Singular Spectrum Analysis: Methodology and Application to Economics Data

Hossein HASSANI · Anatoly ZHIGLJAVSKY

Received: .... 2006/ Revised: 2008

Abstract We describe the methodology of Singular Spectrum Analysis (SSA) and demonstrate that it is a powerful method of time series analysis and forecasting, particulary for economic time series. We consider the application of SSA to the analysis and forecasting of the Iranian national accounts data as provided by the Central Bank of the Islamic Republic of Iran.

Key words Economic time series, Singular Spectrum Analysis (SSA), forecasting, Iranian national accounts.

1 Introduction

Econometric methods have been widely used to forecast the evolution of quarterly and yearly national account data sets. However, many of these structural or time series forecasting models have failed to accurately predict the growth rate of Gross Domestic Product (GDP) or the turning points of business cycles in the industrial economies (see for example, [1]).

Many factors could affect the national economies and hence the national account data which are at best inaccurate representation of the macroeconomic variables because of measurement noise. The exogenous factors that cause instability in macroeconomies including technological changes, government policy changes, changes in the preferences of the consumers, and other events. These shocks cause structural changes in these time series making them nonstationary. Development of a methodology which is robust under these changes is of paramount importance in accurate prediction of macroeconomic time series.

Moreover, many structural econometric and time series models devised for forecasting macroeconomic time series are based on restrictive assumptions of normality and linearity of the observed data. The methods that do not depend on these assumptions could be very useful for modeling and forecasting economics data.

Furthermore, it is well known that noise can seriously limit accuracy of time series prediction. Currently there are not many effective forecasting techniques available when there is significant noise in the time series data. There are two main approaches for forecasting noisy time series.

Hossein HASSANI
1– School of Mathematics, Statistics Group, Cardiff University, UK. Email: HassaniH@cf.ac.uk.
2– Central Bank of the Islamic Republic of Iran, 207/1, Pasdaran Avenue, Tehran, Islamic Republic of Iran.
Anatoly ZHIGLJAVSKY
School of Mathematics, Statistics Group, Cardiff University, UK. Email: ZhigljavskyAA@cf.ac.uk.

According to the first one, we ignore the presence of noise and fit a forecasting model directly from noisy data hoping to extract the underlying deterministic dynamics. According to the second approach, which is often more effective that the first one, we start with filtering the noisy time series in order to reduce the noise level and then forecast the new data points (see, for example, [2, 3]). There are several nonlinear noise reduction methods such as local projective, singular value decomposition (SVD) and simple nonlinear filtering. It is currently accepted that SVD-based methods are very effective for the noise reduction in deterministic time series and correspondingly for forecasting [3].

Additionally, some of the previous research have considered economic and financial time series as deterministic, linear dynamical systems. In this case, the linear models can be used for modeling and forecasting. However, it has been shown that most of the financial time series are nonlinear (see, for example, [2, 3, 4, 5]); in these cases, we should use nonlinear methods. Having a method that works well for both linear and nonlinear time series is ideal for modeling and forecasting. The Singular Spectrum Analysis (SSA) meets all conditions stated above. The SSA technique is a nonparametric technique of time series analysis incorporating the elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing.

The birth of SSA is usually associated with the publication of papers by Broomhead and King (e.g. [6]) while the ideas of SSA were independently developed in Russia (St. Petersburg, Moscow) and in several groups in the UK and USA. A thorough description of the theoretical and practical foundations of the SSA technique (with many examples) can be found in [7, 8]. An elementary introduction to the subject can be found in [9]. Below we describe several applications of SSA and provide a brief discussion on the methodology used. In doing so, we mainly follow [7], chapters 1 and 2.

The basic SSA method consists of two complementary stages: decomposition and reconstruction; both stages include two separate steps. At the first stage we decompose the series and at the second stage we reconstruct the original series and use the reconstructed series for forecasting new data points. The main concept in studying the properties of SSA is 'separability', which characterizes how well different components can be separated from each other. The absence of approximate separability is often observed in series with complex structure. For these series and series with special structure, there are different ways of modifying SSA leading to different versions such as SSA with single and double centering, Toeplitz SSA, and sequential SSA, see [7], Sect. 1.7.

An important feature of SSA is that it can be used for analyzing relatively short series. On the other hand, asymptotic separation plays a very important role in the theory of SSA. It has been observed that in many practical applications the asymptotic features (which hold as the length of the series T tends to infinity) are met for relatively small values of T; it is not uncommon to successfully apply SSA to series with T equal to 20–30. The series considered in this paper have lengths $T = 68$ and $T = 45$.

It is worth noting that although some probabilistic and statistical concepts are employed in the SSA-based methods, we do not have to make any statistical assumptions such as stationarity of the series or normality of the residuals.

In addition, the method has several essential extensions. First, the multivariate version of the method permits the simultaneous expansion of several time series; see, for example [8]. Second, the SSA ideas lead to several forecasting procedures for time series; see [7, 8]. Also, the same ideas are used in [7] and [10] for change-point detection in time series. For comparison with classical methods, ARIMA, ARAR algorithm and Holt-Winter, see [11, 12]. For automatic methods of identification within the SSA framework see [13] and for recent work in 'Caterpillar'-SSA software as well as new developments see [14].

Let us mention some other areas related to SSA. A variety of techniques of time series analysis and signal processing have been suggested that use SVD of certain matrices; for surveys see, for example, [15, 16]. Most of these techniques are based on the assumption that the original series is random and stationary; they include some techniques that are famous in signal processing, such as Karhunen-Loeve decomposition (for signal processing references see, for example [17]). Some statistical aspects of the SVD-based methodology for stationary series are considered, for example, in [18] (Chapter 9) and [19, 20].

In this paper we start with a brief description of the methodology of SSA and apply this technique to 32 original data sets, 16 quarterly and 16 yearly, which are taken from the Central Bank of the Islamic Republic of Iran (CBI) [21]. We use the series of Iranian GDP (quarterly) as the main data set for illustrating details of the practical application of the SSA methodology.

2 Methodology

The main purpose of SSA is to decompose the original series into a sum of series, so that each component in this sum can be identified as either a trend, periodic or quasi-periodic component (perhaps, amplitude-modulated), or noise. This is followed by a reconstruction of the original series. The Basic SSA technique is performed in two stages, both of which include two separate steps as follows:

$$\begin{cases} \text{Stage 1 : Decomposition} & \begin{cases} \text{Step 1 : Embedding} \\ \text{Step 2 : Singular Value Decomposition (SVD)} \end{cases} \\ \text{Stage 2 : Reconstruction} & \begin{cases} \text{Step 1 : Grouping} \\ \text{Step 2 : Diagonal Averaging} \end{cases} \end{cases}$$

2.1 Decomposition

1st step: Embedding

Embedding can be regarded as a mapping that transfers a one-dimensional time series $Y_T = (y_1, \dots, y_T)$ into the multidimensional series $X_1, \dots, X_K$ with vectors $X_i = (y_i, \dots, y_{i+L-1})^T \in \mathbf{R}^L$, where $K = T - L + 1$. Vectors X_i are called *L-lagged vectors* (or, simply, *lagged vectors*). The single parameter of the embedding is the *window length* L, an integer such that $2 \le L \le T$. The window length L should be sufficiently large. The result of this step is the trajectory matrix

$$\mathbf{X} = [X_1, \dots, X_K] = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_k \\ y_2 & y_3 & y_4 & \cdots & y_{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_T \end{pmatrix}.$$

Note that the trajectory matrix $\mathbf{X}$ is a Hankel matrix, which means that all the elements along the diagonal $i+j = const$ are equal. Embedding is a standard procedure in time series analysis. With the embedding performed, future analysis depends on the aim of the investigation. For specialists in dynamical systems, a common technique is to obtain the empirical distribution of all pairwise distances between the lagged vectors X_i and X_j and then calculate the so-called correlation dimension of the series. Note that in this approach, L must be relatively small and K must be very large (formally, $K \to \infty$). The approximation of a stationary series with the help of the autoregression model can also be expressed in terms of embedding: if we deal with the model

$$y_{i+L-1} = a_{L-1} y_{i+L-2} + \cdots + a_1 y_i + \varepsilon_{i+L-1}, \qquad i \geq 1$$

then we search for vector $A = (a_1, \ldots, a_{L-1}, -1)^T$ such that the scalar products (X_i, A) are described in terms of certain noise series.

2nd step: Singular Value Decomposition (SVD)

The second step, the SVD step, makes the singular value decomposition of the trajectory matrix and represents it as a sum of rank-one bi-orthogonal elementary matrices. Denote by $\lambda_1, \ldots, \lambda_L$ the eigenvalues of $\mathbf{X}\mathbf{X}^T$ in decreasing order of magnitude ($\lambda_1 \geq \ldots \lambda_L \geq 0$) and by $U_1, \ldots, U_L$ the orthonormal system of the eigenvectors of the matrix $\mathbf{X}\mathbf{X}^T$ corresponding to these eigenvalues. Set

$$d = \max(i, \text{ such that } \lambda_i > 0) = rank\, \mathbf{X}.$$

If we denote $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$, then the SVD of the trajectory matrix can be written as:

$$\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d, \tag{1}$$

where $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$. The matrices $\mathbf{X}_i$ have rank 1 (thus they are elementary matrices); U_i (in SSA literature they are called 'factor empirical orthogonal functions' or simply EOFs) and V_i (often called 'principal components') are the left and right eigenvectors of the trajectory matrix. The collection $(\sqrt{\lambda_i}, U_i, V_i)$ is called the i-th eigentriple of the matrix $\mathbf{X}$, $\sqrt{\lambda_i}$ $(i = 1, \ldots, d)$ are the singular values of the matrix $\mathbf{X}$ and the set $\{\sqrt{\lambda_i}\}$ is called the spectrum of the matrix $\mathbf{X}$. If all eigenvalues have multiplicity one, then the expansion (1) is uniquely defined.

