Automatic extraction and forecast of time series cyclic components within the framework of SSA

Th. Alexandrov, N. Golyandina

St. Petersburg University, Mathematical Department

Abstract

Application of Singular Spectrum Analysis (SSA) for the purpose of extracting and forecasting time series cyclic components is considered. Automatic methods of identification are applied within the framework of SSA. Their suitability for the class of exponentially modulated harmonic components is numerically investigated.

Introduction

Let us consider a time series $F_N = (f_0, f_1, \ldots, f_{N-1})$ as a sum of additive components (such as a trend, regular oscillations and a noise) and set the task of extraction (approximation) of one of the components and its continuation. Since the model of the signal (i.e., of the deterministic component of the time series) doesn't assumed to be known, the nonparametric method 'Caterpillar'-SSA [2] is therefore appropriate.

We start with brief description of the SSA algorithm for extraction of an additive component S_N of the observed time series $F_N = S_N + R_N$ (R_N is a residual) with length N (the detailed description see in [2]). At the first stage of the algorithm we choose a window length L, $1 < L < N$, and construct a trajectory matrix $\mathbf{X}$ with vectors $X_j = (f_{j-1}, \ldots, f_{j+L-2})^\mathrm{T} \in \mathbb{R}^L$, $j = 1, \ldots, K = N - L + 1$, as columns. The next step is calculation of the matrix $\mathbf{X}\mathbf{X}^T$, its eigenvalues $\{\lambda_j\}_{j=1}^L$ numbered in decreasing order, $d = \max\{j : \lambda_j > 0\}$, eigenvectors $\{U_j\}_{j=1}^d$, and factor vectors $\{V_j\}_{j=1}^d$. The eigentriples $(\sqrt{\lambda_j}, U_j, V_j)$, $j = 1, \ldots, d$, form the Singular Value Decomposition (SVD) $\mathbf{X} = \sum_j \sqrt{\lambda_j} U_j V_j^T$. At the second stage we need to identify the group I of eigentriples that correspond to S_N. Finally, the algorithm produces the reconstructed component $\widetilde{S}_N$, calculated by diagonal averaging (hankelisation) of the matrix $\mathbf{X}^I = \sum_{j\in I} \sqrt{\lambda_j} U_j V_j^T$.

Once the group I is chosen we can construct a recurrent forecast of the reconstructed component: the space spanned by $\{U_j\}_{j\in I}$ generates a linear recurrent formula which can be used for forecasting $\widetilde{S}_N$ (see details in [2]).

Thus extraction/forecast of the component S_N is controlled by choice of window length L and by identification of components which form the group I. Generally speaking, the group I corresponding to S_N exists only if the conditions of (approximate) separability of S_N from the residual R_N are satisfied [2]. Hereinafter we assume that the approximate separability takes place.

There is the software [4], which allows users to produce manual identification of required eigentriples in interactive mode; choice of eigentriples is based on visual information and theoretical results. However workable methods of automatic identification of eigentriples can considerably extend the range of problems solved by SSA.

It is no wonder that automatic identification (AI) produces results worse than that of interactive visual identification (VI). Moreover, AI methods are not fully automatic (they are based on thresholds setting); therefore several questions evidently arise. First of all, we have to be sure that there exists such optimal AI thresholds (certainly their choice is based on the time series) that the accuracy of component reconstruction/forecast by means of AI is comparable to the accuracy of VI results. Secondly, it is necessary to develop an approach to setting the AI threshold, which allows one to process every time series from a prescribed class with appropriate quality. This demand being fulfilled enables one to apply AI to real-life time series with unknown features.

In this paper we investigate the method of AI [1] used for extraction and forecast of exponentially-modulated (e-m from here) harmonics. This method can be applied to extraction of regular oscillations in a rather general case, when these oscillations can be approximated by a sums of e-m harmonic components.

1 Identification of e-m harmonic eigentriples

Automatic identification of eigentriples produced by e-m harmonic S_N with general term $s_n = Ae^{\alpha n}\cos(2\pi\omega n)$, $0 < \omega < 0.5$, is applied to eigenvectors $\{U_j\}_{j=1}^{d}$.

