In particular, it would be interesting to compare the
proposed technique with component data analysis, at least in conditions of low dimensions that would allow such a comparison.

5 Conclusion

Representation of multivariate data matrices in the first singular basis and dimension reduction methods make it possible to undergo transition from a multidimensional time series to an integral curve in a low-dimensional space. This curve can be interpreted as a phase trajectory in a generalized state space. The proposed transformation conforms to the set of constraint characteristic of most other similar approaches [19,25]. In particular, singular representation can be implemented for highly correlated and highly variable series, for which the observation matrix may be poorly conditioned with all the ensuing consequences of incorrectly formulated identification and forecasting problems. We propose a technique whose feasibility highly depends on initial properties of data. Best predictions may be supposed when the dimension of the data set is compared to length of adequate sliding window and when data include some latent periodicities.

At the same time, it should be taken into account that data integration can create uncertainty in reverse transformation, which can lead to ambiguous and paradoxical results of restoring one-dimensional components, especially in forecasting problems.

The proposed technique acquires a significant role in the tasks of proactive management of financial instruments at capital markets. The data model that is closest to reality is based on the concept of stochastic chaos, which means that series are fundamentally non-stationary and non-ergodic. This forces us to operate with these data inside a limited sliding window. Under these conditions, increased stability of singular components to variations in the statistical structure of the source data may increase the effectiveness of the entire asset management strategy.

In particular, SSA can be useful in rapid detection of significant discrepancies in quotations. Furthermore, it appears suitable for precedent analysis that deals with similarity metrics for multidimensional observation segments. These issues, along with selecting a data compression technique for observation segments with different dynamic characteristics, are the subject of our further research.

Acknowledgements: The authors are grateful to participants at the Center for Econometrics and Business Analytics (ceba-lab.org, CEGA) seminar series for helpful comments and suggestions.

Funding information: The research of Alexander Musaev described in this paper is partially supported by the Russian Foundation for Basic Research (grant 20-08-01046), state research FPZ-2022-0004. Dmitry Grigoriev research for this paper was supported by Saint Petersburg State University, project ID: 93024916.

Conflict of interest: The authors state no conflict of interest.

References

[1] Bauwens, L. & Giot, P. (2001). Econometric modelling of stock market intraday activity. Advanced Studies in Theoretical and Applied Econometrics. (Vol. 38). Boston: Kluwer Academic, Springer Science & Business Media.
[2] Bolshakov, A. A., & Karimov, R. N. (2015). Multidimensional data and time series processing methods. Telekom, Moscow: Goryachaya liniya (in Russian).
[3] Corona, I., Giacinto, G., Mazzariello, C., Roli, F., & Sansone, C. (2009). Information fusion for computer security: State of the art and open issues. Information Fusion, 10(4), 274–284.
[4] Dasgupta, D. (Ed.). (2012). Artificial immune systems and their applications. New York: Springer Science & Business Media.
[5] Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. Psychometrika, 3(3), 211–218.
[6] Elsner, J. B., & Tsonis, A. A. (1996). Singular spectrum analysis: A new tool in time series analysis. New York: Springer Science & Business Media.
[7] Golyandina, N., Korobeynikov, A., & Zhigljavsky, A. (2018). Singular spectrum analysis with R. Berlin Heidelberg: Springer.
[8] Kautz, R. (2011). Chaos: The science of predictable random motion. Oxford: Oxford University Press.
[9] Klimontovich, Y. L. (1995). Statistical theory of open systems. (Vol. 67). Dordrecht, Boston, London: Kluwer Academic Publishers.
[10] Lorenz, E. N. (1996). 143 The essence of Chaos. Washington: University of Washington.
[11] Lorenz, E. N., & Haman, K. (1996). The essence of chaos. Pure and Applied Geophysics, 147(3), 598–599.
[12] Makshanov, A., Zhuravlev, A., & Tyndykar, L. (2019). Elaboration of multichannel data fusion algorithms at marine monitoring systems. In International Scientific Siberian Transport Forum (pp. 909–923). Cham: Springer.
[13] Manneville, P. (2004). Instabilities, chaos and turbulence: An introduction to nonlinear dynamics and complex systems. Singapore: World Scientific.
[14] Musaev, A., & Grigoriev, D. (2021). Analyzing, modeling, and utilizing observation series correlation in capital markets. Computation, 9(8), 88.
[15] Musaev, A., & Grigoriev, D. (2022). Numerical studies of statistical management decisions in conditions of stochastic chaos. Mathematics, 10(2), 226.
[16] Nekrutkin, V. (1996). Theoretical properties of the “Caterpillar” method of time series analysis. In Proc. 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing IEEE Computer Society, pp. 395–397.
[17] Peters, E. E. (1996). Chaos and order in the capital markets: A new view of cycles, prices, and market volatility. New York: John Wiley & Sons.
[18] Polukoshko, S., Hilkevica, G., & Gonca, V. (2015). Nonstationary processes studying based on “Caterpillar” – SSA method. In Vibration engineering and technology of machinery (pp. 999–1008). Cham: Springer.
[19] Rao, C. R. (1973). Linear statistical inference and its applications (Vol. 2, pp. 263–270). New York: Wiley.
[20] Schölkopf, B., Smola, A., & Müller, K. R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation, 10(5), 1299–1319.
[21] Smith, L. (2007). Chaos: A very short introduction. Oxford: Oxford University Press.
[22] Takens, F. (1981). Detecting strange attractors in turbulence. In Dynamical systems and turbulence, Warwick 1980 (pp. 366–381). Berlin, Heidelberg: Springer.
[23] Tarakanov, A. O. (2001). Information security with formal immune networks. In International Workshop on Mathematical Methods, Models, and Architectures for Network Security (pp. 115–126). Berlin, Heidelberg: Springer.
[24] Tarakanov, A. O., Skormin, V. A., Sokolova, S. P., & Sokolova, S. S. (2003). Immunocomputing: Principles and applications. New York: Springer Science & Business Media.
[25] Wilks, S. S. (1964). Mathematical statistics. New York: John Wiley & Sons.
[26] Gregory-Williams, J., & Williams, B. M. (2004). Trading chaos: Maximize profits with proven technical techniques (Vol. 161). New York: John Wiley & Sons.