SVD (1) is optimal in the sense that among all the matrices $\mathbf{X}^{(r)}$ of rank $r < d$, the matrix $\sum_{i=1}^{r} \mathbf{X}_i$ provides the best approximation to the trajectory matrix $\mathbf{X}$, so that $\parallel \mathbf{X} - \mathbf{X}^{(r)} \parallel$ is minimum. Here the norm of a matrix $\mathbf{Y}$ is defined as $\sqrt{\langle \mathbf{Y}, \mathbf{Y} \rangle}$, where the scalar product of two matrices $\mathbf{Y} = (y_{ij})_{i,j=1}^{q,s}$ and $\mathbf{Z} = (z_{ij})_{i,j=1}^{q,s}$ is $\langle \mathbf{Y}, \mathbf{Z} \rangle = \sum_{i,j=1}^{q,s} y_{ij} z_{ij}$. Note that $\parallel \mathbf{X} \parallel^2 = \sum_{i=1}^{d} \lambda_i$ and $\parallel X_i \parallel^2 = \lambda_i$ for $i = 1, \ldots, d$. Thus, we can consider the ratio $\lambda_i / \sum_{i=1}^{d} \lambda_i$ as the characteristic of the contribution of the matrix $\mathbf{X}_i$ to expansion (1). Consequently, $\sum_{i=1}^{r} \lambda_i / \sum_{i=1}^{d} \lambda_i$, the sum of the first r ratios, is the characteristic of the optimal approximation of the trajectory matrix by the matrices of rank r.

Another optimal feature of the SVD is related to the properties of the directions determined by the eigenvectors $U_1, \ldots, U_d$. Specifically, the first eigenvector U_1 determines the direction such that the variation of the projections of the lagged vectors into this direction is maximum. Every subsequent eigenvector determines the direction that is orthogonal to all previous directions, and the variation of the projection of the lagged vectors onto this direction is also maximum. Therefore, it is natural to call the direction of the i-th eigenvector U_i the *i-th principal direction*. Note that the elementary matrices $\mathbf{X}_i$ are built up from the projections of the lagged vectors onto the i-th particular directions. This view on the SVD of the trajectory matrix composed of L-lagged vectors and an appeal to association with the *principal component analysis* lead to the following terminology. We shall call the vector U_i the i-th eigenvector, the vector V_i will be called the *i-th factor vector* and the vector $Z_i = \sqrt{\lambda_i} V_i$ the *i-th principal component*.

2.2 Reconstruction

1st Step: Grouping

The grouping step corresponds to splitting the elementary matrices into several groups and summing the matrices within each group. Let $I = \{i_1, \dots, i_p\}$ be a group of indices $i_1, \dots, i_p$. Then the matrix $\mathbf{X}_I$ corresponding to the group I is defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_p}$. The spilt of the set of indices $J = \{1, \dots, d\}$ into disjoint subsets $I_1, \dots, I_m$ corresponds to the representation

$$\mathbf{X} = \mathbf{X}_{I_1} + \cdots + \mathbf{X}_{I_m}. \tag{2}$$

The procedure of choosing the sets $I_1, \dots, I_m$ is called the eigentriple grouping. For a given group I the contribution of the component $\mathbf{X}_I$ in the expansion (2) is measured by the share of the corresponding eigenvalues: $\sum_{i \in I} \lambda_i / \sum_{i=1}^{d} \lambda_i$. If the matrix $\mathbf{X}_I$ is a Hankel matrix, then there exist series $Y_T^{(1)}$ and $Y_T^{(2)}$ such that $Y_T = Y_T^{(1)} + Y_T^{(2)}$ and the trajectory matrices of these series are $\mathbf{X}_I$ and $\mathbf{X}_{J \setminus I}$, respectively. If the matrices $\mathbf{X}_I$ and $\mathbf{X}_{J \setminus I}$ are approximately Hankel matrices then the trajectory matrices of the series $Y_T^{(1)}$ and $Y_T^{(2)}$ are close to $\mathbf{X}_I$ and $\mathbf{X}_{J \setminus I}$. In this case we shall say that the series are approximately separable, see [7] for many more details. Therefore, the purpose of the grouping step (that is, the procedure of arranging the indices $1, \dots, d$ into groups) is to find several groups $I_1, \dots, I_m$ such that the matrices $\mathbf{X}_{I_1}, \dots, \mathbf{X}_{I_m}$ satisfy (2) and are close to certain Hankel matrices. The grouping step is based on the analysis of the eigenvectors U_i and V_i, and eigenvalues λ_i in the SVD expansion (1). The principles and methods of identifying the SVD components for their inclusion into different groups are described in [7], Sect. 1.6. Since each matrix component of the SVD is completely determined by the corresponding eigentriple, we shall talk about the grouping of the eigentriples rather than the grouping of the elementary matrices $\mathbf{X}_i$.

2nd Step: Diagonal averaging

The purpose of diagonal averaging is to transform a matrix to the form of a Hankel matrix which can be subsequently converted to a time series. If z_{ij} stands for an element of a matrix $\mathbf{Z}$, then the k-th term of the resulting series is obtained by averaging z_{ij} over all i, j such that $i + j = k + 1$. This procedure is called diagonal averaging, or Hankelization of the matrix $\mathbf{Z}$. The result of the Hankelization of a matrix $\mathbf{Z}$ is the Hankel matrix $\mathcal{H}\mathbf{Z}$. Note that the Hankelization is an optimal procedure in the sense that the matrix $\mathcal{H}\mathbf{Z}$ is the nearest to $\mathbf{Z}$ (with respect to the matrix norm) among all Hankel matrices of the corresponding size (see [7], Sect. 6.2). In its turn, the Hankel matrix $\mathcal{H}\mathbf{Z}$ uniquely defines the series by relating the value in the diagonals to the values in the series. By applying the Hankelization procedure to all matrix components of (2), we obtain another expansion:

$$\mathbf{X} = \widetilde{\mathbf{X}}_{I_1} + \ldots + \widetilde{\mathbf{X}}_{I_m} \tag{3}$$

where $\widetilde{\mathbf{X}}_{I_1} = \mathcal{H}\mathbf{X}$. This is equivalent to the decomposition of the initial series $Y_T = (y_1, \dots, y_T)$ into a sum of m series:

$$y_t = \sum_{k=1}^{m} \widetilde{y}_t^{(k)} \tag{4}$$

where $\widetilde{Y}_T^{(k)} = (\widetilde{y}_1^{(k)}, \dots, \widetilde{y}_T^{(k)})$ corresponds to the matrix $\mathbf{X}_{I_k}$. A sensible grouping leads to the decomposition (2) where the resultant matrices $\mathbf{X}_{I_k}$ are almost Hankel ones. This corresponds to approximate separability and implies that pairwise scalar products of different matrices $\mathbf{X}_{I_k}$ in (3) are small. The procedure of computing the time series $\widetilde{Y}_T^{(k)}$ (that is, building up the group I_k plus diagonal averaging of the matrix $\mathbf{X}_{I_k}$) will be called *reconstruction* of a series $Y_T^{(k)}$ by the eigentriples with indices in I_k. In relation to the grouping method, it is worthwhile to note that if L is large enough, the eigenvectors in a sense imitate the behavior of the corresponding time series components. In particular, the trend of the series corresponds to slowly varying eigenvectors. The harmonic component produces a pair of left (and right) harmonic eigenvectors with the same frequency, etc.

2.3 Linear Recurrent Formulae (LRF)

Forecasting by SSA can be applied to the time series that approximately satisfy linear recurrent formulae (LRF).

We shall say that the series Y_T satisfies an LRF of order d if there are numbers $a_1, \ldots, a_d$ such that

$$y_{i+d} = \sum_{k=1}^{d} a_k y_{i+d-k}, \qquad 1 \le i \le T-d. \tag{5}$$

The fact that the series y_t satisfies an LRF (5) is equivalent to its representability as a sum of products of exponentials, polynomials and harmonics; that is,

$$y_t = \sum_{k=1}^{q} \alpha_k(t)\, e^{\mu_k t} \sin\left(2\pi\omega_k t + \varphi_k\right). \tag{6}$$

Here $\alpha_k(t)$ are polynomials, μ_k, ω_k and φ_k are arbitrary parameters. The number of linearly independent terms q in (6) is less than or equal to d. The class of series that can be approximated by the series satisfying LRFs of the form (5) (or, equivalently, by the time series of the form (6) with a small number of terms) is very broad.

3 Application of SSA to the analysis of Iranian GDP

The SSA technique can be applied to various time series. Using SSA for analyzing economics time series can be advantageous as these series typically contain periodic components that are difficult to handle with classical techniques. As our main example, let us consider the application of SSA for analyzing and forecasting the quarterly Iranian Gross Domestic Product (GDP) in detail. Fig. 1 shows this series in basic prices (at current price). Visual analysis of Fig. 1 indicates that the depicted series has a trend and this trend can be approximated by a function increasing exponentially fast. A harmonic seasonal component with increasing amplitude is also clearly seen.