Let us describe in brief the so-called Fourier method [1, 3]. We assume hereinafter that $L\omega$ is integer (L is divisible by the cycle size $1/\omega$) and also that L is large enough and α is small enough to give a possibility to use asymptotical results under the assumption that $L \to \infty$, $\alpha \to 0$, $L\alpha \to \mathrm{const} \geqslant 0$. We expect without loss of generality that α is positive. The identification method is based on the fact that an e-m harmonic generates two eigentriples with eigenvectors similar to e-m harmonical sequence with the same frequency and exponential rate where phase shift is close to $\pi/2$. Furthermore the corresponding eigenvalues are close. Here we apply the Fourier method to consequent pairs of eigenvectors U_j, U_{j+1}.

Let us define the periodogram $\Pi_U^L(\omega)$ of a vector $U \in \mathbb{R}^L$ for $\omega \in \{k/L\}_{k=0}^{\lfloor L/2 \rfloor}$ by the formula

$$\Pi_U^L(k/L) = \frac{L}{2}\begin{cases} 2c_0^2, \quad k = 0, \\ c_k^2 + s_k^2, \quad 1 \leqslant k \leqslant (L-1)/2, \\ 2c_{L/2}^2, \quad \text{if } L \text{ is even and } k = L/2, \end{cases}$$

where c_k and s_k are coefficients before cosine and sine with the frequency k/L in the Fourier expansion of the sequence $u_1, \ldots, u_L$. The value $\Pi_U^L(\omega)$ reflects the contribution of a harmonic with frequency ω to the Fourier expansion of $u_1, \ldots, u_L$.

We define $\rho_{j,j+1} = 0.5 \max_{0 \leqslant k \leqslant L/2} \left(\Pi^L_{U_j}(k/L) + \Pi^L_{U_{j+1}}(k/L) \right)$.

We say that an eigenvectors pair (U_j, U_{j+1}) is identified as corresponding to some e-m harmonic if periodograms of U_j and U_{j+1} are peaked at the same frequency and $\rho_{j,j+1} \geqslant \rho_0$ for the given threshold $\rho_0 \in [0,1]$.

Consider the special case of exact separability of e-m harmonic component from the residual and let the pair U_j, U_{j+1} correspond to this component. Then, choosing a pure harmonic ($\alpha = 0$), would yield $\rho_{j,j+1} = 1$. It can be shown under the conditions assumed in Section 1 that the following relation takes place for $\alpha > 0$

$$\rho_{j,j+1} \approxeq \widetilde{\rho}(\gamma) = \frac{2}{\gamma} \frac{(e^\gamma - 1)}{(e^\gamma + 1)}, \tag{1}$$

where $\gamma = \alpha L$. Note that any $\rho_0 \lessapprox \widetilde{\rho}(\gamma)$ leads to identification of the pair of eigentriples corresponding to the e-m harmonic in conditions of exact separability.

2 Numerical results

Consider the time series $F_N = (f_0, f_1, \ldots, f_{N-1})$ with general term

$$f_n = s_n + \sigma e^{\alpha n} \varepsilon_n, \qquad s_n = A e^{\alpha n} \cos(2\pi\omega n), \tag{2}$$

where ε_n is the normal white noise with zero mean and unit variance. Accuracy of separability of the signal S_N from the noise depends on choice of window length L and can be varied by values of σ and of time series length N. Note that the case $\alpha = 0$ corresponds to a pure harmonic. Quality of the identification procedure can be estimated on the base of V simulations of the series (2). Estimates of average error for extraction and Q-term forecast of S_N are examined as characteristics of identification quality.

The model (2) corresponds to multiplicative mode of errors. Therefore it is natural to consider exponentially weighted errors WMSE and WMSD instead of conventional mean square errors (MSE) and square root of that (MSD). The weighted variants use weights $e^{-\alpha n}$ for n-th term of the time series. Surely, weighted errors coincide with MSE and MSD for $\alpha = 0$.

In the below numerical examples we'll use values $A = 3$, $V = 1000$, $Q = 13$, $\omega = 1/12$ (i.e. period of a cycle is equal to 12).

2.1 AI with optimal thresholds

It is known that for the considered model (2) (if the noise is not too big) two leading eigentriples correspond to the signal S_N. That is why the interactive visual identification VI is equivalent to reconstruction/forecast of S_N with $I = \{1, 2\}$.

Denote by $\rho_0^{(\mathrm{opt})}(F_N, L)$ the optimal threshold, which gives minimum average WMSE for the time series F_N defined by (2) and window length $L = L(F_N)$.