3.1 Decomposition: trend, seasonality and residuals

A general descriptive model of the series that is considered in SSA is an additive model where the components of the series are trend, oscillation and noise. In addition, the oscillatory components are subdivided into periodic and quasi-periodic components, while noise components are, as a rule, *aperiodic* series. The sum of all additive components except for the noise

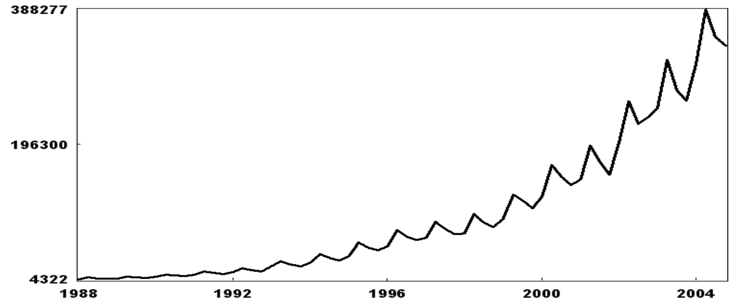

Figure 1: Quarterly Iranian GDP (billion Rials).

will be called the signal. So we decompose the series into the signal (trend and oscillations) and noise. Note that SSA does not require an a priori parametric model for trend and oscillations.

The choice of window length: as was mentioned earlier, the window length L is the only parameter at the decomposition stage. Selection of the proper window length depends on the problem in hand and on preliminarily information about the time series. Theoretical results advise us to choose L large enough but not greater than $T/2$. Knowing that the time series may have a periodic component with an integer period, to achieve a better separability of this periodic component it is advisable to take the window length proportional to that period. For example, the assumption that there is an annual periodicity in the series suggests that we must pay attention to the frequencies $k/12$ ($k = 1, ..., 12$). For quarterly data the period of the seasonal component is equal to 4. Using these recommendations, we take $L = 32$ (in our case T=68). So, based on this window length and on the SVD of the trajectory matrix, we have 32 eigentriples, ordered by their contributions (shares) into the decomposition. The leading eigentriple describes the general tendency of the series. Since in most cases the eigentriples with small shares are related to the noise component of the series, we need to identify the set of leading eigentriples. Let us consider the result of the SVD step. Fig. 2 represents principal components (left eigenvectors) related to the first 9 eigentriples. Note that the form of the factor vectors (right eigenvectors) is almost the same as the form of principal components because L=32 is close to K=37.

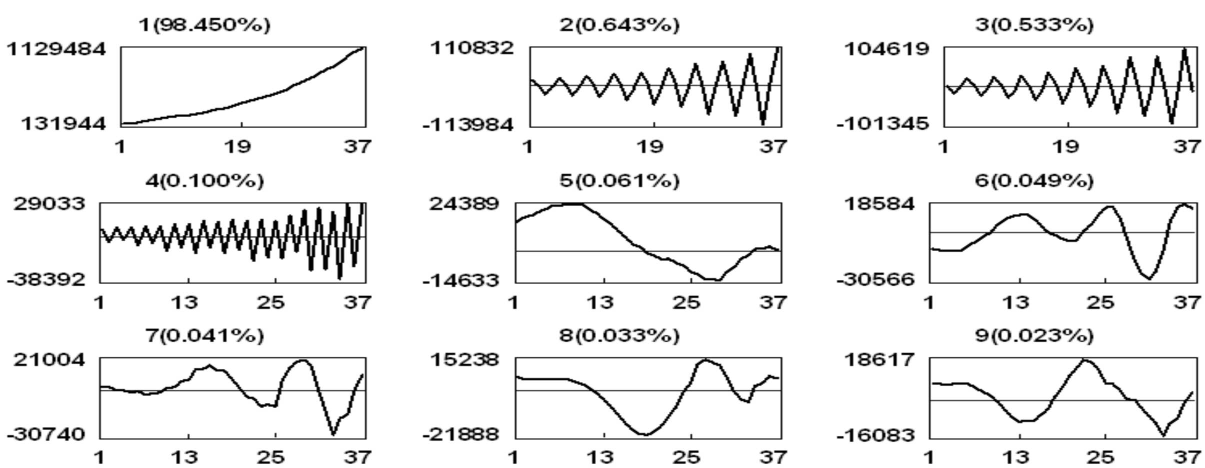

Figure 2: Principal components related to the first 9 eigentriples.

Separability: SSA decomposition of the series Y_T can be successful only if the resulting additive components of the series are approximately separable from each other. The following quantity (called the weighted correlation or *w-correlation*) is a natural measure of dependence between two series $Y_T^{(1)}$ and $Y_T^{(2)}$:

$$\rho_{12}^{(w)} = \frac{\left(Y_T^{(1)}, Y_T^{(2)}\right)_w}{\| Y_T^{(1)} \|_w \| Y_T^{(2)} \|_w}$$

where $\| Y_T^{(i)} \|_w = \sqrt{\left(Y_T^{(i)}, Y_T^{(i)}\right)_w}$, $\left(Y_T^{(i)}, Y_T^{(j)}\right)_w = \sum_{k=1}^{T} w_k y_k^{(i)} y_k^{(j)}, \quad (i, j = 1, 2)$ w_k=$\min\{k, L, T-k\}$ (here we assume $L \leq T/2$).

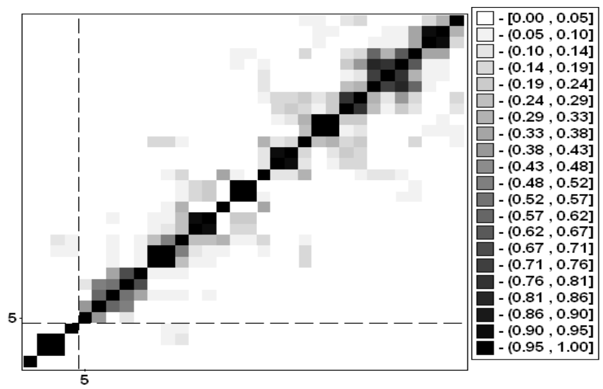

Figure 3: Matrix of w-correlations for 32 reconstructed components.

A useful tool for defining the groups of eigentriples is the matrix of the absolute values of the w-correlations, corresponding to the full decomposition (in this decomposition each group corresponds to only one matrix component of the SVD). If two reconstructed components have zero w-correlation it means that these two components are separable. If the absolute value of a w-correlation is small, then the corresponding series are almost w-orthogonal; if it is large, then the two series are far from being w-orthogonal and are therefore badly separable. Large values of correlations between reconstructed components indicate that they should possibly be gathered into one group and correspond to the same component in SSA decomposition [7]. In Fig. 3 w-correlations for 32 reconstructed components are shown in the 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1. Fig. 3 confirms that the first four eigentriples are well separated from a block of the remaining eigentriples (5–32) which we consider as noise.

Components	C_1	C_2	C_3	C_4	C_5
C_2	0.002				
C_3	0.001	0.960			
C_4	0.000	0.001	0.028		
C_5	0.001	0.001	0.001	0.015	
$C_{(6-32)}$	0.001	0.020	0.016	0.026	0.186

Table 1: w-correlations for components $C_1, \ldots, C_5$ and $C_{(6-32)}$.

The form of the matrix of w-correlations gives an indication of how to make the proper grouping: the leading eigentriple definitely corresponds to the trend (this situation is very common in practice), the three subsequent eigentriples correspond to the harmonics, and the large sparking square (after the fifth or sixth eigentriple) indicates the noise component. We can interpret the fifth component (and components 6–9 as well) as a part of either a trend or noise. Inspection of the matrix of w-correlations (see Table 1) and the quality of the approximation and forecast indicate that we get more stable results if we consider eigentriples above 4 as components of noise.

The components 5–9 are slowly varying and may perhaps be associated with business cycles. For example, the periodograms of these components show that there is definitely a periodic component with period 16 quarters = 4 years. However, the time series we analyze are too short (and too noisy) for reliable identification of the business cycles.

Trend identification: Trend is the slowly varying component of a time series which does not contain oscillatory components. Assume that the time series itself is such a component alone. Practice shows that in this case, one or more of the leading eigenvectors will be slowly varying as well. Exponential and polynomial sequences are good examples of this situation. We know that eigenvectors have a form similar to the form of the corresponding components of the initial time series, thus we should find slowly varying eigenvectors. This can be achieved by the inspection of one-dimensional plots of the eigenvectors.

In our case, the leading eigenvector is definitely of the required form but the eigentriples 2–4 are definitely not. Since we have decided that the eigentriples 5–32 correspond to the noise, the trend is described by the first eigentriple only. This directly implies that the trend of the original series is approximated by an exponential function. Note again, that we can build a more complicated approximation of the trend if we use some other eigentriples. However, the gain in precision will be very small and the model of the trend will become much more complicated. Fig. 4 shows the original series and the extracted trend (which is obtained from the first eigentriple).

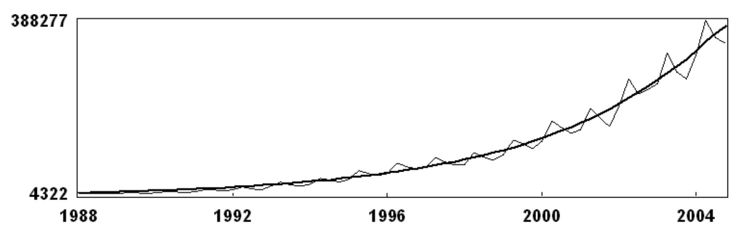

Figure 4: Trend extraction.

Identification of the harmonic components: The general problem here is the identification and separation of the oscillatory components of the series that do not constitute parts of the trend. In parametric form, this problem is extensively studied in classical spectral analysis theory. The statement of the problem in SSA is specified mostly by the model-free nature of the method. In practice, the singular values of two eigentriples of a harmonic series are often very close to each other, and this fact simplifies the visual identification of the harmonic components. An analysis of the pairwise scatterplots of the eigenvectors also helps to visually identify those eigentriples that correspond to the harmonic components of the series, provided these components are separable from the residual component. Fig. 5 depicts scatterplots of paired factor vectors from the Iranian GDP data, corresponding to the harmonics with a small number of frequencies. This figure shows two-dimensional graphs which form two-dimensional trajectories with vertices in a spiral-shaped curve. This indicates that these pairs of eigenvectors are produced by the modulated harmonic components of the initial time series. In that way, the eigentriples 2–3 correspond to the period 4 or frequency 3/12=1/4.

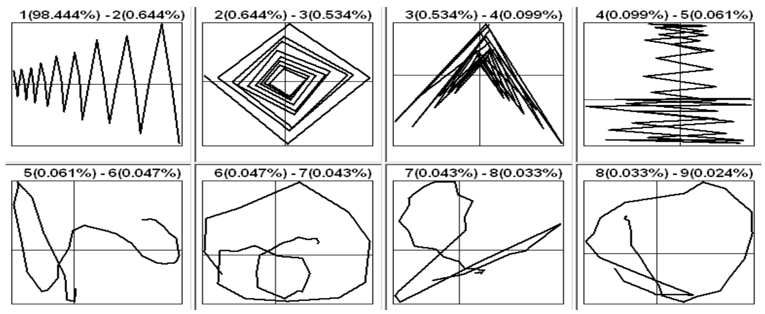

Figure 5: Scatterplots (with lines connecting consecutive points) of the first eight pairs of eigenvectors.

Let us describe additional information, which can help us to identify eigentriples and confirm the grouping of the components. Logarithms of eigenvalues provide such information in the following way: a pair of eigentriples corresponding to a harmonic component produces a plateau in this graph. Analysis of the matrix of w-correlation between reconstructed components of the initial time series is also useful for identification. Certainly, auxiliary information about the initial series always makes the situation clearer and helps in choosing the parameters of the methods. For example, the assumption that there might be a quarterly periodicity in the Iranian GDP data set suggests that the analyst must pay special attention to the frequency 1/4. As shown in Fig. 5, eigentriples 2–4 correspond to some harmonics, since their eigenvectors have a regular periodical form. Fig. 6 shows the oscillation of our data set which is obtained from eigentriples 2–4. Obviously, we do not have oscillation with equal amplitude, it has an increasing rate, similar to what we visually observed earlier for the original series. Also, Fig. 6 confirms that the eigentriple selection for the identification of oscillation of the original series seems correct as eigentriples 2–4 adequately reflect the oscillation behavior of the original series.

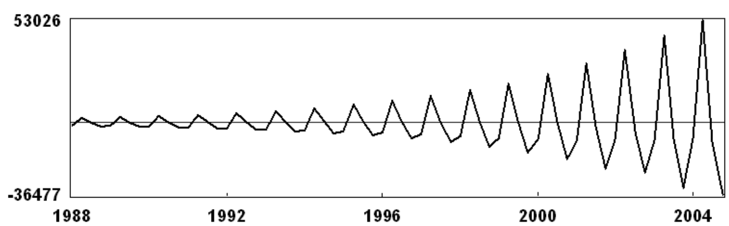

Figure 6: Oscillation extraction.

The components 2–4 correspond to the seasonality components of the series. Looking at the periodorgams of the components 5–9 we may suggest that these components reflect slowly-varying economic cycles. However, the time series we analyze are too short for a conclusion like that and for reliable forecasting of the economic cycles; therefore, we have preferred to omit the corresponding eigentriples on the stage of forecasting.

Separation of the noise from the signal: The problem of finding a refined structure of a series by SSA is equivalent to the identification of the eigentriples of the SVD of the trajectory matrix of this series, which correspond to trend, various oscillatory components, and noise. From the practical point of view, a natural way of noise extraction is the grouping of the eigentriples, which do not seemingly contain elements of trend or oscillation.

Let us make a few remarks concerning the separation of the components corresponding to noise. First, irregular behavior of eigenvectors can indicate that they are part of noise. This irregularity should be distinguished from component mixture, which is caused by lack of separability between these components. The noise component can often be identified as it typically creates a long tail of eigenvalues which are slowly decreasing (almost without jumps). Secondly, the large set of eigentriples which are highly correlated with each other is quite likely to belong to a noise. (Fig. 3 contains such a block of eigentriples with numbers 5–32). The interpretation of the eigentriple 5 (and perhaps 6–9 as well) is unclear. This is a border-line case discussed above; we classify these eigentriples as examples of noise. The fact that the w-correlation between the reconstructed series (the eigentriples 1–4) and the residuals (the eigentriples 5–32) is equal to 0.005 confirms that this grouping is very reasonable. Fig. 7 shows the residuals after extracting the trend and the seasonal component. If we add together the trend and the residuals we come to the original series adjusted for seasonal variations and if we add the series of Figs 4, 6 and 7 we will obtain the original series (Fig. 1).

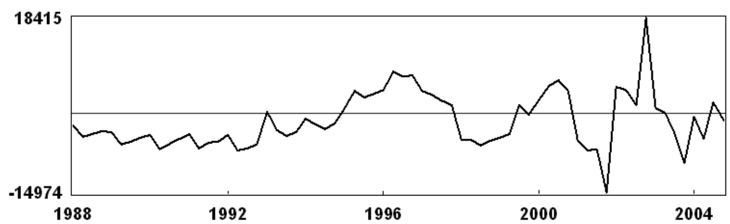

Figure 7: Residual series (eigentriples 5–32.)

3.2 Forecasting

A forecast can be made only if a certain model is built. The model can either be derived from the data or at least checked against the data. In SSA forecasting, this model can be described with the help of a linear recurrent formula. The class of series governed by linear recurrent formula (*LRF*s) is rather wide and important for practical implication.

Assume that we have a series $Y_T = \{y_t\} = Y_T^{(1)} + Y_T^{(2)}$ and the problem of forecasting its component $Y_T^{(1)}$. If $Y_T^{(2)}$ can be regarded as noise, then the problem is that of forecasting the signal $Y_T^{(1)}$ in the presence of a noise $Y_T^{(2)}$. The main assumptions are:
(a) the series $Y_T^{(1)}$ admits a recurrent continuation with the help of an LRF of a relatively small dimension d, and
(b) there exists a number L such that the series $Y_T^{(1)}$ and $Y_T^{(2)}$ are approximately separable for the window length L.

The assumption (b) is important as any time series $Y_T^{(1)}$ is an adaptive component of Y_T in the sense that $Y_T = Y_T^{(1)} + Y_T^{(2)}$ with $Y_T^{(2)} = Y_T - Y_T^{(1)}$. The assumption of (approximate) separability means that $Y_T^{(1)}$ is a natural additive component from the viewpoint of the SSA method.

To check the forecast quality, several methods can be applied. The main way is to truncate several last points of the original series, make the analysis of the reduced series and compare truncated values with the forecasted ones. We cut off the last 4 points of the series and forecast them (that is, we will forecast the Iranian GDP for all quarters of 2004). We used the window length of 32, four leading eigengtriples to decompose the initial series and the so-called vector algorithm (see [7], p. 107) to forecast the most recent data. Fig. 8 shows the initial series and its forecast for the four quarters of 2004. The forecast is basically identical to the data. The vertical line shows the truncation between the last point of original series and the forecast starting point. This forecast is performed using the full LRF produced by the subspace, which is generated by the leading 4 eigentriples.

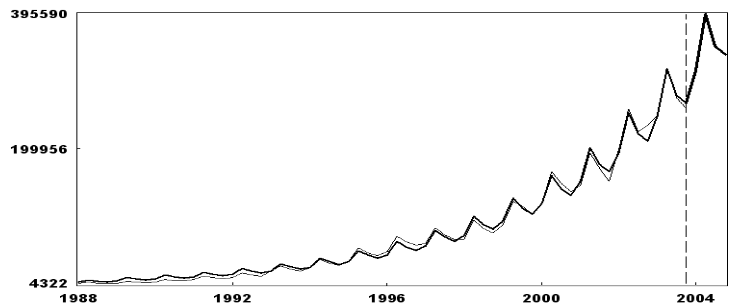

Figure 8: Approximation and Forecasting.

Confidence intervals for the forecasts

Confidence intervals for the forecasts can be calculated by two methods: the empirical method and the bootstrap method. They are calculated using the residuals of the reconstruction.

According to the main SSA forecasting assumption, the component $Y_T^{(1)}$ of the series Y_T has to satisfy an LRF of a relatively small dimension, and the residual series $Y_T^{(2)} = Y_T - Y_T^{(1)}$ has to be approximately separable from $Y_T^{(1)}$. In particular, $Y_t^{(1)}$ is assumed to be a finite subseries of an infinite series $Y^{(1)}$, which is a recurrent continuation of $Y_T^{(1)}$. These assumptions are often hold in practice with high accuracy.

There are two problems related to the construction of the confidence intervals for the forecast. The first problem is to construct a confidence interval for the original series $Y_T = \{y_t\}$ at some future point in time. The second problem is construction of confidence intervals for the signal $Y_T^{(1)} = \{y_t^{(1)}\}$ at some future point in time. These two problems can be solved in different ways. The second requires additional information about the model governing the series $\widetilde{Y}_T^{(2)} = \{\widetilde{y}_t^{(2)}\}$ to perform a bootstrap simulation of the series Y_T. Bootstrap confidence intervals are built for the continuation of the signal $Y_T^{(1)}$ (for more information see [22]).