The inequality

$$\rho_0^{(\mathrm{opt})}(F_N, L) \leqslant \widetilde{\rho}(\alpha L)$$

is reasonable, since eigenvectors form can be distorted due to approximate separability and therefore the AI criterion should be weaken in comparison with the exact separability case. The worse separability of S_N from the residual is, the greater is the difference between $\rho_0^{(\mathrm{opt})}(F_N, L)$ and $\widetilde{\rho}(\alpha L)$.

Numerical experiments show that AI with $\rho_0 = \rho_0^{(\mathrm{opt})}(F_N, L)$ is quite similar to VI, i. e. average errors of AI are very close to average errors of VI (error distributions are close too) if αL isn't too big. These results confirm that the considered AI method is quite appropriate for extraction of e-m harmonics if we can take the window length L divisible by the cycle size.

Let us introduce results of harmonics extraction for $N = 47$ and $L = 24$ (24 is divisible by the period value 12) and take time series (2) with $\alpha = 0, 0.01, 0.02$ and $\sigma = 0, 1, 2$. Tables 1 and 2 contain optimal thresholds for extraction and forecast of the e-m harmonic component and the corresponding errors.

Table 1: Optimal thresholds for extraction (left) and forecast (right): $N = 47$, $L = 24$

α	**0**	**0.01**	**0.02**
$\sigma = 0$	1.00	0.99	0.98
$\sigma = 1$	0.97	0.95	0.91
$\sigma = 2$	0.91	0.90	0.87

α	**0**	**0.01**	**0.02**
$\sigma = 0$	1.00	0.99	0.98
$\sigma = 1$	0.97	0.95	0.91
$\sigma = 2$	0.93	0.92	0.89

Table 2: Minimal WMSD for extraction (left) and forecast (right): $N = 47$, $L = 24$

α	**0**	**0.01**	**0.02**
$\sigma = 0$	0	0	0
$\sigma = 1$	0.34	0.34	0.36
$\sigma = 2$	0.76	0.78	0.84

α	**0**	**0.01**	**0.02**
$\sigma = 0$	0	0	0
$\sigma = 1$	0.52	0.53	0.54
$\sigma = 2$	1.10	1.11	1.15

These tables give us the following observations: optimal thresholds for reconstruction and forecast (almost) coincide; forecast errors are clearly greater than reconstruction errors; weighted errors depend on α slightly.

Optimal values of thresholds produce almost proper quantity of identified eigenvectors with proper numbers. Evidently, thresholds less than optimal ones can lead to the choice of redundant eigenvectors. Thresholds with greater than optimal values can cause a loss of the desired eigenvectors. The last is more crucial. The described effect is confirmed by dependence of WMSD error on threshold value (see Fig. 1 with $\alpha = 0.01$ and $\sigma = 1$).

Thus, we can slightly decrease the threshold ρ_0 (and weaken the AI criterion) with inconsiderable loss of criterion quality, if necessary for whatever reason. However we cannot increase ρ_0, therefore the following inequality should be fulfilled

$$\rho_0 \leqslant \rho_0^{(\mathrm{opt})}(F_N, L). \tag{3}$$

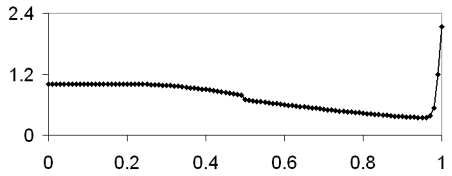

Figure 1: Dependence of WMSD on ρ_0

2.2 Choice of AI thresholds for a set of time series

The previous section shows that simulation allows us to find AI thresholds giving a good quality of identification if we know model of time series in detail. Unfortunately, this way of threshold choice cannot be realized if we process real-life time series and probably know only general features of the set of analyzed time series. Let us study this problem using extraction of e-m harmonic as an example.