Let us consider a method of constructing intervals for the signal $Y_{T+M}^{(1)}$ at the moment $T+M$. In the unrealistic situation, when we know both the signal $Y_T^{(1)}$ and the true model of the noise $Y_T^{(2)}$, a Monte Carlo simulation can be applied to check the statistical properties of the forecast value $\widetilde{y}_{T+M}^{(1)}$ relative to the actual term $y_{T+M}^{(1)}$.

Indeed, assuming that the rules for the eigentriple selection are fixed, we can simulate N independent copies $Y_{T,i}^{(2)}$ $(i = 1, \dots, N)$ of the process $Y_T^{(2)}$ and apply the forecasting procedure

Quarter 2004	Original data	Forecast	Relative Error	Bootstrap Average Forecast	Lower Confidence Interval	Upper Confidence Interval
Q 1	310880	313735	0.017	314052	300528	314052
Q 2	388277	395591	0.012	396032	384484	396032
Q 3	349674	348191	0.005	348933	331456	348933
Q 4	335988	335783	0.002	340786	326411	340786

Table 2: Original, forecast, average, lower and upper 95 % bootstrap confidence intervals.

to N independent time series $Y_{T,i} = Y_T^{(1)} + Y_{T,i}^{(2)}$. Then the forecasting result will form a sample $\widetilde{y}_{T+M,i}^{(1)}$, which should be compared against $y_{T+M}^{(1)}$. In this way the Monte Carlo average series for the forecast can be built up. Since in practice we do not know the signal $Y_T^{(1)}$, we can not apply this procedure. Let us describe the bootstrap variant of the simulation for constructing the confidence intervals for the forecast.

Under a suitable choice of the window length L and the corresponding eigentriples, we have the representation $Y_T = \widetilde{Y}_T^{(1)} + \widetilde{Y}_T^{(2)}$, where $\widetilde{Y}_T^{(1)}$ (the reconstructed series) approximates $Y_T^{(1)}$, and $\widetilde{Y}_T^{(2)}$ is the residual series. Suppose now that we have a (stochastic) model for the residual $\widetilde{Y}_T^{(2)}$ (for instance, we can postulate some model for $Y_T^{(2)}$ and, since $\widetilde{Y}_T^{(1)} \approx Y_T^{(1)}$, we apply the same model for $\widetilde{Y}_T^{(2)}$ with the estimated parameters). Then, simulating N independent copies $Y_{T,i}^{(2)}$ of the series $\widetilde{Y}_T^{(2)}$, we obtain N series $Y_{T,i} = \widetilde{Y}_T^{(1)} + \widetilde{Y}_{T,i}^{(2)}$ and produce M forecasting results $\widetilde{y}_{T+M,i}^{(1)}$ in the same manner as in the Monte Carlo simulation variant.

More precisely, any time series $Y_{T,i}$ produces its own $\widetilde{Y}_{T,i}^{(1)}$ reconstructed series and its own forecasting linear recurrent formula LRF_i for the same window length L and the same sets of eigentriples. Starting at the last $L - 1$ terms of the series $\widetilde{Y}_{T,i}^{(1)}$, we perform M steps of forecasting with the help of its LRF_i, to obtain $\widetilde{y}_{T+M,i}^{(1)}$.

From the sample $\widetilde{y}_{T+M,i}^{(1)}$ $(1 \leq i \leq N)$ we can calculate its (empirical) lower and upper quintiles for a fixed level γ and obtain the corresponding confidence interval for the forecast. This interval (called bootstrap confidence interval) can be compared with the forecast value $\widetilde{y}_{T+M}^{(1)}$ obtained from the initial forecasting procedure. We can also build average bootstrap series. This average can then be compared with the value $\widetilde{y}_{T+M}^{(1)}$ obtained by Basic SSA forecast. Large discrepancy between these two forecast would typically indicate that the original SSA forecast is not reliable.

The simplest model for $\widetilde{Y}_T^{(2)}$ is the Gaussian white noise model. The corresponding hypothesis can be checked with the help of the standard test for randomness and normality. Table 2 presents the original and forecasted data, relative error and the 95 % bootstrap confidence interval, including the average and the lower and upper confidence intervals, of the forecasted data. Confidence intervals are obtained by simulation under the hypothesis that the residuals of the reconstruction form a Gaussian white noise series. This table shows that the forecasted values are very close to the original data and the confidence intervals are narrow.

The Iranian GDP series (thin line) is depicted in Fig. 9 together with its bootstrap confidence interval (dashed line) for both the original (thin line) and forecasted data and the basic vector forecast (thick line). The vertical line corresponds to the truncation point.

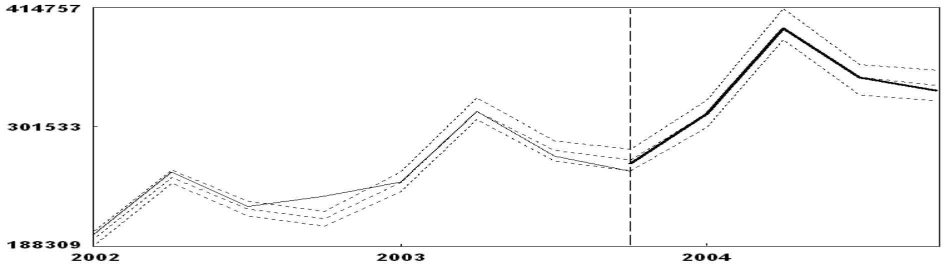

Figure 9: Bootstrap confidence intervals for the original and forecasted data.

4 Analysis of Iranian National Account

In this section we demonstrate the capability of SSA by applying it to the analysis and forecasts of the Iranian national account data. The data sets describe the main economic features of the Islamic Republic of Iran and is provided on the web-site of the Central Bank of the Islamic Republic of Iran, see [21]. The sets of data are quarterly and yearly. There are 16 quarterly data sets each containing 68 data points over the period of 1988 to 2004 (measured in billion rails, the official currency of Iran).

These sets of data are: 1 – Agriculture, 2 – Oil and Gas, 3 – Industries and Mines, 4 – Manufacturing, 5 – Mining, 6 – Electricity, Gas and Water Supply, 7 – Construction, 8 – Services, 9 – Trade, Restaurants and Hotels, 10 – Transportation, Warehousing and Communication, 11 – Financial Services, 12 – Real Estate and Professional Services, 13 – Public Service, 14 – Social, Personal and Domestic Services, 15 – Imputed Bank Services Charge and 16 – Gross Domestic Product (GDP) in Basic Price. We shall refer to these data sets as Series 1 to Series 16, respectively.

Fig. 10 displays Series 1 – 16. In this figure, the series in row i and column j is Series $4(i-1)+j$ $(i,j=1,\ldots,4)$.

It is customary in econometrics to take the logarithms of the data describing economic features. Therefore, we make a parallel analysis of the data taken in the logarithmic scale. Fig. 11 displays Series 1 – 16 in the logarithmic scale (the arrangement of the series is the same as in Fig. 10).

We also consider 16 yearly data sets which contain 45 observations each covering the period of 1959 to 2003 (measured in billion rails). These data describe exactly the same economic features as Series 1–16. We shall refer to these data as Series 17 – Series 32. Fig. 12 displays these series. In this figure, the series in row i and column j is Series $16+4(i–1)+j$ $(i,j=1,\ldots,4)$. Fig. 13 displays Series 17 – 32 in the logarithmic scale.

On the website [21] one can find the Iranian national accounts quarterly data adjusted to seasonal effects. However, we use the original, non-adjusted data since one of our aims is to illustrate the capability of the SSA technique for extracting trend and oscillations from the data. We then use the approximated trend and oscillations for forecasting the data.

4.1 Analysis of quarterly data sets

For each series, we have performed SSA analysis and forecast. Similarly to what we have done with the GDP data in Sect. 2, we have removed the last four points of each series (Q1 – Q4 of 2004), made an SSA approximation for the period 1988 to 2003 and forecasted the data for the four quarters of 2004. In each analysis, we choose the SSA parameters (which are

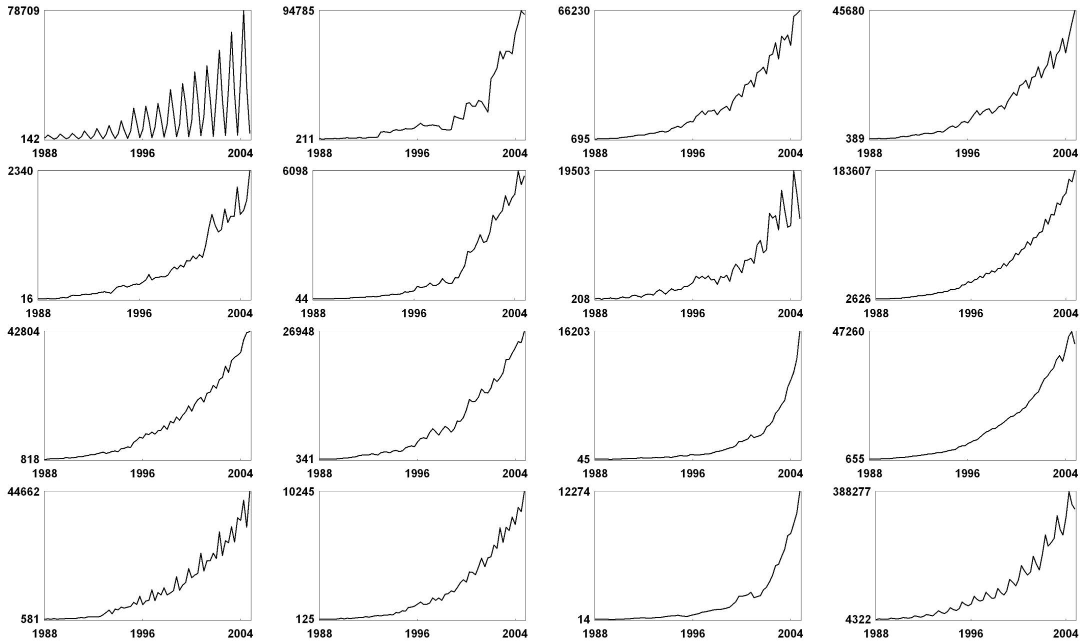

Figure 10: Series 1–16.

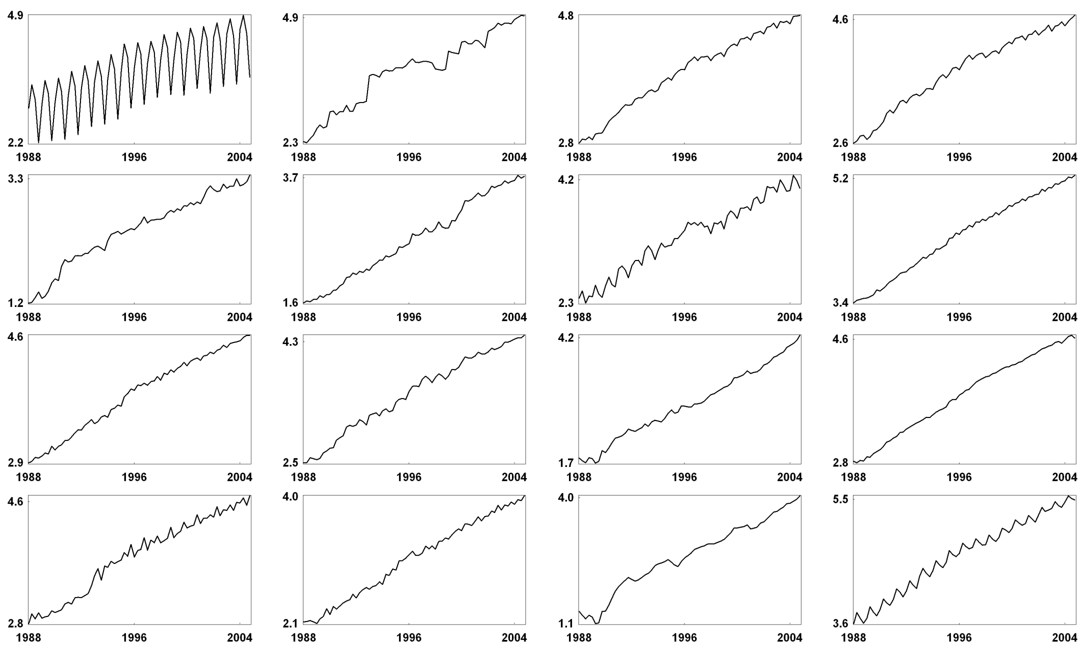

Figure 11: Series 1–16 in the logarithmic scale.

the window length and the number of eigentriples chosen for approximation) to optimize the approximation of the series keeping the window length L large enough.

For each forecasted value (Q1 – Q4 of 2004), we have computed the relative error of the forecast (in percent). To summarize the quality of the forecast, we provide the Mean Relative Error (MRE) which is simply the average of the four relative errors (in percent) for each series.

In parallel, we have performed SSA analysis and forecast for the data taken in the logarithmic scale. All the corresponding results are presented in Table 3 (in brackets). When the

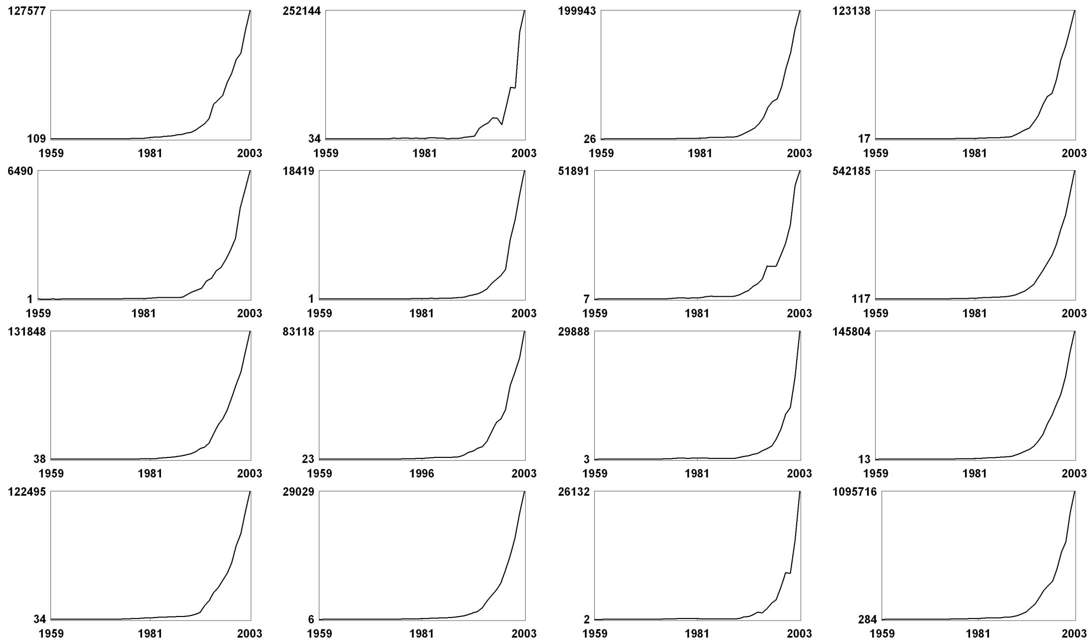

Figure 12: Series 17–32.

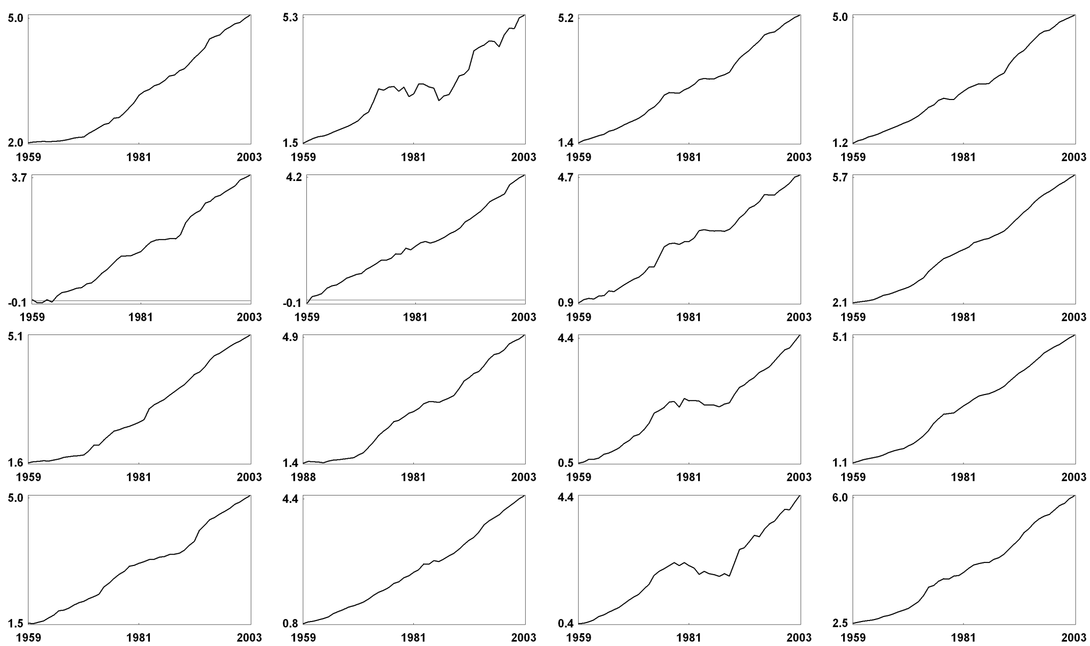

Figure 13: Series 17–32 in the logarithmic scale.

SSA analysis was performed in the log-scale, for computing the relative error of the forecast, we have transformed the forecasted data back to the original scale. We needed to do this in order to be able to compare these results with the results of the original analysis.

Table 3 shows the results. Columns 2 and 3 show the parameters of the SSA algorithm (the window length L, see Stage 1 of the algorithm, and the eigentriples chosen E, see Stage 3). Note that using this information and the SSA-Caterpillar software [14], anyone can repeat the results presented in the table).

In each cell in columns 4–7, there are two numbers: the first one is the relative error of the forecast (in percent) for the original series for a given quarter of 2004 and the second one (in brackets) is the value of the relative error of the corresponding forecast when the analysis was performed after taking the logarithms of the series. In the last column, the bold font indicates the lower of the two values. Table 3 clearly demonstrates that taking logarithms of the data does not improve the quality of the SSA forecast (on the opposite, it typically leads to its deterioration). This is related to the fact that the quarterly data have periodic components which are easier to extract when the data are considered in the original scale (taking logarithms produces additional smoothing and makes extraction of periodic components more difficult).