Consider a set $\mathcal{K}$ of time series consisting of e-m harmonic S_N as an additive component. Our aim is to choose the threshold ρ_0, which gives an appropriate quality of extraction of S_N for any time series $F_N \in \mathcal{K}$. Let us fix such way of choice of window length $L = L(F_N)$ that S_N is approximately separated from the residual. Due to (3) we obtain that the choice

$$\rho_0 = \rho_0^{(\mathrm{opt})}(\mathcal{K}, L) \overset{\mathrm{def}}{=} \min_{F_N \in \mathcal{K}} \rho_0^{(\mathrm{opt})}(F_N, L) \tag{4}$$

gives identification of the desired eigenvector pair for any $F_N \in \mathcal{K}$. Certainly, the smaller threshold value is, the weaker is the AI criterion that leads to possible choice of wrong (redundant) eigentriples pairs. Therefore, application of AI for processing time series from the set $\mathcal{K}$ with fixed rule for choice of $L = L(F_N)$ can induce appropriate results, as long as $\rho_0^{(\mathrm{opt})}(\mathcal{K}, L)$ doesn't differ from 1 too much.

As an example of the set $\mathcal{K}$ let us take time series (2) with $N \geqslant N_0$, $\sigma \leqslant \sigma_0$, $\alpha L \leqslant \gamma_0$.

Let $L \approx N/2$ and $L\omega$ be an integer (this choice gives the best quality of separability). Then the condition $\alpha L \leqslant \gamma_0$ transforms to $\alpha N/2 \leqslant \gamma_0$. This means that the amplitude of S_N cannot increase bigger than $e^{2\gamma_0}$ times. We have $\rho_0^{(\mathrm{opt})}(\mathcal{K}, L) = \rho_0^{(\mathrm{opt})}(F^*_{N_0}, L)$, where $F^*_{N_0}$ is a time series with length N_0 and $\sigma = \sigma_0$ (the case of worst separability) and $\alpha = \alpha_0 = 2\gamma_0/N_0$ (the case of minimal $\widetilde{\rho}(\alpha L)$).

Take $N_0 = 47$, $\sigma_0 = 2$, $\gamma_0 = 0.5$. Then $\rho_0^{(\mathrm{opt})}(\mathcal{K}, L) = 0.87$ (see Table 1, $\sigma = 2$, $\alpha = 0.02$). Table 3 (left) contains errors WMSD obtained with $\rho_0 = 0.87$ (compare with Table 2 (left)).

If we take $N = 95$ and $L = 48$, then $\rho_0 = 0.87$ should guarantee good results for $\alpha \leqslant 0.01$. Table 3 (right) confirms it. Also, this table shows that the threshold 0.87 taken for the worst case gives appropriate results for $\alpha = 0.02$ and $\sigma = 1$. As for $\alpha = 0.03$, AI with threshold equal to 0.87 is failed.

If we consider forecast errors, then these conclusions would remain as before (see Table 4).

Thus the considered method of automatic identification shows its applicability

Table 3: WMSD for extraction: $\rho_0 = 0.87$; $L = 24$, $N = 47$ (left) and $L = 48$, $N = 95$ (right)

α	**0**	**0.01**	**0.02**
$\sigma = 1$	0.38	0.38	0.38
$\sigma = 2$	0.79	0.80	0.84

α	**0**	**0.01**	**0.02**	**0.03**
$\sigma = 1$	0.27	0.27	0.28	1.81
$\sigma = 2$	0.55	0.55	0.84	1.92

Table 4: WMSD for forecast: $\rho_0 = 0.87$; $L = 24$, $N = 47$ (left) and $L = 48$, $N = 95$ (right)

α	**0**	**0.01**	**0.02**
$\sigma = 1$	0.54	0.54	0.55
$\sigma = 2$	1.15	1.15	1.16

α	**0**	**0.01**	**0.02**	**0.03**
$\sigma = 1$	0.34	0.34	0.37	1.78
$\sigma = 2$	0.72	0.73	0.94	1.88

for extraction and forecast of exponentially-modulated harmonics in some restrictions on the analyzed set of time series.

References

[1] Alexandrov, Th., Golyandina, N. (2005) Thresholds for methods of automatic extraction of time series trend and periodical components with the help of the 'Caterpillar'-SSA approach. In: Proceedings of the IV International Conference 'System Identification and Control Problems' SICPRO'05, p. 1849–1864 (in Russian).

[2] Golyandina, N.E., Nekrutkin, V.V., Zhigljavsky, A.A. (2001) *Analysis of Time Series Structure. SSA and Related Techniques*, Chapmap & Hall/CRC.

[3] Vautard, R., Yiou, P., Chil, M. (1992) Singular-spectrum analysis: A toolkit for short, noisy chaotic signals, Physica D 58, p. 95–126.

[4] Time series analysis and forecast: the 'Caterpillar'-SSA method
http://www.gistatgroup.com