We consider the SSA forecasts for all 16 series as very good (an exception is Series 5 and partly Series 7 and 11). The success of the analysis means that in most cases, SSA was able to approximate both the trends and the periodic components with high accuracy. Of course, this is also related to the fact that the economy of Iran was developed steadily during the period 1988 – 2004 (the Iran-Iraq War ended in 1988).

			Relative Error%				MRE %
Ser.	L	E	Q1	Q2	Q3	Q4	
1	17 (32)	1-4 (1-10)	3.55 (6.03)	0.87 (2.50)	4.32 (4.50)	0.81 (12.60)	**2.39** (6.41)
2	32 (5)	1,6-7 (1)	0.06 (2.23)	1.24 (.021)	5.77 (0.62)	4.32 (17.0)	**2.99** (5.05)
3	32 (32)	1-7 (1-7)	1.47 (0.35)	2.16 (1.54)	0.98 (0.68)	5.38 (17.4)	**2.50** (5.01)
4	32 (32)	1-7 (1-7)	2.06 (1.66)	0.73 (4.03)	7.17 (9.56)	1.78 (4.22)	**2.93** (4.85)
5	12 (12)	1,2 (1,2)	3.75 (19.1)	6.01 (19.3)	2.95 (12.2)	13.3 (8.63)	**6.51** (14.8)
6	16 (16)	1,2,4-7 (1-4)	1.96 (0.72)	0.02 (9.04)	2.25 (1.22)	1.92 (4.81)	**1.54** (3.95)
7	32 (8)	1-5 (1-4)	10.4 (15.7)	11.9 (9.02)	1.44 (6.17)	0.34 (5.46)	**6.05** (9.09)
8	32 (32)	1-10 (1-4)	0.44 (1.46)	0.06 (0.25)	0.63 (0.07)	1.10 (5.22)	**0.56** (1.75)
9	32 (32)	1-5 (1-5)	1.02 (3.23)	0.74 (3.75)	4.58 (5.83)	0.60 (3.72)	**1.74** (4.13)
10	32 (32)	1,4-7(1-3)	1.35 (1.24)	0.00 (0.71)	4.78 (6.39)	0.22 (2.20)	**1.59** (2.63)
11	5 (10)	1,2 (1,2)	0.32 (0.17)	3.95 (3.40)	4.40 (2.12)	5.24 (8.22)	**3.48** (3.65)
12	12 (10)	1-4,6 (1,2)	4.29 (0.54)	0.30 (5.88)	0.55 (5.84)	2.01 (8.92)	**1.79** (5.20)
13	32 (32)	1-7 (1-5)	0.77 (2.84)	2.69 (3.04)	1.56 (8.67)	2.21 (1.13)	**1.79** (3.92)
14	32 (32)	1-5 (1-5)	4.10 (2.15)	2.83 (0.81)	2.64 (2.31)	1.29 (0.57)	2.72 **(1.46)**
15	8 (5)	1 (1)	1.12 (6.34)	1.04 (2.29)	0.60 (3.30)	8.17 (0.22)	**2.73** (3.11)
16	32 (24)	1-4 (1-4)	0.91 (0.55)	1.88 (0.03)	0.42 (6.49)	0.06 (0.15)	**0.82** (1.81)

Table 3: Relative Error and Mean Relative Error for Series 1 – 16 before and after taking the logarithm.

Ser.	*L*	E	Relative Error% 2001-2	2002-3	2003-4	MRE %
17	5 (5)	1,2 (1,2)	8.59 (8.48)	1.13 (1.58)	0.86(0.28)	3.52 **(3.45)**
18	3 (12)	1 (1)	24.8 (17.5)	15.9 (19.5)	2.06 (1.51)	14.2 **(12.8)**
19	7 (5)	1,2 (1,2)	1.43 (5.75)	1.64 (6.35)	6.84 (2.35)	**3.30** (4.05)
20	5 (5)	1,2 (1,2)	11.3 (0.12)	3.80 (3.42)	4.83 (6.80)	6.66 **(3.45)**
21	7 (11)	1 (1,2)	3.13 (5.25)	1.50 (1.00)	7.34 (3.20)	3.99 **(3.15)**
22	21 (9)	1 (1-4)	3.77 (4.51)	3.25 (2.68)	20.2 (3.01)	9.09 **(3.40)**
23	21 (3)	1-3 (1,2)	15.1 (0.08)	5.85 (13.8)	3.16 (0.25)	8.04 **(4.71)**
24	6 (4)	1,2 (1,2)	2.25 (0.00)	0.13 (3.64)	0.32 (7.11)	**0.90** (3.58)
25	5 (3)	1,2 (1,2)	3.17 (2.87)	0.98 (0.15)	0.56 (1.16)	1.57 **(1.39)**
26	4 (14)	1,2 (1-5)	13.7 (0.40)	4.18 (3.86)	2.80 (5.51)	6.92 **(3.25)**
27	12 (9)	1,2 (1,2)	4.33 (11.0)	2.44 (1.22)	6.31 (5.27)	**4.36** (5.84)
28	3 (6)	1 (1,2)	1.65 (4.29)	1.18 (1.05)	3.47 (3.56)	**2.31** (2.97)
29	21 (6)	1,2 (1,2)	2.38 (1.90)	1.66 (0.18)	6.31 (2.30)	**3.45** (1.46)
30	10 (10)	1 (1-3)	1.43 (0.60)	1.32 (2.42)	7.55 (5.23)	3.43 **(2.75)**
31	21 (15)	1-5,7 (1)	16.5 (19.6)	2.35 (0.92)	9.27 (16.6)	**9.38** (12.4)
32	11 (11)	1 (1-3)	0.27 (0.59)	7.20 (8.61)	0.96 (82.16)	**2.81** (3.78)

Table 4: Relative Error and Mean Relative Error for Series 17 – 32 before and after taking the logarithm.

4.2 Yearly data sets

In this section we show the results of the application of the SSA technique to 16 yearly data sets (Series 17 –32) described in Sect. 4.1. These data sets cover the period 1959 to 2003. These series contain 45 points and are shorter than the quarterly series. Moreover, the economic features exhibit clear non-stationary behaviour in this period and therefore it is much much harder to forecast the yearly series than the quarterly series.

We cut off the last 3 years of each series and forecast it to consider the precision of the technique (that is, we will forecast the values for 2001–2003). Here we do not have seasonal components so we only need to extract the trend of these data sets.

Table 4 shows the parameters of the SSA algorithm and the results of the forecasts (the structure of this table is the same as that of Table 3). The forecast results for the yearly data are generally worse than that for the quarterly data sets. The main reason for this is the fact that during the period 1959 to 2003 there were significant changes in the dynamics of the Iranian economic features, see Fig. 12 and especially Fig. 13. These changes can be associated with the start and the end of the Iran-Iraq War (1980 – 1988). Note that the changes can easily be detected by SSA, see [23] for information about using SSA for detection of changes in time series.

One may note from Table 4, that contrary to the case of the quarterly data, the forecast based on the analysis of the series in the logarithmic scale often gives better results. This is perhaps related to the fact that the yearly series do not have seasonal components which are easier to extract when the data is in the original scale.

4.3 Inflation rate series

Next, we present the forecasting results for inflation rate based on the monthly Iranian Consumer Price Index (CPI) series for the short and long horizons $h = 1, 3, 6$ and 12. In fact, we used monthly CPI data for the period Mar. 1990 - Sep. 2007. We used Jan. 1990 to Aug. 2004 CPI observations as training set and Sep. 2004 to Sep. 2007 observations for out-of-sample prediction. We select the window length $L = 60$ and the first 19 eigenvalues for reconstructing the original series and consider remaining eigentriples (20–60) as noise for forecasting inflation rate based on the CPI price index over period Sep. 2004 to Sep. 2007. We also use the RW model as a benchmark model in the comparative analyses. The use of the random walk model as a benchmark model should not imply that we believe the model is an optimal forecasting method. We use this model because it is a naive model. The point here is that a superior performance of random walk model would render the analyst's method useless. As a measure of prediction accuracy, we use the following ratio of root-mean-square errors (RMSE):

$$\text{RMSE} = \left(\frac{\sum_{i=1}^{n} (y_{T+i} - \widetilde{y}_{T+i})^2}{\sum_{i=1}^{n} (y_{T+i} - \widehat{y}_{T+i})^2} \right)^{1/2}.$$

Here n represents the number of forecasted points, $\widetilde{y}_{T+i}$ are the forecasted values of y_{T+i} obtained by SSA and $\widehat{y}_{T+i}$ is the forecasted values of y_{T+i} obtained by RW. Note that $\widetilde{y}_{T+i}$ for RW model is y_{T+i-h} for any h-step ahead forecasting. If RMSE < 1, then SSA procedure outperforms alternative prediction method. Alternatively, RMSE > 1 would indicate that the performance of the corresponding SSA procedure is worse than the predictions of the competing method.

Fig. 15 shows the CPI series and also inflation rate series based on the CPI series. Visual analysis of Fig. 15 indicates that the CPI series has a trend and this trend can be approximated by a function increasing exponentially fast. A harmonic seasonal component with decreasing amplitude is also clearly seen in Inflation rate series. In the following, we only consider Inflation rate series.

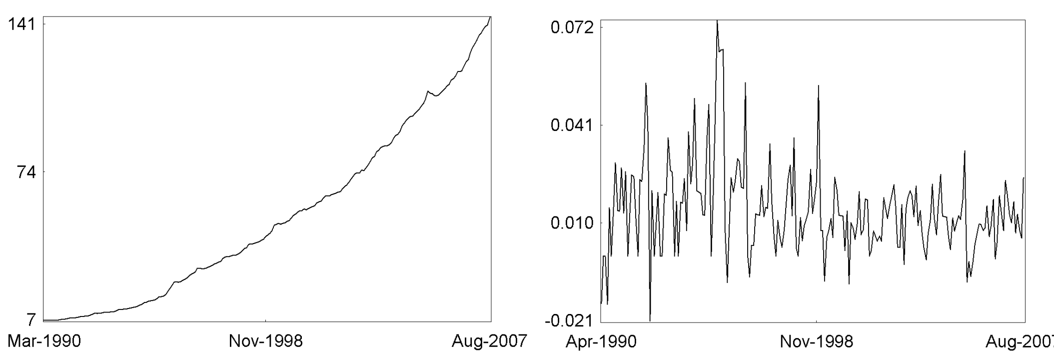

Figure 14: CPI series (left) and inflation rate series based on the CPI series (right) Mar. 1990 - Sep. 2007.

Table 5 shows the RMSEs for SSA/random walk for h-step ahead forecasts of inflation rate based on the CPI series for N forecasted data points. Without exception, SSA outperforms the random walk predictions in all h-step ahead forecasts. In fact, SSA method is up to 27% more efficient compared to the RW method. Table 5 also presents the results of Diebold and Mariano test [24] indicating whether the discrepancies between SSA and RW model forecasting procedures are statistically significant. ** and * imply significance at 1% and 10% confidence levels, respectively. The results of this table confirm that, for all cases, the differences are significant at 1% confidence level.

Additionally, Table 5 presents test results for the null hypothesis of whether the percentages of the direction of changes (DC) are greater than the pure chance (50%). The table shows that all results are statistically significant at 1% and 10% confidence levels. The results of this table also show that MSSA predicts direction of change for 12-step as accurately as it can predict 1-step ahead.

Fig. 15 (left) shows the Iranian GDP deflator series (yearly); the data are taken from http://data.un.org. One can see that this series looks very similar to the GDP series. SSA analysis and forecasting results for these two series are also very similar (the results of SSA analysis for the GDP deflator series are not reported here).

Fig. 15 (right) shows the Iranian GDP series normalized to the Iranian GDP deflator. The results of SSA forecasting (not reported here) show that it is generally more advantageous to analyze and forecast the two series (namely, Iranian GDP series and Iranian GDP deflator series) separately and then compute the ratio of the forecasts rather than to analyze and forecast the ratio only.

$h=1$			$h=3$			$h=6$			$h=12$		
N	RMSE	DC	N	RMSE	DC	N	RMSE	DC	N	RMSE	DC
36	0.81**	0.69**	34	0.78**	0.68*	31	0.73**	0.74**	25	0.84**	0.67*

Table 5: RMSE of SSA forecast results with respect to the RW method, Diebold-Marino significance test results and direction of change test for inflation rate based on the CPI series.

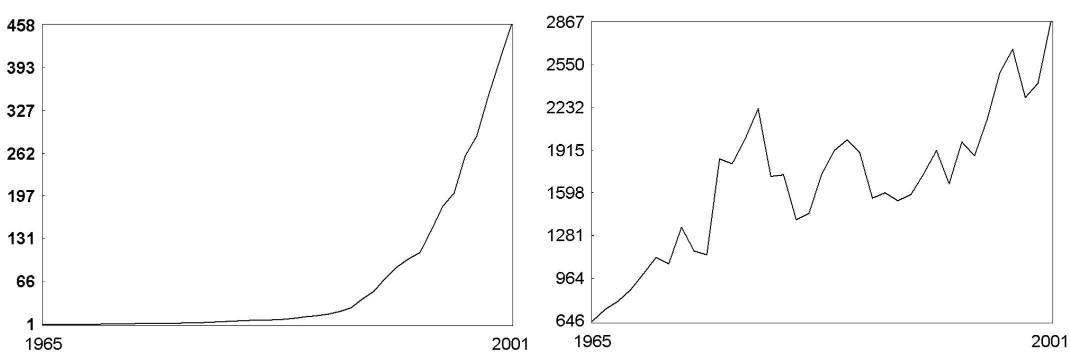

Figure 15: Iranian GDP deflator (left side) and Iranian GDP/Iranian GDP deflator (right side).

5 Conclusion

In this paper we have described the methodology of SSA (Singular-Spectrum Analysis) and demonstrated that SSA can be successfully applied to the analysis and forecasting of economic time series. We have used 32 Iranian national account data sets describing the main economic features of the Islamic Republic of Iran, as provided on the web-site of the Central Bank of the Islamic Republic of Iran [21]. The data are given in a quarterly and yearly format and have different types of non-stationarity. All the data sets are rather short.

The results show that SSA can be successfully used for the analysis and forecasting of short economic time series with different types of non-stationarity. In particular, many quarterly series have periodic components with non-stationary amplitudes but SSA has been able to extract and forecast these periodic components very accurately. Most of the yearly data have clear structural changes which makes the application of standard methods of analysis almost impossible.

Unlike standard methods used for analysis of economics time series, SSA does not require parametric models or transformation of the data into the logarithmic scale. Moreover, our study has shown that in most cases, the transformation of the quarterly series into logarithmic scale has lead to the deterioration of the precision of the forecasts.

Acknowledgement

The authors are grateful to their colleagues Nina Golyandina and Vladimir Nekrutkin (University of St.Petersburg) for constructive remarks and to the referee for useful suggestions. The first author would also like to thank the Central Bank of the Islamic Republic of Iran for its support during his research study at Cardiff University.

References

[1] S. D. Krane, An evaluation of real GDP forecasts: 1996-2001, *Economic Perspectives*, 2003, http://www.chicagofed.org/publications/economicperspectives/2003/1qeppart1.pdf.

[2] L. Y. Cao, and A. Soofi, Nonlinear deterministic forecasting of daily dollar exchange rates, *International Journal of Forecasting*, 1999, **15**(4): 421–430.

[3] A. Soofi, and L. Y. Cao, Nonlinear Forecasting of Noisy Financial Data, in Soofi and Cao (eds.), *Modeling and Forecasting Financial Data: Techniques of Nonlinear Dynamics*, Kluwer Academic Publishers, Boston, 2002.

[4] D. A. Hsieh, Chaos and nonlinear Dynamics: Application to Financial Markets, *Journal of Finance*, 1991, 46: 1839–1877.

[5] J. Scheinkman, and B. LeBaron, Nonlinear Dynamics and Stock Returns, *Journal of Business*, 1989, 62: 311–337.

[6] D. S. Broomhead, and G. P. King, Extracting Qualitative Dynamics from Experimental Data, *Physica D*, 1986, 20: 217–236.

[7] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure: SSA and related techniques*, Chapman & Hall/CRC, New York - London, 2001.

[8] D. Danilov and A. Zhigljavsky (Eds.). *Principal Components of Time Series: the 'Caterpillar' method*, University of St. Petersburg, St. Petersburg. (In Russian.), 1997.

[9] J. B. Elsner and A. A. Tsonis, *Singular Spectral Analysis, A New Tool in Time Series Analysis*, Plenum Press, New York and London, 1996.

[10] V. G. Moskvina, and A. Zhigljavsky, An algorithm based on singular spectrum analysis for change-point detection, *Communication in Statistics - Simulation and Computation*, 2003, **32**(4): 319–352.

[11] H. Hassani, Singular Spectrum Analysis: Methodology and Comparison, *Journal of Data Science*, 2007, **5**(2), 239–257.

[12] H. Hassani, S. Heravi and A. Zhigljavsky, Forecasting European Industrial Production with Singular Spectrum Analysis, *International Journal of Forecasting*, Fortcoming.

[13] Th. Alexandrov, and N. Golyandina, The automatic extraction of time series trend and periodical components with the help of the Caterpillar-SSA approach. *Exponenta Pro 3-4* (In Russian.), 2004: 54–61.

[14] www.gistatgroup.com

[15] S. L. Marple-Jr, *Digital Spectral Analysis*, Prentice Hall, New Jersey, 1987.

[16] M. Bouvet, and H. Clergeot, Eigen and singular value decomposition technique for the solution of harmonic retrieval problems. In E. F. Deprettere (Ed.), *SVD and Signal processing: Algorithm, Applications and Architectures*, North-Holland, Amsterdam, 1988: 93–114.

[17] V. K. Madisetti, and E. Lloyd (Eds.) *The Digital Signal Processing Handbook*, CRC Press, Boca Raton, 1998.

[18] D. Brillinger, *Time Series. Data Analysis and Theory*, Holt, Rinehart and Winston, Inc., New York, 1975.

[19] T. Subba Rao, Canonical factor analysis and stationary time series models. *Sankhya: The Indian Journal of Statistics*, 1976, 38B: 256–271.

[20] T. Subba Rao, and M. M. Gaber, *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Speringer-Verlag, 1984.

[21] http://www.cbi.ir

[22] B. Efron, and R. Tibshirani, Bootstrap methods for standard error, confidence intervals and other measures of statistical accuracy. *Statist. Science*, 1986, **1**(1): 54–75.

[23] http://www.cf.ac.uk/maths/stats/changepoint

[24] F. X. Diebold, and R. S. Mariano. (1995). Comparing predictive accuracy, *Journal of Business, Economics and Statistics*, **13**: 253–263.